Journal of Computers

ISSN 1796-203X

Volume 9, Number 5, May 2014

Contents

A Protocol for Energy Efficient Mechanism for Wireless Sensor Network with Symmetric Cluster Formation

Sanjay Kumar Padhi[1] , Prasant Kumar Pattnaik[2]
[1]Department of Computer Science and Engineering,
Konark Institute of Science and Technology, Jatni, India
sanjaya2004@yahoo.com

[2]School of Computer Engineering, KIIT University, Bhubaneswar, India
patnaikprasantfcs@kiit.ac.in

***Abstract*—The network scalability may be achieved by grouping sensor nodes into a cluster hierarchy. Cluster head is referred as the leader of every cluster. In order to achieve stable clustering for mobile environment many clustering schemes are used for Wireless Sensor Network. This paper, proposeds an extension to Low Energy Adaptive Cluster Head protocol namely a protocol for energy efficient mechanism for wireless sensor network with symmetric cluster formation which is known as Highest Energy Clustering Hierarchy (HECH) protocol. Mathematical simulation studies show the correctness and effectiveness of the protocol.**

***Index Terms*— Cluster, Routing Protocol, Symmetric Cluster Formation, Wireless Sensor Network**

I. INTRODUCTION

Wireless Sensor Networks (WSNs) are network systems containing sensor nodes. The sensor nodes can sense certain physical characteristic and can be used to capture environmental information such as temperature, motion, sound etc[1]. The Sensor was used to send collected data usually via radio transmitted to a command center (sink) either directly or through a data concentration center(gateway).From the technological advances ,the size and cost of sensor has decreased and stimulate our interest to use large set of not reusable unattended sensors[2] .Such interest motivated intensive research in the past few years addressing the potential of collaborations among sensors in data gathering, processing, coordination , managements of the sensitivity activity and data flow to the sink. In adhoc manner a mutual distributed wireless sensor is designed in a network.

II.TYPES OF ROUTING PROTOCOLS

Many routing protocols such as [3, 4, 5, 6] are planned for wireless sensor networks for the problem of routing. The routing mechanism consists of sensor nodes characteristics along with the application and architecture requirements. Almost all of the routing protocols classified as hierarchical, location-based or Data-centric. In subsequent we consider two popular protocols namely Leach and Pegasis.

A. LEACH

A cluster based protocol is Low Energy Adaptive Clustering Hierarchy (LEACH) [4] protocol. The mechanism of LEACH is rotation of cluster Head in a random way and to evenly distribute the energy load among the sensors in a network. The Cluster heads broadcasts once the clusters are constructed. Time Division Multiple Accessing (TDMA) schedules provide the order of transmission for members in the cluster. Time slot is assigned to each node. Within the exclusive timeslot, data is transmitted to the cluster head by the node. The Cluster head will be elected in the next round randomly once the last node in the schedule has transmitted its data. To improve the scalability, it employs localized coordination and balance the energy usage of the network among all the nodes.

B. PEGASIS

The chin-based power efficient protocol called Power-Efficient Gathering in Sensor information System (PEGASIS) [7] protocol was based on LEACH. The assumption has made that each node must know location information about all other nodes at first. From the base station the farthest node is considered first by PEGASIS. By using a greedy algorithm the chain can be constructed easily. The data is aggregated by chain leader and forwards it to base station. Each node in the chain takes turn to be the leader in order to balance the over head involved in communication between the chain leader and the base station. A self-organizing, adaptive cluster protocol that uses randomization to distribute the energy load evenly among the sensors in the network is the routing algorithm. According to the Algorithm the nodes organize themselves in local clusters. Among them one node is organized itself as local base station or cluster head. The cluster head nodes are not fixed rather this position is self elected at different time intervals in order

doi:10.4304/jcp.9.5.1041-1046

to spread the energy usage over multiple nodes. No extra negotiation is required to determine the cluster heads because each node makes its decision about whether to be a cluster-head independently of the nodes in the network.

III. MOTIVATIONS

Based on receiving signal strength in LEACH clusters will be formed and use local cluster heads to routers and to sink but its disadvantages are that there is no uniformity in the battery levels of the nodes due to random election of the cluster head. Particular part of network may die quickly. We propose HECH protocol by the following papers [8,9,10,11,12,13,14,15,16,17,18]. Our proposed protocol provides a cluster constructing method which avoids the uneven member distribution for clusters and also provides a hierarchical routing scheme between cluster heads and base station.

IV. HECH PROTOCOL:

Our approach is given below.
1. Formation of Cluster (For the first round only)
2. Election of Cluster Head.
3. Communication inside the cluster.
4. Cluster Heads and Base station communications.

The features of the proposed protocol are given below.

- **Formation of Cluster:**
  Uniform Cluster formation taking the position information of nodes into consideration is the main key feature of this protocol and the number of nodes in a cluster is limited to cluster number.

- **Election of Cluster Head:**
  Based on the remaining battery level of the nodes cluster heads will be elected for all cluster and this is not happened in LEACH (There is a random election of cluster heads in LEACH).

- **Communication inside the cluster:**
  Initially all cluster members has same energy levels inside the cluster, after one iteration between nodes there will be a difference in the energy levels. The node with maximum battery level is elected as cluster head. With cluster head as root a rooted tree is constructed. The cost of data transmission is depending on the distance of transmission. Except Cluster head inside a cluster a node sends data to its predecessor and the aggregated data will reach to cluster head finally.

- **Cluster Heads and Base station communications:**

In hierarchical tree manner all cluster Heads communicate each other and the data is forwarded to the base station. Two parameters will be taken into consideration here that is hop-count and energy cost of the path (energy cost is given by the total amount of energy consumed if a particular path is followed.)

Description of Algorithm:

STEP-1 The position of the remaining nodes in the network has in the knowledge of every node and the max number nodes inside a cluster are limited to cluster number.

STEP-2: After completion of the previous step the entire network is divided into number of clusters and each node belongs to one cluster exactly.

STEP-3: The node which has highest battery level is elected as cluster head for that round in each cluster.

STEP-4: With cluster head as root trees are constructed inside the cluster and among the cluster heads with base station as root.

STEP-5: By assuming that each node has data to transmit, they will transfer the data by transmitting to their neighbors towards the root (cluster head) and all the cluster heads sends the aggregated data to their neighboring Cluster Head's towards base station. After this each node has different battery levels then we have to go to STEP-3 for next round.

Pseudo Code for Proposed Clustering Algorithm:

```
main ( )
{
formnodes();
// form 100 random sensor nodes in the xy plane.
formclusters();
 //it will construct the clusters as in STEP-1 of the above algorithm.
electclusterheads();
// Clusters heads will be elected based on the battery levels of the nodes.
formtrees();
// Trees will be constructed with cluster head as root.
processtheschedule();
// Data Transfer Phase.
avgclusterenergy():
 // Analysis of the network.
electclusterheads():
 // Elect cluster heads for the next round.
}
```

V EVALUATION OF HECH

The main features of HECH are
1. Reduction of Energy dissipation.
2. Localized coordination along with Self configuration.
3. Maximum energy is with Cluster Head.
4. Load balanced.

However drawbacks in HECH is given. First, our control messages are more than those in LEACH. Because we want to get more information to construct more evenly distributed topology and the hierarchical

routing tree. The cluster-head in LEACH transmits data to base station directly. Our cluster-heads use hierarchical routing to forward the data to the base station.

VI. Simulation Results and Performance Analysis

Our algorithm is implemented in Glomosim Simulator. Every node in the network belongs to some cluster. Our assumptions regarding simulation are as follows:

- For message passing between any two nodes one unit cost is taken.
- For sending request to joining into a cluster one unit of cost is taken.
- For communication between any two nodes two units of cost is taken.

With these assumptions we simulated our algorithm using Glomosim.

Simulation Parameters:

The various parameters considered for simulation were:

- Network size: The network size is considered as 100X100 m^2.
- Area: The radius of the network is a measure of its area. Nodes are randomly deployed in a given area.
- Hop Count: The maximum hop count between cluster head and any node belonging to the cluster.
- Cluster Number: Maximum number of nodes inside a Cluster.

Performance Metrics:

Our aim is to minimize the energy consumption in clustering a network and uniform cluster distribution for uniform load distribution.

- Energy: The most considerable parameter in cluster formation is energy. We calculated the initial energy of whole network and the energy of the network after processing the schedule, which reflects the energy consumed during communication.
- Cluster Number: The Maximum number of nodes inside a cluster.

First, we will discuss the cluster topology distribution

Fig. 1 shows the cluster distribution of LEACH. We can say that the cluster distribution is not balanced from the figure. For making it balanced we will make our clustering approaches. And Figure-2, Figure-3 show that our clustering approaches. All the Figures show that the cluster topology is more balanced. The clustering becomes more balanced as the number of clusters increases but there is an upper limit for number of clusters, otherwise more energy have to be spend for cluster formation itself. So we consider this and find the optimum value for both number of numbers and cluster number.

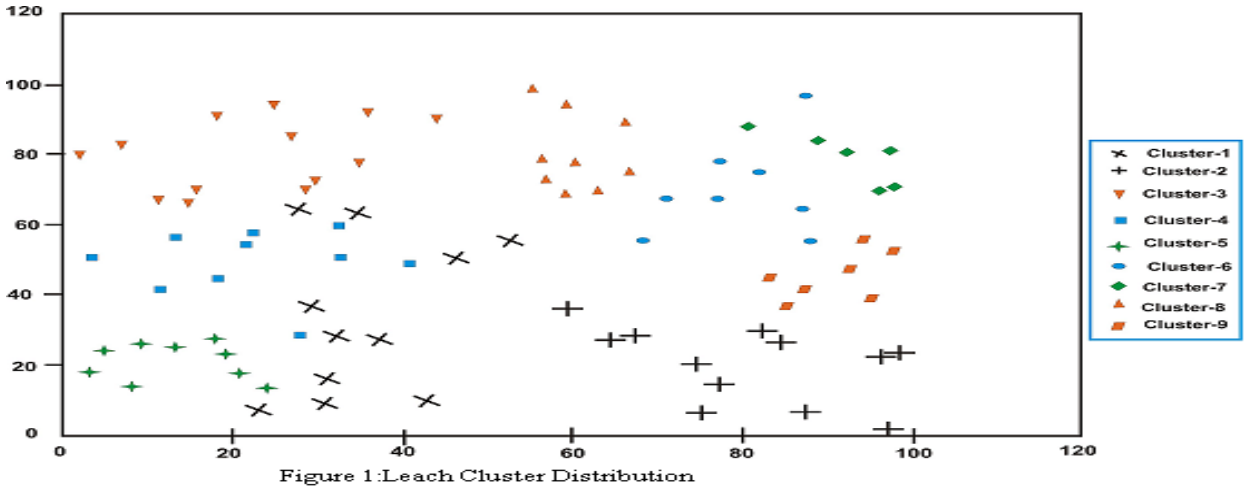


Figure 1 :Leach Cluster Distribution

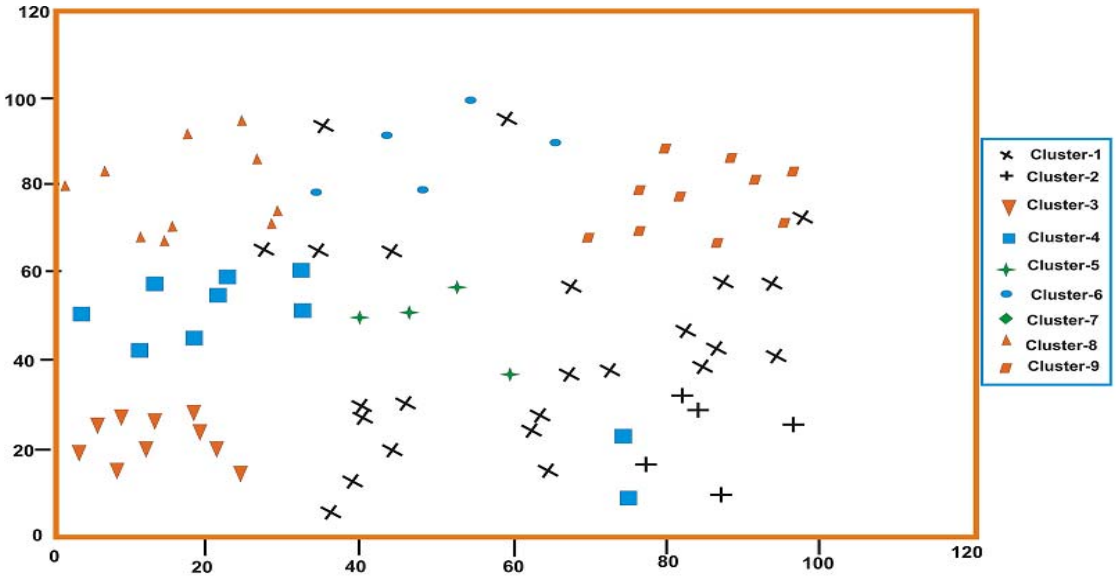


Figure 2:HECH with Cluster Number=12

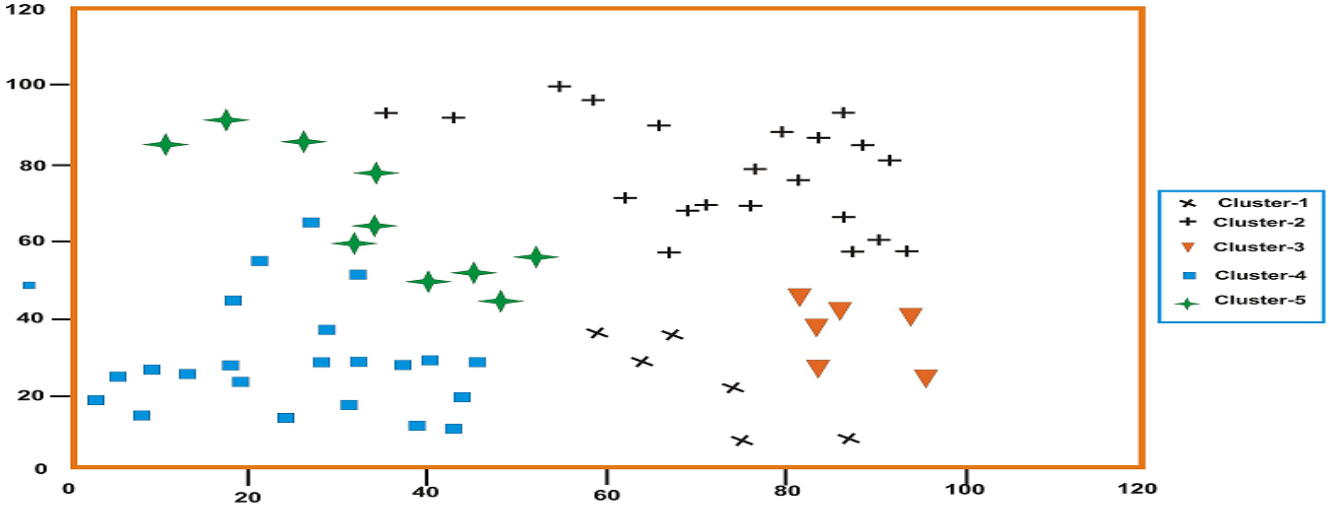


Figure 3:HECH with Cluster Number=15

The following simulation results show that a network containing 9 clusters and the energy levels of each cluster in shown below for both random clustering and HECH. From Fig-4 the average remaining energy per cluster is shown which is more in symmetric clustering than random clustering and at the same time the energy consumed per cluster is more in random clustering.

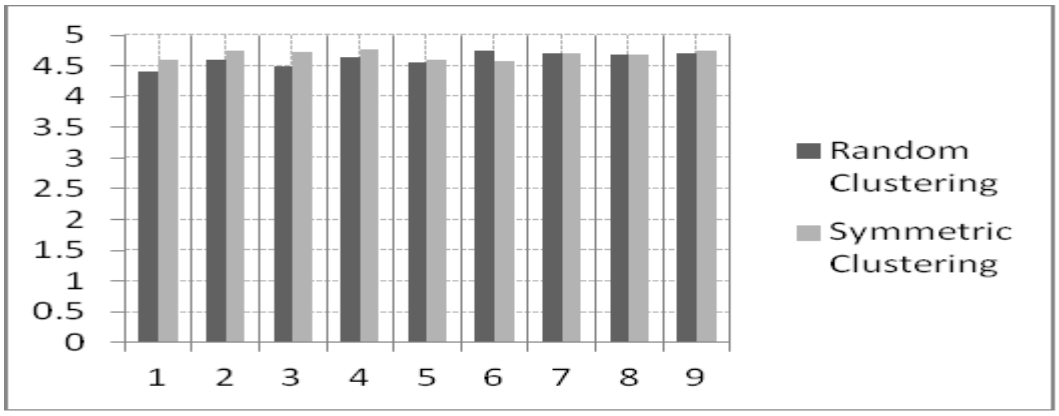


Figure 4:Remaining Energy Levels per Cluster

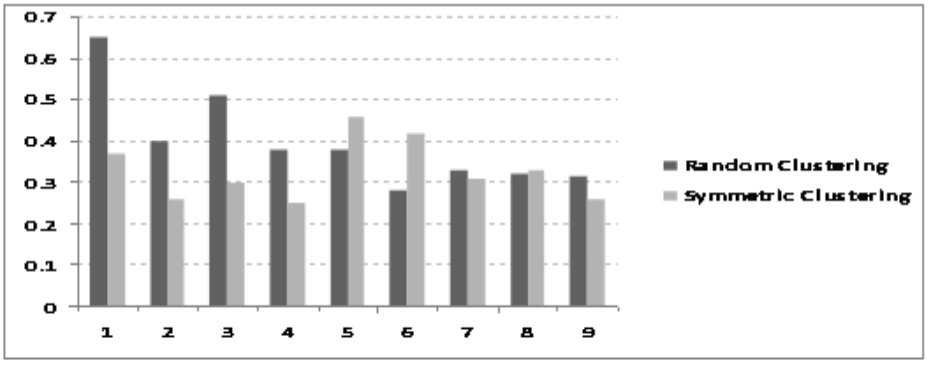


Figure 5:Energy Dissipated per Cluster

.

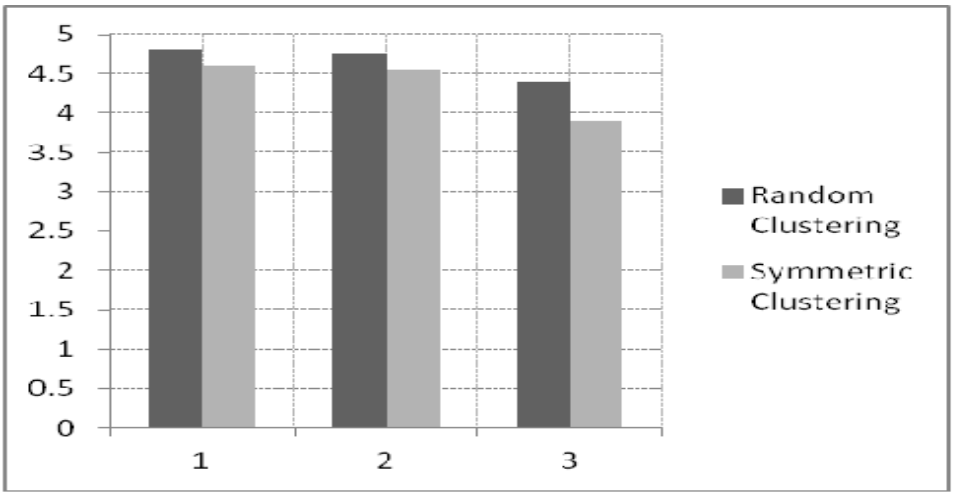


Figure 6:Average Network Energy for Cluster Number=15,12 and LEACH

So our protocol is more efficient in remaining energy than other protocol which we are getting by this simulation results. From figure-5 we will get the energy dissipated per cluster is more in random clustering than symmetric clustering which is shown by the simulation results given below. And the average network energy is compared in the following figure-6 shows average network energy for HECH with 1.CN=15, 2.CN=12 and LEACH .Here we can also see that the number of members in a cluster is an important factor.

VII. Conclusion

Here we proposed a new load balanced and energy-efficient routing protocol HECH outperforms LEACH by a more balanced cluster distribution and by reducing the non uniform cluster topology. It uses the number of cluster members to construct clusters in a certain area. The cluster head for the next round will be elected on the basis of max remaining energy levels of the nodes, so with this cluster topology the network is more energy balanced and net life time of the network increases which shows in the simulation results.

References

[1] Prasant kumar Pattnaik and Rajibmall ,"Fundamentals of Mobile Computing"PHI publication 2012 ,pages 149-169.

[2] Ruay-Shiung Chang and Chia-Jou Kuo. An Energy Efficient Routing Mechanism for Wireless Sensor networks.AINA'06,2006IEEE

[3] Kemel Akhaya and Mohammed younis,"A Survey on energy aware routing protocols in Wireless Sensor Networks".

[4] W.R.Heinzelman, A.Chandrakasan and H.Balkrisnan, "Energy Efficient Communication Protocol for Wireless Micro sensor Networks," Proceedings of the Hawaii International Conference on System Sciences (HICSS), January 2000.

[5] S.K.Padhi, P.K.Pattnaik, B.Puthal,"Review Of Routing Protocols in Sensor and Adhoc Networks", International Journal Of Reviews in Computing, Volume3-2010, page(s)11-17

[6] K. Akkaya, M. Younis, A survey on routing protocols for wireless sensor networks, Elsevier Journal of Ad Hoc Networks 3 (3) (2005),Page(s) 325-349.

[7] Lindesy and C.Raghavendra,"PEGASIS:Power Efficient Gathering in Sensor Information System",Prodeedings of

IEEE Aerospace Conference,Bigsky,Montana,page no.1-6,March 2002.
[8] E.L. Lloyd, Xue Guoliang,"Relay node placement in wireless sensor networks", IEEE Transactions on Computers, Volume 56,Issue 1,2007,Pages:134-138.
[9] Chang Jae-Hwan, L.Tassiulas,"Maximum Lifetime Routing in Wireless Sensor Networks", IEEE/ACM Transactions on Networking, Volume 12,Issue 4,2004,Pages:609-619.
[10] E.L. Lloyd, Xue Guoliang,"Relay node placement in wireless sensor networks", IEEE Transactions on Computers, Volume 56, Issue 1, 2007, Pages: 134-138.
[11] Lusheng-Miao,-karim-Djouani,anish Kurien,G.Noel,"A competing algorithm for gradient based routing protocol in wireless sensor networks",Proceedings of the 2010 Wireless Information networks and Systems(WINSYS),2010,Pages 1-8.
[12] Zheng Rong,"Asymptotic Bounds of Information Dissemination in Power constrained Wireless Networks", IEEE Transactions on Wireless Communications ,Volume 7,Issue 1,2008,Pages 251-259.
[13] Raminder P.Mann,Kamesh R. Namuduri, and Ravi Pendse,"Energy-aware routing protocol for Ad-hoc wireless sensor networks,EURASIP Journal on Wireless Communications and Networking,Vol.5 page(s)635-644,2005.
[14] L.Lin,N.B.Shroff and R.Srikant,"Energy- Aware Routing in Sensor Networks:A Large System Approach,"to appear in Ad Hoc Networks Journal(Elseviar),special issue on Recent Advances in Wireless Sensor Networks 2007.
[15] S.K.Padhi, P.K.Pattnaik,"A novel Distributed protocol For Randomly Deployed Clustered Based Wireless Sensor Network", Journal of Theoretical and Applied Information Technology, Volume15.No.1-2010, pages.39-47
[16] Otgonchimeg Buyanjargal and Youngmi kwon ,"Adaptive and Energy efficient Clustering algorithm for event- driven application in Wireless Sensor Networks(AEEC),"Journal of Networks",Vol.5,No.8,August 2010.
[17] Y.Srinivas,J.Vijayasekhar,Prasad, .R.Satya,Krishna N.Vamsi," Cluster-Id Routing Scheme in wireless SensorNetworks",pageNo.48,volume1,Issue1,IEEE,2011.
[18] Jalil jabari latf,Mehdi Nozad Banab,siavash khorsandi,"A novel Clustered based Routing protocol with extended Life time for Wireless Sensor Networks".IEEE.2008.

Node Reproduction Based Range-free Localization Algorithm in Wireless Sensor Networks

Xiaoming Wu
School of Information Science & Electric Engineering, Shandong Jiaotong University, Jinan, China
Email: sdjtuwxm@163.com

Hua Wu, Yang Liu, Jianping Xing, Mingyue Zhao
School of Information Science & Electric Engineering, Shandong Jiaotong University, Jinan, China
School of Information Science & Engineering, Shandong University, Jinan, China
Cloud, Computing&communications power management, Texas Instruments, USA
Email: wuhua1982111@126.com, ly0314@126.com, sduxingjp@163.com, mzhao@ti.com

***Abstract*—In this paper, a node self-localization algorithm based on node reproduction (NR) for wireless sensor network (WSN) is proposed. This method is adapted to WSN that anchor nodes present a uniform distribution in three dimensional sensor spaces. During the localization process, by reproducing a 3D special node, which is called the reproduced node, unknown nodes can calculate their own positions automatically. This NR algorithm is a three dimensional range-free approach which does not need extra hardware requests and the unknown nodes can calculate reproduced node by position information from three-different anchor nodes. Finally the unknown position coordinates are obtained by the position information of above four nodes. This approach has low complexity and the localization process is much simpler in simulation. The localization error can also reach a low level compared with classic DV-Hop and Centroid which can be found in simulation and the NR algorithm needs the least localization time. In extreme situations, localization error and time are improved by 25%, 84% and 84%, 88% compared with Centroid and DV-Hop algorithms respectively.**

***Index Terms*—WSN, three dimensional sensing space, node reproduction, self-localization**

I. INTRODUCTION

Wireless sensor network (WSN) is an important technology attracting considerable research interest during these years. WSN is composed of thousands of tiny nodes that are deployed in the sensing fields and is widely used in various area, such as intrusion detection, traffic management, space exploration, environmental monitoring, water quality monitoring, precision agriculture design and disaster rescue, etc. Also the advanced hardware design technology have led to the miniaturization of devices which is capable of communication with each other. So these nodes in WSN have limited processing capabilities and energy to be operated. But in these application situations, changing batteries is almost impossible. Meanwhile wireless sensor network is envisioned to allow for the ease of deployment through redundancy and ad-hoc placement. There are several essential issues in wireless sensor networks. Location estimation is one of the most important subjects. Self-localization capability is such a highly desirable characteristic of WSN.

Until now, WSN localization scheme has been widely researched [1-3], a large amount of which can be found in Reference [4] and [5], but there is yet much work to do in the field. So far two main centralized [6-8] and distributed [9-11] localization algorithms have been proposed. All the general localization mechanisms proposed before can be mainly classified as range-based approaches and range-free approaches. The former approaches determine the node position fully based on distance or angular information acquired using the Time of Arrival (TOA), Angle of Arrival (AOA), Time Difference of Arrival (TDOA), or Received Signal Strength Indicator (RSSI) techniques [12-16]. On the contrary, range-free localization schemes merely rely on the existence of radio connectivity to an existing target instead of measuring distance or angle, which decrease the consumption power and hardware requirements [17-20]. Range-free schemes mainly explore the local network topology and the coordinate computation is derived from the locations of the surrounding anchor node position coordinates [21-24].

In this paper, we propose a node self-localization algorithm based on node reproduction (NR) for wireless sensor network (WSN). This method is adapted to WSN that anchor nodes present a uniform distribution in three dimensional sensor spaces. During the localization process, by reproducing a 3D special node, which is called the NR node, unknown nodes can compute their own positions automatically. This NR algorithm is a three dimensional range-free approach which does not need extra hardware requests and the unknown nodes can obtain reproduced node by position information from three different anchor nodes. Finally the unknown position coordinates are obtained by the whole position

doi:10.4304/jcp.9.5.1047-1052

information above. NR algorithm requires no specialized range-determining hardware equipped in the sensors, and relies merely on node-anchor communication to localize the unknown nodes. So it can reduce the computation of the whole network. The main task of the unknown nodes is to listen to the packages which are flooded from the anchor nodes in a fixed time slice. So there is no information exchange between neighboring nodes which is more energy and time efficient. Thus the communication overload is decreased and as a result, the lifetime of the whole network will be prolonged.

This paper is organized as following. Part Ⅱ gives some assumptions of the network model and simulation scenarios. The specific implements of NR are listed in section Ⅲ. Analysis of localization performances and future work are given in part Ⅳ and part Ⅴ respectively. Part Ⅵ depicts the conclusions of the whole work.

II. Assumptions and Network Model

Some assumptions are made as following. The localization space is supposed to be a cube with the edge length 100m which means the whole volume of the 3D localization space is 100×100×100m3. The anchor nodes are in a uniform distribution in sensing cube which is to say there are 216 anchor nodes in all in the localization space. In some application environments, a part of anchor nodes have to be set into space grids to better explore the surrounding circumstance parameters as illustrated above.

Every eight anchor nodes form a cube with the edge length 20m in the space. Unknown nodes are randomly deployed and the number can be changed in this NR method. The other important parameter is the communication range of each unknown node. Once the packages are sent by anchor nodes, enter the communication range of any unknown node, unknown nodes can detect them immediately and record the corresponding anchor information including anchor node ID and the source anchors' coordinate. In the field where sensor nodes are spread the amount of all sensors are very large. In this way the probabilities that not enough anchor nodes form a cube is very low. Once happens the algorithm will go on but with large deviation. But the algorithm will fail if number of anchors is less than 3.

The reproduced node is supposed to be the center of the cube formed by eight anchor nodes. As said above, there is always a cube whose vertexes are anchor nodes. So in this situation, reproduced nodes must exist in this network.

Ⅲ. Realization of NR Algorithm

The core idea of NR algorithm is how to get this reproduced node. This particular process can be found in this part in detail.

A. Definition of Reprodeced Node

In our proposed algorithm, all anchor nodes are supposed to present a uniform distribution in 3D space. Also this sensing space is 100m×100m×100m which means the total number of all anchor nodes is 216. Every eight anchor nodes can establish a small cube with the length of side 20m as shown in Figure1.

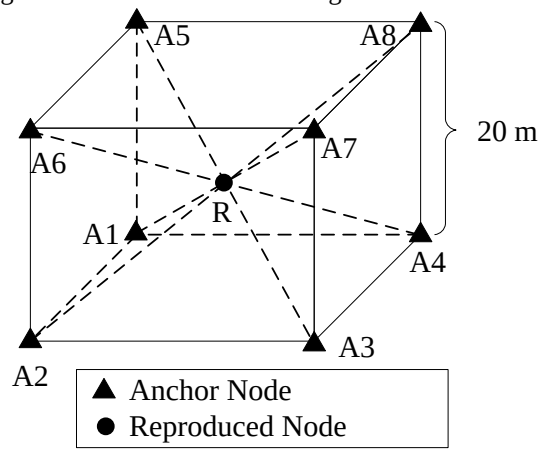


Figure 1. Formation of NR

As shown in Figure 1, A_i (i =1, 2… 8) is an anchor node with known position coordinates. So once the eight anchor nodes are determined, the reproduced node (Node *R* in Figure 1) can also be found out. This node is not only the center but also the centroid of the special cube. Of course this node is not a factual sensor node, but just is reproduced in the sensing space to realize localizing in the next step.

Reproduced node is a key node for realizing the whole localization algorithm. The core problem is how to make it. After the node reproduction is finished, unknown nodes can make self-localization by combining three corresponding anchor nodes where reproduced node is derived from. The position determination of NR algorithm is finished by unknown nodes themselves and the process does not consume too much energy which as a result the lifetime of the network is prolonged. The localization process in detail can be seen in the following *Part B.*

B. Realization of NR Algorithm

In this algorithm each sensor node estimates its position solely based on the information gathered directly from anchor nodes. Since it does not depend on communication between neighboring sensor node, it is independent of network connectivity and is more suitable for all kinds of complicated applications. Now we will describe the NR algorithm in detail. The flow chart of this algorithm is given in Figure 2. The algorithm begins from flooding data packages from anchors to the whole network. Each anchor is able to broadcast information packages periodically. This time slice can be set manually. The data information includes anchor node ID, and coordinates of corresponding anchor nodes. Unknown nodes' task is easy and energy efficient because they are only in charge of listening to these packages in the time slice *T*. Unknown nodes can memory how many packages have received from different anchors. They judge whether the time slice *T* is arrived. If so, the information can be recorded, or go on waiting. As we said in the last part, once the package enters the communication range unknown nodes can detect it immediately and record the information data contained in the packages. At last by the

information provided in these packages, reproduced node is formed and can be computed. Finally unknown position can be derived using the reproduced node above.

Just as the flowchart shows in Figure 2, first a random fixed time slice *T* is produced randomly then unknown nodes begin listening to find out whether there are packages entering its communication range in this predetermined time slice *T* and finally record all the information of three anchor nodes that have sent the most packages. Because three anchors that have sent most packages mean the most nearest to the unknown nodes which can result in least localization error. All the unknown nodes can equally receive packages from all the anchors from the whole network for they are supposed to be the same on the node structure and function. Choosing the three nearest anchors means little localization error. That is because distance error can be accumulated from node to node if too many hops exist.

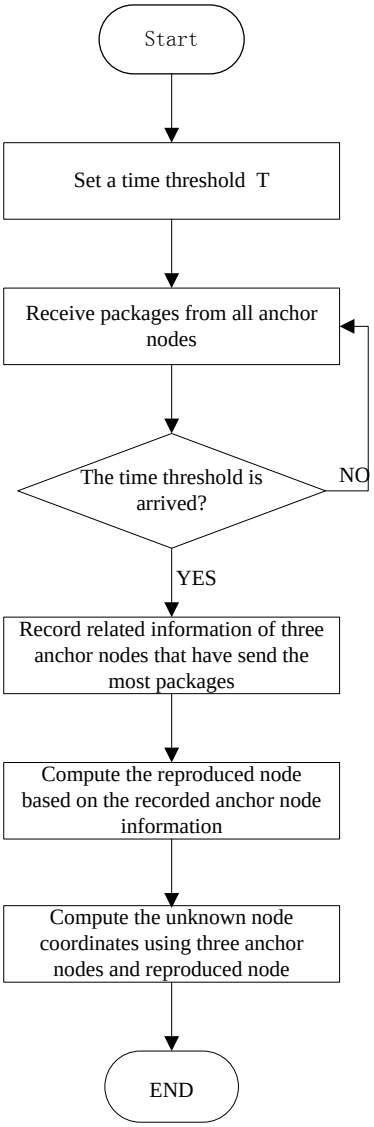


Figure 2. Flow chart of NR algorithm

After the anchor node information is recorded, unknown node begins to analyze that information and finally reproduced node can be determined. Based on the coordinates of three anchor nodes, unknown nodes can deduce a fourth anchor node coordinate that is on the same plane with the three to make up a square with edge length 20m. This step is to fix a universal plane for deducing reproduced node. After the four anchor nodes (three are detected and the other is deduced) are assured, the center of the square can be determined easily by average coordinate value of them. The difference between this center node and reproduced node is that only one ordinate direction is different and the other two is the same (3D coordinates are composed by three coordinate directions).

By using the center of the square, reproduced node can be computed through adding half of communication range on one of three ordinate directions. But there is a problem in this algorithm. As shown in Figure 3, after the fourth node is determined, we cannot make sure the reproduced node is on which side of the plane. It may be on the same side with unknown node which means low estimation error. However if on the opposite side of the unknown node, it will produce localization error. During our localization, the position of reproduced node on which side is decided randomly and it is not a perfect solution induces lots of uncertainties. But how to solve this problem completely is also a research direction in the future to make better localization accuracy.

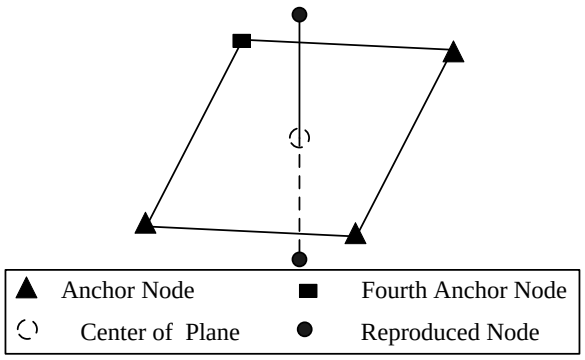


Figure 3. Computation of NR

When the reproduced node is found out, unknown nodes can finish the localization process. As shown in Figure 4, three anchor nodes (*A1*, *A2*, and *A3*) and reproduced node *R* forms a tetrahedron. Of course this tetrahedron is anomalous. The reproduced node could be either side of the plane determined by node *A1*, *A2*, and *A3*. Then we can use the similar way of Centroid algorithm. That is to say the center of the four special nodes is the estimated position of the unknown node. It is necessary to illustrate the feasibility of NR method. As we known, reproduced node is important in this algorithm. Based on this node unknown node could compute the center of a 3D graph.

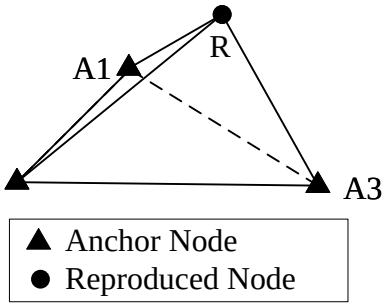


Figure 4. Derivation of unknown node position

In our algorithm, the distances between three anchor nodes are all the same, which is 20m or $20\sqrt{2}$ m. Also the distance between reproduced node and three anchors is $10\sqrt{3}$ m. If we choose the center of the four known nodes as the unknown position, the error could not larger than $10\sqrt{3}$ which means the localization accuracy can be limited below the value above. Simulation results in next part prove this inference.

Ⅳ. Analysis of Localization Performance

Here simulations are made in MATLAB software. In MATLAB we first establish a 3D localization space in which unknown nodes are random deployed for easy use. On the contrary, anchor nodes are deployed in 3D grid. Also in this section we present the performance evaluation of NR as well as the comparisons with classic DV-Hop and Centroid algorithms on estimation error and localization time.

A. Self-performance of NR

Figure 5 shows the accuracy gain of NR algorithm by changing the number of unknown nodes in order to alter the percentage of all anchor nodes in the 3D sensing space. As Figure 5 shows above, the percentage of anchor nodes changes from 5% to 50% in the sensing space and the estimation error is fluctuating between 5m and 8.5m under six different communication ranges. In the figure we can also find that the effects of anchor node number on the estimation error are so small that the fluctuation range of error is less than 1m. Further the six curves are divided into two obvious parts. When the communication range *R* is set as 15m and 20m, the estimation error is much larger than that of the other four simulation curves. When communication range *R* is set as 25m, 30m, 35m, or 40m, the difference is more or less the same. And the fluctuation range is much smaller which is smaller than 2m.

All these performance features are caused by the property of NR itself. Based on finding the special reproduced node, length of communication range for the whole network cover becomes the most critical parameter and the effects of anchor nodes are decreased sharply because in the algorithm their missions are to broadcast beacon information periodically. Also most of the beacon packages from anchors can be sent to unknown nodes efficiently within communication range which assumed in Part Ⅱ. However if the communication range is too small, such as smaller than 10m in the simulating process, the NR algorithm cannot finish the localization task which is proved by the simulation in MATLAB. That is caused for lower communication range means that it cannot cover the whole sensing space which result the destruction of finding corresponding reproduced node. If so lots of information packages from anchors are lost during the transmission.

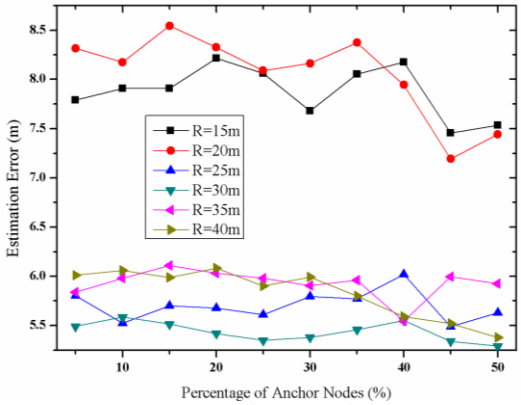


Figure 5. Performance of NR

All these performance features are caused by the property of NR itself. Based on finding the special reproduced node, length of communication range for the whole network cover becomes the most critical parameter and the effects of anchor nodes are decreased sharply because in the algorithm their missions are to broadcast beacon information periodically. Also most of the beacon packages from anchors can be sent to unknown nodes efficiently within communication range which is assumed in Part Ⅱ. However if the communication range is too small, smaller than 10m in the simulating process, the NR algorithm cannot finish the localization task which is proved by the simulation in MATLAB. That is caused for lower communication range means it cannot cover the whole sensing space which results the destruction of finding corresponding reproduced node. Lots of information packages from anchors are lost during the transmission.

However, larger communication hardly means high estimation accuracy. As shown in Figure 5, when communication range *R* is set as 30m, its accuracy is the best. Because the distance between any two anchors is 20m as given in Figure 1, so *R*=30m means unknown nodes can hear most packages from anchors and the fields where information packages from anchors achieves can increase the cover level of all the localization space. When *R* is larger than 30m, the packages from different anchors result in confliction which makes some unknown nodes can hear packages which cannot achieve the best situation.

B. Localization Accuracy Comparisons of NR, DV-Hop, Centroid

In this part comparisons of NR algorithm with classic DV-Hop and Centroid algorithms are given.

Figure 6 and Figure 7 both shows the position estimation errors as a function of the percentage of anchor nodes deployed. We can find that DV-Hop is the worst of the three on the estimation accuracy. Its

localization error is the biggest for too many hop counts induce lots of errors. Localization accuracy of DV-Hop totally depends on average hop size of the network which induces lots of uncertainty which results in higher estimation error.

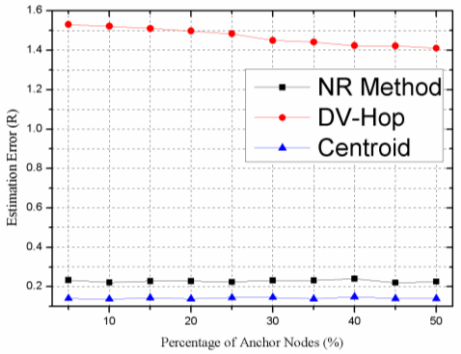


Figure 6. Localization Accuracy Comparisons with R=25m

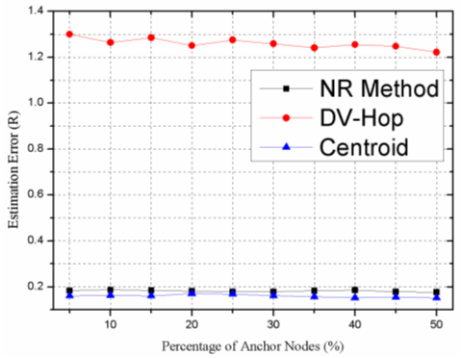


Figure 7. Localization Accuracy Comparisons with R=30m

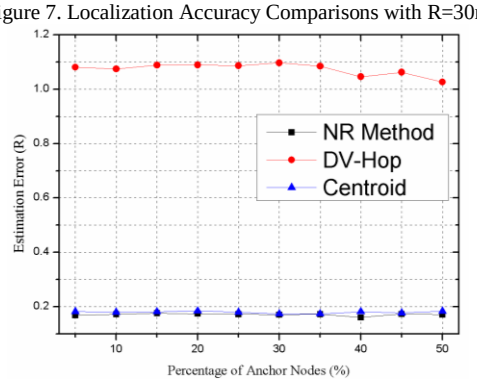


Figure 8. Localization Accuracy Comparisons with R=35m

Our proposed NR and Centroid are much closer to some extent. On the whole Centroid is a little better than NR. Their accuracy can both achieve below 0.25*R*. Especially when *R*=30m, their errors are below 0.2*R*. The two algorithms both have much little relationship with number of anchor nodes as analyzed in the last part but they are critical with communication range.

In Figure 8, communication range *R* is increased to 35m. In this figure, DV-Hop still stands for highest estimation error. However curves of NR and Centroid are almost superposition and all below 0.2*R* which is caused by large communication range *R*. Differently in Figure 9, NR is better than Centroid with error as low as 0.15*R*. So we can find larger communication range under certain of range stands for higher accuracy as also analyzed in last part.

From all the given simulation figures, DV-Hop is the worst on the localization error no matter how the parameters are changed and always stays at a high level. NR and Centroid are more or less the same on the accuracy. The performances of DV-Hop are totally related to average hop size, hop counts and network size which easily induce a lot of error. When communication range is large enough to 30m, NR will be better than Centroid. So in this situation whole network can be covered completely by beacons from anchor nodes. Packages with information in them from anchors can better enter the scope of unknown nodes and reproduced nodes are determined more accurately which means higher education. In all, all of the three localization algorithms are less affected by anchor nodes which are decided by the inner property of those themselves.

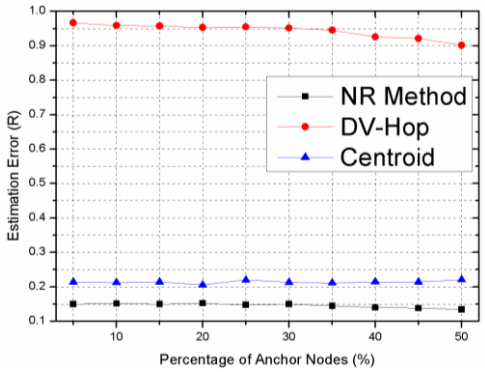


Figure 9. Localization Accuracy Comparisons with R=40m

C. Localization Time Comparisons of NR, DV-Hop, Centroid

Here we record the average localization time with *R*=15m, 20m, 25m, 30m, 35m, and 40m of the three respectively. The localization time is a function of percentage of anchor nodes which is shown in Figure 10.

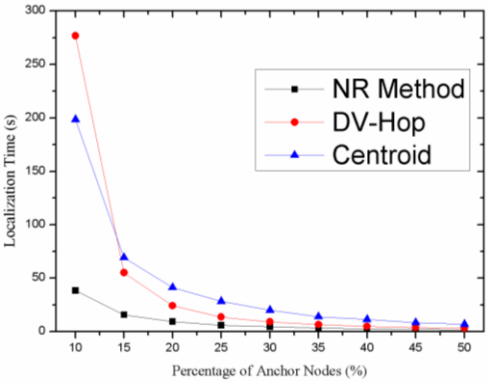


Figure 10. Localization Time Comparisons

The three curves all present a downward tendency. From Figure 10 we can find the proposed NR needs the least localization time no matter how many anchors there are in the 3D space. The lowest point is about 1.5 second and the highest point is 38 second which is both much lower than the other two. The centroid algorithm costs the most although it has the lowest localization accuracy in some situations. DV-Hop is between the two. The two all cost more time than NR to finish a localization task. In DV-Hop algorithm, nodes have to compute average hop distance which costs lots of time and energy. However our NR and Centroid do not need to make this procedure. They just need receive information packages which

reduce time and energy sharply. All in all NR and Centroid is easier in the designing aspect.

We can say our proposed reproduced NR localization algorithm is the most suitable for the rapid positioning situations for there is no complicated computation. So in the future we can think about the use of NR in mobile node scenes in WSN.

Ⅴ. CONCLUSIONS

In this paper, we propose a NR localization algorithm for wireless sensor networks. NR method is a distributed range-free approach which does not require information exchange between neighboring sensors. It has low computation overhead and is simple to implement. NR needs anchor nodes in WSN to flood beacon information periodically which allows each unknown node to record the recording information to realize self-localizing. The estimated position of the unknown node is taken from reproduced node and three other anchor nodes.

Simulation results show NR provides lower localization error than classic DV-Hop and needs little time than DV-Hop, too. Its localization accuracy is similar with that of Centroid algorithm but costs much less time than Centroid which means NR is more suitable for the mutative situations than the other two algorithms in WSN.

REFERENCES

[1] Zhang Youtao, Yang Jun, Li Weijia, Wang Linzhang, Jin Lingling. “An authentication scheme for locating compromised sensor nodes in WSNs,” Journal of Network and Computer Applications, v 33, n 1, p 50-62, January 2010.

[2] Wan Jing; Ghosh, R.K.; Das, Sajal K. “A survey on sensor localization,” Journal of Control Theory and Applications, v 8, n 1, p 2-11, January 2010.

[3] Kulkarni, Raghavendra V; Venayagamoorthy, Ganesh Kumar. “Particle swarm optimization in wireless-sensor networks: A brief survey,” IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, v 41, n 2, p 262-267, March 2011.

[4] Kulkarni, Raghavendra V.; Förster, Anna; Venayagamoorthy, Ganesh Kumar. “Computational intelligence in wireless sensor networks: A survey,” IEEE Communications Surveys and Tutorials, v 13, n 1, p 68-96, First Quarter 2011.

[5] S. Gezici. “A survey on wireless position estimation,” Wireless Personal Communications, vol. 44, pp. 263-282, 2008.

[6] G. Mao, B. Fidan, B.D.O. Anderson. “Wireless sensor network localization techniques,” Computer Networks 51 (10) 2529-2553, 2007.

[7] Ding, Yingqiang; Han, Gangtao; Mu, Xiaomin. “A distributed localization algorithm for wireless sensor network based on the two-hop connection relationship”. Source: Journal of Software, v 7, n 7, p 1657-1663, 2012.

[8] Erdogan, Ayhan; Coskun, Vedat; Kavak, Adnan. “The sectoral sweeper scheme for wireless sensor networks: Adaptive antenna array based sensor node management and location estimation,” Wireless Personal Communications, v 39, n 4, p 415-433, December 2006.

[9] Chen, Hongwei; Zhang, Chunhua; Zong, Xinlu; Wang, Chunzhi. “LEACH-G: An optimal cluster-heads selection algorithm based on LEACH”. Source: Journal of Software, v 8, n 10, p 2660-2667, 2013.

[10] Wang, Tsang-Yi; Han, Yunghsiang S.; Varshney, Pramod K.; Chen, Po-Ning. “Distributed aggregate algorithm for average query based on WSN,” Journal on Selected Areas in Communications, v 23, n 4, p 724-733, April 2005.

[11] D. Niculescu and B. Nath. “Ad-hoc positioning system,” in Proc. of IEEE Globecom, San Antonio, TX, Nov. 2001.

[12] Chen, Zuo; Chen, Kai. “An improved multi-hop routing protocol for large-scale wireless sensor network based on merging adjacent clusters”. Source: Journal of Software, v 8, n 8, p 2080-2085, 2013.

[13] J. Caffery Jr. and G.L. Stuer., “Subscriber Location in CDMA Cellular Networks,” IEEE Trans. Vehicular Technology, vol. 47, no. 2, pp. 406-416, May 1998.

[14] R. Klukas and M. Fattouche, “Line-of-Sight Angle of Arrival Estimation in the Outdoor Multipath Environment,” IEEE Trans. Vehicular Technology, vol. 47, no. 1, pp. 342-351, Feb. 1998.

[15] L. Cong and W. Zhuang, “Hybrid TDOA/AOA Mobile User Location for Wideband CDMA Cellular Systems,” IEEE Trans. Wireless Comm., vol. 1, no. 3, pp. 439-447, July 2002.

[16] J. Aspnes, T. Eren, D. Goldenberg, A.S. Morse, W. Whiteley, Y. Yang, B.D.O. Anderson, P. Belhumeur. “A theory of network localization,” IEEE Transactions on Mobile Computing 5 (12) 1663–1678, 2006.

[17] Wang Yun, Wang Xiaodong, Wang Demin, Agrawal Dharma P. “Range-free localization using expected hop progress in wireless sensor networks,” IEEE Transactions on Parallel and Distributed Systems, v 20, n 10, p 1540-1552, 2009.

[18] Li Mo, Kowloon, Liu Yunhao. “Rendered path: Range-free localization in anisotropic sensor networks with holes,” IEEE/ACM Transactions on Networking, v 18, n 1, p 320-332, February 2010.

[19] Chan, Yiu Wing Edwin; Soong, Boon Hee. “A new lower bound on range-free localization algorithms in wireless sensor networks,” IEEE Communications Letters, v 15, n 1, p 16-18, January 2011.

[20] Wang, Sheng-Shih; Shih, Kuei-Ping; Chang, Chih-Yung. “Distributed direction-based localization in wireless sensor networks,” Computer Communications, v 30, n 6, p 1424-1439, March 26, 2007

[21] A. Savvides, W.L. Garber, R.L. Moses, M.B. Srivastava. “An analysis of error inducing parameters in multihop sensor node localization,” IEEE Transactions on Mobile Computing 4 (6) 567–577, 2005.

[22] F.K.W. Chan, H.C. So, “Efficient weighted multidimensional scaling for wireless sensor network localization,” IEEE Transactions on Signal Processing 57 (11) 4548–4553, 2009.

[23] J.A. Costa, N. Patwari, A.O. Hero. “Distributed weighted-multidimensional scaling for node localization in sensor networks,” ACM Transactions on Sensor Networks 2 (1) 39–64, 2006.

[24] Zhang Shigeng, Cao Jiannong, Li-Jun Chen, Chen Daoxu. “Accurate and Energy-Efficient Range-Free Localization for Mobile Sensor Networks,” IEEE Transactions on Mobile Computing, v 9, n 6, p 897-910, June 2010.

A New Revocation Method for Standard Model Group Signature

Xiaogang Cheng[1, 2]
[1]College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, China
[2]College of Computer Science and Technology, Huaqiao University, Quanzhou, China
Email: cxg@hqu.edu.cn

Jian Wang[1] and Jixiang Du[2]
[1]College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, China
[2]College of Computer Science and Technology, Huaqiao University, Quanzhou, China

***Abstract*—Membership revocation is practically necessary for group signature schemes. At present, most revocable group signature schemes are based on ROM model and many standard model group signature schemes do not support revocation. We present a novel and general revocation method for standard model group signature scheme based on length-reducing commitment and Groth-Sahai proof system, and demonstrate its usage by adding revocation capability to the Groth's full secure group signature scheme. The new method supports two different kinds of linkability for members when signing.**

***Index Terms*—membership revocation, group signature, standard model, dynamic accumulator, linkability**

I. Introduction

Group signature was introduced by Chaum and Heyst in 1991 [1]. Since it combined anonymity and traceability, it soon became a central cryptographic primitive, of which many applications had been found such as e-cash, e-voting and trust computing etc.

As pointed out in [2], one of the major issues for group signature's practical application is the ability to revoke a member when he becomes malicious or when he leaves the group deliberately.

Currently most revocable group signature schemes are based on ROM model such as [3,4,5]. Since ROM model can only provide heuristic security, recently there are many works on constructing cryptographic schemes without ROM such as [6,7]. For standard model group signature schemes such as [8,9,10], revocation is not supported. In [11], an ID-based revocable group signature was presented. However the drawback of the scheme in [11] is that the GM (Group Manager) has to be online when signing.

In this paper, we present a novel and general revocation method for standard model group signature, which is based on a length-reducing commitment scheme [12] and Groth-Sahai proof system [13], and demonstrate its usefulness by adding revocation ability to the well-known Groth's full secure standard model group signature scheme [10].

We note that recently in [14], a highly efficient revocable group signature scheme in standard model was introduced, which was based on broadcast encryption and concise vector commitment. But their revocation is based on assumptions like FlexDHE etc., while our scheme is based on the standard DLIN assumption. Moreover, their method is not easily adapted to other standard model group signatures like [8,9,10], since they use broadcast encryption techniques which is a departure from traditional group signature construction. Our revocation method is conceptually simpler and can be used in more scenarios such as [8, 9].

The general idea is simple. GM publishes a short commitment which contains all the public keys of legal members and gives each member a witness. Then when making group signatures, legal member has to prove that his public key is included in the commitment using the witness he got from the GM. We use Groth-Sahai NIWI/NIZK proof system to achieve anonymity. To revoke a member, GM makes another commitment excluding the member's public key. Hence this member lost its signing right. The whole idea is like the DA (Dynamic Accumulator) revocation method in ROM model [3]. We realize it in standard model, although the cost is not as efficient as the ROM DA method.

This paper is organized as follows. In section II we introduce the tools we used for the construction of this new method. The concrete construction is given in section III. In section IV we analyze its securities, and efficiency issues are discussed and compared in section V. Then we make a conclusion in section VI.

II. Preliminaries

In this section we introduce assumptions and cryptographic tools used for the construction of our new revocation method.

Bilinear groups: Let G_1, G_2 and G_T be groups of order p. A bilinear map e: $G_1 \times G_2 \rightarrow G_T$ must satisfy the following:

(1) For arbitrary $a \in G_1$, $b \in G_2$ and $x \in Z_p$, $y \in Z_p$, $e(a^x, b^y) = e(a,b)^{xy}$.

doi:10.4304/jcp.9.5.1053-1057

(2) *e(g, h)* generates G_T whenever *g* generates G_1 and *h* generates G_2.

(3) There is an efficient algorithm to computer *e(a, b)* for any $a \in G_1$ and $b \in G_2$.

If $G_1=G_2$, we call them symmetric bilinear groups, and use the symbol *G* for both G_1 and G_2. If $G_1 \neq G_2$ and there is no efficiently computable non-trivial homomorphism between them, then we call them asymmetric bilinear groups.

Definition 1. DLIN (Decisional Linear) assumption: In symmetric bilinear groups *G*, G_T of order *p*, given a generator *g* of *G*, and tuple $(g^a, g^b, g^{ac}, g^{bd})$ where *a, b, c, d*$\in Z_p^*$ are random, it is hard to distinguish between a random element $T \in G$ and $T=g^{c+d}$.

Definition 2. DBP (Double Paring) assumption: In asymmetric bilinear groups G_1, G_2, G_T of order *p*, given random elements $G_r \in G_1$, $G_s \in G_1$, it is hard to find non-trivial $R \in G_2$, $S \in G_2$ satisfying $e(G_r, R)e(G_s,S)=1$.

Definition 3. SXDH (Symmetric eXternal Diffie-Hellman) assumption: In asymmetric bilinear groups G_1, G_2, G_T of order *p*, DDH (Decisional Diffie-Hellman) assumption is hard in both G_1 and G_2.

1. Homomorphic, Trapdoor and Length-reducing Commitment

In [12], the authors put forward a homomorphic, trapdoor and length-reducing commitment scheme, which is proved to be both hiding and binding based on the DBP (Double Pairing) assumption.

To set up the scheme, generate groups with bilinear maps $e: \mathbb{G}_1 \times \mathbb{G}_2 \rightarrow \mathbb{G}_T$. The public key for this commitment scheme is $ck=(G_R, G_1=G_R^{x1}, \ldots, G_n=G_R^{xn}) \in \mathbb{G}_1^{n+1}$, the trapdoor is $tk=(x_1, \ldots, x_n)$. To commit to $(M_1, \ldots, M_n) \in \mathbb{G}_2^{n+1}$, return $C = e(G_R, R)\prod_{i=1}^{n} e(G_i, M_i)$. To open this commitment, just give out *R* and $(M_1, \ldots, M_n)$. To open to a different message tuple $(M_1', \ldots, M_n')$, compute $R' = R\prod_{i=1}^{n} (M_i / M_i')^{x_i}$ using *tk*. This trapdoor opening is valid since:

$$e(G_R, R')\prod_{i=1}^{n} e(G_i, M_i') = e(G_R, R)\prod_{i=1}^{n} e(G_i, M_i) = C$$

2. Groth-Sahai Proof System

Groth-Sahai proof [13] is the first efficient NIWI/NIZK proof system for a large class of quadratic equations in bilinear groups, esp. in standard model without ROM. It can be realized in a couple of different assumptions such as DLIN, SXDH and subgroup decision etc. Here we introduce the instantiation based on DLIN assumption.

To set up the proof system, generate prime order groups with bilinear map $e: G \times G \rightarrow G_T$. And set the CRS (Common Reference String) as $\vec{f}_1, \vec{f}_2, \vec{f}_3 \in G^3$, where $\vec{f}_1 = (f_1, 1, g), \vec{f}_2 = (1, f_2, g)$ for $f_1, f_2 \in G$. To commit to a $X \in G$, compute $C = (1,1,X) \odot \vec{f}_1^{\,r} \odot \vec{f}_2^{\,s} \odot \vec{f}_3^{\,t}$ with $r, s, t \leftarrow Z_p^*$, where $\odot$ stands for component wise product. $\vec{f}_3$ can be set in two different yet indistinguishable ways, which give perfect sound setting and WI setting respectively. In the perfect sound setting, set $\vec{f}_3 = \vec{f}_1^{\,\zeta_1} \odot \vec{f}_2^{\,\zeta_2}$ for $\zeta_1, \zeta_2 \in Z_p^*$. So $C = (1,1,X) \odot \vec{f}_1^{\,r} \odot \vec{f}_2^{\,s} \odot \vec{f}_3^{\,t}$ is a DLIN encryption of *X*, and can be decrypted by using $\beta_1 = \log_g f_1, \beta_2 = \log_g f_2$. In the WI setting, $\vec{f}_1, \vec{f}_2, \vec{f}_3$ are linear independent vectors and *C* is a perfectly hiding commitment. Under DLIN assumption, these two different are computationally indistinguishable.

To commit an element *x* in Z_p, compute $C = \psi^x \odot \vec{f}_1^{\,r} \odot \vec{f}_2^{\,s}$ for $r, s \leftarrow Z_p^*$, with a CRS including $\psi, \vec{f}_1, \vec{f}_2$. Similarly as above, ψ can be set up with two different ways to achieve WI and soundness setting respectively. For soundness setting $\psi, \vec{f}_1, \vec{f}_2$ are linear independent. For WI setting, set $\psi = \vec{f}_1^{\,\zeta_1} \odot \vec{f}_2^{\,\zeta_2}$, to give a perfectly hiding commitment since in this scenario *C* is always a DLIN encryption of 1, no matter what the *x* is.

To prove the committed variables satisfy a set of quadratic equations, the prover generates one proof element (which may include a couple of group elements) per equation in a way so that NIWI/NIZK is achieved.

Such NIWI/NIZK proofs are available for PPE (Pairing Production Equations) and multi-exponential equations. Here PPE means:

$$\prod_{i=1}^{n} e(A_i, X_i) \cdot \prod_{i=1}^{n} \cdot \prod_{j=1}^{n} e(X_i, X_j)^{a_{ij}} = t_T$$

For variables $X_1, \ldots, X_n \in G$ and constants $t_T \in G_T$ and $A_1, \ldots, A_n \in G$, $a_{ij} \in Z_p$.

And multi-exponential equations are:

$$\prod_{i=1}^{m} A_i^{y_i} \cdot \prod_{j=1}^{n} X_j^{b_j} \cdot \prod_{i=1}^{m} \cdot \prod_{j=1}^{n} X_i^{y_i \gamma_{ij}} = T$$

For variables $X_1, \ldots, X_n \in G$, $y_1, \ldots, y_m \in Z_p$ and constants $T, A_1, \ldots, A_m \in G$ and $\gamma_{ij}, b_1, \ldots, b_n \in Z_p$.

3. SPS based on DLIN

SPS (Structure Preserving Signature) means that the signature, public key and message are all elements of a bilinear group, and the verification procedure is nothing but checking a couple of PPEs (Pairing Product Equations). The purpose of SPS is to combine with Groth-Sahai proof system [13] to prove that the prover hold a valid signature/message pair in a zero-knowledge way. So SPS can be used in many anonymous and privacy-preserving scenarios.

In [17], a SPS scheme based on Decision Linear assumption was presented. The SPS is based on binary Merkle tree [18] by transforming a one-time signature scheme. On the tree, every node has its own public/private key pair *(vk, sk)*, and parent node's private key is used to sign the public key of child node to authenticate the child node. The leaf node is used to sign messages. The leaf node's public key is authenticated by the nodes on the path from this leaf node up to the root. To verify a message/signature pair, just verify the d (Height of the tree) one-time signatures on path from root to the leaf.

Since it is structure preserving, we can prove that we hold a message/signature pair anonymously by using Groth-Sahai's NIWI/NIZK proof system. It was shown in [17] that signature verification is up to verify a set of PPEs:

$$S_{vk,M}^{TSig} = \{OR(S_{L,i}, S_{R,i})\}_{i\in[d]}$$

where $S_{D,i}$ (for $D\in \{L,R\}$ and $i\in [d]$) is

$$S_{D,i} = \{(\prod_{j=1}^{8} E(U_i, m_{i,j})) \cdot E(G, s_i) \cdot E(H, t_i)$$
$$= E(X_{D,i}, z_{D,i})\}$$

Intuitively this set of PPEs proves that from the root to a leaf, the prover has all the one-time signatures and thus he holds a valid message/signature pair according to the public key of the SPS scheme.

III. CONSTRUCTION

1. Weak Commitment Scheme

The homomorphic, trapdoor and length-reducing commitment scheme mentioned above work in asymmetric bilinear groups, in which G_1, G_2 are different and there are no efficiently computable homomorphism between G_1 and G_2. And its security is based on DBP assumption, which is implied by DDH assumption in G_1 [12].

In our construction we use a variant of this commitment scheme, which works in symmetric bilinear groups, where DDH is easy in the based group G. This variant cannot be proved to be a secure commitment scheme based on any standard assumption. But we do prove that it satisfies a weaker security property which is sufficient for our revocation purpose based on DLIN assumption, as shown in the next section.

To set up this weak commitment scheme, generate prime order bilinear groups G, G_T with $e: G \times G \rightarrow G_T$. The public key is $ck=(G_R, G_R^{x1}, ..., G_R^{xn}) \in G^{n+1}$, where $G_R \leftarrow G$ and $x_1, ..., x_n \leftarrow Z_p^*$. The trapdoor is $tk=(x_{1,...,} x_n)$. To commit to messages $(M_1,...,M_n) \in G^n$, one computes $C = e(G_R, R)\prod_{i=1}^{n} e(G_i, M_i)$. To open this commitment, the committer just gives out R and $(M_1,...,M_n)$. To open to a different message tuple $(M_1',...,M_n')$, compute $R' = R\prod_{i=1}^{n}(M_i / M_i')^{x_i}$ using tk. This trapdoor opening is valid since:

$$e(G_R, R')\prod_{i=1}^{n} e(G_i, M_i') = e(G_R, R)\prod_{i=1}^{n} e(G_i, M_i) = C$$

2. Revocation to the Groth's Full Secure Group Signature

We introduce our new revocation method by adding revocation ability to the well-known Groth's full secure group signature scheme [10]. The main revocation related operations are the following:

Setup: Generate groups with bilinear map $e: G \times G \rightarrow G_T$, and $(g^r, g^{x1}, g^{x2}, g^{x3}, ..., g^{xn}) = (G_R, G_1, G_2, G_3, ..., G_n) \in G^{n+1}$. To commit to $(V_1, V_2, ..., V_n) \in G^n$, let $C = e(G_R, R)\prod_{i=1}^{n} e(G_i, V_i)$. Here the $(V_1, V_2, ..., V_n)$ are public keys of legitimate members and C is the public commitment value. But at the beginning nothing is committed so $C=e(G_R, R)$ where R is a random element in G. The secret key for revocation is $(R, r, x_1, x_2, x_3, ...x_n)$. GM also maintains two lists which are empty at the beginning: A-List and D-List for recording new added user and deleting user's information. Non-revoked members can update their witness according to these two lists.

Generate a SPS scheme public/private key pair, which is introduced above, to sign $G_1, G_2, G_3, G_4, ..., G_n$. The purpose of these signatures is for set membership proof as in [19]. I.e. a user can commit to a $G_i \in \{G_1, G_2, G_3, G_4, ..., G_n\}$, to prove that the committed value is indeed one of the $G_1, G_2, G_3, G_4, ..., G_n$, he can show that he holds a SPS signature of the committed value by Groth-Sahai's NIWI/NIZK proof system. Since the SPS we used is based on Merkle tree, the length and cost of the signature are $O(logN)$, where N is the total number of users.

Join: After the joining process of the Groth's group signature scheme, GM adds the user's public key V_k to the public commitment C by computing $C=C \bullet e(G_k, V_k)$. Thus the new user's public key is accumulated into the public commitment C. Then the GM has to give the user the witness that his public key is indeed in C. GM computes

$$e(G_R, R)\prod_{i=1,i\neq k}^{n} e(G_i, V_i) = e(g, R^r)\prod_{i=1,i\neq k}^{n} e(g, V_i^{x_i})$$
$$= e(g, R^r \prod_{i=1,i\neq k}^{n} V_i^{x_i})$$

and lets $W_k = R^r \prod_{i=1,i\neq k}^{n} V_i^{x_i}$ and gives W_k to User_k as the witness. User_k can verify the witness by checking if $e(g, W_k) \bullet e(G_k, V_k) = C$. Later this user has to prove he is not revoked by showing that he holds such a witness in a zero knowledge way. GM also add V_k^{xk} to the A-List, then non-revoked user can update his witness by $W_j = W_j V_k^{xk}$.

Sign: Before signing, member should update his witness according to A-List and D-list. Then besides the normal signature of the Groth's scheme, User_k has to prove he is not revoked by showing that his public key V_k is in C. This can be done by proving that he has a W_k so that:

$$e(g, W_k) \bullet e(G_k, V_k) = C$$

Here V_k is user's public key such that $V_k = g^{yk}$.

User_k can prove his membership in two different ways with different levels of privacy, i.e. linkable and unlinkable. One is proving he is the K^{th} user by $PK\{(W_k, V_k): e(g, W_k) \bullet e(G_k, V_k) = C\}$, so verifier can know that the signer is the K^{th} user. This way is very efficient, signing cost is constant $O(1)$, and the verification cost is also $O(1)$.

To attain stronger anonymity, the other way is to show $PK\{(W_k, V_k, G_k): e(g, W_k) \bullet e(G_k, V_k) = C$ *and* G_k *is one of the* $G_1, ..., G_n\}$. Note that this time the user doesn't make the G_K public. Instead he makes a commitment to the G_K, then he prove that it is one of $G_1, G_2, G_3, G_4, ..., G_n$ by showing he hold a SPS signature on the committed value. This is the set membership proof trick we mentioned above. But this time the signing and verification cost is $O(logN)$, since the SPS is based on Merkle tree and the height of this tree is $logN$. Note that this stronger

anonymity can also be done by using standard model ring signature techniques such as [15,16], instead of this set membership proof trick. But the performance could be worse to $O(N)$ or $O(N^{1/2})$.

Note that in these proofs, the V_k should be the same with the one used in normal Groth's group signature. This can be done easily by using the same commitment to V_k.

Verify: Verifier should get the latest public commitment C. Then just verify if the ZK proof above is valid or not, after the normal verification procedure of the Groth's scheme.

Revoke: GM can revoke the signing right of User_k by updating the public commitment C with $C=C/e(G_k,V_K)$. Of course the non-revoked user has to update their witness too.

For this purpose, GM adds the value V_k^{xk} to the D-List, now non-revoked user j can update his witness by $W_j=W_j/V_k^{xk}$.

IV. Security

As mentioned above, the trapdoor homomorphic commitment scheme is proved to be secure based on DBP (Double Pairing) assumption.

It is clear that DBP works in asymmetric bilinear groups. We used a variant of the commitment schemes, which worked in symmetric bilinear groups. But we cannot prove its security based on any standard assumption. But we do prove that the commitment scheme satisfy a weaker security based on ODBP (One side DBP) assumption, which is implied by the well-known standard DLIN assumption.

Definition 4. ODBP (One Side DBP assumption) assumption: In symmetric bilinear groups G, G_T. Given G_R, $G_1 \in G$. For a specific random value $V \in G$, it is hard to find $R \in G$ satisfying $e(G_R, R)e(G_1, V)=1$.

Lemma 1. Based on DLIN, the ODBP assumption is hard.

Proof: For a specific value V, it is hard to come up with R to satisfy $e(G_R, R)e(G_1, V)=1$. Suppose an adversary A can break it, we show how to use this to solve DLIN.

We get a DLIN instance $(g, g^a, g^b, g^{ac}, g^{bd}, T)$ to judge whether $T=g^{c+d}$ or T is random. First set $G_R=g^a$, $G_1=g^{ac}$, using A we can get a R_1 s.t. $e(g^a, R_1)e(g^{ac},V)=1$, This means $R_1{}^a V^{ac}=1$, i.e. $R_1 V^c=1$, $R_1=V^{-c}$. Similarly with g^b and g^{bd}, we can get a $R_2=V^{-d}$. So $R_1R_2=V^{-c-d}=(V^{-1})^{c+d}$. With this value plus a single paring computation we can distinguish whether T is g^{c+d} or not (Since this is a Decisional Diffie-Hellman problem which is easy in a group with symmetric bilinear map.) □

Lemma 2. Based on ODBP, the weak-commitment cannot be opened to a chosen message.

Proof: Suppose it is not the case, i.e. an adversary A can find R' for a chosen T satisfy $e(G_R,R)e(G_1,V_1)=e(G_R,R')e(G_1,T)$. We show how to use the adversary A to solve the ODBP problem.

Given an ODBP instance to find R_x for a specific X satisfying $e(G_R,R_x)e(G_1,X)=1$. Set $C= e(G_R,R)e(G_1,V_1)$ for random R and V_1, use A to open C to $Z=V_1/X$, i.e. $C=e(G_R,R)e(G_1,V_1)=e(G_R,R_z)e(G_1,Z)$. Manipulate this equation a little we get $e(G_R, R/R_z)e(G_1,V_1/Z)=1= e(G_R, R/R_z)e(G_1,X)$. So it is obvious that $R_x= R/R_z$ is what we get for the ODBP solution. □

Theorem 1. Based on ODBP, revoked user in the above scheme can no longer generate valid group signatures.

Proof: From Lemma 2, we see that the commitment cannot be opened to a chosen different message. But for a revoked user, his public key V_i is deleted from the accumulator commitment C. For legitimate signature, he needs an opening of C to his public key V_i, which is generated by him and the GM in the joining protocol. But as discussed above, this is impossible as long as the ODBP assumption holds. □

V. Efficiency

See Table I for a comparison of efficiency between different revocable group signatures. Note that in our scheme, verification of linkable signature takes $O(1)$ time, while unlinkable signature takes $O(logN)$ time. It can be seen that our scheme is not very efficient compared with other revocable signatures. But it is the first general way of revocation for standard model group signature and based on standard assumption DLIN, and it can support two different kinds of linkability when signing, which could be an advantage in some scenarios such as trust computing [20].

TABLE I.

Cost of Revocable Group Signatures

RGS	Sign	Verify	History	Public Key	Model
VLA [4]	O(1)	O(R)	No	O(1)	ROM
DA [3]	O(1)	O(1)	Yes	O(1)	ROM
NFHF09 [5]	O(1)	O(1)	No	O(N)	ROM
LPY12 [14]	O(1)	O(1)	No	O(logN)	Std
Ours	O(1)/ O(logN)	O(1)/ O(logN)	Yes	O(N)	Std

VI. Conclusion and Further Research

In this paper, we present a novel and general method for the revocation of membership in standard model group signature. The idea is to mimic DA method which woks in ROM model, though ours works in standard model. But our scheme is not as efficient as the DA method, esp. the public key size is $O(N)$, while it is $O(1)$ in DA. Further research is needed to reduce the public key size to constant first. Second, our scheme is history-dependent, which could be inefficient when lots of joining and revocation happen. A focus of further research could be carried out to remove the dependent on history, while still maintain the generality of our scheme.

ACKNOWLEDGMENT

This work is supported by the grant of the National Science Foundation of China (No.61175121), the grant of the National Science Foundation of Fujian Province (No.2013J06014), Promotion Program for Young and Middle-aged Teacher in Science and Technology Research of Huaqiao University (No.ZQN-YX108). We also thank the anonymous reviewers for helpful advice.

REFERENCES

[1] D. Chaum and E. Van Heyst. “Group Signatures.” Advances in Cryptology - Eurocrypt 91, vol. 547, pp. 257-265, 1991.

[2] G. Ateniese and G. Tsudik. “Some open issues and new directions in group signatures.” In Financial Cryptography, volume 1648 of LNCS, pages 196–211, 1999.

[3] J. Camenisch and A. Lysyanskaya. “Dynamic accumulators and application to efficient revocation of anonymous credentials.” Advances in Cryptology-CRYPTO 2002, vol. 2442 of LNCS, Springer-Verlag, pp. 61-76, 2002.

[4] D. Boneh and H. Shacham. “Group signatures with verifier-local revocation.” In: Proc.11th ACM Conference on Computer and Communications Security (ACM-CCS2004), pp. 168–177 (2004)

[5] T. Nakanishi, et al. “Revocable Group Signature Schemes with Constant Costs for Signing and Verifying,” in 12th International Conference on Practice and Theory in Public Key Cryptograhy, Irvine, CA, 2009, pp. 463-480.

[6] Y. Ming, X. Shen, Y. Peng. "Provably Security Identity-based Sanitizable Signature Scheme Without Random Oracles." Journal of Software, 6(10), Academy Publisher, 2011, 1890-1897.

[7] L. Zhang, Q. Wu, Y. Hu. "New Constructions of Short Signatures in the Standard Model." Journal of Software, 6(10), Academy Publisher, 2011, 1921-1928.

[8] G. Ateniese, J. Camenisch, S. Hohenberger, and B. Medeiros. “Practical group signatures without random oracles.” Cryptology ePrint Archive, Report 2005/385, 2005. http://eprint.iacr.org/

[9] X. Boyen and B. Waters. “Full-domain subgroup hiding and constant-size group signatures.” In: Okamoto, T., Wang, X. (eds.) PKC 2007. LNCS, vol. 4450, pp. 1–15. Springer, Heidelberg (2007)

[10] J. Groth. “Fully anonymous group signatures without random oracles.” In: Kurosawa, K. (ed.) ASIACRYPT 2007. LNCS, vol. 4833, pp. 164–180. Springer, Heidelberg (2007)

[11] X. Cheng, S. Zhou, L. Guo, J. Yu, H. Ma. An ID-Based Short Group Signature Scheme. Journal of Software, 8(3), Academy Publisher, 2013, 554-559.

[12] M. Abe, G. Fuchsbauer, J. Groth, K. Haralambiev and M. Ohkubo. “Structure-preserving signatures and commitments to group elements.” In: T. Rabin(Ed.) CRYPTO 2010, LNCS 6223, pp. 209-236, 2010

[13] J. Groth and A. Sahai. “Efficient non-interactive proof systems for bilinear groups.” In: Smart, N.P. (ed.) EUROCRYPT 2008. LNCS, vol. 4965, pp. 415–432. Springer, Heidelberg (2008)

[14] B. Libert, T. Peters, M. Yung. “Group signatures with almost-for-free revocation.” Crypto 2012, LNCS 7417, pp. 571-589, 2012.

[15] H. Shacham, B. Waters. “Efficient ring signatures without random oracles.” Available at http://eprint.iacr.org/2006/289.pdf, 2006.

[16] N. Chandran, J. Groth, A. Sahai. “Ring signatures of sub-linear size without random oracles.” In ICALP, LNCS 4596, pp. 423-434, 2007

[17] D. Hofheinz and T. Jager. “Tightly secure signatures and public-key encrtyption.” In: R.Safavi-Naini and R.Canetti(Eds.): CRYPTO2012, LNCS 7417, pp. 590-607, 2012

[18] R. C. Merkle. “A Certified Digital Signature.” In: Brassard, G. (ed.)CRYPTO 1989. LNCS, vol.435, pp.218-238, 1990

[19] J. Camenisch, R. Chaabouni, A. Shelat. “Efficient Protocols for Set Membership and Range Proofs.” In: Pieprzyk, J. (ed.) ASIACRYPT 2008. LNCS, vol. 5350, pp. 234-252, 2008

[20] E. Brickell, J. Camenisch, and L. Chen. “Direct Anonymous Attestation,” Proc. 11th ACM Conf. Computer and Comm. Security, pp. 132-145, 2004.

A Vehicle Map-matching Algorithm based on Measure Fuzzy Sorting

Qunyong Wu
Key Lab of Spatial Data Mining and Information Sharing (Fuzhou university), Ministry of Education, Spatial Information Research Center of Fujian Province, P.R.China
Email: qywu@fzu.edu.cn

Xiaoling Gu, Jianping Luo, Panpan Zhang, and Xiaojuan Fang
Key Lab of Spatial Data Mining and Information Sharing (Fuzhou university), Ministry of Education, Spatial Information Research Center of Fujian Province, P.R.China
Email: 812400973@qq.com, 123ljp_123456@163.com, 369102682@qq.com, fangxiaoj@aliyun.com,

***Abstract*— The vehicle position obtained from GPS and dead reckoning is wildly applied to car navigation systems. However, the estimated position has an undesirable error due to the unknown GPS noise. To solve this problem, previous papers presented a method called "map-matching" to correct the position error. In this paper, we proposes a fuzzy ranking map matching algorithm based on measure factor. Comparing with other four algorithms, our algorithm improves in strategies of the error region determination, the road grid index and auto-adapted fuzzy sorting. To be specific, the error rectangle is firstly replaced by the error ellipse to reduce geometrical operation. Secondly, the grid index is adopted to accelerate the speed of filtering candidate road. At last, the relativity function and fuzzy sorting method help to sort the membership degree and to decide the matching road section. For the experiments, we implement a vehicle navigation system of five kinds of vehicle running status to testify the robustness and efficiency of this algorithm. The result shows that 96.7% of the GPS points are matched. In comparison with other algorithms, this algorithm had highest accuracy, which is of importance for vehicle navigation.**

***Index Terms*—fuzzy set, measure fuzzy sorting, map matching, vehicle navigation system**

I. INTRODUCTION

The main purpose of vehicle navigation system isto estimate the running trajectory in electronic map. However, the limitations of positioning technology would lead to the location deviation [1] on the map. Map-matching (MM), as an error correction technology, is precise to overcome the location deviation limitations from the positioning technology [2]. The basic theory of map-matching is to estimate the vehicle position relative to the map by comparing the information obtained from the vehicle positioning device and the road information from the electronic map database.

There are two premises for the application of MM, the continuous moving on the road, and the electronic road data, which is higher accuracy than those of the location estimation from the positioning system. The map-matching problem can be defined as the identification of the road where the vehicle is moving, not as the location of the vehicle position on the road. If there is a measure representing the possibility or certainty of existence of a vehicle on a specific road, it can be easy to locate the vehicle. By Compared the measures of all roads, the road where a vehicle is moving can be decided simply.

II. RELATED WORKS

Map-matching algorithm is actually a pattern identification process. In the past decades, a number of map-matching algorithms have been developed, These algorithms include Kalman filter, fuzzy logic and belief theory [3-4]etc. In general, map-matching algorithms can be categorized into four groups: geometric, topological, probabilistic and other advanced techniques.

The geometric map-matching algorithmwas introduced by Bernstein and Kornhauser[5] first. This algorithm contains point-to-point matching, point-to-curve matching, curve-to-curve matching and improved geometric map matching[6]. Point-to-point and point-to-curve matching don't fully make use of historical information, while curve-to-curve matching constructs piecewise linear curves from the paths that originate from the candidate nodes. whereas it is quite sensitive to outliers and depends on point-to-point matching in result of sometimes yielding unexpected and undesirable results[7]. =Although some improved geometric map-matching algorithm increases the performance, it is still quite sensitive to initial point.

The topological map-matching algorithm aims at the geometry itself, as well as the connectivity and contiguity of the links [8-11]. Topological MM (tMM) algorithms are relatively simple, easy and efficient, enabling them to be implemented in real-time situation [12]. In GIS platform, topology is viewed as the relationship between entities (points, lines, and polygons), such as adjacency, connectivity, containment etc... Because of the difficulty and complexity to estimate the topological of geometry, when the information of spatial topological relations is not sufficient, the result of this algorithm may not work

doi:10.4304/jcp.9.5.1058-1065

well. So in many cases, this algorithm usually works with others.

The probabilistic algorithm is proposed based on the statistics and mathematical [13]. This method constructs elliptical or rectangular error region around position fix [14] to defines the error model in the confidence regions where the vehicles really are, and to identifies a matched road segment by probability and statistics. The weakness of this method is that the more the vehicles deviate from the road, the more uncertain in estimating the location.

The advanced map-matching algorithms, such as Kalmam filter[15-19], Dempster-Shafer's mathematical (D–S) theory of evidence[20-21], fuzzy logic model[22-24], cost function model or computational geometry model, use more refined concepts. Krakiwsky was the first one who proposed the Kalmam filter map-matching algorithm[25]. Then Kim proposed an extended Kalman filter (EKF) [26-27].The inputs of the EKF came from GPS position fixes. The performance of EKF might depend on the quality of electronic map road data. Recently, Dempster developed an algorithm based on D–S theory of evidence which was generalized from traditional Bayesian theory[28-31]. The foundation of fuzzy logic model is fuzzy set that many studied it in different ways.[32-33], the map-matching algorithms based on fuzzy logic had better robustness than geometric's. But the error sources associated with the positioning devices and the electronic maps were not taken into account when estimating the location of the vehicle. The map-matching algorithm based on cost function or computational geometry proposed at the same time might invalid sometimes and measurement errors were not considered as well .

III. MEASURE FUZZY SET THEORY

To take the error sources and measurement errors into account and work out a strategy for invalid moment, we need to develop a new method, which is connected with the measure fuzzy set theory. In this section, the paper makes some explanation of the measure based map matching algorithm and the fuzzy set theory[34].

A. The Fuzzy Set Theory and Membership Function[34]

According to the fuzzy set theory in literature[34], A fuzzy set Ω is characterized by a membership function $f(x)$ which represents the grade of membership of $x \in X$ in Ω and takes values in the interval [0, 1]. Here X is a universal set，x is an object in X. The theory concerns the algorithm of obtaining the relativity function by using the pairwise functions.

Next, we shall introduce a pairwise membership function (or simply a pairwise function) defined in the literature[34].

DEFINITION 1 (Pairwise function): For $x, y \in X$, a pairwise function $f_y(x)$ means the membership function of x and also define that $f_x(x) = 1$ and $f_y(x) = 0$ for $x \notin X$.

DEFINITION 2(Relativity function): For $x, y \in X$, a relativity function *f*(x/y) means the fuzzy measure of choosing *x* over y and x greater than *y,*

$$f(x/y) = fy(x)/\max\{fy(x), fx(y)\}$$

DEFINITION 3(relativity function): For $x, y_i \in X$ (i = 1, 2... n), a relativity function means the fuzzy measure of choosing x over all y_i .

$$f(x/y_1, \ldots, y_n) \equiv f(x/\{y_i\}) \equiv f(x/T)$$

(T= { y_i } is a subset of X)

And

$$f(x/T) = \min[f(x/y_1), \ldots, f(x/y_n)] = \min_{i=1,\ldots,n}[f(x/y_i)]$$

Thus

$T=\{x_i\}=(x_1, x_2, \cdots, x_n) \in X \quad x_i \in X$ (i = 1, 2... n)

Then $f(x_j/T) = \min[f(x_j/x_1), \cdots, f(x_j/x_j), \cdots, f(x_j/x_n)]$

$=\min\{f(x_j/x_1), \cdots, f(x_j/x_{j-1}), f(x_j/x_{j+1}) \cdots, f(x_j/x_n)\} = \min_{i=1,\ldots,n}[f(x_j / x_i)]$

Let T be a set of objects $x_i \in X$ (i = 1, 2... n), and $T_1 \cup T_2 = T$ ($T_1, T_2 \subset T$). Then

$$f(x_j/T) = \min[f(x_j/T_1), f(x_j/T_2)]$$

Then the relation which objects can be ranked in order between the pairwise function and the relativity function will be shown as following:

$$f(x_j/T) = \min\left(\frac{f_{x_1}(x_j)}{f_{x_j}(x_1)}, \cdots, \frac{f_{x_n}(x_j)}{f_{x_j}(x_n)}\right) \quad x_i \in X \quad (i = 1, 2...n)$$

B. The Membership Function for the Candidate Sections

Judging the sections where the vehicle drives is mainly based on the size of the projection distance and the direction angle that is between the current vehicle location point and the candidate section. Generally speaking, the smaller projection distance and direction angle of the candidate road are the more likely to get the matching road sections.

Let $p(x, y)$ and h as the vehicle location and the heading separately, $p_r(x_r, y_r)$ and h_r as the projection point location and the candidate road section direction.

Define distance and direction function:

$$D(p, p_r) = \frac{1}{1 + \frac{\|p - p_r\|}{\sigma^2}} = \frac{1}{1 + \frac{\sqrt{(x - x_r)^2 + (y - y_r)^2}}{\sigma^2}}$$

$$H(h, h_r) = \frac{(h \bullet h_r)^2}{\|h\|^2 \|h_r\|^2} = Cos^2(h - h_r)$$

Here: σ is the standard deviation of the position estimation value; the default is 100m,

We can infer: $D \subset (0,1], \quad H \in [0,1]$

Define the vehicle membership degree of candidate road section:

$$M(r) = \alpha D(p, p_r) + \beta H(h, h_r)$$

Here: $M(r)$ is the vehicle membership function，α、β are the coefficient of the measure

Through simulation testing, the coefficient of α and β is 2 and 1seprately. There are a certain relationship between the sizes of α and β and the vehicle running trajectory, heading, and the geometrical distribution of the candidate road section. When the distance is large and the direction difference is small, α is small and β is bigger, vice versa. Thus, we define measure coefficient:

$$\alpha = \begin{cases} 1 - \cos^2(h_{ri} - h_{rj}) & 0^0 \le h_{ri} - h_{rj} \le 90^0 \\ 1 + \cos^2(h_{ri} - h_{rj}) & 90^0 < h_{ri} - h_{rj} \le 180^0 \end{cases}$$

$$\beta = (\max\{\alpha\} - \min\{\alpha\})/2 = 1$$

Here: $h_{ri} - h_{rj}$ is the direction difference for two candidate sections, and $0° \le |h_{ri} - h_{rj}| \le 180°$； $0 \le \alpha \le 2$.

IV. Measure Fuzzy Sorting Based Map-Matching Algorithm

According to the estimation based map- matching algorithm and the fuzzy set theory, the relativity function, membership function and fuzzy sorting are introduced into the map-matching algorithm and the measure fuzzy sorting map matching algorithm is put forward(Fig.1). The algorithm is described as follows:

(1) acquiring the GPS positioning data including the location and the heading;

(2) using rectangle error region instead of error ellipse region;

(3) determining the current grid through the grid index;

(4) filtering the candidate road sections according to the rectangle error region;

(5) calculating the membership degree for each two candidate sections according to the matching function;

(6) sorting adaptive fuzzy according to the membership degree of all candidate road section;

(7) selecting the candidate with the maximum membership as the matching road section;

(8) calculating projection from the GPS position located directly to the matching road section, which means the projection point is the matching point.

The algorithm compares the road membership value of candidates by fuzzy sorting, and adjusts the measure coefficient to improves the accuracy of map matching.

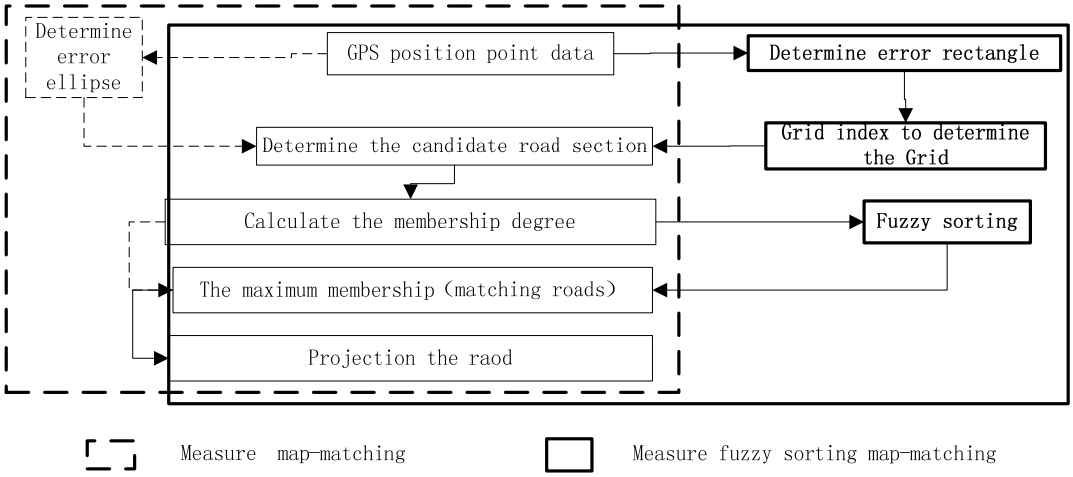


Figure 1. Main flow for map-matching algorithm based on measure fuzzy sorting

From Fig. 1, we can see that the differences between the measure fuzzy sorting based map matching algorithm and other algorithms are the error region determination, the road grid index and fuzzy sorting etc.

A. Determine the Error Rectangle

GPS positioning error ellipse is used to express the positioning error region. According to the references, the calculating formula of the error ellipse is shown as follows:

$$a = \delta_0 \sqrt{\frac{1}{2}(\delta_x^2 + \delta_y^2 + \sqrt{\delta_x^2 + \delta_x^2 + 4\delta_{xy}^2})}$$

$$b = \delta_0 \sqrt{\frac{1}{2}(\delta_x^2 + \delta_y^2 - \sqrt{\delta_x^2 + \delta_x^2 + 4\delta_{xy}^2})}$$

Here: **a** is the major semi-axis of the ellipse**,** and **b** is the minor semi-axis of the ellipse. δ_0 is the extend factors, δ_x and δ_y is the standard deviation of the GPS measuring error. δ_x^2 and δ_y^2 are the variances； δ_{xy} and δ_{yx} is the covariance.

It should be noted that The calculation of locating candidate road sections from the error ellipse is more difficult than rectangle region. The external rectangle of the error ellipse is substituted for the error ellipse. As the GPS position error is about 100m, the length of the error region will be set 100m in our study.

B. Calculate Roads Grid Index and the Candidate Road Section

The performance indicators of grid index include: the sizes of the grid, the record numbers of the grid index table, the ratios between the record numbers of grid index table and the entity record numbers, the average entities

numbers of every grid et al. The key of these indicators is the size of the grid. If the size of the grid is too large, the amount of the road sections within the grid is overmuch, which can reduce the efficiency of spatial query. Otherwise, if the size of the grid is too small, there are large numbers of duplicate records between the grids, which causing a large amount of redundancy. Taking the single frequency GPS for the example, if the size of the grid is set to 100m, the accuracy exceeds 95%. So in our study, we set the size of the grid to 100m.

(1)Roads grid division and grid index construct

Taking 100m as the step size, the road network was divided equally into M * N grids , abbreviated as Grid (M * N), from top to bottom, from left to right, where M and N respectively means the number of rows and columns. For each grid, it needs to record all the code of sections including or crossing this grid, and to preserve the coordinate of the grid upper left corner as its initial coordinates. And then, we need to index the grid in order to locate quickly. This step needs to be completed when loading road layer data in initialization phase.

(2)Candidate road sections determination

In order to decide the candidate road sections, the grid must firstly be determined in which the GPS position point lies, and then all the roads in the gird must be selected as the candidate road sections. According to the location of the GPS position point, there are four relations between the GPS position point and the road grid (as shown in Fig2). First is the GPS position point at the nodes of the grid. The second is the GPS position point in the grid. The third is the GPS position point in the column edge of the grid. And the last is the GPS position point in the row edge of the grid.

In the relation and the error rectangle, we expand the rectangle error region according to the gird in order to simplify the candidate road sections selection. For the first circumstances, the rectangle error region expands for 4 grids (the bold black line grid as shown in Fig2). For the second, the rectangle error region expands for 9 grids (the bold black line grid as shown in Fig2). For the third and the fourth, the rectangle error region expands in horizontal and in vertically separately for 6 grids (the bold black line grid as shown in Fig2).

All the roads in the rectangle error region are the candidate road sections.

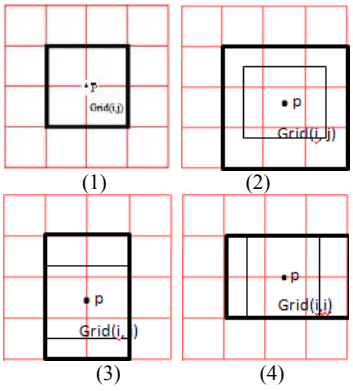


Figure 2. relation between the GPS position point and the Grid

(3)Candidate section measure fuzzy sorting

According to the membership function, we can calculate the membership values M_i (r) and $M_j(r)$ for each two candidate road sections i and j, and the value reflects the measure degree between the two candidate road sections separately.

The definition of the vehicle relativity function:

$f(x_i \mid x_j) = M_i(r) / \max\{M_i(r), M_j(r)\}$

Calculate the vehicle relativity function for all candidate roads, and define comparison matrix as:

$$C = \begin{Bmatrix} f(x_1/x_1) & f(x_1/x_2) & \cdots & f(x_1/x_n) \\ \cdots & \cdots & \cdots & \cdots \\ f(x_n/x_1) & f(x_1/x_2) & \cdots & f(x_n/x_n) \end{Bmatrix}$$

And then calculate the minimum value for every row,

$C_i' = \min f(x_i \mid X) \quad i = 1, 2, \cdots, n$

C_i' is the membership value for the candidate road section i.

According to the introduction in chapter[3.1], The road section with the largest membership sort value is the matching road section(CMR). That is :

$\mathrm{CMR} = \max(C_i') = [\min f(x_1 \mid x_i), \min f(x_2 \mid x_i), \cdots, \min f(x_n \mid x_i)]^{\mathrm{T}} \quad i = 1, 2, \cdots, n$

V. Algorithm Adaptability Analysis and Processing for Vehicle Navigation

Any algorithm has its adaptability. Therefore, to improve the robustness and the efficiency of the algorithm, different processing methods is designed for different vehicle cases. For example, when the match road section is very clear, calculating the matching road section is not necessary and a simple projection can implement map matching. According to the requirement of the vehicle navigation, five kinds of vehicle running statuses are designed(shown as fig 3), which are road search status, normal driving status, delay matching status , parking or low-speed status and signal errors status..

A. Road search Status

Map-matching algorithm starts when the matching road is not determined, such as in the initial state and the intersection state. In this case, fuzzy sorting matching algorithm can be used directly and the GPS position point can be projected into the matching road section that has the maximum membership degree.

B. Vehicle Normal Running Status

If the GPS position point is not in the range of intersection (not entered or left), and at the same time the current matching road section is determined, or the GPS position point is in the intersection, and road search has been completed. In above state, according to the road connectivity rule, it is not need to search the matching road section but to project the GPS position point into the identified road directly.

C. Delay Matching Status

In the fuzzy sorting algorithm, if the gap between the maximum matching value and large matching values is

very small, for example, less than the threshold condition (generally set to 0.1-0.2), it means that the road condition is complex and it is difficult to judge which one is the suitable matching road section. In this case, the delayed matching is needed until the matching road section can be determined.

D. Vehicle Parking or Low-speed

When the vehicle is stopped or low-speed, the vehicle location should be in motionless state in theory. However, because of the random drift of GPS positioning error, the vehicle position information given by the GPS is often randomly shift within a region. In this case, the current matching point can be seen as the current location of the vehicle until the vehicle is on.

E. GPS Data Exception

If the distance between two continuous GPS position point is significantly higher than the normal range, it can be seen as the GPS data exception. In our design, it is twice than the normal value. In this case, it is required to calculate the vehicle location though the linear interpolation according to the historical vehicle trajectory.

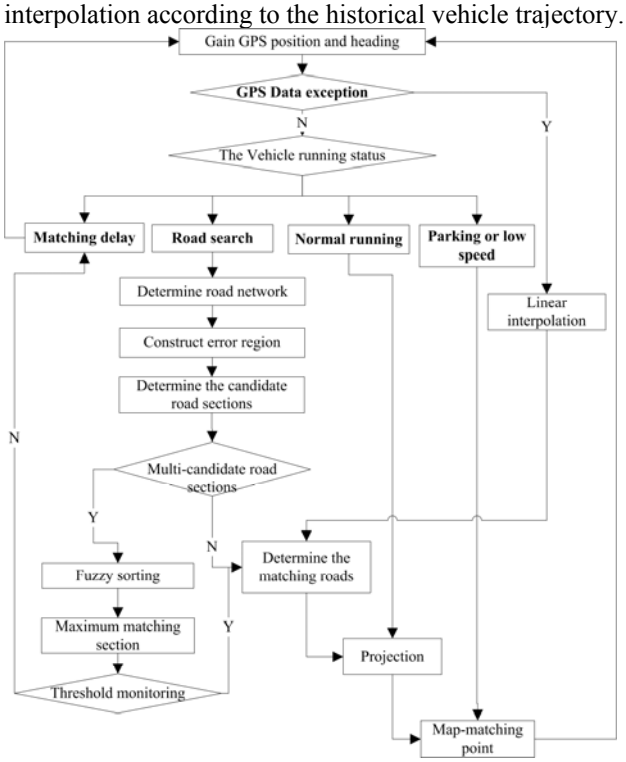


Figure 3. Map-matching flowchart with vehicle different running statuses

VI. Implement of Measure Fuzzy Sorting based Map-matching Algorithm

Taking the roads of Fuzhou city as an example, there are 91 GPS position points. the algorithm is implemented according these points.

A. Data Pre-processing

The coordinate of the electronic map is the Gauss projection coordinate system under the Xi'an-80 geographic coordinate system. But the coordinate of GPS position point is WGS-84 coordinate system, which has the different ellipsoid and the geodetic datum with Xi'an 80 coordinate system. To make sure they are in same coordinate system, the coordinate of GPS position point need to be converted into the Gauss projection coordinate system under Xi'an- 80 geodetic coordinate system. There are two steps of the data pro-processing. The first is to transform the WGS-84 coordinate system of GPS position point into Xi'an-80 geodetic coordinate system. The second is to transform the geographic coordinate system into projection coordinate system though the Gauss projection.

B. Algorithm Design Patterns according to the Vehicle Running Status

The Strategy Pattern defines a family of algorithms, encapsulates each one, and makes them interchangeable. Strategy lets the algorithm vary independently from clients that use it. The structure of the strategy is shown in Fig4. It is very suitable to use strategy pattern to deal with the vehicle running status changes.

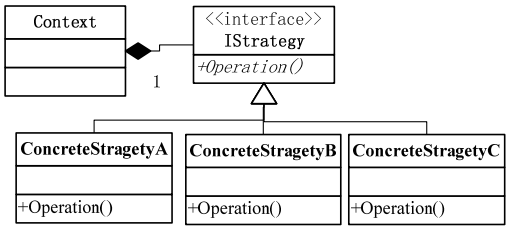


Figure 4. The structure of strategy pattern

From the chart we can see: strategy pattern is actually to package the algorithm separately, such as the ConcreteStrategyA, ConcreteStrategyB and ConcreteStrategyC. They are inherited from an interface in order to realize the algorithm calls with polymorphic manner when Context called them.

According the vehicle's five running conditions, which include road search status, normal driving status, delay status, parking or low-speed status and signal error status, we built five algorithms separately. In order to call the different algorithm, we use the strategy patterns to package them. The UML class diagram is shown in Fig5.

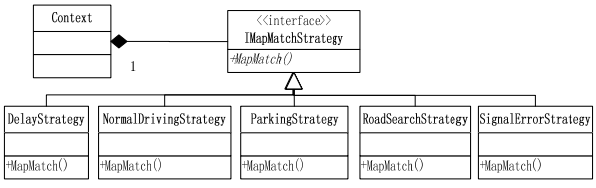


Figure 5. The UML class diagram of strategy patterns

C. Fuzzy Sorting Algorithm Implementation

RoadSearthStrategy class is built to implement fuzzy sorting. First, MembershipCalculator class is used to calculate the membership degree between the GPS position point and the candidate road sections. The MembershipCalculator class uses the location (MatchingPoint) , heading (MatchingPointdirection) of GPS position point , the projection location (ProjectedPoint) and heading (CandidateSectionDirection)

of the candidate road section to calculate the membership degree. Then, FuzzyRank class sorts the candidate road sections according the membership degree calculated by MembershipCalculator class and selects the candidate road section with the maximum membership degree as the matching road section.

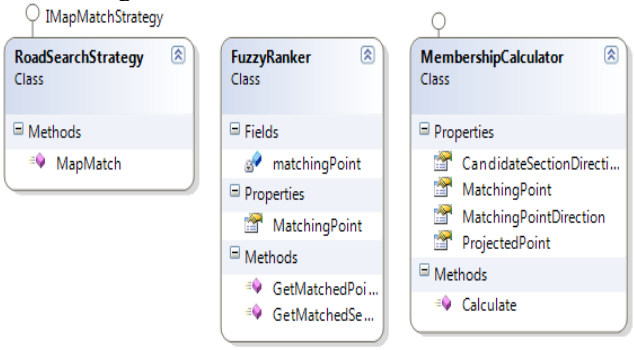


Figure 6. The fuzzy sorting class interfaces

D. Grid Index Construct

GridsBuilder class is to implement the Grid index, (shown as fig7) which gets access road entity and sets through the IPolylineFeatureCollection interface. Then the class divides the grid by the grid size(Width) and generation Grid type list. ,It also stores the road entity (IPolylineFeature) within each grid in the corresponding grid object road entity set (Features).

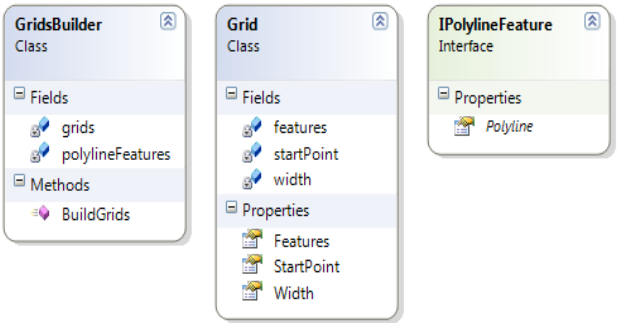


Figure 7. The grid index construct class interfaces

E. Candidate Road Sections Filter

The RouteFilter class is to implement the filter of candidate road sections.

Firstly, according to the GPS position point, Rect Class determines the grid location. The binary search algorithm is used to accelerate the search speed. Secondly, ErrorRegion class constructs the error region rectangle. Finally, the candidate road sections set (IPolylineFeatureCollection) will be generated though filtering all the road sections within the error region rectangle.

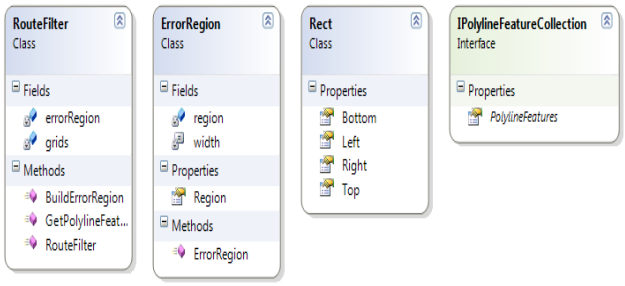


Figure 8. The candidate road filter class interfaces

VII. Experiment and Result Analysis

Due to the complexity of algorithm, it would take a little more time than others, but the result got the highest accuracy, which is the most important for terminal users.

In this experiment, there are total 91 GPS positing point. Through the calculating, the matching-map result is shown in Fig9. The number of matching point is 88, which means the accuracy of our alogorithm is 96.7%, the total time cost is 9000ms, and the single point time cost is 99ms.

The comparison with other algorithms is list in table 1. Compared with other algorithms, our algorithms accuracy is obviously improved, the algorithm complexity is elevated which leads to the computation time is longer, so we need to be further improved to reduce time-consuming.

TABLE 1
Comparison in Accuracy Degree and Time Cost

Map matching algorithm	Accuracy degree/%	Time cost/ms
Based on computational geometry	96.24	0.038
Projection algorithm	87.86	53
Based on D-S reasoning	93.97	85
Based on fuzzy logic	92.31	78
Based on Cost function	92.78	81
Based on the measure fuzzy sorting	**96.7**	**99**

*The accuracy degree and cost time of other algorithm comes from references [27]

VIII. Conclusion

In this paper, we have proposed a fuzzy sorting map-matching algorithm. The grid index was used for candidate road sections filtering, and the rectangle error region substituted for the error ellipse. Relativity function and fuzzy sorting method were used to sort the membership degree so as to determine the matching road section. For the experiments, different running conditions of the vehicle was considered and implemented,, t the results demonstrate that the accuracy was better than others. In our further research, the algorithm complexity needs to be improved, but for now it is a better choice among others for most users, which does make a difference in this field.

IX. ACKNOWLEDGEMENT

This paper is supported by FP7-PEOPLE-IRSES (No.247608), Fujian Province Nature Science Foundation (No: 2012J01168), National Technology Support Project (No: 2013BAH28F00). I would like to thank the members of my committee for their support.

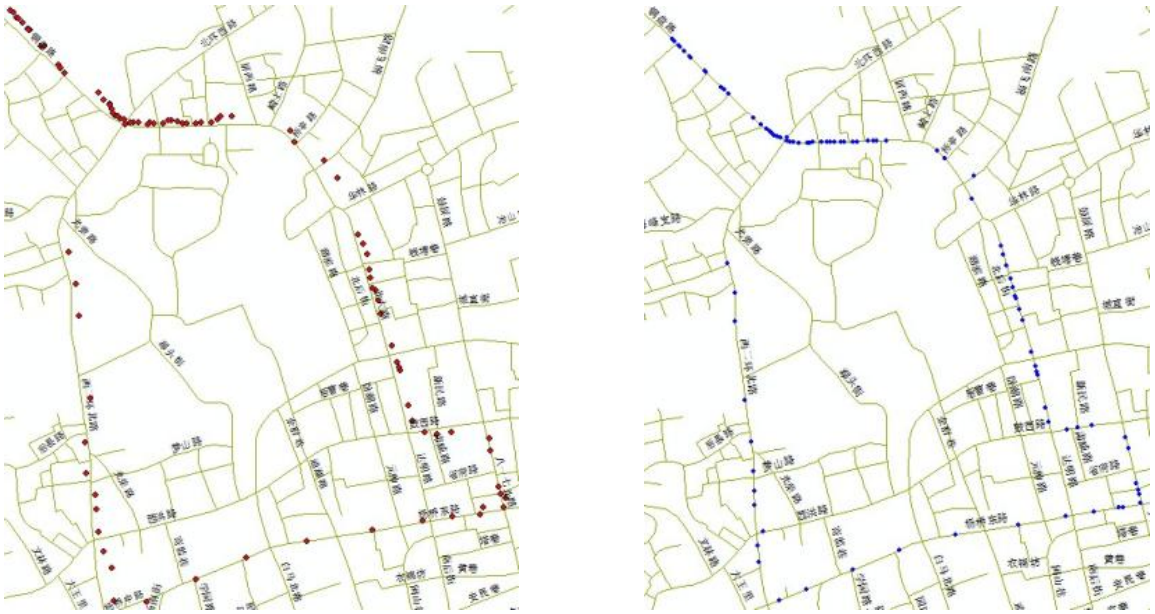


Figure 9. The original GPS position point and the results of map-matching

REFERENCES

[1] J. Su, D.F. Zhou, C.S. Yue. "Real-time map-matching algorithm in GPS navigation system for vehicles," Acta GeoDaeticaet Cartographica Sinica, vol. 30(3), pp.252-256, 2001.

[2] M.A. Quddus, W.Y. Ochieng, R.B. Noland. "Current map-matching algorithms for transport applications: State-of-the art and future research directions," Transportation Research, vol. 15(3), pp.312–328, 2007.

[3] M.A. Quddus, W.Y. Ochieng, L. Zhao, R.B.Noland. "A general map-matching algorithm for transport telematics applications," GPS Solutions, vol. 7 (3), pp.157-167, 2003.

[4] C.E. White, D. Bernstein, A.L. Kornhauser. "Some map matching algorithms for personal navigation assistants," Transportation Research Part C, vol. 8, pp.91-108, 2000.

[5] D. Bernstein, A. Kornhauser. "An introduction to map matching for personal navigation assistants," Technical report, New Jersey TIDE Center Technical Report, 1996.

[6] B. Phuyal. "Method and use of aggregated dead reckoning sensor and GPS data for map-matching," In: proceedings of the Institute of Navigation (ION) annual conference, 2002.

[7] M.A Quddus.. "High integrity map-matching algorithms for advanced transport telematics applications," PhD Thesis. Centre for Transport Studies, Imperial College London, UK, 2006.

[8] Y. Meng. "Improved Positioning of Land Vehicle in ITS Using Digital Map and Other Accessory Information. PhD Thesis," Department of Land Surveying and Geoinformatics, Hong Kong Polytechnic University, 2006.

[9] C. Wu, Z.L. Li, M. Yu, Y.Q. Chen. "Effects of sensor errors on the performance of map-matching ," Journal of Navigation, 2005.

[10] H. Yin, O. Wolfson. "A Weight-based map-matching method in moving objects databases, Scientific and Statistical Database Management ," In: Proceedings of the International Working Conference, vol. 16, pp.437-438, 2004.

[11] C.A. Blazquez, A.P. Vonderohe. "Simple map-matching algorithm applied to intelligent winter maintenance vehicle data," Transportation Research Record, 2005.

[12] R.V. Nagendra, A.Q. Mohammed, L.B. Abigail. "Developing an enhanced weight-based topological map-matching algorithm for intelligent transport systems ," Transportation Research Part C: Emerging Technologies, vol. 17(6), pp.672–683,2009.

[13] S.K. Honey, W.B. Zavoli, K.A. Milnes, A.C. Phillips, White MS, Loughmiller GE. "Vehicle navigational system and method ," United States Patent, 1989.

[14] F. Chen, M.Y. Shen, Y.N. Tang. "Local Path Searching Based Map Matching Algorithm for Floating Car Data," Procedia Environmental Sciences, vol. 10(A), pp.576-582, 2011.

[15] E.J. Krakiwsky, C.B. Harris, R.V.C. Wong. "A Kalman filter for integrating dead reckoning, map matching and GPS positioning," In: Proceedings of IEEE Position Location and Navigation Symposium, pp. 39-46, 1988.

[16] W. Kim, G. Jee, J. Lee. "Efficient use of digital road map in various positioning for ITS," In: IEEE Symposium on Position Location and Navigation, San Diego, CA, 2000.

[17] M.E. El Najjar, P. Bonnifait. "A road-matching method for precise vehicle localization using belief theory and Kalman filtering ," Autonomous Robots, vol. 19(2), pp.173-191, 2005.

[18] J.C. Fang, G.X Shen., D.J. Wan. "Establishment of an Adaptive Extended Kalman Filter Model for a GPS /DR Integrated Navigation System ," Control and decision, vol.

14(5), pp.448-452, 1999.
[19] Y.J. Cui. “Autonomous vehicle positioning with GPS in urban canyon environments ,” IEEE Transactions on Robotics and Automation, vol. 19 (1), pp. 15-25, 2003.
[20] D.K. Yang, B.G. Cai, Y.F. Yuan. “An improved map-matching algorithm used in vehicle navigation system ,” IEEE Proceedings on Intelligent Transportation Systems, vol. 2, pp.1246-1250, 2003.
[21] X.Y. Yang, S.G. Huang. “Map Matching Algorithm Based on Road Reduction Filter in Integrated Vehicle Navigation System ,” Journal of remote sensing, vol. 9 (2), pp.215-219, 2005.
[22] M.Y. Fu, J. Li, M.L. Wang. “A hybrid map-matching algorithm based on fuzzy comprehensive Judgment ,” IEEE Proceedings on Intelligent Transportation Systems, pp.613-617, 2004.
[23] M. Cossaboom, J. Georgy, T. Karamat, A. Noureldin. “Augmented Kalman Filter and Map Matching for 3D RISS/GPS Integration for Land Vehicles,” International Journal of Navigation and Observation, 2012.
[24] M.A. Quddus, R.B. Noland, W.Y. Ochieng. “A high accuracy fuzzy logic-based map-matching algorithm for road transport,” Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, vol. 10 (3), pp.103-115 , 2006.
[25] E. J. Krakiwsky, C. B. Harris, R. V. C. Wong. A Kalman filter for integrating dead reckoning, map matching and GPS positioning. Position Location and Navigation Symposium, 1988. Record. Navigation into the 21st Century. IEEE PLANS '88., IEEE (06 August 2002), pp. 39-46.
[26] S.J. Julier; J.K. Uhlmann."Unscented filtering and nonlinear estimation". Proceedings of the IEEE(2004): 401–422.
[27] E. Courses; T. Surveys. "Sigma-Point Filters: An Overview with Applications to Integrated Navigation and Vision Assisted Control". Nonlinear Statistical Signal Processing Workshop, 2006 IEEE: 201–202.
[28] P. Davidson, J. Collin, J. Takala. “ Application of particle filters to a map-matching algorithm. Gyroscopy and Navigation, vol.2(4), pp.285-292, 2011.
[29] Sh. Sun, J. Gao, M. Chen, B. Xu, Zh. Ding. “FS-DS based Multi-sensor Data Fusion”, Journal of Software, Vol. 8(5),pp. 1157-1161, 2013.
[30] M.Y. Fu, Z.H. Deng, T. Liu. “Intelligent vehicle navigation technology,” Science Press, 2009.
[31] A.B. Carola. “A Decision-Rule Topological Map-Matching Algorithm with Multiple Spatial Data,” Global Navigation Satellite Systems: Signal, Theory and Applications, InTech, 2012.
[32] M. Yan. “Dominance-based Rough Interval-valued Fuzzy Set in Incomplete Fuzzy Information System”. Journal of Software, Vol. 7(6),pp. 1375-1384. 2012.
[33] J. Li, G. Liao, F. Wang, J.i Li. “Maximum Lifetime Routing Based on Fuzzy Set Theory in Wireless Sensor Networks”. Journal of Software, Vol. 8(9) pp.2321-2328, 2013.
[34] S. Masamichi. “Fuzzy sets concept in rank-ordering objects,” Journal of Mathematical Analysis and Applications, vol. 43 (3), pp.717 -733, 1973.

Qunyong Wu, male, was born in 1973, in Shandong China. .He received the Bachelor's degree and the Master's degree in Computer Application respectively in 1996 and in 2002, from Fuzhou University, Fujian, China. He received Doctor's degree in Cartography and geographic information system in 2006, from Key Lab of Graduate School, Geographic Sciences & Natural Resources Research Institute, and Resources & Environment Information System in China Academic of Sciences, Beijing, China.

He used to involve in education and research of computer application in Computer Department of Fuzhou University between 1996 and 2000. During 2000 and 2001, he took part in the research of Geographic Information Science and Technology Institute. Since then, he has been working as assistant researcher, vice researcher and instructor in Spatial Information Research Center of Fujian, Fuzhou University, China. Besides, he has been also a member of Key Lab of Spatial Data Minging and Information Sharing (Fuzhou university) from 2004. His research interests are network geographic information sharing and service, and geographic information system software development.

Xiao-ling Gu, femal , was born in Jiangxi, China, in 1989.She received the Bachelor's degree in Geographical Information System in 2011, form China University of Petroleum(East China), Qingdao, China. She is currently a postgraduate at Spatial Information Research Center of Fujian in Fuzhou University, Fujian, China. Her research interest is geographic information sharing.

Jianping Luo, male, was born in Sichuan, China, in 1986. In 2009, He received the Bachelor's degree in Resource Environment & Urban and Rural Planning and Management in Anhui Construction Industry Institute in Hefei, Anhui. In 2013, he earned the Master’s degree of Cartography and geographic information system in Spatial Information Research Center of Fujian in Fuzhou University, Fujian, China. His majority researches are geographic information sharing and service technology, and moving objects database.

Panpan Zhang, male, born in 1986 in Henan, China, received his Bachelor’s degree in School of Resources and Environment of Henan Polytechnic University, Henan, China, in 2008 and got Master’s degree of Cartography and geographic information system in Spatial Information Research Center of Fujian in Fuzhou University, Fujian, China in 2011.
Now he is working in Beyon Information Technology Co. Ltd. in Beijing, China, as a software engineer. His research interests are geographic information service and spatial database management system.

Xiaojuan Fang, femal ,was born in Xiamen, China, in 1988. She received the Bachelor's degree in Engineering in 2011, from Xiamen University of Technology, Xiamen, China. She is currently a postgraduate at Spatial Information Research Center of Fujian in Fujian university, Fujian, China. Her primary research interest is geographic information sharing.

Optimizing Data Distribution for Loops on Embedded Multicore with Scratch-Pad Memory

Qiuyan Gao[a], Qingfeng Zhuge[b], Jun Zhang[a,b], Guanyu Zhu[d], and Edwin H.-M. Sha[b,c]
[a] College of Information Science and Technology
Hunan University, Changsha, China.
[b] College of Computer Science
Chongqing University, Chongqing, China.
Email: { angela.gao8908, qfzhuge, jeffjunzhang}@gmail.com
[c] Department of Computer Science
University of Texas at Dallas, Richardson, Texas 75080, USA.
Email: edsha@utdallas.edu
[d] Huawei Technologies Co. Ltd., Shenzhen, China.
Email: zhuguanyu@huawei.com

Abstract— **Software-controlled Scratch-Pad Memory (SPM) is a desirable candidate for on-chip memory units in embedded multi-core systems due to its advantages of small die area and low power consumption. In particular, data placement on SPMs can be explicitly controlled by software. Therefore, the technique of data distribution on SPMs for multi-core system becomes critical in exploiting the advantages of SPM. Previous research efforts on data allocation did not consider the placement of array data accessed in loops. Loops are the most time-consuming and energy-consuming part for most of the computation-intensive applications. In this paper, we propose a high-performance, low-overhead data distribution technique, the Iterational Optimal Loop Data Distribution Algorithm based on dynamic programming. It optimizes data allocation of both scalar and array data for embedded multi-core systems with SPMs. The experimental results show that the IOLDD algorithm reduces the energy consumption by 30.12% and 14.52% on average compared with random data distribution and greedy stretagy, respectively. It also reduces the memory access time by 18.45% and 18.38% on average compared with the random distribution strategy and the greedy strategy, respectively.**

Index Terms— **Data distribution, multi-core, scratch-pad memory, embedded systems**

I. INTRODUCTION

MULTI-CORE design becomes the mainstream of high-performance embedded systems because of the ever-increasing demand on performance for applications such as digital signal processing, wireless communication, and mobile computing. Meanwhile, the design of multi-core systems usually has to satisfy strict requirements on low power consumption and small die area. Therefore, Scratch-Pad Memory becomes an effective design alternative to replace cache as on-chip memory in embedded multi-core systems. Software-controlled SPM guarantees a single-cycle access time with low energy consumption and small die area compared with hardware-controlled cache. In particular, data on SPM can be precisely controlled by software during system design. Many digital signal processing systems such as Analog Devices ADSP-BF534/6/7 [1] and TI's TMS370CX7X [2], as well as multicore architectures such as NVIDIA GeForce 8800 [3], employ SPM as on-chip memory [4]. Therefore, how to efficiently distribute data items to on-chip SPMs to minimize the memory access cost becomes one of the key problem for fully exploiting the advantages of SPMs in embedded multi-core systems.Because loops are the most time and energy consuming code section in most of the computation-intensive applications, it is desirable to have efficient techniques to allocate array data, as well as scalar data, to multiple SPMs in a multi-core system.

A lot of multi-core systems employs symmetric multiprocessing (SMP) architecture. Multiple cores share a centralized main memory. Each core is equipped with a small and fast on-chip SPM to speed up data accesses. Usually, there is only one copy of each data item. Data items accessed by multiple cores can spread on SPMs for multiple cores. The cost for searching and moving data items around, however, is high for multicore systems. In this paper, we propose a technique for keeping and updating one copy of array data in main memory efficiently with a minimized backup cost. Furthermore, we propose a data duplication method to replicate local copies for read-only data to further reduce the data access cost.

Because of the capacity of SPMs is limited, only the commonly used data or data quite critical should be loaded into SPMs. Other data items are stored in off-chip main memory [5]. In this paper, a dynamic programming approach is used to produce optimal results for embedded multi-core systems with SPMs. The approach also achieves the goal of distributing both array and scalar data items in loops on multi-core systems and minimizing the time cost and energy consumption.

doi:10.4304/jcp.9.5.1066-1076

In this paper, we make the following contributions:

1) We propose a polynomial-time data distribution algorithm, the Iterational Optimal Loop data Distribution algorithm with Duplication (IOLDD), to minimize the total cost of memory access on multi-core systems equipped with SPMs for both arrays and scalar variables in loops.
2) We present a data duplication technique and integrate it into the data distribution algorithm. It further reduces the total cost of memory accesses by replicating multiple copies for read-only data items.

The rest of this paper is organized as follows. Related works are discussed in Section II. Models and some basic concepts are introduced in Section III. A motivational example is discussed in Section IV to illustrate some basic ideas of our algorithm. The problem definitions used in the paper are given in Section V. In Section VI, details of our improved dynamic approach IOLDD are presented. Section VII presents our experiments and Section VIII concludes the whole paper and mentions the future work.

II. Related Works

There are a lot of works tackling the data distribution problem. Some of the works proposed static data distribution methods. The data distribution is determined for the whole program and will not change during the execution of the program [6] [7] [8]. The drawback of static methods is that it cannot explore the benefit of varying data locality in a running environment. The other category of previous techniques is dynamic data distribution [9] [10] [11]. For those dynamic methods, program will be divided into different regions. Data movement instructions are inserted before each region to generate data distribution for a program region. The data distribution remains the same in the execution of a particular region. Greedy strategy, for example, is used to find a data distribution for each region by Udayakumaran in [12] [13]. Since dynamic data distribution takes advantage of the data locality of each program region, they have better performance than the static ones.

Array data is different from scalar data. Elements in array occupy contiguous memory locations. A single iterative statement in loop may process arbitrarily many elements of an array. Distributing array data is, therefore, quite different from the way we handle scalar data. To the best of the authors' knowledge, there is not much research work conducted on data distribution for arrays in a loop, and some methods greatly rely on the loop's characteristics. O. Ozturk *et al.* proposed algorithms to manage data for array-intensive nested loops with regular data access patterns [14] [15]. R. Thakur *et al.* proposed efficient algorithms to manage dynamic redistribution of arrays [16]. W. Huan *et al.* proposed algorithms to optimize all of the data segments, including global data, heap and stack data in general [17]. These methods do not consider the case of distributing both array and scalar data items in a loop on multi-core systems.

Research efforts have also been taken on the data distribution problem for SPMs. R. Banakar *et al.* proposed a simple SPM data management algorithm. But the algorithm cannot guarantee to achieve optimal results and is only applicable to scalar data [4]. Jun Zhang *et al.* proposed an algorithm for loops on single-core systems instead of multi-cores [18]. Y. Guo and Q. Zhuge *et al.* proposed a polynomial-time algorithm to solve the data distribution problem on multiple types of memory units [19] [20]. However, they only consider distribution for scalar data items and do not mention data distribution for loops. The data distribution problem for array data is very important for most of the applications. In this paper, we focus on developing a dynamic programming approach for both array and scalar data items on multi-core systems.

III. Models and Basic Concepts

In this section, we first introduce the hardware architecture. Then, we will present the program execution model we use in this paper.

A. Hardware Architecture

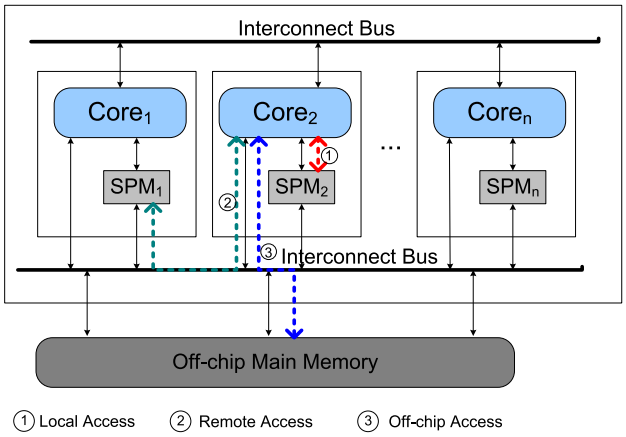


Figure 1. A multi-core hardware architecture.

The organization of on-chip memories of our hardware architecture is shown in Fig. 1. Every core has its own on-chip SPM, while all cores share the DRAM main memory. Each core can access its own local SPM. It can also access data items on other cores' SPMs by the interconnect bus. Scratch-pad-memory here can be organized as a Virtually Shared SPM (VS-SPM) architecture for on-chip memory that takes advantage of both shared and local SPM [21]. Distinct from the local SPM, SPMs of other cores are referred to as remote SPMs. There is no limit for the number of remote SPMs that a core can access in the architecture.

In Fig. 1, we show three types of memory access pattern as depicted by three types of dot lines. They are local access, remote access and off-chip access, corresponding to the accesses for cores to local SPM, remote SPM and off-chip memory, respectively. Due to the communication cost of the interconnected data bus, remote access incurs

longer latency than the local access. Objectively, for the low performance of DRAM and the high communication cost, latency of off-chip access is much longer than the latencies of both local and remote accesses [22]. In our architecture, each core can access all remote SPMs. Let $Dist$ be the distance between two cores. The cost of remote access is a non-decreasing function f of $Dist$.

B. Execution Model

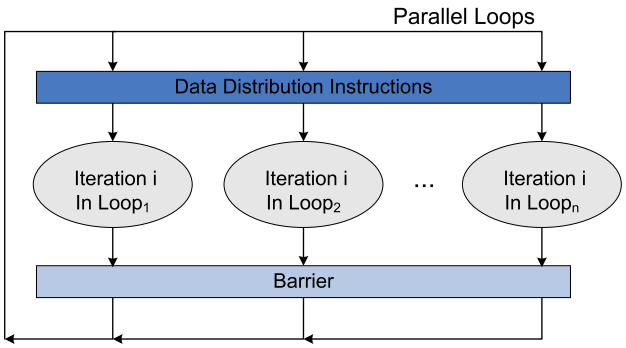


Figure 2. Demonstration of execution model for Loops .

In this paper, we consider the data distribution problem for loops can be paralleled in the program. A barrier is used to synchronize the execution of each iteration for all the loops. We also consider each basic block of a program as a program region. The execution model of loops executed in parallel on a multi-core system is shown in Fig. 2. Assume that there is no conditional branch in loop body. Each iteration is regarded as a program region by the compiler. The number of accesses on each data item in a program region can be obtained through profiling. Compiler inserts data distribution instructions at the beginning of a loop iteration. Therefore, data items are allocated to various memory units before parallel regions are executed. In case of conditional branch existing in loop body, each branch then should be considered as a program region. Data distribution instructions should be inserted at the beginning of each region. The data distribution problem considered in this paper tries to explore the opportunity of the optimal data placement on SPMs in multi-core systems. It aims to improve the performance and reduce the cost of memory accesses for our execution model.

IV. Motivational Example

In this section, a motivational example is presented to illustrate the main idea of the proposed algorithm (IOLDD). The goal of optimization is to minimize the total memory access cost of a parallel iteration in loops.

In this motivational example, we assume all data items have the same size. Thus, the size of SPM is denoted by the number of data items that can be stored in SPM. Focus on the architecture shown in Fig. 1, we assume the system has only two cores $Core_1$ and $Core_2$. Each core is equipped with an on-chip SPM marked as SPM_1 and SPM_2, respectively. For the purpose of simplicity and illustration, we assume that SPM_1 has a capacity of two, and SPM_2 can hold three data items in the motivational example. All cores can access the shared main memory, which is large enough to store all data items.

```
for(i= 3; i< 10; i++)
{
    B[i]= A[i-1];
    A[i]= A[i-1]+ A[i-2]+ B[i];
    A[i-2]= A[i]+ A[i-1]- i;
}
```
(a) $Loop_1$ for $Core_1$

```
for(i= 3; i< 10; i++)
{
    A[i-3]= A[i-1]+ A[i-2];
    d2= A[i-1]+ A[i-2]- A[i-3];
    A[i-3]= d1+ A[i-1]* A[i-2];
}
```
(b) $Loop_2$ for $Core_2$

Figure 3. The loop programs for $Core_1$ and $Core_2$ in the example

Given two loop programs, as shown in Fig. 3, $Loop_1$ is assigned to $Core_1$ and $Loop_2$ is assigned to $Core_2$. There are seven data items to be accessed: two scalar data items d_1, d_2 and five array data items $A[i]$, $A[i-1]$, $A[i-2]$, $A[i-3]$ and $B[i]$. Each data item can be assigned to SPM_1 or SPM_2. Besides, each iteration is regarded as a program region, and in the execution model, the iteration in loop body forms a parallel region.

TABLE I.
NUMBERS OF DATA ACCESSES FOR EACH CORE

Data	d_1	d_2	d_3 (A[i])	d_4 (A[i-1])	d_5 (A[i-2])	d_6 (A[i-3])	d_7 (B[i])
$Core1_{Access}$	0	0	2	3	2	0	2
$Core2_{Access}$	1	1	0	3	3	3	0

The number of seven data accesses in one iteration is shown in Table I. Both $Core_1$ and $Core_2$ can access these data items and they run in parallel. The "Access" operation includes both "Read" and "Write" operations for the core. In Table I, take $Core_1$ for example, row "$Core1_{Access}$" shows the access times of $Core_1$ for the data items. The corresponding loop program in $Core_1$ is depicted in Fig. 3(a). For data d_1, we know $Core1_{Access}(d_1) = 0$ and $Core2_{Access}(d_1) = 1$. In $Loop_1$, $Core_1$ has neither "Read" operation nor "Write" operation for data d_1. While in $Loop_2$, data d_1 is read once. In this paper, we define a data which is "Read" by some cores and not "Written" by any core as a "Read-only" data. The problem of this example is how to find a data distribution for the seven data items in each iteration such that their total cost of memory access is minimized.

TABLE II.
NOTATIONS USED IN EXAMPLES

Notation	Time (μs)	Definition
$Read_{spm_i}$	1	the cost of reading from $Core_i$'s SPM.
$Write_{spm_i}$	1	the cost of writing to $Core_i$'s SPM.
$Read_{spm_j \to spm_i}$	20	the cost of reading from $Core_j$'s SPM to $Core_i$'s SPM.
$Write_{spm_j \to spm_i}$	20	the cost of writing from $Core_j$'s SPM to $Core_i$'s SPM.
$Read_M$	60	the cost of reading from main memory.
$Write_M$	60	the cost of writing to main memory.
$Migration_{spm_i \to spm_j}$	21	the cost of moving data from $Core_i$'s SPM to $Core_j$'s SPM.
$Migration_{spm_i \to M, M \to spm_i}$	61	the cost of moving data between $Core_i$'s SPM and main memory.

Table II shows the notation, the time cost (in μs) and the definition of memory operations. All data items are supposed in the main memory in the initial data distribution. In this example, we assume the non-decreasing

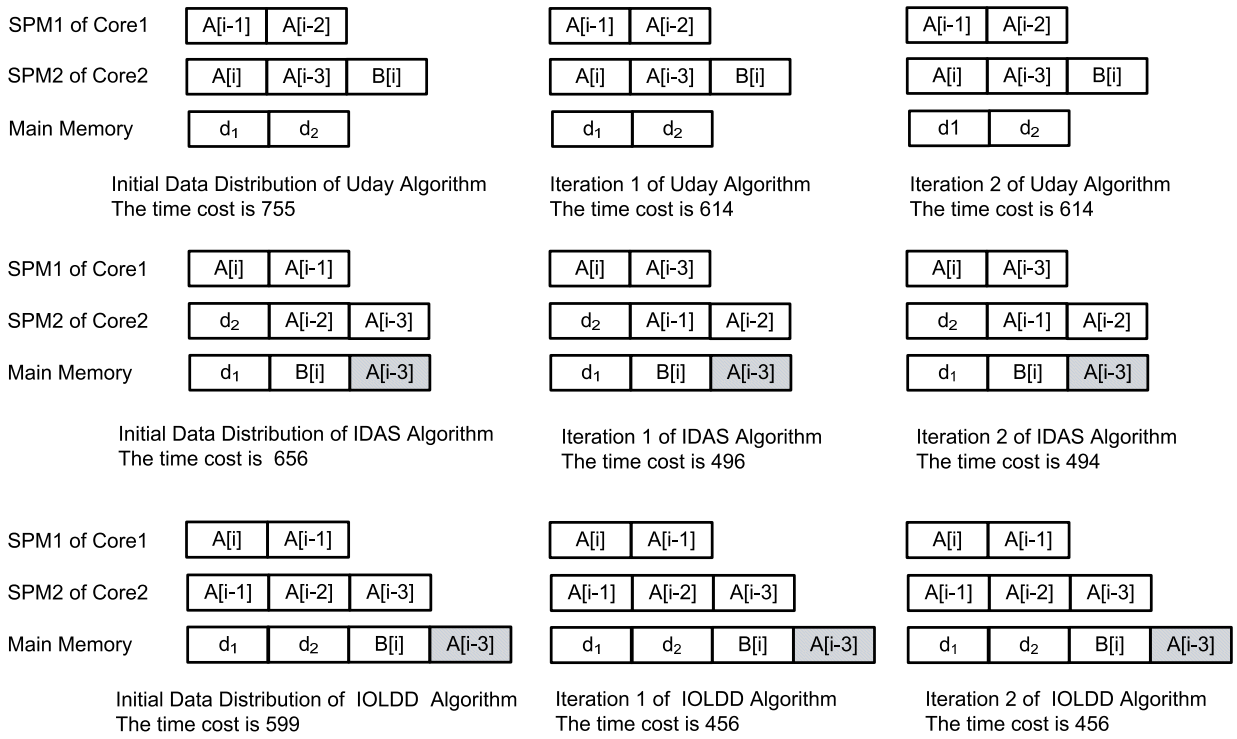


Figure 4. Costs and Distribution of each iteration with three different methods

remote SPM access cost function is a linear function $f(d) = 20*d$. Since there are two cores in the system, the distance d equals 1. Therefore, the remote SPM accessing costs $Read_{spm_j \to spm_i}$ and $Write_{spm_j \to spm_i}$ are 20.

We solve the problem with three strategies. One is the greedy strategy (Uday), which is derived from Udayakumaran's algorithm [12] [13] on single-core systems. The other is dynamic programming strategy on single-core (IDAS), which is derived from Zhang's paper [18]. Finally, it is our iterational optimal strategy (IOLDD), which will be presented in detail in Section VI.

The greedy algorithm is derived from Udayakumaran's algorithm in [12] [13], and we call it "Uday" for short. The algorithm is a greedy algorithm, it distributes data items according to their read and write access times of all cores. However, the approach only targets single-core processors. For the purpose of comparison, we adopt the algorithm and apply it on multi-core systems.

The derived Uday algorithm works as follows: To begin with, data items are sorted according to their total number of accesses which is the sum of the number of accesses from all cores to this data. After that, data with the most total number of accesses is picked by the compiler. Then the compiler distributes the data into the available SPM of the core that accesses the data most times. When all the SPMs of the cores are full, the data should be distributed into the main memory.

As to the IDAS algorithm, it is derived from Zhang's algorithm in [18]. Though the algorithm is a dynamic programming algorithm, it is limited in single-core systems. Besides, it can not fully utilize the benefits of the private local SPM on each core. On the other side, the IOLDD algorithm propose a duplication mechanism to utilize the distinct costs of data to/from local SPM, remote SPM and main memory. It is a technique that achieves higher time efficiency at the cost of space. In multi-core systems, traditionally, a data item has only one copy in either one of the SPMs or main memory. Multiple cores may access the same data in one parallel region. It is sometimes beneficial to duplicate data and place multiple copies of the same data in different SPMs. However, some of the data items cannot be duplicated because of high synchronization cost. Therefore, we only allow read-only data items to be duplicated.

TABLE III.
COST COMPARISON AMONG THREE DIFFERENT METHODS

Tech.	Initial Data Distribution		Iteration 1		Iteration 2	
	Cost	Imprv.	Cost	Imprv.	Cost	Imprv.
IOLDD	599	—	456	—	456	—
Uday	755	20.66%	614	25.73%	614	25.73%
IDAS	656	8.69%	496	8.06%	494	7.69 %

Results of the example with three strategies can be seen in both Table III and Fig. 4. The final data distribution for Uday algorithm results in a time cost of 614 μs. To avoid array data items in loop being distributed in different SPMs and hard to find, we use the technique of update and have the backups of arrays in main memory. As shown in Fig. 4, these updated data items are distinguished by gray boxes. For IDAS algorithm, data $A[i-3]$ needs to be updated to main memory. Since data $B[i]$ has already been in main memory, it does not need to be updated. Time cost can finally be reduced to 494 μs. For IOLDD algorithm, data $A[i-1]$ is located in both SPM_1 and SPM_2, we say data $A[i-1]$ is duplicated. The final result of time cost is stable at 456 μs, less than the costs of

other two algorithms. Thus, we can conclude that IOLDD algorithm works better than both the Uday algorithm and IDAS algorithm in the motivational example.

V. Problem Definition

In this section, definitions for the problem of loop data distribution on multi-core systems are presented. Firstly, the following notations and definitions are used:

Notation	Definition
D	a set of distinct data items for loops $\{d_1, d_2, ..., d_N\}$, where data can be either a scalar data or an array data item.
SPM	a set of SPMs on each core $\{spm_1, spm_2, ..., spm_T\}$, where T is a constant. Besides, m_0 is the main memory.
$Core$	a set of cores in hardware architecture $\{Core_1, Core_2, ..., Core_T\}$. Each core $Core_i$ has its own on-chip local SPM spm_i.
$CAcc_i(d_j)$	the number of access times for data d_j by $Core_i$.
$dist^P(d_j)$	a loop data distribution function for data d_j defined from D to SPM in one iteration P.
$Cost_{spm_i}(d_j)$	a cost of memory access for data d_j in spm_i without duplication.
$Cost_{spm_{i+k+...}}(d_j)$	a duplication cost of memory access for data d_j where d_j is in (spm_i, spm_k, ...).

For function $dist^P$: $D \rightarrow SPM$, if $dist^P(d_j)$ = spm_i, it indicates data d_j is distributed to spm_i on $Core_i$ in iteration P. Therefore, $dist^{P-1}(d_j)$ represents the location of data d_j before the execution of iteration P in loop body.

Loop Data Distribution Problem. Given a set of data items D in iteration P, a set of SPMs SPM. $Size_{spm_i}$ represents the capacity of $spm_i \in SPM$, $Size_{d_j}$ means the size of data d_j. The loop data distribution problem is to find a mapping between data $d_j \in D$ and SPM $spm_i \in SPM$. The total cost of memory access $Cost_{dist^P(d_j)}(d_j)$ is minimized, and inequality $\sum_{\forall d_j \in D, dist^P(d_j)=spm_i} Size_{d_j} \leq Size_{spm_i}$ is satisfied.

Definition 1: **Migration Cost**. The data migration cost is the cost of retrieving one data item from its original location and writing it to another memory unit. Hence, the migration cost is defined as the sum of one read operation from the original location and one write operation to the destination memory unit, as shown in Equation 1.

$$Migrate(d_j) = \begin{cases} Read_{spm_i} + Write_M \\ Read_{spm_i} + Write_{spm_j} \quad i,j = 1,2,...,T \\ Read_M + Write_{spm_i} \end{cases} \tag{1}$$

Related Array. An array is related when the array data item accessed in the previous iteration is still accessed in the current iteration. Take array A for example, if $A[i]$, ..., $A[i-j]$, ..., and $A[i-k]$ ($k \geq j \geq 1$) are accessed in loop at the same iteration i, We say that array A is related. The relation of data items in array A is defined as: $relation_{A[i]}$ = 0, ..., $relation_{A[i-j]}$ = 0, ..., $relation_{A[i-k+1]}$ = 0, and $relation_{A[i-k]}$ = k. The relation of array A is defined as $relation_A$ = k.

Unrelated Array. An array is unrelated when the array data item accessed in the previous iteration is not accessed in the current iteration. The relation of an unrelated array B is also defined as $relation_B$. It equals to -1. All data items in unrelated array are also unrelated. We define: $relation_{B[i]}$ = -1, here $B[i]$ is one data item from array B in iteration i.

Definition 2: **Updating Cost**. The updating cost is computed only for array items to ensure that there is a copy of that array with newest data value in main memory. It refers to the cost of reading array data item from its distributed location (except main memory), and writing it back to main memory. With the array updated into main memory, we do not need to seek scattered data items on different SPMs. It is convenient to access array data items because they are stored continuously in the main memory.

Data items in array may come from related arrays or unrelated arrays. To eliminate extra cost, only array data $A[i-relation_A]$ is updated to main memory. When array data is already in main memory, it does not need to be updated.

Theorem 1: In loop data distribution problem, array data d_j will be distributed to the main memory, if and only if $relation_{d_j} \neq 0$ and $dist^{P-1}(d_j) \neq m_0$.

$$Update(d_j) = \begin{cases} Read_{spm_i} + Write_M & \text{if } dist^{P-1}(d_j) \neq m_0 \\ & \text{and } relation_{d_j} \neq 0 \\ 0 & \text{else} \end{cases} \tag{2}$$

In Equation 2, spm_i is the location of array data d_j. Suppose in iteration i, array data $A[i]$, $A[i-1]$ and $A[i-2]$ are used: $relation_{A[i]}$ = 0, $relation_{A[i-1]}$ = 0 and $relation_{A[i-2]}$ = 2. After each iteration, we need to update $A[i-2]$, if it is not in the main memory, both $A[i]$ and $A[i-1]$ do not need to be updated. The updating cost is a sum of one read operation from local SPM spm_i and one write operation to main memory m_0, as shown in Equation 2.

Definition 3: **Cost of Memory Access for One Data Item without Duplication**. Let spm_i be the location of data item d_j, which is decided by the data distribution function $dist^P(d_j)$. Without duplication mechanism, the total cost of memory access for d_j is represented by a cost function $Cost_{spm_i}(d_j)$. It is the sum of data item's accessing cost for program execution, the migration cost of data d_j to a memory unit before the execution and the updating cost. The cost is computed by Equation 3.

Definition 4: **Moving Cost**. The moving cost is computed when data is going to be duplicated. It equals to the sum of migrating the data item from its original distributed location to the destination SPMs. We can compute the cost by Equation 4.

According to Equation 4, we can see that the moving cost is related to $dist^{P-1}(d_j)$. If data d_j was in one SPM before the execution of iteration P in loop, which is $dist^{P-1}(d_j) \neq m_0$. The moving cost is the sum of each

$$Cost_{spm_i}(d_j) = \sum_i CAcc_i(d_j) \times (Read_{spm_i} + Write_{spm_i}) + Migrate(dist^{P-1}(d_j), dist^P(d_j)) + Update(d_j) \quad (3)$$

$$Move(d_j) = \begin{cases} \sum_i Migrate(dist^{P-1}(d_j), spm_i) & \text{if } dist^{P-1}(d_j) \neq m_0 \\ Migrate(m_0, spm_i) + \sum_{i-1}(Migrate(spm_i, spm_k) & \text{if } dist^{P-1}(d_j) = m_0 \end{cases} \quad (4)$$

$$Cost_{spm_{i+k+...}}(d_j) = \begin{cases} \sum_i CAcc_i(d_j) \times (Read_{spm_i} + Write_{spm_i}) + Move(d_j) + Update(d_j) & \\ \infty & CWrt_i(d_j) \neq 0 \end{cases} \quad (5)$$

cost of migrating data from the original SPM $dist^{P-1}(d_j)$ to the destination SPMs. On the other hand, if data d_j was in main memory before the execution of iteration P, say $dist^{P-1}(d_j) = m_0$. The moving cost is the cost of migrating data from main memory to one destination SPM spm_i, plus a sum of other cost of migrating d_j from this spm_i, which has one copy already, to other destination SPMs spm_k (spm_i, $spm_k \in SPM$ and $k \neq i$).

Definition 5: **Duplication Cost**. The duplication cost is a cost of copying one data item to other SPMs on multi-core systems. It can be computed as the sum of accessing cost for "Read" or "Write" operations in local SPM, the moving cost to the destination SPMs and the updating cost for array data items.

Without the duplication mechanism, if a data item is intensively accessed by multiple cores, a lot of remote accesses will be incurred. Wherever the data is distributed, only one core is benefited from the local SPM. Data duplication technique will solve the problem by distributing a copy of the data item to each SPM that may be benefited. As a result, the time and energy cost incurred by remote accesses is reduced.

For exclusive copy mode, there is no worry about the data consistency. However, for data duplication mode, the data consistency problem becomes a key issue. Since it is common for multiple cores to access the same data in one parallel region, inconsistency of this data will occur if multiple cores have "Write" activities. Though in write heavy applications, duplicating to-be-written data may be beneficial with a well-designed data consistency protocol, the overhead caused by maintaining data consistency may offset the benefits of duplicating written data. Therefore, only read data is allowed to be duplicated. We define $CWrt_i(d_j)$ to be the number of "Write" times for data d_j by $Core_i$. If a data item d_j is updated by a core, it should not be replicated on other cores because of synchronization issues. Hence, the duplication cost of a data item when $CWrt_i(d_j) \neq 0$ is ∞.

No matter whether an array data can be duplicated or not, it should be updated into the main memory. The updating cost can be computed in Equation 2. The total duplication cost is computed as Equation 5.

VI. A Dynamic Algorithm for Loop Data Distribution With Duplication on Multi-core

In this section, we present details of the IOLDD algorithm on multi-core systems. The algorithm is a dynamic programming method and uses the technique of duplication.

Definition 6: **Total Cost.** Let $Tcost[j, i_1, i_2, ..., i_T]$ be the current total cost of all data items when data d_j, (j=1, 2, ..., N), is considered to be distributed in a certain SPM when there are still i_k available space in SPM spm_k, $spm_k \in SPM$ and k=1, 2, ..., T.

The recursive function of dynamic programming is shown in Equation 6. Suppose the distribution of the first $j-1$ data items (from d_1 to d_{j-1}) have been optimally determined. The value of $Tcost[j, i_1, i_2, ..., i_T]$ is computed as the minimal total cost of memory access for the iteration when data d_j is distributed to a certain SPM spm_i or duplicated. Assume that the number of cores allowed to share a SPM is t, we compare duplication costs on 1 to t cores, i.e. $Cost_{m_0}(d_j)$, $Cost_{spm_i}(d_j)$, $Cost_{spm_{i+k}}(d_j)$,..., and $Cost_{spm_{i+k+...+t}}(d_j)$, and choose the minimum cost for accessing data item d_j. If $Tcost[j, i_1, i_2, ..., i_T]$ is minimum when d_j is just in spm_k, we say d_j is distributed to spm_k. The available space size of spm_k reduces 1. Or if $Tcost[j, i_1, i_2, ..., i_T]$ is computed as the minimal total cost when data d_j is in both spm_k and spm_m, we consider d_j is duplicated and the available space size of both spm_k and spm_m minus 1. The remaining data items $(d_{j+1}, d_{j+2}, \ldots, d_N)$ reside in their previous locations.

In our IOLDD algorithm, each iteration includes four steps. First, it computes the cost of each memory access. Second, it uses dynamic programming with duplication to decide optimal loop data distribution. Third, it redistributes data items, and the fourth, it stops the algorithm. The algorithm is shown in Algorithm 1.

> **Step 1.** Lines 1-4 build a cost table of memory access $Cost_{spm_i}(d_j)$ and duplication cost $Cost_{spm_{i+k+...}}(d_j)$ for $\forall d_j \in D$ on main memory and each SPM of each core. At the beginning, $Tcost[j, Size_{spm_1}, Size_{spm_2}, ..., Size_{spm_T}]$ is the sum of $Cost_{m_0}(d_j)$, means all data items are stored in main memory.

The cost of memory access $Cost_{spm_i}(d_j)$ is computed for every item d_j on all SPM spm_i and main memory m_0 by Definition 3. Data d_j is in the set of data items $D = (d_1, d_2, \ldots, d_N)$, and spm_i is in the list of SPMs $SPM = (spm_1, spm_2, \ldots, spm_T)$. For array items, we add an updating cost (Definition 2) to ensure that in main memory the copy of newest data in array can

$$Tcost[j,i_1,...,i_T] = \begin{cases} \sum_j Cost_{m_0}(d_j), & \text{if } j=0, \\ & \forall k=1,2,...,T,\ i_k = Size_{spm_k}, \\ \infty & \text{if } \sum_{k=1}^{T} i_k + t < \sum_{k=1}^{T} Size_{spm_k} - j, \\ & \text{or } \exists k \in \{1,2,3,\ldots,T\}\ i_k > Size_{spm_k}, \\ \min(Tcost[j-1,i_1,i_2,...,i_T] + Cost_{m_0}(d_j), & \\ Tcost[j-1,i_1,...,i_p+1,...,i_T] + Cost_{spm_p}(d_j), & \\ ..., & \\ Tcost[j-1,i_1,...,i_p+1,...,i_q+1,...,i_T] + Cost_{spm_{p+q}}(d_j), & \\ ..., & \text{if } \sum_{k=1}^{T} i_k + t \geq \sum_{k=1}^{T} Size_{spm_k} - j. \\ Tcost[j-1,i_1,...,i_p+1,...,i_q+1,...,i_r+1,...,i_T] + Cost_{spm_{p+q+r}}(d_j), & \\ ..., & \\ ) - Cost_{m_0}(d_j) & \end{cases} \tag{6}$$

Algorithm 1 Iterational Optimal Loop Data Distribution Algorithm with duplication (IOLDD)

Require: A set of loop data items $D = (d_1, d_2, \ldots, d_N)$, a set of SPM units $SPM = (spm_1, spm_2, \cdots, spm_T)$, m_0 is main memory, $Size_{spm_i}$ $\forall spm_i \in SPM$, and the initial distributions for all d_j in D.

Ensure: A data distribution under which the total cost of the execution for each iteration in parallel loops is regionally minimized.

1: Compute cost function $Cost_{spm_i}(d_j)$, $\forall d_j \in D$ and $\forall spm_i \in SPM$
2: **for** $j \leftarrow 1$ to $|D|$ **do**
3: Tcost[j,$Size_{spm_1}$,$Size_{spm_2}$,...,$Size_{spm_T}$]$\leftarrow\sum Cost_{m_0}(d_j)$
4: **end for**
5: **for** $j \leftarrow 1$ to $|D|$ **do**
6: **for** $i_1 \leftarrow Size_{spm_1}$ to 0 **do**
7: ...
8: **for** $i_T \leftarrow Size_{spm_T}$ to 0 **do**
9: Compute Tcost[j,i_1,i_2, ...,i_T] according to the recursive Equation 6 to get the minimum Tcost[];
10: **end for**
11: ...
12: **end for**
13: **end for**
14: **for** $j \leftarrow |D|$ to 1 **do**
15: **if** Tcost[j,i_1,i_2,...,i_T] = Tcost[j-1,i_1,i_2,...,i_T] **then**
16: $location_{d_j} \leftarrow 0$
17: BackPath[j,i_1,i_2,...,i_T] = (j-1,i_1,i_2,...,i_T)
18: Continue
19: **end if**
20: **for** $k \leftarrow 1$ to $|T|$ **do**
21: **if** Tcost[j,i_1,i_2,...,i_T] = Tcost[j-1,i_1,...,i_k+1,...,i_T]-$Cost_{m_0}(d_j)$+$Cost_{spm_k}(d_j)$ **then**
22: $location_{d_j} \leftarrow k$
23: BackPath[j,i_1,i_2,...,i_T] = (j-1,i_1,...,i_k,...,i_T)
24: $i_k \leftarrow i_k + 1$
25: Continue
26: **end if**
27: **end for**
28: **for** $k \leftarrow 1$ to $|T-1|$ **do**
29: **for** $m \leftarrow$ k+1 to $|T|$ **do**
30: **if** Tcost[j,i_1,i_2,...,i_T]= Tcost[j-1,i_1,...,i_k+1,...,i_m+1,...,i_T]-$Cost_{m_0}(d_j)$ +$Cost_{spm_{k+m}}(d_j)$ **then**
31: $location_{d_j} \leftarrow j$
32: $location_{d_j} \leftarrow m$
33: BackPath[j,i_1,i_2,...,i_T]= (j-1,i_1,...,i_k+1,...,i_m+1,...,i_T)
34: $i_k \leftarrow i_k + 1$
35: $i_m \leftarrow i_m + 1$
36: Continue
37: **end if**
38: **end for**
39: **end for**
40: **end for**
41: Redistribute the data items' locations for the next iteration.
42: **if** Current iteration distribution = Previous iteration distribution **then**
43: Stop IOLDD algorithm forever.
44: **end if**

be accessed continually. Besides, the duplication cost $Cost_{spm_{i+k+...}}(d_j)$ is computed by Definition 4 and 5. The cost table in initial data distribution iteration for the motivational example is shown in Table IV.

TABLE IV.
COST TABLE OF IOLD ALGORITHM IN INITIAL DATA DISTRIBUTION

Data	$Cost_{m_0}(d_j)$	$Cost_{spm_1}(d_j)$	$Cost_{spm_2}(d_j)$	$Cost_{spm_{1+2}}(d_j)$
d_1	60	81	62	62
d_2	60	81	62	∞
A[i]	120	63	101	∞
A[i-1]	360	124	124	67
A[i-2]	300	123	104	∞
A[i-3]	180	182	125	∞
B[i]	120	124	162	∞

Since we have only two cores, the duplication cost can be written as $Cost_{spm_{1+2}}(d_j)$. The result of optimal data distribution is shown in "Iteration 2" of Fig. 4. Based on the number of accesses in Table I, if all data items have their initial locations in the main memory, costs of memory access are computed in Table IV. Besides, according to the loop programs assigned to $Core_1$ and $Core_2$ in Fig. 3, we can see that data d_1 and $A[i-1]$ are read-only data items. Data d_1 and $A[i-1]$ then can be duplicated on multi-core systems, the others have their duplication costs to be ∞. For example, data d_1 is one scalar data item, then the updating cost of d_1 is 0. According to Equation 3,

$Cost_{m_0}(d_1)$=$1 \times 60 + 0 + 0 = 60$,
$Cost_{spm_1}(d_1)$=$0 \times 1 + 1 \times 20 + 61 + 0 = 81$,
$Cost_{spm_2}(d_1)$=$0 \times 20 + 1 \times 1 + 61 + 0 = 62$,
$Cost_{spm_{1+2}}(d_1)$=$0 \times 1 + 1 \times 1 + 61 + 0 = 62$.

Data d_6 $(A[i-3])$ is one array data item with its $relation_{A[i-3]} = 3$, then it needs to be updated to main memory, then in the "Initial Data Distribution" iteration, costs of d_6 can be computed as:

$Cost_{m_0}(d_6)$=$0 \times 60 + 3 \times 60 + 0 + 0 = 180$,
$Cost_{spm_1}(d_6)$=$0 \times 1 + 3 \times 20 + 0 + 61 + 61 = 182$,
$Cost_{spm_2}(d_6)$=$0 \times 20 + 3 \times 1 + 61 + 61 = 125$,
$Cost_{spm_{1+2}}(d_6)$=∞.

Step 2. Lines 5-40 determine the optimal loop data distribution with dynamic programming and duplication technique. The cost of memory access during the execution for one iteration is minimum.

Since a cost table is built in step 1, the optimal data distribution can be determined by using a multi-dimensional dynamic programming table. The structure of the table is as following: the first dimension of the table is represented by data d_j, each one of the other dimensions is represented by the available space of $spm_i \in SPM$ except m_0. We assume m_0 is main memory and large enough to hold all data items in the program.

The IOLDD algorithm is presented in Algorithm 1. During the computation, $location_{d_j}$ is used to keep the intermediate data, and the array $BackPath[j, i_1, i_2, ..., i_T]$ is introduced to trace back an optimal solution. The array $BackPath[j, i_1, i_2, ..., i_T]$ keeps a list of the state of data distribution in one iteration before data d_j is distributed.

TABLE V.
THE DYNAMIC PROGRAMMING TABLE THAT COMPUTES THE COST ARRAY C$[j, i_1, i_2]$ FOR THE EXAMPLE IN INITIAL DATA DISTRIBUTION OF SECTION IV

i_2 = 3

j / i_1	d_1	d_2	A[i]	A[i-1]	A[i-2]	A[i-3]	B[i]
2	***1200***	***1200***	1200	1200	1200	1200	1200
1	1221	1221	***1143***	964	964	964	964
0	∞	1242	1164	907	787	787	787

i_2 = 2

j / i_1	d_1	d_2	A[i]	A[i-1]	A[i-2]	A[i-3]	B[i]
2	1202	1202	1181	964	964	964	964
1	∞	1202	1143	907	768	768	768
0	∞	∞	1145	***850***	711	711	711

i_2 = 1

j / i_1	d_1	d_2	A[i]	A[i-1]	A[i-2]	A[i-3]	B[i]
2	∞	1204	1183	945	768	768	768
1	∞	∞	1145	888	711	711	711
0	∞	∞	∞	850	***654***	654	654

i_2 = 0

j / i_1	d_1	d_2	A[i]	A[i-1]	A[i-2]	A[i-3]	B[i]
2	∞	∞	1185	947	749	713	713
1	∞	∞	∞	890	692	656	656
0	∞	∞	∞	∞	654	***599***	***599***

To illustrate our IOLDD algorithm, we construct a dynamic programming table for the example's final data distribution result in initial data distribution in Section IV with duplication. Since there are two SPMs assumed in the example, the dynamic programming table is constructed as a 3-D table of $Tcost[j, i_1, i_2]$. The 3-D table consists of four 2-dimensional tables as shown in Table V. The value of i_1 and i_2 indicates the available space on spm_1 and spm_2. The 2-D table of "$i_2 = 3$" computes the costs for all data items when available space on spm_2 equals 3. For example, we compute the cost cell $Tcost[1, 1, 3]$ for data d_1 when there is one available space in spm_1 of $Core_1$ and three available space in spm_2 of $Core_2$. The cost is 1221, as shown in row "$i_1 = 1$" and column "d_1" in the first 2-D table. Then, we compute the cost cell $Tcost[1, 0, 3]$. According to equation 6: $\sum_{k=1}^{T} i_k < \sum_{k=1}^{T} Size_{spm_k} - j$, we know $i_1 + i_2 < Size_{spm_1} + Size_{spm_2} - j$, then $Tcost[1, 0, 3] = \infty$. We compute costs for data item d_1 in columns "d_1" in all 2-D tables in a similar way. The final total cost of memory access with the optimal data distribution in the initial data distribution is 599. The backtracking path indicated by underlined cell in Table V shows one of the optimal distributions as follows: data $A[i]$ and $A[i-1]$ are assigned in spm_1 of $Core_1$, data $A[i-1]$, $A[i-2]$ and $A[i-3]$ are assigned in spm_2 of $Core_2$, data d_1, d_2 and $B[i]$ are assigned in main memory, data $A[i-3]$ is updated to the main memory.

Step 3. Line 41 redistributes data items for the current iteration to make sure that at the beginning of the next iteration in loop, locations of data items still remain accurate.

Redistributing array items can guarantee these array items' locations still being right in the next iteration. Take array $A[i]$ for example, in iteration i, data $A[i]$ and data $A[i-1]$ are accessed; In iteration t, data $A[t]$ and data $A[t-1]$ will be accessed. Obviously, if $t = i + 1$, then $A[i]$ becomes $A[t-1]$, $A[i-1]$ will never be accessed, and $A[t]$ is a new data item to be accessed. Certainly, memory location of $A[t-1]$ should be the location of $A[i]$ instead of $A[i-1]$. Therefore, we have to distribute data items to the right location before the next iteration.

Step 4. Lines 42-44 compare the results with previous iteration to decide when to stop the algorithm.

At the end of our IOLDD algorithm in each iteration, we compare the distribution results with the results in previous iteration. If they are the same, stop doing the algorithm in loops. The data distribution is optimal and we could use that distribution for later iterations.

With the four steps mentioned above, we can conclude that the IOLDD algorithm has five features. Firstly, it computes to obtain the cost table of each memory access and duplication cost. Besides, array items in loops can be handled by the algorithm. Furthermore, the relations of array items are used to reduce the updating costs. Fourthly, it redistributes array items. Finally, duplication is proposed on multi-core systems.

The IOLDD algorithm's time complexity is O($N \times Size_{spm_1} \times Size_{spm_2} \times \ldots \times Size_{spm_T}$). N is the number of data items in D, and T is the constant number of SPMs.

VII. EXPERIMENT

In this section, we compare our dynamic IOLDD algorithm with greedy algorithm and random algorithm, respectively. Benchmarks are chosen with both scalar and array data items in loop programs. The following benchmarks are used in our experiments: 2IIR, 4-lattice, ellENC, ellfilter, 8-lattice, allpole, C-sehwa, diff2, diff-ct1 and voltera. Three algorithms are evaluated by time cost and energy consumption for memory access with their generated distributions.

A. Experimental Setup

The architecture in the experiment has two types of memory units: four on-chip SPMs made with SRAM of four cores and a block of main memory made with SRAM. In this paper, we assume these four SPMs have

TABLE VI.
SYSTEM SPECIFICATION FOR SPM, L2 MEMORY AND MAIN MEMORY

Component		Description
CPU	**Core**	Number of cores: 4, frequency: 1.0 GHz
SRAM	**Main memory**	Size: 2.56 MB, access latency: 1.3882 ns, access energy: 0.7189 nJ
SRAM	**SPMs**	Size: 8 KB, Local SPM access latency: 0.2109 ns, Local SPM access energy: 0.0252 nJ
		Remote SPM access latency: 0.5275 ns, Remote SPM access energy: 0.0441 nJ

the same size, with the capacity of 8 KB. We also assume the capacity of the off-chip main memory is 2.56 MB, which is large enough to store all data items. A set of parameters collected from CACTI tools provided by HP for these two memory types is shown in Table VI.

A custom simulator based on SimpleScalar is developed to simulate the process of data distribution and obtain costs of memory accesses for the program. With the CACTI tools [23] provided by HP, we can obtain the latency and energy consumptions for memory accesses, then we use them as parameters on our benchmark programs. In the experiment, we run random, greedy and the dynamic programming algorithm IOLDD. Compute the time cost and energy consumption for all data accesses. Our program is easy to compatibly integrate into any compiler.

B. Experimental Results

In Fig. 5, algorithms are compared via ten benchmarks include: 2IIR, 4-lattice, ellENC, ellfilter, 8-lattice, allpole, C-sehwa, diff2, diff-ct1 and voltera. Time costs and energy consumptions of data distributions on multi-core systems are also presented. In Tab. VII, the column "Random" represents the algorithm that data items are randomly picked and distributed to four on-chip SPMs of the cores. Due to the randomness of the technique, the experiment is precessed 10 times to get an average number. The column "Uday" is a greedy algorithm proposed by the Udayakumaran's algorithm [12]. The column "IOLDD" denotes our IOLDD algorithm. It is a dynamic programming algorithm applied into both array and scalar data items in loop on multi-core systems. The percentage of improvement for IOLDD algorithm over the "Random" technique is shown in column "Imprv (IOLDD/Random)". The average improvement of time costs of the IOLDD algorithm is 18.45%. Moreover, "Imprv (IOLDD/Uday)" displays the improvement for "IOLDD" algorithm over "Uday" algorithm, and the average improvement is 18.38%. As shown in the experimental results, our IOLDD algorithm achieves the best improvement of time costs on average among all other techniques. In the best case, e.g. C-sehwa, the percentage of improvement in time cost is 52.12% over "Random" algorithm and 54.26% over "Uday" algorithm.

Not only data accesses time is reduced because of the effective solution of data distribution, but also the energy consumption lessens. Both Fig. 5 and Tab. VIII show the comparison of energy consumption among various data distribution solutions that are generated by various techniques. Accordingly, the average improvement of IOLDD algorithm over random technique is 30.12%, and the average improvement of IOLDD algorithm over Uday algorithm is 14.52%, our IOLDD algorithm also achieves the best improvement of energy consumption. In the best case, e.g. C-sehwa, the percentage of improvement in energy consumption is 64.63% over "Random" algorithm and 57.99% over Uday algorithm.

Experiments indicate that IOLDD algorithm obtains better improvements in most of the time and energy costs compared with the Uday algorithm which employs a greedy strategy. The major advantages in techniques compared with the Uday algorithm can be seen in three parts. First, the IOLDD algorithm considers the initial data distribution of the loops and the effect of migrating data items. Therefore, it can generate the optimal solution for iterational data distribution. Second, the IOLDD algorithm takes the cost of updating array items into account and uses the relations of these array items to minimize the cost, while Uday algorithm does not. Third, the IOLDD algorithm uses the technique of duplication to reduce the cost.

VIII. Conclusion and Future Work

In this paper, we achieve the minimum cost of loop data distribution on multi-core systems by developing an Iterational Optimal Loop data Distribution algorithm with Duplication (IOLDD). The algorithm is used on multi-core systems, while the algorithm Zhang *et al*. proposed is limited to single-core systems and cannot fully utilize the benefits of the private SPM on each core [18]. The IOLDD algorithm improves the performance on multi-core systems by taking the relations of an array, the updating cost and the duplication technique into consideration.

To explore more, in the future, we will further consider the problem of distributing loop data and avoiding contention when there are multiple "Write" activities on the same data for multiple cores. As to the mechanism of duplication, we will explore whether data still can be duplicated on condition that data is not read-only, and how to duplicate it. Besides, we will consider the problem when the main memory is a mixed structure of DRAM and non-volatile main memory (NVM). We will also deal with the problem of reducing the "Write" activities to the NVM in a heterogeneous system.

Acknowledgment

This work is partially supported by National 863 Program 2013AA013202, Chongqing cstc2012ggC40005, NSFC 61173014, NSF CNS-1015802, Texas NHARP 009741-0020-2009.

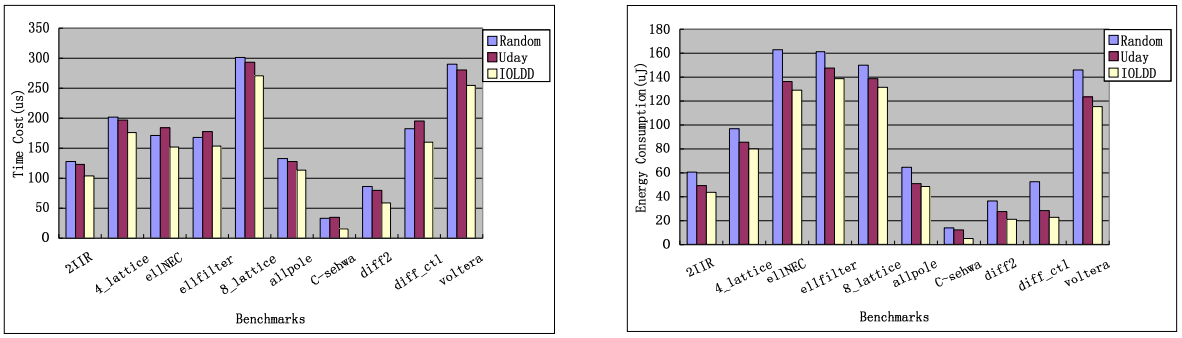


Figure 5. The corresponding time and energy results in the experiment.

TABLE VII.
COMPARISON OF TIME COSTS FOR VARIOUS DATA DISTRIBUTION ON MULTI-CORE SYSTEMS

Bench.	*Uday(μs)*	*IOLDD(μs)*	*Random(μs)*	*imprv(IOLDD/Random)*	*imprv(IOLDD/Uday)*
2IIR	127.529	123.253	103.512	18.83%	16.02%
4-lattice	201.754	196.25	175.542	12.99%	10.55%
ellNEC	170.923	184.623	151.371	11.44%	18.01%
ellfilter	167.907	177.599	153.177	8.45%	13.45%
8-lattice	301.091	293.28	270.973	10.00%	7.61%
allpole	133.276	127.262	113.757	14.65%	10.61%
C-sehwa	33.708	35.276	16.135	52.12%	54.26%
diff2	86.786	80.113	59.106	31.89%	26.22%
diff-ct1	182.001	194.498	160.334	11.90%	17.57%
voltera	289.184	280.551	253.902	12.20%	9.50%
average				18.45%	18.38%

TABLE VIII.
COMPARISON OF ENERGY COSTS FOR VARIOUS DATA ALLOCATION ON MULTI-CORE SYSTEMS

Bench.	*Uday(μJ)*	*IOLDD(μJ)*	*Random(μJ)*	*imprv(IOLDD/Random)*	*imprv(IOLDD/Uday)*
2IIR	60.907	49.595	43.800	28.09%	11.68%
4-lattice	96.981	85.855	79.562	17.96%	7.33%
ellNEC	162.593	136.499	129.285	20.49%	5.29%
ellfilter	160.928	147.844	138.316	14.05%	6.44%
8-lattice	150.268	138.562	131.525	12.47%	5.08%
allpole	64.624	50.679	131.525	12.47%	5.08%
C-sehwa	14.295	12.037	5.057	64.63%	57.99%
diff2	36.404	27.315	21.060	41.15%	22.90%
diff-ct1	52.201	28.084	23.122	55.71%	17.67%
voltera	145.837	123.716	115.588	20.74%	6.57%
average				30.12%	14.52%

REFERENCES

[1] http://www.analog.com/static/imported-files/data_sheets/ADSP-BF534_BF536_BF537.pdf, Analog Devices Inc., 2010.

[2] http://www-s.ti.com/cs/psheets/spns034c/spns034c.pdf, Texas Instruments Inc., 2012.

[3] http://www.nvidia.com/page/8800_tech_briefs.html, nVIDIA Co., 2013.

[4] R. Banakar, S. Steinke, B.-S. Lee, M. Balakrishnan, and P. Marwedel, "Scratchpad memory: a design alternative for cache on-chip memory in embedded systems," in *Hardware/Software Codesign, 2002. Proceedings of the Tenth International Symposium on*, 2002, pp. 73–78.

[5] L. Zhang, M. Qiu, W. Tseng, and E. Sha, "Variable partitioning and scheduling for mpsoc with virtually shared scratch pad memory," *Journal of Signal Processing Systems*, vol. 58, no. 2, pp. 247–265, 2010.

[6] O. Avissar, R. Barua, and D. Stewart, "An optimal memory allocation scheme for scratch-pad-based embedded systems," *ACM Trans. Embed. Comput. Syst.*, vol. 1, no. 1, pp. 6–26, Nov. 2002.

[7] P. R. Panda, N. D. Dutt, and A. Nicolau, "Efficient utilization of scratch-pad memory in embedded processor applications," in *Proceedings of the 1997 European conference on Design and Test*, ser. EDTC '97, 1997.

[8] P. Panda, D. Dutt, and A. Nicolau, "On-chip vs. off-chip memory: the data partitioning problem in embedded processor-based systems," *ACM Trans. Des. Autom. Electron. Syst.*, vol. 5, no. 3, pp. 682–704, Jul. 2000.

[9] Y. Zhou, Q. Zhu, and Z. Y., "Spatial data dynamic balancing distribution method based on the minimum spatial proximity for parallel spatial database," *Journal of Software*, vol. 6, no. 7, pp. 1337–1344, 2011.

[10] Q. Zhu and Y. Zhou, "Automatic connection of components for dynamic distributed applications," *Journal of Software*, vol. 8, no. 6, pp. 1419–1427, 2011.

[11] Y. Xu, Z. Zhao, W. Wu, and Y. Zhao, "Rppa: A remote parallel program performance analysis tool," *Journal of Software*, vol. 6, no. 12, pp. 2399–2406, 2011.

[12] S. Udayakumaran, A. Dominguez, and R. Barua, "Dynamic allocation for scratch-pad memory using compile-time decisions," *ACM Trans. Embed. Comput. Syst.*, vol. 5, no. 2, pp. 472–511, May 2006.

[13] S. Udayakumaran and R. Barua, "Compiler-decided dynamic memory allocation for scratch-pad based embedded systems," in *Proceedings of the 2003 international*

conference on Compilers, architecture and synthesis for embedded systems, ser. In CASES '03, 2003, pp. 276–286.
[14] O. Ozturk, M. Kandemir, and I. Kolcu, "Shared scratch-pad memory space management," in *Quality Electronic Design, 2006. 7th International Symposium on*, Mar. 2006, pp. 576–584.
[15] O. Ozturk, M. Kandemir, and S. Narayanan, "A scratch-pad memory aware dynamic loop scheduling algorithm," in *Quality Electronic Design, 2008. 9th International Symposium on*, Mar. 2008, pp. 738–743.
[16] R. Thakur, A. Choudhary, and J. Ramanujam, "Efficient algorithms for array redistribution," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 7, no. 6, pp. 587–594, 1996.
[17] W. Huan, Z. Yang, M. Chen, and L. Ming, "Energy-oriented dynamic spm allocation based on time-slotted cache conflict graph," in *Design, Automation Test in Europe Conference Exhibition (DATE), 2010*, 2010, pp. 598–601.
[18] J. Zhang, T. Deng, Q. Gao, Q. Zhuge, and E.-M. Sha, "Optimizing data allocation for loops on embedded systems with scratch-pad memory," in *Embedded and Real-Time Computing Systems and Applications (RTCSA), 2012 IEEE 18th International Conference on*, aug. 2012, pp. 184 –191.
[19] Y. Guo, Q. Zhuge, J. Hu, M. Qiu, and E.-M. Sha, "Optimal data allocation for scratch-pad memory on embedded multi-core systems," in *Parallel Processing (ICPP), 2011 International Conference on*, Sep. 2011, pp. 464–471.
[20] Y. Guo, Q. Zhuge, J. Hu, and E.-M. Sha, "Optimal data placement for memory architectures with scratch-pad memories," in *Trust, Security and Privacy in Computing and Communications (TrustCom), 2011 IEEE 10th International Conference on*, nov. 2011, pp. 1045 –1050.
[21] O. Ozturk, M. Kandemir, G. Chen, M. Irwin, and M. Karakoy, "Customized on-chip memories for embedded chip multiprocessors," in *Design Automation Conference, 2005. Proceedings of the ASP-DAC 2005. Asia and South Pacific*, vol. 2, 2005, pp. 743–748 Vol. 2.
[22] M. Kandemir, J. Ramanujam, and A. Choudhary, "Exploiting shared scratch pad memory space in embedded multiprocessor systems," in *Proceedings of the 39th annual Design Automation Conference*. ACM, 2002, pp. 219–224.
[23] "Cacti model." http://www.hpl.hp.com/research/cacti/.

Qiuyan Gao received her BSc in information security from Hunan University, China, in 2011. She is currently a Master Candidate at Hunan University of Communication Engineering, China. Her current research interest includes parallel architecture, embedded multi-core systems and optimization algorithms.

Qingfeng Zhuge received her Ph.D. from the Department of Computer Science at the University of Texas at Dallas in 2003. She obtained her BS and MS degrees in Electronics Engineering from Fudan University, Shanghai, China. She is now a full professor at Chongqing University, China.

She received Best Ph.D. Dissertation Award in 2003. She has published more than 60 research articles in premier journals and conferences. Her research interests include parallel architectures, embedded systems, supply-chain management, real-time systems, optimization algorithms, compilers and scheduling.

Jun Zhang received his BSc and MSc in computer science from Hunan University, China, in 2011 and 2013. He is currently a PhD candidate in Chongqing University, China. His research interests include high-performance computer and multi-core embedded systems, real-time systems and information security.

Guanyu Zhu is currently a researcher and leader of storage and software group at Huawei Research Center. His research interests include clustered storage architectures, design of memory architecture with non-volatile memory, storage management, and file systems.

Edwin Hsing-Mean Sha received Ph.D. degree from the Department of Computer Science, Princeton University, Princeton, NJ, in 1992. Since 2000, he has been a tenured full professor in the Department of Computer Science at the University of Texas at Dallas. Since 2012, he served as the Dean of College of Computer Science at Chongqing University, China. He has published more than 300 research papers in refereed conferences and journals. He has served as editors for many journals, and as program committee and Chairs for numerous international conferences.

He received Oak Ridge Association Junior Faculty Enhancement Award, Teaching Award, Microsoft Trustworthy Computing Curriculum Award, NSF CAREER Award and NSFC Overseas Distinguished Young Scholar (B) Award, Chang Jiang Honorary Chair Professorship and China Thousand-Talent Program. His web page can be found in http://www.utdallas.edu/ edsha.

Adaboost Face Detection Based on Improved Covariance Feature

Rui Li
Lanzhou University of Technology//School of Computer and Communication, Lanzhou, China
Email: ruililut@126.com

Changfeng Li
Lanzhou University of Technology//School of Computer and Communication, Lanzhou, China
Email: changfengli2013@sina.cn

***Abstract*—Excessive number of Haar-like features and the complex threshold calculation of covariance matrix feature are two key issues of Adaboost face detection. In this paper, an efficient feature named covariance feature is proposed. The novel method divides the face image into several regions and it calculate covariance feature of any two regions. Then optimal weak classifiers will be picked out by Adaboost algorithm and they will be composed to a strong classifier. The experiments result in MIT+CMU data sets shows that the feature extraction times of the novel method is slightly slower than covariance matrix feature. However, the feature threshold is obtained much faster than covariance matrix feature, leading the significant reduction of the training time of Adaboost algorithm. Comparing with the Haar-like feature, the detection rate and speed improved obviously.**

***Index Terms*—face detection, covariance feature, Adaboost, feature extraction**

I. INTRODUCTION

Face detection(FD) refers to the process of defining the position, size, posture of face(if it exists) on the input image. As a special case of object detection, face detection problem has attracted extensive attention of researchers in last two decades due to its widely application in various fields. The fields are human-computer interfaces, content retrieval, digital video processing, visual monitoring[1] and so on. In recent years a lot of face detection methods appeared[2-3], and the statistical learning methods has become mainstream in pattern recognition field gradually. In 2001, Viola proposed Adaboost face detection based on Haar-like feature(HLF)[4]. This method can obviously improve the detection rate and speed .

Adaboost is an iterative algorithm and the core idea is training weak classifiers from a training set, then gather some suitable weak classifiers to a strong classifier. The calculation of sample weight is according to classification is correct or not and the overall classification accuracy of last time. Then training the modified new data sets with the next layer. Finally grouping the best classifiers of each layer into the final decision classifier. Using Adaboost classifier can eliminate some unnecessary features and focus on key data training.

HLF calculate face feature by integral image which can boost computational efficiency. However, HLF is a coarse feature, it is sensitive to edge, line and can only describe specific graph structure. Recently, covariance matrix feature(CMF) is proposed[5], it can reflect the intrinsic relevance of image pixels, and overcome the shortcomings of HLF . Nonetheless, the detection time is three times slower than HLF because the threshold calculation is too complex.

Therefore, this paper proposed the covariance feature (CF) to improve detection speed based on CMF. Comparing with CMF, the detection speed was greatly improved and CF is about CMF detection rate. Comparing with HLF, the face detection rate was improved and elapsed time was less than HLF. This novel feature is concerned with two regions instead of matrix feature of four regions. It vastly reduces features and simplify the calculation of matrix feature threshold, thereby it improve speed of Adaboost face detection .

II. RELATED WORKS

In classical Adaboost algorithm the weak classifier used Haar-like features which named after similarity to Haar wavelet. HLF is a kind of simple rectangle feature and it extract face feature fast. HLF value refers to the difference of gray level sum between the black rectangle and white rectangle on the test image. It reflects the differences of local image gray.

Figure 1 gives five basic templates of HLF feature. We can matching the feature template in every position of detection and achieve a HLF feature. The matrix in image window should meet the S-T criteria[6]. The formula of quantity calculation is as the follow:

$$\Omega^{n}_{(s,t)} = \left(\left\lceil \frac{m}{s} \right\rceil + \left\lceil \frac{m-1}{s} \right\rceil + \cdots \left\lceil \frac{s+1}{s} \right\rceil + 1 \right) \bullet \left(\left\lceil \frac{m}{t} \right\rceil + \left\lceil \frac{m-1}{t} \right\rceil + \cdots \left\lceil \frac{t+1}{t} \right\rceil + 1 \right) \quad (1)$$

If we just use this five feature templates as show , there are 78460 matrix features in 20x20 pixel image. However, we may use more expansion HLF templates[7] in practical applications. The detector with slightly better performance may contain thousands of HLF.

doi:10.4304/jcp.9.5.1077-1082

Figure 1. Five basic kinds of HLF template

HLF can describe the differences of local gray, most of the time it can achieve a fine detection result. But HLF describe the structure coarsely, it is sensitive to certain graph structure, such as edges, lines,etc. Figure 2 embodies the defect of HLF, Figure 2(a) shows two basic HLF and in Figure 2(b) the two features can match the facial image ideally, thus the face image can detect by the rectangle feature. However, the image in Figure 2(c) may greatly mistaken for a face because the features also can match the image.

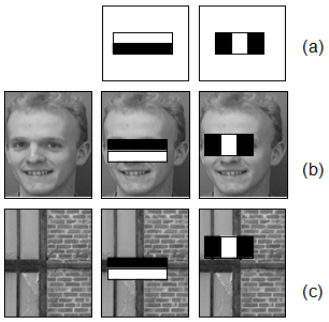


Figure 2. Example of matching HLF in two typical images

With the above problem, the CMF was proposed by Hua et.al.[5]. CMF itself is a covariance matrix[8], each value of matrix is a covariance between each sample subregion. The detection process is as follow, firstly, the covariance between each sample are calculated, then combine a portion of covariance value into covariance matrix according to some rule. Each covariance matrix is a CMF.

Hua divide the image into 5×5 same size matrix region as shown in Figure 3. Taking any four different sub-regions and extract the CMF, 12650 covariance matrix features are extracted if we don't consider the positional relationship. Every CMF is a feature template, The appropriate weak classifiers are selected by Adaboost algorithm, then a face classifier is combined by the weak classifiers. As displayed in Figure 4, we select four typical regions p(1,2), p(2,4), p(4,3), p(5,4) and calculate the covariance between each region pixel, then it constitute a 4×4 covariance matrix, it also call a CMF. The matrix value is shown in Table 1.

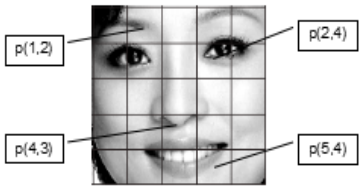


Figure 3. Dividing face into 5x5 regions

Figure 4. Face sub-regions

TABLE 1.
VALUE OF A CMF AS THE SUB-REGIONS OF FIGURE 4 SHOWN

866.8	-758.1	-231.4	23.15
-758.1	3905.8	14.32	603.5
-231.4	14.32	494.7	-176.6
23.15	603.5	-176.6	224.3

CMF need calculate the threshold of weak classifier, a CMF corresponds to a week classifier. The calculation of CMF threshold is complex, firstly, calculating the Euclidean distance between feature matrix s and template matrix t, the Euclidean distance is as follow:

$$d_j^a(s,t) = sqrt\left[\sum_{i=1}^{n}(s_j^a[i]-t[i])^2\right] \qquad (2)$$

In the formula(2), j is the current CMF and a is the current sample, $s_j^a[i]$ is the jth covariance value of sample a, t[i] is the corresponding template value. Then calculating the distance mean value of every CMF, the mean value is the final threshold for judgment. Variable k is the sum of samples. The threshold formula of feature j is as follow:

$$\mu(j) = \frac{1}{k}\sum_{a=1}^{k} d_j^a(s,t) \qquad (3)$$

After the threshold of a test sample is achieved, if the Euclidean distance between covariance matrix and template matrix is less than the threshold, the sample is considered to be a face,otherwise it considered a non-face. The weak classifier is as follow:

$$h_j(x) = \begin{cases} 1, f_j(x) < \mu(j) \\ 0, otherwise \end{cases} \qquad (4)$$

The process of training CMF weak classifiers is complex, furthermore, a 5×5 image contains 12650 CFMs, the Adaboost algorithm based on CMF leads to a time-consuming result.

There are 78460 HLFs in a 20×20 pixel image, it is time-consuming even we calculate the feature by integral image. There are 12650 CMFs in a 5×5 sub-region image and the CMF threshold calculation is complex, the entire training process is more time-consuming than HLF. Focus on these problems, this paper proposed the covariance feature based on CMF, it can reduce the number of feature and overcome the time-consuming problem, and remain the detection rate steady. In the experiment we use the 8×8 sub-region covariance feature and extract only 4064 features, the number of feature is far less than HLF and CMF. The feature threshold calculation is similar to HLF, it is more simple and faster than CMF threshold.

CMF and CF all have to calculate the covariance between each sub-region in common, they all can describe correlative degree of each sub-region. The difference of CMF and CF is as follows:

(1)The representation of CF value is a two-relative region covariance, but the representation of CMF value is Euclidean distance between feature matrix and template matrix.

(2)CF only need to calculate a covariance matrix to express all the CF value in each test image. However, a covariance matrix is a CMF. Four relative sub-regions covariance matrix combine into a CMF.

III. ADABOOST FACE DETECTION BASE ON CF

A CF is a covariance coefficient[9], and it is a low dimensional feature that can calculate easily. On a given W×W pixel image, dividing sample image into d×d same size matrix sub-region, then we can calculate the covariance coefficient C(x,y) of any two matrix. x,y represent the ith, jth matrix respectively. The number of pixels in each matrix is n, and $\mu(j)$ represent the mean value of all ith matrix pixels:

$$\mu(i)=\frac{1}{n}\sum_{k=1}^{n} f_k(i) \qquad (5)$$

The covariance coefficient is similar to covariance formular and defined to be:

$$C(i,j)=\frac{1}{n-1}\sum_{k=1}^{n}(f_k(i)-\mu(i))(f_k(j)-\mu(j)) \quad (6)$$

In the formula (6), the ith, jth sub-region is interpreted as a two dimension random variable and evaluate the covariance coefficient C(i,j). In virtue of the covariance symmetry and the covariance of sub-region with itself is insignificant, we can extract $[(d^4+d^2)/2]-d^2$ significative CFs on d×d matrices. For illustration purposes, as shown in figure 5, dividing a test image into 3×3 sub-region. We can obtain a covariance matrix and the covariance of any two different coordinate region is a CF. The 3×3 specification only generate 36 CFs as the table 2 shown. For example, p(1,3) and p(3,2) represent the 3rd and 8th sub-region, and Via formula (6) we can calculate the (3,8) coordinate value of the CF matrix is 14.76 .

TABLE 2.
3×3 MATRIX OF CF VALUE

1575.2	140.56	1268.8	-182.7	-85.57	-51.36	-52.16	-107.0	-94.62
140.56	364.77	241.23	20.62	101.78	2.66	343.67	79.82	-60.09
1268.8	241.23	1819.1	-45.76	182.63	-57.37	328.69	14.76	57.39
-182.7	20.62	-45.76	277.02	136.82	143.35	176.14	133.89	104.17
-85.57	101.78	182.63	136.82	283.69	65.49	255.91	96.61	8.20
-51.36	2.66	-57.37	143.35	65.49	164.15	154.25	65.78	-21.84
-52.16	343.67	328.69	176.14	255.91	154.25	1148.0	162.27	-73.69
-107.0	79.82	14.76	133.89	96.61	65.78	162.27	307.38	135.62
-94.62	-60.09	57.39	104.17	8.20	-21.84	-73.69	135.62	463.62

Figure 5. 3×3 cutting face image

The experiment result showed that the extracting speed of CF is faster than CMF with the same specification, but there is a large gap with the desired detection. Boosting the detection performance is necessary. The purpose of Adaboost algorithm is improve the classification accuracy of any given learning algorithm, the next process is discussing how to improve the performance of CF classifier through Adaboost algorithm.

The example of 3×3 specification is too rough and can't primely describe the internal relation of face sub-region. To refining the extracting feature, we choose the 8×8 specification to extract the CF in the experiment. Figure 6 shows the 8×8 diving face, choosing any two sub-region and calculating the CF value C(x,y) by formula (6), we can extract 4064 CFs $C_n(1,2,\cdots,4064)$ in total. Through the Adaboost algorithm, we can training a appropriate number of CFs from the 4064 CFs that constitute a strong classifier.

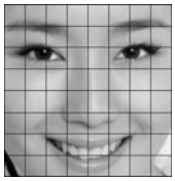


Figure 6. 8×8 cutting face image

Adaboost is a self-adapting Boosting algorithm, the training process is greatly associated with sample probability distribution. When the sample classified correctly, reduce the sample weight, and when the sample classified falsely, increase the sample weight while the classifier receive more attention. Then combining the optimum weak classifier of each round into a strong classifier and output the value of judgment. So long as the classify accuracy of CF classifiers is greater than 50% and training enough CF weak classifiers, we can obtain a strong classifier which the error rate approach 0. The form of the jth weak classifier is given as follows:

$$h_j(x)=\begin{cases}1, p_j f_j(x) < p_j\theta_j \\ 0, otherwise\end{cases} \qquad (7)$$

$h_j(x)$ is the judge value of sample *x*, the threshold offset is p_j, the value only can be ± 1, the threshold is θ_j, $f_j(x)$ is the jth CF value of sample *x*.

Following are the procedure of Adaboost classifier learning:

(1)Given training example $(x_1,y_1),(x_2,y_2)\cdots(x_n,y_n)$ where y_i =0,1 for negative and positive examples respectively.

(2)Initialize weights $\omega_{1,i}=1/2m,1/2t$ for y_i =0,1 respectively, where *m* and *l* are the number of negatives and positives.

(3)For $t=1,\cdots T$

(a)Normalize the weights, $\omega_{t,i}=\omega_{t,i}/\sum_{j=1}^{n}\omega_{t,j}$, so that ω_t is a probability distribution.

(b)For each feature *j*, training a classifier h_j which is restricted using single feature. The error is

evaluated with respect to ω_t, calculate the error rate: $\varepsilon_j = \sum_{j=1}^{n} \omega_{t,j} \left| h_j(xi) - y_i \right|$.

(c)Choose the classifier h_t, with the lowest error ε_t.

(d)Update the weights: $\omega_{t+1} = \omega_{t,i} \beta_t^{1-e_i}$, where e_i =0 if example x_i is classified correctly, e_i =1 otherwise, and $\beta_t = \varepsilon_t / (1-\varepsilon_t)$.

(4)The final strong classifier is:

$$h(x) = \begin{cases} 1, \sum_{t=1}^{T} \alpha_t h(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0, otherwise \end{cases} \quad (8)$$

Where $\alpha_t = \log \frac{1}{\beta_t}$.

We chose 356 classifiers of higher detection rate from 4064 weak classifiers after training, the detection sub-regions are distribute mainly beside eyes, eyebrows, nose, mouth. These region can distinguish the face and non-face effectively.

Figure 7. 6 CFs which are easiest to distinguish.

Below is the detection algorithm:

Dividing the test image $k(1,2,\cdots,n)$ into 8×8 sub-region and mark for 1, then detect the image in each layer classifier of the cascade classifier[10-11] orderly.

For $i = 1:T$ //the ith layer classifier

For *j*=1:*M*//the jth weak classifier of ith layer

Calculating the CF value of two sub-region appointing by order jth, from the jth classifier we can obtain the judgment value of the current sample, and the value is H_{ij}.

End

The weight of jth weak classifier in layer ith is α_{ij}.

$$h(x) = \begin{cases} 1, \sum_{j=1}^{M} \alpha_{ij} H_{ij}(x) \geq \frac{1}{2} \sum_{j=1}^{M} \alpha_{ij} \\ 0, otherwise \end{cases} \quad (9)$$

Mark 0 if $h(x)$ =0 and eliminate the sample which is unnecessary to detect on the next layer. Otherwise, mark 1 and detecting sequentially.

End

After T layers selection, the samples which are marking 1 are estimated faces and remain are non-faces.

Ⅳ.EXPERENENT RESULTS

The training samples chose from MIT face test set, we selected 2106 faces and 2781 non-faces for training CF classifiers. In the pre-process, we normalized the sample into 40×40 pixels, and enhanced the image by histogram equalization. We divided the normalized image into 8×8 specification and every sample image can extract 4064 CFs. The test set select from MIT+CMU, there are 1236 face samples and 1276 non-face test samples. Furthermore, face image from Internet were used, and preprocessed the image with skin detection method[12]. The operating environment is AMD 3.0GHZ PC, Windows XP operating system. Our experiment compared the 20×20 pixels HLF, 5×5 specification CMF and 8×8 specification CF.

Figure 8. Part of face and non-face samples from MIT set

The experiment results of feature extracting are as follow in table 2. A large number of HLF lead time-consuming extraction. Though the covariance calculation is more complicated, the 5×5 specification CMF only need calculate covariance value 300 times per sample image, then combine any four different value into a CMF. There are 12650 combinations and only spend 143s. The 8×8 specification CF calculate the covariance values 4064 times and the calculation complexity of covariance value is same as the CMF. CF spend 473s in feature extraction. The speed of CMF extraction is slightly higher than CF. However, the speed of CMF threshold calculate is too more complex than CF. It is the biggest defect of CMF and this problem lead to a time-consuming detection process.

TABLE 2.
EXTRACTING TIME-CONSUMING OF THREE FEATURE

Feature	Number of feature	Time-consuming(s)
HLF	78460	1584
CMF	12650	143
CF	4064	473

In the testing process, , the detection rate is not always increscent while the weak classifiers increase. Comparing the three curves in the figure 9, CMF reach the maximum firstly, CF second and HLF last. When the weak classifiers increase unceasingly, HLF detection rate will declines obviously because the sample weight has become distorted and HLF has degraded. Adaboost algorithm based on CMF and CF can avoid the degradation preferably, CMF and CF enhance the detection rate and lower the false positive rate.

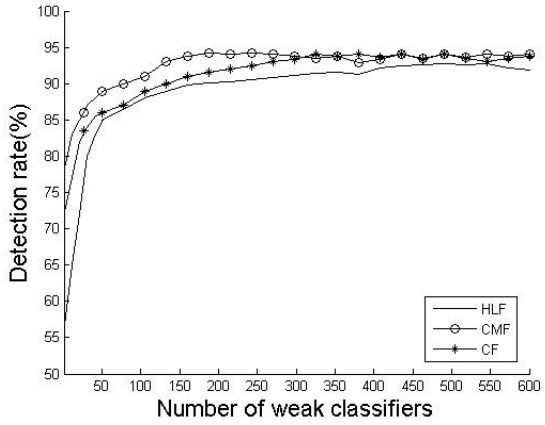


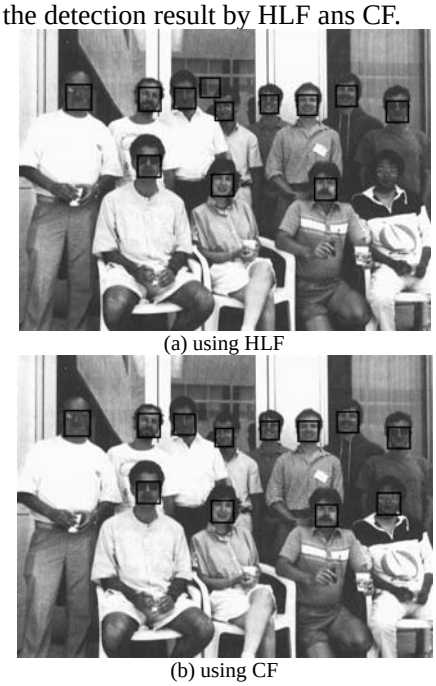
Figure 9. The relation of detection rate and number of weak classifiers

The CF can overcome the shortcomings of time-consuming as shown in table 3, the time spending of CF training and detection is 2562s and significantly less than CMF and HLF. The CF detection rate is 93.7% which is nearly CMF and higher than HLF. After tests, using 356 CF weak classifier can reach the best detection rate.

TABLE 3. DETECTION RATE AND TIME SPENDING OF WHOLE PROCESS

Feature	Detection rate(%)	Number of weak classifier	Average time spending(s)
Haar-like	92.5	528	13647
CMF	93.6	237	37459
CF	93.7	356	2562

The following picture is from CMU test set, and we can see the detection result by HLF ans CF.


(a) using HLF

(b) using CF

Figure 10.Face detection result

V .CONCLUSIONS

In this paper, the covariance feature is proposed. It combines the advantages of traditional Haar-like feature and covariance matrix feature. This paper also related the combination of CF and Adaboost algorithm. The test results on MIT+CMU test set showed CF can greatly improve training and test speed while the detection rate is similar to CMF. In recent years, many researches[13-14] focus on multi-view face feature and the research results showed these improvements can detect non-front face preferably. In the future, the multi-view CF is a novel direction worthy of studying.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China(No. 61263019), and supported by the Natural and Science Foundation of Gansu Province(No. 1208RJZA212), and supported by the scientific research project of Gansu Finance Department(No. 1114ZTC144).

REFERENCES

[1] L.H. Liang, H.Z. Ai, G.Y. Xu and B. Zhang, "A survey of human face detection," *Chinese Journal of Computers*, vol.25, No.5, pp.132-137, May 2002.

[2] M.H. Zhang and L.H. Gang, "A Detection Method of Driver's Face Orientation Based on Visual Cues and SVM," Journal of Software, Vol 8, pp924-931, Apr 2013.

[3] Y. Ouyang and N. Sang, "Robust Automatic Facial Expression Detection Method, " Journal of Software, Vol 8, pp1759-1764, Jul 2013.

[4] Viola P, Jones M, "Rapid object detection using a boosted cascade of simple feature," in Proceedings of the IEEE *Computer Society Conference on Computer Vision and Pattern Recognition*, pp.511-518, 2002.

[5] Y. Hua, H.J. Peng and J.L. Gu, "Face detection method by boosting covariance feature," *Computer Engineering and Applications*, Vol.46, No.18, pp.186-189, 2010.

[6] B. Wu, C. Huang, H.Z. Ai and S. Lao, "A multi-view face detection based on real adaboost algorithm." *Journal of computer research and development*, Vol.42, No.9, pp.1612-1621. 2005.

[7] Lienhart R, Maydt J, "An extended Set of Haar-like Features for Rapid Object Detection," [C]//Proc. Of ICIP'02. New York, USA:[s.n.], pp.155-162, 2002.

[8] Moakher M, "A differential geometric approach to the geo-metric mean of symmetric positive-definite matrices," *SIAM J Matrix Anal Appl*, Vol.26, No.3, pp.735-747, 2005.

[9] Tuzel O, Porikli F, Meer P, "Region covariance: A fast descriptor for detection and classification," //Proc 9th *European Conf on Computer Vision*, Graz, Austria, pp.589-600, 2006.

[10] Viola P, Jones M, "Robust real-time face detection," *Computer Vision*, Vol.57, No.2, pp.137-154, 2004.

[11] W.H. Li, Y.F. Lin, B. Fu, M.Y. Sun and W.T. Wu, "Cascade Classifier Using Combination of Histograms of Oriented Gradients for Rapid Pedestrian Detection," Journal of Software Vol 8, pp.71-77, Jan 2013.

[12] Erdem C, Ulukaya S, Karaali A and Erdem A, "Combining Haar Feature and skin color based classifiers for face detection," In *Acoustics.Speech and Signal Processing (ICASSP)*, 2011 IEEE International Conference, pp.1497-1500, may 2011.

[13] X.K.. Liu, X.H. Sun and Y.X. Zhou, "Multi-angle Face Detection Based on New Haar-like Feature," *Computer Engineering*, Vol.35, No.19, pp.195-197, 2009.

[14] Z.X. He, X.Q. Ding, C. Fang and D. Wen, "Multiview face detection based on LBP and CCS-Adaboost," *Journal of Zhejiang University(Engineering Science)*, Vol.47, No.4, pp.622-629, 2013.

Rui Li was born in Gansu Province, China in 1971. She received the Master of Science degree in computer application in 2005. She is currently a professor at school of computer and communication of Lanzhou University of Technology. Her research interests include pattern recognition, picture processing and digital watermarking.

Changfeng Li was born in Hubei Province, China in 1991. He received the B.E degree in 2011.He is currently a postgraduate student at School of Computer and Communication of Lanzhou University of Technology. His research interests include pattern recognition and picture processing.

Researching and optimizing key technology of HDR analog signal chain in the satellite communications

Shuai Zhang
School of Electronic Information, Wuhan University, Wuhan, China, 430072
E-mail: shuai173@whu.edu.cn

Shilei Sun* Yongqiang Cui and Dexiang Deng
Institute of Microelectronics and Information Technology, Wuhan University, Wuhan, China, 430072
School of Electronic Information, Wuhan University, Wuhan, China, 430072
School of Electronic Information, Wuhan University, Wuhan, China, 430072
E-mail: sunsl@whu.edu.cn, cuiyq@whu.edu.cn, whuddx@gmail.com

Abstract— **Due to the low dynamic range of the analog front end, low respond speed, bad flatness of the Amplitude-frequency characteristics and the poor noise suppression capability of the signal conditioning in the analog signal chain of the traditional digital down converter (DDC) system, an improved analog signal chain scheme used in high data rate receiver (HDR) of the satellite communication system has been put forward. This scheme has 55dB dynamic range, high respond speed and good noise suppression capability. Moreover, it has been realized on the platform of the ground terminal HDR for demodulator system. 720MHz and 1.2GHz intermediate frequency conversion is an important component of the satellite ground monitoring X-band data channel receiver system. The scheme consists of a band-pass filter module, a digital control variable gain amplifier (VGA), a low noise amplifier (LNA), a radio or intermediate frequency gain block and an ADC which has 1.7GHz input bandwidth. It realizes sampling and conditioning of the 720MHz and 1.2GHz intermediate frequency signal and automatic gain control (AGC). Experimental results show that the scheme has the characteristics: automatic and fast gain control, wide dynamic range and good amplitude-frequency characteristics of the gain points in the pass-band, low power consumption and good noise suppression capability.**

Index Terms— Down conversion system, digital down converter, Analog signal conditioning, Automatic gain control, High speed ADC sampling

I. INTRODUCTION

With the rapid development of the world communication standards and protocols, flexibility and adaptability has become one of the most important characteristics of modern communication. Designing more efficient architectures which support multiple communication protocols and enhancing the data transmission speed and the data capacity amount have become an inevitable trend. To reduce the restrictions of the analog signal chain in the communication system, researchers have to introduce greater amounts of digital signal processing and optimizing algorithms. Future development of the communication system will face great challenges [1]. In addition, with the rapid development of the semiconductor technology, analog integrated circuits, signal processing technology and the production process of the analog chips and printed circuit boards, key parameters of low noise amplifier (LNA) and analog digital converter (ADC) in RF and IF filters have been improved significantly. It provides reliable technical support for us to improve the traditional analog front-end of receiver circuit.

The design of UWB receivers exposes to unique challenges [2]. According to the related material of the aerospace remote sensing, when the ground resolution of the image approaches 1 meter, the real-time data capacity reaches 1.7Gbps. Considering compression ratio is 1:5, real-time pure data transmission from satellite to ground is 340Mbps. In USA, the 300Mbps satellite transmission system has been instantiated. 650Mbps and 1Gbps satellite high data rate transmission system is being researched or experimented. In 2002, NASA proposed to use OFDM system to achieve a 622Mbps modem system [3]. Japan is also researching on the 1.2Gbps data transmission system.

In terms of the current situation of space data transmission, the transmission rate from satellite to ground or between satellites is far from enough. HDR technology has been used in satellite communication navigation and monitoring communication network. In recent years, satellite ground demodulation receiving system of 720Mbps and 1.2Gbps IF have been researched or experimented. Based on the key technology of front-end analog chain of the traditional HDR system, improved scheme is proposed and the result is shown in this paper.

Figure 1 shows the structure diagram of the traditional ground demodulation receiving system. The scheme consists of three parts: RF down converter, IF analog signal conditioning and digital down converter in Field Programmable Gate Array (FPGA). Part 1: RF down

doi:10.4304/jcp.9.5.1083-1089

converter. The antenna receives the C / X / Ku / Ka band satellite signal and send it to the down converter. Then the down converter gets the expected IF signal. This part is separated from the HDR system and composes RF down converter. In the traditional satellite communication, the data rate of down converter is generally no more than 375Mbps. Rear-end analog signal chain and digital data processing are based on this data rate. However, in recent years, in the research of higher rate satellite communication, research and optimization on traditional key technology for higher speed transmission needs to be done.

Part 2 and Part 3 compose HDR demodulator system (this paper focuses on research and optimization analog signal chain in these two parts). Part2: IF analog signal conditioning. IF signal connects to the conditioning circuit of the analog front-end signal through dedicated coaxial interfaces, such as SMA or BNC. IF signal is output by different satellite down converters. To sample the IF signal, the LNA module conditions it by fixed gain and sends it to ADC in single-ended signal or differential signal. In the traditional receiver, the ADC performance is more dependent on the power and SNR of IF signal in the down converter. It may leads to many problems, such as too low signal amplitude, low resolution, signal saturation and distortion. In addition, the analog signal chain that is composed of a fixed gain amplifier and an ADC has narrow signal bandwidth and poor anti-jamming capability. Part 3: Digital down converter in FPGA. After sampling the signal entering the FPGA platform-based DDC processing module. The traditional DDC system consists of a local oscillator (NCO), a mixer, filter, a half-band filter (FIR&HBF) and a extractor (CIC). Its main function is to change the IF signal to zero-IF signal and to reduce the sample rate [4]. The spectrum shows that the digital down converter transforms the signal sampled by ADC from RF signal to baseband. It is completed by two steps: Firstly, the quadrature carrier multiplies the input signals. Then the digital filter eliminates the unwanted frequency components. If the sampling rate is higher than 600MHz, it is difficult to do real-time processing in FPGA [5].

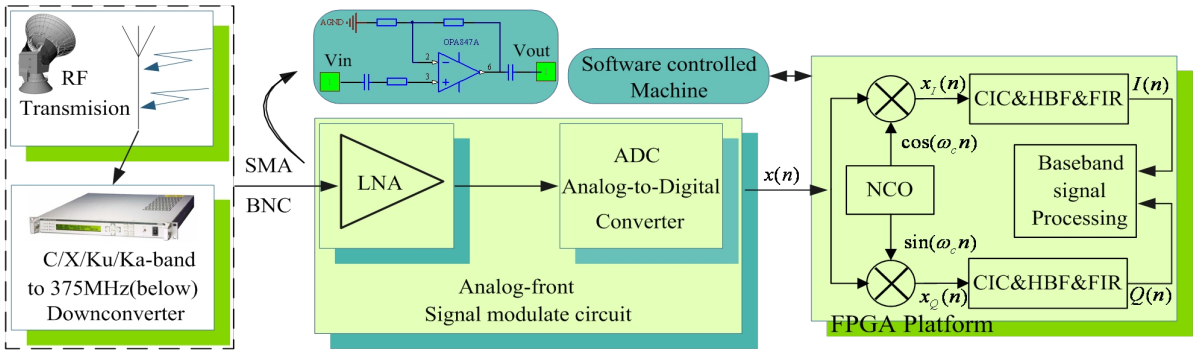


Figure 1. Traditional ground demodulation receiving system

The traditional DDC system has the shortcomings of large amount of computation, low utilization of logical resource and is difficult to process the RF signal simultaneously. Though many researchers have proposed improved schemes of the multi-channel digital down conversion to solve these problems, huge challenge still exists due to the rapid development of demodulation rate [6]. This paper presents an optimization of the analog front-end signal chain to reduce the difficulty of the back-end FPGA processing. It significantly reduces resource consumption of the back-end digital AGC and FIR in FPGA. It also reduces the error rate brought by multipath transmission and deep fading and demodulation loss of the receiver.

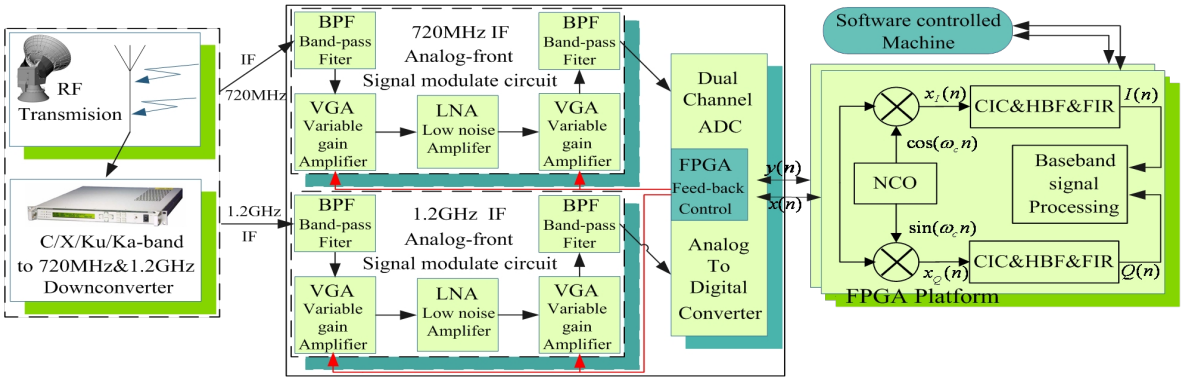


Figure 2. Ground demodulation receiver system

II. System Structure of the HDR Receiver

Traditionally, the signal is digitized at a low frequency. However, with the rapid development of the fields such

as radar, communication and electronic instruments, the frequency has gradually changed to IF and even to RF [7]. For example, the improvement of data transmission rate, carrier frequency and IF signal frequency has higher demand for the HDR system.

For the 720Mbps and 1.2Gbps IF satellite ground demodulation receiving system that was proposed recently, this paper puts forward an improved scheme of the key technology of the analog signal chain. Figure 2 shows the design diagram of the demodulation system. Compared with the third parts of the traditional receiver, the front-end IF signal in part1 increases from 375MHz to 720MHz and 1.2GHz. Two channels $x(n)$ and $y(n)$ are processed simultaneously on FPGA-based platform in part3. The analog signal chain is an improved scheme proposed in this paper. Through BPF, variable gain amplifier (VGA), LNA and high-speed ADC, the scheme realizes automatic gain control of IF signal, band-pass filtering and high-speed ADC sampling [8].

III. Comparison of Three HDR Front-End Analog Signal Chain Schemes

The output IF frequency of the RF down converter (in door) is 720±200MHz or 1.2GHz±300MHz. The impedance of it is 50Ω. The first scheme uses RF/IF Gain Block. For example, the ADL5542 integrates channel impedance matching circuit and 20dB fixed gain amplifier circuit. This scheme has simple structure and good amplitude-frequency characteristic. Its fluctuation from 500MHz to 1.5GHz is less than 1dB. However, the fatal defect is that it cannot adjust the amplitude of the IF signal, which may leads to the problems of too low amplitude of the IF signal, bad resolution signal saturation and distortion. It makes an unrecoverable effect on the subsequent ADC sampling and FPGA processing [9]. In the second scheme, discrete AGC module is used to automatically adjust the gain according to the amplitude of the input signal, so that the output voltage of this module maintains in the optimum range of the ADC's input power [10]. Figure 3 shows the structure diagram of the second scheme. The core module is the gain controlled amplifier adjusted by the voltage signal. The voltage signal is automatically generated by a loop that is composed of an electric level detector (peak value detection circuit), a low pass filter, a DC amplifier, a voltage comparator and a control voltage generator.

The input signal of the discrete AGC $u_i = \cos \omega t$. The output signal $u_o = U_{im} \cos \omega t$. The gain $A_u = U_{om} / U_{im}$. The gain A_u is controlled by the control voltage u_c . The control voltage is obtained after the conversion of the error voltage generated by the voltage comparator via a control voltage generator. Independence is the advantage of discrete AGC. The internal closed-loop control amplitude of the output signal adjusts the IF signal to the most suitable power for ADC sampling. However, the structure is complex and has too many discrete components. In addition, when the processing analog signal is around 1 GHz, the amplitude-frequency characteristic within the pass band has high jitters and the different gain characteristics have poor consistency.

Considering the advantages and disadvantages of the first two schemes, this paper proposes a new analog channel conditioning scheme. Figure 2 shows the design of analog signal chain. This scheme has the advantages of low structure complexity and improves the jitter performance of the amplitude-frequency characteristic and the consistency under different gain. In addition, it realizes automatic gain control, solving the problems in the first scheme that the amplitude of the IF signal cannot be controlled flexibly. Moreover, this scheme has the advantages of wide dynamic range of gain, high respond speed of conditioning, low power consumption and good capability of noise suppression [11].

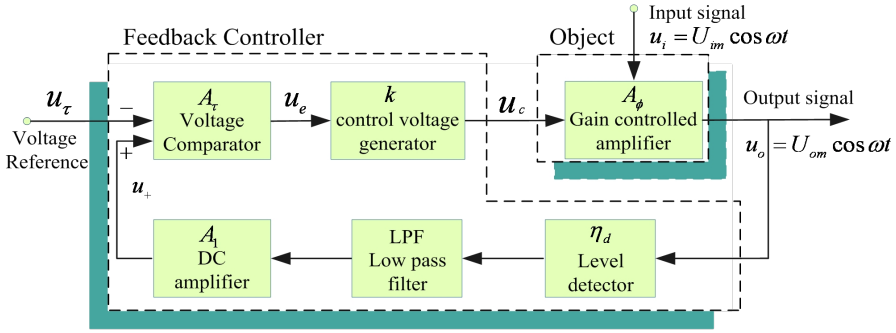


Figure 3. Structure diagram of discrete AGC

To support synchronous demodulation of dual-channel satellite data, the scheme uses two different center frequencies and bandwidths of the BPF and handles IF signal of two channels. In terms of one channel, when the IF signal enters the HDR demodulation system, firstly, the IF signal filters the low frequency and high frequency interference through a BPF. Secondly the output signal enters the next level of digital control VGA. Its gain is controlled by FPGA signal processing end, realizing 30dB gain conditioning [12]. Thirdly, it realizes fixed gain of 20dB through the LNA. Then the HF noise and external interference that is brought in the

front-end analog chain are filtered through VGA and BPF. Finally, it is changed to differential signal through the 30~1800MHZ RF transformer and enters dual-channel 8-bit ADC. The ADC supports dual-channel simultaneous sampling. The maximum speed of the inside sample-and-hold element reaches 1.7Gbps in the case of dual-channel sampling and reaches up to 3.4Gbps in the case of single-channel sampling. The digital signal sampled decreases its speed by 1:2 inside the ADC. Then it is sent to FPGA as level mode of 32 pairs differential LVDS signal. Considering that the maximum transfer rate of the 32 pairs of LVDS signal is still around 1.5Gbps, to maintain synchronization and low crosstalk between signals, a differential serpentine long wire is often used for connecting FPGA when designing the PCB. To better match with the differential bank receiver end of FPGA, 100Ω characteristic impedance is used for designing transmission wire Figure 4 shows the circuit of the scheme.

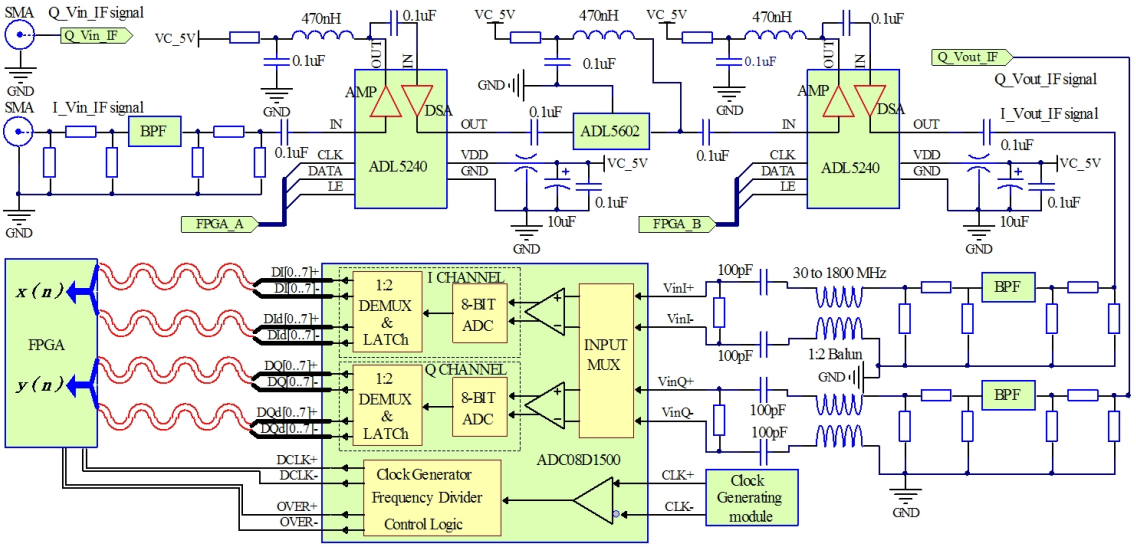


Figure 4. New scheme diagram of analog channel conditioning

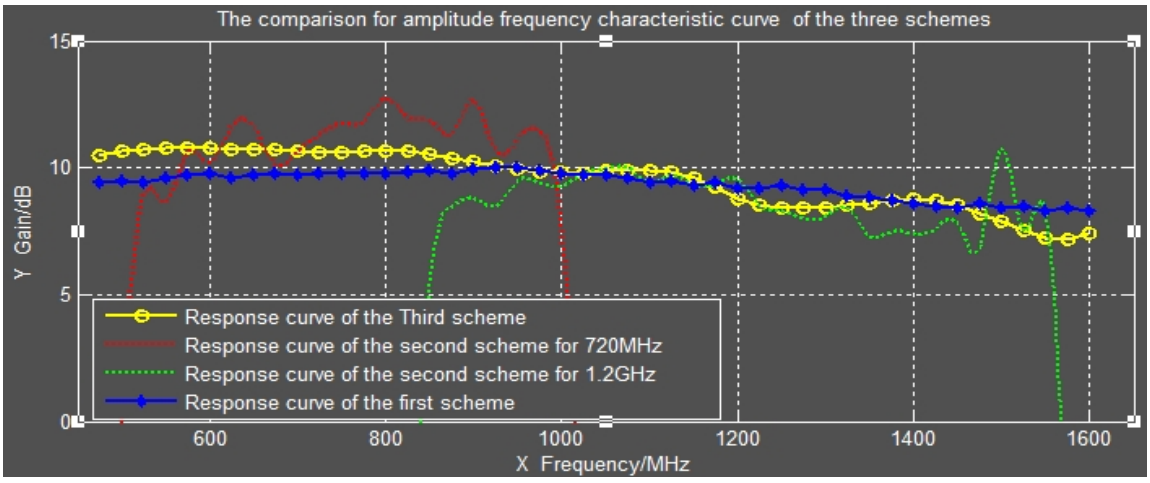


Figure 5. The amplitude-frequency characteristics of these three schemes

Through PCB test of these three analog signal chains, the comparison chart of the key indicators of the analog signal chain is achieved by simulating actual data. Figure 5 shows the response curves of the amplitude-frequency characteristics of these three schemes. The input signal is $^{-10}$ dBm and the frequency ranges from 500MHz ~ 1.6GHz. Simulate and compare the gain of these three schemes in MATLAB. It can be seen clearly that the blue curve that stands for the first scheme has the best flatness within the pass band, only 1.7dB. The third scheme also has good flatness within the pass band. In particular, within the band range of 720 ± 200MHz or 1.2GHz ± 300MHz, its fluctuations are less than 0.73dB or 2.07dB. Fluctuation within the pass band of the third scheme is the largest and irregular. Within the band range of 720 ± 200MHz or 1.2GHz ± 300MHz, its fluctuation is 4.1dB or 3.9dB. From the spectral response, the third scheme overcomes the defects in the first scheme that the amplitude cannot be adjusted flexibly. Furthermore, it improves the intolerable band jitter in the second scheme. So it is the best analog signal chain for HDR demodulation system.

For HDR demodulation system, consistency of the respond curve under different gain is also very important.

Figure 6 shows the respond curve of the amplitude under different gain. Set different power of the input signal. The gain of the analog front-end ranges from -5dB to 50dB. The frequency ranges from 500MHZto 1.6GHZ. Simulating data is achieved by MATLAB. Good consistency of the respond curve under different gain is shown clearly. Conclusion can be made that the max actual dynamic range measured reaches 55dB. Consistency under different gain is good.

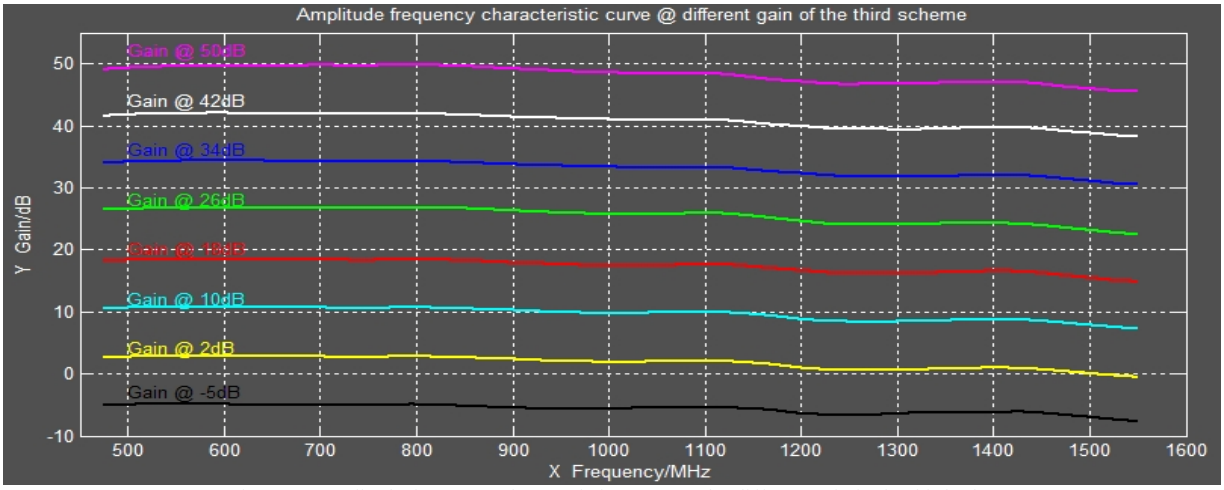


Figure 6. Amplitude-frequency characteristic curves under different gain in the third scheme

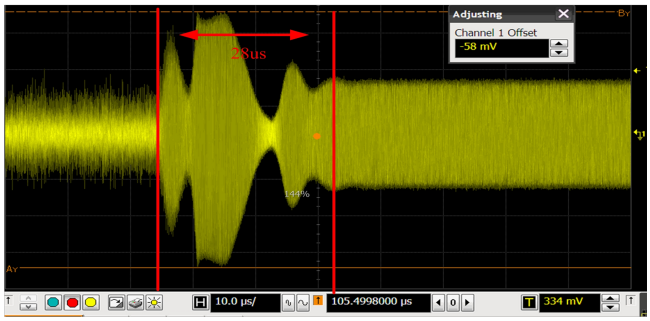


Figure 7. Respond time of AGC in the second scheme

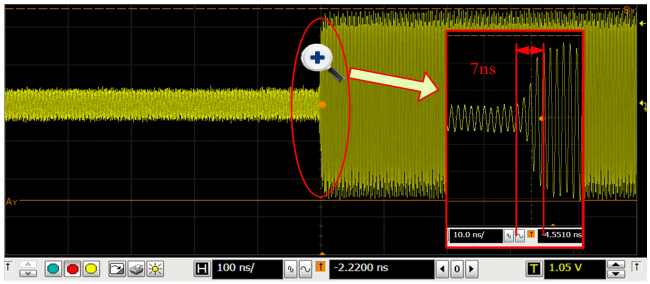


Figure 8. Respond time of AGC in the third scheme

In wireless communication, due to the impact of various factors such as climate, environment and distance, the amplitude of the received signal fluctuates randomly [13]. For a better conditioning of the signal of analog front-end, the chain is required to have a quick response characteristic of the AGC. Figure 7 and Figure 8 show the comparison figures of the respond time in the second and third scheme. There is a big difference between them. The respond time of the second scheme is 28us, which has an obvious process of adjusting signal output. The process is an inevitable result of co-adaptation of the inside peak detection circuit of Figure 3, voltage comparator and control voltage generator. While the third scheme controls the gain of digital signal directly and has high sensitivity. In addition, the respond time is only 7ns. So it has more advantages than the second scheme.

The third scheme has 2.9dB noise figure at 2.14 GHz, single supply operation from 4.75 V to 5.25 V, low quiescent current of 275mA and power consumption of 1.4W. The second scheme has actual test data of supply from 12V to 15V, working current 810mA and power consumption 19.7W.

However, the AGC in the third scheme is not a pure combination of analog circuits. It has digital peak detection of ADC. Therefore, the third scheme can be used in the HDR system or similar systems, other systems without ADC are not appropriate.

The test results show that the new scheme can automatically adjust the signal gain. It has 55dB wide dynamic range, 7ns respond time, and only 1.7dB amplitude-frequency characteristics of the gain points in the pass-band.

IV. Research of several key points in the PCB designing for high-speed analog signal chain

It is known that circuit design and detection of high-speed analog signal chain has great difference with that of the low-speed analog signal chain, especially when supporting a maximum 1.5GHZ analog chain band. In this paper, research and exploration are done from three directions.

(1). How will the material of PCB impact on the high frequency characteristic
(2). Impedance matching, 90° and R wiring of PCB
(3). Split of the power supply and ground

Dielectric constant of FR4(R-1766) @ 1MHz is 4.7.

In the type of low dielectric constant, dielectric constant of R-5715J and R-5755 @1MHz is 3.8 and 3.5 respectively. The smaller the dielectric constant tangent ($\tan\delta$) is, the more suitable for wiring in HF and high-speed circuit. The bigger dielectric constant of the material is, in the same frequency the more blunt the eye diagram is. This difference is obvious when the frequency is 2.5GHz and 10GHz [14].

The impedance design of the analog signal chain is key to signal integrity. Any impedance mutation can cause the signal reflection and distortion [15]. Characteristic impedance Z_0 of the micro strip line is

$$Z_0 = \frac{87}{\sqrt{\varepsilon_r + 1.41}} \ln\left(\frac{5.98h}{0.8w + t}\right) \quad (1)$$

In formula (1), ε_r is the material dielectric constant. h is the dielectric thickness between the wire and the base level. w is the width of the wire. t is the thickness of the wire. Each RF trace on our test board has a characteristic impedance of 50 Ω and is fabricated on FR4(R-5755) material. In addition, each trace is a coplanar waveguide (CPWG) with a width of 25 mils, a spacing of 18 mils, and a dielectric thickness of 10 mils. Then Z_0 =50.16Ω, which matches well with single-ended analog channel transmission.

PCB wiring of the key signal chain is very important to impedance matching. Figure 12 shows the TDR wiring for two kinds of wiring [16]. Conclusion can be made that impedance of R wiring almost has no change. So all of the analog signal-chains adopt R wiring on the test PCB, reducing the uncertain signal reflection brought by impedance change.

When designing the test PCB, the impact of power supply and segment of signal grounds on analog signal chain has been considered carefully [17]. Power supply of the analog chain and ADC digital supply circuit should be independent of each other in the spatial layout of the circuit board and the current path. The module of power supply single-point common ground a used. To get better test results, cellular shield shell and full shielding are used. In addition, the architecture of daughter board and the mother board is designed to reduce the interference between the external environment and the power supply. Figure 13 shows the two test PCBs that are used to test the analog signal chain. The lower half part is the physical picture described in the second scheme and the upper half part is the physical picture described in the third scheme. After the test, we draw the conclusion that R wiring reduces the in-band fluctuations by 2.3 dB compared with the 90 ° wiring.

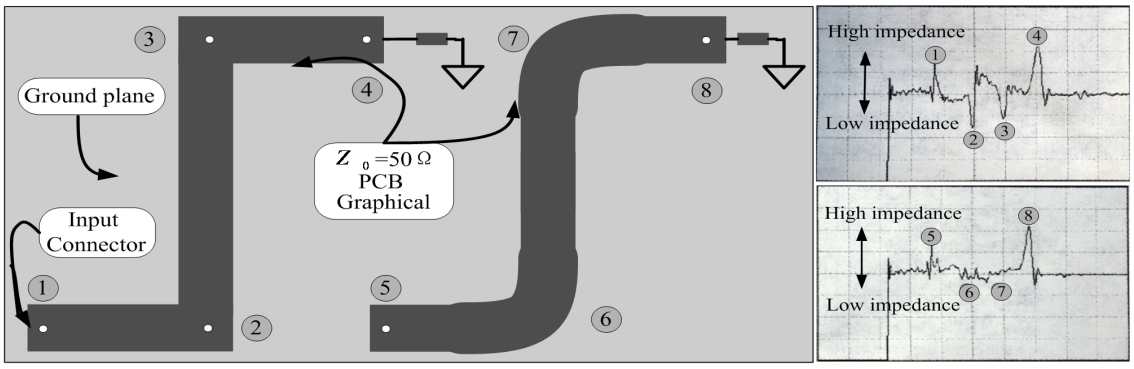


Figure 12. TDR simulation of different using 90 ° and R wiring

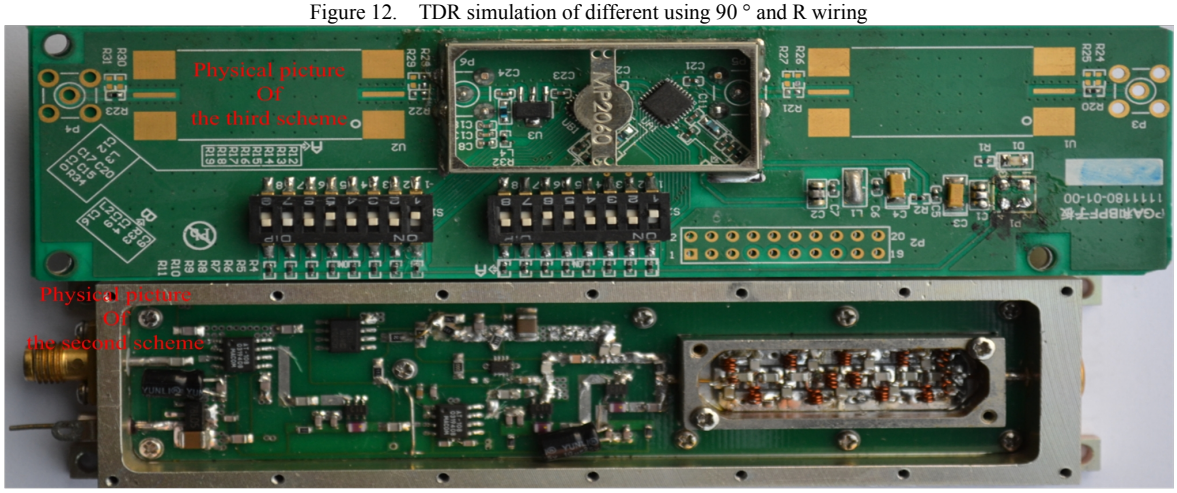


Figure 13. Physical pictures of the two schemes

V. CONCLUSION

In this paper, an optimization scheme of analog signal chain that has wide dynamic range, high respond speed

and good noise suppression capability is realized. Sampling, conditioning and automatic gain control (AGC) of the 720MHZ and 1.2GHZ IF analog signal is achieved through a band-pass filter module, a digital control variable gain amplifier (VGA), a low noise amplifier (LNA), a RF and IF gain block and 1.7GHZ wide input band. The amplifier can be controlled fast and automatically. It has wide dynamic range, good amplitude characteristic at the gain point within the pass band, low power and good noise suppression capability. Moreover, the scheme can be extended to the satellite broadband signal receivers and the analog signal chain of related application of signal detection, modulation, demodulation and identification.

REFERENCES

[1] Paolo Zicari, Emanuele Sciagura, Stefania Perri, Pasquale Corsonello, "A programmable carrier phase independent symbol timing recovery circuit for QPSK/OQPSK signals", Microprocessors and Microsystems, vol. 32, pp.437-446, 2008.

[2] Y. Vanderperren, Geert Leus, W. Dehaene, "An approach for specifying the ADC and AGC requirements for UWB digital receivers", IET Seminar on Ultra Wideband Systems, Technologies and Applications, vol. 10, no. 4, pp.169-200, 2006.

[3] MuliKifeThoms P. Bizon Dale J. Mortensen IMPLEMENTATION OF A 622 MBPS OFDM DIGITAL MODEM 20th AIAA International communication satellite Systems Conference and Exhibit 12-15 may 2002 Montreal, Quebec, Canada

[4] Zhang Qingxiang, Su Xiaoxiao, "The Design of Digital Down Converter Based On FPGA", 2012 International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2012, Conference article (CA).

[5] Deng Xiaoping, Tian Mao, Luo Yijun, Wang Yuhao，"Optimization and implementation of universal multi-channel digital down-converter", Chinese Journal of Scientific Instrume t, vol. 32, no. 9, pp.1993-1997, 2011.

[6] Xiaoting Chen, Guodong Zhao, Ping Jin, Wenbin Gong, Huijie Liu, Xuwen Liang, "QPSK/OQPSK signal detection with IFAGC on LEO satellite", IEEE 2007 International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, MAPE, pp.1179-1182, 2006, IEEE 2007 International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications.

[7] Chang Liu, Yue-Peng Yan, Wang-Ling Goh, Yong-Zhang Xiong, Li-Jun Zhang, Mohammad Madihian, "A 5-Gb/s automatic gain control amplifier with temperature compensation", IEEE Journal of Solid-State Circuits, vol. 47, no. 6, pp.1323-1333, 2012.

[8] Park Jeong I.L, Park Hwa Se, "A study on demodulation system design of the VOR receiver", Communications in Computer and Information Science, vol. 184 CCIS, no. Part I, pp.343-349, 2011.

[9] Lu Hao, Wang Zhen-Zhan, "Design of 1.2GHz bandwidth digital-correlator in fully polarimetric microwave radiometer", 2011 International Conference on Electric Information and Control Engineering, ICEICE 2011 - Proceedings, pp.4084-4087, 2011, 2011 International Conference on Electric Information and Control Engineering.

[10] Deng Qing, Wu Jian-Hui, Shi Long-Xing, Sun Wen, Yun Ting-Hua "007 9th International Symposium on Signal Processing and its Applications", ISSPA 2007, Proceedings, 2007.

[11] Xuejun Li, Ke Wang, Lingli Jiang, "The Application of AE Signal in Early Cracked Rotor Fault Diagnosis with PWVD and SVM", Journal of Software, vol. 6, No.10, pp.1969-1976, 2011.

[12] Zhang Dawei, Huang Fengyi, Tang Xusheng, Sun Xiaopeng,"Design of VGA for 6 GHz radio frequency communication system", IEEE MTT-S International Microwave Workshop Series on Millimeter Wave Wireless Technology and Applications, IMWS 2012 - Proceeding, p 32-35, 2012.

[13] Wenxian Xiao, Zhen Liu, Wenlong Wan, Xiali Zhao, "Research on the Expected Synchronization of Autonomous System", Journal of Software, vol. 8, No.1, pp.85-92, 2013.

[14] Tadashi Kubodera, Fen Jie, Gai Qiang, Pan Dong, "Kousoku Digital Kairo Jissou Knowhow", CQ Publishing Co., Ltd., Tokyo, pp.4-9, 2002.

[15] Zhou Xiang, Liu Xue-Guan, Guo Hui-Ping, Shao Lv-Xia, "Design of broadband impedance transformer using coupled microstrip transmission lines", Proceedings - 2009 3rd IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, MAPE 2009, p 994-997, 2009.

[16] Umenyiora Chidubem Andrew, Druce R.L, Curry Randy D., Norgard Peter, McKee T., Bowders J.J., Bryan D.A, "Dielectric constant of sand using TDR and FDR measurements and prediction models", IEEE Transactions on Plasma Science, vol. 40, no.10 PART 1, pp.2408-2415, 2012.

[17] Hui Feng, Zeqi Yu, "The Correction Method for Power Noise in Digital Class D Power Amplifiers", Journal of Software, vol. 8, No 2, pp.488-494, 2013.

Shuai Zhang, Inner Mongolia, 1987, School of Electronic Information, Wuhan University, Wuhan, China, Ph.D Candidate.
Interest research: Signal Process, High Speed Circuit

Shilei Sun, Hubei, China, 1981, Institute of Microelectronics and Information Technology, Wuhan University, Wuhan, China, Associate Professor.
Interest research: Signal Process, Artificial Intelligence

Yongqiang Cui, Shanxi, China, 1988, School of Electronic Information, Wuhan University, Wuhan, China, Ph.D Candidate.
Interest research: Signal Process, Satellite Communications

Dexiang Deng, Electronic Information School, Wuhan University, Wuhan, China, Professor, Ph.D. supervisor.

A New Method for Text Location in News Video Based on Ant Colony Algorithm

Ming Jiang[1,2] , Taotao Zha[1,2] , Xingqi Wang[1,2] , Jingfan Tang[1,2] , Chunming Wu[3]

(1. Institute of Software and Intelligent Technology, Hangzhou Dianzi University, Hangzhou, 310018, China)
(2. Zhejiang Provincial Engineering Center on Media Data Cloud Processing and Analysis, Hangzhou Dianzi University, Hangzhou, 310018, China)
(3. College of Computer Science and Technology, Zhejiang University, Hangzhou, 310027, China)

Email: jmzju@163.com, zhataotao1989@126.com, xqiwang@163.com, tangjf@hdu.edu.cn, wuchunming@zju.edu.cn

***Abstract*—Text in video is a very compact and accurate clue for video indexing and summarization. The paper presents a new method for text location in news video with ant colony algorithm. Three features of characters are extracted as a basis for the formation of heuristic function. In order to balance the weight of the three features, three functions are introduced to transform them. The ants will be randomly put on sub-blocks of video frames for searching text areas. Therefore, ants would leave pheromone in each sub-block. After the ant colony algorithm is finished, it produces a pheromone matrix. By binarizing the pheromone matrix, the text blocks can be located. The result has proved that the binarization method proposed in this paper is more accurate than otsu's method. At last, to reduce the false detection rate, the different directions of edge intensity ratio of text areas are computed, as the real text areas' edge intensity ratio is much smaller than the false one.**

***Index Terms*—video indexing, text location, ant colony algorithm, binarization, edge intensity ratio.**

I. INTRODUCTION

With the widely application of the digital information technology and multimedia technology, all walks of life have a lot of digitized information. In the face of massive video data, how quickly and easily to obtain the required information has become a research focus in the field of image analysis, data mining etc. To extract the text from image, first the text area should be located in image with complex background. Existing text detection techniques is categorized into four types: the method based on edge, the method based on region, the method based on texture, the method based on machine learning. *The method based on edge*[1-2] takes the advantage of the feature that text areas always have high sharpness, whose edge components is more than the non-text area. This method has a high speed as well as high false detection ratio. *The method based on texture*[3-4] takes the characters as a special kind of texture. The characters are usually made up by many fine stroke, hence the area with more strokes are the area rich in texture, which can be used to decide the text blocks. *The method based on region*[5-6], which is about connecting the similar or the same color's pixels and filtering these areas with knowledgeable rules, has high speed in processing but not fitting the case that text words have different colors. There are many problems to locate the text in video, for example the different size of characters, the diversity of character styles, the complex of the background. *The method based on machine learning*[7] is a new algorithm that can deal with these unstable factors. But the only disadvantage is that the choosing of the training samples and testing samples have a great affect on the result, while we haven't had a good system produce the reasonable samples until now.

According to the analysis above, most of the recent methods have their own advantages as well as disadvantages. To solve the problem of the text location in video with complex background in a better way, this paper proposes a method based on ant colony algorithm. The ant colony algorithm has its advantages of intelligent searching, global optimization, robustness, positive feedback, distributed computing and so on. It has reduced the influence of human factors because this method neither needs to set experiential threshold nor needs to choose test samples. Three typical features about characters are selected as the reason to update the pheromone in ant colony algorithm. The pheromone matrix has the same size of the test image. By binarizing the pheromone matrix with the new binarization algorithm, the text blocks and non-text blocks in an image can be divided. The effectiveness of this method has been proved by the experimental results.

II. OUR METHOD

The primary process of the method of text location using ant colony algorithm is as follows:

doi:10.4304/jcp.9.5.1090-1095

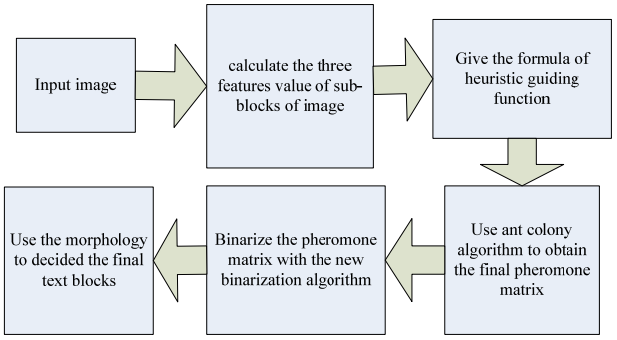


Fig 1. The process of ant colony algorithm

A. The Three Features of Characters

1. Gray-level Co-occurrence Matrix

GLCM, a joint distribution of two pixels' grayscale in different directions within an image, is a common method to analysis texture, which can well reflect the correlation law of grayscale about texture. According the 'Textural Features for Image Classification' written by Haralick[12], there are 11 characteristics of GLCM to describe texture. Here one of these, the variance whose value indicates the changing speed of texture and the length of period of texture, is adopted in our formula. The experimental result has proved that it is more effective and efficient than other features.

The formula to calculate the variance is

$$\sum_{k=2}^{2N}\sum_{i=1}^{k-1}(2i-k)^2 P_\delta \,. \tag{1}$$

where N is the grayscale of image, P_δ is the GLCM of image.

2. Wavelet Transform

A single can be decomposed into sub-bands at various scales and frequencies through wavelet transform. In the case of images, the wavelet transform is useful to detect edges with different orientations. The wavelet transformation can be implemented using filter banks consisting of high-pass and low-pass filters. The application to an image consists of a filtering process in horizontal direction and a subsequent filtering process in vertical direction. For example, when applying a 2-channel filter bank (L: low pass filter, H: high-pass filter), four sub-bands are obtained after filtering: LL, HL, LH and HH. The three high-frequency sub-bands (HL, LH, HH) enhance at most edges in horizontal, respectively vertical or diagonal direction. Since text areas are commonly characterized by having high contrast edges, high valued coefficients can be found in the high-frequency sub-bands.

Gllavata[4] proposed 2 characteristics from wavelet transform in both sub-band HL, LH: variance and histogram variance of wavelet coefficient.

$$VarCoef = \frac{1}{M\times N}\sum_{i=1}^{M}\sum_{j=1}^{N}(Coef(i,j)-MeanCoef)^2 \tag{2}$$

$$HistVarCoef = \sum_{\min_{coef}}^{\max_{coef}}(Bin\_new(j)-Mean_{Bin\_new})^2$$
$$\begin{cases} if\ Bin(j)>0, Bin\_new(j)=j; \\ if\ Bin(j)=0, Bin\_new(j)=0; \end{cases} \tag{3}$$

where *Coef(i,j)* is the wavelet coefficient, *MeanCoef* is the average of all coefficients. *Bin(j)* is the *jth*'s number of wavelet coefficient; *M, N* is the number of the ranks of sub-band HL, LH. These two sub-band focus on the horizontal and vertical texture of the characters. As the great correlation between the variance and histogram variance, we here just pick up the variance of coefficient with higher value between sub-band HL and LH.

3. The Corners

The corners, existing in the edge or outline of object, can reflect the important information of the local image with little redundancy. The text words in image have more and intenser corners, especially Chinese characters. Text in video or image always has sharp contrast with the background in color and brightness. Therefore, the text area will contain more corners. Harris[13], using the method of differential operation and sub-matrix correlation, is a common corner detection algorithm. This algorithm has the advantage of simple calculation, extracting the feature points with predefined number and stable operator. This paper used Harris algorithm for corner detection.

B. Text Detection in Video with Ant Colony Algorithm

The ant colony algorithm has been used into edge detection of image that takes images as undirected graphs [14] .The principle of ant colony algorithm used in text blocks detection is based on it: Split the input image into blocks with the size of 8×8, then put the ants into these blocks randomly. Each ant will search the image blocks according to the intensity of the pheromone in blocks. At the end of the algorithm, the text blocks will have more pheromone than the non-text ones. The specific process can be summarized as the following four steps:

1) Initialization. Image I is taken as a graph, the splited blocks can be taken as nodes of graph. *m* ants are randomly put into nodes ; *m* is $\sqrt{\frac{M}{8}\times\frac{N}{8}}$, where *M, N* is the rows' and columns' number of *I*. Initialize parametric variables and the pheromone in each block. The pheromone τ should be set with a positive that tends to 0.

2) Path Finding. The probability of an ant move from the *ith*'s to the adjacent *jth*'s block is:

$$p_{(i,j)}(t)=\frac{\tau_{i,j}(t-1)^{\alpha}{\eta_{i,j}}^{\beta}}{\sum_{j\in\Omega_i}\tau_{i,j}(t-1)^{\alpha}{\eta_{i,j}}^{\beta}}. \tag{4}$$

where $\tau_{i,j}(t-1)$ is the intensity of pheromone in *jth*'s block next to *ith*'s in (t-1)th's time. $\tau_{i,j}(0)$ is initialized with 0.0001 here. $\eta_{i,j}$ is the heuristic information α, β are the relatively importance of pheromone and heuristic information. If β is high in value with respect to α, the algorithm will converge early. Here, we set $\alpha=10, \beta=0.1$. Ω_i is the set of the adjacent nodes of *i*. The definition of the heuristic guiding information will be given in section C.

3)Pheromone updating. In this paper, the pheromone will be updated in both locally and globally. In the local updating, the pheromone will be updated as formula (5) describes when the *kth*'s ant takes one step at *tth*'s time.

$$\tau_{i,j}(t)=\begin{cases}(1-\rho)\cdot\tau_{i,j}(t-1)+\rho\cdot\Delta\tau_{i,j}^{k} & \text{if } (i,j) \text{ visited by the kth ant}\\ \tau_{i,j}(t-1) & \text{else}\end{cases} \quad (5)$$

where ρ is an evaporation coefficient; $\Delta\tau_{i,j}^{k}$ has a relation with the heuristic guiding function, we here define $\Delta\tau_{i,j}^{k}=\eta_{i,j}$. After all ants finish their own travel, we update the global pheromone with (5).

$$\tau(t)=(1-\phi)\cdot\tau(t-1)+\phi.\Delta\tau \quad (6)$$

Where φ is the attenuation coefficient of pheromone.

4) Text Location. When all ants finish the iterations of travel, each block of image has a value of pheromone. If the value exceeds the threshold *T*, it will be taken as a text block. Otherwise, as the non-text block. The *T* is obtained by the method proposed in section D.

C. The Definition of Heuristic Guiding Function.

It can be seen from the above analysis, the function of heuristic information is to guide ants to choose the text blocks, it can be decided with 3 features talked above in section A. 3 windows are selected to compute the value of features. Their size are $16\times48, 48\times16, 16\times16$, where the experiment has shown that 16×48 is suitable to locate the horizontal characters, 48×16 is suitable to locate the vertical characters, while 16×16 is suitable to locate both directions, but the result is less effective than the formers. This paper is mainly focus on news video, where the most text areas are horizontal. Hence, the window-size is 16×48, step-size is 8, that means each sub-block of 8×8 size in the image has a different feature value. We respectively normalize the data of the 3 features. The statistical result has shown that the three

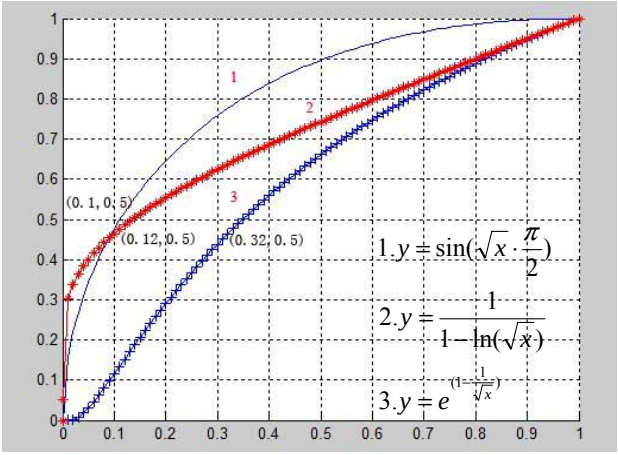


Fig 2 Coordinate graph of 3 conversion function

features have the different classification lines to split the text and non-text, which is respectively about 0.1, 0.12 and 0.32. If these data were directly used, the variance, whose value in (0.1, 1) can be taken as text, would have a greater impact on the result than corners with text recognition interval in (0.32, 1). To balance the weight of each feature, 3 conversion functions are made up to change their classification lines all to about 0.5.

It can be seen from the Fig 2 that the expected result can be achieved, through converting variance with function 1, converting wavelet coefficient with function 2, and converting corners with function 3. Therefore the heuristic guiding function is defined by (7):

$$\eta_{i,j}=\sin(\frac{\pi}{2}\cdot\sqrt{SD_{i,j}})\cdot\frac{1}{1-\ln(\sqrt{VarCoef_{i,j}})}\cdot e^{(1-\frac{1}{\sqrt{Corner_{i,j}}})} \quad (7)$$

where *SD* is the variance of normalization, *VarCoef* is the wavelet coefficient of normalization; *Corner* is the number of corners of normalization.

D. The New Binarization Method

There are many classic binarization algorithm like Bersen[16], LEVBB[17], OTSU[15]. The OTSU is a more popular one. It is a method called maximum variance between clusters which can automatic find a maximum variance between background and foreground to divide images into two parts. The otsu method considers the whole parts value, while isolated points would affect the final result. Here we proposed a new method with iterations, some isolated points can be removed in the first few iterations. The specific process is as follows:

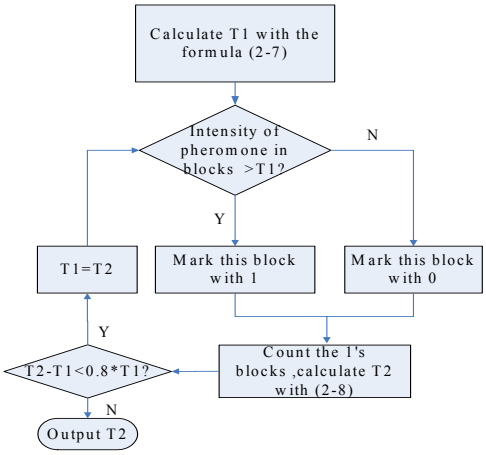


Fig 3 Process of the proposed binarization method

$$T1 = mean + 0.8 * std \ . \quad (8)$$

where *mean* and *std* is the mean and standard deviation of all blocks.

$$T2 = \theta \cdot (mean2 + std2) \ . \quad (9)$$

where *mean2* and *std2* is the mean and standard deviation of the blocks marked with 1. $\theta \in [0,1]$.

The algorithm can converge after only 2-3 iterations, so its speed is faster than the otsu algorithm. At end, the pheromone matrix is binarized with T2, and then close operation is applied using 80×2 structuring element.

E. Text Area Verifing Using Edge Intensity

After all these process, the text areas are initially located. But the false detection rate will be high when the background of the image contains leaves, grass, line-upped team which has the similar intensive edge as characters. Like the case shows in fig 4. To solve this problem, a method about calculating the edge intensity ratio of different directions is proposed. As usual, the Chinese character has similar edge intensity in different directions, but the edge intensity of non-text in different

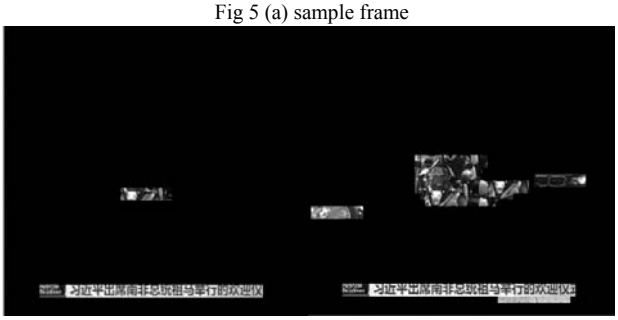
Fig 4. False detection in text location

directions is unpredictable. Firstly, use canny operator to detect the text area's edge that was found above. Then do with the edge graph using Transverse differential, Longitudinal differential, and 45 degree and -45 degree differential (Transverse differential is that move the image to the left with 1 pixel, subtracting it from original image. The others are in the same way). Now we have the horizontal, vertical, 45 degree and -45 degree information of edge. We compute these 4 ratios: horizontal to 45 degree, horizontal to -45 degree, vertical to 45 degree, vertical to -45 degree. Take the maximum of them to be the final edge intensity. The result of edge intensity of fig 3, from up to down, is 2.69, 10, 1.54. Obviously, an appropriate threshold can be to remove the non-text area.

III. Experimental and Comparative Results

The proposed ant colony algorithm is evaluated with 500 news video frames. Fig 4 shows a sample of original frames; fig 6 shows the result of ant colony algorithm text location, with a contrast between the binarization algorithm of otsu method and our method. The *recall* and

Fig 5 (a) sample frame



(b). Result using our method (c). Result using otsu method

Fig 6

precision are stated as follows:

$$recall = \frac{Net}{Tn} \quad (10)$$

$$precision = \frac{Net}{Td} \ . \quad (11)$$

where, *Net* is the exactly detected number of text lines. *Tn* is total number of text lines in test images. *Td* is the total number of text lines that the method detected. Text

line is the one that contains text blocks with the height of 8 pixels.

To prove the effectiveness of the verification of edge intensity, the initial text location and the location with verification are compared. The result is showed in table I.

It can be seen above that the precision is higher after using the verification while recall is a little lower. It is because the method we proposed can remove the most

TABLE I.
Recognition rate of text

Method	*Precision*	*Recall*
Location	65%	91%
Location+verification	80%	89%

false detections as well as some real text blocks under complex background.

The comparison of the proposed method by this paper with the ones proposed in reference [1]: extract the edge feature then using SVM to recognize. And reference [4]: use the wavelet transform to recognize (for comparison, here the window-size is also 16×48, step-size is 8) is listed in table II.

Table II has shown that the proposed method is better than the ones in reference [1] and [4]. The reason is that ref[1] uses SVM while the result of recognition is depended on the train sample. It is difficult to find a good

TABLE II.

THE DIFFERENT ALGORITHM COMPARISON

Method	*Precision*	*Recall*
Ref[1]	75%	83%
Ref[4]	71%	80%
The Proposed method	80%	89%

sample to train when the style of text and non-text area are not unique. Ref[4] uses only one feature, the wavelet coefficient, while our method uses three. Above all, the method proposed can handle more complex cases.

IV. CONCLUSION

In this paper, a text location method based on ant colony algorithm is proposed. The reason why use the ant colony algorithm is that text detection can be seen as a binary-class problem to differentiate text blocks and non-text blocks. Firstly, split the image into blocks, then extract 3 features of the sub-blocks to use ant algorithm for unsupervised classification. Finally, verify it using the edge intensity in different directions to reduce the false detection rate. This paper also compares it with other methods to prove its efficiency.

ACKNOWLEDGMENT

This work is supported by the National High Technology Development 863 Program of China (No.2011AA01A107) and the Zhejiang Provincial Technical Plan Project (No. 2011C13008).

REFERENCES

[1] CHEN D, BOURLARD H, THIRAN J P. Text identification in complex background using SVM[C]. Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2001, 2: 621-626.

[2] PALAIAHNAKOTE S, HUANG W, LIM T. An efficient edge based technique for text detection in video frames [C].The Eighth IAPR Workshop on Document Analysis Systems. Washington, DC: IEEE Computer Society, 2008: 307-314.

[3] WU V, MANMATHA R, RISEMAN E M. Text finder: An automatic system to detect and recognize text in images [J].IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(11):1224-1229.

[4] GLLAVATA J, EWERTH R, FREISLEBEN B. Text detection in images based on unsupervised classification of high-frequency wavelet coefficients[C]. Proceedings of 17th International Conference on Pattern Recognition. Washington, DC: 2004:425-428.

[5] SRIVASTAVI A, KUMAR J. Text detection in scene images using stroke width and nearest-neighbor constraints[C]. Proceedings of IEEE TENCON 2008. Washington, DC:IEEE Computer Society,2008: 1-5.

[6] KIM P-K. Automatic text location in complex color images using local color quantization[C]. TENCON 99: Proceedings of the IEEE Region 10 Conference. Washington, DC: IEEE Computer Society, 1999: 629-632.

[7] Huiping Li, David Doermann, Omid Kia. Automatic text detection and tracking in digital video[C]. IEEE Transactions on Image Processing. 2000, 9(1):147-156.

[8] Haralick R, Shanmugam K, Dinstein I, Textural features for image classification[C], IEEE Transactions on Systems, Man and Cybernetics, 1973, 3(6):610-621.

[9] Michael R.Lyu, Jiqiang Song, Min Cai. A comprehensive method for multilingual video text detection, location and extraction[C]. IEEE Transactions on Systems, 2005,15(2):243-256.

[10] J. Ohya, A. Shio. Recognition characters in scene images. IEEE Trans. Pattern Analysis and Machine Intelligence, 16(2):1994,214-220.

[11] Sara Saatchi and Chih Cheng Hung. Hybridization of the ant colony optimization with the K-means algorithm for clustering [J] .Springer-Verlag Lecture Notes in Computer Science, 2005:511-520.

[12] Haralick R, Shanmugam K, Dinstein I, Textural features for image classification, IEEE Transactions on Systems, Man and Cybernetics, 1973, 3(6):610-621.

[13] Harris C G, Stephens M J. A Combined Corner and Edge Detector. Proceedings of the 4th Alvey Vision Conference. 1988.

[14] LEE M, KIM S, CHO W, et al. Segmentation of brain MR images using an ant colony optimization algorithm [C]. Processing of the 9th IEEE International Conference on Image Processing. 2006, 1:985-988.

[15] Ostu N. A threshold selection method from gray-level histoSystems Man Cybernetic, 1978 (8): 62- 65IEEE Trans.

[16] Bernsen J. Dynamic Thresholding of Gray-level Images. Proc of 8thlntel Conf.on Patt. Recog. Paris, France: IEEE Computer SocietyPress, 1986: l251-l255.

[17] Ridler T W, Calvard S. Picture thresholding using an iterative selection methodIEEE-SNC.1978, 81:630-632.

[18] Zeng F, Zhang G, Jiang J. Text Image with Complex Background Filtering Method Based on Harris Corner-point Detection[J]. Journal of Software, 2013, 8(8): 1827-1834.
[19] Luo Q, Gao Y, Luo J, et al. Automatic Identification of Diatoms with Circular Shape using Texture Analysis[J]. Journal of Software, 2011, 6(3): 428-435.
[20] Wong F, Chao S, Chan W K. Cyclops-Snapshot Translation System Based on Mobile Device[J]. Journal of Software, 2011, 6(9): 1664-1671.

Ming Jiang He received the B.S. degree and M.S. degree in science in 1996 and 2001 respectively, and Ph.D. degree in Computer Science in 2004, all from Zhejiang University, China. He is currently a Professor in college of Computer Science, Hangzhou Dianzi University, China. His research interests include network virtualization, Internet QoS provisioning, and network multimedia processing.

Taotao Zha He received his BSc in Software Engineering from Hangzhou Dianzi University in 2012. Currently he is a master student in this university. His primary research area focuses on image and video processing, text recognition .

Xingqi Wang He received his Bachelor and Master degree from Harbin Institute of Technology in 1997 and 1999, respectively, and Ph.D degree from Zhejiang University in 2002. He is an associate professor in college of Computer Science, Hangzhou Dianzi University, China. All his major are Computer Science. As a researcher, he visited CERCIA, University of Birmingham, UK from 2005 to 2006. His research interests include machine learning, data mining and multimedia content analysis.

Jingfan Tang He received the Ph.D. degree in Computer Science in 2005 from Zhejiang University, China. He is currently an Associate Professor in college of Computer Science, Hangzhou Dianzi University, China. His research interests include network virtualization, quality assurance, process improvement and legacy system re-engineering..

Chunming Wu He received the B.S. degree, M.S. degree and Ph.D. degree in Computer Science from Zhejiang University, China, in 1989, 1992 and 1995 respectively. He is currently a Professor in college of Computer Science, Zhejiang University, and the director of NGNT laboratory. His research fields include Network Multimedia processing, reconfigurable network technology, network virtualization and artificial intelligence

Probability Hypothesis Density Filter Based on Gaussian-Hermite Numerical Integration

Jinguang Chen
School of Computer Science, Xi'an Polytechnic University, Xi'an 710048, China
School of Electronic Engineering, Xidian University, Xi'an 710071, China
Email: xacjg@163.com

Ni Wang, Lili Ma, Tiantian Zhao
School of Computer Science, Xi'an Polytechnic University, Xi'an 710048, China
Email: 807904390@qq.com

Abstract—**This work addresses the multi-target tracking problem in the nonlinear Gaussian system. One probability hypothesis density filtering algorithm based on Gaussian-Hermite numerical integration is proposed. In order to calculate integrations in the Gaussian mixture probability hypothesis density filter, the Gaussian-Hermite numerical integration method is used to approximate the integration. In the filtering stages of prediction and update, we calculate the corresponding Gaussian-Hermite integral points and weights, employ the method of numerical accumulation to approximate the integrations of the Gaussian mixture probability hypothesis density filter. Then the corresponding Gaussian items are calculated and the recursions of Gaussian mixture are implemented. The new algorithm can estimate not only the state vector effectively but also the number of targets accurately. Moreover, its time complexity increases in a low level. The simulation results show that the new algorithm can improve the accuracy of target tracking, and its time complexity keeps the same order of magnitude as the extended Kalman Gaussian mixture probability hypothesis density filter.**

Index Terms—**probability hypothesis density filter, random finite sets, Gaussian-Hermite numerical integration, multi-target tracking, state estimation**

I. Introduction

In recent years, the theory of random finite sets (RFS) is widely used in the fields of information fusion which is dealing with point estimation [1] and target tracking. More and more scholars pay attention to this theory. Random sets theory mainly refers to the random finite sets theory, it can solve the problem of multi-target tracking effectively under a complex environment. It already becomes one of the most popular direction in the multi-target tracking research field [2]. Mahler's probability hypothesis density (PHD) filter [3] is a filtering method based on the framework of RFS. This method represents multi-target state and measurement as random finite sets, and adopts an approach which is similar to the Bayesian theory to implement in a unified style. The complexity of the data association problem is avoided in this progress. Because formulas of the PHD filtering recursion contain integrals, it is generally difficult to obtain the analytical solution. To solve this problem, Vo et al. propose the Gaussian mixture PHD (GM-PHD) filter [4] which is applicable to the linear Gaussian systems. This algorithm assumes that the multi-target PHD could be written as the form of Gaussian mixture, then at each time step, the prediction and update PHD can also subject to the distribution of Gaussian mixture. Thus the recursive Gaussian mixture PHD filter algorithm is derived. Furthermore, Clark et al. prove that the GM-PHD filter is convergence, and present the error boundary in trim and merge phases [5]. For multi-target tracking problems with nonlinear Gaussian assumption, the extended Kalman PHD (EK-PHD) filtering algorithm is given in [4]. EK-PHD uses the method of Taylor series expansion to get the local linearization of nonlinear functions, and then the GM-PHD is employed directly. However, only under the condition of weak nonlinear system, the EK-PHD can get satisfied filtering accuracy. If system is strong nonlinear, due to the large linear truncation errors, the filtering accuracy become low. In order to improve the accuracy in the strong nonlinear system, article [6] which is based on particle filter [7] presents the particle PHD (P-PHD) filter, also known as sequential Monte Carlo PHD (SMC-PHD). The algorithm uses a large number of particles and weights to approximate the nonlinear transformation of random variables. It can obtain higher filtering accuracy and can also apply to the condition of the system which is nonlinear and non-Gaussian. Then the convergence of P-PHD is analyzed and proved in [8], and the boundary of the mean square error (MSE) is derived. Literature [9] adopts unscented particle filter (UPF) to implement P-PHD filter, which uses unscented Kalman filter to get better importance density function and sample particles from it. In this way, good error performance is acquired. However, time complexity of this algorithm is very high

Manuscript received on September 7, 2013.
Corresponding author, Jinguang Chen, Email: xacjg@163.com.

doi:10.4304/jcp.9.5.1096-1102

and real-time performance in filtering stage is poor. So, people try to find some other PHD filters whose complexities are not so high. Such as the unscented Kalman PHD (UK-PHD)[4] filter with nonlinear Gaussian assumption, which uses unscented transform to determine the sampling points to characterize the statistical properties of Gaussian random vector, and approximates the state of the system's posterior probability. In [10], the central difference Kalman (CDK) filter is combined with PHD, and the CDK-PHD filter is proposed. This method uses the Stirling interpolation formula to approximate the polynomial of nonlinear function. It improves the estimation accuracy. Cubature Kalman Filter (CKF) [11] can be combined with the PHD, and the cubature PHD filter is given in [12]. This algorithm adopts the sampling rules of third-order spherical cubature-radial to compute the probability distribution of target state, which solves the computing problems of nonlinear state equation and observation equation.

As the Gaussian-Hermite numerical integration is applied to the nonlinear Gaussian filtering system, it is the Gaussian-sum quadrature Kalman (QKF) filter [13]. Under the inspiration of the work above, we apply the Gaussian-Hermite numerical integration method to the PHD filter process, and obtain a new PHD filter which can deal with multi-target nonlinear tracking system, namely the Gaussian-Hermite probability hypothesis density filter (GH-PHD). Compared with the EK-PHD filter algorithm, new algorithm improves the filtering accuracy. Although time complexity of this algorithm increases, it keeps the same order of magnitude with EK-PHD filter algorithm. So, it is acceptable in many practical applications.

II. Problem Formulation

In the process of multi-target tracking, old targets may be disappearing and new targets may be appearing in one time step so that the number of targets is changing over time. Suppose the number of targets is $M(k)$ at time k, by using the random finite sets, the state set of targets can be represented as $\boldsymbol{X}_k = \{\boldsymbol{x}_{k,1}, \cdots \boldsymbol{x}_{k,M(k)}\}$ and the measurement set can be represented as $\boldsymbol{Z}_k = \{\boldsymbol{z}_{k,1}, \cdots \boldsymbol{z}_{k,N(k)}\}$, where $N(k)$ is the measurement number at time k.

Suppose that the state set is $\boldsymbol{X}_k$ at time k, then at time $k+1$ the state set of targets can be expressed as

$$\boldsymbol{X}_{k+1} = S_{k+1|k}(\boldsymbol{X}_k) \bigcup B_{k+1|k}(\boldsymbol{X}_k) \bigcup \Gamma_{k+1} \tag{1}$$

where $S_{k+1|k}(\boldsymbol{X}_k)$ is the RFS of targets survival, $B_{k+1|k}(\boldsymbol{X}_k)$ is the RFS of targets spawned at time $k+1$ from previous targets with state $\boldsymbol{X}_k$, Γ_{k+1} is the RFS of spontaneous birth at time $k+1$. Usually we use $f(\boldsymbol{X}_{k+1} | \boldsymbol{X}_k)$ to express the transition density of multi-target state.

Similarly, at time k the set of measurements can be expressed as

$$\boldsymbol{Z}_k = \boldsymbol{K}_k \bigcup \Theta_k(\boldsymbol{X}_k) \tag{2}$$

where $\boldsymbol{K}_k$ is the measurement random set of false measurements or clutter. $\Theta_k(\boldsymbol{X}_k)$ is the measurement random set produced by the real targets. Usually we use $g(\boldsymbol{Z}_k | \boldsymbol{X}_k)$ to express the measurement likelihood function.

In the target tracking system, it is usually assumed that the dynamic model and measurement model of a single target are represented as follows

$$\boldsymbol{x}_k = f(\boldsymbol{x}_{k-1}) + \boldsymbol{\omega}_k \tag{3}$$

$$\boldsymbol{z}_k = h(\boldsymbol{x}_k) + \boldsymbol{\upsilon}_k \tag{4}$$

where $\boldsymbol{\omega}_k$ and $\boldsymbol{\upsilon}_k$ are both the additive Gaussian noises. $\boldsymbol{x}_k$ is the state vector, $f(\cdot)$ and $h(\cdot)$ are the transition function and the measurement function, respectively. Assume $p(\boldsymbol{x}_0)$ is the initial state distribution, the purpose of target tracking is to estimate the posterior distribution recursively, thus to estimate the target state and the target number.

III. Gaussian-Hermite Numerical Integration

Gaussian-Hermite filter is a nonlinear Bayes filter under the assumption of Gaussian distribution. It is a kind of recursive filtering method based on Gaussian-Hermite numerical integration [13], which is implemented by choosing integral points and the corresponding weights to enhance the accuracy of the system state mean and the variance estimate [14].

Assume that $g(x)$ is a weighted integral function on an interval (a,b), then the integral can be expressed as

$$I(g) = \int_a^b W(x) g(x) dx \tag{5}$$

where $W(x)$ is a weighted function. If we use m numerical points to integrate, formula (5) can be approximated as

$$I(g) \approx \sum_{i=1}^{m} w_i g(\xi_i) \tag{6}$$

where ξ_i is the standard integral point, w_i is its corresponding weight.

Firstly, we consider one-dimensional situation, it is assumed that a random variable x with a Gaussian probability density is $p(x)=\mathrm{N}(x;0,1)$. The expectation of the function $g(x)$ can be approximated as

$$\begin{aligned} \mathrm{E}[g(x)] &= \int_R g(x) \mathrm{N}(x;0,1) dx \\ &\approx \sum_{l=1}^{m} w_l g(\xi_l) \end{aligned} \tag{7}$$

The method which uses root-finding to calculate the integral points is unstable on the arithmetic, so we adopt the method which is proposed in literature [13] to get integral points and weights. This method exploits the relationship between orthogonal polynomials and the tridiagonal matrix. Assume that $\boldsymbol{J}$ is a symmetric

tridiagonal matrix with zero diagonal elements and other elements are

$$J_{i,i+1} = \sqrt{i/2}, 1 \le i \le (m-1) \tag{8}$$

Here, m is the number of standard integral points. In fact, the specific number of integral points is determined by m and the dimension of state vector. For example, if $m = 5$, then the dimension of state vector is 4, it will have 625 points. If $m > 3$, through experiments we found that the filtering precision of algorithm improvement is not big, but its time complexity increases significantly. For these reasons, we select $m = 3$ in the experiment. The standard integral point is taken to be $\xi_l = \sqrt{2}\varepsilon_l$, where ε_l is the l-th eigenvalue of matrix $\boldsymbol{J}$; the corresponding weight is $w_l = (v_l)_1^2$, where $(v_l)_1$ is the first element of the l-th normalized eigenvector of matrix $\boldsymbol{J}$.

Furthermore, we can extend one dimensional case to multi-dimensional case. Assume a random vector $\boldsymbol{x}$ has a Gaussian density $p(\boldsymbol{x}) = \mathrm{N}(\boldsymbol{x}; \boldsymbol{0}, \boldsymbol{I}_{n_x})$, where $\boldsymbol{I}_{n_x}$ is the identity matrix with $n_x \times n_x$ dimensions. Since each element of $\boldsymbol{x}$ is mutually uncorrelated, integral rule of n_x dimensions Gaussian-Hermite is as follows:

$$\begin{aligned} \mathrm{E}[g(\boldsymbol{x})] &= \int_R g(\boldsymbol{x})\mathrm{N}(\boldsymbol{x}; \boldsymbol{0}, \boldsymbol{I})d\boldsymbol{x} \\ &\approx \sum_{l_{n_x}=1}^{m} w_{l_{n_x}} \cdots \sum_{l_1=1}^{m} w_{l_1} g(\xi_{l_1} \cdots \xi_{l_{n_x}}) \\ &= \sum_{l=1}^{m^{n_x}} w_l g(\boldsymbol{\xi}_l). \end{aligned} \tag{9}$$

where $\boldsymbol{\xi}_l = [\xi_{l_1} \cdots \xi_{l_{n_x}}]^{\mathrm{T}}, w_l = \prod_{j=1}^{n_x} w_{l_j}$.

Moreover, we further assume that a random vector $\boldsymbol{x}$ has a Gaussian density $p(\boldsymbol{x}) = \mathrm{N}(\boldsymbol{x}; \hat{\boldsymbol{x}}, \boldsymbol{P})$, do Cholesky decomposition of $\boldsymbol{P}$, and get $\boldsymbol{P} = \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}}$, $\boldsymbol{y} = \boldsymbol{S}^{-1}(\boldsymbol{x} - \hat{\boldsymbol{x}})$, then

$$\begin{aligned} \mathrm{E}[g(\boldsymbol{x})] &= \int g(\boldsymbol{x})\mathrm{N}(\boldsymbol{x}; \hat{\boldsymbol{x}}, \boldsymbol{P})d\boldsymbol{x} \\ &= \int g(\boldsymbol{S}\boldsymbol{y} + \hat{\boldsymbol{x}})\mathrm{N}(\boldsymbol{y}; \boldsymbol{0}, \boldsymbol{I}_{n_x})d\boldsymbol{y} \\ &\approx \sum_{l_{n_x}=1}^{m} w_{l_{n_x}} \cdots \sum_{l_1=1}^{m} w_{l_1} g(\boldsymbol{S}[\xi_{l_1} \cdots \xi_{l_{n_x}}]^{\mathrm{T}} + \hat{\boldsymbol{x}}) \\ &= \sum_{l=1}^{m^{n_x}} w_l g(\boldsymbol{S}\xi_l + \hat{\boldsymbol{x}}) \\ &= \sum_{l=1}^{m^{n_x}} w_l g(\boldsymbol{x}_l). \end{aligned} \tag{10}$$

Now, integral points can be obtained by

$$\boldsymbol{x}_l = \boldsymbol{S}\xi_l + \hat{\boldsymbol{x}} \tag{11}$$

IV. Probability Hypothesis Density Filter

Traditional multi-target tracking algorithms are related to data association, which means we need to determine the corresponding relationship between tracks and measurements. The computational complexity of data association grows exponentially along with the increase of the number of targets and measurements. Mahler's PHD filter is a kind of target tracking algorithm based on random sets theory [3]. This algorithm can avoid the complex progress of data association and can deal with multi-target tracking problem in an effective manner. The traditional Bayes filter propagates global probability density, but the calculation of global probability density in multi-target tracking is very difficult. Aiming at this problem, PHD propagates first-order statistics of the random finite sets via the posterior probability density. Also because of PHD propagation is posterior intensity of the state space, its integral in any state space is the expectation of targets' number. Therefore, PHD filter can not only track the multi-target state when target number is unknown or changing over time, but also estimate the target number. Similar to the Bayes filter in multi-target tracking, the recursions in PHD filter also include prediction stage and update stage.

At time k, the PHD prediction formula is

$$\begin{aligned} D_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{Z}_{1:k-1}) &= \gamma_k(\boldsymbol{x}_k) + \int [p_{S,k}(\boldsymbol{x}_{k-1}) f_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{x}_{k-1}) \\ &\quad + \beta_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{x}_{k-1})] D_{k-1}(\boldsymbol{x}_{k-1} | \boldsymbol{Z}_{1:k-1}) d\boldsymbol{x}_{k-1} \end{aligned} \tag{12}$$

where $\gamma_k(\boldsymbol{x}_k)$ is intensity of the spontaneous birth RFS at time k, $\beta_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{x}_{k-1})$ is intensity of the RFS spawned at time k by a target with previous state $\boldsymbol{x}_{k-1}$, $p_{S,k}(\boldsymbol{x}_{k-1})$ is the probability that a target with previous state $\boldsymbol{x}_{k-1}$ still exists at time k. $f_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{x}_{k-1})$ is the transition probability density of target state. $D_{k|k-1}(\boldsymbol{x}_{k-1} | \boldsymbol{Z}_{1:k-1})$ is the posterior intensity of target at time $k-1$.

The integral of PHD prediction function $D_{k|k-1}(\cdot)$ is the estimation of target number, i.e.,

$$N_{k|k-1} = \int D_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{Z}_{1:k-1}) d\boldsymbol{x}_k = N_{k|k-1}^{\Gamma} + N_{k|k-1}^{S} + N_{k|k-1}^{B} \tag{13}$$

where

$$N_{k|k-1}^{\Gamma} = \int \gamma_k(\boldsymbol{x}_k) d\boldsymbol{x}_k \tag{14}$$

$$N_{k|k-1}^{S} = \int p_{S,k}(\boldsymbol{x}_{k-1}) f_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{x}_{k-1}) D_{k-1}(\boldsymbol{x}_{k-1} | \boldsymbol{Z}_{1:k-1}) d\boldsymbol{x}_{k-1} \tag{15}$$

$$N_{k|k-1}^{B} = \int \beta_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{x}_{k-1}) D_{k-1}(\boldsymbol{x}_{k-1} | \boldsymbol{Z}_{1:k-1}) d\boldsymbol{x}_k \tag{16}$$

Here, formulas (14)-(16) express the expectation of spontaneous birth target number, the expectation of survival target number, and the expectation of spawned target number, respectively.

At time k, the PHD update formula is

$$\begin{aligned} D_k(\boldsymbol{x}_k | \boldsymbol{Z}_{1:k}) &= (1 - p_{D,k}(\boldsymbol{x}_k)) D_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{Z}_{1:k-1}) + \\ &\sum_{z_k \in \boldsymbol{Z}_k} \frac{p_{D,k}(\boldsymbol{x}_k) g_k(z_k | \boldsymbol{x}_k) D_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{Z}_{1:k-1})}{\kappa_k(z) + \int p_{D,k}(\boldsymbol{x}_k) g_k(z_k | \boldsymbol{x}_k) D_{k|k-1}(\boldsymbol{x}_k | \boldsymbol{Z}_{1:k-1}) d\boldsymbol{x}_k} \end{aligned} \tag{17}$$

where $\kappa_k(z)=\lambda_k c_k(z)$ is the intensity of clutter RFS at time k, $c_k(z)$ is probability density of clutter, λ_k is the average number of clutter. Assume that the number of clutter which appears at each time obeys Poisson distribution, $g_k(\boldsymbol{z}_k|\boldsymbol{x}_k)$ is the measurement likelihood function of target, $p_{D,k}(\boldsymbol{x}_k)$ is the detection probability.

Similarly, the update formula of the expectation value of the target number is

$$\begin{aligned} N_k &= \int D_k(\boldsymbol{x}_k|\boldsymbol{Z}_{1:k})d\boldsymbol{x}_k \\ &= N_{k|k-1} - \int p_{D,k}(\boldsymbol{x}_k)D_{k|k-1}(\boldsymbol{x}_k|\boldsymbol{Z}_{1:k-1})d\boldsymbol{x}_k \\ &\quad + \sum_{z\in Z_k}\frac{\int p_{D,k}(\boldsymbol{x}_k)g_k(z_k|\boldsymbol{x}_k)D_{k|k-1}(\boldsymbol{x}_k|\boldsymbol{Z}_{1:k-1})d\boldsymbol{x}_k}{\kappa_k(z)+\int p_{D,k}(\boldsymbol{x}_k)g_k(z_k|\boldsymbol{x}_k)D_{k|k-1}(\boldsymbol{x}_k|\boldsymbol{Z}_{1:k-1})d\boldsymbol{x}_k} \end{aligned} \tag{18}$$

From the prediction equation (12) and update equation (17) above, there are integrals among them. These integrations have no analytical solution in generally. In fact, the integral form can be written as an integral form of a nonlinear function multiplied by a Gaussian distribution, i.e.,

$$\int \text{nonlinear function} * \text{Gaussian distribution}\, d\boldsymbol{x} \tag{19}$$

This kind of integral form can be approximated by using the method of formula (10), thus gaining a Gaussian-Hermite PHD filter algorithm.

V. GAUSSIAN-HERMITE PHD FILTER

Assume that the collection for Gaussian mixture components is $\{w_{k-1}^{(i)},\boldsymbol{m}_{k-1}^{(i)},\boldsymbol{P}_{k-1}^{(i)}\}_{i=1}^{J_{k-1}}$ at time $k-1$, and at time k the measurement set is Z_k. For birth targets, the prediction stage is the same as that of the GM-PHD filter in which the numerical integrations are not needed. For survival targets, it uses numerical integration in recursions. In time update stage, for the variance of every Gaussian items, do Cholesky decomposition, i.e.,

$$\boldsymbol{P}_{k-1}^{(i)}=\boldsymbol{S}_{k-1}^{(i)}(\boldsymbol{S}_{k-1}^{(i)})^{\mathrm{T}} \tag{20}$$

Then, use formula (10) to calculate the integral points of each Gaussian component

$$\boldsymbol{x}_{l,k-1}^{(i)}=\boldsymbol{S}_{k-1}^{(i)}\xi_l+\boldsymbol{m}_{k-1}^{(i)} \tag{21}$$

Next, predict each integral point according to the state transition equation

$$\boldsymbol{x}_{l,k|k-1}^{(i)}=f(\boldsymbol{x}_{l,k-1}^{(i)}) \tag{22}$$

Finally, one-step prediction mean and variance of survival targets are given as

$$\boldsymbol{m}_{k|k-1}^{(i)}=\sum_{l=1}^{m}w_l\boldsymbol{x}_{l,k|k-1}^{(i)} \tag{23}$$

$$\boldsymbol{P}_{k|k-1}^{(i)}=\sum_{l=1}^{m}w_l\boldsymbol{x}_{l,k|k-1}^{(i)}(\boldsymbol{x}_{l,k|k-1}^{(i)})^{\mathrm{T}}-\hat{\boldsymbol{x}}_{k|k-1}^{(i)}(\hat{\boldsymbol{x}}_{k|k-1}^{(i)})^{\mathrm{T}}+\boldsymbol{Q}_k \tag{24}$$

where w_l is the corresponding weight of Gaussian-Hermite integral point and $\boldsymbol{Q}_k$ is the variance of process noise.

Assume that the prediction result of the component collection for Gaussian Mixture is $\{w_{k|k-1}^{(i)},\boldsymbol{m}_{k|k-1}^{(i)},\boldsymbol{P}_{k|k-1}^{(i)}\}_{i=1}^{J_{k|k-1}}$, in the measurement update, we do Cholesky decomposition for $\boldsymbol{P}_{k|k-1}^{(i)}$ as well as formula (20), i.e., $\boldsymbol{P}_{k|k-1}^{(i)}=\boldsymbol{S}_{k|k-1}^{(i)}(\boldsymbol{S}_{k|k-1}^{(i)})^{\mathrm{T}}$, then the new integral point is

$$\boldsymbol{x}_{l,k|k-1}^{(i)}=\boldsymbol{S}_{k|k-1}^{(i)}\xi_l+\boldsymbol{m}_{k|k-1}^{(i)} \tag{25}$$

Then, calculate integral point of measurement prediction

$$z_{k|k-1}^{(i)}=\sum_{l=1}^{m}w_l h(\boldsymbol{x}_{l,k|k-1}^{(i)}) \tag{26}$$

Next, state update and covariance update of the integral point are calculated as

$$\boldsymbol{x}_k^{(i)}=\boldsymbol{m}_{k|k-1}^{(i)}+\boldsymbol{K}_k^{(i)}(z^{(i)}-z_{k|k-1}^{(i)}) \tag{27}$$

$$\boldsymbol{P}_k^{(i)}=\boldsymbol{P}_{k|k-1}^{(i)}-\boldsymbol{K}_k^{(i)}\boldsymbol{P}_{zz}^{(i)}(\boldsymbol{K}_k^{(i)})^{\mathrm{T}} \tag{28}$$

where $h(\cdot)$ is the measurement function of target, $z^{(i)}$ is the actual measurement value of the corresponding time, $\boldsymbol{K}_k^{(i)}$ is the filter gain, it is calculated as follows:

$$\boldsymbol{K}_k^{(i)}=\boldsymbol{P}_{xz}^{(i)}(\boldsymbol{P}_{zz}^{(i)})^{-1} \tag{29}$$

$$\boldsymbol{P}_{zz}^{(i)}=\boldsymbol{R}_k+\sum_{l=1}^{m}w_l\boldsymbol{z}_{l,k|k-1}^{(i)}(\boldsymbol{z}_{l,k|k-1}^{(i)})^{\mathrm{T}}-\boldsymbol{z}_{k|k-1}^{(i)}(\boldsymbol{z}_{k|k-1}^{(i)})^{\mathrm{T}} \tag{30}$$

$$\boldsymbol{P}_{xz}^{(i)}=\sum_{l=1}^{m}w_l\boldsymbol{x}_{l,k|k-1}^{(i)}(\boldsymbol{z}_{l,k|k-1}^{(i)})^{\mathrm{T}}-\boldsymbol{m}_{k|k-1}^{(i)}(\boldsymbol{z}_{k|k-1}^{(i)})^{\mathrm{T}} \tag{31}$$

where $\boldsymbol{z}_{l,k|k-1}^{(i)}=h(\boldsymbol{x}_{l,k|k-1}^{(i)})$.

From the description above, the pseudo code for the GH-PHD filter is summarized as below:

Step 1: prediction for spontaneous birth targets

$i=0$

for $j=1,\cdots,J_{\gamma,k}$

$i:=i+1$

$\boldsymbol{m}_{k|k-1}^{(i)}=\boldsymbol{m}_{\gamma,k}^{(j)}$, $w_{k|k-1}^{(i)}=w_{\gamma,k}^{(j)}$, $\boldsymbol{P}_{k|k-1}^{(i)}=\boldsymbol{P}_{\gamma,k}^{(j)}$

end

Step 2: prediction for existing targets

$\boldsymbol{P}_{k-1}^{(j)}=\boldsymbol{S}_{k-1}^{(j)}(\boldsymbol{S}_{k-1}^{(j)})^{\mathrm{T}}$

for $j=1,...,J_{k-1}$

$i:=i+1$

for $l=1,\cdots,m$

$\boldsymbol{x}_{l,k-1}^{(i)}=\boldsymbol{S}_{k-1}^{(j)}\xi_l+\boldsymbol{m}_{k-1}^{(j)}$

$\bar{\boldsymbol{x}}_{l,k|k-1}^{(i)}=f(\boldsymbol{x}_{l,k|k-1}^{(i)})$

end

$w_{k|k-1}^{(i)}=p_{S,k}w_{k-1}^{(j)}$, $\boldsymbol{m}_{k|k-1}^{(i)}=\sum_{l=1}^{m}w_l\bar{\boldsymbol{x}}_{l,k|k-1}^{(i)}$

$\boldsymbol{P}_{k|k-1}^{(i)}=\sum_{l=1}^{m}w_l\bar{\boldsymbol{x}}_{l,k|k-1}^{(i)}(\bar{\boldsymbol{x}}_{l,k|k-1}^{(i)})^{\mathrm{T}}-\boldsymbol{m}_{k|k-1}^{(i)}(\boldsymbol{m}_{k|k-1}^{(i)})^{\mathrm{T}}+\boldsymbol{Q}_k$

end

$J_{k|k-1} = i$

Step 3: each component in the measurement update

for $j = 1,..., J_{k|k-1}$

$$\hat{\boldsymbol{z}}_{k|k-1}^{(j)} = \sum_{l=1}^{m} w_l \boldsymbol{z}_{l,k|k-1}^{(j)}$$

$$\boldsymbol{P}_{zz} = \boldsymbol{R}_k + \sum_{l=1}^{m} w_l \boldsymbol{z}_{l,k|k-1}^{(j)} (\boldsymbol{z}_{l,k|k-1}^{(j)})^{\mathrm{T}} - \hat{\boldsymbol{z}}_{k|k-1}^{(j)} (\hat{\boldsymbol{z}}_{k|k-1}^{(j)})^{\mathrm{T}}$$

$$\boldsymbol{P}_{xz} = \sum_{l=1}^{m} w_l \boldsymbol{x}_{l,k|k-1} (\boldsymbol{z}_{l,k|k-1}^{(j)})^{\mathrm{T}} - \boldsymbol{m}_{k|k-1}^{(j)} (\hat{\boldsymbol{z}}_{k|k-1}^{(j)})^{\mathrm{T}}$$

$$\boldsymbol{K}_k^{(j)} = \boldsymbol{P}_{xz} \boldsymbol{P}_{zz}^{-1}, \quad \boldsymbol{P}_{k|k}^{(j)} = \boldsymbol{P}_{k|k-1}^{(j)} - \boldsymbol{K}_k^{(j)} \boldsymbol{P}_{zz} (\boldsymbol{K}_k^{(j)})^{\mathrm{T}}$$

end

Step 4: measurement update

for $j = 1,..., J_{k|k-1}$

$$w_k^{(j)} = (1 - p_{D,k}) w_{k|k-1}^{(j)}$$

$$\boldsymbol{m}_k^{(j)} = \boldsymbol{m}_{k|k-1}^{(j)}, \quad \boldsymbol{P}_k^{(j)} = \boldsymbol{P}_{k|k-1}^{(j)}$$

end

$i := 0$

for $z \in \boldsymbol{Z}_k$

$i := i + 1$

for $j = 1,..., J_{k|k-1}$

$$\boldsymbol{P}_{k|k-1}^{(i)} = \boldsymbol{S}_{k|k-1}^{(i)} (\boldsymbol{S}_{k|k-1}^{(i)})^{\mathrm{T}}$$

for $l = 1, \cdots m$

$$\boldsymbol{x}_{l,k|k-1}^{(i)} = \boldsymbol{S}_{k|k-1}^{(i)} \xi_l + \boldsymbol{m}_{k|k-1}^{(i)}$$

$$\boldsymbol{z}_{l,k|k-1}^{(i)} = h(\boldsymbol{x}_{l,k|k-1}^{(i)})$$

end

$$w_k^{(i)} = p_{D,k} w_{k|k-1}^{(j)} \mathrm{N}(\boldsymbol{z}^{(j)}; \hat{\boldsymbol{z}}_{k|k-1}^{(j)}, \boldsymbol{P}_{zz}),$$

$$\boldsymbol{m}_k^{(i)} = \boldsymbol{m}_{k|k-1}^{(j)} + \boldsymbol{K}_k^{(j)} (\boldsymbol{z}^{(j)} - \hat{\boldsymbol{z}}_{k|k-1}^{(j)})$$

$$\boldsymbol{P}_k^{(i)} = \boldsymbol{P}_{k|k}^{(j)}$$

end

$$w_k^{(i)} = \frac{w_k^{(j)}}{\kappa_k(z) + \sum_{i=1}^{J_{k|k-1}} w_k^{(i)}}, \quad j = 1,..., J_{k|k-1}$$

end

$J_k = i J_{k|k-1} + J_{k|k-1}$

Output: $\{w_k^{(i)}, \boldsymbol{m}_k^{(i)}, \boldsymbol{P}_k^{(i)}\}_{i=1}^{J_k}$.

The number of Gaussian items for the posterior probability density will be increasing as time passing. As a result, a large number of calculation time is wasted to update the Gaussian items which have small weights. In order to control the number of Gaussian items, we can use pruning and merging method which is described in [4]. This can be done by setting a truncation threshold T and a merging threshold U. In pruning stage, Some Gaussian items are abandoned as their weights are smaller than truncation threshold. As a result, the Gaussian items whose weights are greater than truncation threshold are kept. In merging stage, give two Gaussian items $\{w_k^{(i)}, \boldsymbol{m}_k^{(i)}, \boldsymbol{P}_k^{(i)}\}$ and $\{w_k^{(j)}, \boldsymbol{m}_k^{(j)}, \boldsymbol{P}_k^{(j)}\}$, when the means and variances of the two Gaussian items meet $(\boldsymbol{m}_k^{(i)} - \boldsymbol{m}_k^{(j)})^{\mathrm{T}} (\boldsymbol{P}_k^{(i)})^{-1} (\boldsymbol{m}_k^{(i)} - \boldsymbol{m}_k^{(j)}) \le U$, the two Gaussian items will be merged.

For multi-target state extraction, the clustering method is commonly used in the P-PHD filter[15]. Under the condition of Gaussian mixture, we generally use the method mentioned in [4] to extract targets' state. Given a threshold, the Gaussian component with weight which is greater than the threshold can be regarded as a target and the corresponding state is the estimation of the target.

VI. Simulation Experiment and Result Analysis

In the simulation region, assume that there are two targets. Each target is moving via constant velocity model or constant turn model. Target 1 appears at time $k = 1$, and dies at time $k = 40$; target 2 appears at time $k = 6$, and dies at time $k = 49$. Two targets both travel in straight lines before time $k = 16$, then they are making turns until time $k = 34$, and the targets resume straight trajectories after time $k = 34$.

The state vector of targets is $\boldsymbol{x}_k = [x \quad \dot{x} \quad y \quad \dot{y}]^{\mathrm{T}}$, it consists of position component (x, y) and velocity component $(\dot{x}, \dot{y})$. Each target has survival probability $p_{S,k} = 0.99$ and detection probability $p_{D,k} = 0.98$. The corresponding motion equation is

$$\boldsymbol{x}_k = \boldsymbol{F} \boldsymbol{x}_{k-1} + \boldsymbol{G} \boldsymbol{\omega}_k \tag{32}$$

where transition matrix of state noise is $\boldsymbol{G} = \begin{bmatrix} T^2/2 & T & 0 & 0 \\ 0 & 0 & T^2/2 & T \end{bmatrix}^{\mathrm{T}}$. When targets do constant velocity motion, $\boldsymbol{F} = \mathrm{blkdiag}\left(\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)$; when targets do constant turn motion, $\boldsymbol{F} = \begin{bmatrix} 1 & \frac{\sin \Omega T}{\Omega} & 0 & -\frac{1 - \cos \Omega T}{\Omega} \\ 0 & \cos \Omega T & 0 & -\sin \Omega T \\ 0 & \frac{1 - \cos \Omega T}{\Omega} & 1 & \frac{\sin \Omega T}{\Omega} \\ 0 & \sin \Omega T & 0 & \cos \Omega T \end{bmatrix}$. $\boldsymbol{\omega}_k$ is white Gaussian noise, state noise matrix is $\boldsymbol{Q}_k = \mathrm{diag}([0.5 \quad 0.5])$, sampling period is $T = 1$ s, and the turn rate is $\Omega = (5\pi/80)$rad/s.

Measurement equation of the system is

$$\boldsymbol{z}_k = \begin{bmatrix} \theta_k \\ r_k \end{bmatrix} = \begin{bmatrix} \arctan(\dfrac{y}{x}) \\ \sqrt{x^2 + y^2} \end{bmatrix} + \boldsymbol{v}_k \tag{33}$$

where $\boldsymbol{v}_k \sim \mathrm{N}(\cdot; \boldsymbol{0}, \boldsymbol{R}_k)$, and $\boldsymbol{R}_k = \mathrm{diag}([\sigma_{\theta_k}^2 \quad \sigma_{r_k}^2])$, $\sigma_{\theta_k} = 2\times(\pi/180)$ rad/s, $\sigma_{r_k} = 8$ m. Measurement vector $\boldsymbol{z}_k$ includes two components, the bearing θ_k and range r_k.

For convenience, assume that there are no spawned targets. And suppose PHD of spontaneous birth target random sets is

$$\gamma_k(\boldsymbol{x}) = 0.1\mathrm{N}(\boldsymbol{x}; \boldsymbol{m}_\gamma^{(1)}, \boldsymbol{P}_\gamma) + 0.1\mathrm{N}(\boldsymbol{x}; \boldsymbol{m}_\gamma^{(2)}, \boldsymbol{P}_\gamma)$$

where

$\boldsymbol{m}_\gamma^{(1)} = [-1000 \quad 60 \quad 500 \quad 0]^{\mathrm{T}}$,

$\boldsymbol{m}_\gamma^{(2)} = [1050 \quad -62 \quad 1070 \quad 0]^{\mathrm{T}}$,

$\boldsymbol{P}_\gamma = \mathrm{diag}([100\ 40\ 100\ 40])$. Clutter is uniformly distributed in the surveillance region, and the number of clutter subjects to a Poisson distribution whose mean is $r = 5$. Pruning threshold is $T_{\mathrm{prun}} = 1\mathrm{e}-5$, merging threshold is $U_{\mathrm{prun}} = 4$. $c_{\mathrm{ospa}} = 70$ is the adjustment factor of state error and cardinality error, $p_{\mathrm{ospa}} = 2$ is the distance of OSPA, $J_{\max} = 100$ is the largest number of Gaussian components. The entire time of the simulation is 49 s, the surveillance region is $[-\pi/2, \pi/2]\mathrm{rad}\times[0,1600]\mathrm{m}$.

Results of the GH-PHD filter and the EK-PHD filter are shown in Fig 1. From Fig 1, we are easily known that EK-PHD filter and GH-PHD filter both have better estimate of targets. The difference is that the tracking performance of GH-PHD filter is better than that of EK-PHD. EK-PHD filter cannot accurately estimate the location of targets, it will also leak with some targets in the turn stage, and GH-PHD filter can accurately estimate the location of targets. Target trajectories and measurements in the area of the surveillance are shown in Fig 2. The marks of blue " o " express the clutter distribution, the marks of red "*" are the true trajectory of targets. $r = 5$ is the number of clutter and it subjects to the Poisson distribution, the clutter uniformly distributes in the whole surveillance area.

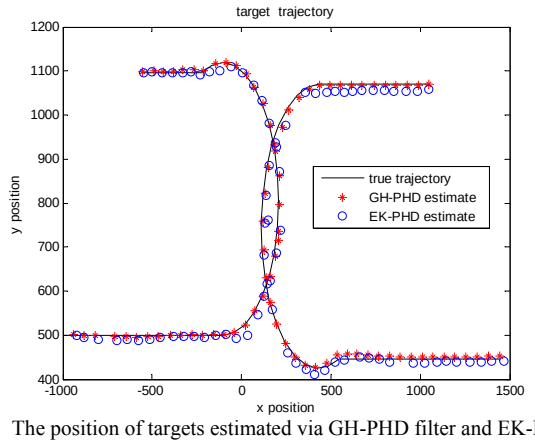


Fig 1 The position of targets estimated via GH-PHD filter and EK-PHD filter

Fig 3 displays the true number and the estimation number of targets of GH-PHD filter and EK-PHD filter throughout the simulation by time step. From Fig 3, we know that target number estimated by GH-PHD algorithm matched with the true target number well. It appears that the deviation of estimation only happened at time $k = 18$ and $k = 19$, at other time it can accurately estimate the number of targets; but the target number which is estimated by EK-PHD filter algorithm and the true target number have a few differences.

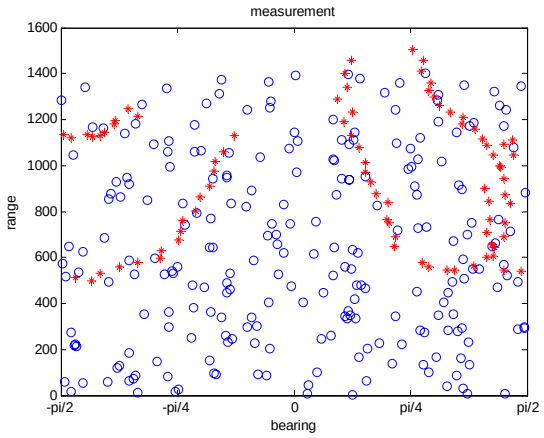


Fig 2 Clutters and measurements

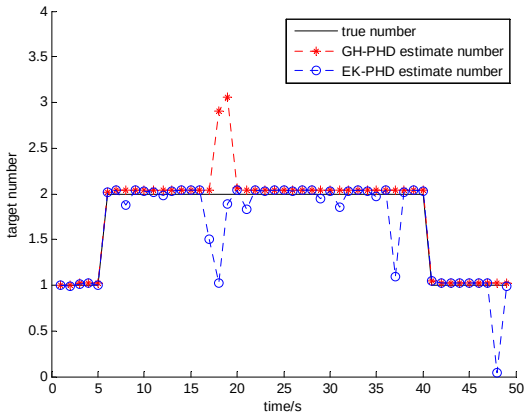


Fig 3 The true number and the estimation number of targets

In order to evaluate the accuracy of multi-target tracking filter, optimal subpattern assignment distance (OSPA)[16] is used, which can measure the difference between sets well, and it is one of the most popular evaluation criteria which has been used by many scholars in recent years. OSPA distance in the simulation is displayed in Fig 4. We are easy to know that the OSPA of GH-PHD filter is smaller than that of EK-PHD, although a few steps' OSPA is high, like the case at time $k = 18$, however, the performance is better than EK-PHD's as a whole.

From the experimental results above, no matter what the target moving model is straight line or making turns, GH-PHD filter can both obtain higher filter accuracy than EK-PHD's. At the same time, GH-PHD filter needs to calculate numerical integrations and the corresponding weight for each Gaussian item, its calculation is inevitably bigger than that of EK-PHD filter. One simulation experiment shows that GH-PHD filter

consumes 10.0900 seconds and EK-PHD filter consumes 0.9521 seconds.

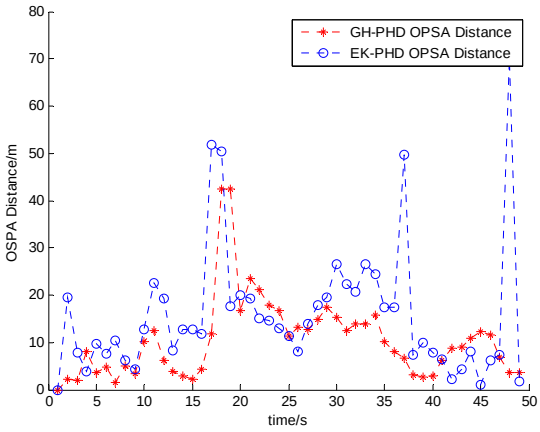


Fig 4 OSPA distance

VII. Conclusion

Aiming at the probability hypothesis density filter based on the theory of random finite sets in the multi-target tracking problem, this paper combines Gaussian-Hermite numerical integral with Gaussian mixture PHD filter, and one algorithm to deal with nonlinear Gaussian system is presented, i.e., GH-PHD filter. In this approach, the joint distribution of targets and the number of targets can be well estimated, even if the target number is unknown or changing with time. It has good satisfied accuracy in multi-target tracking system. The proposed algorithm is suitable for the nonlinear clutter environment, and it breaks through the limitation of GM-PHD which is only suitable for linear system. For solving the multi-target tracking under a nonlinear system, the new algorithm provides a new implemented approach. The new algorithm calculates the integral points and weights of every Gaussian items in PHD recursion so that the computation of new algorithm is larger than that of EK-PHD filter. But the computational complexity still keeps in the same order of magnitude, so it is acceptable in many engineering applications.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (61201118), the China Postdoctoral Science Foundation (2013M532020), the Scientific Research Program Funded by Shaanxi Provincial Education Department (12JK0529), and the National Training Programs of Innovation and Entrepreneurship for Undergraduates (201310709006).

References

[1] Xintao Xia, Leilei Gao, Jianfeng Chen, “Fusion method for true value estimation based on information poor theory”, Journal of Software, 7(5), pp. 1014-1021, 2012.
[2] Mahler R.P.S, “Statistical Multisource Multitarget Information Fusion”, Artech House Publishers, 2007.
[3] Mahler R.P.S, “Multitarget Bayes filtering via first-order multitarget moments”, IEEE Transactions on Aerospace and Electronic Systems, 39(4), pp. 1152-1178, 2003.
[4] Ba-Ngu Vo, Wing-Kin Ma, “The Gaussian mixture probability hypothesis density filter”, IEEE Transactions on Signal Processing, 54(11), pp. 4091-4104, 2006.
[5] Clark D, Ba-Ngu Vo, “Convergence analysis of the Gaussian mixture PHD filter”, IEEE Transactions on Signal Processing, 55(4), pp. 1204-1212, 2007.
[6] Ba-Ngu Vo, Singh S, Doucet A, “Sequential Monte Carlo methods for multi-target filtering with random finite sets”, IEEE Transactions on Aerospace and Electronic Systems, 41(4), pp. 1224-1245, 2005.
[7] Junying Meng, Jiaomin Liu, Juan Wang, et al, “Target tracking based on optimized particle filter algorithm”, Journal of Software, 8(5), pp. 1140-1144, 2013.
[8] Daniel Edward Clark, Judith Bell, “Convergence results for the particle PHD filter”, IEEE Transactions on Signal Processing, 54(7), pp. 2652-2661, 2006.
[9] Fan-bin Meng, Yan-ling Hao, Chong-meng Zhang, et al, “Sequential fusion algorithm based on unscented particle probability hypothesis density filter”, Systems Engineering and Electronics, 33(1), pp. 30-34, 2011 (in Chinese).
[10] Li-ming Chen, Zhe Chen, Fu-liang Yin, et al, “Central difference Kalman-probability hypothesis density filter for multi-target tracking”, Control and Decision, 28(1), pp. 36-42+48, 2013 (in Chinese).
[11] Arasaratnam I, Haykin S, “Cubature Kalman filters”, IEEE Transactions on Automatic Control, 54(6), pp. 1254-1269, 2009.
[12] Pin Wang, Wei-xin Xie, Zong-xiang Liu, et al, “A novel Gaussian Mixture PHD filter for nonlinear models”, Acta Electronica Sinica, 40(8), pp. 1597-1602, 2012 (in Chinese).
[13] Arasaratnam I, Haykin S, “Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature”, Proceedings of the IEEE, 95(5), pp. 953-977, 2007.
[14] Challa S, Bar-Shalom Y, Krishnamurthy V, “Nonlinear filtering via generalized edgeworth series and Gauss–Hermite quadrature”, IEEE Transactions on Signal Processing, 48(6), pp. 1816-1820, 2000.
[15] Weifeng Liu, Chongzhao Han, Feng Lian, et al, “Multitarget state extraction for the PHD filter using MCMC approach”, IEEE Transactions on Aerospace and Electronic Systems, 46(2), pp. 864-883, 2010.
[16] Schuhmacher D, Ba-Tuong Vo, Ba-Ngu Vo, “A consistent metric for performance evaluation of multi-object filters”, IEEE Transactions on Signal Processing, 56(8), pp. 3447-3457, 2008.

Dr. Jinguang Chen (1977-): He was born in Henan Province, China. He obtained Bachelor and Master degree from the Xi’an Polytechnic University in 2000 and 2005, respectively. In 2011, He graduated from School of Electronic Engineering, Xidian University, and obtained PHD. His interests include information fusion, target tracking, etc. Now, he is an associate Professor of Xi’an Polytechnic University.

Clustering Unsynchronized Time Series Subsequences with Phase Shift Weighted Spherical *k*-means Algorithm

Tiantian Yang
School of Digital Media, Jiangnan University, Wuxi, China
Email: taliayoung@163.com

Jun Wang
School of Digital Media, Jiangnan University, Wuxi, China
Email: wangjun_sytu@hotmail.com

***Abstract*—Time series have become an important class of temporal data objects in our daily life while clustering analysis is an effective tool in the fields of data mining. However, the validity of clustering time series subsequences has been thrown into doubts recently by Keogh et al. In this work, we review this problem and propose the phase shift weighted spherical *k*-means algorithm (PS-WSKM in abbreviation) for clustering unsynchronized time series. In PS-WSKM, the phase shift procedure is introduced into the clustering process so that the phase problem is solved effectively. Meanwhile, the subsequences weights are assigned to subsequences to make the algorithm more robust. Experimental results on ECG datasets show that our approach is effective for the problem of unsynchronized time series subsequences clustering, which makes contributions to a wide range of applications, particularly in intelligent healthcare.**

***Index Terms*—time series clustering, unsynchronized time series subsequences, phase shift weighted spherical *k*-means algorithm**

I. Introduction

Time series data is an important kind of temporal data, which has initiated various research and development attempts in the fields of data mining. Clustering is one of the most frequently used methods in the fields of machine learning [1-4, 11-16]. Recently, time series clustering has aroused great interest among researchers. However, Keogh et al. declared that clustering time series subsequence is meaningless [5]. In [5], E. Keogh et al. conducted several clustering experiments with some of the commonly used clustering algorithms, such as k-means, hierarchical, EM, SOMs and other variants of k-Means, and found that the center subsequences obtained by the clustering algorithms are seriously distorted. This work invalidated the contributions of dozens of previously published papers. The reason for this phenomenon comes from the fact that the phase problem is not effectively handled in the clustering process and each subsequence has the equal contributions to the center subsequences. As a result, the final center subsequences are seriously distorted and the clustering algorithms lose effectiveness.

In order to further study the problems posed by Keogh and link up the clustering techniques with the time series applications, in this work, we integrate the principle of phase shift into the clustering process and then propose a novel clustering algorithm, i.e., phase shift weighted spherical k-means algorithm (PS-WSKM in abbreviation). We aim at providing an effective and robust approach for the problem of clustering times series subsequences.

The rest of the paper is organized as follows. In section 2, we discuss the principle of unit vectors. In section 3, we propose the phase shift weighted spherical k-means algorithm PS-WSKM and investigate its properties. Section 3 reports the experimental results.

II. Properties of Unit Vectors

A. Definitions

Definition 1: Subsequence: Given a time series T of length m, a subsequence C of T is a sampling of length $s \le m$ of contiguous position from T, that is, $C = t_p, \ldots, t_{p+n-1}$ for $1 \le p \le m-s+1$.

Definition 2: Optimal phase shift: Given subsequence X and C of length s, the optimal phase shift τ_{opt} of X relative to C is defined as

$$\tau_{opt}(X,C) = \arg\min_{\tau} d\left(C, X^{(\tau)}\right)$$

$$= \arg\min_{\tau} \sqrt{\sum_{i=1}^{s} \left(c_{i+\tau} - x_i\right)^2}$$

τ=1, 2, …, s

in which $X^{(\tau)}$ denotes the subsequence resulting from shifting X with phase shift τ. If $i+\tau > s$, the subsequence wraps around to its end and uses the value at $i+\tau-s$..

B. Some Properties of Unit Vectors

In this section, we will show some important properties of the unit vectors. Suppose we are given n unit vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ in $\boldsymbol{R}^s$ and their weights $w_1, \ldots, w_n$, $w_i>0$, $i = 1, 2, \ldots, n$. The weighted mean vector of the unit vectors can be computed as:

doi:10.4304/jcp.9.5.1103-1108

$$\mathbf{m} = \frac{\sum_{i=1}^{n} w_i \mathbf{x}_i}{\sum_{i=1}^{n} w_i} \quad (1)$$

Note that the weights of $\mathbf{x}_i$'s should be nonnegative real numbers. In addition, the mean vector **m** need not have a unit norm. One may capture its direction via the following calculation:

$$\mathbf{c} = \frac{\mathbf{m}}{\|\mathbf{m}\|} \quad (2)$$

The weighted mean vector **c** with unit norm computed by Eq.(2) may be thought as the vector that is closest in cosine similarity (in an average sense) to all the unit vectors in dataset *X*. This provides us a solid theoretical foundation for the proposed algorithm.

It is a hard work to cluster time series subsequences that are not strictly synchronized and many solutions have been proposed. One straightforward solution is to adjust the phase while the algorithm runs so that the cosine similarity between two subsequences is maximized. This procedure requires finding the optimal phase shift between two subsequences. However, for two subsequences with length *s*, brute force search for the optimal phase shift between them involves $O(s^2)$ computation complexity. This will become the speed bottleneck when the algorithm runs on large datasets. In the following, we will show that, using the convolution theorem, the time complexity to find the optimal phase shift τ_{opt} between **x** and **c** is O(*s*log*s*), which provides us an efficient approach for finding the optimal phase shift.

Let **x** and **y** be two normalized vectors whose length equals to 1 in the Euclidean space, r_{xy} be the consine similarity and d_{xy} be Euclidean distance between **x** and **y**, respectively, then we could have the following relationship:

$$1 - r_{xy} = \frac{1}{2} d_{xy}^{\ 2} \quad (3)$$

On the other hand, the cosine similarity between two subsequences $X^{(\tau)} = \{x_1, x_2, \ldots, x_s\}$ and $C = \{c_1, c_2, \ldots, c_s\}$ can be computed as follows:

$$r_{x,c}(\tau) = \sum_{i=0}^{s} x_i c_{i+\tau} \quad (4)$$

where *s* is the length of the two subsequences. Obviously, the cosine similarity between two subsequences is similar in nature to the convolution of two discrete series. Whereas convolution involves reversing the series, then shifting it and multiplying by another one, the cosine similarity defined in Eq.(4) only involves shifting it and multiplying, without the reversing step.

Theorem 1: Assuming *X* and *C* are subsequences of length *s*, the time complexity to find the optimal phase shift τ_{opt} between *X* and *C* is O(*s*log*s*).

Proof: According to the convolution theorem, under suitable conditions, the Fourier transform of a convolution is the pointwise product of Fourier transforms. Let *F* denotes the Fourier operator and F^{-1} as the inverse Fourier transform, $F\{X\}$ and $F\{C\}$ are the Fourier transforms of time series *X* and *C*, the consine similarity defined in (4) can be computed as follows:

$$r_{x,c}(\tau) = F^{-1}\{F\{X\} \cdot F^{*}\{C\}\}(\tau)$$

in which $F^{*}(C)$ denotes the complex conjugate of the Fourier transform of *C*. With Fast Fourier Transforms (FFT), $r_{x,c}(\tau)$ s for different values of τs can be computed together, thus the time complexity of computing $r_{x,c}(\tau)$ s for different τ values is identical to that of FFT, which is O(*s*log*s*). On the other hand, according to the relationship between consine similarity and Euclidean distance revealed in Eq.(3), the optimal phase τ_{opt} of *X* relative to *C* can be computed with

$$\tau_{opt}(X, C) = \arg\max_{\tau} r_{x,c}(\tau)$$

Thus the Theorem is proved. ∈

The properties discussed above provide us useful theoretical tools to develop an effective clustering algorithm for clustering unsynchronized time series data.

III. Phase Shift Weighted Spherical k-Means Algorithm

Let *X* be the set of subsequences with the length of *s*, $\mathrm{B} = [\beta_1, \beta_2, \ldots, \beta_n]$ be the weight vector, $V = [v_1, \ldots, v_s]$ be the center subsequence, $d(x_i, v_j)$ be the Euclidean distance between subsequence x_i and v_j, and *r* be a parameter for the weight β_i. For the *j*th cluster, the learning criterion can be defined as follows:

$$J_j(X, v_j, \mathrm{B}) = \sum_{i=1}^{n} \frac{1}{\beta_i^r} d^2(x_i, v_j) + \varepsilon \sum_{i=1}^{n} \frac{1}{\beta_i^r} \quad (5a)$$

$$\beta_i > 0, \ \sum_{i=1}^{n} \beta_i = 1 \quad (5b)$$

The penalty term $\varepsilon \sum_{i=1}^{n} \frac{1}{\beta_i^r}$ in Eq.(5) is introduced into the objective function to avoid zero division errors. We can minimize Eq.(5) by iteratively solving the following optimization problems:

Problem P_1: Fix $\mathrm{B} = \hat{\mathrm{B}}$ and $X = \hat{X}$, solve the reduced problem $J(\hat{X}, v, \hat{\mathrm{B}})$;

Problem P_2: Fix $V = \hat{V}$ and $X = \hat{X}$, solve the reduced problem $J(\hat{X}, \hat{v}_j, \mathrm{B})$;

Problem P_3: Fix $\mathrm{B} = \hat{\mathrm{B}}$ and $V = \hat{V}$, solve the reduced problem $J(X, \hat{v}_j, \hat{\mathrm{B}})$.

To solve problem P_1 and find the center that makes Eq.(5) minimized, the cluster center v_j can be computed as follows:

$$v_j = \frac{\sum_{i=1}^{n} \frac{x_i}{\beta_i^r}}{\sum_{i=1}^{n} \frac{1}{\beta_i^r}}, \ l = 1, \ldots, s \quad (6)$$

To solve problem P_2 and find the fuzzy feature weight B that makes the objective function minimized under constraints Eq.(5b), we use Lagrange multipliers. By using Lagrange multipliers, we have

$$\beta_i = \frac{\left(d^2\left(x_i, v_j\right)+\varepsilon\right)^{\frac{1}{r+1}}}{\sum_i^n \left(d^2\left(x_i, v_j\right)+\varepsilon\right)^{\frac{1}{r+1}}} \tag{7}$$

Unlike problem P_1 and problem P_2, problem P_3 is defined on a discrete domain and the function $J(X, v_j, \hat{\mathrm{B}})$ is uncontinuous. Thus the partial derivatives cannot be used here. However, we can shift the phase for each subsequence x_i so that $d^2\left(x_i, \hat{v}_j\right)$ is minimized. In this way, the phase problem involved in time series data is solved in this step. For fixed $\hat{\mathrm{B}}$ and $\hat{v}_j$,

$$J_j(\hat{X}, \hat{v}_j, \mathrm{B}) = \sum_{i=1}^{n} \frac{1}{\hat{\beta}_i^r}\left(d^2\left(x_i, \hat{v}_j\right)+\varepsilon\right) \tag{8}$$

is also minimized.

In the following, we will extend the above procedure to the case of multiple clusters. Let $W = [w_1, w_2, \ldots, w_k]$ be the weights of each cluster, $\mathrm{B} = [\beta_{ji}]_{c\times n}$ be the weights of each subsequence in each cluster, V be center subsequences of clusters, the objective function can be formulated as follows:

$$\begin{aligned} Q_U(X, V, W, \mathrm{B}) &= \sum_{j=1}^{k} \frac{1}{w_j^{\alpha}} J_j\left(X, v_j, \mathrm{B}\right) \\ &= \sum_{j=1}^{k} \frac{1}{w_j^{\alpha}} \sum_{i\in C_j} \frac{1}{\beta_{ji}^r}\left(d^2\left(x_i, v_j\right)+\varepsilon\right) \end{aligned} \tag{9a}$$

subject to

$$\sum_{i\in C_j} \beta_{ji} = 1, \beta_{ji} > 0,\ j = 1, 2, \ldots, k \tag{9b}$$

$$w_j > 0,\ \sum_{j=1}^{k} w_j = 1 \tag{9c}$$

It can be illustrated as the combination of several clusters with objective functions Eq.(5) weighted by $\frac{1}{w_j^{\alpha}}$. Its minimization implies that the weighted quadratic sum of the within-cluster distance should be minimized. We have the following iteration equations:

$$\beta_{ji} = \begin{cases} \dfrac{\left(\frac{1}{w_j^{\alpha}}\left(d^2\left(x_i, v_j\right)+\varepsilon\right)\right)^{\frac{1}{r+1}}}{\sum_{i\in C_j}\left(\frac{1}{w_j^{\alpha}}\left(d^2\left(x_i, v_j\right)+\varepsilon\right)\right)^{\frac{1}{r+1}}} & i \in C_j, \\ 0 & i \notin C_j \end{cases} \tag{10a}$$

$j = 1, 2, \ldots, k$

$$w_j = \frac{\left(\sum_{i\in C_j} \frac{1}{\beta_{ji}^r}\left(d^2\left(x_i, v_j\right)+\varepsilon\right)\right)^{\frac{1}{\alpha+1}}}{\sum_{j=1}^{k}\left(\sum_{i\in C_j} \frac{1}{\beta_{ji}^r}\left(d^2\left(x_i, v_j\right)+\varepsilon\right)\right)^{\frac{1}{\alpha+1}}},\ j = 1, 2, \ldots, k \tag{10b}$$

$$v_j = \frac{\sum_{i\in C_j} \frac{1}{\beta_{ji}^r} x_i}{\sum_{i\in C_j} \frac{1}{\beta_{ji}^r}}, j = 1, 2, \ldots, k \tag{10c}$$

Now we discuss the phase adjustment and cluster assignment procedure. For each subsequence, we find the cluster center which has the maximal similarity with it and the corresponding optimal phase shift τ_{opt} required to produce this maximal similarity. Then we assign the subsequence to this cluster and adjust the phase by the optimal phase shift τ_{opt}.

We state the process of the algorithm PS-WSKM as follows:

Algorithm: PS-WSKM

Input: periodic time series T, period s, cluster number k

Output: the vector of the subsequences weights B, partition matrix U, cluster weights W and the center subsequence V.

Step 1: Segment the time series T with period s to obtain the set of the subsequences $X^{(0)}$.

Step 2: Randomly generate a set of initial cluster weights $W^{(0)}$ and subsequences weights in each cluster $B^{(0)}$. Randomly choose k subsequences as initial centers $V^{(0)}$. Set $t = 0$.

Step 3: Let $\hat{\mathrm{B}} = \mathrm{B}^{(t+1)}$, $\hat{W} = W^{(t+1)}$ and $\hat{V} = V^{(t+1)}$. For each subsequence $X_i^{(t)}$, find the center subsequence $V_l^{(t+1)}$ which has the maximal similarity with it and the corresponding optimal phase shift τ_{opt} required to produce this similarity. Assign the subsequence to this cluster with $u_{li}^{(t+1)} = 1$ and $u_{ji}^{(t+1)} = 0$, for $j = 1, \ldots, k$ and $j \neq l$. Adjust the phase of $X_i^{(t)}$ with $X_i^{(t+1)} = \{X_i^{(t)}\}^{(\tau opt)}$.

Step 4: Let $\hat{X} = X^{(t+1)}$, $\hat{W} = W^{(t)}$ and $\hat{v} = v^{(t)}$, compute $B^{(t+1)}$ using equation Eq.(10a).

Step 5: Let $\hat{W} = W^{(t)}$, $\hat{\mathrm{B}} = \mathrm{B}^{(t+1)}$ and $\hat{X} = X^{(t+1)}$, compute $v^{(t+1)}$ using equation Eq.(10c).

Step 6: Let $\hat{X} = X^{(t+1)}$, $\hat{\mathrm{B}} = \mathrm{B}^{(t+1)}$ and $\hat{v} = v^{(t+1)}$, compute $W^{(t+1)}$ using equation Eq.(10b).

Step 7: Compute $Q_{U(t+1)}(X^{(t+1)}, v^{(t+1)}, W^{(t+1)}, B^{(t+1)})$ using equation Eq.(10a). If $|\ Q_{U(t+1)}(X^{(t+1)}, v^{(t+1)}, W^{(t+1)}, B^{(t+1)}) - Q_{U(t)}(X^{(t)}, v^{(t)}, W^{(t)}, B^{(t)})| < \varepsilon$, output $(U^{(t+1)}, v^{(t+1)}, W^{(t+1)}, B^{(t+1)})$ and stop. Otherwise, $t = t+1$ and goto Step 3.

The newly introduced weighting vector W is used to determine discords clusters. According to Eq.(10b), the cluster weight w_j reflects the weighted within-cluster distance of cluster j. If most of the subsequences within cluster j are close to their cluster center and the size of cluster j is much smaller than others, the w_j will have

small value. This shows that the cluster with small size and small within-cluster distance is more likely to be a discord cluster.

IV. Experimental Study

Electrocardiograms (ECG) are typical periodic time series we encounter in our daily life. Up to now, there are many test datasets available from Internet. In this work, we test our algorithm using ECG datasets.

In our experiments, we split ECGs into subsequences without overlapping with each other, each of which contained one cycle. Then we normalized the subsequences to make each of them has the length of 1. The resulted subsequences were the input of our algorithm.

A. Robustness Experiments on PS-WSKM

This experiment was conducted to verify the robustness of the proposed algorithm PS-WSKM when the cluster number is set to 1 and compared it with Pk-means (k=1) [6].

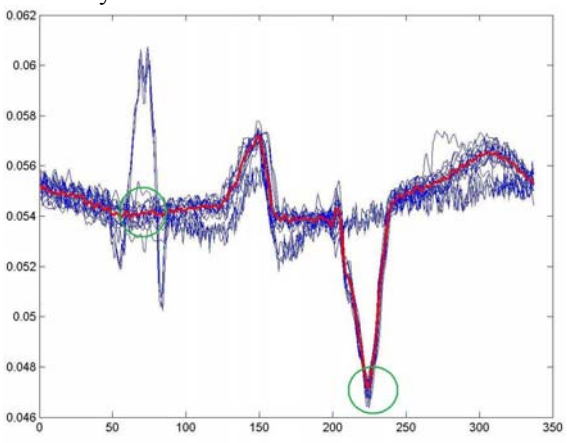
In order to intuitively show how the discords affect the center subsequence found by the algorithm, we created a synthetic dataset by introducing known number of discords into the *xmitdb_x108_0* dataset. In our dataset, the discords were more than 10% percent such that the negative effects of the discords to the whole dataset cannot be ignored. We run the proposed PS-WSKM (k=1) and Pk-means on it respectively and plot the resulting subsequences and the center subsequences in Fig 1(a) and Fig 1(c), respectively. Meanwhile, we also run the both algorithms on the original *xmitdb_x108_0* dataset without discords involved and the results are plotted in Fig 1(b) and Fig 1(d). In Fig 1, the subsequences after phase adjustment by both algorithms are drawn in blue and their center subsequences are highlighted in red. The primary differences between the center subsequences are highlighted by green circles. One point needs to be mentioned here is that the phases of the four resulted center subsequences in Fig 1 are not synchronized. This is due to the random initialization of the algorithms. However, this does not prevent us coming to our conclusions.

Comparing Fig 1(a) with Fig 1(b), we can observe that the center subsequence generated by PS-WSM changes slightly after discords are introduced. This implies that PS-WSKM is robust to the discords in the dataset. In contrast, the center subsequence generated by Pk-means is sensitive to the discords in the dataset when we compare Fig 1(c) with Fig 1(d). The robustness of PS-WSKM comes from the introduction of the weight vector B. Recall the center update equation in Eq.(6) and we can easily infer that the discords with larger β_is have less contribution to the center subsequence computation. However, for Pk-means, each subsequence makes identical contribution to the center subsequence such that the center subsequence is influenced greatly if enough discords are introduced.

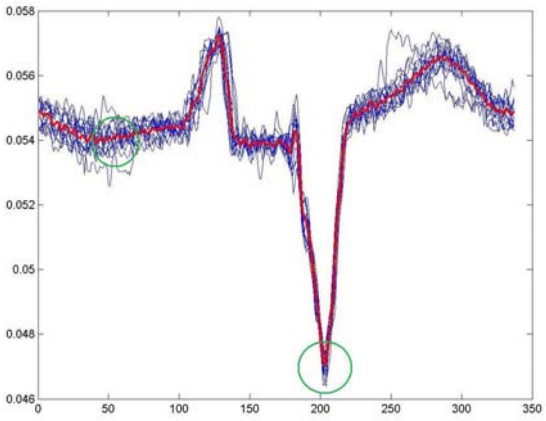
B. Multiple-cluster Datasets Experiments on PS-WSKM

To evaluate the performance of PS-WSKM, we conducted two experiments on the dataset containing multiple clusters.

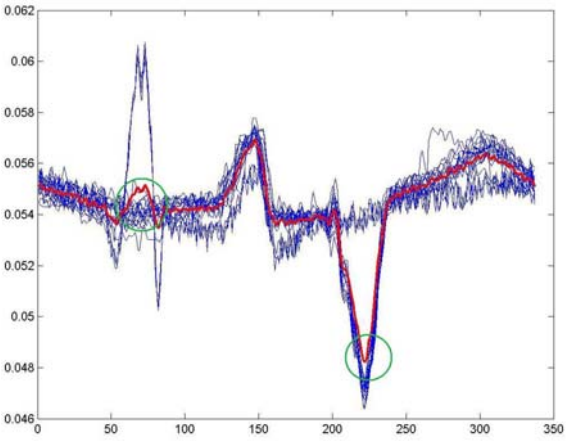
The first part of PS-WSKM experiment was conducted on MIT-BIH datasets labeled 102, 104, 106 and 108. Fig 2 shows the results obtained by PS-WSKM running on these datasets when the optimal k values are properly assumed. As can be seen from Fig 2, PS-WSKM can discover the center subsequence of each cluster successfully, which agree with the clustering results obtained by Pk-means.



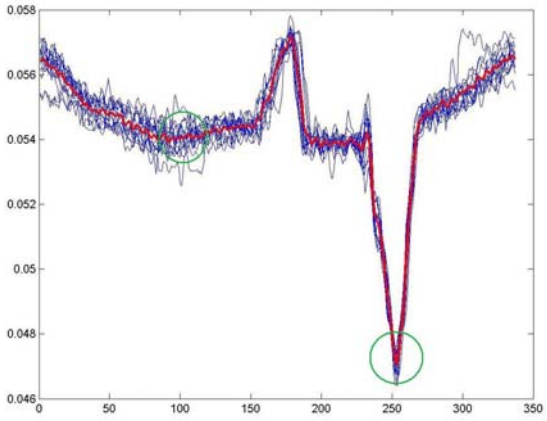
(a)



(b)



(c)



(d)

Fig.1 the center subsequences obtained by different algorithms

(a) center subsequences resulted from PS-WSKM (k=1) running on the dataset with discords introduced (b) center subsequences resulted from PS-WSM running on the original *xmitdb_x108_0* dataset (c) center subsequences resulted from Pk-means running on the dataset with discords introduced (d) center subsequences resulted from Pk-means running on the original *xmitdb_x108_0* dataset

	PS-WSKM (r=1, α=3)			Pk-means
	Cluster weight	Center subsequence of each cluster	Distribution of the subsequences weights	Center subsequence of each cluster
102	0.6994			
	0.3006			
104	0.4084			
	0.5916			
106	0.7177			
	0.2823			
108	0.4363			
	0.5637			

Fig.2 Comparison of PS-WSKM and Pk-means results on several MIT-BIH datasets

IV CONCLUSIONS

The effectiveness of clustering time series subsequences has been thrown into doubts by Keogh et al. recently. In this work, we investigate this problem by introducing subsequences' weights and a phase shift procedure into the clustering process. We proposed PS-WSKM to cluster time series subsequences so that the unsynchronized time series subsequences can be clustered effectively.

On one hand, this work has opened up new opportunities for applying clustering techniques to the unsynchronized time series subsequences clustering tasks. On the other hand, the introduction of subsequences weights makes the algorithm more robust. However, we only use the raw time series data. In our future work, we will integrate the dimensionality reduction or feature extraction techniques into the clustering process. In this way, the learning algorithm will become more smart and effective.

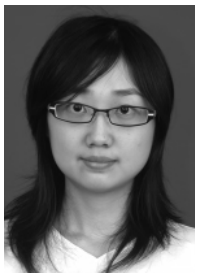
ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grants 61300151 and by the Natural Science Foundation of Jiangsu Province under Grant BK20130155.

REFERENCES

[1] T. W. Liao. Clustering of time series data—a survey, Pattern Recognition, vol.38, no.11, 2005, pp.1857-1874
[2] X. Zhang, J. Liu, Y. Du, et al. A novel clustering method on time series data, Expert Systems with Applications, vol. 38, no. 9, 2011, pp. 11891-11900.
[3] Pierpaolo D'Urso and Elizabeth Ann Maharaj. Autocorrelation-based fuzzy clustering of time series, Fuzzy Sets and Systems, vol.160, no.24, 2009, pp.3565-3589.
[4] E. A. Maharaj and P. D'Urso. Fuzzy clustering of time series in the frequency domain. Information Sciences, vol.181, no.7, 2011, pp.1187-1211
[5] Eamonn Keogh and Jessica Lin. Clustering of time-series subsequences is meaningless: implications for previous and future research, Knowledge and Information Systems, vol.8, no. 2, 2005, pp 154-177.
[6] Umaa Rebbapragada, Pavlos Protopapas and Carla E. Brodley, et al. Finding anomalous periodic time series. Machine Learning, vol.74, no.3, 2009, pp. 281-313.
[7] E. Keogh, S. Lonardi, and B. Y. Chiu. Finding Surprising Patterns in a Time Series Database in Linear Time and Space. In Proceedings of the Fifth IEEE International Conference on Data Mining, pp 226-233, 2005
[8] MIT-BIH Arrhythmia Database Directory, http://www.physionet.org/physiobank/database/html/mitdbdir/intro.htm. 2010.
[9] L. Xu, A. Krzyzak, E. Oja. Rival Penalized Competitive learning for clustering analysis, RBF net and curve detection. IEEE Transactions on Neural Networks, vol. 4, no. 4, pp. 636-649, 1993
[10] Z. Bar-Joseph, G. Gerber, D. Gifford, et al. A new approach to analyzing gene expression time series data. In proceedings of the 6th Annual International Conference on Research in Computational Maolecular Biology. Washington, D.C., 2002, pp.39-48.

[11] J. Xie, Shuai Jiang, Weixin Xie, Xinbo Gao. An Efficient Global K-means Clustering Algorithm, Journal of Computers, vol. 6, no. 2, 2011, pp. 271-279
[12] J. Wu, Jie Xia, Jian-ming Chen, Zhi-ming Cui. Moving Object Classification Method Based on SOM and K-means, Journal of Computers, vol.6, no. 8, 2011, pp.1654-1661
[13] T. Li, Yan Chen, Jinsong Zhang. Logistics Service Provider Segmentation Based on Improved FCM Clustering for Mixed Data, Journal of Computers, vol.7, no.11, 2012, pp.2629-2633
[14] Hongfen Jiang, Junfeng Gu, Yijun Liu, et al. Study of Clustering Algorithm based on Fuzzy C-Means and Immunological Partheno Genetic, Journal of Software, vol. 8, no. 1, 2013, pp.134-141.
[15] Hongjie Jia, Shifei Ding, Hong Zhu, et al. A Feature Weighted Spectral Clustering Algorithm Based on Knowledge Entropy. Journal of Software, vol.8, no.5, 2013, pp.1101-1108.
[16] Jiashun Chen, Dechang Pi, Zhipeng Liu. An Insensitivity Fuzzy C-means Clustering Algorithm Based on Penalty Factor, vol.8, no.9, 2013, pp.2379-2384.



Tiantian Yang received her Ph.D. degree from Shanghai Institute of Microsystem and Information Technology. She is currently a lecture at the School of Digital Media, Jiangnan University. Her research interests include pattern recognition, digital image processing and MEMS.

Jun Wang received his Ph.D. degree from the Nanjing University of Science and Technology in January, 2011. He is currently an associate professor at the School of Digital Media, Jiangnan University. He has published nearly 20 papers in international/national authoritative journals. His research interests include pattern recognition, data mining and digital image processing.

SVM-based Automatic Annotation of Multiple Sequence Alignments

Jiansi Ren*
School of Computer Science, China University of Geosciences, Wuhan, China
Email: renjiansi@hotmail.com

***Abstract*—Multiple Sequence alignments are a critical step in phylogeny inference. There is a lack of an appropriate approach which is capable of 1) finding the best global alignment and 2) automating and reproducing manual editing. Progressive alignment is an effective method for multiple Sequence alignments. However, its application in practice has also long been largely hampered because the alignment regions are not homologous to maximize the alignment score. The standard practice in phylogenetics involves manual editing of alignments and manual editing is a non-trivial task. Aiming at these problems, this study 1) uses SVM to capture the neighborhood of a site to automate and reproduce manual editing, and 2) builds the procedure of SVM Model Training and Automatic Annotation. Experimental results demonstrate that a SVM-based classifier can reproduce the manual editing tasks with an accuracy of 95.5%. This method is stable to both RBF parameters (Gamma and C) and clearly outperforms GBLOCKS and AL2CO, which are conventional editing/annotating methods. The classification accuracy achieved by the proposed method is always much higher than those achieved by the counterpart methods.**

***Index Terms*—Multiple Sequence Alignments, machine learning, automatic annotation**

I. Introduction

Sequence phylogeny is used by biologists to reconstruct the series of events that have led to the distribution and diversity of life. Evolutionary patterns can be found by aligning the sequence of bio-molecules such that homologous positions are aligned into columns.

Biological sequences are supplemented with previously released sequences that are collected from databases using algorithms such as BLAST [1]. Sequence alignments are a critical step in phylogeny inference. However, finding the best global alignment is a computationally complex operation. If a region of the global alignment isn't properly aligned, the phylogeny reconstruction will attempt to accommodate the erroneously alignment character into the phylogeny, leading to a decrease in resolution. Obtaining biologically accurate alignments is a challenge, as the best methods sometimes fail to align readily apparent conserved motifs [2]. The exhaustive solution has the order of $O(n^k)$ where n is the length of the longest sequence and k is the number of sequences, a prohibitive constraint with only a few sequences [3]. Heuristics have been developed, the most famous of which is probably the progressive alignment method [4].

The progressive alignment algorithm is based on the idea that sequences to be aligned are phylogenetically related and these relationships are used to guide the alignment. Using this approach a tree is inferred by performing alignments [5] between each possible pairs of sequences. The distance between each pairs of sequence is computed as the number of mismatched positions in an alignment divided by the total number of matched position. A neighbor joining [6] "guide tree" is generated from these pair-wise distances, which gives the order of the generation of progressive alignment. The alignment continues with each step treated as a pair-wise alignment between a cluster and the next closest sequence. Gaps are added to an existing multiple sequence alignment and a gap will always be a gap.

A penalty is incurred by introducing and extending a gap. For a linear gap penalty this amounts to scoring each column of the alignment by the sum of the amino acid pair scores in this column. The corresponding score is called the sum of pairs (SP) score [7]. Although progressive alignment enjoys immense popularity and is used in multiple alignment programs like ClustalW [8], it has some weaknesses such as it will attempt to align regions that are not homologous to maximize the alignment score. Furthermore there is no ultimate way of quantifying whether or not the alignment is good.

The standard practice in phylogenetics involves some level of manual editing of alignments. The whole process of manual editing is a time consuming and a non-trivial task. Our aim is to automate and reproduce manual editing using artificial intelligence. The method of choice in this study is neural networks, although we have tested a selection of alternative strategies in the past [9]. In previous work [9], decision tree induction (C4.5), Naïve Bayes, and support vector machine methods were applied to the same dataset. There was no clear winner among the different approaches. SVM [10-13] recorded high precision for the classification of inadequate sites where as for the prediction of valid sites C4.5 was the best. Because the manual editing process often considers the neighborhood of a site, we finally chose to use SVM to capture this important factor.

Reproducing the manual editing of multiple sequence alignment has two aims: 1) to automate the process to improve the quality of the input data for large-scale

doi:10.4304/jcp.9.5.1109-1116

phylogenetic studies, and 2) to improve the repeatability of the process of editing. We believe that this process may outperform manual editing because it also considers the general phylogenetic structure of the data by using site likelihood computed on a preliminary tree.

The rest of this paper is organized as follows. Section II briefs related work in uses of methods for alignment editing. Section III introduces the methods and the process of implementation. This section also details an application of this system. System accuracy and stability are presented in Section IV. This section also provides a performance evaluation by comparing the approach with the existing editing tools. Section V concludes the paper with a summary.

II. Related Work

Other work have been done on alignment editing: GBLOCKS [14], AL2CO [15] are a few software implementations of alignment editing programs. GBLOCKS is a program that is designed to take as input a multiple protein sequence alignment and perform editing to produce a similarly formatted output with the putative "inadequate" sites removed. GBLOCKS claims to be based on the improvement of phylogenetic results and takes into account homology rather than sequence similarity. While GBLOCKS can function as an alignment editing program and was shown to yield improved results for phylogenetic analysis [14], it is not the one that emulates the manual editing process. The criterion was chosen by the user indicating the amount of variability that will be tolerated at a site. This approach effectively removes columns corresponding to the highest site rates with the argument that they contain multiple hidden substitutions and are then ill-suited for phylogenetic analysis. However, these fast-evolving sites may contain valuable phylogenetic signal to resolve closely related sequences. In the AL2CO implementation, the concept of conservation index was introduced and recommended for use as a parameter for refinement of multiple sequence alignment [15].

Our method was trained on multiple sequence alignments extracted from the PFAM database [16]. About 13,000 sites were classified as valid, inadequate or ambiguous. The latter class was used in the design in hope that the classifier could perceive elements in the alignment that are not obvious to the human eye. Using this annotated corpus, training and testing were performed to create an automated annotator of multiple sequence alignment that can be used for editing.

III. Methods

This section first addresses the dataset and parameterization. After that, the procedure of SVM Model Training and Automatic Annotation is proposed .

A. Dataset

Thirty-six multiple sequence alignments of protein domains were arbitrarily retrieved from PFAM [16], a database containing a collection of multiple alignments of protein domains or conserved protein regions. A total of about 13,000 sites of multiple sequence alignments were manually annotated by the authors. Two classes were identified during manual annotation, inadequate and valid sites. Sites were classified as valid where there was evidence that the variability in residue identity within the site was solely due to a substitution process occurring over time. Inadequate sites appeared to be the results of alignment artifacts or contain gap characters for most sequences in the alignment. The natural distribution of the data set is 23%-77% inadequate valid.

B. Parameterization

Five parameters were gathered from the multiple sequences. The first parameter derived from the alignments is called gap ratio g. For each site, we use N-gram analysis (default size=3) and the gap ratio of a site is calculated by dividing the number of N-grams (C) that contain gap characters (-) by the total number of N-grams in the given site (T). Thus, the following equation is used to find the gap ratio.

$$g = \frac{C}{T} \tag{1}$$

The possible values of gap ratio, then, lie between 0.00, where none of the sequence in a column have a gap at the site, and 1.00, when all sequences in the column contain a gap character.

The Normalized Site Likelihood Ratio (NSLR) is the site log likelihood (log (l)) considering the data in a column of the alignment, the JTT substitution model [17] and a Neighbor joining tree created from an unedited alignment, minus the site log likelihood (assuming that all sequence are unrelated) of base states picked at random from a set of residues frequencies in the JTT evolutionary model: log (r) normalized by the number of sequences in the alignment without a gap at that position ((1-g)*t).

$$NSLR = \frac{\log(l) - \log(r)}{(1-g)\times t} \tag{2}$$

Where, l = site for a given a preliminary tree, r = the site likelihood if the sequences were unrelated (i.e. independent, or random), g = gap ratio, and t = number of sequences. The value of the normalized site likelihood ratio is not bounded, except by zero as a minimum.

Third, parsimony count (PC) is the gap to no-gap transitions given a preliminary tree. The Parsimony count means the minimum number of character changes observed on the tree. The parsimony count was calculated by converting each alignment column into a binary vector (gap/no-gap character). NJ tree was used as a guide to count PC. Site rate, the fourth parameter, is the measure of the rate of evolution at a site relative to other sites in the alignment.

The 4th parameter, site rate, was evaluated using the NJ tree of the unedited alignment, the JTT model and the libcov library [18]. The alpha parameter was estimated from the data.

The 5th parameter is the Normalized Similarity Score (NSS), which was calculated based on N-gram analysis

using the CNG formula [19] for calculating similarity score as follows:

$$\text{Similarity}= \sum_{g\cup D1\cup D2}\left(\frac{2\cdot(f_1(g)-f_2(g))}{f_1(g)+f_2(g)}\right)^2 \quad (3)$$

where $f_i(g)$=0 if $g \notin D2$. Once we got the similarity score, we have NSS=Similarity/T, where T is the total number of N-grams for the given site. Default size used for N-gram analysis is 3. For each site, only the next contiguous site was selected to do the similarity analysis.

After parameterization, we have the following output for every multiple sequence alignment and here is an example (Table I):

TABLE I.
OUTPUT OF A SAMPLE PARAMETERIZATION

class variable	gap ratio	NSLR	PC	site rates	NSS
1	0.692308	0.517516	0	1.58326	7.3333333
1	0.692308	0.590195	0	2.53675	6.6666667
1	0.769231	0.555369	0.1603	1.19501	4.2314815
1	0.307692	0.649415	0	2.53675	4.0012472
1	0	0.656452	0	2.53675	5.1067823
1	0	0.546849	0.3544	0.927187	2.6718751
0	0	0.53055	0	1.58326	4.3541667
0	0.692308	0	0	2.53675	1.3333333
0	0.692308	0	0.7143	2.53675	2.88
1	0	0.522676	0	0.71405	3.037037
1	0	0.678862	0	0.160314	5.666667

Alignment file used here is a2m.ann.fta. The first column is the class variable, where 1 indicates a valid site and 0 indicates an invalid site. The 2nd column is gap ratio, the 3rd column is Normalized Site Likelihood Ratio (NSLR), the 4th column is Parsimony Count (PC), the 5th column is the site rates and the last column is the Normalized Similarity Score (NSS).

C. SVM-based Implementation

1) LibSVM package

LibSVM package was employed to build our application. It is an integrated software stack for supporting vector classification, regression and distribution estimation. This package includes source code written in different languages such as C++, Java, Python, Perl, and Matlab and so on.

2) SVM model training

Before doing classification or annotation for a multiple sequence alignment, we need to obtain a model by training our machine learning system [20-22]. The procedure of training a model was illustrated in Fig 1.

In the first place, the training data set was prepared by manually annotating a set of sites and then doing parameterization. The output of our parameterization is then transformed to a standard format (Table II) to feed the SVM application so that we can get a trained model and save it for later use (classification or annotation).

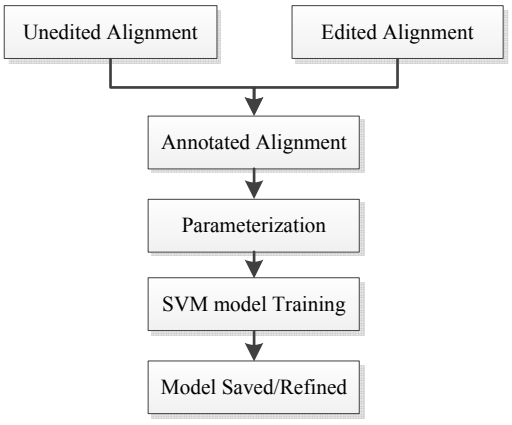


Figure 1. Procedure of SVM Model Training.

Annotated alignments initially used for model training could be prepared by end user by simply providing a pair of unedited and edited alignment, and then this system will generate the annotated alignment. In this regard, the SVM model could be refined.

TABLE II.
STANDARD FORMAT OF PARAMETERIZATION FOR LIBSVM

class variable	gap ratio	NSLR	PC	site rates	NSS
1	1:0.692308	2:0.517516	3:0	4:1.58326	5:7.3333333
1	1:0.692308	2:0.590195	3:0	4:2.53675	5:6.6666667
1	1:0.769231	2:0.555369	3:0.1603	4:1.19501	5:4.2314815
1	1:0.307692	2:0.649415	3:0	4:2.53675	5:4.0012472
1	1:0	2:0.546849	3:0.3544	4:0.927187	5:2.6718751
0	1:0	2:0.53055	3:0	4:1.58326	5:4.3541667
0	1:0.692308	2:0	3:0	4:2.53675	5:1.3333333
0	1:0.692308	2:0	3:0.7143	4:2.53675	5:2.88
1	1:0	2:0.522676	3:0	4:0.71405	5:3.037037
1	1:0	2:0.678862	3:0	4:0.160314	5:5.666667

The only difference between this format and the original output of our parameterization is that for each parameter, there is a serial number followed by a colon preceding it.

3) Procedure of Automatic Annotation

Once we have the saved training model, we can build our automatic annotator easily. The procedure of automatic annotation is illustrated in Fig 2. Given an unedited alignment, we did the same process of calculations to get those 5 parameters and fed them to the SVM classifier based on libsvm and loaded the saved model got from the previous step, and then the classifier can predict the classification label for each site, therefore the final annotated alignment or edited alignment will be obtained. An annotation process is to classify each site in an alignment and labelled with its class variable (For instance, "X" indicates a valid site while "C" indicates an invalid site). Here is the example of annotation process as shown in Fig 3.

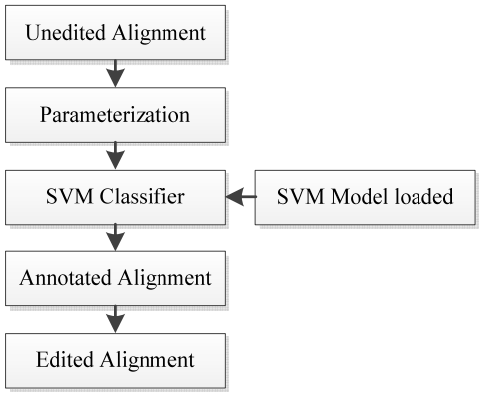


Figure 2. Procedure of Automatic Annotation.

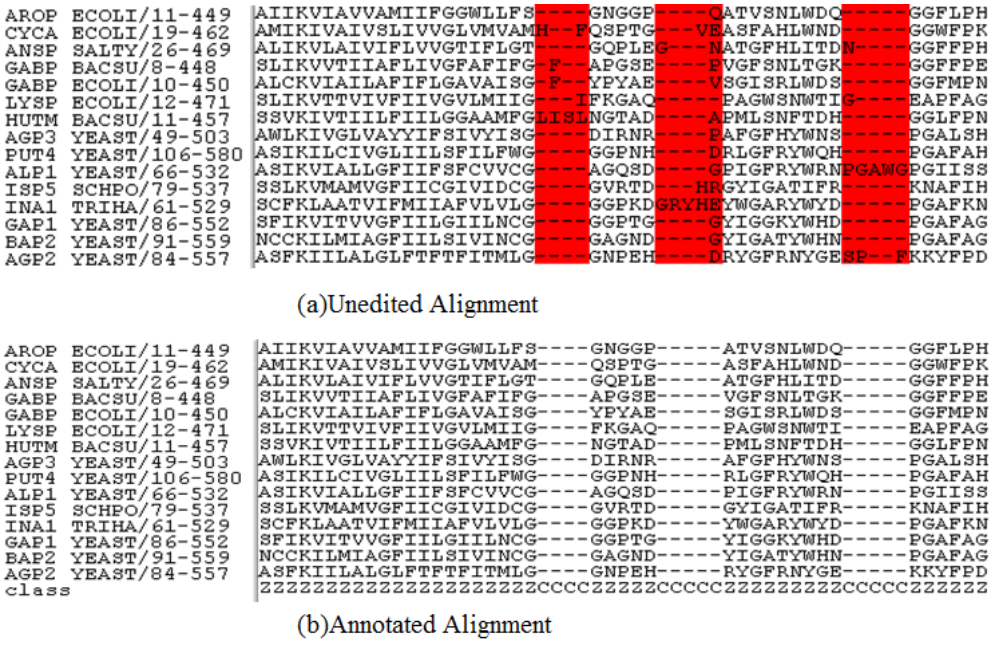


Figure 3. Example of alignment annotation and editing.

"C" in class row indicates the corresponding site is invalid while "Z" stands for a valid site. The only difference between an annotated alignment and an edited alignment is that all sites that contain only gap characters which are marked as "C" will be removed from the alignment file.

IV. Results

System accuracy and stability were tested in this section.Experimental results were evaluation by comparison with existing editing tools.

A. System Accuracy

The accuracy of automatic annotation (or site classification) of this system is 95.5% by using 10-fold cross validation testing on the current data set (about 13,000 sites in total).

B. System Stability

This system was tested by using different combinations of SVM parameters (C, gamma, and Kernel Type and so on). The final Kernel type was used is radical basis function (RBF), since the system performs the best with this kernel type. For RBF kernel, a set of C and gamma were tested and here is the output(Table III, Table IV):

TABLE III.
OUTPUT OF C

C	accuracy
1	95.0542%
2	95.1097%
5	95.0859%
10	95.1493%
20	95.1810%
25	95.1889%
50	95.2523%
100	95.2602%
200	95.2840%
500	95.3077%
1000	95.3791%
5000	95.5613%
50000	95.5692%

TABLE IV.
OUTPUT OF GAMMA

Gamma	Accuracy
1	95.3846%
5	95.6151%
10	95.6943%
20	95.5574%
50	95.579%
100	95.5358%
200	95.4638%

As we can see from the result above, this system is stable for both C and Gamma parameters, since the system accuracy is always around 95.5% with different choices of C or Gamma.

C. Evaluation by Comparison with Existing Editing Tools

As mentioned previously in the introduction section, some existing tools were frequently used to do multiple sequence alignment editing in phylogenetical analysis such as GBLOCKS [14] and AL2CO [15]. To evaluate the performance of our system, the best way is to do comparisons between our application and those existing tools.

1) Comparison with GBLOCKS

Annotation performed by GBLOCKS is to take as input a multiple protein sequence alignment and perform editing to produce a similarly formatted output with the putative "inadequate" sites removed, where valid sites were marked by blue blocks, the rest part of the alignment is considered as inadequate sites as illustrated in Fig 4.

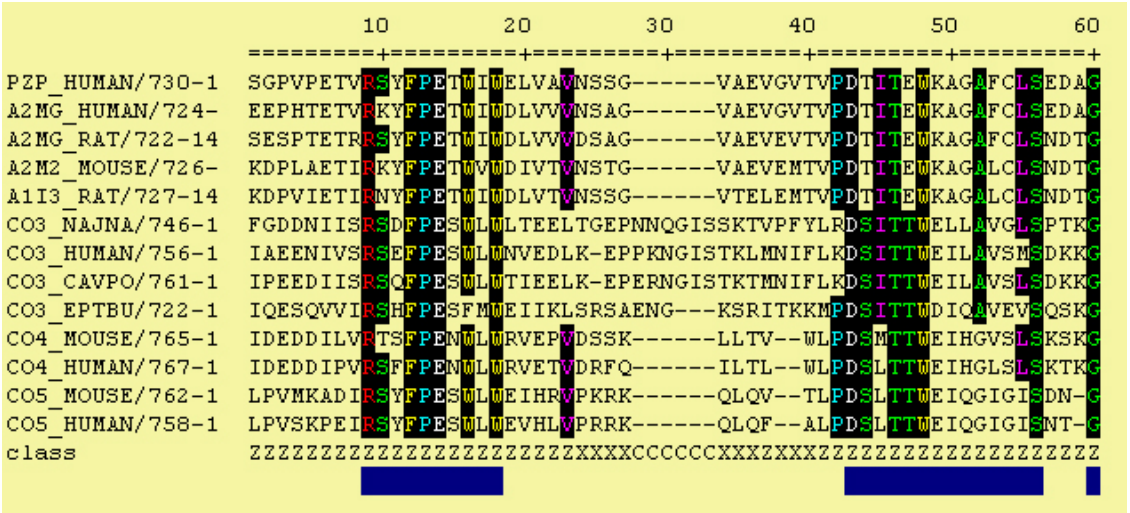


Figure 4. Annotation by GBLOCKS

The alignment used here is a2m.ann.fta. The last row in the alignment is our manually annotated class labels (training data) for each site. Blue blocks are the output of GBLOCKS which indicates the corresponding sites are valid sites.

Using GBLOCKS as a classifier, we calculated its accuracy for site classification and here is the result of comparison (Table V). Obviously, our system shows higher accuracies in comparison with GBLOCKS.

TABLE V.
COMPARISON BETWEEN OUR SYSTEMS TO GBLOCKS

Alignment	This work	GBLOCKS
a2m.ann	87.3%	39.7%
malic.ann	91.5%	30.6%
MotA_ExbB.ann	90.7%	23.9%
aa_permease.ann	91.2%	44.9%
bac_export_1.ann	94.7%	21.6%
bunya_g1.ann	97.6%	94.56%
rubisco_large.ann	96.4%	98.1%

The accuracy for the last alignment is higher than our system. This is because GBLOCKS reserves almost all the sites and considers them as valid.

2) Comparison with AL2CO

In the AL2CO implementation, the concept of conservation index was introduced and recommended for use as a parameter for refinement of multiple sequence alignment. Here is an example of output by AL2CO as shown in Table VI.

TABLE VI.
OUTPUT OF AL2CO

serial number	protein name	conservation score
18	W	2.534
19	E	-0.718
20	L	0.062
21	V	0.032
22	A	-1.578
23	V	1.260
24	N	-0.804
25	S	0.079
26	S	-0.718
27	G	-0.447
28	-	-1.000 *
29	-	-1.000 *
30	-	-1.000 *
31	-	-1.000 *
32	-	-1.000 *
33	-	-1.000 *
34	V	-1.124
35	A	-0.434
36	E	-0.400

Alignment used here is a2m.ann.fta.

Followed the procedure proposed above, we obtained the accuracies for AL2CO as a classifier and here is the result compared to our system (Table VII).

Alignment used here is malic.ann.fta. The first column is the serial number of each site, the second column is the representative protein name for each site and the last column is the conservation score generated by AL2CO.

Since AL2CO didn't classify each site implicitly, we may use an alternative way to make it a classifier. The idea is to divide the output of AL2Co into 2 groups by choosing a threshold. If the conservation score is higher than the threshold, we then considered it as 1 (valid site) otherwise 0 (invalid site). The problem for this idea is that what the best splitter (threshold) will be?

By taking information theory into consideration, we can figure out a reasonable way of solving this problem:

1. Choose each of these conservation scores as thresholds and build a confusion matrix with four values TP (True Positive), TN(True Negative), FP(False Positive) and FN(False Negative).

2. Calculate information gain (IG) for each threshold using the following formula(Equation 4):

$$\mathrm{IG} = -\sum p_i \log p_i \qquad (4)$$

Where $pi = (TP, TN, FP, FN)/(TP + TN + FP + FN)$

3. Find the highest value of IG (Fig 5) which is a relatively easy job to do after the IG distribution plot was generated. Choose the conservation score with highest IG as the best threshold and then calculate the corresponding accuracy for the given alignment and compare to that of our system.

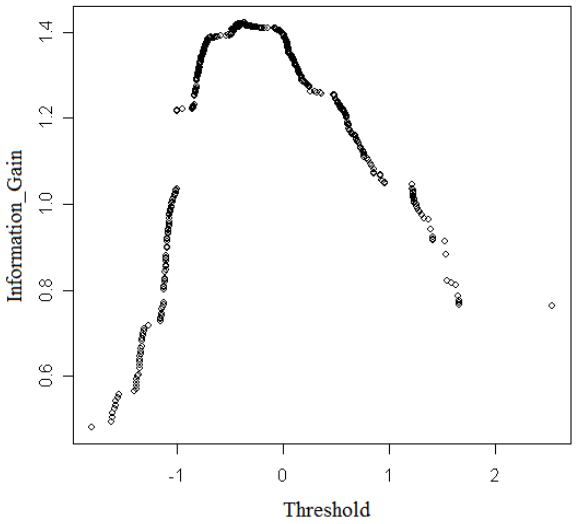


Figure 5. Information Gain values distribution.

TABLE VII.
COMPARISON BETWEEN OUR SYSTEM TO AL2CO

Alignment	This work	AL2CO
a2m.ann	87.3%	55.1%
malic.ann.	91.5%	53.3%
MotA_ExbB.ann	90.7%	52.7%
aa_permease.ann	91.2%	54.6%
bac_export_1.ann	94.7%	51.0%
Bunya_g1.ann	97.6%	50.1%
Rubisco_large.ann	96.4%	50.2%

The result above also shows that our system outperforms AL2CO with higher accuracies.

V. CONCLUSIONS

In order to cater for the needs of Multiple Sequence Alignments, this study explores an approach to 1) automate and reproduce manual editing, and 2) enable efficient and scalable Automatic Annotation. The first issue had been addressed using SVM to capture the neighborhood of a site. The Automatic Annotation problem had been tackled by building the procedure of SVM Model Training and Automatic Annotation. Comparison with existing editing tools had been carried out and revealed this method can facilitate the process of multiple alignments annotation. It is stable for both of RBF parameters (c & gamma). This system outperforms some of the existing annotation methods with higher accuracy. Most importantly, this method allows individual users to refine or redefine the training set used to build the classifier by simply providing example pairs of annotated and original MSA in order to reproduce the editing criteria of individual phylogeneticists. This refine/redefine process does not require any knowledge of SVM-based machine learning classification from the end-user. It provides an ideal tool for Multiple Sequence Alignments.

REFERENCES

[1] Altschul, S.F., Gish, W., Miller, W., Myers, E. and Lipman, D. "A basic local alignment search tool, " J. Mol. Biol., Vol. 215, pp. 403-410, 1990.

[2] Edgar, R.C. "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic Acids Res 32(5), 1792-1797, 2004.

[3] Thompson, J.D., Plewniak, F. and Poch, O. "A comprehensive comparison of multiple sequence alignment programs." Nucleic Acids Research, 27(13), 2682-2690, 1999.

[4] Feng, D. and Doolittle, R. F. "Progressive sequence alignment as a prerequisite to correct phylogenetic trees." J. Mol. Evol. 60, 351-360, 1987.

[5] Needleman, S. B. and Wunsch, C. D. "A general method applicable to the search for similarities in the amino acid sequence of two proteins." J. Mol. Biol. 48, 443-453, 1970.

[6] Swofford, D.L. and OlsenG.J. "Phylogenetic inference." Hillis, D.M., Moritz, C., and Mable, B. (eds.) Molecular Systematics (2nd ed.). 407-514. Sinauer Associates, Sunderland, Massachusetts, 1996.

[7] Gusfield, D. "Algorithms on Strings, Trees, and Sequences." Cambridge University Press, 1997.

[8] Thompson, J.D., Higgons, D.G. and Gibson, T.J. "CLUSTAL W: improving the sensitivity of progressive multiple alignment through sequence weighting, position specific gap penalities and weight matrix choice." Nucleic Acids Research, 22(22), 4673-4680, 1994.

[9] Shan, Y., Milios, E.E., Roger, A.J., Blouin, C. and Susko, E. "Automating Recognition of Regions of Intrinsically Poor Multiple Alignment for Phylogenetic Analysis using Machine Learning." Proceedings of the 2003 IEEE Bioinformatics Conference. 482-483, 2003.

[10] Zhu Fang, Wei Junfang and Shi Wenbo. "SVM Fast Classification Algorithm of Based on Similarity Analysis", International Journal of Digital Content Technology and its Applications(JDCTA), AICIT, vol. 7, no. 2, pp. 10-16, 2013.

[11] Liqin Fu, Haiguang Zhai, Yongmei Zhang and Dan Yu. "Binary tree SVM –based Emotion Recognition from Speech Signal", International Journal of Advancements in Computing Technology(IJACT), AICIT, vol. 5, no. 1, pp. 224-232, 2013.

[12] Zheng Xiaomei. "A Novel Method for Foreign Language Teaching Evaluation Based on Feature Selection", International Journal of Digital Content Technology and its Applications(JDCTA), AICIT, vol. 7, no. 2, pp. 133-140, 2013.

[13] Tao Dongli, Xiao Zhitao, Zhang Fang, Geng Lei and Wu Jun. "Cloth Defect Classification Method Based on SVM", International Journal of Digital Content Technology and its Applications(JDCTA), AICIT, vol. 7, no. 3, pp. 614-622, 2013.

[14] Castresana, J. "Selection of Convserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis." Molecular Biology and Evolution. 17(4), 540-552, 2000.

[15] Pei, J. and Grishin, N.V. "AL2CO: calculation of a positional conservation in a protein sequence alignment." Bioinformatics 17(8), 700-712, 2001.

[16] Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffith-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L.L., Studholme, D.J., Yeats, C. and Eddy, S.R. "The Pfam protein families database." Nucleic Acids Research. 32, D138-141, 2004.

[17] Jones,D.T., Taylor,W.R. and Thornton,J.M. "The rapid generation of mutation data matrices from protein sequences." Comp. Appl. Biosci., 8, 275–282, 1992.

[18] Butt, D., Roger, A.J., and Blouin, C. "libcov: A C++ bioinformatic library to manipulate protein structures, sequence alignments and phylogeny." BMC Bioinformatics, in press, 2005.

[19] Vlado Keselj, Fuchun Peng, Nick Cercone and Calvin Thomas. "N-gram-based Author Profiles for Authorship Attribution" In Proceedings of the Conference Pacific Association for Computational Linguistics, PACLING'03, Dalhousie University, Halifax, Nova Scotia, Canada, 255-264, 2003.

[20] Baldi, P. and S. Brunak. "Bioinformatics: the Machine Learning Approach." MIT Press, 1998.

[21] Kohavi, R. and Provost, F. "Special Issue on Applications of Machine Learning and the Knowledge Discovery Process". Machine Learning, 30, 271-274, 1998.

[22] Mitchell, T.M. "Machine Learning." WCB/McGraw-Hill, 1997.

Jiansi Ren received the B.Sc. degree from Northeast Normal University, Changchun, China, and the M.Ed. degree from Central China Normal University, Wuhan, China. He received the Ph.D. degree from China University of Geosciences, Wuhan, China. He is currently a lecturer in the School of Computer Science at China University of Geosciences, Wuhan, China. His research interests include computer-based modelling and simulation, bioinformatics,and computation intelligence.

GPU Implementation of Parallel Support Vector Machine Algorithm with Applications to Intruder Detection

Xueqin Zhang, Yifeng Zhang
East China University of Science and Technology, Shanghai, China
Email: zxq@ecust.edu.cn, outshine-george@163.com

Chunhua Gu
Shanghai University of Electric Power, Shanghai, China
Email: guchunhua@shiep.edu.cn

***Abstract*—The network anomaly detection technology based on support vector machine (SVM) can efficiently detect unknown attacks or variants of known attacks, however, it cannot be used for detection of large-scale intrusion scenarios due to the demand of computational time. The graphics processing unit (GPU) has the characteristics of multi-threads and powerful parallel processing capability. Based on the system structure and parallel computation framework of GPU, a parallel algorithm of SVM, named GSVM, is proposed. Extensive experiments were carried out onKDD99 and other large-scale datasets, the results showed that GSVM significantly improves the efficiency of intrusion detection, while retaining detection performance.**

***Index Terms*—Network Intrusion Detection, Support Vector Machine, GPU, Parallel Algorithm**

I. INTRODUCTION

[1] Intrusion detection is essentially a problem of classification. The anomaly detection technology based on machine learning can efficiently detect unknown attacks or variants of known attacks on the host or network. Support Vector Machine (SVM) is a powerful statistical learning algorithm based on VC dimension (Vapnik-Chervonenkis Dimension) and the structural risk minimizing theory [1]. SVM has been applied to many classification problems, such as speech reorganization and face gaze detection, with some degree of success [2-3]. However, solving the SVM problem involves a quadratic programming (QP) problem where memory and time complexity increase as a function of l^2, where l is the number of training samples. Consequently, for large-scale network intrusion detection problems, with a rapid increase of l, the SVM training becomes impractical due to huge costs. To reduce the training time of SVM, the original optimization problem of the SVM can be decomposed into a series of sub-problems with heuristic methods such as the Osuna decomposition algorithm [4], the Sequential Minimal Optimization (SMO) algorithm [5], the SVM^{light} algorithm [6], and the LIBSVM algorithm [7].However, a large-size SVM problem still requires a large amount of computational time.

The graphics processing unit (GPU) has a powerful computational capability with its high degree of parallelism. It has been widely used in general-purpose computing in recent years, such as in image and video processing, computational biology, power estimating [8-11]. In order to solve a large-scale network intrusion detection problem, a parallel algorithm of support vector machine, namely GSVM, is presented in this paper. This algorithm is based on the system structure and parallel implementation framework of the GPU. Extensive experiments were carried out on large-scale dataset and the results showed that GSVM can improve the training and testing speed of SVM Dramatically while retaining classification accuracy.

II. BASIC THEORY OF SUPPORT VECTOR MACHINE

A. Support Vector Machine for Classification

SVM seeks an optima hyper plane to separate two-class samples with the maximal margin.

Given training samples: $(x_1, y_1),...,(x_1, y_1), x \in R^d, l$ is the number of training examples, d is the dimension of feature vector, and $y_1$is the class label. SVM derives the weight vector $\boldsymbol{w}$ and offset b by solving the following optimization problem:

$$\min_{\mathbf{w},b,\xi} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{l}\xi_i \qquad (1)$$

$$\text{subject to } y_i(\mathbf{w}^T\phi(x_i) + b) \geq 1 - \xi$$

$$\xi_i \geq 0, i = 1, 2, ..., l .$$

Where function $\phi()$ maps vector x_i from the input space into a high-dimensional feature space, C is the penalty coefficient, and ξ is a relax variable, $\xi_i \geq 0$.

[1]Manuscript received September 29, 2013; revised October 29, 2013; accepted October 30, 2013.
corresponding author: Chunhua Gu (guchunhua@shiep.edu.cn)

doi:10.4304/jcp.9.5.1117-1124

Based on the Wolfe Dual theory, the dual optimization problem of the above formulation can be expressed as:

$$\min_{\alpha} \frac{1}{2}\boldsymbol{\alpha}^T Q\boldsymbol{\alpha} + \boldsymbol{p}^T\boldsymbol{\alpha} \quad (2)$$

$$\text{subject to } \boldsymbol{y}^T\boldsymbol{\alpha} = 0,\ 0 \le \alpha_i \le C, i = 1,2,...,l .$$

Where α_i is the Lagrange multiplier, $\boldsymbol{p} = [-1,-1,...,-1]^T$, Q is a $l \times l$ semidefinite matrix, $Q_{ij} \equiv y_i y_j \mathbf{K}(x_i, x_j)$, and $\mathbf{K}(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$ is the kernel function. Common kernel functions include linear kernel, polynomial kernel, radial basis function (RBF) kernel and sigmoid kernel. When the kernel function is decided, $(a_1^*, ..., a_l^*)^T$ can be obtained by solving this problem.

The decision function is:

$$\text{sgn}(\boldsymbol{w}^T\phi(\boldsymbol{x}) + b) = \text{sgn}(\sum_{i=1}^{l} y_i \alpha_i K(\boldsymbol{x}_i, \boldsymbol{x}) + b) \quad (3)$$

Among them, sgn() is a sign function, $\boldsymbol{w}$ satisfies

$$\boldsymbol{w} = \sum_{j=1}^{i} y_i \alpha_i \phi(\boldsymbol{x}_i) .$$

B. The Sequential Minimal Optimization Algorithm

When solving equation (2), as matrix Q is generally nonzero, a large amount of computer memory is required to store the kernel matrix for a large size problem, because the number of elements of $\boldsymbol{K}$ is equal to l^2 .Cortes and Osuna etc. used a decomposition algorithm to solve this problem. They decomposed the primal large QP problem into a series of small QP problems and only a subset of training data are optimized at each step according to a certain strategy. The extreme case of this method is sequential minimal optimization (SMO) [5]. The SMO algorithm only allows a subset containing two samples to be optimized at each step. Although this approach increases the number of iterations, the problem size as well as the computational time is greatly reduced at each step, so that the total performance of the algorithms is improved dramatically.

The SMO algorithm is described as follows:

Step 1: Set $\boldsymbol{\alpha}^1 = [0,0,...,0]^T$, $\mathbf{G}^1 = [-1,-1,...,-1]^T$ as the initial feasible solution, and set the number of iterations $k = 1$. The gradient of $f(\boldsymbol{\alpha})$ is $\mathbf{G} = \nabla f(\boldsymbol{\alpha})$.

Step 2: If $\boldsymbol{\alpha}^k$ is an optimal solution of equation (2), then the algorithm terminates. Otherwise, find a subset of two elements to form work set $\mathbf{B} = \{i, j\} \subset \{1,...,l\}$ with WSS working set selection algorithm. Let $N \equiv \{1,...,l\}$, and define $\boldsymbol{\alpha}_B^k$ and $\boldsymbol{\alpha}_N^k$ as sub vectors corresponding to $\mathbf{B}$ and $\mathbf{N}$, respectively.

Step 3: Update $a_i, a_j, G_t (t = 1,...,l)$.

Step 4: Set $k \leftarrow k+1$, and jump back to step 2.

1. Algorithm termination conditions

According to the Karush-Kuhn-Tucker (KKT) conditions, the termination condition of the SMO is:

$$m(\boldsymbol{\alpha}^k) - M(\boldsymbol{\alpha}^k) \le \varepsilon . \quad (4)$$

Here, ε is the error threshold,

$$m(\boldsymbol{\alpha}) \equiv \max_{i \in I_{up}(\boldsymbol{\alpha})} -y_i \nabla f(\boldsymbol{\alpha})_i,\ M(\boldsymbol{\alpha}) \equiv \min_{i \in I_{low}(\boldsymbol{\alpha})} -y_i \nabla f(\boldsymbol{\alpha})_i ,$$

$$I_{up}(\boldsymbol{\alpha}) \equiv \{t \mid \alpha_t < C, y_t = -1 \text{ or } \alpha_t > 0, y_t = -1\}$$

$$I_{low}(\boldsymbol{\alpha}) \equiv \{t \mid \alpha_t < C, y_t = -1 \text{ or } \alpha_t > 0, y_t = -1\}, t = 1,...,l$$

2. The working set selecting algorithm WSS

For the working set $\mathbf{B}$, considering the maximal violating pair, the working set selecting algorithm WSS is:

Step 1: For all of $t, s, t = 1,...,l, s = 1,...,l$, define:

$$a_{ts} \equiv K_{tt} + K_{ss} - 2K_{ts}, b_{ts} \equiv -y_t \nabla f(\boldsymbol{\alpha}^k)_t + y_s \nabla f(\boldsymbol{\alpha}^k)_s > 0 \quad (5)$$

and $$\bar{a}_{ts} \equiv \begin{cases} a_{ts}, & \text{if } a_{ts} > 0 \\ \tau, & \text{otherwise} \end{cases}, \quad (6)$$

Then, select:

$$i \in \arg\max_t \{-y_t \nabla f(\boldsymbol{\alpha}^k)_t \mid t \in I_{up}(\boldsymbol{\alpha}^k)\}$$

$$i \in \arg\max_t \left\{-y_t \nabla f(\boldsymbol{\alpha}^k)_t \mid t \in I_{up}(\boldsymbol{\alpha}^k)\right\}$$

$$j \in \arg\min_t \left\{-\frac{b_t^2}{\bar{a}_{it}} \mid t \in I_{low}(\boldsymbol{\alpha}^k) - y_t \nabla f(\boldsymbol{\alpha}^k)_t < -y_i \nabla f(\boldsymbol{\alpha}^k)_i\right\}$$

Step 2: Get $\mathbf{B} = \{i, j\}$.

The working set selecting is essentially a process to seek the extreme value of function.

3. Lagrange multiplier $\boldsymbol{\alpha}$ and Gradient $\nabla f(\boldsymbol{\alpha})$ updating

In each iteration, the SMO algorithm selects two Lagrange multipliers α_i, α_j and fixes others to make optimization. At the same time, it updates one of the two multipliers by adjusting the other one to follow the linear constraints. Lagrange multiplier α_i, α_j can be obtained as follows:

$$\alpha_j^{k+1} = \alpha_j^k - \frac{y_j[(\nabla f(\alpha^k)_i - \nabla f(\alpha^k)_j)]}{K(x_i, x_i) + K(x_j, x_j) - 2y_i y_j K(x_i, x_j)} \quad (7)$$

Applying constraint condition, then:

$$\alpha_j^{k+1} = \begin{cases} M & \text{if } \alpha_j^{k+1} \ge m \\ \alpha_j^{k+1} & \text{if } m < \alpha_j^{k+1} < M \\ m & \text{if } \alpha_j^{k+1} \le M \end{cases}$$

M is on the upper boundary of α_j^k and m is on the lower boundary of α_j^k. Then:

$$\alpha_i^{k+1} = \alpha_i^k + y_i y_j (\alpha_j^k - \alpha_j^{k+1}) .$$

The gradients of all Lagrange multipliers are updated as follows:

$$\nabla f(a^{k+1})_t = \nabla f(a^k)_t + \Delta\alpha_i y_i K(x_i, x_t) + \Delta\alpha_j y_j K(x_j, x_t) \quad (8)$$

Here, $t = 1,...,l$, $\Delta\alpha_i = \alpha_i^{k+1} - \alpha_i^k, \Delta\alpha_j = \alpha_j^{k+1} - \alpha_j^k$.

4. Calculating value b

If α_i satisfies $0 < \alpha_i < C$, considering the KKT

conditions, $b = -y_i \nabla f(\alpha)_i$, and to avoid the numerical error, let:

$$b = -\frac{\sum_{0<\alpha_i<C} y_i \nabla f(\alpha)_i}{\sum_{0<\alpha_i<C} 1}. \quad (9)$$

Otherwise, the condition becomes:

$$m(\alpha) \le b \le M(\alpha)$$

Here,

$$M(\alpha) = \max\{y_i \nabla f(\alpha)_i \mid \alpha_i = 0, y_i = -1 \text{ or } \alpha_i = C, y_i = 1\}, i = 1,...,l$$

$$m(\alpha) = \min\{y_i \nabla f(\alpha)_i \mid \alpha_i = 0, y_i = -1 \text{ or } \alpha_i = C, y_i = 1\}, i = 1,...,l$$

Then the value of midpoint in this range is b:

$$b = \frac{M(\alpha) + m(\alpha)}{2}. \quad (10)$$

III. PARALLELANALYSIS AND DESIGN OF SMO ALGORITHM

A. Parallel Analysis

According to the above analysis, it can be observed that:

1) In the sequential SMO algorithm, the optimization process accounts for most of the computation time. In particular, over 90% of the total computational time is used for updating $\nabla f(\boldsymbol{\alpha})$ [12]. However, according to (9), updating $\nabla f(\boldsymbol{\alpha})$ is evaluated one at a time. This is suitable for parallelization. Therefore, updating can be implemented in a parallel design.

2) In the SMO algorithm, kernel evaluations consist of a number of steps (e.g. the calculation of $\boldsymbol{\alpha}$, $\nabla f(\boldsymbol{\alpha})$ and the decision function). The calculation involves examining all of the training data points (e.g. dot product operation).Although there are a very large amount of floating-point calculations, the instructions are very simple. As GPU is very effective in dealing with this kind of SIMD problems, the kernel evaluations can also be performed in a parallel fashion.

3) The WSS algorithm is a process to seek the maximal or the minimal value, which requires comparison of all the data points one by one in order to obtain the extremum. The efficiency of the CPU is very low in dealing with this kind of problem. By using the GPU parallel reduction algorithm, the global problem can be decomposed into local problems for acceleration.

4) In the SMO algorithm, a lot of matrix operations, such as multiplication and cumulative sum, are involved. These operations also belong to simple instruction problems and can be executed in a parallel design by the GPU to a high degree.

5) The efficiency of the GPU is poor in processing conditional branch problems. However, in the SMO algorithm, each iteration is terminated by conditional decisions. For large-size problems, the number of executing branch statement will increase greatly. Therefore this part should still be implemented on the CPU.

B. Parallel Design of SMO Algorithm based on GPU

In order to improve the efficiency in the determination of intrusion with large scale problems, a parallel SVM algorithm based on GPU and CUDA (Compute Unified Device Architecture) [13] named GSVM is developed. GSVM consists of two independent parts: training and testing. The framework of these two parts based on the GPU is shown in Figures 1 and 2.

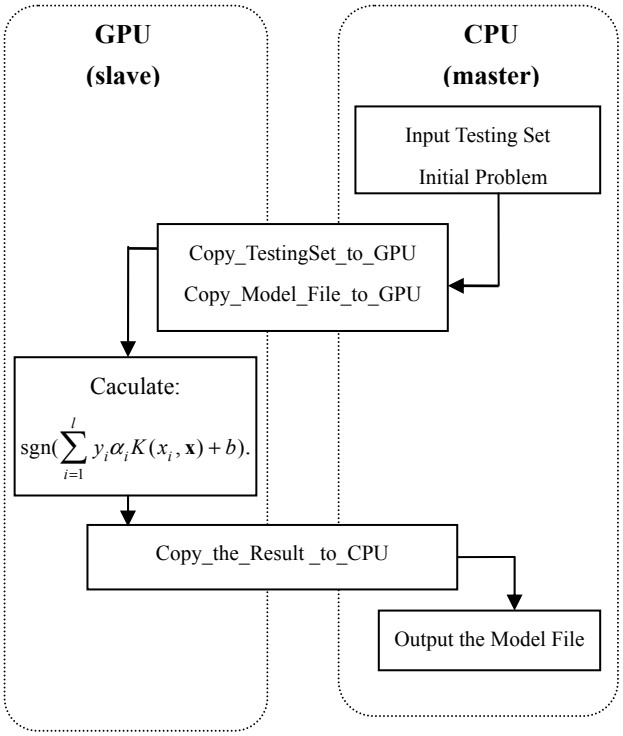


Figure 1. GSVM classification algorithm process

C. Time Complexity Analysis

For sequential SMO algorithm, updating $\nabla f(\boldsymbol{\alpha})$ and selecting working set B are the most time consuming steps.Cao and co-workers showed that,in each of iteration, if the time complexity of kernel calculating is $O(l)$,the time complexity of the whole algorithm is $\#Iter \times O(ld)$ when the most of columns of matrix Q are not in the cache; otherwise it is $\#Iter \times O(l)$ [12]. Here, *#iter* denotes the number of iterations and is proportional to*l*.

For parallel SMO algorithm, let p denotes the number of threads used, #iter' represents the number of iterations in GPU end. If the most columns of matrix Q are in the cache, calculation is insignificant. Therefore the time complexity is $\#Iter' \times O(l)$. Otherwise, it is $\#Iter' \times O(kl + \Delta)$ .Here, Δ is the cost of threads synchronization and k is a coefficient associated with d and p.

IV. APPROACHES USED TO IMPROVE GPU PARALLEL COMPUTING PERFORMANCE

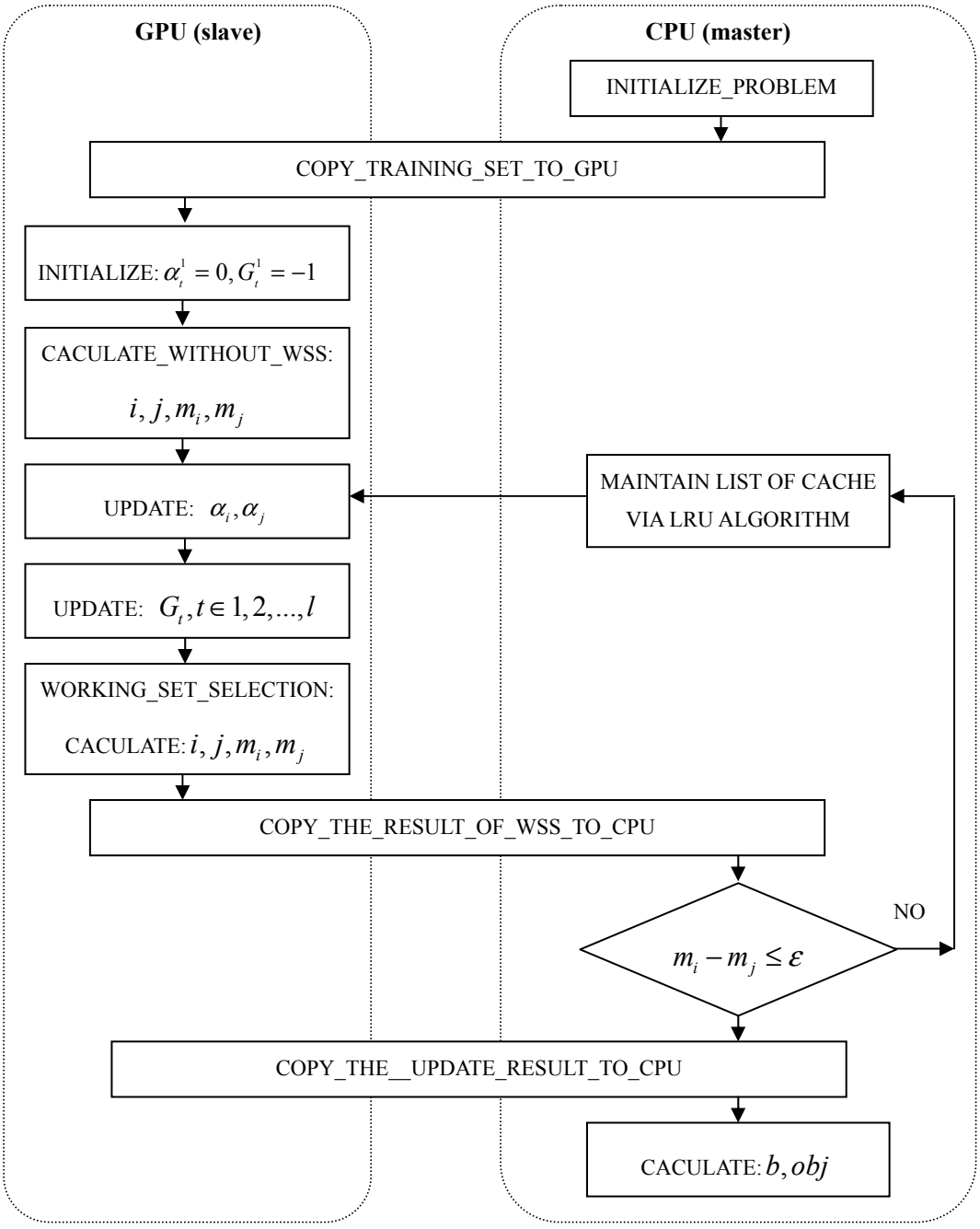


Figure2. GSVM training algorithm process

A.SPRG Parallel Computing Method

SVM training and testing processes involve matrix operations such as multiplication, cumulative sum and seeking extreme value. The computational complexities of these operations are proportional to the size of data sets. GSVM algorithm uses parallel reduction methods to optimize these calculation processes. The basic idea of this method is described as follows. First, the data are divided into *n* parts and transferred to parallel computing nodes. Second, each computing node summarizes its data and executes corresponding operations, such as multiplication and addition. Finally, each parallel computing node transmits its operation results to the aggregation node for implementing the last operation. Because all of the data are split and computed in parallel, the total computational time is reduced.

To be specific, in this paper, the SPRG (scatter-parallel-reduce-gather) parallel reduction method is designed as follows.

1) Scatter: Data sets are partitioned into smaller subsets according to the number of blocks and threads used and moved from global memory to shared memory in line with the corresponding address sequence. Considering that GPU usually accesses the shared memory in 4 clock cycles, far less than the 400-600 clock cycles that are needed to access the global memory, this operation also improves the subsequent access speed.

2) Parallel-reduce: All of the threads in blocks execute the same instruction and obtain the reduction sub-results.

These results are stored in the head address units of blocks.

3) Gather: Considering that the data in the GPU blocks are unable to communicate with each other, the sub-results in each block need to be transferred from the shared memory back to the global memory according to the corresponding address sequence in order to do the further reduce.

4) Repeat steps 1 and 2 until the sub-results are reduced in one block. The final result is output to the global memory for ease of communication with CPU end.

Figure 3 shows the process of solving the extreme value based on the SPRG method.

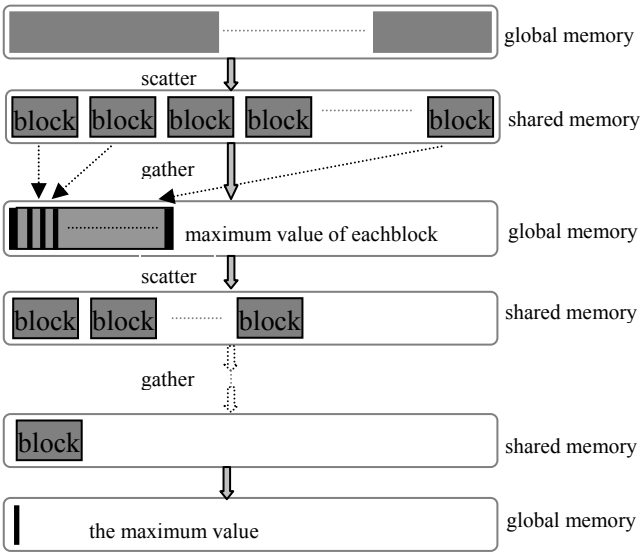


Figure 3.extreme value solving based on SPRG

In the GSVM algorithm, the SPRG method is widely used, such as to solve the gradient extremum in the training period and achieving the cumulative sum in the testing period. If the time required for data transmission and synchronization can be ignored, the computational complexity can then be decreased by one order of magnitude compared with the serial algorithm. This has a great advantage in processing large-scale arrays.

B. Optimization Methods

1) Cache technology: In sequential SMO, cache technology is used to store calculated results of kernel function and LRU (Least Recently Used) algorithm is used to manage a two-way circular linked list of cache. In GSVM, considering the complexity of the list structure and LRU algorithm, the task of maintaining the list is still executed by the CPU, while caches are allocated in the GPU. In each iteration, when the working set selection is completed, a judgment will be made by CPU end (Host end), and then inform GPU end (Device end) to execute the subsequent operations.

2) Algorithmic Optimizations: ① In the shared memory, when applying the reduction technology with tree-based interleaved addressing mode, bank conflicts in the shared memory usually occur and result in memory access operation serialization. In this case, the effective bandwidth decreases exponentially. Therefore the sequential addressing principle should be followed to unify the operations of all the threads in warps to avoid this problem.②In addition, when branch statements(e.g. "if-then-else") are executed by threads in one warp, the CUDA GPU deals with these threads sequentially. As such branch statements should be avoided in one warp.

3) Instruction Optimizations: ①The ability of GPU's integer processing unit is weak, therefore the integer modulo and division operation costs are usually very high. Thus the bit operation is adopted instead.②In the CUDA programming, all the threads in a block typically need to be synchronized in each loop. However, since the operations in one warp always meet consistency, this step is not needed at this case. Loop unrolling is a common instruction optimization method to improve the execution speed, even for very long expression. In many cases, loop unrolling can improve performance by 20%. Thus for an inner loop, the loop unrolling method can be used to deal with the last few iterations in order to accelerate the speed.

V. EXPERIMENT AND CONCLUSIONS

A. The Experiment Environment

In our experiment, the CPU is Intel Pentium dual-core E5500 with a frequency of 2.80GHz. The GPU is NVIDIA GeForce GTS 250.The memory capacity is 512 MB, the number of multiprocessors is 16, the number of cores (stream processors) is 128, multiprocessors shared memory size is 16KB, and memory bandwidth is 70.4 GB/S. The integrated development environment is Microsoft Visual Studio 2008 with CUDA SDK 2.3. LIBSVM is used as the sequential SVM algorithm.

B. Experiment and Results

1. Experiment 1

(1) Description of the Experimental Data

Nine groups of large-scale datasets from the field of machine learning are used to verify the validity and superiority of the GSVM algorithm. The size and dimension of these datasets are shown in TABLE I. Here, #Tr is the number of training samples, and #Te is the number of testing samples.

(2) Experimental in SVM Training Period

In the SVM training period, the LIBSVM uses the default parameters. Set the RBF kernel as the kernel function, $C = 1, \varepsilon = 0.001$.

The LIBSVM and GSVM training algorithms were executed for the same training data set to derive the training models. The number of the support vectors, iterations and the training time, as shown in TABLE II. Here, *#SV* indicates the number of the support vectors, *#iter* denotes the number of iterations, *T* represents training time, and *ratio* is the speed-up ratio. In order to display the result visually, the comparison of training time is also shown in Figure 4.

(3) Experiment in testing period

The LIBSVM and GSVM testing algorithms were

executed for the same test dataset with the two models built above to compare the classification accuracy and testing time. Classification results are shown in TABLE III. In this table, *T* denotes testing time, and *Acc* is the accuracy. The comparison of the testing time is also listed in Figure 5.

(4). Experimental Analysis and Conclusions

From the above results, it can be seen that:

1) The numbers of support vectors, iterations and the classification accuracy obtained from GSVM and LIBSVM are virtually the same. These demonstrate the correctness of the GSVM algorithm.

2) Comparing the experiment result of GSVM with LIBSVM on large-scale datasets, the training time decreased 2 to 43 times, and the classification time was reduced 40 to 349 times. These results demonstrate the superiority of the GSVM algorithm in computational speed.

3) Comparing the parallel training algorithm with the testing algorithm, the performance improvement of the later is remarkable due to the absence of conditional branch operations in the testing period. The testing part in SVM is more suitable to be parallelized.

4) Cost for communication between CPU and GPS is always considered as the bottleneck for GPU parallel implementation. It is noted that the much improved computational speed from GSVM already included this cost. As such, this algorithm could find valuable practical applications.

2. Experiment 2

(1) Description of Experiment Data

In order to verify the effectiveness of GSVM in intrusion detection, KDD99 dataset is adopted in this set of experiments. The numbers of the training samples and the test samples are 49407 and 49405, respectively.

(2) Results and conclusions

For better understanding of the cost of various subparts in the sequential and parallel SMO, the computation time in different steps (initialization, optimizing $\boldsymbol{\alpha}$, updating $\nabla f(\boldsymbol{\alpha})$ and selecting work set **B**) are listed in TABLE III. The table shows that updating $\nabla f(\boldsymbol{\alpha})$ and selecting work set **B** cost large amount of training time, and is better performed in parallel.

The experiment evaluates the intrusion detection system based on GSVM from the training time, testing time, the accuracy (Acc), the detection rate (DR) and the false positive rate (FP). The experiment results are shown in TABLE IV, V and VI.

VI. CONCLUSIONS

In order to solve the SVM time-complexity problem, a parallel SVM algorithm based on the GPU, namely GSVM, is presented in this paper. According to the parallel implementation architecture of GPU, the training and testing part of the sequential SVM algorithm are parallelized. Additional optimization measures are designed to improve the performance of the GSVM. Experimental results derived from large-scale datasets show that the GSVM algorithm can speed up SVM training and testing process dramatically without loss of classification accuracy. The speedup effectiveness becomes more obvious with the increase of the dataset size. The GSVM algorithm has good real-time processing capabilities on massive data and can be utilized on abnormality detection effectively.

From the experimental results on KDD99, it was concluded that the intrusion detection system based on the GSVM parallel algorithm has virtually the same accuracy, detection rate and false positive rate compared with results obtained from the LIBSVM algorithm. Moreover, the training and classification time are decreased by 13.3 times and 64.92 times, respectively, in the GSVM parallel algorithm. Thus, the GSVM parallel algorithm based on the GPU is very suitable for constructing intrusion detection classifier.

TABLE I.
THE DATA SET DESCRIPTION

Dataset	#Tr	#Te	Dim	Source
cod-rna	59,535	271,617	8	BMC
covtype	300,000	281,012	54	UCI
epsilon	20,000	20,000	2000	PASCAL Challenge2008
ijcnn1	49,990	91,701	22	IJCNN
webspam	150,000	200,000	128	CEAS
face	1,996,201	2,005,601	10	FRGC
sonar	10,000	9,375	60	UCI
adult	32,561	16,281	123	UCI
w8a	49,749	14,951	300	UCI

TABLE II.
COMPARISON OF SVM TRAINING RESULTS

Data set	CPU Training			GPU Training			ratio
	#SV	#iter	T (sec)	#SV	#iter	T (sec)	
cod-rna	16,179	65,638	1,301	16,214	65,712	123	10.542
covtype	299,997	489,969	128,770	299,998	490,404	2,968	43.381
epsilon	19,833	10,038	4,575	19,834	10,102	107	42.766
ijcnn1	8,970	5,563	119	8,986	5,713	11	11.912
webspam	63,930	35,854	8,867	63,926	35,791	254	34.936
face	3,842	2,178	1,396	3,841	2,277	45	30.737
sonar	1,430	1,271	7	1,434	1,718	3	2.33
adult	11,951	7,848	300	11,951	8,186	16	18.75
w8a	2,949	2,843	208	2,944	2,824	9	23

TABLE III.
COMPARISON OFTHE COMPUATION TIME IN DIFFERENT STEPS OF TRAINING PERIOD

CPUTime (sec)			GPU Time (sec)		
Initialization	$\boldsymbol{\alpha}$	$\nabla f(\boldsymbol{\alpha})$ +WSS	Initialization	$\boldsymbol{\alpha}$	$\nabla f(\boldsymbol{\alpha})$ +WSS
0.0	0.0	24.8	0.0	0.0	1.63

TABLE IV.
COMPARISON OF TRAINING RESULTSON KDD99

CPU			GPU		
Acc(%)	DR	FP	Acc(%)	DR	FP
98.92	98.33	0.188	98.92	98.32	0.188

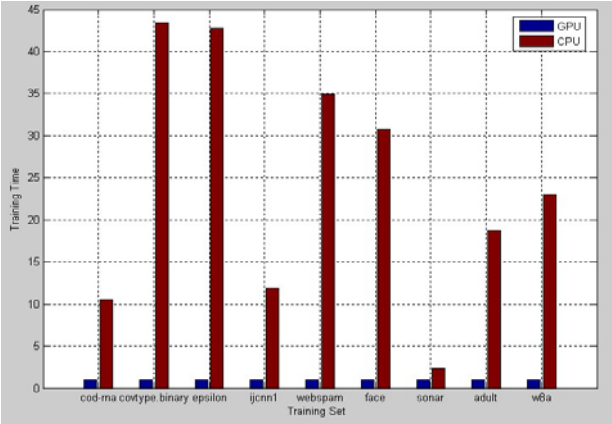


Figure 4.comparison of the training time of GPU and CPU

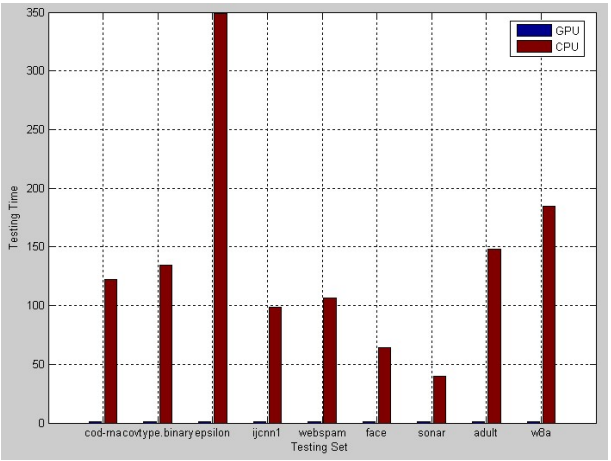


Figure 5.comparison of the classification time of GPU and CPU

TABLE VI.
COMPARISON OF SVM CLASSIFICATION RESULTS

Data set	CPU Training			GPU Training			ratio
	#SV	#iter	T (sec)	#SV	#iter	T (sec)	
KDD99	975	849	26.1	976	998	2.03	13.3

TABLE V.
KDD99 CLASSIFICATION PERFORMANCE

Data set	CPU Classification		GPU Classification		ratio
	T (sec)	Acc (%)	T (sec)	Acc (%)	
KDD99	17.4	98.917	0.27	98.915	64.96

REFERENCES

[1] N. Cristianini, J. Shawe-Taylor, *An Introduction to Support Vector Machines*, Cambridge, 2000.

[2] B.Yu, H.F.Li, C.Y. Fang."Speech Emotion Recognition based on Optimized Support Vector Machine", *Journal of Software,* vol. 7, no. 12, pp.2726-2733, 2012.

[3] M.h. Zhang, L.h. Gang."A Detection Method of Driver's Face Orientation Based on Visual Cues and SVM",*Journal of Software*, vol. 8, no. 4,pp.924-931,2013.

[4] E.Osuna, R.Freund, F.Girosi, "*An improved training algorithm for support vector machines*",*Proc. IEEE Workshop on Neural Networks and Signal Processing*, Piscataway: IEEE Press,1997, pp. 276-285.

[5] J. Platt,"Fast training of support vector machines using sequential minimal optimization", *Advances in Kernel Methods-Support Vector Learning*, Cambridge, MA, 1998, pp. 185-208.

[6] T. Joachims,"Advances in kernel methods-support vector learning", Cambridge, USA, 1998.

[7] C. Chih-Chung, L. Chih-Jen, LIBSVM: a Library for Support Vector Machines.[online].Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[8] K. Yadav,A. Mittal, and M. A. Ansar, et al., "Parallel Implementation of Compressed Sensing Algorithm on CUDA- GPU", *International Journal of Computer Science and Information Security*, vol. 9,no.3,2011.

[9] B. Pieters,D. Van,R.Wesley,et al., "Performance Evaluation of H.264/AVC Decoding and Visualization using the GPU," *Applications of digital image processing* , vol.6696,no.1,2007.

[10] Meredith, J.S.Alam, S.R.Vetter, et al. , "Analysis of a Computational Biology Simulation Technique on Emerging Processing Architectures. Parallel & Distributed Processing Symposium (IPDPS)",*2007 IEEE International*, Long Beach, CA. 2007,vol. 1.

[11] H.f.Wang, Q.k. Chen. "Power Estimating Model and Analysis of General Programming on GPU". *Journal of Software*, vol. 7, no. 5, pp.1164-1170,2012

[12] L.Cao, S.Keerthi, and Ong, et al., "Parallel sequential minimal optimization for the training of support vector machines", *IEEE Transactions on Neural Networks*, vol.17,pp.1039-1049,2006.

[13] "NVIDIA Corporation NVIDIA CUDA C Programming Guide3.2," 2010,[online].Available: http://developer.download.nvidia.com/compute/cuda/3_2 /toolkit/docs/CUDA_C_ Programming_Guide.pdf.

Xueqin Zhang received the Ph.D.degreein Detection Technology and Automation Devices from East China University of Science and Technology (ECUST), Shanghai, China, in 2007. Since 1998, she has been in the Electrical and Communication Engineering Department, ECUST, where she is currently an ASSOCIATE PROFESSOR.

At 2006, she worked as a visiting scholar in University of Wisconsin Madison. Her research interests include pattern classification, information security and data mining etc.

Yifeng Zhang received the M.S. degree in Electrical and Communications from East China University of Science and Technology (ECUST), Shanghai, in 2012. His research interests are parallel computation, pattern classification etc.

Chunhua Gu received the Ph.D. degree in Control Science and Engineering from East China University of Science and Technology (ECUST), Shanghai, China, in 2007. From 1992 to 2013, he was in the Computer Science and Engineering Department, School of Information Science and Engineering, ECUST, where he is currently a PROFESSOR. Now he is in the School of Computer Science and Technology at Shanghai Institute of Electric Power.

At 2002, he worked as a visiting scholar in School of Computing and Information Sciences, Florida International University. His research interests include intelligent computing, software engineering and information security.

A Greedy Algorithm for Constraint Principal Curves

Shiying Yang
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044，China
Email: 12120291@bjtu.edu.cn

Dewang Chen*
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044，China
Corresponding author: dwchen@bjtu.edu.cn

Xiangyu Zeng
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044，China
Email: 11120319@bjtu.edu.cn

Peter Pudney
Center for Industrial and Applied Mathematics, University of South Australia, Mawson Lakes, Australia
Email: Peter.Pudney@unisa.edu.au

***Abstract*—Principal curves can learn high-accuracy data from multiple low-accuracy data. However, the current proposed algorithms based on global optimization are too complex and have high computational complexity. To address these problems and in the inspiration of the idea of divide and conquer, this paper proposes a Greedy algorithm based on dichotomy and simple averaging, named as KPCg algorithm. After that, three simulation data sets of sinusoidal, zigzag and spiral trajectories are used to test the performance of the KPCg algorithm and we compare it with the k-segment algorithm proposed by Verbeek. The results show that the KPCg algorithm can efficiently learn high-accuracy data from multiple low-accuracy data with constraint endpoints and have advantages in accuracy, computational speed and scope of application.**

***Index Terms*—Principal curves algorithm; principal of nearest neighbor; adaptive radius; dichotomy; simple averaging**

I. Introduction

A principal curve is a smooth curve that passes through the middle of a data set. In statistics, the principal curve should be self-consistent; each point on the principal curve is the expected average of the data that projects to this point. Principal curves have been widely used in a growing number of areas, such as fingerprint skeleton extraction, hydraulic machinery, image processing [1-3]. We are particularly interested in automatic generation of railway track profiles from GPS data [4-6].

In 1904, Spearman proposed the linear principal component analysis method. This method is simple, and is now one of important tools for statistical analysis of data [7]. But not all data is linear. Hastie proposed the concept of principal curves in 1984; these smooth curves should pass through the "center" of the data distribution and satisfy a "self-consistency" property [8-9]. In 1992, Banfield and Raftery principal curves (BR principal curves) improved the previous principal curves. But BR principal curves bring the numerical instability. So BR principal curves algorithm may obtain smooth but false principal curves [10]. In 1997, Kegl proposed the concept of length constraint principal curves [11], and proved the existence and uniqueness of the K principal curves. He also proposed a polygonal-time algorithm to obtain K principal curves.

In 2000, Verbeek proposed the K-Segment principal curves (KPCv) algorithm [12]. In his paper, Verbeek uses an incremental method when finding principal curves. Line segments are fitted and connected to form polygonal lines. New segments are inserted until a performance criterion is met. However, when we add segments to improve the accuracy of a principal curve, the algorithm may fail as shown in Figure1, and the generated principal curve needs to be reprocessed in the next stage.

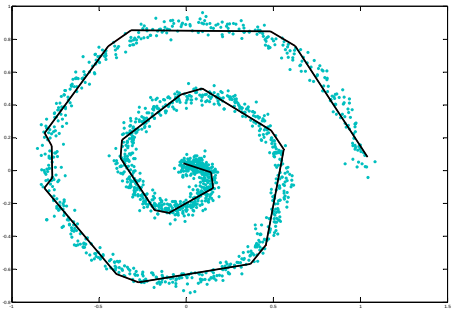
There are many principal curves algorithms based on global optimization. In 2008, Chen proposed a heuristic algorithm based on continual split and non-linear optimization to estimate the path of a railway line from multiple runs collecting GPS data [13]. Still, this algorithm has some defects including long computation times and narrow application scope. After that, Zhang and Chen proposed a principal curves algorithm [14] which resolved the generation and adaptability of principal curves with constraint points. But the algorithm is too complex to be practical. In 2011, Jia proposed the MPM (Max Point Method) optimization algorithm for constraint principal curves [15]. However, the fitness and robustness need to be improved. Also in 2011, Zhang proposed two principal curve algorithms for partitioning

doi:10.4304/jcp.9.5.1125-1130

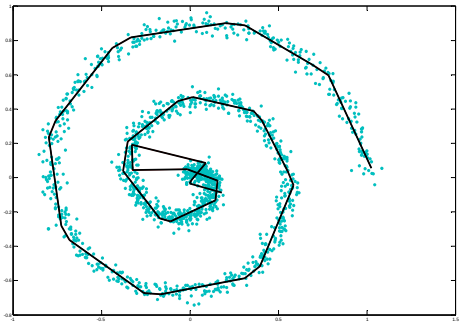
high-dimensional data spaces [16]. And both of them can show a better performance than some other competing partitioning algorithms in some ways.

To increase the computation efficiency, we propose the Greedy algorithm based on dichotomy and simple averaging, named as KPCg algorithm. As the start point and end point of the principal curve are known, we draw circles from the start point to the endpoint with adaptive radii, where the radii values are determined by dichotomy. And we search for the vertexes by simple averaging. Finally, we can get the principal curve by connecting the vertexes in order. To test the performance of the proposed principal curves algorithm, we compare it with the K-Segment algorithm proposed by Verbeek (named as KPCv algorithm in this paper). The results show that our algorithm can efficiently generate good approximations from multiple low-accuracy data and has advantages in accuracy, computational time and scope of application.

The structure of the paper is as follows: Section II defines a constraint principal curve and describes how we calculate errors; In Section III, we describe the KPCg algorithm in detail; and Section IV shows the results of the verification and analyzes the results; we end this paper with a conclusion.



(a) KPCv principal curve with 12 segments



(b) KPCv principal curve with 17 segments

Figure1. Results for KPCv algorithm

II. Constraint Principal Curves and Error Model

A. Constraint Principal Curves

In many practical applications, some points can be fixed as constraints when generating a principal curve and we need to take them into account. Therefore, a constraint principal curve is a principal curve with several fixed points (i.e. points which have been measured accurately). In Figure2, the green curve is a principal curve and the black curve with two fixed red endpoints is a constraint principal curve where V_s is the start point and the V_e is the end point.

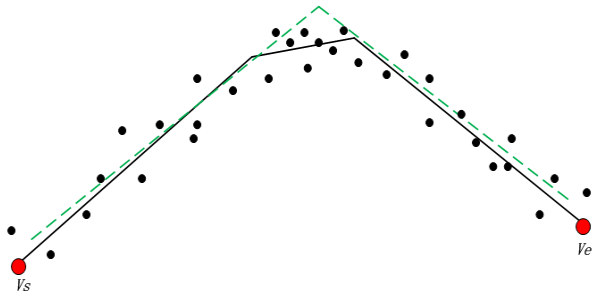


Figure2. Principal curve and constraint principal curve

When generating a principal curve, making full use of the fixed points can help us obtain a constraint principal curve of higher precision.

B. Principle of Nearest Neighbor and Error Model

Principal of nearest neighbor has been used in some areas, such as querying, filtering [17, 18]. Here, it can be used practically and extensively when establishing the error model.

Similar to K-Segment principal curves, we will represent a principal curve by a set of vertices and the line segments between adjacent vertices. By comparing the projection distance from the data points to each line segment and each vertex, the data points can be divided into different parts: one line segment or one vertex. As illustrated in Figure3, the constraint principal curve is composed of vertexes V_j (j = 1,…, n), and line segments $S_{j,j+1}$, j = 1, …, n-1. Here, n is the number of vertexes.

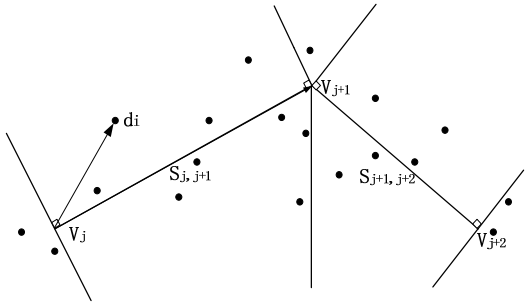


Figure3. Principle of nearest neighbor

The distances from data point d_i (i = 1, .., m, m is the number of data points), to vertex V_j can be computed by Formula (1).

$$D_{P_{i,j}} = \left|\overrightarrow{V_j d_i}\right| = \sqrt{(X_{d_i} - X_{V_j})^2 + (Y_{d_i} - Y_{V_j})^2} \quad (1)$$

The distance from data point d_i to corresponding line-segment $S_{j,j+1}$can be computed by Formula (2).

$$D_{L_{i,j}} = \left|\overrightarrow{V_j V_{j+1}} \times \overrightarrow{V_j d_i}\right| \Big/ \left|\overrightarrow{V_j V_{j+1}}\right| \quad (2)$$

The smallest one of all $D_{P_{i,j}}$ and $D_{L_{i,j}}$ (j = 1,…, n) is defined as the projection distance of each data point to the principal curve, show in the following Formula (3). So E_i is the minimum distance from point d_i to the principal curve.

$$E_i = \min\{D_{P_{i,j}}(j=1,2,\cdots,n), D_{L_{i,j}}(j=1,2,\cdots,n)\} \quad (3)$$

$\overline{E}$ is the mean of all E_i (i = 1, .., m, m is the number of data points), and E is the error of the principal curve, shown in Formula (4).

$$E = \overline{E} = \sum_{i=1}^{m} E_i \Big/ m \quad (4)$$

When we have more than two fixed points through which the principal curve must pass, we can treat each pair of adjacent fixed points as the start and end points of a separate principal curve. This reduces the problem to that of finding a principal curve where the first and last vertexes are specified. Our objective is to define vertexes V_1, …, V_n (and the corresponding line segments) so that the mean distance E is as smaller as possible.

III. The Greedy Algorithm: KPCg Algorithm

In the inspiration of the idea of divide and conquer, a Greedy algorithm based on dichotomy and simple averaging is developed. We name it the KPCg algorithm. The KPCg algorithm progresses from the start endpoint towards the end point with adaptive radius. It uses simple averaging to find new vertexes.

In the following parts, V_s is the start point, V_e is the end point. n is the number of vertexes and it will be unknown until we find all the vertexes. We set the upper and lower bounds of local error E^l_j (j=1,2,...,n-1): E_{min} and E_{max}, where E_{min}=0 usually. E is the overal error of the principal curve we obtain by the principal curves algorithm. For each E^l_j (j=1,2,...,n-1), if E^l_j <=E_{max}, e<=E_{max}. In the KPCg algorithm, simple averaging is used to calculate the new vertex V_j (j=2, 3,..., n-1, V_1 is V_s and V_n is V_e), and we use dichotomy to make the radii adaptive.

A. Dichotomy

We can continuously approximate the target by adaptive algorithm like the adaptive genetic algorithm [19]. When drawing circles, adaptive radius can be obtained according to the dichotomy to make E^l_j (j=1, 2,..., n-1) ∈ [E_{min}, E_{max}]. Specific methods are as follows:

step1. The center of a circle is V_j (j=1, 2,..., n-1), the initial interval of the radius is [R_d R_u]. R_d=0 and R_u=2* d^e_j, where d^e_j is defined as the distance from V_j to V_e;

step2. R_j= R_d + (R_u - R_d)/2, j=1,2...,n-1;

step3. Draw the circle. We can find a vertex V_{j+1} by simple averaging introduced in the next part and compute the local error E^l_j (j=1, 2,..., n-1);

step4. Judge: if E_{min}<=E^l_j<=E_{max}; the value of R_j is what we need; if E^l_j <=E_{min}, then R_d =R_j and turn to the second step; else, R_u =R_j and turn to Step2.

Then, we can get a suitable value of R_j (j=1, 2,..., n-1). By connecting V_j and V_{j+1}, we can get a sub principal curve whose local error E^l_j ∈ [E_{min}, E_{max}].

The flowchart of dichotomy is shown in Figure4.

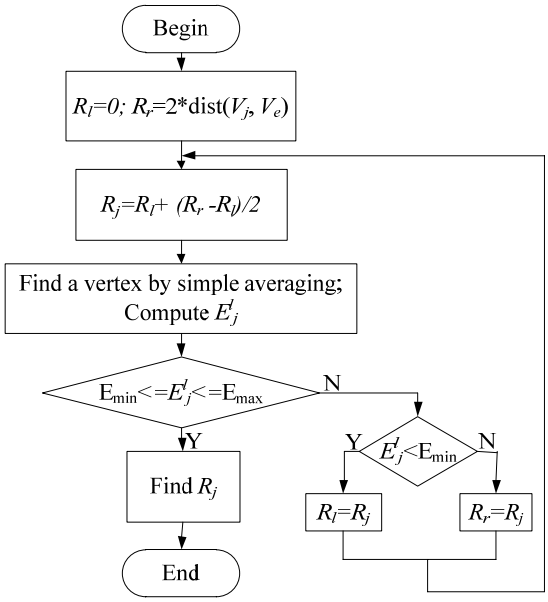


Figure4. The flowchart of dichotomy

B. Simple Averaging

When solving the problems, simple averaging can also help us make the best choice in the current view. That is to say, we can only use simple averaging to get a local optimum in some sense, but without the whole consideration. This can speed up the computation.

As we see, a principal curve is constructed by different lines that are sub principal curves. For each line, when the start point is known (that is the end point of the previous line), we need to find this line's end point, and then a line is obtained. In the KPCg algorithm, we use simple averaging to calculate the new vertexes: every vertex's coordinates are the mean value of the point coordinates within a ring. Here, the ring's inner radii R_j' (j=1,2,...,n-1) can be obtained by R_j'=R_j'*0.9 and the outer radii R_j (j=1,2,...,n-1) is obtained by dichotomy. For better displaying the idea of simple averaging, we use sectors to represent circles in Figure5. We will make a specific introduction of simple averaging by searching for V_2.

V_s is the first vertex, that is V_1, and we need to find V_2. The initial value of R_1 is the distance from V_e to V_1, that is R_1= R_u/2= d^e_1. And R_1' is obtained by R_1'=R_1*0.9. So we get the first ring, and the coordinates of V_2 are the mean value of the coordinates of the points within the ring. Assuming that V_1, V_2 and the line between them construct a sub principal curve, we compute the local error E^l_1 of the sub principal curve according to the error model introduced in Section II part B. If E_{min}<=E^l_1<=E_{max}, we get V_2. Otherwise, we change the value of R_1 by dichotomy introduced in the previous part. Finally, we can find V_2 and obtain a sub principal curve which can pass through the middle of the data set whenever possible.

Since V_2 has been obtained, we are now searching for V_3 in the same way. After that, V_3, V_4,..., V_{n-1} can also be

found in this way and V_e is the last vertex V_n (n is the number of vertexes).

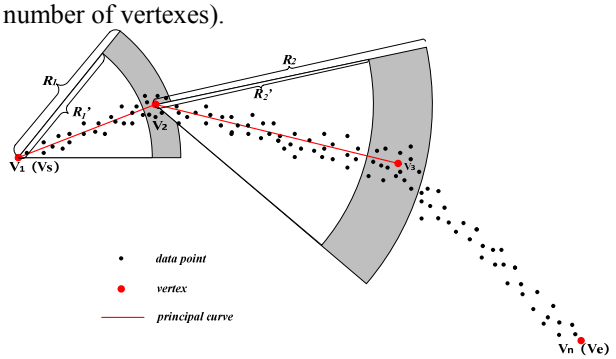


Figure5. Simple averaging

C. The KPCg Algorithm

The KPCg algorithm is a constraint principal curves algorithm based on dichotomy and simple averaging. We draw circles from the start point to the end point, where the radii' values are determined by dichotomy according to the upper and lower limits of local error, and we fit the unused data in each circle by simple averaging.

The key steps are as follow:

step1. n=2, V_s'=V_s;

step2. Connecting V_s' and V_e. Assuming that V_s', V_e and the line between them construct a sub principal curve, we compute the local error E^l_{n-1} of the sub principal curve. If E_{min}<= E^l_{n-1}<=E_{max}, turn to Step6. Otherwise, turn to step3;

step3. n=n+1;

step4. Use dichotomy to find a value of R_j and use simple averaging to find V_{n-1}, where $E^l_{n-1} \in [E_{min}, E_{max}]$;

step5. V_s'=V_{n-1}, turn to Step2;

step6. Line V_1, V_2,..., V_n. Then we get the principal curve and calculate the overall error which will be smaller than E_{max}.

The flowchart of the KPCg algorithm is shown in Figure6.

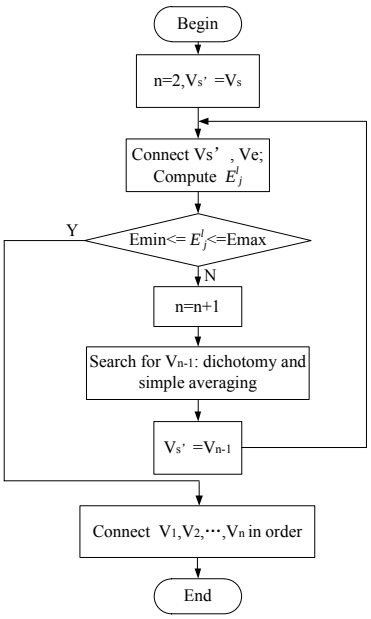


Figure6. The flowchart of the KPCg algorithm

IV. Verification and Analysis

A. Acquisition of Simulation Dataset

To test the fitness of the KPCg algorithm, we conducted experiments on three artificial datasets of sinusoidal trajectory, zigzag trajectory and spiral trajectory. And the acquisition of these simulation dataset is: Firstly, we generate an accurate simulated trajectory such as the sinusoidal trajectory. Then, some data points are randomly generated around the accurate trajectory.

In this way, we get the simulation datasets: sinusoidal dataset with 1000 data points, zigzag dataset with 2000 data points and spiral dataset with 1500 data points.

B. Evaluation Indexes

In the experiments, we do not need to compare the generated principal curves with the exact trajectory for that the data are corrupted by the noise we add.

To make the comprehensive comparision on the KPCg algorithm and the KPCv algorithm, we define the following indices: 1) N is the number of the vertexes generated by each algorithm. The smaller the N is, the less storage it consumes; 2) T represents the time each algorithm uses when generating a k-segment principal curve (KPC); 3) and E is defined in Eq.(4).

C. Verification

1. Sinusoidal data set

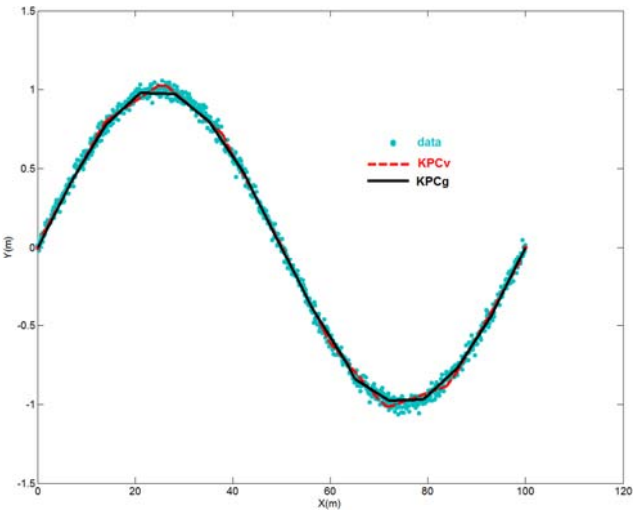


Figure7. Results on sinusoidal data set

TABLE I.
COMPARISON ON SINUSOIDAL DATA SET

Algorithm	N	T/s	E/m
KPCv	16	17.5	0.0295
KPCg	**15**	**2.8**	**0.0250**

2. Zigzag data set

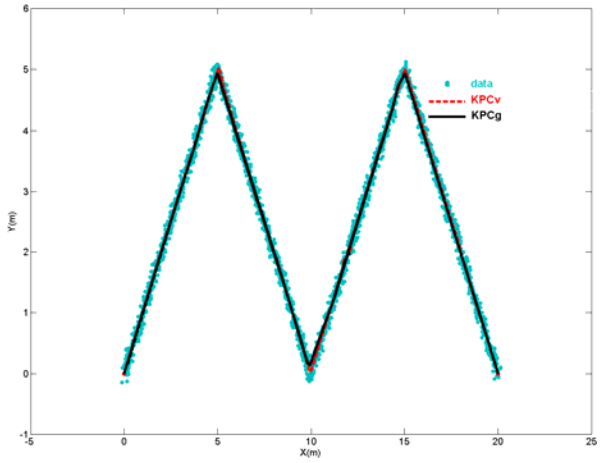


Figure8. Results on zigzag data set

TABLE II.
COMPARISON ON ZIGZAG DATA SET

Algorithm	N	T/s	E/m
KPCv	18	15.2	0.0742
KPCg	18	**7.4**	**0.0731**

3. Spiral data set

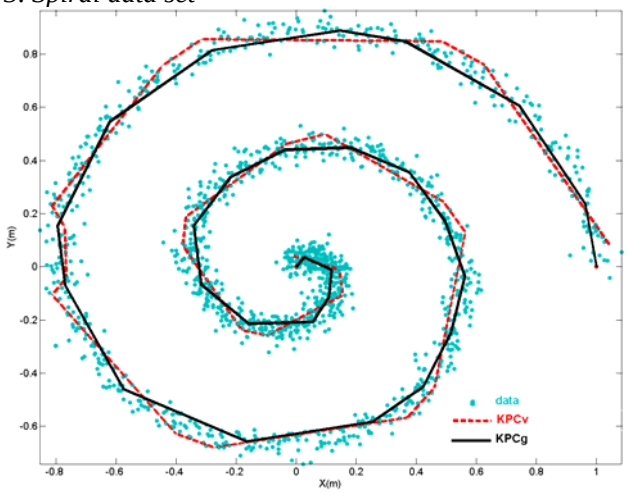


Figure9. Results on spiral data set

TABLE III.
COMPARISON ON SPIRAL DATA SET

Algorithm	N	T/s	E/m
KPCv	**24**	30.1	0.0355
KPCg	28	**9.0**	**0.0269**

D. Analysis

Figure7-9 show that the KPCg algorithm can learn high-accuracy data from multiple low-accuracy data. What's more, the KPCg algorithm also has a large scope of application.

When comparing with the KPCv algorithm in TABLEI-III, we can find that the KPCg algorithm is faster because of the simple averaging. Owing to the constraint endpoints and adaptive radius, the error of the KPCg algorithm is smaller than that of the KPCv.

V. CONCLUSIONS

In this paper, the history and defects in principal curves are introduced firstly. Then, in the inspiration of the idea of divide and conquer, we proposed the KPCg algorithm, a Greedy algorithm based on dichotomy and simple averaging. At last, we compare it with the KPCv algorithm by simulation data.

For the usage of dichotomy, simple averaging and the constraint endpoints, the results show that the KPCg algorithm has a good performance on accuracy, computational efficiency and the scope of application.

In future research, we will study the constraint principal curves algorithms deeply, and improve them so that the precision and robustness can be enhanced.

ACKNOWLEDGMENT

This work is partially supported by New Scientific Star Program of Beijing under grant 2010B015, by the Fundamental Research Funds for the Central Universities under grant 2012JBM016, by the independent research project from the State Key Laboratory of Rail Traffic Control and Safety under grant RSC2011ZT001 and by the National High Technology Research and Development Program ("863" Program) of China under grant 2012AA112800,.

REFERENCES

[1] C. Ma, H.Y. Zhang, D.Q. Miao, Improvement of principal curves algorithm and its application in fingerprint skeleton extraction, Computer Engineering and Applications, 46(16), 2010, pp. 170-173.

[2] Y.H. Wang, Z.Y. Shen, Y.Sun, Numerical simulation of characteristic curves of hydraulic turbine based on principal curves, Journal of Hydroelectric Engineering, 28(3), 2009, pp. 181-186.

[3] H. Su, F.G. Huang, An image segmentation method based on AEP and K principal curves, Journal of Harbin Engineering University, 25(6), 2004, pp. 756-760.

[4] D.W Chen, T. Tang, F. Cao, B.G. Cai, An integrated error-detecting method based on expert knowledge for GPS data points measured in Qinghai–Tibet Railway, Expert Systems with Applications, 39(2), 2012, pp. 2220-2226.

[5] D.W Chen, Y.S. Fu, B.G. Cai, Modeling and Algorithms of GPS Data Reduction for the Qinghai–Tibet Railway, IEEE Transactions on Intelligent Transport System, 11(3), 2010, pp. 753-758.

[6] G.G. Gao, B.G. Cai, Research on the Automatic Electronic Map Generation Algorithm for the Train Supervision System, Journal of the China Railway Society, 28(1), 2006, pp. 63-67.

[7] J.P. Zhang, J.Wang, Overview of principal curves, Chinese Journal of Computers, 26(2), 2003, pp.129-146.

[8] T. Hastie, Principal Curves and Surfaces: Stanford University doctoral dissertation (1984).

[9] T. Hastie, W. Stuetzle, Principal curves, Journal of the American Statistical Association, 84(406), 1988, pp. 502-516.

[10] J.D. Banfield, A.E. Raftery, Ice floe identification in satellite images using mathematical morphology and clustering about principal curves, Journal of the American Statistical Association, 87(417), 1992, pp. 7-16.

[11] B. Kegl, A. Krzyzak, A polygonal line algorithm for constructing principal curves, Proceedings of Neural Inform ation Processing Systems, 1999, pp. 501-507.
[12] J.J. Verbeek, N. Vlassis, B. Krose, A k-segments algorithm for finding Principal Curves, Pattern Recognition Letters, 23, 2002, pp. 1009-1017.
[13] D.W. Chen, B.C. Cai, T. Tang, An Information Fusion Algorithm for Multiple GPS Track Data, CA: 2008 Fourth International Conference on Natural Computation, 2008.
[14] J.P. Zhang, D.W. Chen, U. Kruger, Adaptive Constraint K-Segment Principal Curves for Intelligent Transportation Systems, IEEE Transactions on Intelligent Transportation Systems, 9(4), 2008, pp. 666-677.
[15] X.Z. Jia, D.W. Chen, Study on Information Fusion Algorithm for Multiple GPS Railway Tack Data, Journal of the China Railway Society, 33(9), 2011, PP.72-74.
[16] J.P. Zhang, X.D. Wang, U. Kruger, F.Y. Wang, Principal Curve Algorithms for Partitioning High-Dimensional Data Spaces, IEEE Transactions on Neural Networks, 22(3), 2011, pp. 367-380.
[17] C. Zhang, J.Y. Yang, D.C. Yan, S.Q. Yang, Y.T. Chen, Automated Breakpoint Generation for Debugging, Journal of Software, 8(7), 2013, pp. 603-616.
[18] P.F. Li, J.X. Huang, L.X. Ye, Y. Wang, Z.J. Li, D.W. Li, Directional Fuzzy Data Association Filter, Journal of Software, 7(10), 2012, pp. 2286-2293.
[19] C. J. Li, W. L. Jia, Y.Y. Yang, X. W, Adaptive Genetic Algorithm for Steady- State Operation Optimization in Natural Gas Networks, Journal of Software, 6(3), 2011, pp. 452-459.

Shiying Yang was born in 1989. She received the B.S. degree in Electronics and Information Engineering, Beijing Jiaotong University, China in 2012. She is working toward the M.S. degree with the Electronics and Information Engineering, Beijing Jiaotong University. Her research interests include machine learning, principle curve.

Dewang Chen was born in 1976. He received the B.S. degree in Mechanical and Electrical Engineering and the M.S. degree in Control and Automation from Harbin Engineering University in 1998 and 2000, respectively, and the Ph.D. degree in Control Theory and Control Engineering from the Institute of Automation, Chinese Academy of Sciences in 2003. He is a Professor with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University since 2011. His current research interests include intelligent control, machine learning, soft computing, optimization and their applications in intelligent transportation systems and railway systems.

Xiangyu Zeng was born in 1989. He received the B.S. in Electronics and Information Engineering, Beijing Jiaotong University, China in 2011. He is working toward the M.S. degree with the Electronics and Information Engineering, Beijing Jiaotong University. His research interests include machine learning, data mining.

Peter Pudney is a Senior Research Fellow in the Centre for Industrial and Applied Mathematics and the Institute for Sustainable Systems and Technologies at the University of South Australia. He received the B.S. degree and the M.S. degree in Applied Science in Computer Studies from South Australian Institute of Technology, and the Ph.D. degree in Mathematics from University of South Australia. He is currently working on projects related to railway train and crew scheduling, as well as helping with trials of the Freightmiser driver advice system by several railways around the world.

An Architecture Independent Packing Method for LUT-based Commercial FPGA

Meng Yang
State Key Lab of ASIC and Systems, Fudan University, Shanghai, China
Email: mengyang@fudan.edu.cn

Jinmei Lai
State Key Lab of ASIC and Systems, Fudan University, Shanghai, China
Email: jmlai@fudan.edu.cn

A.E.A. Almaini
School of Engineering and the Built Environment, Edinburgh Napier University, Edinburgh, UK
Email: a.almaini@napier.ac.uk

***Abstract*—This paper proposes an efficient architecture independent packing method for commercial FPGA. All specific logics of commercial FPGA such as carry chain arithmetic, x-LUT, are pre-designed into reference circuits according to its architecture. Due to complex architecture of contemporary FPGA, to enumerate all reference circuits in a fine-grain manner is impractical. To overcome this problem, coarse-grain manner is adapted in the approach. By using constraint satisfaction problem technique the proposed method matches pre-designed reference circuits from the given user logic circuit. Transformation from the reference circuit to the pre-packed cluster is simplified by using several specifically designed instructions. In the next stage, those directly connected FFs are absorbed into the pre-packed clusters. The Last stage packs LUTs and FFs into clusters in a delay-based manner. This method is architecture independent and can be applied for any other commercial FPGAs as long as the pre-designed reference circuits are modified accordingly. The results obtained and compared with commercial tool, ISE MAP, and academic tool, PAM MAP, have shown the effectiveness of the proposed method.**

***Index Terms*—Packing, Algorithm, Computer-aided design, FPGA**

I. Introduction

Contemporary commercial field-programmable gate arrays (FPGAs) consist of a cluster of configurable logic blocks (CLBs) formed by look-up tables (LUTs) and flip-flops (FFs) as well as arithmetic circuitry, configurable I/O blocks (IOBs) and specialised hard IP blocks. For example, a SLICE, a half of CLB, in the latest Xilinx Virtex-7 FPGA family device contains four six-input LUTs, eight FFs, carry chain arithmetic logic and other circuitry. It is widely acknowledged that FPGAs are slower, less area-efficient and less power efficient than custom ASICs [1]. However, the programmability of FPGAs, gives them the advantage of short time to market. As a result, they have been widely used in a variety of applications such as domestic communications and automotive electronics.

Packing, which falls between technology mapping and placement, is an extremely important step of the FPGA computer aided design (CAD) flow. This step is most commonly regarded as packing LUTs and FFs together to form clusters [2]. However, in commercial FPGAs, packing is the step that the various logic gates of technology mapped circuit including not only LUTs and FFs but also other logic gates are mapped to FPGA fabric according to the available hardware resources. Packing algorithms are well-studied in the literature for the academic FPGA model, which consists of several basic logic elements (BLEs). Each BLE has one LUT and one FF. The FF can be optionally bypassed for implementing combinational logic only. Local interconnect is available for realising fast paths within the cluster. The output of LUT/FF drives both local interconnect and general interconnect. Inputs to the cluster come from general interconnect [2].

The earliest work based on the academic FPGA model proposed an area-driven packing algorithm (VPack) in the earlier version of versatile placement and routing (VPR) CAD tool [3]. This used the simplest graph pattern match to pack LUTs and registers into BLEs in the first step and packs BLEs into clusters in the second step. Marquardt further extended the previous work carried out by Betz to perform timing-driven packing (T-VPack) [4] and improve speed and density. Recently, Verilog-to-routing (VTR) [5], the latest version of VPR was proposed, in which hardcore IPs are supported in the packing stage.

Tom et al [6] proposed a non-uniform depopulation technique, (Un/DoPack), which runs the FPGA CAD flow twice. First iteration is the regular CAD flow. In the second iteration, packing uses the layout result of the first iteration and depopulates the congested regions. While reducing the channel width, Un/DoPack, similar to the other depopulation-based packing approaches, observes an increase in total area and critical path delay.

T-NDPack [7] proposed an objective cost function with consideration of the criticality in terms of delay and

doi:10.4304/jcp.9.5.1131-1137

routability simultaneously, which consequently reduces the channel width requirements and the depth of the critical path. However, it incurs logic area overhead. It was claimed that minimum channel width and critical path delay were reduced by 11.07% and 2.89% respectively while increasing the number of CLBs by 13.28% compared to T-VPack.

Easwaran et al proposed a routability driven power-aware packing method (W-T-VPack) [8] with introduction of a new packing cost function based on predicted individual net length. It claimed that W-T-VPack outperforms T-RPack [9] and iRAC [10] in terms of energy by 11.23% and 9.07%, respectively.

Rajavel et al proposed a many-objective FPGA circuit packing strategy (MO-Pack) [11] that minimised the channel width and the energy of a circuit implementation without incurring any overhead on critical path delay.

Yang et al proposed a yet another many-objective FPGA packing method (YAMO-Pack) [12]. It claimed that YAMO-Pack outperforms iRAC and MO-Pack in terms of channel width by 38.8% and 42.2%, respectively and in terms of delay by 11.8% and 11.5%, respectively. However, it requires acceptably more CPU time.

All methods mentioned above target the academic FPGA model, which is significantly simpler than that used for commercial FPGAs. Ahmed et al [13] from Xilinx reported an architecture-specific packing for Virtex-5 FPGAs. However, it can only be used for Xilinx FPGA devices. Moreover, Shao, et al developed an area-driven architecture independent PAM MAP algorithm [14]. The architecture they used differs from the academic model, but it targets area reduction only. To our best knowledge, no timing-driven architecture independent packing method has ever been published for commercial FPGA. The remainder of the paper is organized as follows. Section II gives details of Virtex-7 FPGA architecture, which will be used in the experiment for demonstration. Constraint satisfaction packing techniques and specific designed instructions are given in Section III. Section IV discusses comparison results between the proposed method and other tools. Conclusion is then given in Section V.

II. VIRTEX-7 FPGA CLB ARCHITECTURE

To show the complexity of the contemporary commercial FPGA architecture, a virtex-7 FPGA is reviewed in this section. This architecture will be used for evaluation experiment for demonstration purpose. A Virtex-7 logic block, which is referred to as a CLB, comprises two SLICEs (SLICEL and SLICEM) and a switch matrix. SLICEL and SLICEM are exactly identical, except that LUT in SLICEL is used for logic only and SLICEM can be used for implementing memory cells. The switch matrix allows for connections from a SLICE back to the same SLICE, between the two SLICEs, as well as into rows and columns of general interconnect. Each SLICE contains four 6-input LUTs and 8 flip-flops. The LUTs in Virtex-7 are implemented as what Xilinx called true 6-LUTs, rather than being constructed using smaller LUTs that can be optionally combined together via multiplexers. The output of two true 6-LUTs, either in top half of a SLICE or bottom half of a SLICE, can be constructed as one 7-LUT via multiplexer F7MUX. Two 7-LUTs can function in one SLICE at the same time. Besides, two 7-LUTs can be further combined together via multiplexer F8MUX to form an 8-LUT in one SLICE. Both outputs of 7-LUT and 8-LUT can be registered individually. Fig. 1 shows the architecture of a SLICE of Virtex-7 FPGAs.

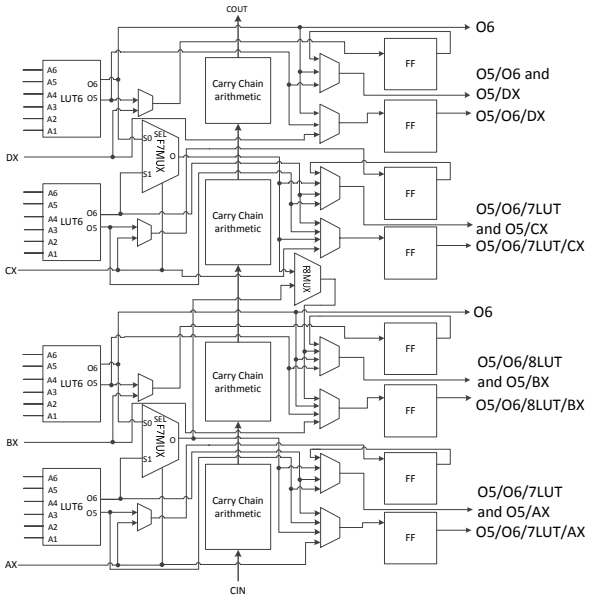


Fig. 1 Virtex-7 SLICE architecture.

III. CONSTRAINT SATISFACTION PROBLEM TECHNIQUE FOR FPGA PACKING

A constraint satisfaction problem [15] is defined by an ordered set of n variables $X=(1,2,\cdots,n)$, a finite domain D_i of possible values for each variable i, and a set of constraints among variables. A constraint $R_{j_1,j_2,\cdots,j_r}$ on the ordered set of variables $(j_1,j_2,\cdots,j_r)$ is a subset of $D_{j_1}\times D_{j_2}\cdots\times D_{j_r}$, which only contains the allowed combinations of values for variables $j_1,j_2,\cdots,j_r$.

An isomorphism of a graph $G_1=(V_1,E_1)$ with a sub-graph of a graph $G_2=(V_2,E_2)$ is equivalent to the constraint satisfaction problem [16]. A variable i is associated with each vertex $v_i\in V_1$, and all variables take values on domain V_2. Let n be the cardinality of V_1. Finding a sub-graph isomorphism is then equivalent to finding a complete assignment satisfying the following structure constraint:

$$R_{i,j}=\left\{\begin{array}{c}(v_a,v_b)\in V_2\times V_2 \mid v_a\neq v_b\wedge edge(G_1,i,j)\\ \Rightarrow edge(G_2,v_a,v_b)\end{array}\right\} \quad (1)$$

$$\text{for all } i,j=1,2,\cdots,n \text{ with } i\neq j$$

Packing problem is similar to isomorphic match problem. A user circuit C can be described by a directed graph $G_1=(V_1,E_1)$, where each vertex $v_i\in V_1$ in G_1 corresponds to a component or a primary input or primary

output in C, and each directed edge $e_i \in E_1$ corresponds to a wire connecting between two different vertexes in C. The set of given circuits is a set of configurable circuits implementing different types of logic functions, which is known as reference circuits from packing point of view and can also be described by directed graphs respectively. Each directed graph $G_2 = (V_2, E_2)$ corresponds to a reference circuit. These configurable circuits are pre-constructed manually according to available FPGA hardware logic resources. Packing algorithm identifies all isomorphic matches in a user design circuit according to a set of given reference circuits.

In order to match reference circuits in a user design circuit, several constraints should be applied. Type constraint should be satisfied for the purpose of matching exact type of vertex in the circuit such as LUT and FF. Start constraint is used for the outgoing edge from a vertex. Similarly, end constraint is for the incoming edge from a vertex. These two constraints are used for matching one particular edge of graph, i.e., from one type of logic gate to another. Input constraint and output constraint are used for primary input and primary output respectively. Shared input constraint identifies shared inputs which is used in the case of more than one sink net shared by two pins.

As long as the reference circuits represent all the functionalities that FPGA hardware resources can implement, it can always find a feasible solution for packing result. However, it is impossible to enumerate all reference circuits for a complex contemporary FPGA, which makes isomorphism packing impractical. Let us consider a case of two 6-LUTs and a 2to1 multiplexer F7MUX forming one 7-LUT in one SLICE. If ignoring sequential outputs, there are 4 cases already, as shown in Fig. 2. Hence, four reference circuits must be constructed in order to match all these patterns. If considering sequential outputs, the number of combination patterns can be increased significantly. It is therefore crucial to select the proper reference circuits, achieving not only less number of reference circuits but also covering all the functionalities a SLICE of contemporary FPGA can implement.

In order to reduce the number of reference circuits, the construction of reference circuits in the proposed method only considers combinational logic. Although the sequential logic is not included in the reference circuits, it will be dealt with after graph pattern match in the second step of packing method. By doing so, it can not only reduce the complexity of the individual reference circuit but the count number of the reference circuits as well. Those different logic functions that behave a similar function are categorized as one function type. For example, there are four different ways to form 7-LUT in one SLICE, as shown in Fig.2(a), Fig.2(b), Fig.2(c) and Fig.2(d), respectively. The graph, shown in Fig. 2(a), is the subset of the graph shown in Fig.2(d). The graphs, shown in Fig.2(b) and Fig.2(c), are also the subset of the graph shown in Fig.2(d). Therefore those four graphs are considered as one function type. One function type accordingly has only one reference circuit. The directed graph of reference circuit is modified by inserting a virtual primary input (VPI) at the input of the vertex and inserting a virtual primary output (VPO) at the output of the vertex, as shown in Fig. 3.

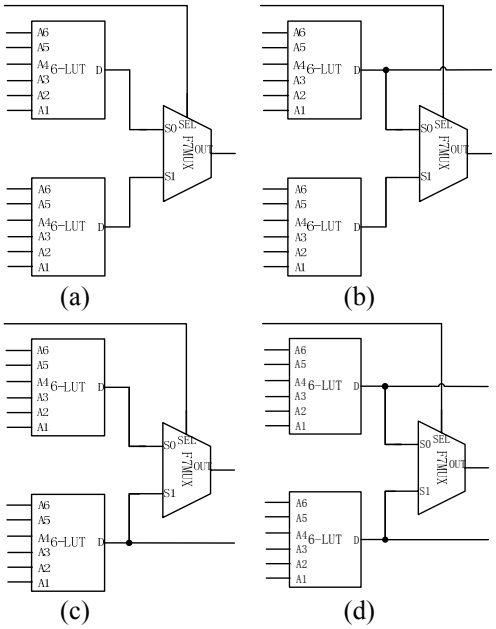


Fig. 2 (a) 7-LUT with single output, (b) 7-LUT with 2 outputs, in which one output is from 7-LUT and the other is from the top 6-LUT, (c) 7-LUT with 2 outputs, in which one output is from 7-LUT and the other is from the bottom 6-LUT, (d) 7-LUT with 3 outputs, in which two outputs are from 6-LUTs and one is from 7-LUT.

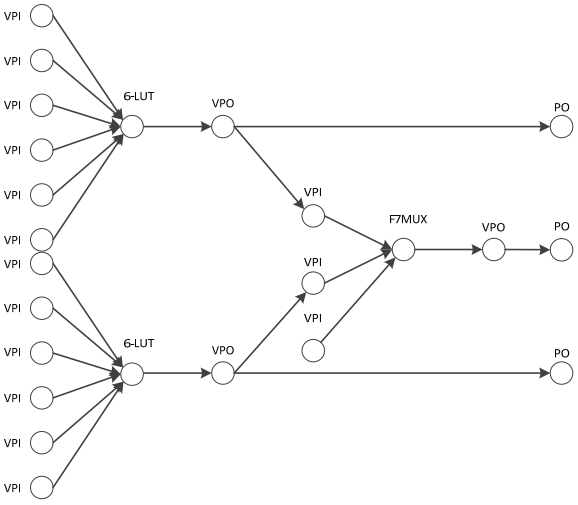


Fig. 3 Directed graph for 7-LUT reference circuit

After a reference circuit is matched from a given user design circuit by utilising graph constraint satisfaction technique, transformation from the reference circuit to the pre-packed cluster process is required. The process for the newly created cluster involves creating a new cluster, wires connection, wires disconnection and specifying configurations such as buffer, MUX, LUT and FF. A key observation is that for a given reference circuit wire connections for the newly created cluster and the configuration settings never alter. In addition, the transformation processes of different reference circuits

are identical. The net connections and the configuration values for different created clusters are different. Therefore each step in the process can be used as an instruction. As a result, the whole process works as executing instruction one after another. For a different specific architecture, reference circuits are different and those reference circuits must be modified accordingly. However, the execution of the instruction is the same for a different architecture. The designed instructions are architecture independent, simple but effective, as shown in TABLE I.

TABLE I
SUMMARY OF INSTRUCTIONS

Instructions	Description
create_instance (inst, type)	Creates a new instance according to its type.
unhook (insta.p1)	Disconnects the pin with name p1 of instance from its net.
connect (inst1.p1, inst2.p1)	Connects the pin with name p1 of instance1 to the net which has pin with name p1 of instance2.
reconnect (inst1.p1, inst2.p1)	Disconnects pin with name p1 of instance1 from its net and connects it to the net which has pin with name p1 of instance2.
xconnect (inst.p1, inst.p2)	Exchanges the net connections of two different pins, p1 and p2, which is used for two pins swap.
set_configuration(inst, value)	Sets one configuration of the instance.
copy_property (inst1,value1, inst2,value1)	Copies value1, which is one property of instance1, to instance2.
set_ property (inst, value)	Sets one property of the instance.
create_instance (inst, type)	Creates a new instance according to its type.
unhook (inst.p1)	Disconnects the pin with name p1 of instance from its net.
connect (inst1.p1, inst2.p1)	Connects the pin with name p1 of instance1 to the net which has pin with name p1 of instance2.

Example 1: Use designed instructions to create a SLICE with functionality of 7-LUT by combining two 6-LUTs and F7MUX. Assume 6-LUT has six inputs A1, A2, A3, A4, A5 and A6 as well as two outputs O5 and O6. F7MUX has three inputs I0, I1 and S as well as one output O. The SLICE has the same architecture as Xilinx-7 FPGA.

1. Create a slice with name slice_a by using instruction create_slice (slice_a, SLICE)
2. Reconnect function generator A6LUT input connections to the newly created slice inputs by using following instructions.

```
reconnect (A6LUT.A1, slice_a.A1)
reconnect (A6LUT.A2, slice_a.A2)
reconnect (A6LUT.A3, slice_a.A3)
reconnect (A6LUT.A4, slice_a.A4)
reconnect (A6LUT.A5, slice_a.A5)
reconnect (A6LUT.A6, slice_a.A6)
```

3. Reconnect function generator B6LUT input connections to the newly created slice inputs by using following instructions.

```
reconnect (B6LUT.A1, slice_a.B1)
reconnect (B6LUT.A2, slice_a.B2)
reconnect (B6LUT.A3, slice_a.B3)
reconnect (B6LUT.A4, slice_a.B4)
reconnect (B6LUT.A5, slice_a.B5)
reconnect (B6LUT.A6, slice_a.B6)
```

4. Reconnect wires to the newly created slice outputs and internal connections by using following instructions, in which A, B, AX and AMUX are SLICE pin name of Xilinx Vertex 7 series family FPGA device.

```
reconnect (A6LUT.O6, slice_a.A)
reconnect (B6LUT.O6, slice_a.B)
reconnect (F7MUX.S, slice_a.AX)
reconnect (F7MUX.O, slice_a.AMUX)
connect (A6LUT.O6,F7MUX.I0)
connect (B6LUT.O6,F7MUX.I1)
```

5. Copy properties from 6-LUTs and set properties by using following instructions, in which "INIT" is the 6-LUT initial value and "NAME" is the 6-LUT instance name.

```
copy_property (A6LUT,INIT, slice_a,A6#LUT)
copy_property (B6LUT,INIT, slice_a,B6#LUT)
copy_property (A6LUT,NAME, slice_a, ANAME)
copy_property (B6LUT,NAME, slice_a, BNAME)
set_property (slice_a, FXLUT::TRUE)
```

6. Set configurations by using following instructions, in which AOUTMUX, AUSED, BUSED are SLICE configurations of Xilinx Vertex 7 FPGA device.

```
set_configuration (slice_a, AOUTMUX::F7)
set_configuration (slice_a, A6#LUT::A6#LUT)
set_configuration (slice_a, AUSED::0)
set_configuration (slice_a, BUSED::0)
```

For the consideration of timing issue, the constraint satisfaction problem technique of graph matching mentioned in early sections is only used for the first stage of the proposed packing. In this stage, only combinational specific logics are matched and packed for a given user design. As a result, the input to the second stage is a netlist consisting of pre-packed combinational clusters, hard IP blocks, LUTs and FFs. In the second stage it packs selected FFs to pre-packed combinational clusters, in which the FF directly driven by the output of the cluster is selected. In other words, if the FF is driven by the output of other block such as a LUT or a FF, this FF is ignored. It is known as FF absorption stage. In the

same way, it repeatedly packs the selected FFs into the cluster until no more FF can be selected for packing. After this stage completes, the netlist consists of pre-packed combinational clusters, pre-packed sequential clusters, hard IP blocks, LUTs and FFs. Final stage deals with LUTs and FFs in a delay-based manner, which is similar to MO-Pack [11] and YAMO-Pack [12], to pack them into clusters.

The pseudo code of proposed algorithm is outlined as follows.

```
Inputs: a netlist of the user design circuit after technology mapping C and a set of
pre-designed reference circuits S
Output: the packed circuit netlist Cp consisting of SLICEs and IP cores
    Convert C to graph Gc
    for all reference circuits r in R do {
      Convert r to graph Gr
      Using constraint satisfaction technique to match a pattern Gr from Gc
      If matched
          Map Gr to SLICE using the pre-designed corresponding instructions
          Update Gc
          update netlist Cp
      endif
    }
    for all SLICEs s in S do {
      for all outputs o in s {
        Trace the output o of s to the data input of FF dff
        If dff exists
            Pack dff to the SLICE s
          modify s configurations
          update netlist Cp
        endif
      }
    }
    unpackedBLEs = use simple pattern match to pack LUTs and FFs to BLEs
    for all BLEs ble in unpackedBLEs do
        calculate ble criticality as in T-VPack
    while (unpackedBLEs != NULL) {
      ble = selected the most critical BLE in unpackedBLEs
      create a new SLICE sn and pack ble to sn
      modify sn configurations
      update netlist Cp
      unpackedBLEs = unpackedBLEs – selectedBLE
      while (sn is not full) {
        selectedBLE = select max attraction BLE in unpackedBLEs
        pack selectedBLE to sn
        modify sn configurations
        update netlist Cp
        unpackedBLEs = unpackedBLEs – selectedBLE
      }
    }
    Output netlist Cp
```

IV. RESULTS

The proposed method is developed under Microsoft Visual Studio 2010 and implemented in C++. The results have been run on the PC with an INTEL CPU 2.4 GHz and 4 GB RAM.

To verify the effectiveness of the proposed method, design circuits in register transfer level Verilog format from the benchmark suite in Quartus II university interface program (QUIP) [17] are chosen. The selected designs are architecture independent and those circuits can be logically optimised by Xilinx commercial logic synthesis tool XST. Xilinx ISE MAP and the proposed method are then applied to the output of XST to pack logic into Xilinx Virtex-7 FPGA SLICE. The device xc7k160t-fbg676-3 is chosen for demonstration. TABLE II shows the comparison of the mapping results.

TABLE II
COMPARISON OF ISE MAP AND OURS IN TERMS OF DELAY

Benchmarks	ISE MAP (ns)	Ours (ns)	Improve-ment (%)
barrel16	2.243	2.164	3.52
barrel16a	2.829	2.706	4.35
barrel32	2.995	2.978	0.57
barrel64	3.495	3.420	2.15
fip_cordic_cla	4.762	4.555	4.35
fip_cordic_rca	4.017	3.841	4.38
fip_risc8	8.243	8.454	-2.56
mux32_16bit	2.268	2.205	2.78
mux64_16bit	3.114	3.038	2.44
mux8_128bit	2.595	2.685	-3.47
mux8_64bit	1.364	1.314	3.67
oc_aes_core	4.885	4.964	-1.62
oc_aes_core_inv	7.73	8.144	-5.36
oc_des_area_opt	2.389	2.262	5.32
oc_des_des3area	3.115	3.260	-4.65
oc_des_des3perf	3.392	3.305	2.56
oc_des_perf_opt	2.423	2.356	2.77
oc_minirisc	6.612	6.381	3.49
oc_miniuart	1.812	1.753	3.26
oc_mips	14.64	14.924	-1.94
oc_rtc	3.312	3.222	2.72
oc_ssram	1.609	1.500	6.77
oc_video_dec	3.223	3.207	0.50
oc_video_enc	2.089	2.164	-3.59
oc_video_jpeg	3.723	3.677	1.24
Average	3.96	3.94	0.51

TABLE III
COMPARISON OF PAM MAP AND OURS

Benchmarks	Area (No. of SLICE)			Delay (ns)		
	PAM MAP	Ours	Imp (%)	PAM MAP	Ours	Imp (%)
barrel16	24	20	-17	2.473	2.164	-12.5
barrel16a	49	41	-16	2.961	2.706	-8.6
barrel32	110	103	-6	3.384	2.978	-12.0
barrel64	140	133	-5	3.842	3.42	-11.0
fip_cordic_cla	120	118	-2	4.945	4.555	-7.9
fip_cordic_rca	74	68	-8	3.961	3.841	-3.0
fip_risc8	134	122	-9	8.454	8.454	0.0
mux32_16bit	136	135	-1	2.225	2.205	-0.9
mux64_16bit	268	259	-3	3.874	3.038	-21.6
mux8_128bit	279	270	-3	2.952	2.685	-9.0
mux8_64bit	140	136	-3	1.456	1.314	-9.8
oc_aes_core	170	158	-7	6.023	4.964	-17.6
oc_aes_core_inv	300	295	-2	8.465	8.144	-3.8
oc_des_area_opt	160	147	-8	2.845	2.262	-20.5
oc_des_des3area	297	285	-4	3.856	3.26	-15.5
oc_des_des3perf	300	280	-7	3.505	3.305	-5.7
oc_des_perf_opt	580	560	-3	2.969	2.356	-20.6
oc_minirisc	180	165	-8	7.023	6.381	-9.1
oc_miniuart	36	34	-6	1.965	1.753	-10.8
oc_mips	1100	1055	-4	16.4	14.92	-9.0
oc_rtc	130	118	-9	3.762	3.222	-14.4
oc_ssram	36	33	-8	1.865	1.5	-19.6
oc_video_dec	134	130	-3	3.723	3.207	-13.9
oc_video_enc	104	96	-8	2.476	2.164	-12.6
oc_video_jpeg	476	465	-2	3.937	3.677	-6.6
Average	219	209	-6	4.374	3.939	-11

It can be seen that the proposed method can achieve comparable results compared to Xilinx ISE MAP. It should be noted that since the proposed method is architecture independent it can be used for Altera FPGA architecture as well as long as the pre-designed reference circuits are modified accordingly to be suitable for Altera FPGA architecture.

Other published methods such as iRAC [9], MO-Pack [11], YAMO-Pack [12] etc are not comparable because they are targeting academic FPGA model. The method presented in [13] is not comparable either, because the test suite used is from industry and not available. Therefore, PAM MAP [14] is chosen for comparison, since PAM MAP is architecture independent and it can target Virtex-7 as well. The comparison results are shown in TABLE III. It can be seen that the proposed method can outperform PAM MAP in terms of area and delay in all tested cases, achieving, on average, 6% and 11% improvement, respectively.

V. Conclusions

The latest FPGAs contain composite logic blocks with LUTs, FFs, MUXs and other arithmetic circuitry. Packing design elements into the available logic resources is an extremely complex problem. In this paper, an architecture independent packing method for the commercial FPGA device is proposed. The proposed method has three stages. In the first stage, the constraint satisfaction problem technique of graph matching is utilised to implement specific logic such as 7-LUT, 8-LUT and carry chain arithmetic logic from the given user design circuit. Second stage packs the selected FFs to pre-packed combinational clusters. In the third stage, the delay-based method is carried out to deal with unclustered LUTs and FFs. The experimental results show that the proposed approach achieves similar performance in terms of speed compared with Xilinx commercial tool ISE MAP. The proposed algorithm also outperforms area-driven architecture independent PAM MAP, which can achieve on average, 6% and 11% in terms of area and speed, respectively.

Acknowledgment

This work was supported by a grant (No. 11MS011) from State Key Lab of ASIC and System, China and the National High Technology Research and Development (863) Thematic Program of China (No. 2012AA012001).

References

[1] I. Kuon, J. Rose, "Measuring the gap between FPGAs and ASICs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol.26, pp. 203–215, 2007.

[2] V. Betz, J. Rose, A. Marquardt, *Architecture and CAD for Deep-Submicron FPGAs*, Kluwer Academic Publisher, 1999.

[3] V. Betz, J. Rose, VPR, "A new packing, placement and routing tool for FPGA research," *Proceedings of the 7th International Workshop on Field-Programmable Logic and Applications*, pp. 213-222, 1997.

[4] A. Marquardt, V. Betz, J. Rose, "Using cluster-based logic blocks and timing-driven packing to improve FPGA speed and density," *Proceedings of the ACM/SIGDA 7th International Symposium on Field Programmable Gate Arrays*, pp. 37-46, 1999.

[5] J. Rose, J. Luu, C. Yu, et al, "The VTR project: architecture and CAD for FPGAs from Verilog to routing," *Proceedings of the ACM/SIGDA 20th International Symposium on Field Programmable Gate Arrays*, pp. 77-86, 2012.

[6] M. Tom, D. Leong, G. Lemieux, "Un/DoPack: reclustering of large system-on-chip designs with interconnect variation for low-cost FPGAs," *Proceedings of the IEEE/ACM 2006 International Conference on Computer-Aided Design*, pp. 680-687, 2006.

[7] H. Liu and A. Akoglu, "Timing-driven nonuniform depopulation-based clustering," *International Journal of Reconfigurable Computing*, vol. 2010, pp. 1-11, 2010.

[8] L. Easwaran, and A. Akoglu, "Net-length-based routability-driven power-aware clustering," *ACM Transaction on Reconfigurable Technology and Systems*, vol. 4, pp. 38:1-16, 2011.

[9] A. Singh, G. Parthasarathy, and M. Marek-Sadowska, "Efficient circuit clustering for area and power reduction in FPGAs," *ACM Transaction on Design Automation Electronic Systems*, vol. 7, pp. 643-663, 2002.

[10] E. Bozorgzadeh, S. O. Memik, X. Yang, and M. Sarrafzadeh, Routability-driven packing: metrics and algorithms for cluster-based FPGAs, *Journal of Circuits, Systems and Computers*, vol. 13, pp. 77–100, 2004.

[11] S. Rajavel, and A. Akoglu, "MO-Pack: Many-objective clustering for FPGA CAD," *Proceedings of the 48th ACM/IEEE Design Automation Conference*, pp. 818-823, 2011.

[12] M. Yang, J.M. Lai and J.R. Tong, "Yet Another Many-Objective Clustering (YAMO-Pack) for FPGA CAD," *Proceeding of the 23rd International Conference on Field Programmable Logic and Applications*, pp. 1-4, 2013.

[13] T. Ahmed, P. D. Kundarewich, J. H. Anderson, at al, "Architecture-Specific Packing for Virtex-5 FPGAs," *Proceedings of the ACM/SIGDA 16th International Symposium on Field Programmable Gate Arrays*, pp. 5-13, 2008.

[14] Y. Shao, J.M. Lai, J. Wang and J.R. Tong, "PAM Map: an architecture-independent logic block mapping algorithm for sram-based FPGAs," *Proceedings of the 5th Southern Conference on Programmable Logic*, pp. 15-19, 2009.

[15] L.P. CordelIa, P. Foggia, C. Sansone and M. Vento, "A (sub)graph isomorphism algorithm for matching large graphs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 1367-1372, 2004.

[16] B.N. Tran, T.D. Nguyen, "An Efficient Algorithm for Isomorphic Problem on Generic Simple Graphs," *Proceedings of the Second Asia International Conference on Modeling Simulation*, pp. 824-829, 2008.

[17] Quartus II University Interface Program. Available: http://www.altera.com.cn/education/univ/research/unvquip.html

Meng Yang received Bachelor of Engineering (Honor) degree in Electrical Engineering from Shanghai University, Shanghai, China, in 1999. He received Master of Science with distinction in Electronics and Communication Engineering and Ph.D. in Electronics from School of Engineering Edinburgh Napier University, Edinburgh, UK, in 2002 and 2006, respectively.

Currently he is a lecturer of State Key Lab of ASIC and System and Department of Microelectronics, School of Information Science and Technology, Fudan University, Shanghai, China. His research interests include algorithms in FPGA design automation, logic synthesis, and dynamic reconfigurable FPGA automation design. He has published more than 30 research papers.

Jinmei Lai received PhD degree in Shanghai Jiaotong University, Shanghai, China, in 1998. She was a Post-Doctor in Zhejiang University and Fudan University.

Currently she is a full professor of State Key Lab of ASIC and System, Fudan University, Shanghai, China. Her research interests include low power and reconfigurable architecture of FPGA and SOC, embedded IP core generation automation, logic synthesis and dynamic reconfigurable FPGA automation design, SOC testing automation. She has published more than 80 research papers and holds dozens of Chinese patents.

A.E.A. Almaini was born in Baghdad where he completed his school education. He received the B.Sc. (Eng), M.Sc., and Ph.D. degrees in electrical & electronic engineering from universities in England. He published a book, Electronic Logic Systems, and over 100 research papers.

Currently he is a Professor Emeritus at Edinburgh Napier University, Edinburgh, UK. His main research interests include the synthesis, optimization and automation in the field of digital electronics.

A Novel Power Amplifier Behavior Modeling Based on RBF Neural Network with Chaos Particle Swarm Optimization Algorithm

Mingming Gao
School of Electrics and Information Engineering, Liaoning Technical University, Huludao, China
E-mail:gaomingming2080@163.com

Jingchang Nan
School of Electrics and Information Engineering, Liaoning Technical University, Huludao, China
E-mail: nanjc886@sina.com

Surina Wang
University of British Columbia, Vancouver, Canada
E-mail: surina.wang@gmail.com

***Abstract*—In order to design and optimize high-linearity power amplifier (PA), which with nonlinear and memory effect, it is very important to build power amplifier behavior modeling accurately. This paper proposes a power amplifier behavior modeling based on RBF neural network with improved chaos particle swarm optimization algorithm. To make the particles evenly distribute in the problem search space, a novel Chaos Particle Swarm Optimization (CPSO) is proposed based on the analysis of the ergodicity of chaos and inertia weight of Particle Swarm Optimization (PSO). Based on circle model, the new model is introduced to avoid PSO from getting into local optimum. This paper uses free scale semiconductor chip MRF6S21140 to carry on amplifier circuit design in the ADS and the MATLAB fitting simulation of the extracted data, by improved CPSO-RBF algorithm. Its accuracy is assessed by comparing RBF modeling with voltage RMS error (RMSE), epochs, and fitting time. The result shows that improved CPSO-RBF has better fitting function.**

***Index Terms*—Power Amplifier; CPSO; Neural Network; Behavioral Model**

I. INTRODUCTION

With the development of modern communication technology, many communication systems use high efficient spectrum techniques. These techniques have a high peak to average power ratio, in order to increase the transfer rate and channel capacity, that is, efficient digital modulation formats (such as M-QAM and OFDM) and new multiple access methods (such as OFDMA, MC-CDMA, and WCDMA). The power amplifier device will produce nonlinearity and memory effects because of these techniques. In order to linearize the nonlinear power amplifier system, an accurate power amplifier behavior model is to be obtained. Thus, the behavior modeling techniques, which can deal with both nonlinear and memory effect, have become one of the hot topics in interdisciplinary research, including: microwave applications, wireless communications, radar, semiconductor physics, nonlinear control, and instrumentation.

In behavioral modeling, for example, one of the challenges of the modeling of nonlinear behavior of RF microwave modules is to precisely describe both strong nonlinearity and memory effects, which have more time on the dynamic characteristics of a constant or combination [1-3]. Power amplifier is a critical nonlinear module in various radio frequency communication systems. In an efficient modulation system, the fact demonstrates that power amplifier is not only characteristic of nonlinearity but also of strong memory effects. Therefore, how to precisely model memory effect in dynamic nonlinear amplifier is an important problem. There are lots of reports on RF power amplifier modeling both at home and abroad, which include memoryless model, Volterra-Series model and its Simplified model, different types of neural network model, and so on [4-6]. Compare to Volterra series model, neural network model has good approximation capabilities. It can better describe the behavior of weak and strong nonlinear amplifier models, and the result can be generalized. One of the advantages of RBF neural network model is that it can be applied to any nonlinear function, and it is also suitable for power amplifier model building. In RBF neural network model, we need to determine the structure parameters, which are the center bits of basis function, the variance, and network weights [7,8]. These parameters will determine the performance of the neural network. Particle Swarm Optimization (PSO) algorithm uses the same speed model as the position search, which is fast, accurate, and concise, and has low computational complexity and easy implementation. Particle Swarm Optimization (PSO) algorithm will search for global

Project number: No.61372058; LR2013012; MTKJ2011-339
Contact author: Mingming Gao

doi:10.4304/jcp.9.5.1138-1143

optimal solution through groups of particles in cooperation and competition. In order to improve the performance of neural network training, researchers at home and abroad used Particle Swarm Optimization (PSO) algorithm to train neural network weights and topology. However, the basic PSO algorithm in slow fitting is close to the optimal solution, which is easy to appear and even to a standstill, and makes the network training difficult to achieve the desired effect. Hence, many scholars proposed a modified PSO algorithm. Currently several improved algorithms have been made, such as adaptive PSO algorithm, hybrid PSO algorithm, collaborative PSO algorithm, discrete PSO algorithm, and immune PSO algorithm. Chaos in a nonlinear phenomenon is widespread in nature: it appears to be chaotic, but has exquisite internal structure; it has randomness, ergodicity, and regularity characteristics, and it is extremely sensitive to initial conditions; it can change according to its own laws within a certain range and will not repeat to loop through all state. These properties can be optimized using chaotic motion search.

This paper is organized as follows: in Section II, the paper proposes a novel chaotic Particle Swarm Optimization algorithm, by combining with the RFB neural network to build power amplifier behavioral modeling. In Section III, the paper proposes a power amplifier behavior modeling based on RBF neural network with improved chaos particle swarm optimization algorithm (CPSO-RBF). In Section IV, the paper shows the simulation results of CPSO-RBF power amplifier modeling has higher precision. Finally, in Section V, a conclusion is presented.

II Power Amplifier Neural Network Model of RBF and the Learning Algorithm

A. The Expression of RBF Neural Network Model of Power Amplifier

This model is the selection of General RBF Neural networks. RBF Neural network is consisted of input layer, hidden layer, and output layer. In this model, input amplifiers of complex signals are converted to amplitude and phase of a polar form, and then are trained on the real-valued amplitude and phase. According to the amplifier's nonlinear characteristics and memory effects of power amplifier, the amplifier output expression is:

$$\begin{aligned} y(l) &= g(|x(l)|,|x(l-1)|,\cdots,|x(l-L)|)\times \\ &\exp\{j[\Phi(l)+f[|x(l)|,|x(l-1)|,\cdots,|x(l-L)|]]\} \end{aligned} \quad (1)$$

$|x(l)|$ and $\Phi(l)$ are the input signal amplitude and phase respectively, L is the memory effect in a memory depth, which is the number of models in the previous sample. Nonlinear power amplifier can be expressed by AM/AM and AM/PM characteristic curves. There are two corresponding output nodes for AM/AM and AM/PM nonlinear distortion functions, $g(\cdot)$ and $f(\cdot)$, that can describe the dynamic characteristics of power amplifier's AM/AM and AM/PM.

RBF Neural network has L+1 entries input notes, which is $X=[|x(l)|,|x(l\text{-}1)|,\ldots,|x(l\text{-}L)|]^T$. It supposes based training samples for n, and it's hidden layers have M ($M<N$) neurons. Any one of the neurons indicated by i, $\varphi(X,C_i)$ is the primary function which for the i^{th} motivation of hidden units in output. Hidden output layer weights can be provided by ω_{ji}. Output unit also set a threshold φ, in order to suppress floor G_0 of a neuron's output as 1, the output unit is attached to the right value for ω_{0i}. The model structure is shown in Figure 1:

Function $\varphi(X,C_i)$ generally selectes basis on Green's

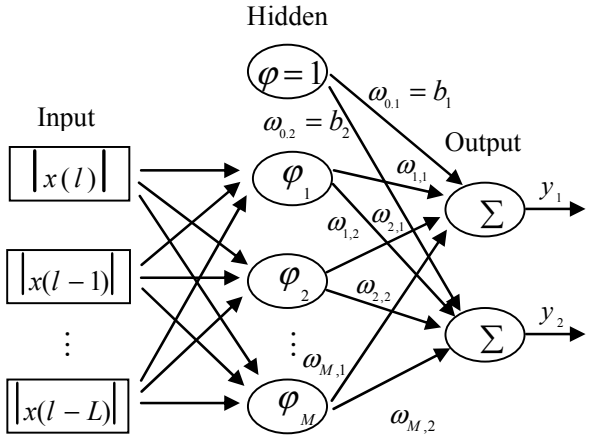


Figure1. RBF Neural Network model block diagram

function, using the definition of Green's functions $G(\|x-c_i\|)=\exp(-\frac{\|x-c_i\|^2}{2\sigma_i^2})$ to indicate the hidden layer of non-linear function $\{\varphi_i(x)=G(\|X\text{-}c_i\|),\ i=1,2,\ldots,M\}$. M is the number of hidden layer of units, X is the input vector, $\{c_i|i=1,2,\ldots,M\}$ is the central point of $G(\cdot)$, and σ_i is the field width of the i^{th} hidden node. The j^{th} (j=1,2) output node of the output is:

$$y_j(x)=\sum_{i=1}^{L}\omega_{ji}G(\|x-c_i\|)+b_j\,. \quad (2)$$

Among them, ω_{ji} is the weight that connects between the hidden layer of i^{th} nerve cell and output layer of the j^{th} neuron, and b_j is the ground term. Smoothness of the approximation is depended on σ_i^2.

The number of real parameters to M(L+3)+2 of RBF Neural network can show the nonlinear dynamic behavior of the power amplifier.

B. Improved Chaos Particle Swarm Optimization Algorithm

Set in the group, the i^{th} Particles is x_i (x_{i1}, x_{i2}, …, x_{id}) with experienced location p_i (p_{i1}, p_{i2}, …, p_{id}), and for individuals the best location is pbest. Currently, all particles that make up the groups have experienced the best location pgbest, and grain i^{th}'s speed can provide by v_i(v_{i1}, v_{i2}, …, v_{id}). On each iteration, the grain i^{th} in d-dimensional ($1\le i\le d$) space the following equations:

$$v_{id}^{k+1} = \omega v_{id}^{k} + c_1 Rand()(p_{id} - x_{id}^{k}) + c_2 Rand()(p_{gbest}^{k} - x_{id}^{k}) . \quad (3)$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k} . \quad (4)$$

In formula (3), ω is the inertia weights, which keeps the particles movement inertia, and ω gives it the ability to explore new areas; *c1* and *c2* are acceleration constants, their values are usually between 1.5-2, and the algorithm takes the value of 2. They allow each particle's accelerated motion to pbest and pgbest locations. *Rand* () represents random numbers, which ranges form (0, 1).

Chaos in a nonlinear phenomenon is widespread in nature, and the more commonly used model is the chaos model of logistic model, whose expression is:

$$X_{n+1} = \mu X_n (1 - X_n) \qquad n = 1, 2, \cdots . \quad (5)$$

The following figure is a performance chart image of logistic, when μ=4. Logistic map is iterated 1000 times within the range of (0, 1) map. Figure 2 shows 0, 0.1, and 0.9, 1 with high probability interval value, where the highest probability point could reach 212 times. However, between 0.1 to 0.9, the average probability point is 76 times. When the optimal value falls between 0.1, and 0.9, we need a number of iterations to get the optimal solution, and this greatly reduces the efficiency of algorithms.[9]

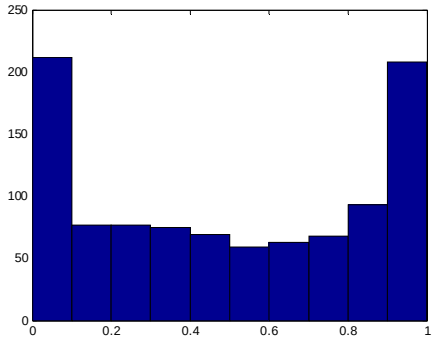


Figure 2. Logistic distribution by iterated 1000 times

However, logistic model produces uneven distribution of chaos, and round mapping model Traverse with good uniformity. The equation is as follows:

$$X_{n+1} = X_n + b - \left[\frac{a}{2\pi}\right] \sin\left[2\pi X_n\right] \bmod 1 . \quad (6)$$

Following figure 3 is the allocation plan of Circle map by iterated 1000 times. Among them, a=0.5 and b=0.2. The probability of 0-1 point is as shown in figure 3. It can be perceived that the maximum number is 151 times, minimum is 63 times, and the average is about 100 times, which is higher than logistic mapping but less than uniform probability distribution.

In order to further improve the adequacy and traversal of chaotic search, this paper presents a new map, which can be written as the following formula:

$$X_{n+1} = (1 * X_n + b - \left[\frac{a}{4\pi}\right] \sin\left[4\pi X_n\right]) \bmod 1 . \quad (7)$$

Figure 4 is the new model, mapping iterative 1000 times and 0-1 range of distribution. You can see from the figure, the maximum value is 115 times and minimum is 94 times. The new model is better than the basic Circle

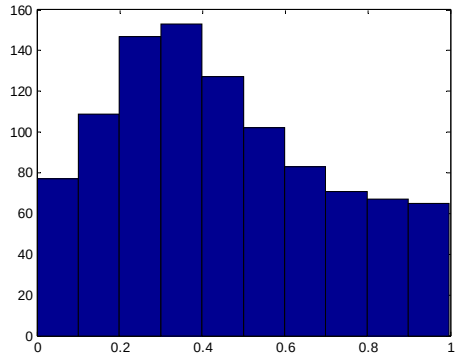


Figure 3. Circle distribution by iterated 1000 times

map and the logistic map on mapping efficiency, and it distributes more evenly.

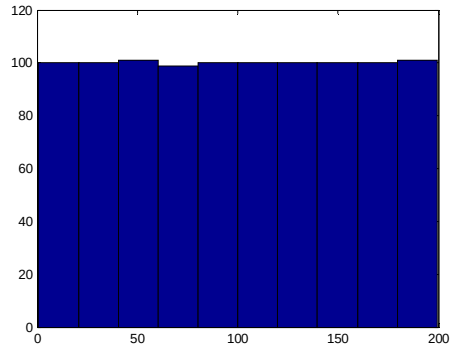


Figure 4. The new map distribution by iterated 1000 times

Chaos Particle Swarm Optimization is mainly reflected in the application of the paper: By using the new model, it can have a uniform distribution of chaos, initialized to population, and inertia weight. $\omega$is an important parameter in Particle Swarm Optimization algorithm, with evolutionary adaptive adjustment of the value of ω, formula sets as follows:

$$X_{n+1} = 1 * X_n + b - \left[\frac{a}{4\pi}\right] \sin\left[4\pi X_n\right] \bmod[1] + 0.2 . \quad (8)$$

In formula (8), K is the number of iterations.

The detailed algorithm is as follows:

Step 1. Initial population. Set the population size to be N, the particle dimension to be D, and assign initial values with small differences to the chaotic equation (7) of i, then we can get a chaotic variable xi; and by changing the ith variable on the interval of Xmin, Xmax map to the location variable values, we can build location variable.

Step 2. Calculate the fitness value d for each particle.

Step 3. Set the individual extreme pbest as the current position of each particle; select the particle with the best

fitness value to be the corresponding global extreme value, gbest.

Step 4. Use formula (8) to calculate the inertia weight value ω, in formula (3).

Step 5. Update the speed and position of the particle according to formulas (3) and (4).

Step 6. The fitness value of each particle will be compared with the existing pbest value. If the new value is better, then replace the existing pbest value, otherwise, retain the original values; select the individual with the optimal fitness in pbest to be gbest.

Step 7. Determine whether the fitting criteria is met, if yes, terminate the optimization process and output the result; otherwise return to step 2.

III THE ALGORITHM OF RBF NEURAL NETWORK WITH IMPROVED CHAOS PARTICLE SWARM OPTIMIZATION

The performance of RBF network is determined by the parameters of the network, which is the center and variance of the basis function as well as network weights [10]. If the chaotic particle swarm algorithm is used for neural network training, then it is not easy for particles to get into the local optimum [11-12]. This algorithm can also expand the search space, search the global optimal solution, and speed up the fitting rate of the neural network training algorithm. Specific optimization (improved CPSO-RBF) procedures are as follows:

a) Preprocess the sample, and normalize the sample data value to the range between 0 and 1.

b) Initialize the network structure; then the parameters of w_i, c_i, σ_i will constitute particles and give them random values to initialize the size, location, and speed of the particle swarm.

c) After getting the input/output response value of RBF neural network, calculate the fitness value of particle swarm according to fitness value formula of fitness to determine the optimal value of individual and population. N denotes the number of the training samples, D denotes the number of output neurons, y_{ij} and t_{ij} denote the output value, and the expected output value of the j^{th} component of the i^{th} sample respectively.

d) Update the position and speed of population particles according to the formulae (3) and (4), and then produce new particle swarm.

e) Judge whether the result meets the optimization goal or maximum number of training, if it meets the termination conditions, then terminate the algorithm; otherwise returns to c.

IV THE SIMULATION RESULTS OF POWER AMPLIFIER BEHAVIOR MODELING

It uses the MRF6S21140 semiconductor transistor of free scale for power amplifier circuit design. Based on the design of the circuit, CDMA2000 source is used for enveloping simulation. The amplifier circuit schematic is shown in figure 5. Figure 6 shows that due to the inherent nonlinear of power amplifier, there is a certain degree of distortion of output signal corresponding to the original input signal of power amplifier. In order to modeling power amplifier, we must take the waveforms of input/output voltage amplitude. In the process of the neural network training, it needs to use the amplifier's input voltage amplitude as neural network input and the amplifier's output amplitude as expected output of the neural network.

Choose the input/output of the power amplifier with the greatest amplitude to build the new model. The experiment uses 90 data points to facilitate the simulation of neural network, and we select memory depth M=3.

It extracts 200 groups of input and output voltage amplitude values from the power amplifier circuit design, compares the fitting degree of input/output voltage amplitude of RBF, and improves CPSO_RBF power amplifier models according to the selected voltage from amplifier circuit. Analyze the simulation results and fit the simulation results. The results include: voltage amplitude output value and output error. It shows the curves of fitting result of voltage range based on RBF model in figure 7, the curves of fitting result of voltage

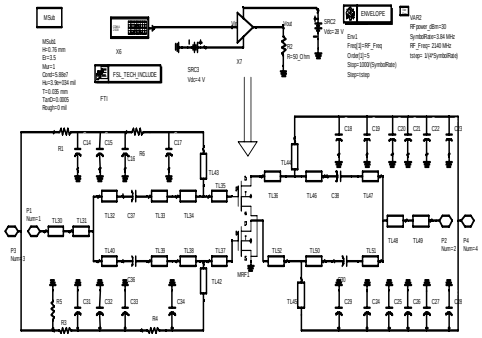


Figure5. Power amplifier circuit schematic

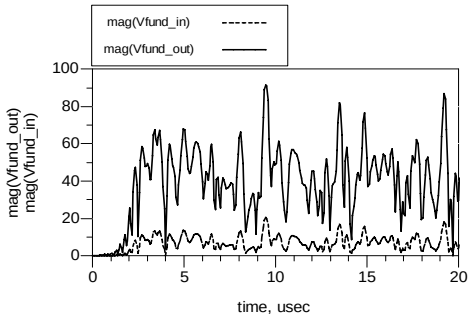


Figure 6. The input and output voltage of power amplifier

range based on improved CPSO-RBF model in figure 8, fitness curve and epoch curve of improved CPSO-RBF model in figure 9 and figure 10. The specific simulation data are shown in table 1.

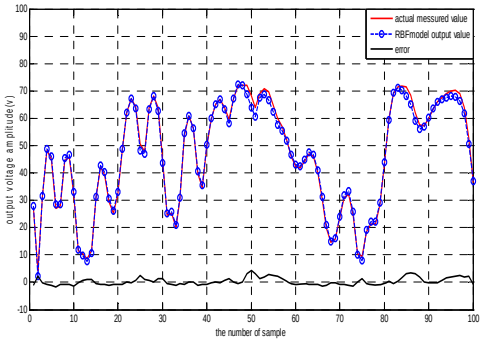


Figure 7. Fitting result of voltage range based on RBF model

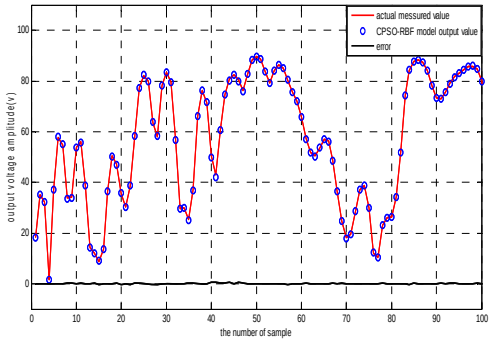


Figure 8. Fitting result of voltage range based on CPSO-RBF model

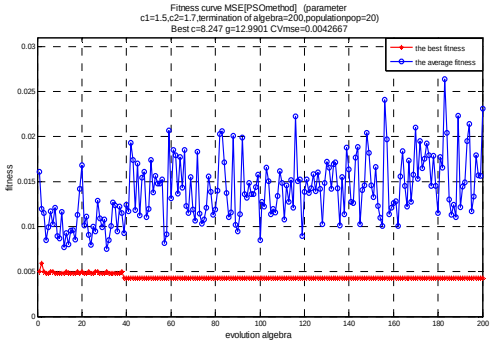


Figure 9. Fitness curve of improved CPSO-RBF model

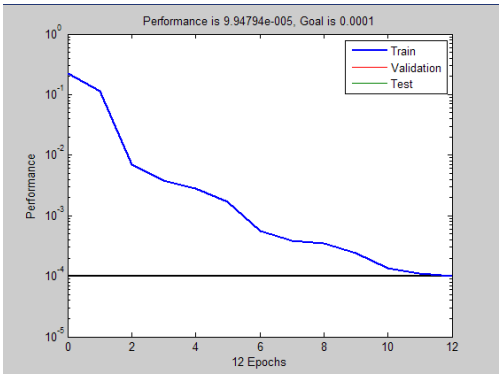


Figure 10. Epoch curve of improved CPSO-RBF model

TABLE I.
THE SIMULATION RESULTS OF RBF AND CPSO-RBF MODELING BASE ON VOLTAGE AMPLITUDE

Items	RBF	CPSO-RBF
RMSE	0.0042	0.0027
Epochs	12	20
Fitting time(s):	2.00	3.00

By contrasting figure 7 and figure 8, we can see that improved CPSO-RBF modeling has better fitting effect than RBP modeling. Improved CPSO-RBF modeling can simulate the power amplifier more accurately. It optimizes RBF neural network weights for global by improved CPSO algorithm, and overcomes the problems in training accuracy and speed of RBF. By comparing the actual output voltage with the simulated output voltage map, the simulated data of improved CPSO-RBF modeling is close to the measurement and is quite consistent. The effect of memory is also being considered in the simulation process, and thus, this modeling can simulate the nonlinear and memory effects of the power amplifiers. Improved CPSO-RBF modeling, which increases training frequency, has longer time-consumption. Voltage RMS error (RMSE) is compared between improved CPSO-RBF modeling output and RBF modeling output of amplifier to validate the model accuracy. Also, we compare epochs and fitting time to validate the modeling training speed. Obviously, this error can be controlled in a small numerical range, and fitting time can be maintained in a short time.

The output power spectrum diagram of the amplifier model of improved CPSO-RBF modeling and RBF modeling is shown in figure 5 (a), and figure 5 (b) is the local enlarged spectrum diagram of figure 5 (a). In figure 5 (b), curve ① is for actual output power spectrum of amplifier, curve ② is output power spectrum for RBF model calculation, and curve ③ is the output power spectrum for improved CPSO-RBF model calculation. It can be seen that improved CPSO-RBF model's output power spectrum of amplifier is better and closer to the actual power spectra model performance by the simulation results.

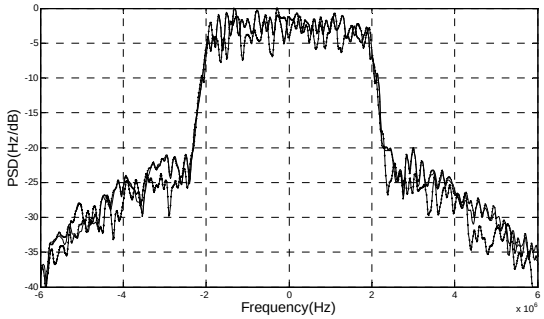


Figure 11. The output power spectrum diagram

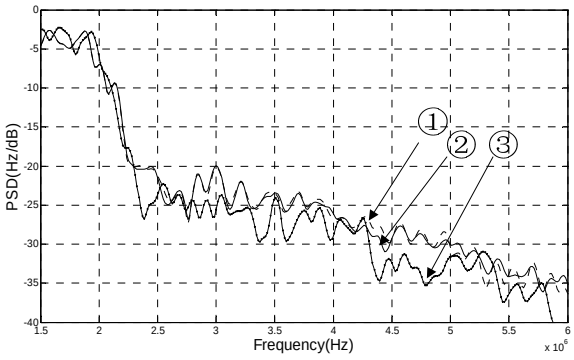


Figure 12. The local enlarged spectrum diagram

V CONCLUSIONS

Building a precise power amplifier behavior modeling is important for building simulating models, and it is crucial for practical designing purpose. This paper extends from RBF power amplifier behavior modeling, and by improving the calculation of PSO and network weights of RBF, it builds and simulates the improved CPSO_RBF modeling. In conclusion, we can get great and precise results from the improved CPSO_RBF, and it can describe nonlinear and memory effects of the power

amplifiers precisely. Hence, this modeling could be used for the linearization of power amplifier system, and it is crucial for practical designing purpose.

ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China (No.61372058), Talent fund projects in Liaoning province (LR2013012), China coal industry association of science and technology research plan projects(MTKJ2011-339)and the earlier stage funding for the natural science foundation of the Liaoning Technical University in 2011.

REFERENCES

[1] J. C. Pedro et al, "A Comparative Overview of Microwave and Wireless Power-Amplifier Behavioral Modeling Approaches," IEEE Transactions on Microwave Theory and Techniques, Vol.53, No.4, April 2005.

[2] Zhang Jing, He Song-bai and Gan Lu. Design of a memory polynomial predistorter for wideband envelope tracking amplifiers [J]. Journal of Systems Engineering and Electronics, 2011, 22(2): 193-199.

[3] Hui Feng, Zeqi Yu. "The Correction Method for Power Noise in Digital Class D Power Amplifiers", Journal of Software, Vol. 8, No. 2, pp:488-494,2013.

[4] Farsaei A R, Safian R. An effective method for generating initial condition in harmonic balance analysis using method of nonlinear currents[C]. Microwave Conference, Dec.2009:1501-1504.

[5] O. Hammi, F. M. Ghannouchi, and B. Vassilakis, A compact envelope memory polynomial for RF transmitters modeling with application to baseband and RF-digital predistortion [J], IEEE Transactions on Microw.Wireless Compon.Lett., 2008, 18(5):359–361.

[6] NAN Jing-chang, GAO Ming-ming; LIU Yuan-an, "Analysis and Comparison of Behavioral Models for Nonlinear RF Power Amplifier, " Journal of Microwaves, 2008.S1

[7] LI Jiu-chao; NAN Jing-chang, LIU Yu-an, "Study of Memory Effects on Power Amplifier for Communication System," Computer Simulation, 2010.7.

[8] WANG Yan, SUN Xiang feng, "LI Ming, Training Method for Support Vector Machine Based on Chaos Particle Swarm Optimization," Computer Engineering, Vol. 36, No. 23, 189-191, December 2010.

[9] Zhihui Zhan, Jun Zhang, Yun Li, et al. Adaptive Particle Swarm Optimization [J]. Transctions on Systems, Man and Cybernetics, 39(6), 2009:1362-1381.

[10] TIAN Yubo, LI Zhengqiang, ZHU Renjie, "Selective neural network ensemble methods based on chaos PSO," Computer Applications, Vol. 28, No. 11, 2844-2846, Nov. 2008.

[11] Liangyou Shu, Lingxiao Yang." A Modified PSO to Optimize Manufacturers Production and Delivery" ,Journal of Software, Vol. 7, No.10, pp:2325-2332,2012.

[12] Xuewen Xia, Jingnan Liu, Yuanxiang Li." Particle Swarm Optimization Algorithm with Reverse-Learning and Local-Learning Behavior", Journal of Software, Vol. 9, No.2, pp:350-357,2014.

Mingming Gao was born in 1980, China. She is a Ph.D. candidate in Liaoning Technical University. Also she is a lecturer in School of Electronic and Information Engineering, Liaoning Technical University. She obtained her Master Degree from Liaoning Technical University in 2008. She is interested in wireless communications, coalmine disaster forecast and communication, etc.

Jingchang Nan was born in 1971, China, a professor from Liaoning technical university. He received the B.Eng degree in industrial electric antomation from Liaoning technical university, Fuxin City, Liaoning Province, China, in 1993, and received the M.S (signal and information processing) and Ph.D degrees (electromagnetic field and microwave technology) from Liaoning technical university and Beijing university of post and telecommunication, in 2003 and 2007, respectively.

He began teaching in Liaoning technical university from 1993, and was engaged in teaching and research of communication engineering in 1994. So far, he authored and co-authored over 40 papers in international conferences and journals, including ICCP, MAPE, ICMMT, J. of electromagn. Wave and appl., J. of electronic and information, High technique letters, and so on. He has presided over some important subjects and researches, including the national natural foundation of China, Dr start fund project of science and research for Liaoning province. He is interested in RF circuit and system, communication system and simulation, adaptive signal and information processing.

Surina Wang, was born in 1983, Canada, she is currently a Master's student in University of British Columbia, Vancouver, Canada, and her focus is Math Education and Pedagogy. She she completed both her secondary and post secondary studies in Canada as well. She received two bachelors degrees in University of British Columbia, a B.Sc. majored in Mathematics and a B.Ed. majored in Secondary Mathematics. She is interested in Mathematics.

An Abnormal Crowd Behavior Detection Algorithm Based on Fluid Mechanics

Xiaofei Wang
College of Electronics and Information Engineering, Sichuan University, Chengdu, China
Arts & Sciences College of Sichuan Normal University, Chengdu, China
Email:scu_xfwang@163.com

Mingliang Gao
School of Electrical & Electronic Engineering, Shandong University of Technology, Zibo, China
Email: sdut_mlgao@163.com

Xiaohai He, Xiaohong Wu, Yun Li
College of Electronics and Information Engineering, Sichuan University, Chengdu, China
Email: hxh@scu.edu.cn, wxh@scu.edu.cn, 645291027@qq.com

***Abstract*— Abnormal crowd behavior detection is an advanced topic researched in fields of computer vision and digital image processing. The problems such as diversity of monitoring scene, different crowd density and mutual occlusion among crowds etc result in a low recognition rate for abnormal crowd behavior detection. In order to solve these problems, this paper combines a streakline model based on fluid dynamics with an abnormal behavior detection method presented by Hassner et al., and proposes a modified algorithm to improve the recognition accuracy of abnormal crowd behavior. Finally, the validity and accuracy of the algorithm are verified via a large amount of challenging real-world surveillance videos.**

***Index Terms*—Streakline, Streak flow, Abnormal crowd behavior detection, Support vector machine**

I. INTRODUCTION

With the development of social economy, density of population in the city is much higher. The occurrence of abnormal crowd behaviors such as group fight and illegal gathering will pose a hazard to social public security. Therefore, abnormal crowd behavior detection becomes a topic in video monitoring with great value for study. However, it is difficult to detect abnormal crowd behavior owing to the following factors: large numbers of crowd targets, different velocity of crowd objects, occlusion among objects, difficulty in eliminating background interference, obvious variances in different crowd objects, etc. At present, scholars at home and abroad have made many achievements on abnormal crowd behavior detection. Wang et al [1] applies improved spatial-temporal characteristics with adaptive sizes to describing the crowd behavior. Kratz et al [2] makes use of a gradient-based spatial-temporal model to describe motion information of the scene and realizes the local abnormal crowd behavior detection using the HMM. Xu et al [3] extracts crowd behavior characteristics using LBP-TOP algorithm, builds a LDA model through training the texture characteristics, and finally detects the local abnormal crowd behavior. Algorithms above which mainly adopt spatial-temporal characteristics, gradient characteristics or texture characteristics to describe crowd behaviors for local abnormal crowd behavior detection, but they lack of description on global characteristics of crowd behavior and cannot represent the motion information of crowd scene completely. Therefore, understanding the crowd behaviors in whole scene, without knowing the actions of individuals, is often advantageous. Recently, some approaches based on hidden Markov model [4, 5], Lagrangian coherent structures [6], social force model [7], Markov random field [8] ,chaotic invariants [9] and kinematic features[10] have been proposed to detect the abnormal behavior in crowded scenes and can provide excellent performance on some benchmarks [7, 9]. However, in the situations where the video has very low resolution, camera jitter, or the speed of objects in the video is too fast or too slow, etc., they may fail to detect the abnormal behavior in crowded scenes. The abnormal crowd behavior recognition method based on Violence Flows descriptor proposed in Ref. [11] begins with the computation of optical flow between consecutive frames and well adapts to the above video set, but the recognition accuracy needs to be enhanced. Comparing with the optical flow computation method, the streakline model proposed by Mehran et al. [12] can represent spatial and temporal changes of flow in crowded scenes more accurately. With the rapid development of crowd behavior analysis, this model has drawn a great deal of attention recently [13-20].

The aim of this work is to devise an abnormal crowd behavior detection algorithm that can perform well in

Corresponding author: Xiaohai He. E-mail: hxh@scu.edu.cn

doi:10.4304/jcp.9.5.1144-1149

challenging real-world surveillance videos with high crowd density (as shown in Fig. 1). For the sake of model with the method in Ref. [11] and propose an improved method for detecting abnormal crowd behavior. The remainder of this paper is organized as follows. Section Ⅱ describes streakline model briefly. In Section Ⅲ, we present the process of our improved abnormal crowd behavior detection algorithm. Experimental results and analysis are provided in Section Ⅳ, and Section Ⅴ concludes this paper.

Figure 1. Examples of normal (top-right) and abnormal (bottom-left) crowd behavior in real-world videos.

II. STREAKLINE MODEL

In fluid mechanics and flow visualization [21], streakline is well known as a tool for measurement and analysis of the flow. A streakline is the collection of all particles which are initialized at a particular pixel. With regard to fluid mechanics, a streakline can be defined as locations of all particles at a given time passing through a particular point. It can be computed by initializing a set of particles at every time instant in the field and propagating them with time based on the optical flow field. This results in a set of paths, each belonging to one point of initialization.

To explain how streaklines are calculated, let $(x_i^p(t), y_i^p(t))$ be a particle position at time t, initialized at point p and frame i for i, $t = 0,1,2,...,T$. The position of particle through point p at any time instant t is computed as described in Ref. [12].

$$\begin{aligned} x_i^p(t+1) &= x_i^p(t) + u(x_i^p(t), y_i^p(t), t) \\ y_i^p(t+1) &= y_i^p(t) + v(x_i^p(t), y_i^p(t), t), \end{aligned} \tag{1}$$

where u and v are optical flow field. This brings a set of curves, all beginning with point p. In order to represent the streaklines more clearly, Fig. 2 gives the example of streaklines showing on every 10th row and column.

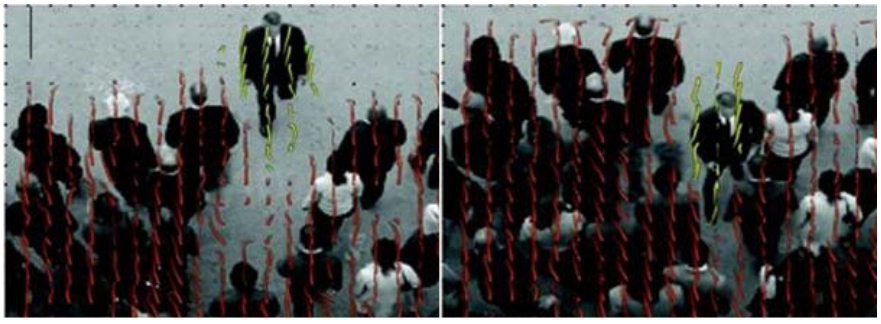


Figure 2. A visualization of the strealines.

For the purpose of fluid visualization, streaklines can transport a color material along the flow, and propagate changes in the flow along their path. Similarly, the streaklines are allowed to propagate velocities by the instantaneous optical flow $\Omega = (u, v)^T$ at the time of initialization, along the flow like a material. To this end, an extended particle i as a set of position and initial velocity is defined as follows

$$P_i = \{x_i(t), y_i(t), u_i, v_i\}, \tag{2}$$

where $u_i = u(x_i^p(i), y_i^p(i), i)$ and $v_i = v(x_i^p(i), y_i^p(i), i)$.

To represent the flow more completely in the whole scene, a new motion field, named as streak flow ($\Omega_s = (u_s, v_s)^T$), is constructed based on streaklines. Streak flow can provide motion information of the flow for a period of time and capture crowd motions better in a dynamically changing flow. The computation of u_s is described as follows, and the computation of v_s is similar. Given data in the vector

$$U = [u_i] \tag{3}$$

where $u_i \in P_i$, $\forall i, p$, the streak flow in the x direction at each pixel is computed.

Equation (1) implies that each particle P_i has three neighboring pixels (nearest neighbors). It is reasonable to consider u_i being the linear interpolation of the three neighboring pixels. Hence, the definition of u_i is as follows

$$u_i = a_1 u_s(k_1) + a_2 u_s(k_2) + a_3 u_s(k_3) \tag{4}$$

where k_j is the index of a neighboring pixel, and a_j is the known basis function of the triangulation of the domain for the j-th neighboring pixel [12]. Each $u_s(k_i)$ is computed using a triangular interpolation formula. For all the data points in U, a linear system of equations is formed using (4)

$$Au_s = U \tag{5}$$

where a_i are entries of the matrix A, and u_s is the least square solution of (5). More details can be found in Ref. [12].

III. Improved Abnormal Crowd Behavior Detection Algorithm

In the Ref. [11], for a given sequence of video frames $Seq = \{f_1, f_2, \cdots, f_T\}$, the optical flow between consecutive frames is firstly calculated. Compared with the optical flow, the streak flow with strong capacity of resisting disturbance [12] described in section Ⅱ is better able to describe information of the whole scene accurately. Therefore, in our algorithm, we substitute the optical flow vector in the original with the streak flow vector $(u_{s_{x,y,t}}, v_{s_{x,y,t}})$ so as to describe the motion information of the whole scene. x, y, t respectively represent location and coordinate of pixel $p_{x,y,t}$ as well as subscript of the video frame, $u_{s_{x,y,t}}$ denotes the value of streak flow at $p_{x,y,t}$ in x direction and $v_{s_{x,y,t}}$ denotes the value of streak flow at $p_{x,y,t}$ in y direction. Refer to Ref. [11], only the magnitudes of these vectors are considered:

$$m_{x,y,t} = \sqrt{(u^2{}_{s_{x,y,t}} + v^2{}_{s_{x,y,t}})} \tag{6}$$

Although flow vectors encode meaningful temporal information, their magnitudes are arbitrary quantities. In order to describe the scene information better, the similarities of flow-magnitudes is considered. For each pixel, we obtain a binary indicator $B_{x,y,t}$, reflecting the significance of the change of magnitude between frames:

$$B_{x,y,t} = \begin{cases} 1 & \text{if } \left| m_{x,y,t} - m_{x,y,t-1} \right| \geq thr \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where thr is the mean of the magnitude variations of the streak flow between frames, shown as follows:

$$thr = \frac{\sum_{x=0,y=0}^{x=rows,y=cols} (|\, m_{x,y,t} - m_{x,y,t-1} \,|)}{rows \times cols} \tag{8}$$

where *rows* and *cols* respectively indicates the number of rows and columns.

Equation (7) provides us with a binary, magnitude-change, significance map for each frame. For each pixel, a mean magnitude-change map by simply averaging these binary values is computed as follows

$$\overline{B}_{x,y} = \frac{1}{T} \sum_t B_{x,y,t}. \tag{9}$$

For a given sequence of frames Seq, our improved descriptor is a vector of frequencies of quantized values $\overline{B}_{x,y}$. Each such vector is then classified as representing an either abnormal or normal video using support vector machines (SVM). In order to understand the detection process better, the architecture of our method is illustrated in Fig. 3. Detailed experimental results and the accuracy comparison will be discussed in section Ⅳ.

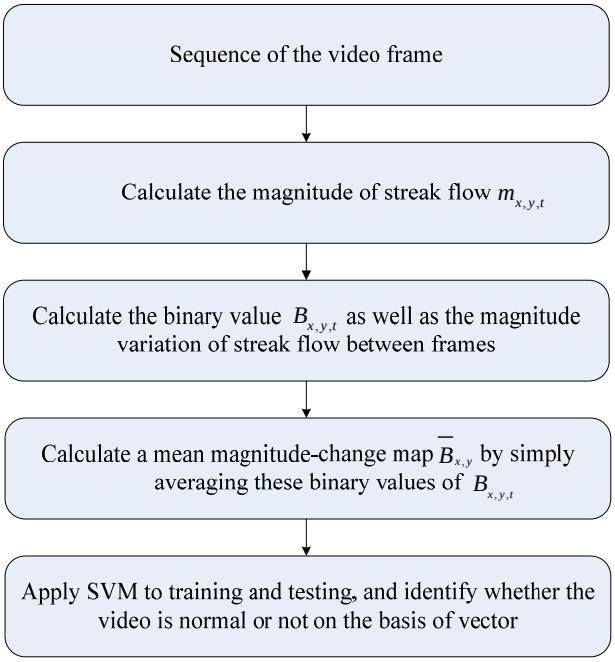


Figure 3. The architecture of our improved method

IV. Experiments And Discussions

Some experiments are done to verify the validity and accuracy of our improved method for abnormal crowd behavior detection by comparing with the algorithm in Ref. [11]. The experiments are conducted on an Intel(R) Core(TM) 2 Duo 2.93GHz with 2GB memory and the Windows XP operating system. The software platform is built in Visual studio 2008. The VIF database [22] is used, in which video is artificially divided into video clips including normal behaviors and abnormal behaviors. For the accomplishment of target proposed in section Ⅰ, 100 video clips including 50 normal videos and 50 abnormal videos are selected from the database. In the experiment, we choose 30% of normal and abnormal videos respectively as a training set, and rest 70% as a test set. The mean prediction accuracy (ACC) is adopted as an assessment index for the algorithm, according to ROC curve theory,

$$ACC = \frac{TP + TN}{P + N}. \quad (10)$$

where TP is the number of true positive samples, TN is the number of true negative samples, P is the number of positive samples and N is the number of negative samples.

The accuracy comparison of different algorithms according to (10) is shown in table 1. From this table, it is shown that the algorithm in our paper can get a higher accuracy on abnormal crowd behavior detection. Recognition results of partial normal videos obtained by using our algorithm for detecting abnormal behaviors are shown in Fig. 4, marked as "Normal", while recognition results of partial abnormal videos acquired by using our algorithm for detecting abnormal behaviors are shown in Fig. 5, marked as "Abnormal".

TABLE 1:
ACCURACY COMPARISON OF DIFFERENT ALGORITHMS

Method	Accuracy
Ref. [11]	75.1 %
Our method	92.3 %

Figure 4. Recognition results of partial normal videos

Figure 5. Recognition results of partial abnormal videos

V. Conclusion

In order to improve the accuracy of abnormal crowd behavior detection, an improved algorithm combining the streakline model based on fluid dynamics with Hassner's method is proposed in our paper. The algorithm adopts magnitude variation of the streak flow to describe the crowd behavior in the scene, and combines with SVM to detect and identify abnormal crowd behavior. Finally, many challenging real-world videos are used to validate our improved algorithm. Experimental results show that our method can perform better in the accuracy of abnormal crowd behavior detection.

Acknowledgement

This work is supported by NSAF under Grant No. 11176018.

References

[1] B. Wang, M. Ye, X. Li et.al, "Abnormal crowd behavior detection using size-adapted spatio-temporal features," *International Journal of Control Automation and Systems*, vol. 9, pp. 905-912, 2011.

[2] K. kratz Lnishino, "Anomaly detection in extremely crowded scenes spatio-temporal motion pattern models," in: *Proc. CVPR*, pp. 1446-1453, 2009.

[3] J. Xu, S. Denman, C. Fookes et.al, "Unusual event detection in crowded scenes using bag of LBPs in spatio-temporal patches," in: *Proc. DICTA*, pp. 549-554, 2011.

[4] E. L. Andrade, S. Blunsden, and R. B. Fisher, "Hidden markov models for optical flow analysis in crowds," in: *Proc. ICPR*, pp. 460-463, 2006.

[5] E. L. Andrade, S. Blunsden, and R. B. Fisher, "Modeling crowd scenes for event detection," in: *Proc. ICPR*, pp. 175-178, 2006.

[6] S. Ali and M. Shah, "A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," in: *Proc. CVPR*, pp. 1-6, 2007.

[7] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in: *Proc. CVPR*, pp. 935-942, 2009.

[8] J. Kim and K. Grauman, "Observe locally, infer globally: A space-time MRF for detecting abnormal activities with incremental updates," in Proc. *CVPR*, pp. 2921-2928, 2009.

[9] S. Wu, B. E. Moore and M. Shah, "Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes," in: *Proc. CVPR*, pp. 2054-2060, 2010.

[10] S. Ali and M. Shah, "Human action recognition in videos using kinematic features and multiple instance learning," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 32, pp. 288-303, 2010.

[11] T. Hassner, Y. Itcher and O. Kliper-Gross, "Violent flows: Real-time detection of violent crowd behavior," in: *Proc. CVPRW*, pp. 1-6, 2012.

[12]R. Mehran, B. E. Moore and M. Shah, "A streakline representation of flow in crowded scenes," *in Proc. ECCV*, 439-452, 2010.

[13] N. M. Nayak, B. Song and A. K. Roy-Chowdhury, "Dynamic modeling of streaklines for motion pattern analysis in video," in: *Proc. CVPRW*, pp.39-46, 2011.

[14] N. M. Nayak, A. T. Kamal and A. K. Roy-Chowdhury, "Vector field analysis for motion pattern identification in video," in: *Proc. ICIP*, pp. 2089-2092, 2011.

[15] X. M. Zhao and G. Medioni, "Robust unsupervised motion pattern inference from video and applications," in: *Proc. ICCV*, pp. 715-722, 2011.

[16] S. Wu and H. S. Wong, "Joint segmentation of collectively moving objects using a bag-of-words model and level set evolution," *Pattern Recogn.*, vol. 45, pp. 3389-3401, 2012.

[17] N. Scherf, C. Ludborzs, K. Thierbach et al, "A fluid registration approach to cell tracking," *Curr. Med. Imaging Rev.*, vol. 9, pp.102-112, 2013.

[18] N. M. Nayak, Y. Y. Zhu and A. K. Roy-Chowdhury, "Vector field analysis for multi-object behavior modeling," *Image Vision Comput.*, vol. 31, pp. 460-472, 2013.

[19] Y. Hu, Y. Zhang and L. S. Davis, "Unsupervised abnormal crowd activity detection using semiparametric scan statistic," in: *Proc. CVPRW*, pp. 767-774, 2013.

[20] X. F. Wang, X. M. Yang, X. H. He, et al, "A high accuracy flow segmentation method in crowded scenes based on streakline," *Optik*, vol. 125, pp. 924-929, 2014.

[21] J. J van Wijk, "Image based flow visualization," in: *Proc. CGIT*, pp. 745-754, 2002.

[22]http://www.openu.ac.il/home/hassner/data/violentflows/index.html.

Xiaofei Wang is currently pursuing a doctor degree at College of Electronics and Information Engineering of Sichuan University. His main research interests include analysis of motion and image processing.

Mingliang Gao a faculty member of School of Electrical & Electronic Engineering at Shandong University of Technology. His main research interests include meta-heuristic optimization, analysis of motion, and tracking in image sequences.

Xiaohai He received his Ph.D. degree in Biomedical Engineering in 2002 from Sichuan University, China. He is a professor in College of Electronics and Information Engineering at Sichuan University. Dr. HE is a senior member of the Chinese Institute of Electronics. He is an Editor of both the Journal of Information and Electronic Engineering and the Journal of Data Acquisition & Processing. His research interests include image processing, pattern recognition, image communication.

Xiaohong Wu received Ph.D. degrees in school of electronics and information engineering from the Sichuan University in 2008. She is an associate professor in school of electronics and information engineering from the Sichuan University. Research interests include image processing, computer vision, statistical pattern recognition and circuit and system.

Yun Li is pursuing a master's degree at the College of Electronics and Information Engineering of Sichuan University, Chengdu, China. The main research area is analysis of abnormal motion, and object tracking in image sequences.

Mining Frequent Closed Patterns using Sample-growth in Resource Effectiveness Data

Lihua Zhang[1,2], Miao Wang[2,3,*], Zhengjun Zhai[1], Guoqing Wang[1,2,3]
[1]School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an, China, 710072
[2]Science and Technology on Avionics Integration Laboratory, Shanghai, China, 200233
[3]China National Aeronautical Radio Electronics Research Institute, Shanghai, China, 200233
Email: {zhang_lihua, wang_miao,wang_guoqing}@careri.com
[*]Corresponding author

***Abstract*—As the occurrence of failure of electronic resources is sudden, real-time record analysis on the effectiveness of all resources in the system can discover abnormal resources earlier and start using backup resources or restructure resources in time, thus managing abnormal situations and finally realizing health management of the system. This paper proposed an algorithm: *MFPattern*, for mining frequent closed resource patterns in resource effectiveness matrix. In order to improve the efficiency, *MFPattern* algorithm uses sample-growth method and effective pruning strategies to guarantee mining all frequent closed patterns without candidate maintenance. Different from the traditional frequent closed pattern, *MFPattern* algorithm can mine resource combination patterns with all resources very effectively during work, those with simultaneous failure of resources and combination patterns in which some resources are very effective while some others have failure. The experimental result shows that our algorithm is more effective than existing algorithms.**

***Index Terms*—frequent pattern, closed, resource**

I. Introduction

Since resources are the physical support of system, the effectiveness of resources will directly influence the effectiveness of the system. Due to the influence of environmental factors and performance degradation of materials, electronic resources in the system will inevitably experience the process of capability loss. The influence of external stress in the process of system operation might cause acceleration of the degradation process of electronic resources and even damage and failure in extreme situations. Due to the aggravation of degradation, if not found in time, resource fault might finally cause system failure. Therefore, research on resource effectiveness is the footstone of prognostics and health management [1] of the system. As the occurrence of failure of electronic resources is sudden, real-time record analysis on the effectiveness of all resources in the system can discover abnormal resources earlier and start using backup resources or restructure resources in time, thus managing abnormal situations and finally realizing health management of the system.

Due to the great number of resources in system and the huge size of effectiveness value matrix of resources sampled in a period of time, how to mine the required resource pattern from these data efficiently is vital for health management of the system. The widely used data mining technology can mine various required knowledge from a lot of complex data. Frequent pattern mining [2-10] is an important branch of data mining technology. It can mine patterns meeting some rules and occurring frequently from huge data. However, this method has a disadvantage, i.e. existence of excessive redundancy modes. Literatures [11-14] reduce redundant patterns for mining frequent closed pattern. The important feature of frequent closed pattern algorithm is reserving and detecting strategies with candidate item. However, when there are a great number of item-sets in data set, it will produce a lot of candidate item-sets and thus cause the breakdown of internal memory. Therefore, *BIDE*[15] algorithm proposed by Han et al. uses backward checking method to avoid frequent item reservation, thus it can improve the mining efficiency. However, when the data set is sparse, the method of backward checking has a low pruning efficiency, thus increasing the complexity of the algorithm. Then *WIBE*[16] algorithm proposed by Wang et al. judges frequent closed pattern based on item-set weight information detection. Specifically, all items in the current candidate item-set have the same weight, this pattern is the subset of a frequent closed gene pattern that has been output and can be pruned. However, if the data set is large or sparse, this algorithm needs to store a lot of weight information in the internal memory, which will influence its efficiency and cause the breakdown of internal memory. Pan et al. [17,18] proposed a sample enumeration method of exploring enumeration space through recursive establishment and transposition of project database. Due to the lack of effective pruning strategy, the mining efficiency of this algorithm will be influenced when the number of samples is great.

To discover abnormity or failure in time, there should not be excessive sampling sites when the effectiveness value of resources is collected. There are two reasons as follows: (1) excessive sampling sites will influence the mining efficiency of the algorithm, thus causing the failure to report abnormal situations of resources in time;

doi:10.4304/jcp.9.5.1150-1158

(2) excessive sampling sites will influence the real-time property of the system, thus causing the failure to ensure normal operation of the system and finally influencing the health of the whole system. Based on the analysis above, this paper proposes an algorithm: *MFPattern*, for mining frequent closed resource pattern in resource effectiveness matrix using sample-growth method. In order to improve the mining efficiency, this algorithm uses effective pruning strategies to ensure the mining of all frequent closed patterns without candidate maintenance. *MFPattern* algorithm can mine the following three patterns: (1) resource combination pattern with very effective resources during system operation; (2) resource combination pattern with simultaneous failure of resources during system operation; (3) resource combination pattern in which some other resources will inevitably have failure during efficient operation of some resources.

II. Problem Description

Resource effectiveness matrix is defined as a two-dimensional real matrix $D = R \times S$, where row collection R represents the resource names set and column collection S refers to a set of different sampling sites. Element D_{ij} of matrix D is a real-valued number which refers to the effective value (e.g. BIT value) of resource i under sampling site j. $|R|$ is the number of resources in data set D and $|S|$ is the number of sampling sites in data set D. For the convenience of mining, the original effective value in resource effectiveness matrix is generally discretized into 1, -1 and 0, where 1 represents resource health, 0 represents sub-health and -1 represents resource failure, as shown in table 1.

Discrete resource effectiveness matrix

	S_0	S_1	S_2	S_3	S_4
R_1	1	-1	1	-1	1
R_2	1	-1	1	-1	-1
R_3	0	1	-1	1	1
R_4	1	1	0	1	1
R_5	-1	-1	0	-1	1

The relationship between two resources R_1 and R_2 can be defined as follows: (1) if R_1 and R_2 are both very effective (the value is 1), they have effective positive correlation, expressed as R_1R_2. The support is $sup(R_1R_2) = \frac{|R_1R_2|}{|S|}$, where $|R_1R_2|$ is the number of sampling sites when R_1 and R_2 are both very effective; (2) if R_1 and R_2 both have failure (the value is -1), they have failure positive correlation, expressed as $-R_1$-R_2. The support is $sup(-R_1 - R_2) = \frac{|-R_1 - R_2|}{|S|}$, where $|-R_1\text{-}R_2|$ is the number of sampling sites when R_1 and R_2 are both failure; (3) if R_1 is very effective while R_2 has failure, they have effective negative correlation, expressed as R_1-R_2; if R_2 is very effective while R_1 has failure, they also have effective negative correlation, expressed as $-R_1$-R_2; when R_1 and R_2 have effective negative correlation, the support is $sup(R_1 - R_2) = \frac{|R_1 - R_2|}{|S|} + \frac{|-R_1R_2|}{|S|}$, where $|R_1\text{-}R_2|$ is the number of sampling sites when R_1 is very effective and R_2 is failure, and $|-R_1R_2|$ is the number of sampling sites when R_1 is failure and R_2 is very effective.

Among three relationships above, the first one mines resource combination patterns in which resources are all very effective, i.e. there is a high efficiency when these resources work together; the second one mines resource combination patterns in which resources have failure together; in such patterns, potential reasons of failure can be found and decisions can be made for the use of backup resources; the third one mines resource patterns when some resources are very effective while some other resources have failure, in which potential unstable resources can be found and replaced earlier.

The mining purpose of *MFPattern* algorithm is to mine all frequent closed resource patterns in resource effectiveness data set, i.e. mine resource patterns without superset and with the same support which satisfies the threshold of support. To improve the mining efficiency of the algorithm, *MFPattern* algorithm mines frequent closed patterns using sample-growth without candidate maintenance. The mining process of this algorithm will be introduced in the next section.

III. The MFPattern Algorithm

Most traditional methods for frequent closed pattern mining are based on row extension. There are two reasons as follows: (1) with row extension, apriori principle can be used to prune the current extended resource that not satisfying the support threshold in real time; (2) according to the definition of frequent closed pattern: if there is no superset with the same support as the current extended frequent pattern, the current extended frequent pattern is frequent closed pattern; it is more convenient to make closed judgment through mining with row extension; if there is no candidate item-set with the same support in all candidate item-sets of the current extended frequent pattern, the current frequent pattern must be frequent closed pattern.

However, there are some disadvantages of using row extension as follows: (1) resources might have three relations simultaneously. If the method of row extension is used for mining, the number of resource patterns produced will present 3^n increase in extreme situations, thus row extension increases the complexity of mining; (2) if there are many resources and a few sampling sites, the use of row extension will increase the number of iterations of the algorithm, thus influencing the time efficiency and space efficiency of the algorithm; (3) as resource sampling is dynamic, with the use of row extension, the support of resources will change when the column of original data increases, thus making the design of incremental frequent closed pattern mining algorithm

relatively complex. Therefore, *MFPattern* algorithm uses the method of sample-growth for mining.

Before introducing how *MFPattern* algorithm mines frequent closed resource patterns without candidate maintenance in detail, we will first analyze mining strategies of existing frequent closed pattern mining algorithms briefly. Generally speaking, there are three strategies for mining frequent closed patterns: (1) it mines all frequent patterns firstly, and then outputs frequent closed patterns meeting the condition according to the definition of frequent closed pattern. This mining method is simple and visual, but has a low efficiency; (2) it mines frequent closed patterns with the method of real-time monitoring, i.e. store frequent closed patterns produced currently in the internal memory and compare each new pattern produced with frequent closed patterns stored in memory. If any frequent closed pattern in memory has the same support with the current extended pattern and is also the superset of the current extended pattern, then it stops the current pattern extension; if it is a new frequent closed pattern, update frequent closed patterns stored in the internal memory till finishing the mining of all frequent closed patterns. Such method can make pruning judgment for the current extended pattern in real time. However, as it is required to store all frequent closed patterns produced in the internal memory, it wastes the storage space. Meanwhile, each new frequent pattern produced is compared with frequent closed patterns in the memory, thus it reduces the mining efficiency; (3) it makes pruning judgment for the current extended frequent pattern with the method of backward checking. If the current candidate pattern is the subset of a frequent closed pattern that has been mined, the current extended pattern must have a prior candidate resource (which is candidate resource that has been extended) satisfying the definition of frequent closed pattern for pruning judgment. This method can avoid storing frequent closed patterns in the internal memory for pruning. However, if original data are sparse, the success rate of pruning will decrease and this backward checking method is time-consuming.

As resources in the system are effective in most cases, health management will be conducted in the case of failure. Therefore, the resource effectiveness matrix used in this paper is dense. *MFPattern* will mine frequent closed resource patterns with the method of sample-growth and backward checking. The mining process of this algorithm can be divided into two steps: first, construct a sample weighted graph; then, mine all frequent closed resource patterns with the method of sample-growth.

A. Construct Undirected Sample Relational Weighted Graph

The method of mining patterns with sample weighted graph was first used in the mining of bicluster in *MicroCluster* algorithm [19]. Then, Wang et al. [20,21] used sample weighted graph to mine bicluster and fault-tolerant bicluster. *MFPattern* algorithm proposed in this paper will use undirected sample relational weighted graph (sample weighted graph for short, the same below) to mine frequent closed resource patterns.

Definition 1. Sample weighted graph can be expressed as $G = \{E, V, W\}$. Each vertex in vertex set V in the weighted graph represents a sampling site; if an edge exists between a pair of vertices, it means that there are resources with effectiveness relationship under two sampling sites represented by this pair of vertices. The collection of edge is expressed with E; the weight on each edge is resource collection with effectiveness relationship under two sampling sites connected with this edge. The weight set is expressed with W.

Fig.1 shows the sample weighted graph corresponding to table 1. Weight on each edge is denoted as the set of effective resources under two sampling sites connected with this edge. The effective relationship among resources is obtained through the third relationship among resources defined in section two of this paper.

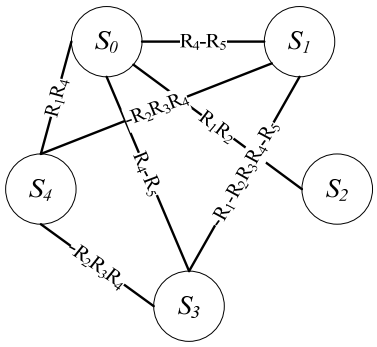


Figure 1. Undirected sample relational weighted graph corresponding to table 1

B. Mine Frequent Closed Resource Pattern

This section will introduce how *MFPattern* algorithm mines frequent closed resource patterns in the sample weighted graph. For the convenience of designing effective pruning strategies, *MFPattern* algorithm adopts the depth-first principle. It is assumed that frequent closed resource patterns are mined from the sample weighted graph in Fig.1. Firstly, *MFPattern* algorithm extends the edge connected with S_0 node, i.e. S_0->S_0S_1->$S_0S_1S_2$; then, it extends S_1 after the extension of S_0 branch. The support of resource pattern under sample $S_0S_1S_2S_3S_4$ is 1; that under samples $S_0S_1S_2S_3$, $S_0S_1S_2S_4$, $S_0S_2S_3S_4$ and $S_1S_2S_3S_4$ is 0.8; that under samples $S_0S_1S_2$, $S_0S_1S_3$, ..., and $S_2S_3S_4$ is 0.6. It can be seen from the extension method above that the method based on sample-growth has a different process from the mining of frequent closed patterns with resource extension.

When using resource extension to mine, it produces the support of two item-sets and then extends to three and four item-sets, etc. According to Apriori principle, the support decreases with the increase of the number of resources in item-set. The size of frequent patterns (the number of resources in the pattern) mined under the same sample branch with the method of sample-growth arranges from large to small and the support of pattern increases progressively. As there are only two samples during the extension, with the increase of the number of samples, the number of resources satisfying the definition together under the sample must decrease. However, the

support of resource pattern increases (because the number of samples increases). Therefore, when frequent closed resource patterns are mined based on resource extension, extension can terminate as long as the support of resource does not meet the threshold value. However, when patterns are mined based on sample-growth, it is necessary to continue the extension of sample set so long as the weight under the sample set extended is not null or not more than the minimum number of resources defined by users.

Definition 2. P is the current extended resource pattern. If all samples in S_i and P have a link in the sample weighted graph and the number of intersections (expressed as S_i.*Resources*) of weights of all join edges satisfies the number defined by users, S_i is the candidate sample of P.

Definition 3. P is the current extended resource pattern. If S_i and S_j are candidate samples of P, but S_i has been extended by P and S_j has not been extended by P yet. For S_j, S_i is prior candidate sample of P.

It can be known from the descriptions of the definition above that, a candidate sample becomes prior candidate sample if it has been extended. According to the principle of mining frequent closed patterns with backward checking, if the resource under the current extended sample set is the subset of resources under a prior candidate sample or candidate sample, the resource pattern under the current extended sample must not be frequent closed resource pattern. For example, assuming that the order of sample extension is from S_0 to S_4 based on the name and the current extended sample set is S_0S_1, S_3 and S_4 are its candidate samples and S_2 is its prior candidate sample (i.e. mining of $S_0S_1S_2$ branch has been completed). If the current sample to be extended is S_3, for $S_0S_1S_3$, S_2 is its prior candidate sample. It means that $S_0S_1S_2S_3$ must have been extended before $S_0S_1S_3$ is extended. If the resource pattern under $S_0S_1S_3$ is a subset of that under $S_0S_1S_2$, the resource pattern that can be obtained by extension of $S_0S_1S_3$ can definitely be obtained by extension of $S_0S_1S_2$ as well as $S_0S_1S_2S_3$ (because the resource pattern under $S_0S_1S_2S_3$ is the same as that under $S_0S_1S_3$ at this time); as the support of $S_0S_1S_2S_3$ is higher than that of $S_0S_1S_3$, $S_0S_1S_3$ can be pruned according to the definition of frequent closed pattern.

Lemma 1. Assume that P is the current extended resource pattern, M is the candidate sample set of P and N is the prior candidate sample set of P, if a prior candidate sample $N_j(N_j \in N)$ making PM_i. *Resource* a subset of PN_j.*Resource* exists for candidate sample $M_i(M_i \in M)$, PM_i. *Resource* should be pruned.

Proof. Proof by contradiction. Assuming that the resource set of the current candidate sample M_i is not a subset of resource collection of a prior candidate sample N_j before it, M_i can be pruned. It can be known from the assumption that resources not belonging to PN_j.*Resource* exists in PM_i.*Resource*. As frequent closed pattern mining uses sample depth-first extension method and N_j is extended earlier than M_i, there might be another sample S_m making PM_iS_m.*Resource* is not equal to $PM_iN_jS_m$.*Resource*. Therefore, M_i cannot be pruned, which is contradictory with the assumption. Thus, the original evidence is true.

Candidate samples meeting theorem 1 should be directly pruned and those not meeting pruning conditions should continue extension. However, whether the current extended pattern is output, should be judged according to the following output strategy (which actually meets the definition of frequent closed pattern):

Output strategy. Assuming that P is the current frequent closed pattern to be extended, M is candidate sample set of P and N is prior candidate sample set of P, if there is no candidate sample M_i making P.*Resource* a subset of PM_i.*Resource*for all candidate samples $M_i(M_i \in M)$ of P, P.*Resource* can be output.

Based on the analysis above, *MFPattern* algorithm can directly mine frequent closed resource pattern with the method of sample-growth without storing candidate frequent pattern in memory. Fig.2 illustrates the mining process of *MFPattern* algorithm. Example data is shown in table 1 and the threshold of support is 0.4 and the minimum number of resources in pattern is 2.

Algorithm 1: *MFPattern* algorithm

Input: threshold of support: r_{min}; resource effectiveness data: D

Output: all frequent closed resource patterns satisfying the threshold value

Initial value: sample weighted graph: G =Null; current pattern to be extended Q =Null，S_i=Null，S_j=Null

Algorithm description: MFPattern(r_{min}, D, Q, S_i, S_j)

(1) If G is null, scan data set D and make its weighted graph. S_i is the first sample in the weighted graph;

(2) For each sample S_j connected with sample S_i,

(3) If all resources linked lists in S_j satisfy pruning conditions,

(4) Continue;

(5) Else

(6) For resource linked lists not satisfying pruning conditions, $Q.Sample = Q.Sample \cup S_j$; $Q.Resource = Q.Resource \cap S_iS_j.Resource$;

(7) MFPattern(r_{min}, D, Q, S_i, S_j->next);

(8) end if

(9) end for

(10) if Q satisfies output conditions, then

(11) Output Q;

(12) end if;

(13) $S_i = S_i$->next;

(14) return

IV. Experimental Result and Analysis

In this section, we will make an experimental comparison on the mining efficiency and result of the algorithm above and existing algorithms. The hardware environment of the experiment is desktop computer: Intel(R) Core(TM)2 Duo 2.53GHz CPU and 4G internal memory; the software environment is Microsoft Windows 7 SP1 operating system; the algorithm programming and operating environment is Microsoft Visual C++ 6.0 SP6. Experimental data used in this paper is simulation data. To fully test the performance of the

algorithm, we generate six data sets randomly. Each data set contains 35 sampling sites and 800 resources. Table 2 describes the proportion of 1, 0 and -1 in each row in each data set.

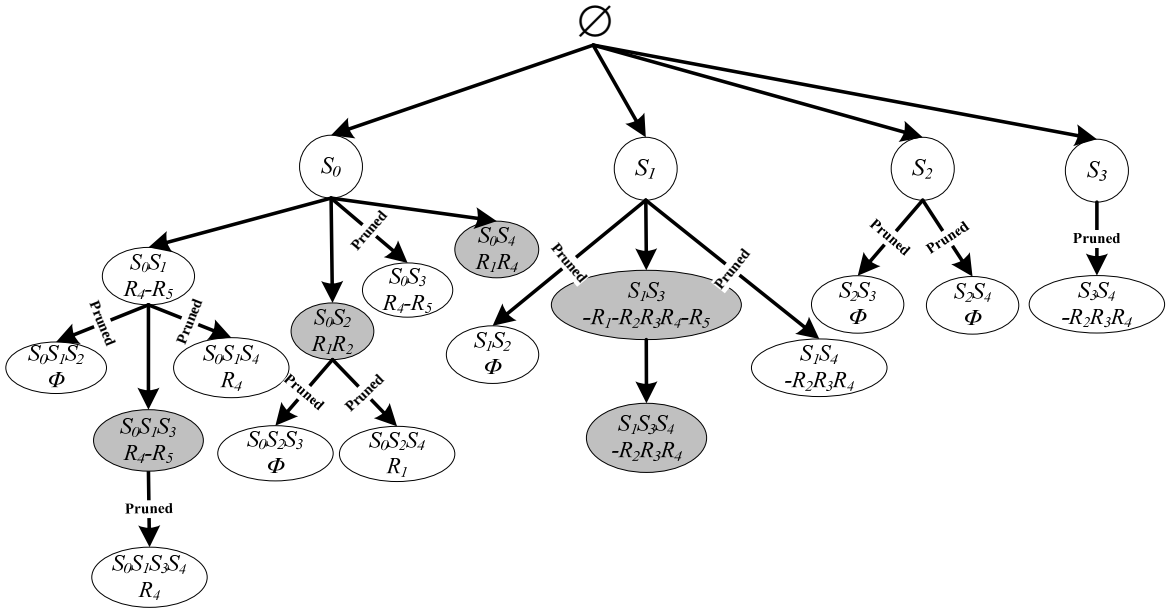


Figure 2. Example mining process of MFPattern algorithm

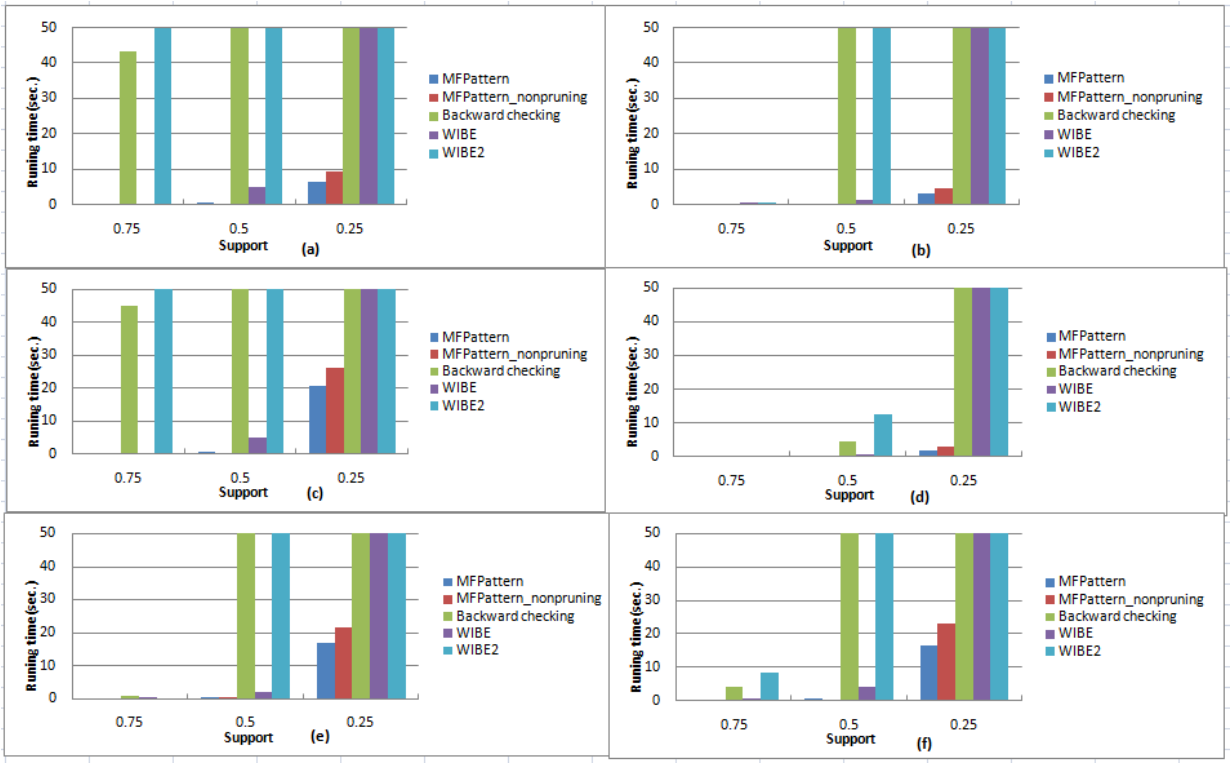


Figure 3. Comparison of operating time of five algorithms under different data sets with 200 resources and 20 sampling sites: (a) D_1; (b) D_2; (c) D_3; (d) D_4; (e) D_5; (f) D_6

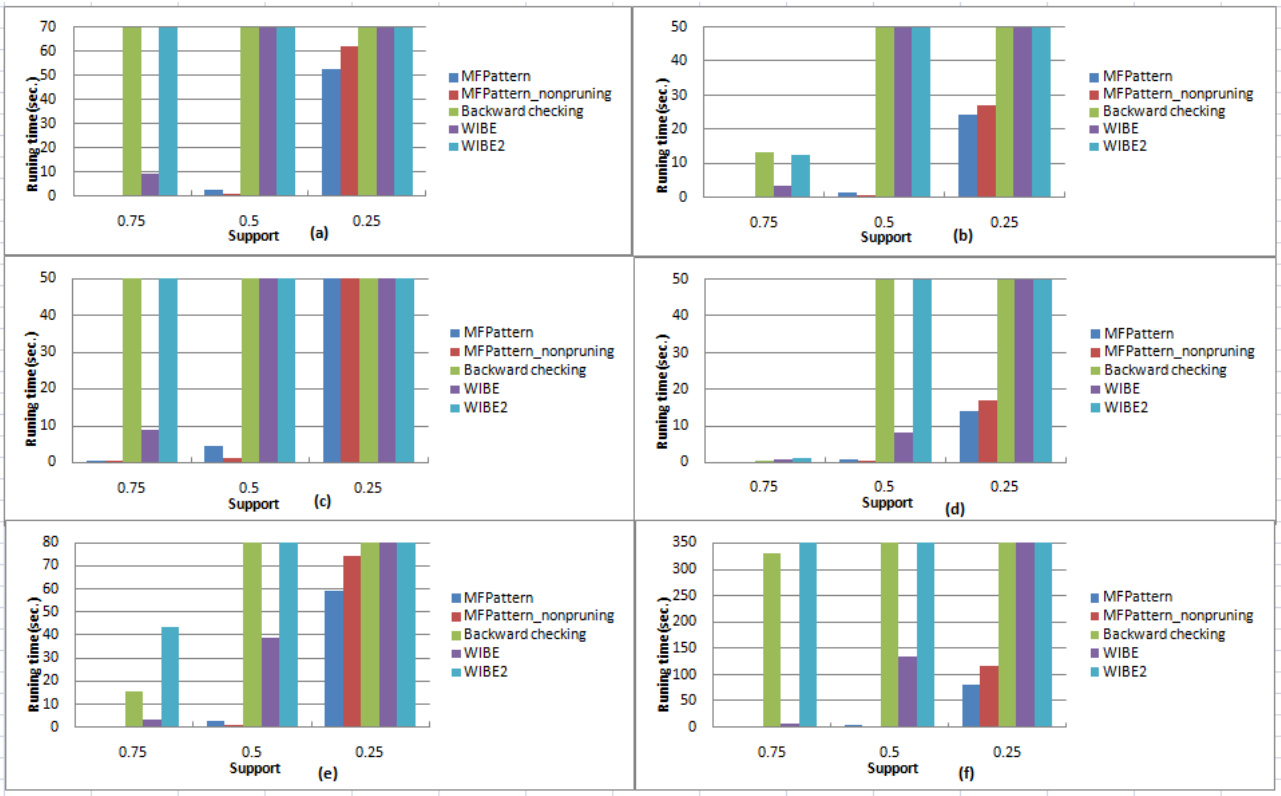


Figure 4. Comparison of operating time of five algorithms under different data sets with 500 resources and 20 sampling sites: (a) D_1; (b) D_2; (c) D_3; (d) D_4; (e) D_5; (f) D_6

DISTRIBUTION OF NUMERICAL VALUE PROPORTION IN SIX DATA SETS

	S_0	S_1	S_2	S_3	S_4
R_1	1	-1	1	-1	1
R_2	1	-1	1	-1	-1
R_3	0	1	-1	1	1
R_4	1	1	0	1	1
R_5	-1	-1	0	-1	1

In this section, *MFPattern* algorithm will be compared with *MFPattern* algorithm without pruning (denoted as *MFPattern_nonpruning*), *Backward Checking* algorithm, *WIBE* algorithm and *WIBE2* algorithm. *MFPattern_nonpruning* algorithm uses the same mining method as *MFPattern* algorithm, i.e. mine frequent closed patterns with the method of sample-growth. Different from *MFPattern* algorithm, *MFPattern nonpruning* algorithm does not use pruning strategies described in Lemma 1, but mines with the method of full extension. *Backward Checking* algorithm uses backward checking method described in literature [15] and mines frequent closed patterns without storing frequent item-sets. *WIBE* algorithm uses pruning strategies described in literature [16] and can mine frequent closed patterns without candidate maintenance. *WIBE2* algorithm uses the same pruning strategies as *WIBE* algorithm to mine frequent closed patterns. Different from *WIBE* algorithm, *WIBE2* algorithm uses co-expression support for mining.

The mining efficiency of five algorithms above will be compared. To fully compare the extendibility of algorithms, we produce multiple groups of data sets with different numbers of resources and sampling sites in allusion to six data sets in Table 2. Resources and sampling sites are selected according to the order of resources and sampling sites in data sets. Figs 3(a)-3(f) provide the comparison of operating time of five algorithms above under different data sets with 200 resources and 20 sampling sites. It can be seen from these figures that *MFPattern* and *MFPattern_nonpruning* algorithms can complete the mining process within 1 second and have more mining efficiency than other algorithms when the support is 0.75 and 0.5 under each data set. As *Backward Checking* algorithm judges frequent closed patterns with the method of backward checking, when the data set is dense (the proportion of 0 is low), it is necessary to frequently use backward checking strategy. Thus, the mining efficiency is low. Therefore, even when the support is 0.75, *Backward Checking* algorithm cannot complete the mining process in limited memory space under dense D_1 and D_3 data sets. When the support is 0.5, *Backward Checking* algorithm can only complete the mining process under data set D_4 with the lowest density of data. When the support is 0.25, only *MFPattern* and *MFPattern_nonpruning* algorithms can complete the mining process and other three algorithms cannot complete mining in limited memory space. As more frequent closed patterns can be produced when the support is low, pruning strategies used in

MFPattern algorithm can reduce the mining space of the algorithm. Thus, it has a higher mining efficiency than *MFPattern_nonpruning* algorithm. It is thus clear that the mining of frequent closed patterns with sample-growth is highly efficient and meanwhile pruning strategies used in *MFPattern* algorithm can improve the mining efficiency of the algorithm.

To further verify the extendibility of algorithms, Figs.4(a)-4(f) provide the comparison of operating time of five algorithms above under different data sets with 500 resources and 20 sampling sites. It can be seen that, similar to descriptions in Fig.3, *MFPattern* and *MFPattern_nonpruning* algorithms have more efficiency than other algorithms under different supports in each data set. When the support is 0.25, pruning strategies used in *MFPattern* algorithm have more obvious advantages than other supports. Figs.5(a)-5(f) provide the comparison of operating time of *MFPattern* algorithm and other three algorithms under different data sets with 500 resources and 35 sampling sites. It can be seen that *MFPattern* algorithm has the fastest mining process under most data sets and supports. However, compared to the operating time in Fig.3 and Fig.4, the operating time of *MFPattern* algorithm increases with the increase of the number of sampling sites. When the support is 0.55, as shown in Fig.5(c) and Fig.5(3), *MFPattern* algorithm has a lower mining efficiency than *WIBE* algorithm. As both data sets are dense, *MFPattern* algorithm needs frequent pruning judgment, thus influencing the mining efficiency. When the support is 0.4, *Backward Checking* and *WIBE2* algorithms cannot complete the mining process under all data sets in limited memory space. *MFPattern* algorithm can complete the mining process under sparse data sets D_2 and D_4.

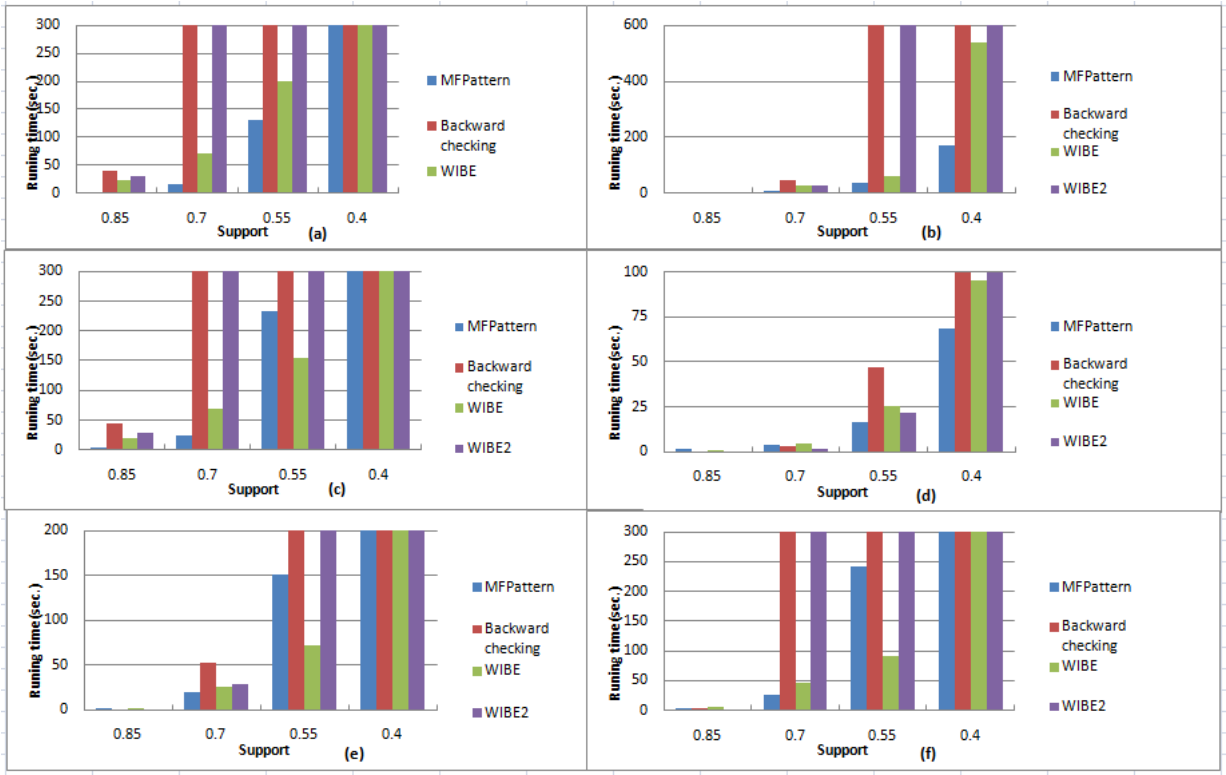


Figure 5. Comparison of operating time of five algorithms under different data sets with 800 resources and 35 sampling sites: (a) D_1; (b) D_2; (c) D_3; (d) D_4; (e) D_5; (f) D_6

V. CONCLUSION

This paper proposed an algorithm mining frequent closed resource patterns from data effectiveness matrix with the method of sample-growth: *MFPattern*, which uses effective pruning strategies to guarantee the mining of all frequent closed patterns without maintaining candidate item-sets. Different from the traditional frequent closed pattern, *MFPattern* algorithm can mine resource combination patterns with all resources effectively during work, those with simultaneous failure of resources and combination patterns in which some resources are very effective while some other resources have failure. The experimental result shows that this algorithm is more efficient than existing mining methods of frequent closed pattern. However, mining on discrete data will cause the loss of original data information. Our next research direction is to mine frequent closed patterns related to resource health from real resource effectiveness data.

ACKNOWLEDGMENT

This paper is Supported by National Key Basic Research Program of China under Grant No. 2014CB744900.

REFERENCES

[1] Michael Pecht,et al.. A prognostics and health management roadmap for information and electronics-rich systems. Microelectronics Reliability, 2010:317–323.

[2] J. Han, J. Pei, B. Mortazavi-Asl, Q. Chen, U. Dayal, and M.-C. Hsu, FreeSpan: Frequent Pattern-Projected Sequential Pattern Mining, Proc. 2000 ACM SIGKDD Int'l Conf. Knowledge Discovery in Databases (KDD '00), pp. 355-359, Aug. 2000.

[3] Pei Jian, Han Jiawei. Mining Sequential Patterns by Pattern-growth: The PrefixSpan Approach. IEEE Transactions on Knowledge and Data Engineering, 2004, 6(10): 1-17.

[4] Nohuddin P N E, Coenen F, Christley R, et al. Finding "interesting" trends in social networks using frequent pattern mining and self organizing maps. Knowledge-Based Systems, 2012, 29: 104-113.

[5] Mishra S, Mishra D, Satapathy S K. Fuzzy Frequent Pattern Mining from Gene Expression Data using Dynamic Multi-Swarm Particle Swarm Optimization. Procedia Technology, 2012, 4: 797-801.

[6] Pyun G, Yun U. A Frequent Pattern Mining Technique for Ranking Webpages Based on Topics. Multimedia and Ubiquitous Engineering. Springer Netherlands, 2013: 121-128.

[7] Glatz E, Mavromatidis S, Ager B, et al. Visualizing big network traffic data using frequent pattern mining and hypergraphs. Computing, 2013: 1-12.

[8] Yun U, Lee G, Kim S J. Analyzing Efficient Algorithms of Frequent Pattern Mining. IT Convergence and Security 2012. Springer Netherlands, 2013: 937-945.

[9] Xufei Zheng, Yanhui Zhou, Yonghui Fang. The Dual Negative Selection Algorithm Based on Pattern Recognition Receptor Theory and Its Application in Two-class Data Classification. Journal of computers, Vol 8, No 8, 2013, p:1951-1959.

[10] Haichao Luo, Xiaomin Li, Shujing Zheng, Mei Li, Lili Song. Study on Synthesis Evaluation of Intensive Land Use and Growth Pattern Transformation of Towns. Journal of computers, Vol 7, No 8, 2012, p:1959-1966.

[11] Pasquier N, Bastide Y, Taouil R, Lakhal L. Discovering frequent closed itemsets for association rules. In: Beeri C, et al, eds. Proc.of the 7th Int'l. Conf. on Database Theory. Jerusalem: Springer-Verlag, 1999. 398~416.

[12] Pei J, Han J, Mao R. CLOSET: An efficient algorithm for mining frequent closed itemsets. In: Gunopulos D, et al, eds. Proc. of the 2000 ACM SIGMOD Int'l. Workshop on Data Mining and Knowledge Discovery. Dallas: ACM Press, 2000. 21-30.

[13] Meng Han, Zhihai Wang, Jidong Yuan. Closed Sequential Pattern Mining in High Dimensional Sequences. Journal of Software, Vol 8, No 6, 2013, p:1368-1373.

[14] Qunhui Wu, Shilong Ma, Hao Wang. Extracting Feature Sequences in Software Vulnerabilities Based on Closed Sequential Pattern Mining. Journal of Software, Vol 8, No 8, 2013, p:1809-1817.

[15] Wang, J, Han, J. BIDE: Efficient Mining of Frequent Closed Sequences, Data Engineering, 2004. Proceedings. p: 79 – 90.

[16] Miao Wang, Xuequn Shang, Jingni Diao, Zhanhuai Li. WIBE: Mining Frequent Closed Patterns Without Candidate Maintenance in Microarray Dataset. DMIN 2010: 200-205.

[17] Cong, G., Tan, K., Tung, A. et al.: Mining Frequent Closed Patterns in Microarray Data. ICDM'04. IEEE Press, 2004, 363–366.

[18] Pan, F., Cong, G., Tung, K., Yang, J., Zaki, M.. Carpenter: Finding closed patterns in long biological datasets. In: Proc. ACM SIGKDD Intl. Conf. Knowledge Discovery and Data Mining (KDD), 2003, pp. 637–642.

[19] Lizhuang Zhao, Mohammed J. Zaki, MicroCluster: An Efficient Deterministic Biclustering Algorithm for Microarray Data, in IEEE Intelligent Systems, special issue on Data Mining for Bioinformatics, 2005, Vol. 20, No. 6, pp: 40-49.

[20] Miao Wang, Xuequn Shang, Miao Miao, Zhanhuai Li, Wenbin Liu. FTCluster: Efficient Mining Fault-Tolerant Biclusters in Microarray Dataset. Proceedings of ICDM 2011 workshop on Biological Data Mining and its Applications in Healthcare, p 1075-1082.

[21] Miao Wang, Xuequn Shang, Shaohua Zhang, Zhanhuai Li. FDCluster: Mining frequent closed discriminative bicluster without candidate maintenance in multiple microarray datasets. ICDM 2010 workshop on Biological Data Mining and its Applications in Healthcare, p 779-786.

Lihua Zhang is a doctoral student at the School of Computer Science and Engineering at the Northwestern Polytechnical University, Xi'an China. She completed her master degree from northwestern polytechnical university in 2008. Her current research interests are PHM, avionics, data mining and safety. Since 2013, she has been studying at science and technology on avionics integration laboratory.

Miao Wang is an engineer at Science and Technology on Avionics Integration Laboratory. He completed his doctor degree and master degree from Northwestern Polytechnical University in 2013 and 2008, respectively. He is a member of China Computer Federation. His research interests mainly include data mining, PHM, avionics and safety.

Zhengjun Zhai is a professor at the School of Computer Science and Engineering at the Northwestern Polytechnical University, Xi'an China. He is vice chairman of NPU youth association for science and technology, distinguished expert of aerospace electrical & electronics and weapon system Standardization technology committee, distinguished expert of AAMRI and premium member of china computer federation. His research interests include experiment and testing systems Integration, remote maintenance and fault diagnosis and virtual visualization.

Guoqing Wang is a professor and a supervisor of Ph.D. student in Northwestern Polytechnical University. He was born in 1956 and received his M.S. and Ph.D. degrees in Computer Science and Technology from the Northwestern Polytechnical University in 1984 and 1991 respectively. He is the institute director of China aeronautical radio electronics research institute, and the director of science and technology on avionics integration laboratory. He has long been engaged in the related technical research of avionic system integration, distributed parallel processing, high reliable fault-tolerant system, network and bus system etc. He serves as the vice director of national serve environment computer academy.

Efficient Mining Maximal Variant Usage and Low Usage Biclusters in Discrete Function-Resource Matrix

Lihua Zhang[1,2], Miao Wang[2,3,*], Zhengjun Zhai[1], Guoqing Wang[1,2,3]
[1]School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an, China, 710072
[2]Science and Technology on Avionics Integration Laboratory, Shanghai, China, 200233
[3]China National Aeronautical Radio Electronics Research Institute, Shanghai, China, 200233
Email: {zhang_lihua, wang_miao,wang_guoqing}@careri.com
[*]Corresponding author

***Abstract*—The functional layer is the pillar of the whole prognostics and health management system. Its effectiveness is the core of system task effectives. In this paper, we proposed a new bicluster mining algorithm: *DoCluster*, to effectively mine all biclusters with maximal variant usage rate and low usage rate in the discrete function-resource matrix. In order to improve the mining efficiency, *DoCluster* algorithm constructs a sample weighted graph firstly; secondly, all biclusters with maximal variant usage rate and low usage rate satisfying the variant usage rate and low usage rate definition are mined using sample-growth and depth-first method in the constructed weighted graph. *DoCluster* algorithm also uses several pruning strategies to ensure the mining of maximal bicluster without candidate maintenance. The experimental results show *DoCluster* algorithm is more efficient than other two algorithms.**

***Index Terms*—bicluster, variant usage rate, low usage rate, function, resource**

I. INTRODUCTION

The function is the foundation of task realization and also the basis of improving and guaranteeing quality, performance and effectiveness of system task information. The functional layer is the pillar of the whole system. Its effectiveness is the core of system task effectives. The health of the functional layer includes the status of functional components in the hierarchy range and overall health status of the whole functional layer. Health management objective of the functional layer is the effectiveness of the functional components and the hierarchy and to form function self-organizing platform based on the effectives of functional components. Although studying the effectiveness degree of resources is the base to construct a prediction and health management system [1]. The health degree of resources directly influences functional health. So, analysis of the call relation between functions and resources can excavate the health relation between them so as to complete the functions through using healthy resources and improve the health degree of functions.

The call relation of functions and resources can be abstracted as a matrix. In other words, each row means a resource and each column means a function, the value in the matrix is the use degree of a function to a resource. This value is defined during functional design, i.e. resource dependence degree of this function in aircraft system in order to complete a function. For example, for the resource whose storage spaces are 100K, function F_1 needs 60K storage spaces to store some temporary variables. The dependence degree of this function on this storage resource is 0.6. Through mining the above function-resource matrix, the usage relation between a group of functions and a group of resources can be gained. For instance, for a group of functions $F_1F_2F_3$, the resource relations called by each function are as follows: F_1==>$R_1R_2R_3$, F_2==>$R_2R_4R_5$ and F_3==>R_6R_7. Suppose $F_1F_2F_3$ need to cooperate to complete a task T. All above three functions may be called at the same time. For resource R_2, it supports F_1 and F_2 simultaneously. There may have two conditions: (1) R_2 has high effectiveness for F_1, but has low effectiveness for F_2; (2) R_2 has high effectiveness for both F_1 and F_2. The health degree of the first condition is higher than that of the second one. The reason is that, resource R_2 can serve F_1 and F_2 simultaneously in the first condition; while in the second condition, resource R_2 needs to serve for two functions. From the perspective of functional health, if resource R_2 has defects, its influence on the first condition is lower than the second one. So, through function-resource matrix mining, in order to achieve a group of functions, the resources which can satisfy all functional demands simultaneously and the resources which can satisfy all functional demands through multiple accesses can be mined, i.e. mine bicluster with variant usage rate or low usage rate from function-resource matrix.

The above mining concept complies with the bicluster in data mining field. Biclustering concept was first proposed by Cheng and Church [2]. As a special clustering method [3-9], bicluster does not generate cluster in overall experimental conditions, but only finds out the item sets with special significance for specific matrix sample. Thus, biclustering algorithm can mine

doi:10.4304/jcp.9.5.1159-1167

bicluster with variant usage rate and low usage rate described above from function-resource matrix. Currently, large quantities of algorithms based on greedy strategy or exploratory strategy are applied in mining bicluster. Cheng and Church proposed an algorithm based on greedy strategy [2]. This algorithm adopts a low square root residue to delete redundant nodes step by step. After that, many algorithms based on greedy strategy were raised [10-17]. All the above algorithms adopt the following two mining strategies: 1) produce cluster overall according to traditional clustering method and then optimize gradually; 2) mine bicluster in two types of data respectively and then gain the result through comparison and integration. But for the above two strategies, the efficiency of algorithms are not well. Thus, to design a high-efficiency bicluster mining algorithm is current research hotspot. So, Wang et al. came up with the mining algorithm to mine the maximal bicluster in discretized data [18].

The existing differential bicluster mining methods can be classified into two groups. One is to construct a difference matrix to mine discriminative biclusters. [19] developed a methodology for differential co-expression on a global scale. [20] proposed an algorithm to extract differential biclusters from the two gene expression datasets. [21] aims to mine subspace differential co-expression patterns. And it can also be used for mining differential biclusters. Another recent proposed algorithm called *DeBi* [22] uses frequent pattern mining approach for discovering maximum size homogeneous bicluster in which all genes are co-expressed under a subset of samples. However, this algorithm cannot effectively mine bicluster with variant usage rate meeting difference restraint from function-resource matrix.

We can see through the above analysis that existing bicluster algorithm has some shortcomings during mining a bicluster with variant usage rate and with low usage rate. In order to improve mining efficiency, this paper proposed a new bicluster mining algorithm - *DoCluster* algorithm which can effectively mine all biclusters with maximal variant usage rate and low usage rate from discrete function-resource matrix. Since the number of functions is far lower than that of resources in function-resource matrix, this algorithm uses sample-growth method for mining. First, a sample weighted graph is constructed, which includes all resource collections between both samples that satisfy the definition of variant usage rate or low usage rate; then, all biclusters with maximal variant usage rate and low usage rate satisfying the definition are mined with the mining method of using depth-first sample-growth method in the weighted graph. To improve the mining efficiency of the algorithm, *DoCluster* algorithm uses several pruning strategies to ensure the mining of maximal bicluster without candidate maintenance.

II. PROBLEM DESCRIPTION

Function-resource matrix is defined as a two-dimensional real matrix $D = R \times F$, in which row set R represents the set of resources and column set F refers to the set of functions. Element D_{ij} of matrix D is a real number which represents the ability validity or usage rate of resource i supporting function j. $|R|$ is the number of resources in data set D and $|F|$ is the number of functions in data set D. For the convenience of mining, the original effective values in function-resource matrix are usually dispersed as 1, -1 and 0, where -1 means the usage rate of the resource is the minimum during the implementation of some function; 0 means the usage rate of the resource is moderate during the implementation of some function; 1 means the usage rate of the resource is the maximal during the implementation of some function, as shown in Table 1.

The significance of bicluster to be mined from function-resource matrix as shown in Table 1 is to mine a group of functions executed; under this group of functions, the usage rate of the resource is the maximal, i.e. which resources can reach the maximal usage rate when used together. In other words, the resources have the highest effectives when all functions are executed. For example, for a group of functions F_1F_2 (F_1==>$R_1R_2R_3$, F_2==>R_2R_4), these three functions may be called simultaneously. For resource R_2, there are three situations for supporting F_1 and F_2: (1) for F_1, the usage rate of R_2 is high, while it is low for F_1, as shown in Table 2; (2) for both F_1 and F_2, the usage rate of R_2 is high, as shown in Table 3; (3) for both F_1 and F_2, the usage rate of R_2 is low, as shown in Table 4, the health degree in the first and the third conditions is higher than the second condition. The reason is that R_2 can serve F_1 and F_2 at the same time in the first and the third conditions resource. In the third condition, resource R_2 needs to serve the two functions respectively. This paper puts forward that bicluster mined by *DoCluster* algorithm aims at the first and third conditions.

TABLE I.
AN EXAMPLE OF FUNCTION-RESOURCE MATRIX

	F_1	F_2	F_3	F_4	F_5
R_1	1	-1	-1	-1	1
R_2	-1	1	-1	-1	1
R_3	1	-1	-1	-1	0
R_4	0	1	-1	-1	1

TABLE II.
AN EXAMPLE OF VARIANT USAGE RATE

	F_1	F_2
R_1	1	-1
R_2	1	-1
R_3	1	0
R_4	0	-1

TABLE III.
AN EXAMPLE OF NON-VARIANT USAGE RATE

	F_1	F_2
R_1	1	-1
R_2	1	1
R_3	1	0
R_4	0	-1

TABLE IV.
AN EXAMPLE OF LOW USAGE RATE

	F_1	F_2
R_1	1	-1
R_2	-1	-1
R_3	1	0
R_4	0	-1

Definition 1. In order to facilitate description of bicluster with variant usage rate and low usage rate, suppose the use values of resource R_1 after discretization under the functions F_1 and F_2 are V_1 and V_2. There are four representations for R_1 under F_1 and F_2: (1) if V_1=1 and V_2=-1, or V_1=-1 and V_2=1, the contribution rate of R_1 to F_1 and F_2 satisfies diversity requirement, expressed as 'R_1' and '$*R_1$' respectively; (2) if V_1=-1 and V_2=-1, the contribution rate of R_1 to F_1 and F_2 satisfies diversity requirement, expressed as '$-R_1$'; (3) if V_1=1 and V_2=1, the contribution rate of R_1 to F_1 and F_2 does not satisfy diversity requirement, so no record is given; (4) if V_1=0 or V_2=0, the contribution rate of R_1 to F_1 and F_2 does not meet diversity requirement, so no record is given.

Thus, in bicluster mined by *DoCluster* algorithm, each resource can satisfy the first or the second conditions described above under all functions. To improve mining efficiency of the algorithm, *DoCluster* algorithm mines biclusters with maximal variant usage rate and maximal low usage rate by using sample-growth method without candidate maintenance. The mining process of this algorithm will be introduced in the next section.

III. THE DOCLUSTER ALGORITHM

The mining steps of *DoCluster* algorithm can be divided into two steps: firstly, scan original function-resource matrix, according to the definition of biclusters with maximal variant usage rate and maximal low usage rate, all sample weighted graphs satisfying the above definition are produced; then, use sample-growth method to mine all biclusters with maximal variant usage rate bicluster and maximal low usage rate bicluster.

A. Construct Sample Relational Weighted Graph

The method of mining modes with sample relational weighted graph was used in *MicroCluster* algorithm [12] to mine bicluster firstly. Then, Wang et al. [18, 23] also used sample relational weighted graph to mine bicluster and fault-tolerant bicluster. *DoCluster* algorithm in this paper will adopt undirected sample relational weighted graph (hereinafter referred to as sample weighted graph) to mine biclusters with maximal variant usage rate and maximal low usage rate.

Definition 2. Sample weighted graph can be expressed with the set $G = \{E, V, W\}$. Each node in the vertex set V in the weighted graph represents a function. If an edge exists between a pair of vertices, this means the resource with variant usage rate or low usage rate exists below two functions represented by this pair of vertices. The set of the edges is denoted as E. The weights of each edge are the resource set satisfying the definition of variant usage rate or the definition of low usage rate under the two functions connected with this edge. The set of the weights is denoted as W.

According to the description in Definition 1, when the resources among functions satisfy the definition of variant usage rate, the weight between two functions does not satisfy commutativity. For instance, the weight under F_1F_2 is $R_1*R_2R_3$, while the weight under F_2F_1 is $*R_1R_2*R_3$. So, in Definition 2, the weight of each edge is the weight under F_iF_j, where $i<j$. Fig.1 shows the weighted graph corresponding to Table 1. For the convenience of follow-up description, Fig.2 provides storage structure of Fig.1.

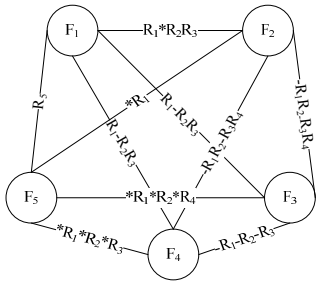


Figure 1. The sample weighted graph constructed from Table 1

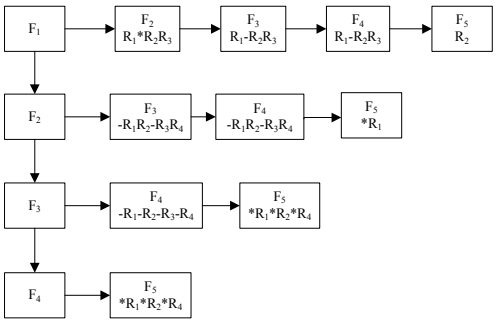


Figure 2. The storage structure of Fig.1

B. Mining Maximal Bicluster

After the sample weighted graph is constructed, this section will introduce how *DoCluster* algorithm mines all biclusters with maximal variant usage rate and maximal low usage rate from sample weighted graph without candidate maintenance in detail. According to the description in Definition 2, biclusters with variant usage rate and low usage rate extended satisfy anti-monotonicity, i.e. if the bicluster obtained by extension of $F_1F_2…F_n$ does not satisfy constraint conditions, neither does any superset $F_1F_2…F_nF_m$. Therefore, biclusters with a greater scale can be obtained by extension of the weight on each edge in the weighted graph in terms of intersection. But the bicluster mined by *DoCluster* is different from the extension mode described in [18]. In *FDCluster* algorithm, if S_3 is gained through extending S_1S_2, $S_1S_2S_3$ set can be obtained through calculating the intersection of the weight of S_1S_2 and the weight of S_1S_3. However, this scheme cannot be used in this algorithm. It is required to calculate the intersection of the edges of F_1F_2, F_1F_3 and F_2F_3 in order to gain the resource set meeting conditions under $F_1F_2F_3$. Only in this way, such situation can be avoided that two or more '1' occur

simultaneously. For example, in Table 1, for resource R_2, when extended to F_5 from $F_1F_2F_3F_4$, R_2 is included in the weights under F_1F_5. However, R_2 is not included in the weights under F_2F_5. If the intersection of R_2 and the weight of F_2F_5 is not calculated when extended to F_5 from $F_1F_2F_3F_4$, a wrong $F_1F_2F_3F_4F_5$ bicluster including R_2 will emerge. So, when a new function is introduced in bicluster, it is necessary to calculate the intersection of all edges of the function newly introduced and the resource collection of bicluster extended. When calculating the intersection of the weights, it is only necessary to calculate the intersection of the resources, not necessary to consider '*' or '-' symbols before resources. With different symbols before resources, the intersection can also be calculated. These symbols are only used in pruning design.

We will introduce how *DoCluster* algorithm uses pruning strategies to mine all biclusters with maximal variant usage rate and maximal low usage rate from sample relationship weight graph without candidate maintenance in detail. This paper will judge maximal bicluster with the method of backward checking proposed in [24] without candidate maintenance. That is to say, if resources under the current candidate sample and some prior candidate sample (mined sample) have some inclusion relation, i.e. all biclusters produced by the current candidate sample can be produced by some prior candidate sample, the current candidate sample can be pruned. When calculating the intersection of the weights, it is just necessary to calculate the intersection of resources, and the intersection can also be calculated with different symbols before resources. But, in accordance with the description (1) in Definition 1, since resource expression forms of V_1=1 and V_2=-1 or V_1=-1 and V_2=1, resource expression forms under F_1F_2 and F_2F_1 may be different. For example, when mining F_2, the candidate functions are F_3(-R_1R_2-R_3R_4), F_4(-R_1R_2-R_3R_4) and F_5(*R_1), and the prior candidate function is F_1(*R_1R_2*R_3). Since currently F_2 is extended, F_1 is its prior candidate function. At this moment, F_2F_1(*R_1R_2*R_3) should be produced, instead of F_1F_2(R_1*R_2R_3). As resource expression forms under F_1F_2 and F_2F_1 are different, the weighted graph made by this algorithm is a directed graph rather than undirected graph. For F_n and F_m, it is necessary to build edges on F_nF_m and F_mF_n respectively. For F_nF_m and F_mF_n, the difference of weights on the edge is the interchange of resource expression forms "R_1" and "*R_1". Therefore, for saving the storage space, the storage of weight is only that of weight on F_iF_{i+1} edge. The weight on $F_{i+1}F_i$ edge can be calculated with F_iF_{i+1}. For instance, the storage structure of Table 1 is as shown in Fig.2. F_2F_1(*R_1R_2*R_3) can be gained through "complementing" F_1F_2(R_1*R_2R_3) ("R_i" and "*R_i" interchange, and "-R_i" remains unchanged).

During function extension, the resource "symbol" is not considered. But during candidate function pruning, it is necessary to judge according to resource symbols under the candidate functions. Here, resource symbols under the candidate functions are decided by candidate functions at current layer and resource symbols of the weights on the edge of initial extension function. For example, according to the storage structure shown in Fig.2, assuming the bicluster extended currently is F_2F_3(-R_1R_2-R_3R_4), its candidate functions are F_4(-R_1R_2-R_3R_4) and F_5(*R_1); its prior candidate function is F_1(*R_1R_2*R_3). Resource (-R_1R_2-R_3R_4) under candidate function F_4 is gained through calculating the intersection of the weights of edges F_2F_4, F_3F_4 and F_2F_3. The "symbol" of each resource is the resource "symbol" on the edge F_2F_4. Because function F_2 is extended currently, resource symbols of candidate functions are decided by F_2F_4. Similarly, for prior candidate function F_1 of F_2F_3, its resource is also gained through calculating the intersection of F_1F_2, F_1F_3 and F_2F_3. Its resource symbols are decided by resource symbols on F_2F_1.

TABLE V.
AN EXAMPLE OF PRUNING USED MATRIX

	F1	F2	F3	F4
R1	1	-1	-1	-1
R2	-1	1	-1	-1
R3	1	-1	-1	-1

Resource R_1 is respectively expressed as 'R_1' and '*R_1' above when the form of expression of resources is illustrated, just for the convenience of design of pruning strategies. For example, assuming there is the sole resource R_1 in Table 5, for R_1, when the extension starts from F_1F_2, according to the above description, the expression form of R_1 is 'R_1'. Assuming all functions extended from F_1F_2 have been extended, when extending F_1F_3, the expression form of R_1 on the edge of F_1F_3 is also 'R_1'. At this moment, F_1F_3 can be pruned. It is known from the expression form of R_1 that '1' must exist under F_1. According to previous variance definition, '1' impossibly exists under other functions extended from F_1. That is to say, R_1 can only be '-1' under F_3. Therefore, functions which F_1F_3 can extend must be gained through F_1F_2 extension. Meanwhile, F_1F_2 can extend F_3. So, F_1F_3 can be pruned.

If a resource in the current candidate function to be extended meets the form of 'R_1', this resource can be pruned according to the Lemma 1 below.

Lemma 1. Assuming that P is the bicluster with variant usage rate to be extended currently; M is the candidate function set of P and N is the prior candidate function set of P. If the expression form is 'R_j' for any resource R_j in candidate function M_i ($M_i \in M$) and there is a prior candidate function N_j ($N_j \in N$) under which resource R_j also exists, resource R_j in M_i can be obtained by extension of prior candidate function N_j.

Proof. Proof by contradiction is adopted. Resource expression form of current candidate function M_i is 'R_j'; a prior candidate function N_j ($N_j \in N$) exists; resource R_j also exists under N_j. Thus, M_i can be pruned. In line with description (1) in Definition 1, for resource R_j, '1' is under some function in P. In accordance with the definitions of variant usage rate and low usage rate,

resource R_j must be '-1' under candidate function M_i and prior candidate function N_j. So, the bicluster extended currently must be a bicluster with variant usage rate. As only one '1' can exist for each resource under all functions in the bicluster with variant usage rate, the bicluster with variant usage rate gained through extension of PM_i can be obtained through extension of PN_jM_i. Thus, M_i can be pruned. This contradicts the assumption, so the original proof is established.

However, for the expression form of '*', the above pruning strategy is not applicable. For example, assuming there is the sole resource R_2 in Table 5, for R_2, when the extension starts from F_1F_2, according to the previous description, the expression form of R_2 is '*R_2'. Assuming all functions extended from F_1F_2 have been extended, when extending F_1F_3, the expression form of R_2 on the edge of F_1F_3 is also '-R_2'. At this moment, F_1F_3 can not be pruned. It is known from the expression form of R_2 that '1' must exist under F_2. According to previous variance definition, '1' likely exists under other functions extended from F_1. That is to say, '1' likely appears under the functions extended by F_1F_3. For R_2, $F_1F_3F_4F_5$ can be gained through extension of F_1F_3, but $F_1F_2F_3F_4F_5$ cannot be gained through extension of F_1F_2. Therefore, functions which F_1F_3 can extend may not be gained through F_1F_2 extension. So, For R_2, F_1F_3 can not be pruned.

According to the above analysis, if a resource in the current candidate function to be extended satisfies the form of '*R_1', it should be judged whether this resource can be pruned according to the weight of prior candidate function. Therefore, the following Lemma can be used for pruning.

Lemma 2: assuming that P is the bicluster with variant usage rate to be extended currently; M is the candidate function set of P and N is the prior candidate function set of P. If the expression form is '*R_j' for any resource R_j in candidate function M_i ($M_i \in M$) and there is a prior candidate function N_j ($N_j \in N$) under which resource R_j with the expression form of '-R_j' also exists, resource R_j in M_i can be obtained by extension of prior candidate function N_j.

Proof: Proof by contradiction is adopted. When resource expression form of current candidate function M_i is 'R_j'; a prior candidate function N_j ($N_j \in N$) exists; resource R_j also exists under N_j with the expression form of '-R_j', M_i can be pruned. In line with description (1) in Definition 1, for resource R_j, '1' is under current candidate function in M_i. In accordance with the definitions of variant usage rate and low usage rate, resource R_j under all functions in P must be '1'. Since resource R_j also exists under N_j with the expression form of '-R_j', the bicluster extended currently must be a bicluster with low usage rate. As only one 1 can exist for each resource in the bicluster with variant usage rate, the bicluster PN_jM_i with variant usage rate can be gained through extension of PN_j. Thus, the bicluster with variant usage rate gained through extension of PM_i can be obtained through extension of PN_jM_i. Thus, M_i can be pruned. This contradicts the assumption, so the original proof is established.

Similarly, if a resource in the current candidate function to be extended meets the form of '-R_1', it should be judged whether this resource can be pruned according to the weight of prior candidate function. Therefore, the following Lemma can be used for pruning.

Lemma 3: assuming that P is the bicluster with variant usage rate to be extended currently; M is the candidate function set of P and N is the prior candidate function set of P. If the expression form is '-R_j' for any resource R_j in candidate function M_i ($M_i \in M$) and there is a prior candidate function N_j ($N_j \in N$) under which resource R_j with the expression form of '-R_j' also exists, resource R_j in M_i can be obtained by extension of prior candidate function N_j.

Proof: Proof by contradiction is adopted. When resource expression form of current candidate function M_i is 'R_j'; a prior candidate function N_j ($N_j \in N$) exists; resource R_j also exists under N_j with the expression form of '-R_j', M_i can be pruned. In line with description (1) in Definition 1, for resource R_j, '-1' is under current candidate function in M_i. In accordance with the definitions of variant usage rate and low usage rate, resource R_j under all functions of in P may be '-1' or '1' under some functions. Since resource R_j also exists under N_j with the expression form of '-R_j', the bicluster extended currently may be a bicluster with low usage rate or a bicluster with variant usage rate. As the expression form of resource R_j under current candidate function M_i is '-1', the bicluster PN_jM_i with variant usage rate or low usage rate can be gained through extension of PN_j. Thus, the bicluster with variant usage rate gained through extension of PM_i can be obtained through extension of PN_jM_i. Thus, M_i can be pruned. This contradicts the assumption, so the original proof is established.

Lemma 4: assuming that P is the bicluster with variant usage rate to be extended currently; M is the candidate function set of P and N is the prior candidate function set of P. If the same prior candidate function $N_j(N_j \in N)$ exists for each resource R_j in candidate function $M_i(M_i \in M)$, making each resource R_j in candidate function M_i meet the conditions in Lemma 1 or 2 or 3, candidate function M_i can be pruned.

Proof: the process of proof can be gained through merging the processes of proof in Lemma 1, 2 and 3, so it is omitted here.

It can be seen from Lemma 4 that, the candidate function can only be pruned if all resources in the candidate function can be obtained by resource extension in the same prior candidate function; otherwise, this candidate function will be extended. If no successor or prior is its superset, it can be outputted. We will explain the algorithm mining process through an example. The data in the example are function-resource use relationship matrix shown in Table 1. Firstly, construct the weight graph among functions, as shown in Fig.1; then,

DoCluster algorithm deeply mines according to function extension.

(1) Firstly, the extension starts from F_1F_2, and all candidate functions of F_1F_2 are produced: $F_3(R_1\text{-}R_2R_3)$ and $F_4(R_1\text{-}R_2R_3)$. The resource conditions after the intersection is calculated are shown in the brackets. When candidate functions are produced, the resource set under candidate function F_3 of F_1F_2 can be gained only after the intersection of the weights of F_1F_2, F_1F_3 and F_2F_3 currently extended is calculated. Then, the candidate function $F_4(R_1\text{-}R_2R_3)$ is produced through mining $F_1F_2F_3(R_1\text{-}R_2R_3)$. Here, when producing the resource set under candidate function F_4, since the intersection of $F_1F_2F_3$ and $F_1F_2F_4$ has been worked out, it can be obtained through calculating the intersection of the weights of $F_1F_2F_3$, $F_1F_2F_4$ and F_3F_4, without the need of calculating the intersection of each edge. So, the maximal bicluster $F_1F_2F_3F_4(R_1\text{-}R_2R_3)$ can be gained through extending F_1F_2 deeply and preferentially. Then, prepare to extend $F_1F_2F_4(R_1\text{-}R_2R_3)$. For F_1F_2, when all resources in current candidate function meet pruning conditions in Lemma 1, 2 or 3 for prior F_3. Therefore, $F_1F_2F_4$ can be pruned according to Lemma 4.

(2) Next, branch F_1F_3 is produced. All candidate functions of $F_1F_3(R_1\text{-}R_2R_3)$ are $F_4(R_1\text{-}R_2R_3)$ and $F_5(R_2)$. For F_4, a prior candidate function $F_2(R_1{*}R_2R_3)$ of F_1F_3 can be found. All resources of $F_4(R_1\text{-}R_2R_3)$ are the subset of resources in $F_2(R_1{*}R_2R_3)$, but resource R_2 does not meet pruning conditions (Lemma 3). So, $F_1F_3F_4$ can continue to be extended, but cannot be outputted. Then, the candidate function $F_5(R_2)$ is generated through extending $F_1F_3F_4(R_1\text{-}R_2R_3)$. Since $F_5(R_2)$ does not meet pruning conditions, $F_1F_3F_4F_5(R_2)$ can be outputted. When preparing to extend $F_1F_3F_5(R_2)$, since a prior F_4 makes $F_5(R_2)$ satisfy the pruning conditions in Lemma 1, $F_1F_3F_5(R_2)$ should be pruned. Similarly, the branches of F_2, F_3 and F_4 can be mined respectively.

(3) When F_2 is mined, its candidate functions are $F_3(\text{-}R_1R_2\text{-}R_3R_4)$, $F_4(\text{-}R_1R_2\text{-}R_3R_4)$ and $F_5(*R_1)$; its prior candidate function is $F_1(*R_1R_2{*}R_3)$. As resources under $F_2F_3(\text{-}R_1R_2\text{-}R_3R_4)$ do not satisfy pruning conditions, it is necessary to continue to extend $F_2F_3(\text{-}R_1R_2\text{-}R_3R_4)$ whose candidate functions are $F_4(\text{-}R_1R_2\text{-}R_3R_4)$ and $F_5(*R_1)$ and prior candidate function is $F_1(*R_1R_2{*}R_3)$. The candidate function $F_4(\text{-}R_1R_2\text{-}R_3R_4)$ dissatisfies pruning conditions, so it is necessary to continue extending to generate $F_2F_3F_4(\text{-}R_1R_2\text{-}R_3R_4)$. The candidate function is $F_5(*R_1)$ and the prior candidate function is $F_1(*R_1R_2{*}R_3)$. Then, $F_2F_3F_4F_5(*R_1)$ can continue to be generated and outputted. According to pruning conditions, F_2F_4 and F_2F_5 can be pruned.

(4) When extending to F_3, the candidate functions of F_3 are produced: $F_4(\text{-}R_1\text{-}R_2\text{-}R_3\text{-}R_4)$ and $F_5(*R_1{*}R_2{*}R_4)$. Its prior candidate functions are $F_1(*R_1\text{-}R_2{*}R_3)$ and $F_2(\text{-}R_1{*}R_2\text{-}R_3{*}R_4)$. As resources in $F_4(\text{-}R_1\text{-}R_2\text{-}R_3\text{-}R_4)$ are the subset in prior candidate function $F_2(\text{-}R_1{*}R_2\text{-}R_3{*}R_4)$, $F_3F_4(\text{-}R_1\text{-}R_2\text{-}R_3\text{-}R_4)$ cannot be outputted. But $F_3F_4(\text{-}R_1\text{-}R_2\text{-}R_3\text{-}R_4)$ dissatisfies pruning conditions, so it is necessary to continue extending $F_3F_4(\text{-}R_1\text{-}R_2\text{-}R_3\text{-}R_4)$ to produce the candidate function $F_5(*R_1{*}R_2{*}R_4)$ and prior candidate functions: $F_1(*R_1\text{-}R_2{*}R_3)$ and $F_2(\text{-}R_1{*}R_2\text{-}R_3{*}R_4)$. At this moment, F_5 dissatisfies pruning conditions, so $F_3F_4F_5(*R_1{*}R_2{*}R_4)$ can be produced and outputted. For $F_3F_5(*R_1{*}R_2{*}R_4)$, a prior $F_3F_4(\text{-}R_1\text{-}R_2\text{-}R_3\text{-}R_4)$ exists, making F_3F_5 satisfy pruning conditions, so $F_3F_5(*R_1{*}R_2{*}R_4)$ is pruned.

(5) When extending $F_4F_5(*R_1{*}R_2{*}R_4)$, a prior $F_4F_3(\text{-}R_1\text{-}R_2\text{-}R_3\text{-}R_4)$ exists, making F_4F_5 satisfy pruning conditions, so F_4F_5 is pruned.

The above mining process is shown in Fig.3. The specific description of *DoCluster* algorithm is as follows:

Algorithm 1: *DoCluster* algorithm
Input: number threshold: r_{min}; function-resource matrix: D
Output: all biclusters with maximal variant usage rate or maximal low usage rate meeting the threshold
Initial value: sample weight graph: G =Null, current bicluster to be extended Q =Null, S_i=Null and S_j=Null.
Algorithm description: DoCluster(r_{min}, D, Q, S_i, S_j)
(1) If G is null, scan data set D and construct its weighted graph. S_i is the first sample in the weighted graph;
(2) For each sample S_j connected with sample S_i
(3) If all resource linked lists in S_j satisfy pruning conditions in Lemma 4, then
(4) Continue;
(5) Else
(6) For resource linked lists not satisfying pruning conditions, $Q.Sample = Q.Sample \cup S_j$; $Q.Resource = Q.Resource \cap S_iS_j.Resource$;
(7) DoCluster(r_{min}, D, Q, S_i, S_j->next);
(8) Endif
(9) Endfor
(10) If Q satisfies maximal definition, then
(11) Output Q
(12) Endif;
(13) $S_i = S_i$->next;
(14) Return

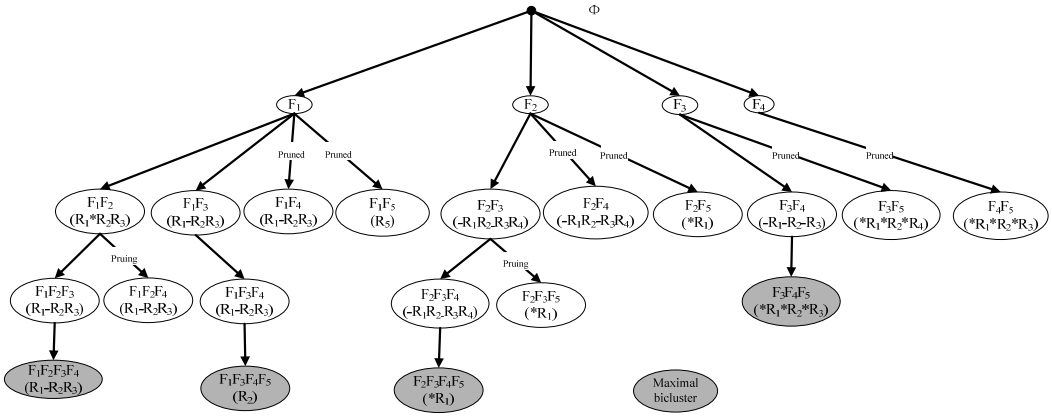


Figure 3. Example mining process of DoCluster algorithm

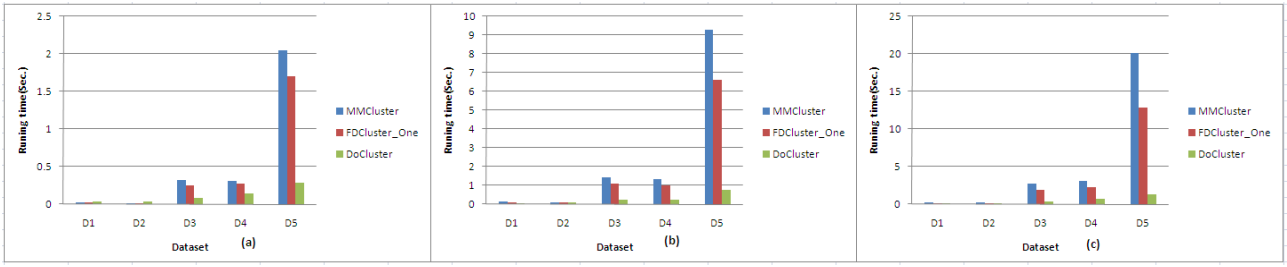


Figure 3. The comparison of performance periods of the above three algorithms under each data set when the number of functions is 20: (a) 200 resources；(b) 500 resources；(c) 800 resources

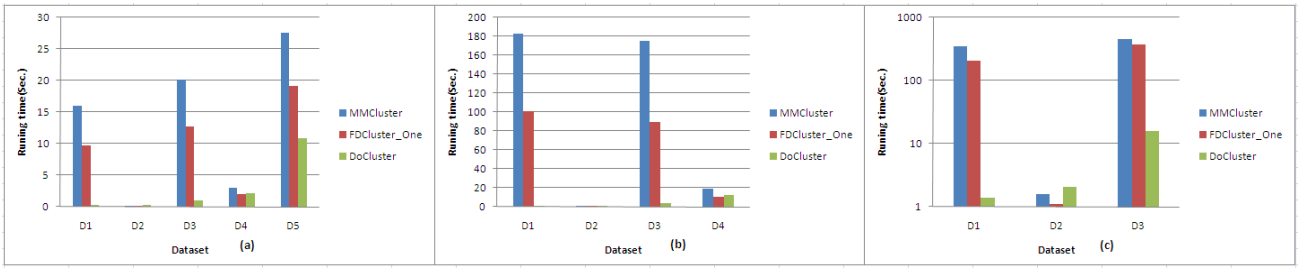


Figure 4. The comparison of performance periods of the above three algorithms under each data set when the number of functions is 35: (a) 200 resources；(b) 500 resources；(c) 800 resources

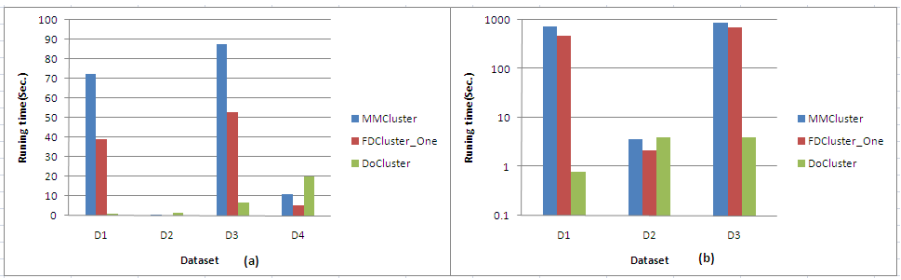


Figure 5. The comparison of performance periods of the above three algorithms under each data set when the number of functions is 50: (a) 200 resources；(b) 500 resources

IV. Experimental Result and Analysis

In this section, we will make an experimental comparison on the mining efficiency and result of the algorithm above and existing algorithms. The hardware environment of the experiment is desktop computer: Intel(R) Core(TM)2 Duo 2.53GHz CPU and 4G memory; the software environment is Microsoft Windows 7 SP1 operating system; the algorithm programming and operating environment is Microsoft Visual C++ 6.0 SP6. Experimental data used in this paper are simulation data. To fully test the performance of the algorithm, we produce five data sets randomly, each of which contains 50 functions and 800 resources. Table 6 describes proportions of 1, 0 and -1 in each row in each data set.

TABLE VI.
THE PROPORTION OF EACH VALUE IN FIVE DATA SET

	1	0	-1
D_1	0.2	0.7	0.1
D_2	0.3	0.6	0.1
D_3	0.1	0.7	0.2
D_4	0.2	0.6	0.2
D_5	0.1	0.6	0.3

In this section, the comparison will be made on the mining efficiency of *DoCluster* algorithm *FDCluster_One* algorithm and *MMCluster* algorithm. *FDCluster_One* algorithm adopts prior detection method described in literature [18]: mine maximal bicluster from discretized matrix data without candidate maintenance. The mining process of *MMCluster* algorithm and *FDCluster_One* algorithm is basically the same. The difference is that during design of pruning strategy, *MMCluster* algorithm first judges whether the gene set of current potential samples is the subset of a prior candidate sample set, while *FDCluster_One* algorithm first calculates the intersection and then judges prior samples.

The mining efficiency of the above three algorithms is compared as follows. To fully compare the scalability of algorithms, we produce multiple groups of data sets with different numbers of resources and functions in allusion to five data sets in Table 6. The selection of resources and functions are based on the order of resources and functions in data set. Figures 4(a)-4(c) provide the comparison of performance periods of the above three algorithms under each data set when the number of functions is 20 and the number of resources is 200, 500 and 800 respectively. It can be seen from these figures that the mining time of each algorithm increases progressively with the increase in the proportion of '-1' in the data set. This is because the biclusters with variant usage rate and low usage rate do not restrain the number of '-1'. Thus, as the number of '-1' increases in the data set, the scale of the bicluster mined will increase continuously, thus increasing mining complexity of each algorithm. It thus can be seen, for mining of function-resource matrix, the number proportion of '-1' in the data set directly influences the complexly of the algorithm. But, when the proportion of '-1' is certain, as the proportion of '1' increases in the data set, the complexly of the algorithm also increases. For data sets D_1 and D_2, the three algorithms can complete mining within 0.5s. The efficiency superiority of *DoCluster* algorithm is not obvious. However, as the proportion of '-1' in the data set increases, in data sets D_3, D_4 and D_5, the pruning strategy of *DoCluster* algorithm displays efficiency superiority.

To further test and verify the scalability of algorithms, figures 5(a)-5(c) provide the comparison of performance periods of the above three algorithms under each data set when the number of functions is 35 and the number of resources is 200, 500 and 800 respectively; figures 6(a)-6(c) provide the comparison of performance periods of the above three algorithms under each data set when the number of functions is 50 and the number of resources is 200 and 500, respectively. It can be seen from these figures that the mining efficiency of the three algorithms declines significantly compared with Fig.4 with the increase in the number of functions. This is because the three algorithms adopt row extension for mining. As the number of samples in the data set increases, mining depth and complexity of the algorithms increase. Meanwhile, the number of prior candidate samples for pruning judgment will also increase, thus increasing pruning complexity. In most data sets shown in Fig.5 and 6, the mining efficiency of *DoCluster* algorithm is the highest. However, in the data sets with large proportion of '1', *DoCluster* algorithm fails to show the advantage of mining efficiency. This may be because multiple biclusters including '1' can exist simultaneously in the biclusters mined by *MMCluster* algorithm and *FDCluster_One* algorithm, while at most one '1' can be included in a bicluster mined by *DoCluster* algorithm due to the restraint of the bicluster with variant usage rate. So, the number of biclusters mined by *DoCluster* algorithm is greater than the above two algorithms, thus including the pruning efficiency of the algorithm.

V. CONCLUSION

This paper proposed an efficient algorithm - *DoCluster* algorithm which can effectively mine all biclusters with maximal variant usage rate and low usage rate from the discrete function-resource matrix. First, this algorithm constructs a sample weighted graph which includes all resource collections between both samples that satisfy the definition of variant usage rate or low usage rate; then, all biclusters with maximal variant usage rate and low usage rate meeting the definition are mined with the mining method of using sample-growth and depth-first method in the constructed weighted graph. To improve the mining efficiency of the algorithm, *DoCluster* algorithm uses several pruning strategies to ensure mining maximal bicluster without candidate maintenance. However, original data information will be lost if the mining is conducted in discrete data. Our next research direction is to mine biclusters with variant usage rate and low usage rate in real function-resource matrix.

ACKNOWLEDGMENT

This paper is supported by National Key Basic Research Program of China under Grant No. 2014CB744900.

REFERENCES

[1] Michael Pecht,et al.. A prognostics and health management roadmap for information and electronics-rich systems. Microelectronics Reliability, 2010:317–323.

[2] Y. Cheng, G.M. Church, "Biclustering of Expression Data," Proc. 8th Int'l Conf. Intelligent Systems for Molecular Biology (ISMB00), ACM Press, 2000, pp: 93–103.

[3] Cui Xiang, Yin Guisheng, Zhang Long, Kang Yongjin. Method of Collaborative Filtering Based on Uncertain User Interests Cluster. Journal of computers, Vol 8, No 1, 2013, p:186-193.

[4] Jing Zhang, Gongqing Wu, Xuegang Hu, Shiying Li, Shuilong Hao. A Parallel Clustering Algorithm with MPI –

MKmeans. Journal of computers, Vol 8, No 1, 2013, p:10-17.
[5] Olovnikov I, Le Thomas A, Aravin A A. A Framework for piRNA Cluster Manipulation. PIWI-Interacting RNAs. Humana Press, 2014: 47-58.
[6] Fotso H, Yang S, Hafermann H, et al. Extended Correlation in Strongly Correlated Systems, Beyond Dynamical Cluster Approximation. Bulletin of the American Physical Society, 2012, 57.
[7] Xiao Xue, Zhe Wei, Zhifeng Zeng. The Design of Service System for SMEs Collaborative Alliance: Cluster Supply Chain. Journal of Software, Vol 6, No 11, 2011, p:2146-2153.
[8] Ling-ling Pei, Zheng-xin Wang. An Optimized Grey Cluster Model for Evaluating Quality of Labor Force. Journal of Software, Vol 8, No 10, 2013, p:2489-2494.
[9] Yu Wang, Youfang Huang, Huiqiang Zheng, Daofang Chang. Quay Crane Allocation of Container Terminal Based on Cluster Analysis. Journal of Software, Vol 8, No 5, 2013, p:1201-1208.
[10] Ben, et al. Discovering local structure in gene expression data: the order-preserving submatrix problem. J. Comput. Biol, 2003; 10: 373-384.
[11] Cheng et al. Bivisu: software tool for bicluster detection and visualization. Bioinformatics, 2007, 23: 2342-2344.
[12] Lizhuang Zhao, Mohammed J. Zaki, MicroCluster: An Efficient Deterministic Biclustering Algorithm for Microarray Data, in IEEE Intelligent Systems, special issue on Data Mining for Bioinformatics, 2005, Vol. 20, No. 6, pp: 40-49.
[13] U. Maulik, A. Mukhopadhyay, M. Bhattacharyya, L. Kaderali, B. Brors, S. Bandyopadhyay, and R. Eils. Mining Quasi-Bicliques from HIV-1-Human Protein Interaction Network: A Multiobjective Biclustering Approach. IEEE-ACM Transactions on Computational Biology and Bioinformatics, vol.10, 2013, pp.423-435.
[14] de Sousa Filho G F, dos Anjos F Cabral L, Ochi L S, et al. Hybrid Metaheuristic for Bicluster Editing Problem. Electronic Notes in Discrete Mathematics, 2012, 39: 35-42.
[15] Király A, Abonyi J, Laiho A, et al. Biclustering of High-throughput Gene Expression Data with Bicluster Miner. Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference on. IEEE, 2012: 131-138.
[16] Desai B, Andhale P, Rege M, et al. Biclustering and feature selection techniques in bioinformatics. Data Engineering and Management. Springer Berlin Heidelberg, 2012: 280-287.
[17] Pio G, Ceci M, D'Elia D, et al. A novel biclustering algorithm for the discovery of meaningful biological correlations between miRNAs and mRNAs. EMBnet. journal, 2012, 18(A): pp. 43-44.
[18] Miao Wang, Xuequn Shang, Shaohua Zhang, Zhanhuai Li. FDCluster ： Mining frequent closed discriminative bicluster without candidate maintenance in multiple microarray datasets. ICDM 2010 workshop on Biological Data Mining and its Applications in Healthcare, p 779-786.
[19] Lucinda K. Southworth, Art B et al, Aging Mice Show a Decreasing Correlation of Gene, PLoS Genetics, December 2009, Volume 5,Issue 12.
[20] O. Odibat, C. K. Reddy and C. N. Giroux. Differential biclustering for gene expression analysis. In Proceedings of the ACM Conference on Bioinformatics and Computational Biology (BCB), 2010, p: 275–284.
[21] G. Fang, R. Kuang, G. Pandey, M. Steinbach, Chad L. Myers and V. Kumar. Subspace Differential Coexpression Analysis: Problem Definition and A General Approach. Proceedings of the 15th Pacific Symposium on Biocomputing(PSB), 2010, 15:145-156.
[22] A. Serin and M. Vingron. Debi: Discovering differentially expressed biclusters using a frequent itemset approach. Algorithms for Molecular Biology, vol. 6, no. 1, 2011, p:18-29.
[23] Miao Wang, Xuequn Shang, Miao Miao, Zhanhuai Li, Wenbin Liu. FTCluster: Efficient Mining Fault-Tolerant Biclusters in Microarray Dataset. Proceedings of ICDM 2011 workshop on Biological Data Mining and its Applications in Healthcare, p 1075-1082.
[24] Wang, J, Han, J. BIDE: Efficient Mining of Frequent Closed Sequences, Data Engineering, 2004. Proceedings. p: 79 – 90.

Lihua Zhang is a doctoral student at the school of computer science and engineering at the northwestern polytechnical university, Xi'an China. She completed her master degree from northwestern polytechnical university in 2008. Her current research interests are PHM, avionics, data mining and safety. Since 2013, she has been studying at science and technology on avionics integration laboratory.

Miao Wang is an engineer at science and technology on avionics integration laboratory. He completed his doctor and master degree from northwestern polytechnical university in 2013 and 2018, respectively. He is a member of China computer federation. His research interests mainly include data mining, PHM, avionics and safety.

Zhengjun Zhai is a professor at the school of computer science and engineering at the northwestern polytechnical university, Xi'an China. He is vice chairman of NPU youth association for science and technology, distinguished expert of aerospace electrical & electronics and weapon system Standardization technology committee, distinguished expert of AAMRI and premium member of china computer federation. His research interests include experiment and testing systems Integration, remote maintenance and fault diagnosis and virtual visualization.

Guoqing Wang is a professor and a supervisor of Ph.D. student in Northwestern Polytechnical University. He was born in 1956 and received his M.S. and Ph.D. degrees in computer science and technology from the Northwestern Polytechnical University in 1984 and 1991 respectively. He is the institute director of China aeronautical radio electronics research institute, and the director of science and technology on avionics integration laboratory. He has long been engaged in the related technical research of avionic system integration, distributed parallel processing, high reliable fault-tolerant system, network and bus system etc. He serves as the vice director of national serve environment computer academy.

A Dynamic Architecture for Mobility Management in Hierarchical Mobile IPv6

Jianmin Chen[1,2], Zhongyang Xiong[1], Peng Yang[2], Yuanbing Zheng[3] , Chunyong Liu[1],Guangyong Li[1]
1 College of Computer Science, Chongqing University,Chongqing, China
2 School of Information Engineering, Nanchang Hangkong University,Nanchang, China
3 Chongqing Electric Power Information & Communication Branch Company; Chongqing, China;
Email:jm_chenn@163.com, zyxiong@cqu.edu.cn

***Abstract*—Hierarchical Mobile IPv6 (HMIPv6) is an enhanced Mobile IPv6 for reducing signaling cost of location management. Multi-level Hierarchical Mobile IPv6 (MHMIPv6) can organize mobile region as a multi-level hierarchy architecture, which is more flexible to support scalable services. However, MHMIPv6 will bring additional packet processing overhead, and produce negative impact especially on some mobile nodes (MNs) with relatively low movement characteristics. This paper proposes a dynamic hierarchical Mobile IPv6 (DHMIPv6) management, in which different hierarchies are dynamically set up to minimize the total cost for different MNs according to their movement characteristics respectively. Under such management MNs can select the monolayer or two-layer mobility anchor point(MAP) structure when they occur the handover at any time. Experimental results show that compared with HMIPv6 and MHMIPv6, DHMIPv6 achieves high adaptability with lower total cost under various scenarios.**

***Index Terms* —Hierarchical Mobile IPv6, Multi-level Hierarchical Mobile IPv6, dynamic hierarchical architecture, Adaptability**

I. INTRODUCTION

Hierarchical Mobile IPv6 management (HMIPv6) [1] divides mobile node's (MN) mobility [2] into micro-mobility and macro-mobility. When a MN moves within a particularly hierarchical domain, then micro-mobility; In this case, HMIPv6 utilize local mobility management to reduce the amount of signaling generated by the registration to the correspondent nodes (CNs) and to the home agent (HA).when the MN moves out to a new domain, then macro-mobility, the mobility of the MN will be managed by the standard Mobile IPv6 management (MIPv6) [3]. Mobile Anchor Point (MAP) is a substitute of "Home Agent" (HA) in each domain of the network which hides user's mobility from the outer domain. Then the binding updates are sent from MN directly to MAP rather than more distant HA or CNs when the MN stays in a specific region; meaning that MN's exact position is hidden from outer region and the signaling overhead is reduced. The MN needs to register its position to HA and CNs when it moves out of the specific region, just like the standard MIPv6.

Shengling Wang et al proposed a model to analyze the application scopes of MIPv6 and HMIPv6 [4]. In [4] ,Wang presented HMIPv6 does not always outperform MIPv6 , The analytic model in [4] can choose the better alternative between MIPv6 and HMIPv6 according to the mobility and service characteristics of users and can choose the best mobility anchor point and regional size when HMIPv6 is adopted, addressing how to hierarchize the network..

There are also some works considered the MN's characteristics to get a better performance [5] [6] [7].Xie et al. proposed an analytic model for Mobile IP regional registration which is one of hierarchical mobility management schemes [5]. The proposed analytic model focused on the determination of the optimal size of regional networks, given the average total location update and packet delivery cost. Besides, Ma introduced a dynamic hierarchical mobility management strategy for mobile IP networks, in which different hierarchies are dynamically set up for different users and the signaling burden is evenly distributed among the network [6]. [7] introduced two static MAP selection schemes: the furthest MAP selection scheme, the nearest MAP selection scheme and two dynamic MAP selection schemes: the mobility-based and the adaptive MAP selection schemes. [7] proposed that the dynamic schemes are better than the static schemes since the dynamic schemes can select the serving MAP depending on the MN's characteristics In addition, the dynamic MAP selection schemes achieve better load balancing.

To reduce signaling cost further, Multilevel Hierarchical Mobile IPv6 management (MHIPv6) [8] subdivide mobile region of MNs to multiple levels. The MAPs are placed at all levels of hierarchical structure of network, The MAP of the lower layer is responsible for location management and data delivery of MNs in a smaller range reign, However the higher MAP has a wider service domain through the lower layer MAPs beneath itself. So an MN moves out to another, which MAP Will be selected to register depend to the MN's changing mobility. MHMIPv6 is more flexible than

doi:10.4304/jcp.9.5.1168-1176

HMIPv6, and the signaling cost will be reduced even further. But it also presents new problems. First, the high level MAP is responsible for the management and data forwarding of all mobile nodes in a large area. The Center fixed structure may result in high load for the high level MAP, and the failure of the high level MAP cause the communication interrupt for all MNs. Secondly , packets from the communication nodes sent to mobile node need be stored and forward through the MAPs at all levels, that significantly increased the network load.

This paper proposes a dynamic two-level MAP hierarchical mobile IPv6 model (DHMIPv6). The model of physical structure is the monolayer MAP, but can implement the two levels of structure logically. Compare to HMPIv6, the model reduce the number of home registration for the mobile node, so reduce the costs. At the same time it avoids the heavy load and system robustness problems brought by center fixed structure of the MHMIPv6. In this paper, we also introduce a new mathematical model to calculate the smallest overall costs of location registration and packet delivery for determining the proper location registration strategy when a mobile node moves among subnets. This model considers the factors such as various network parameters, movement characteristics of mobile nodes etc.

This paper is organized as follows: In Section II, the introduction of HMIPv6 and MHMIPv6. In Section III, the mobility model (DHMIPv6) is described and a method for deriving the total location update and packet delivery cost is introduced. In Section IV, we analyzed the total costs of location registration and packet delivery in DHMIPv6; determine the proper location registration strategy to minimize the total costs. In Section V. Comparison results of system performance between DHMIPv6 and HMIPv6, MHIPv6 is presents. Section VI gives the conclusions.

II. HMIPv6 AND MHMIPv6

A. HMIPv6

The system architecture of HMIPv6 is shown in Fig.1. The HMIPv6 uses the MAP to provide agent service for MNs at foreign subnet, The MN will receive router notice which Contains one or more MAP information option when it moves from one subnet to another. The MN checks if the MAP domain has changed according to the MAP information. If the MAP domain has changed, The MN can obtain a new link care-of address (LCOA) and a regional care-of address (RCOA). An MN needs to register its LCOA with the MAP for the routing purpose, and register the new RCOA to the home agent and correspondent nodes. Then, all packets to MN will through the new MAP forward. If the MN find MAP domain has not changed, it only needs to register its LCOA with the MAP. So the home registration is not need when switchover happened between two subnets of the same MAP domain. In this case, the mobility of MN is transparent to home agent and correspondent node, it reduce relatively registration cost and time delay. [9] proposed an adaptive network mobility support protocol based on hierarchical mobile IPv6. [10][11] improved the performance through reduce the handover latency. In HMIPv6.Hierarchical mobile management is also used in other mobility management scheme, such as Cellular IP [12], HAWAII [13], Wireless Ad Hoc Networks [14] etc.

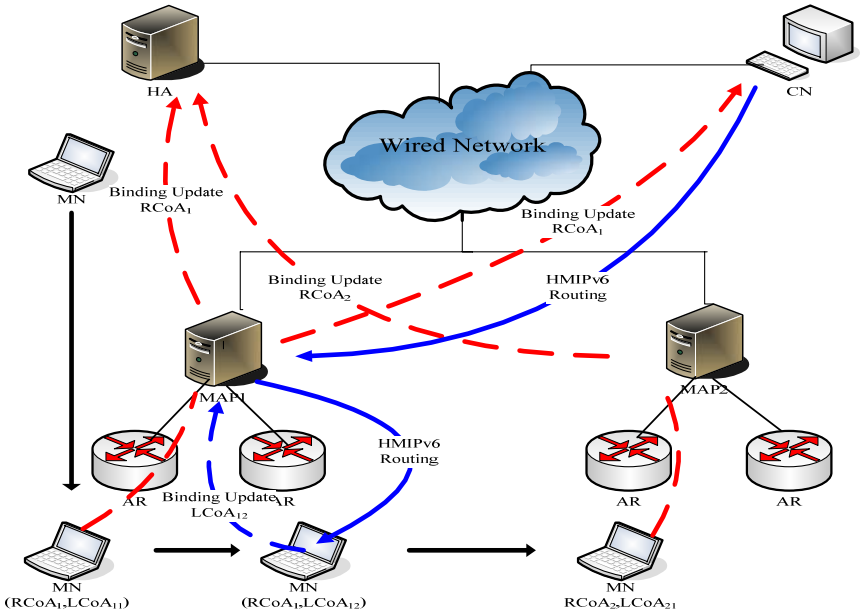


Figure 1. Position registration and packet routing in HMIPv6

B. MHMIPv6

MAPs are organized into a tree structure in MHMIPv6, as shown in Fig.2, The top MAP of the tree is the root MAP (RMAP). The MAPs at lowest layer are the leaf MAPs (LMAP), When a switchover of a MN happened, the MN register its LCOA with the leaf MAP at first, then the LMAP check if the MN has existed in its binging list. If existed, the LMAP need only to reply the binging update, else the LAMP need to register to the higher level MAP, Analogously, if a MAP has received a binging update, it will check if the MN has existed in its binging list. If the MN has registered, the MAP will finish the registration process after it confirm the binging update, else the MAP send sequentially registration Binding Update to the higher level MAP until RMAP. This process is repeated in each MAP in the hierarchy until a MAP having the MN in its mapping table can be found. In this way, the binding update delivery through multiple levels MAP to RMAP even correspondent nodes and the home agent. Therefore, the first binding update in a foreign network, the BU message is forwarded up to the RMAP in the foreign network and the HA. When a packet is delivered to an MN from a correspondent node, it will carry forwarded by all levels MAPs. It is possible to provide more scalable services and to support a larger number of MNs in MHMIPv6, and the MHMIPv6 is more flexible than HMIPv6 which only divides MN's mobility into two situations: global and regional. However, the MHMIPv6 results in a higher processing cost than the HMIPv6 when a packet is delivered to an MN. This is because the packet goes through more intermediate MAPs and the encapsulation/decapsulation procedures are repeated at each MAP, and if a MAP failed, all the tree structure under the MAP will fail, it will also affect the robustness of the entire network[15].[15] developed an analytic model based on MHMIPv6 architecture to calculate optimal hierarchy too.

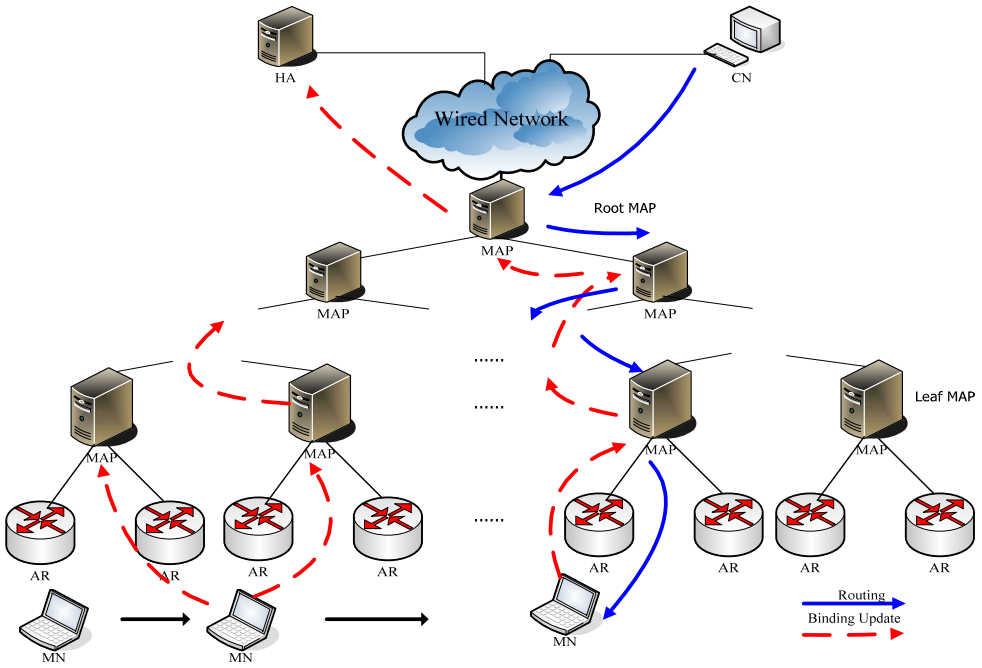


Figure 2. Position registration and packet routing in MHMIPv6

III OUR PROPOSED DHMIPV6

In MHMIPv6, the high level MAP is responsible for packets forwarding of all mobile nodes in a large domain. So, it is likely that they will meet the problem of load. The centralized fixed scheme of the MHMIPv6 has single point invalidation that the failure of a top MAP will lead to communication interrupt of a large area. And the packet sent to the MN from correspondent node through more intermediate MAPs and the encapsulation/decapsulation procedures are repeated at each level MAP which also increased the load on the network .To solve these problems, we propose a new mobility management scheme which has the same physical structure with HMIPv6 but has dynamic two levels MAP logically.In the mobility scheme each MAP can function either as an Root MAP or a leaf MAP or a single level MAP , and an MAP should act as Which types of MAP depends on the user mobility. Thus, the traffic load in a regional network is evenly distributed to

each MAP. Through this approach, the system robustness is enhanced. We also propose how to adjust the number of LMAPs under a RMAP for each MN according to the user-variant and time-variant user parameters in the dynamic scheme. In this dynamic system, there is no fixed regional network boundary for each MN. An MN decides when to perform a home location update according to its changing mobility, packet arrival pattern and the number of correspondent nodes. The detailed analysis will be given in Section IV.

As shown in Fig.3, MN1 First enters the MAP1 domain, Passes through MAP2 domain and MAP3 domain next, finally reaches MAP4 domain. At first MN1 enter the MAP1 domain, it performs a home registration through MAP1, and obtains RCOA (regional care of address) and LCOA (Online care to address).MN1 perform the procedure of location management and packet routing same with the standard HMIPv6 at t-his time.MN1 do not perform the home registration when it enters MAP2 domain or MAP3 domain to reduce the cost of registration. But it registers to the MAP1 with the regional care-of ad-dress obtained in new MAP as a online care-of address domain. Packets from CNs forwarding to current MAP domain of MN1 through MAP1 too and finally submitted to the MN1. So we implement two levels of structure MAP logically. With this method, when MN1 turn crossed k different MAP domain (in this case, the value of k is 3) to reach a new MAP (MAP4) domain, MN1 performs a home registration, and repeat the above steps.

As shown in Fig.3, When MN1 and MN2 in MAP3 domain at the same time. For MN1, MAP1 is MAP3's Upper MAP, but MAP5 is MAP3's Upper MAP for MN2, Such the system assigned functions of upper MAP on o MAP to the two MAPs. So that the influence brought by the failure of MAP1 or MAP5 is relatively limited; And the functions of high-level MAP allocated to each MAP, so that it can avoid single location invalidation and load balancing problem in fixed center structure effectively.

The protocol descriptions of packet delivery and location management in DHMIPv6 are shown in Table I and Table II respectively.

TABLE I

DH-MIPV6 LOCATION DH-MIPV6 PACKET DELIVERY PROTOCOL

```
if (a correspondent node deliver the packet to an MAP which address
registered in the correspondent node with a MN binding?)
   if(The MAP is the LMAP of the MN？ )
      Packet submitted   to the   LMAP address registered by MN
      through the tunnel
   else
      Packet submitted to the MN through the tunnel from the LMAP
end
end.
```

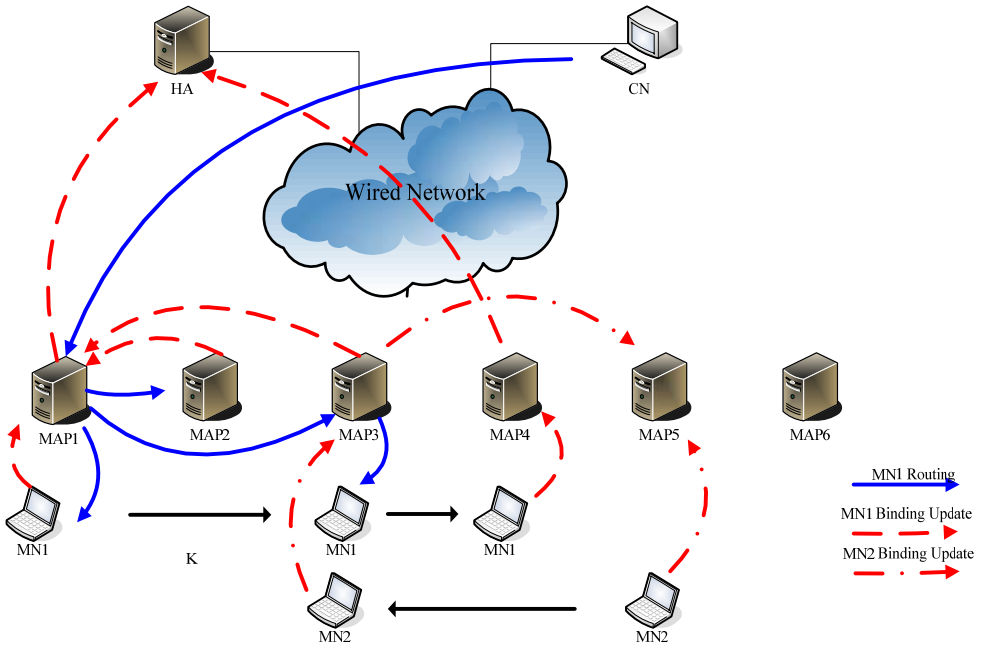


Figure 3. Position registration and packet routing in DHMIPv6

TABLE II

DH-MIPV6 LOCATION MANAGEMENT PROTOCOL

```
if (MN enters a new subnet?)
  if(The MN enters a new MAP domain?)
    compare the address of the new MAP to the addresses in buffer
      if(the new address is one of addresses in buffer?)
      MN perform a regional registration to the selected Previously
      RMAP through the new MAP
  else
      if (The number of addresses in buffer>=Kopt?)
     compute the new Kopt according to the movement characteristics
     of MN
     delete all the addresses in buffer
     Select the MAP as the new RMAP, perform a home registration
     through it
      else
     MN perform a regional registration to the selected Previously
     RMAP through the new MAP
      end
      record the new map address in buffer
    end
else
    Perform a region registration to the Current MAP
  end
end
```

IV ANALYTICAL MODE

A. Location Update Cost

In our model, a MN may perform a home registration or may perform a regional registration to logical RMAP or logical LMAP when it enters a new subnet. Then these location cost can be expressed respectively as Eq. (1), Eq. (2), Eq. (3).

the location update cost of home registration is:

$$C'_{UH} = \omega + \eta(d_{ar-MAP} + d_{MAP-HA}) + 2a_{MAP} + a_{HA} \quad (1)$$

the location update cost of regional registration to logical RMAP is:

$$C'_{URl} = \omega + \eta(d_{ar-LMAP} + d_{LMAP-HMAP}) + 2a_{LMAP} + a_{HMAP} \quad (2)$$

the location update cost of regional registration to logical LMAP is:

$$C'_{UR2} = \omega + \eta d_{ar-MAP} + a_{MAP} \quad (3)$$

the location update cost of registration to CN can be expressed as Eq. (4):

$$C'_{UC} = \omega + \eta(d_{ar-MAP} + d_{MAPCN}) + 2a_{ar} + 2a_{MAP} + a_{CN} \quad (4)$$

Where ω and η are the unit costs when a location update procedure is performed in a wireless and a wired link. d_{ar-MAP} denotes a distance between AR and MAP，d_{MAP-HA} denotes a distance between HA and MAP. a_{MAP} is the processing cost of location update at the MAP, a_{HA} is the processing cost of location update at the HA, a_{CN} is the processing cost of location update at the CN, $d_{LMAP-HMAP}$ denotes a distance between upper MAP and lower MAP. a_{HMAP} is the processing cost of location update at the upper MAP，a_{LMAP} is the processing cost of location update at the lower MAP.

We call the action an MN moving out of a subnet "a movement". Define a random variable M so that an MN perform a home registration at movement M. we assume that the total number of MAPs is N, the number of subnets located within a MAP domain is L. In order to facilitate the analysis, we assume the MN will move out to the other subnets with equal probability, so when the MN perform a handover, it moves into the others subnets with equal probability $P = 1/NL-1$. For the standard HMIPv6 scheme, the probability of performing a home registration at movement M is:

$$P_M = \frac{NL-L}{NL-1} * (\frac{L-1}{NL-1})^{M-2} \quad (5)$$

It can be shown that the expectation of M is:

$$E \mid M \mid = \sum_{M=2}^{\infty} MP_M = \frac{2NL-L-1}{NL-L} \quad (6)$$

For DHMIPv6 system architecture, the MN may not need to perform the home registration when it occur the handover between two MAP domains. A MN take a MAP_0 that it passed before as the upper agent to reduce the number of home registrations according to the movement characteristics of the MN. So there are a MAP group that the MAP_0 is the center. Then the MN do not need to perform the home registration when it moves into a MAP domain from another MAP domain inside the MAP group domain, but need to perform the regional registration to MAP_0. The DHMIPv6 system operates similarly the standard HMIPv6 when the MN stay in the domain of MAP_0 .But when the MN stay in the domain of the others k-1 MAPs, it operates similarly two level MHMIPv6 that MAP_0 is the upper MAP. Define a random variable D so that an MN perform a home

registration at movement D in DHMIPv6 system architecture, in other words, the MN move out of the domain of the MAP group at movement D.

Define the expectation of the number of movements it takes an MN moving from its first MAP domain to its second new MAP domain as $E \mid D \mid_{1\to 2}$:

$$E \mid D \mid_{1\to 2} = 1 \tag{7}$$

Similarly, when an MN has visited two different MAP domains, define the expectation of the number of movements it takes an MN moving to its third new MAP domain as $E \mid D \mid_{2\to 3}$:

$$E \mid D \mid_{2\to 3} = \sum_{n=1}^{\infty} n \cdot (\frac{1}{N-1})^{n-1} \frac{N-2}{N-1} = \frac{N-1}{N-2} \tag{8}$$

the expectation of an MN moving to its K new MAP domain and an MN moving out of the MAP group domain can be shown respectively:

$$E \mid D \mid_{k-1\to k} = \sum_{n=1}^{\infty} n \cdot (\frac{k-1}{N-1})^{n-1} \frac{N-k+1}{N-1} = \frac{N-1}{N-k+1} \tag{9}$$

$$E \mid D \mid_{k\to k+1} = \sum_{n=1}^{\infty} n \cdot (\frac{k}{N-1})^{n-1} \frac{N-k}{N-1} = \frac{N-1}{N-k} \tag{10}$$

So the expectation of D be obtained as Eq.(11):

$$\begin{aligned} E|D| &= E|D|_{1\to 2} + E|D|_{2\to 3} + \cdots + E|D|_{k-1\to k} + E|D|_{k\to k+1} \\ &= 1 + \frac{N-1}{N-2} + \cdots + \frac{N-1}{N-k+1} + \frac{N-1}{N-k} \\ &= (N-1)\sum_{i=1}^{k} \frac{1}{N-i} \end{aligned} \tag{11}$$

Define a random variable R to describe the number of the MN stay in the domain of MAP_0 before it moves out of the domain of the MAP group , the expectation of R can be obtained as Eq.(12):

$$\begin{aligned} E \mid R \mid &= 1 + \frac{1}{2} E \mid D \mid_{2-3} + \cdots + \frac{1}{k} E \mid D \mid_{k\to k+1} \\ &= \sum_{i=1}^{k} \frac{N-1}{(N-i)i} \end{aligned} \tag{12}$$

Hence, the location update cost in our model can be obtained from the following equations:

$$\begin{aligned} C'_{LU} = C'_{UH} &+ (E \mid D \mid \cdot E \mid M \mid -E \mid D \mid + E \mid R \mid) C'_{UR1} \\ &+ (E \mid D \mid - E \mid R \mid) C'_{UR2} + \sum_{j=1}^{y} C'_{UCj} \end{aligned} \tag{13}$$

The total location update cost in standard HMIPv6 can be obtained as Eq.(14):

$$\begin{aligned} C_{LU} &= (E \mid D \mid -1)(C_{UH} + \sum_{j=1}^{y} C_{UC_j}) \\ &+ (E \mid D \mid \cdot E \mid M \mid -E \mid D \mid) C_{UR} \end{aligned} \tag{14}$$

Where y is the number of correspondent nodes, C'_{UCj} denotes the location update cost that the MN registered to the jth correspondent Node. So, $\sum_{j=1}^{y} C'_{UCj}$ denotes the location update cost that the MN registered to all correspondent Nodes.

B. Packet Delivery Cost

The cost for packet delivery procedure through two MAPs or single MAP can be expressed respectively as:

$$\begin{aligned} C_{P1} &= \rho \cdot \alpha \cdot (d_{CN-HMAP} + d_{HMAP-LMAP} + d_{MAP-AR}) \\ &+ \beta \cdot \alpha \cdot d_{AR-MN} + P_{HMAP} + P_{LMAP} \end{aligned} \tag{15}$$

$$\begin{aligned} C_{P2} &= \rho \cdot \alpha \cdot (d_{CN-MAP} + + d_{MAP-AR}) \\ &+ \beta \cdot \alpha \cdot d_{AR-MN} + P_{MAP}) \end{aligned} \tag{16}$$

Where β the unit transmission cost is in a wireless link, ρ is the unit transmission cost in a wired link. α is the CMR(call-to-mobility ratio) of the MN, In this paper, λ is the arrival rate of packets, the average time a MN stays in each subnet before making a movement is $1/\mu$,the CMR of the MN can be shown as : $\alpha = \lambda/\mu$. P_{MAP} is the processing cost at the MAP. The processing cost at the MAP includes a lookup cost and a packet encapsulation/decapsulation cost. It is assumed that the lookup cost is proportional to the logarithm of the number of MNs located in the MAP domain and the encapsulation/decapsulation cost is a constant value. In the DHMIPv6 system architecture, there are a group of lower MAPs under the management of a upper MAP for a MN, but the function of upper MAPs is assigned to all MAPs. So, in addition to manage $o \cdot L$ MNs in its domain, Each MAP manages others $o \cdot L$ MNs through other MAPs too. So P_{MAP} can be expressed as:

$$P_{MAP} = P_{HMAP} = P_{LMAP} = \delta \log 2oL + g \tag{17}$$

Then, the overall packet delivery cost is the sum of all packet delivery costs can be expressed as:

$$C_P = E \mid M \mid \cdot (E \mid D \mid - E \mid R \mid) \cdot C_{P1} + E \mid M \mid E \mid R \mid \cdot C_{P2} \tag{18}$$

the total cost is the sum of the location update cost and the packet delivery cost is:

$$C_{total} = C_L + C_P \tag{19}$$

The average total cost per unit time is:

$$C_T = (C_L + C_P) \Big/ E|M| \cdot E|D| \tag{20}$$

The unit time is the average time an MN stays in each subnet.

The optimal number of MAPs beneath a MAP for a MN, k_{opt} is defined as the value of k that minimizes the total cost. To investigate the impact of the optimal number and CMR, we formulate the total cost as a function of the optimal number and CMR. The difference function is also defined order to find the optimal number as shown in Eq. (21):

$$\Delta(k,\alpha) = C_T(k,\alpha) - C_T(k-1,\alpha) \tag{21}$$

Using the difference function, it is possible to find the optimal number when the CMR are given:

$$k_{opt} = \begin{Bmatrix} 1 \quad , \quad \Delta(2,\alpha) > 0 \\ \max(k : \Delta(k,\alpha) \le 0) \quad , \quad otherwise \end{Bmatrix} \tag{22}$$

V Experimental Results and Analysis

In this section, we demonstrate the performance improvement of the DHMIPv6 scheme to HMIPv6 scheme and MHMIPv6 scheme of two levels. We have compared our DHMIPv6 scheme with HMIPv6 scheme and MHMIPv6 scheme in terms of Location update cost and packet delivery cost. Table III lists some of the parameters used in our performance analysis. Compared to HMIPv6 scheme, our DHMIPv6 scheme reduces the location update cost but bring the increase of the packet delivery cost. The total costs are affected by several factors: in the case of the unit location update cost, which procedures are performed for location registration determines the update cost. On the other hand, in the case of packet delivery cost, the unit transmission cost is one of the important factors to be considered. Therefore, we analyze various results for different cost sets. Table IV shows the different cost sets.

Fig.4 plots the Kopt as a function of CMR for the DHMIPv6.As shown in the figure, Kopt decreases as CMR increases for DHMIPv6 systems. When the CMR is low, Registration cost accounted for a large proportion of the total cost, a MN of DMHMIPv6 can choose a MAP as the higher MAP to manage a large range of MAPs domain (in this case, Kopt is greater) to reduce the number of home registration. When the CMR is high, the packet delivery cost dominates and the saving in packet delivery cost becomes significant. The MN can choose a smaller Kopt to reduce the packet delivery cost. When the CMR is large enough, the packet delivery cost is far greater than the registration cost .the Kopt can be set to 1 to reduce the packet delivery cost the MN is actually carried out in the standard HMIPv6 management at the time. Fig.4 also shows when the unit cost of the impact on the calculation of Kopt, large unit packet delivery cost or small unit registration costs prompting Kopt to close 1 faster.

TABLE III

Performance Analysis Parameters

d_{ar-MAP}	d_{MAP-CN}	d_{MAP-HA}	d_{MAP-HA}
1	10	10	2
a_{MAP}	a_{HA}	a_{CN}	
2	4	4	
g	N	L	P
2	8	2	5

TABLE IV

Different Cost Sets

	ω	η	β	ρ
Se1	2.0	1.0	1.0	2.0
Set2	10.0	5.0	1.0	2.0
Set3	2.0	1.0	5.0	10.0

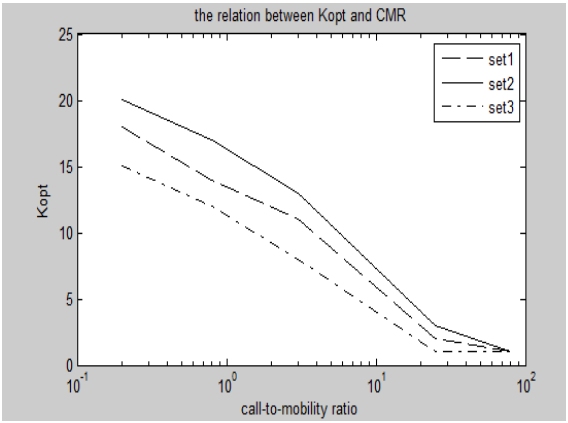


Figure 4. Kopt for DHMIPv6

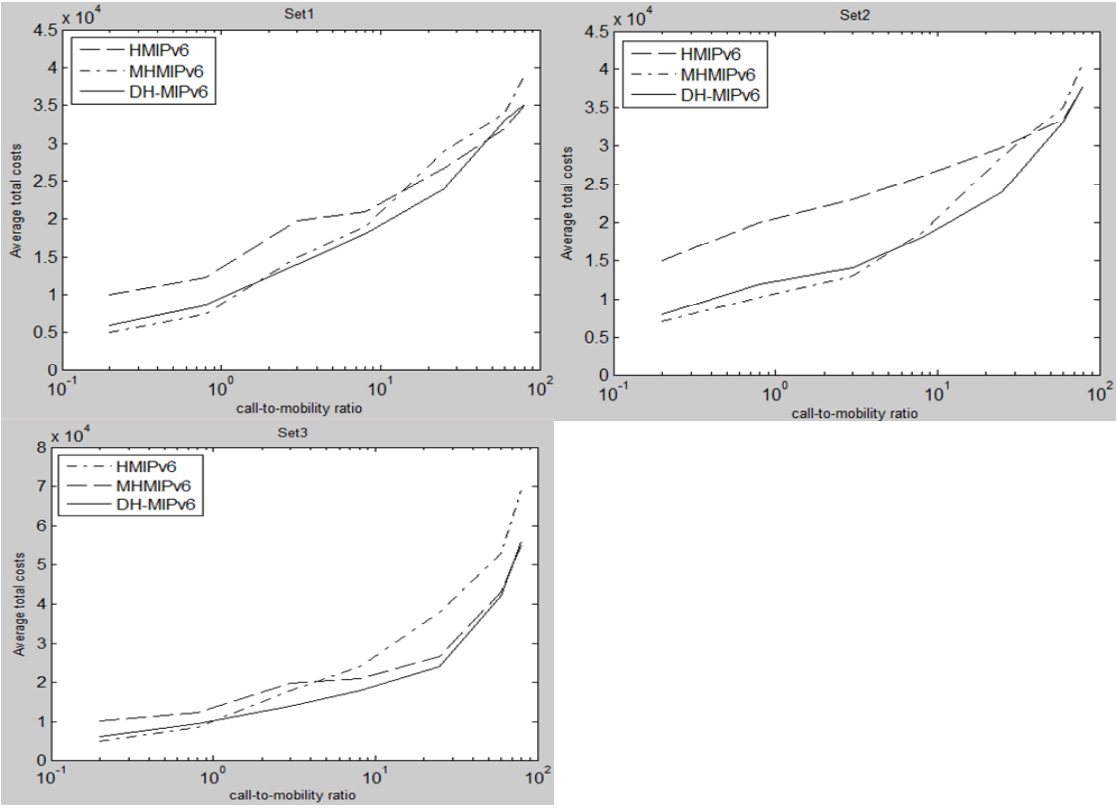


Figure 5. Comparison of the total costs for different cost parameters.

In Fig.5, We compared the average total cost of DHMIPv6, HMIPv6 and MHMIPv6 in three different network environments. We can observe that the average total cost of MHMIPv6 is smaller than the average total cost of HMIPv6 because of the reduction of home registrations when the CMR is small, at the same time DHMIPv6 choose a logical MAP as the high level MAP of a large area, the average total cost reach or near the average total cost of HMIPv6. With the increase of CMR, the MN can adjust their Kopt to get the better average total cost in DHMIPv6.So we can observe the average total cost of DHMIPv6 is smaller than HMIPv6 and MHMIPv6 sometimes. When CMR is large enough, the location management cost act as a small proportion of total cost, the average total cost of HMIPv6 is smaller than the average total cost of MHMIPv6 because of the encapsulation/decapsulation procedure occurred during forwarding process at each MAP. At this time, DHMIPv6 tend to similitude HMIPv6, it has roughly the same average total cost of HMIPv6. We can also get a Conclusion from the figure, in the three different network environments, DHMIPv6 have a good adaptability in comparison with HMIPv6 and MHMIPv6. It have a relatively good total average cost regardless of the CMR is large or small.

VI CONCLUSIONS

HMIPv6 using MAPs as the foreign agents for MN, MHMIPv6 divided the foreign network into multiple levels of domains. MHMIPv6 is a more flexible mobility management but it also brings new problems: high-level load and single point of failure problem and the extra cost of forwarding data between multi-level MAPs. This paper proposes a dynamic two-level MAP hierarchical mobile IPv6 model (DHMIPv6), When a MN moved into a new MAP domain, the home registration is not necessarily, the MN may choose the previous MAP as the high-level MAP, and register to it through the current MAP, or may perform a home registration, choose a one level or two levels MAP structure adaptively. Then, it dose not only reduce the number of home registration, but also avoid load and single point of failure problem. In case of large data flow, DHMIPv6 using the monolayer structure, it reduce the packet delivery cost of data forwarding process between multiple MAPs, and Well adapted to the variation of the movement characteristics of MNs.

REFERENCES

[1] H.Soliman, C.Castelluccia, K.E.Malki, and L.Bellier, “Hierarchical Mobile IPv6 Mobility Management (HMIPv6),”IETF RFC 4140,Aug.2005.
[2] C.Perkins,“IP Mobility Support in IPv4,”IETF RFC3344,Aug.2002.
[3] D.Johnson, C.Perkins, and J.Arkko,“Mobility Support in IPv6,” IETF RFC 3775,June2003.
[4] Shengling Wang; Yong Cui; Das, S.K.; Wei Li and Jianping Wu” Mobility in IPv6: Whether and How to Hierarchize the Network?” IEEE Transaction on parallel and distributed systems, vol 22, no10, October 2011.

[5] J.Xie, and I.F.Akyildiz, "A Distributed Dynamic Regional Location Management Scheme for Mobile IP," IEEE Trans. Mobile Computing, Vol.1, No.3, July 2002.
[6] W. Ma and Y. Fang, "Dynamic hierarchical mobility management strategy for mobile IP networks," IEEE Journal on Selected Area in Communications, vol. 22, no.4, May 2004.
[7] S.Pack, T.Kwon, and Y.Choi, "Aperformance comparison of mobility anchor point selection schemes in hierarchical mobile ipv6 networks," ComputerNetwork, vol.51, 2007.
[8] Kawano K., Kinoshita K, and Murakami K., "Multilevel Hierarchical Mobility Management Scheme in Complicated Structured Networks," Proc. of 29th Annual IEEE International Conference on Local Computer Networks, pp. 34-41, November 2004.
[9] SangheonPack, Taekyoung Kwon, YangheeChoi, and Eun Kyoung Paik," An Adaptive Network Mobility Support Protocol in Hierarchical Mobile IPv6 Networks" IEEE Transaction on vehicular technology, vol 58,no7 ,September 2009.
[10] N. Montavont et al., "Handover Management for Mobile Nodes in IPv6 Network," IEEE Communications Magazine, August 2002.
[11] Wei Kuang Lai and Jung Chia Chiu, "Improving Handoff Performance in Wireless Overlay Networks by Switching Between Two-Layer IPv6 and One-Layer IPv6 Addressing", IEEE journal on selected areas in communications,vol.23,NO.11,November 2005.
[12] A.Campbell, J.Gomez, S.Kim, A.Valko, C.Wan,andZ.Turanyi,"Design, Implementation and Evaluation of Cellular IP," IEEE Personal Comm. Magazine, vol.7,no.4,pp.42-49,Aug.2000.
[13] R.Ramjee, T.Porta, S.Thuel, K.Varadhan, and S.Wang, "HAWAII: A Domain-Based Approach for Supporting Mobility In Wide-Area Wireless Networks," IEEE/ACM Trans. Networking, vol.6, no.2, pp. 396-410, June2002.
[14] Tie Qiu, Wei Wang, Feng Xia, Guowei Wu, and Yu Zhou, "A Failure Self-recovery Strategy with Balanced Energy Consumption for Wireless Ad Hoc Networks". Journal of Computers, Vol 7, No 1 (2012), 116-123, Jan 2012.
[15] Sangheon Pack, Minji Nam, and Yanghee Choi "A Study On Optimal Hierarchy in Multi-Level Hierarchical Mobile IPv6 Networks". IEEE Communications Society Globecom 2004.
[16] Dongbo Liu, Peng Xiao, "User-oriented Mobile Filesystem Middleware for Mobile Cloud Systems". Journal of Computers, Vol 8, No 9 (2013), 2209-2216, Sep 2013.

Jianmin Chen is currently a Ph.D. candidate at College of Computer Science, Chongqing University, Chongqing, China. He is also a lecturer at Nanchang Hangkong University, China. His research interests include mobility management of wireless networks and data mining.

Zhongyang Xiong received the Ph.D. degree in computer science from Chongqing University, Chongqing, China, in 2004. He is a professor in College of Computer Science, Chongqing University. He had been a Visiting Professor with University of Kentucky, during 2005. His major field of study includes image processing, the next generation network and data mining.

Monitoring of Surface Subsidence of the Mining Area Based on SBAS

Yufeng Zhu[1,2]
[1] Fundamental Science on Radioactive Geology and Exploration Technology Laboratory, East China Institute of Technology, NanChang, China
[2] Central South University, Changsha, China
Email: yfzhu@ecit.cn

Xiaoli Ding[2], Zhiwei Li[2], Yan Luo[1]
[1] Fundamental Science on Radioactive Geology and Exploration Technology Laboratory, East China Institute of Technology, NanChang, China
[2] Central South University, Changsha, China
Email: lsxlding@polyu.edu.hk, zwli@mail.csu.edu.cn, yluo@ecit. cn

***Abstract*—This paper has collected 7 scenes of L band PALSAR sensor radar data of a mine in FengCheng city, jiangxi province, using the Small-baseline Subset (SBAS) method to invert the surface subsidence of the mine. Baselines of interference less than 800m has been chosen to constitute short baseline differential interference atlas, using pixels whose average coherent coefficient was larger than or equal to 0.3 as like high coherent point target, using singular value decomposition (SVD) method to calculate deformation phase sequence based on these high coherent points, and the accumulation of settlements of study area of different period had been obtained, so as to reflect the ground surface settlement evolution of the settlement of the area. The results of the study has showed that: SBAS technology has overcome coherent problem of the traditionality D-InSAR technique, continuous deformation field of surface mining in time dimension of time could been obtained, characteristics of ground surface settlement of mining subsidence in different period has been displayed, so to improve the accuracy and reliability of the monitoring results.**

***Index Terms*—D-InSAR, Small-baseline Subset, Mining subsidence, deformation monitoring**

I. INTRODUCTION

Ground settlement and subsidence caused by mining is a destruction disasters often happening in mining area, and it is one of the most important geological disasters in China. Ground subsidence of mined-out area not only destroys all kinds of architectural, engineering, water conservancy, transportation facilities and farmland, it is also a threat to people's life and property safety, and it has become an important factor which influences and restricts the sustainable development of mining cities[1,2]. Therefore, it is necessary to use the advanced technology to monitor and control damage caused by the ground surface settlement, which can guarantee the sustainable development of mining area.

Differential Interferometric Synthetic Aperture echoes(D-InSAR) is a surface deformation monitoring method recently developed, monitoring the coal mine area of all-weather, non-contact, planar, fast and accurate, with centimeters or mm level. However D-InSAR technology focuses on the single deformation within short time interval, so SAR image which can be used is little, without redundant observations. When the deformation phase is not far from phase noise, deformation phase can't be separated effectively, and reliability of deformation calculation is poor [3]. Because of incoherent of the time and the baselines, orbit error and the influence of the atmosphere, D-InSAR single discrete settlements cannot be connected in most cases, and it is difficult to detect the deformation sequence, and the development and evolution of settlement of the study area coal mine can't be obtained.

According to the D-InSAR technology limitations, in recently years, some scholars abroad calculated the deformation rate effectively by using least squares estimation method in the high coherence point target , through virtual observation method, and in which accuracy and reliability of D-InSAR technology deformation calculation has been improved [4,5]. Ferretti, etc. (2000) first proposed Permanent Scatter technology, choosing a series of phase stable point as PS point according to a certain criterion in a group of radar interference image given, and after removing DEM error, atmospheric delay phase, orbit error and other noise of PS points according to the given phase model, then the exact deformation of surface shape can be obtained [6,7].

Berardino [8] (2002) and Lanari [9] (2004) proposed SBAS method, using small baselines combination to get interference measurement, calculating of the combination of much small baselines using SVD method can effectively to solve discontinuous time problems caused by space between baseline of different SAR data which is too long to improve resolution of monitoring time, so that accuracy of D-InSAR technology deformation calculation has been improved further.

doi:10.4304/jcp.9.5.1177-1184

In this paper, a coal mine area in FengCheng city, Jiangxi province has been selected as the research area, the L-band PALSAR data of the region has been chosen to do time series analysis using SBAS technology, and time series deformation field could be obtained, and mining subsidence funnel could be detected and positioned. Development situation of mining subsidence funnel has been analyzed, to reveal the rule of evolution of subsidence funnel according with time.

II. SMALL-BASELINE SUBSET THEORY

In 2002, Berardion and others proposed that SBAS method was used to study low resolution and large scale of deformation in order to get deformation time sequence diagram [8]. According to the principle of SBAS technology——baseline distance sets in SAR image is small, baseline distance between the set of SAR image is big , the SAR data acquired would be combined into lots of combinations, and then baselines combinations could be solved together by using SVD method or Least squares (LS, Least Square) method. This method can effectively weak the influence incoherent of the space, and at the same time, overcome DEM error and inconsistent of atmospheric phase, so that deformation diagram got could be more continuous in space. This specific principle can be following as three steps:

A. Least Squares (LS)

Suppose that there are N + 1 scenes covering the same area of SAR image, obtaining in time $t_0, t_1, \cdots, t_N$, and every image can interfere with another image at least, this means that each short baseline could be combined by at least two images. Based on the above assumptions, M numbers of differential interferograms could be obtained, in which M meets the following inequality (N is assumed to odd):

$$\frac{N+1}{2} \le M \le N(\frac{N+1}{2}) \tag{1}$$

Suppose that the j scene interferogram could be produced by SAR image obtained from t_A and t_B. Removing topographic phase, and assuming $t_B > t_A$, then in bearing - distance pixel coordinate system (x, r), the interference phase in (x, r) of the interference figure can be expressed as:

$$\begin{aligned}\delta\phi_j(x,r) &= \phi(t_B,x,r) - \phi(t_A,x,r) \\ &\approx \frac{4\pi}{\lambda}\left[d(t_B,x,r) - d(t_A,x,r)\right]\end{aligned} \tag{2}$$

In formula (2), λ is a radar center wavelength, $d(t_B,x,r)$ and $d(t_A,x,r)$ are accumulated form variables of LOS respectively to the reference time t_0, thus $d(t_0,x,r) \equiv 0$; Of course $d(t_i,x,r)$ (I = 1,..., N) is used to respect the deformation time series obtained, and the corresponding phase was set up as $\phi(t_i,x,r)$, and then $\phi(t_i,x,r) \approx 4\pi d(t_i,x,r)/\lambda$. It should be pointed out that, in the formula (2) atmospheric phase, residual topographic phase and incoherent phenomenon are not be considered in phase $\phi(t_i,x,r)$; And assume that all phase signals are unwrapping, and a pixel with a type variable known and be corrected is used as starting point unwrapped.

Vector of N number of the unknown phase value corresponded with the pixel shape variables analyzed could be expressed as:

$$\phi^T = \left[\phi(t_1),\ldots,\phi(t_N)\right] \tag{3}$$

Vector consists from M phase unwrapping from the differential interference graphical could be expressed as:

$$\delta\phi^T = \left[\delta\phi_1,\ldots,\delta\phi_M\right] \tag{4}$$

The corresponding time series of IE image and IS image could be expressed as respectively:

$$IE = \left[IE_1,\ldots,IE_M\right] \quad IS = \left[IS_1,\ldots,IS_M\right] \tag{5}$$

Suppose that IE image and IS image could be arranged according to time, just $IE_j > IS_j$, and j=1,⋯, M, then the phase of j scene of interferogram could be expressed as:

$$\delta\phi_j = \phi(t_{IE_j}) - \phi(t_{IS_j}), \quad \text{j=1,}\cdots\text{, M} \tag{6}$$

For all interferogram, linear model in formula (6) could be expressed as M number equations of N unknown values, the matrix form is:

$$A\phi = \delta\phi \tag{7}$$

Among them A is $M \times N$ matrix, each row corresponds to an interferogram, each column corresponds to a SAR image of one time, the column of IE image is + 1, the column of IS image is - 1, the rest is zero. If M $\geqslant$ N and the rank of A is N, then do it according to the least square method:

$$\phi = A^{\#} \cdot \delta\phi, \quad A^{\#} = (A^T A)^{-1} A^T \tag{8}$$

B. Singular Value Decomposition (SVD)

When M < N, matrix A is rank-defect, $A^T A$ is a singular matrix; suppose that there are L number of different baseline sets, the rank of matrix is N - L + 1, and the equations will have an infinite number of solutions. In order to solve this problem, it needs to use the singular value decomposition of matrix method, just doing SVD of the coefficient matrix A, and the minimum norm of the sense of least squares could be obtained. The concrete implementation steps are as follows:

$$A = USV^T \tag{9}$$

In formula, U is orthogonal matrix $M \times M$ dimensional, composed by the characteristic vector u_i of AA^T; V is orthogonal matrix $N \times M$ dimensional, composing by the characteristic vector v_i of AA^T; S is diagonal matrix $M \times M$ dimensional, and diagonal element pixel is characteristic value σ_i of AA^T. Assume that rank of A is R, the former R characteristic value of AA^T is nonzero value, then M - R characteristic value is 0. A pseudo inverse matrix is defined as A^+, then for valuation, the value of $\hat{\phi}$ acquired by squares constraints can be expressed as follows:

$$\hat{\phi} = A^+\delta\phi, \quad A^+ = VS^+U^T \tag{10}$$

So least squares phase estimate value in minimum norm could be obtained as:

$$\hat{\phi} = \sum_{i=1}^{R} \frac{\delta\phi^T}{\sqrt{\sigma_i}} v_i u_i$$

C. Extraction of Linear Deformation Rate

In order to obtain the physical meaning of the sedimentation sequence, phase is used to respect the product of average phase velocity and time in formula (9), namely:

$$v^T = \left[v_1 = \frac{\phi_1}{t_1 - t_0}, \ldots, v_N = \frac{\phi_N - \phi_{N-1}}{t_N - t_{N-1}} \right] \quad (11)$$

Instead the phase of formula (6),

$$\sum_{k=IS_J+1}^{IE_J} (t_k - t_{k-1}) v_k = \delta\phi_j \ , \quad \text{j=1,}\cdots\text{,M} \quad (12)$$

So a new matrix equation could be obtained, namely:

$$Bv = \delta\phi \quad (13)$$

In formula (13), B is a $M \times N$ matrix. For the first j line, columns located in acquisition time of IE and IS images, $B(j,k) = t_k - t_{k-1}$,and the other $B(j,k) = 0$,the average deformation rate phase value v_k (k=1,⋯,N) of each time interval of N + 1 number images could be worked out by the least squares method or SVD decomposition, and cumulative surface linear deformation phase acquisition time from the first image to the other image could be calculated by multiplication accumulation of v_k and interval times.

After acquisition of linear deformation phase, a new linear model can be established by other linear phase contribution parameters (error and track error caused by DEM), so that deformation valuations inversion could be obtained more accurately. In addition, based on the linear model, the nonlinear deformation phase and atmospheric phase could be separated by filter according to the different characteristics of the residual phase in time and space performance [6,7].

III. Experimental Research

In order to verify the validity and practicability of the technology in surface subsidence monitoring, a coal mine area in FengCheng city, Jiangxi province has been chosen as the research object, using SBAS technology for time series analysis, sedimentation sequence diagram from June 2007 to May 2008 could be obtained.

A. The Experimental Datas

In this paper, 7 scenes the l-band SAR data of PALSAR ALOS satellite sensors covered plateau lake coal mine area and a series of disposal of no "empty" global SRTM V4 DEM data with 3 seconds accuracy (resolution for 90 m) which nominal absolute elevation accuracy is ±16m and nominal absolute plane precision is ±20m provided by spatial information association (CSI) of CGIAR were used. The specific parameters of image information are shown in table 1. In the table image of FBD mode should be 2 times over sampling in the distance in order to keep consistent resolution with FBS mode. In removing topographic phase information with collecting DEM, it should do 6 times over sampling of DEM.

TABLE1

THE SCHEDULE OF SAR DATA

Image Number	Date (YYYYMMDD)	Track Number	Observation Mode	Polarization Mode
1	20070614	07389	FBD	HH+HV
2	20070730	08060	FBD	HH+HV
3	20070914	08731	FBD	HH+HV
4	20071030	09402	FBD	HH+HV
5	20071215	10073	FBS	HH
6	20080130	10744	FBS	HH
7	20080501	12086	FBD	HH+HV

For the l-band of ALOS satellite, less than or equal to 800 m of the baseline can be used as a small baseline, the space of incoherent could be improved, and the coherence increased obviously [3]. Similarly, L band PALSAR data were selected in this experiment, baseline less than 800 m were selected to produce 11 small baselines interfere pairs, and with the use of SBAS method, vertical baseline set distribution could be seen as figure 1.

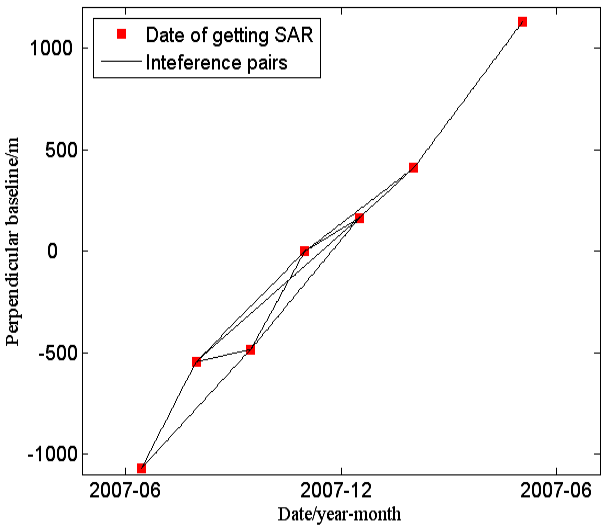


Figure 1. Vertical baseline set pattern

B. Data Processing

Using SBAS method to process data, firstly SAR image data acquired were done with differential interference preproccess of D-InSAR method, and time series D-InSAR analysis model were established to calculate settlement rate and cumulative type variable of each time based on selected high coherent points. Data processing flow of SBAS method was shown in figure 2.

1) Pretreatment

Before the D-InSAR time series analysis, small base set interferogram sequence was generated by SAR data "two rail" method differential interference processing in GAMMA software of Swiss GAMMA Company in this paper.

(1) Selection of IE image

Because SBAS method is used to analyze the time series of identical point, as PS technology, interferograms

should be registered to the same coordinate system in this method. Therefore, a picture of IE image should be chosen, and other images should be registered to the IE image, and then single visual interference figure could be generated based on the same main image. The image of the vertical baseline less than 800 m of interference and the time base, vertical baseline and doppler centroid frequency baseline of minimizing was chosen as IE image. IE image selected in this paper was SAR image obtained in October 30, 2007. Three baseline statistical results could be seen as table 2.

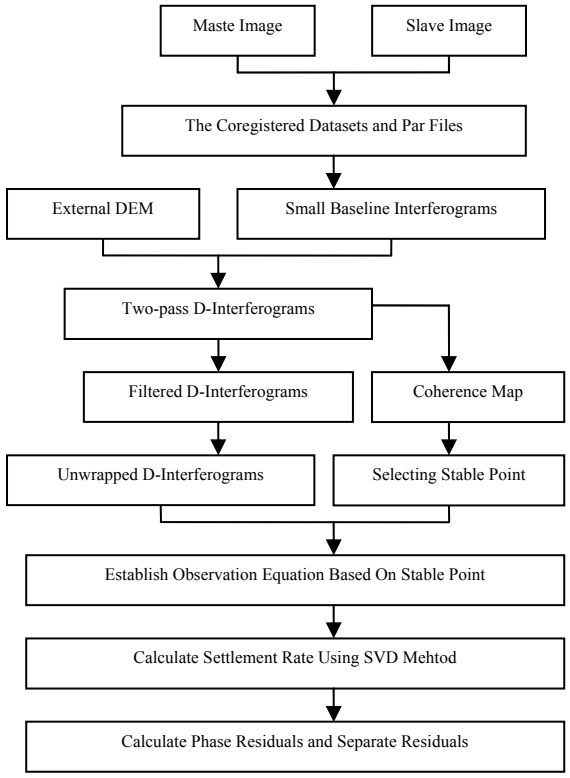


Figure 2. The flowchart of data processing in SBAS

TABLE2

THE STATISTICAL TABLE ABOUT THREE BASELINES

Image Number	Sum of Time	Sum of Vertical Baseline	Sum of Doppler frequency	Sum of Three Baselines
20070614	1012	7085.1524	45.3596	8142.5120
20070730	782	4467.0160	44.5616	5293.5776
20070914	644	4288.3927	74.4414	5006.8341
20071030	598	3803.8635	51.6112	4453.4747
20071225	644	3967.3606	86.5655	4697.9261
20080130	782	4701.1704	66.9226	5550.0930
20080501	1242	8304.1775	44.1424	9590.3199

For the l-band of ALOS satellite, less than or equal to 800 m of the baseline can be used as a small baseline, the space of incoherent could be improved, and the coherence increased obviously [3]. Similarly, L band PALSAR data were selected in this experiment, baseline less than 800 m were selected to produce 11 small baselines interfere pairs, and with the use of SBAS method, vertical baseline set distribution could be seen as figure 1.

(2) Generation of differential interference pattern

Interferogram was generated by SLC differential interference in GAMMA commercial software Swiss GAMMA Company. This process was the foundation of D - InSAR time series analysis. Specific operation processment was as follows:

IE image should be cut, the scope should include the study area and have obvious relief change, and it is an advantage to the registration of study area of SAR intensity map simulated by DEM and IE image intensity map. In this paper the size of the cut was 7000 pixel in the distance, 17500 pixels in the bearing;

The other image was registered to IE image, and single visual interferogram was generated based on the same IE image;

The small base set interference was registered another, generating more apparent small baseline interference atlas of 5 bearing ,2 distance (image size: 3500 pixel * 3500 pixel); Then the small baseline interference atlas was interfered using two rail differential interference and filtering, and then differential interference atlas could be obtained after filter; Further utilizing the MCF method for unwrapping differential interference graph, then unwrapping differential interference atlas under radar coordinates could be obtained;

Due to orbit data of PALSAR radar image was not precise, similar "the ground effect" of the interference fringes (track error phase) was concluded in the filtered differential interference. Unwrapping differential interference figure of orbit error phase removed could be obtained by difference interferogram unwrapping minus orbit error phrase, using the least squares fitting method in this paper. The least squares fitting method of orbit error removed was as follows:

From the filtered difference interferogram, it could be found that orbit error was in obvious linear trend distribution, in radar system, the linear trend was shown with type double linear model [11,12] :

$$\phi_{orbit\_error} = a_0 + a_1 \cdot x + a_2 \cdot y + a_3 \cdot x \cdot y \qquad (14)$$

In formula, ϕ_{oribt_error} is track error phase; X is the distance coordinate of radar coordinate system, y is azimuth coordinates of radar coordinates system(x and y are pixel position of differential interference image); a_0 、 a_1 、 a_2 and a_3 are undetermined coefficient of model, which can obtain through the least squares method. Formula of matrix form was:

$$\phi = M \cdot A \qquad (15)$$

Among them $\phi = [\phi_1 \ \ \phi_2 \ \ \cdots \ \ \phi_n]$; $A = [a_0 \ \ a_1 \ \ a_2 \ \ a_3]$;

$$M = \begin{bmatrix} 1 & x_1 & y_1 & x_1 \cdot y_1 \\ 1 & x_2 & y_2 & x_2 \cdot y_2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & y_n & x_n \cdot y_n \end{bmatrix};$$

The least squares solution was:

$$A = (M'M)^{-1}M'\phi \quad (16)$$

The specific processment of Least squares fitting method to remove orbit error phase was:

Step 1: some points were selected evenly in the differential interferogram of filtering and unwrapping, in order to improve the accuracy of estimation, it should be ensured that the chosen points were located in a relatively stable area and they were of high coherence.

Step 2: The formula model (16) was established, and the undetermined coefficients a_0, a_1, a_2 and a_3 could be calculated through the selected points in step 1 by the least squares.

Step 3: the undetermined coefficient calculated were substituted into formula 14 to get the orbit error phase of the whole image. Then removing the orbit error phase from the differential interference figure unwrapping, differential interference figure unwrapping could be obtained.

Differential interference atlas unwrapping in the geographic coordinates could be obtained by geocoding of the differential interference graph unwrapping which has removed the orbit error.

Due to the small mining area, in order to reflect the study area of the surface settlement clearly and reduce the workload, data processing using only a small image map (550 pixel* 400 pixel) was inversion calculated later.

The unwrapping phase atlas of small range geographic coordinates was shown in figure 3.

From the unwrapping phase figure 3, the subsidence area of mining area surface was clear seen, and there was unwrapping deviation in the dotted area from the interference of 20070914 _20070614, and it was due to the bigger baseline of the interference. And in the same time the field was in summer, the surface vegetation was lush growth, which was leading to the poor coherence of interference. Also there was unwrapping deviation in the dotted area from the interference of 20080130 _20071030, this interference was in winter, plant was rare, which would not be affected by time incoherent, but during 15 December 2007 to 2008 on January 30, the study area was appeared to "frozen" disaster, which affected the radar echo signal this period, causing that coherence of interference was very poor, interference phase was discontinuity, thus there were some errors in unwrapping results.

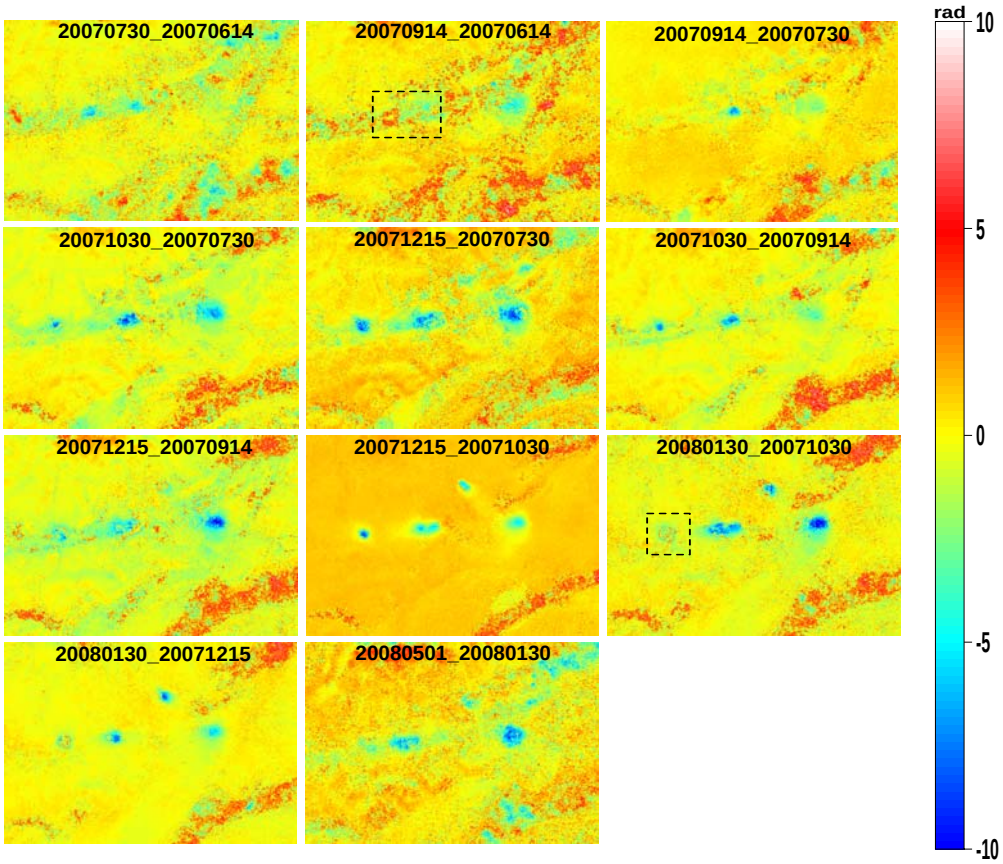


Figure 3. The unwrapped interferograms in MCF

2) Choice of high coherent targets

Before calculating deformation rate and DEM error by use of SBAS time series analysis model, the first was to choose the high coherent target. The chosen of the high coherent points would be able to make data of practical processing later reduced greatly, from a few bit down to hundreds of bits. At present the methods of selection of

high coherent target are amplitude threshold value method and correlation coefficient threshold value method. From the experimental, density of high coherent point by the coherent coefficient threshold value method was bigger than density of high coherent point by the amplitude threshold value method [12]. Amplitude threshold value method is used to estimate phase stability using pixel strength stability instead of coherence, so this often needs more SAR image (generally not less than 30 picture), and the statistical characteristic of the amplitude can be estimated correct. And in time series, stable high coherent point could be found by analyzing the strength of the pixel, which involves complicated mathematics method, and need do a large number of simulation experiments to determine the threshold value. Due to the small amount of data in this paper was small, simple algorithm coherent coefficient threshold value method was chosen to selected high coherent point. Selection criteria: coherent map of time series was taken average, the threshold value was chosen by average coherent coefficient of the cumulative images. Namely:

$$\gamma_{mean} = \frac{1}{M}\sum_{j-1}^{M}\gamma_j \geq \gamma^T \qquad (17)$$

In formula (17), γ_j was coherent coefficient of coherent map of the first j amplitude difference interferogram corresponding; λ^T was coherent coefficient threshold value. In the determination of the correlation coefficient threshold value, it should not only consider the number of the pixel, but also consider the quality of pixel [13]. Considering actual coherence of the study area, if the choice of threshold value is large, it would lose the advantages of large area coverage, so in this paper 0.3 was taken as a threshold value, the average coherent value which was greater than and equal to 0.3 point would be chosen as high coherent target. The study area average coherent diagram was as shown in figure 4.

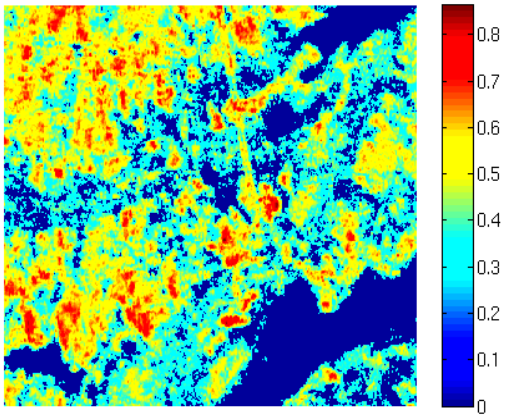


Figure 4. Mean coherence map of study region

3) SBAS processing

In the selection of high coherent target, a linear deformation rate of the linear model and constructs the system of equations is adopted in this paper, the method of singular value decomposition (SVD) to solution equations, it is estimated that interval linear settlement rate each time. Each time settlement rate in time domain integral can get on the accumulation of each time form variables

C. Results Analysing

Differential interference atlas unwrapping of study region after geocoding has been analyzed by SBAS method, and each time of settling diagram of mining area surface could be obtained, as shown in figure 5.

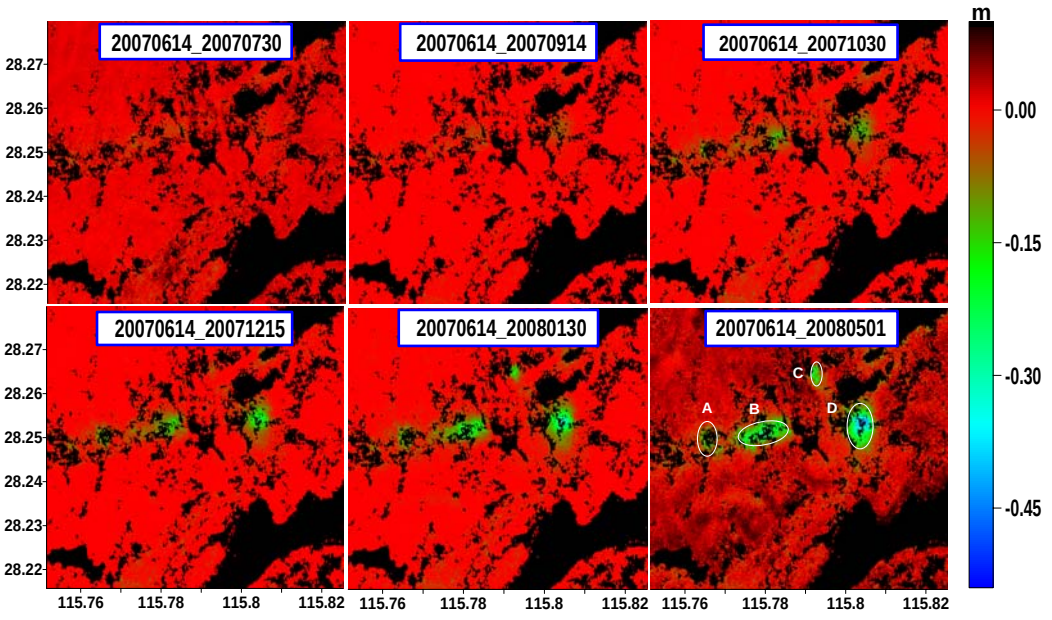


Figure 5. Time series settlement map of study region

From figure 5 it could be seen that the settlement and settlement area of mining area surface has been growing, forming subsidence funnel A, B, C and D during this period of time. The accumulated settlement and subsidence area of subsidence funnel A and C were small in this period of time, because of slow settlement in early

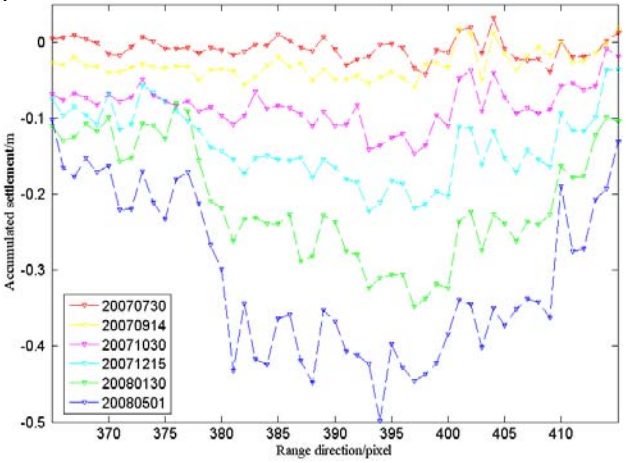
production period; Accumulated settlement of subsidence funnel B and C were bigger in this period of time, because of sedimentation acceleration in mining active phase. In order to analyze the evolution of subsidence funnel and subsidence area, the accumulation of subsidence area in different period should be statistical analyzed, and the results of statistical could be seen in table 2. It could be clearly seen in table 2 that subsidence area of the study area was gradually expanding as times goes on. Due to the influence of ice disaster the SAR image of January 30, 2008, overall coherence of this group of image combination of the differential interferogram was very poor, leading to a certain deviation interfere with the unwrapping results, so there were unsmooth shown in subsidence area of January 2008, 30.

TABLE III.

ACCUMULATED SUBSIDENCE AREA

Date	Subsidence area (Unit: m^2)				
	Settlement >0.1m	Settlement >0.2m	Settlement >0.3m	Settlement >0.4m	Settlement >0.5m
20070730	591.0588	0	0	0	0
20070914	1418.5000	0	0	0	0
20071030	104970	0	0	0	0
20071215	219050	9693.4000	0	0	0
20080130	408890	7565.6000	7683.8000	0	0
20080501	762580	190680	59106	12058	470.8470

In order to analyze the development of subsidence funnel further, subsidence funnel A, C and D has been selected to do profile analysis, in order to observe their time series of the changes of the settlement. The profile of settlement according to time series of center section of funnel D in the column and row direction has been shown in figure 6. From the figure it could been clearly reflected that the settlement of the subsidence funnel gradually increase as time increases , so it could be concluded that it was doing the coal mining in this work face during this period of time.



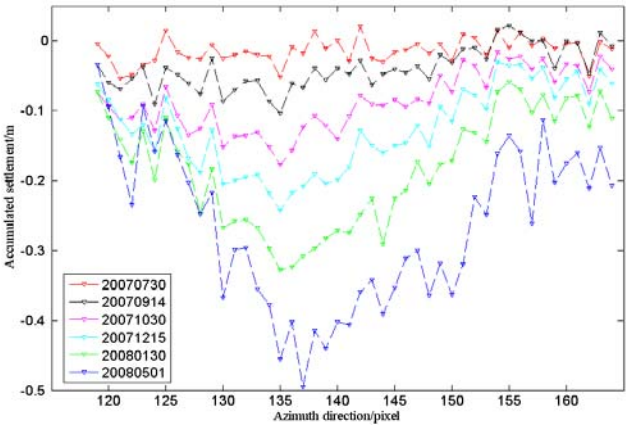


Figure 6. Time series settlement profiles of subsidence funnel D in distance and azimuth direction

Also funnel A and C area has been analyzed respectively, as shown in figure 7 and figure 8. The settlement of funnel C and funnel D area were continuous subsiding as times goes by, as funnel B area. But there was discontinuity of settlement in the space in the figure 7 and 8, due to the poor coherence caused by near to water and crops.

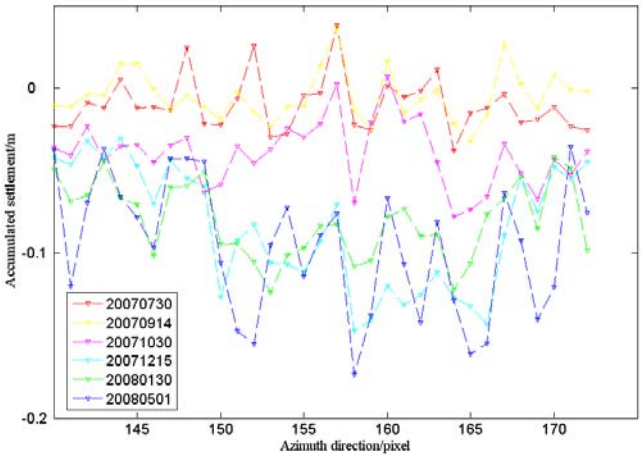


Figure 7. Time series settlement profile of subsidence funnel A in bearing direction

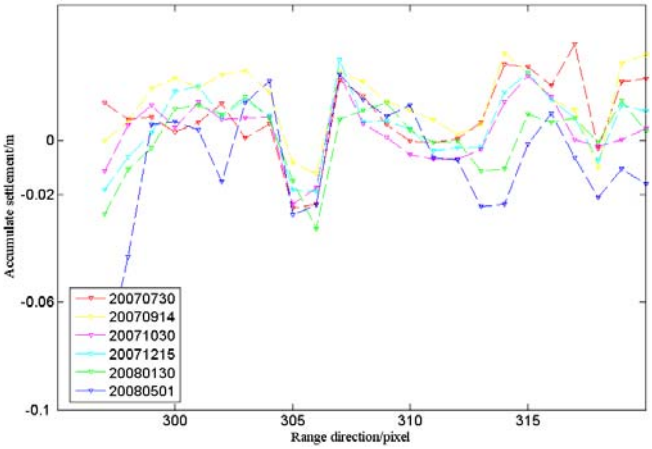


Figure 8. Time series settlement profile of subsidence funnel C in distance direction

IV. CONCLUTION

In this paper SBAS method has been tested with l-band PALSAR sensor radar data, and the experiment results showed that this method can effectively get the subsidence sequence diagram of coal mine area surface. Although SBAS technology has improved D-InSAR technology results well, the quantity of using data was limited, leading to a certain error in results. As the data have two scene data influenced by ice disaster, the overall coherence was poorer, and there were some deviation in unwrapping results, and the accuracy of inversion deformation has been affected. Due to the smaller mining area, and compared to the atmospheric error phase of space low frequency characteristics, settlement has higher spatial frequency, the differential interferometric phase was led by the phase signal of settlement, so in this paper the influence of atmospheric error hasn't been considered[14]. In addition, the phase error caused by inaccuracy DEM has not been considered in the data processing, and it is the further step to do research on how will DEM error can be used as a parameter to estimate and separate effectively.

ACKNOWLEDGMENTS

This paper is supported by Natural Science Foundation of China (40874010), Natural Science Foundation of Jiangxi China (2012ZBAB216001), State Key Laboratory Breeding Base of Nuclear Resources and Environment, East China Institute of Technology (101115).

REFERENCES

[1] Y.F. Zhu, X.L. Ding, Z.W. Li, S.J. Zhou, "Discussion on the Application of GPS Using in Marine Construction Survey," Journal of Computers. Finland, Vol.7, pp. 1663-1670, 2012.

[2] Y.F. Zhu, X.L. Ding, Y. Luo, Z.W. Li, "Analysis on the Application of GPS-RTK Technology using in Yangkou man-made Island in Nantong City," Journal of Networks. Finland, Vol.6, pp. 1444-1451, 2011.

[3] H.J. Yin, J.J. Zhu, Z.W. Li, X.L. Ding, C.C. Wang, "Ground Subsidence Monitoring in Mining Area Using DInSAR SBAS Algorithm," Acta Geodaeticaet Cartographica Sinica. China, Vol.40, pp.52-58, 2011.

[4] Usai S,KLEES R, "SAR Interfermetry on a Very long Time Scale:a study of the Interferometric Characteristics of Man_made Features," IEEE Transactions on Geoscience and Remote Sensing.U.S, Vol.37, pp.2118-2123. 1999.

[5] Usai S, "A Least Squares Database Aproach for SAR Interfermetry Data," IEEE Transactions on Geoscience and Remote Sensing. U.S, Vol.41, pp.753-760, 2003.

[6] Ferretti A,Prati C,and Rocca F, "Nonlinear Subsidence Rate Estimation Using Permanet Scatters in Differential SAR Interferometry," IEEE Transactions on Geoscience and Remote Sensing. U.S, Vol.38, pp.2202-2212, 2000.

[7] Ferretti A,Prati C,and Rocca F, "Permanent Scatters in SAR interferometry," IEEE Transactions on Geoscience and Remote Sensing. U.S, Vol.39, pp.8-20, 2001.

[8] Berandino P,Fornaro G,Lanari R.A new algorithm for surface deformation monitoring based on small baseline differential interferograms[J].IEEE Transactions on Geoscience and Remote Sensing,2002,40(11):2375-2383.

[9] Lanri R,Mora O,Manunta M, "A small_baseline approach for investigating deformation on full_resolution differential SAR interferograms," IEEE Transations on Geoscience and Remote Sensing. U.S, Vol.42, pp.1377-1386, 2004

[10] H.Zhang, C.Wang, T.Wu, Y.X.Tang, Research on DInSAR method Based on the coherent target. Science publications, China 2009.

[11] H.L.Yang, J.H. Peng, D.X. Zhang, S.H. Li, "Infuluence of orbital errosr on InSAR data processing," Journal of Geomatics Science and Technology. China, Vol.29, pp.118-121, 2012.

[12] X.Y. Dong, Research on earthquake deformation monitoring Based on image matching technology, Central South University, China 2011.

[13] X.F.He, M.He, Earth observation data processing method and comprehensive measurement based on InSAR Technology. Science publications, China 2012.

[14] B.W.Huang, X.G.Song, Z.J.Wang, X.J.San, "Research on ground subsidence monitoring of Ge Ting coal mine based on D-InSAR and GIS techniques," Geotechnical Investigation & Surveying. China, Vol.4, pp.55-60, 2012.

Yufeng Zhu was born in Jinxian, Jiangxi, China, in 1981. He received the Master degree in cartography and geographic information engineering from East China Institute of Technology (ECIT), Fuzhou, China, in 2005. He is currently studying toward the Doctor degree in geodesy and surveying engineering, School of Geosciences and Infor-physics, Central South University (CSU), Changsha, China.

As a teacher of East China Institute of Technology, Fuzhou, China, his research topics include intelligence algorithm, research of GIS, mining surveying, and application of InSAR technology.

A Fuzzy Evaluation and AHP based Method for the Energy Efficiency Evaluation of EV Charging Station

Hanwu Luo
School of Electrical Engineering, Wuhan University, Wuhan, China
Email: hanwu_luo@yahoo.com

Jiangjun Ruan and Fang Li
Henan Electric Power Company, Songshan South Road, Zhengzhou, China

***Abstract*—To promote the development of electric vehicles (EVs), many charging stations have been built. The construction of charging stations costs a lot, hence it's necessary to evaluate their energy efficiency due to their main aims of energy conservation. In this paper, a fuzzy model based on the fuzzy comprehensive evaluation (FCE) and analytic hierarchy process (AHP) is presented to evaluate the energy efficiency of charging station. Firstly the evaluation system for the charging station is set up according to AHP and expert surveys, and the evaluation indices are determined by the Argument Delphi method. Then the method to establish the judgment matrix is described, and 7 judgment matrices are established. The process to calculate the weights of all indices in the evaluation system is formulated based on AHP, including the weights for the criterions in the second layer and indices in the third layer. Finally the energy efficiency evaluation on the charging station is conducted according to FCE. The energy efficiency of a charging station in Chongqing was evaluated, and results indicate that the fuzzy model in this paper is effective for the energy evaluation of charging station.**

***Index Terms*—energy efficiency evaluation, charging station, analytic hierarchy process, fuzzy comprehensive evaluation, evaluation indices**

I. INTRODUCTION

As environmental pressure and energy depletion are increasingly severe, more and more attention has been paid to electric vehicles (EVs) because of their high energy efficiency and low off-gas emission compared to conventional internal combustion engine based vehicles [1]. Ultimately, EVs will shift energy demands from crude oil to electricity for the personal transportation sector. To promote the development of EVs, many charging stations have been built, and more and more will be built [2-5].

Meanwhile many researches regarding EVs are being conducted, including the design and optimization of charging stations, investigation on the control of vehicle-to-grid (V2G)[6-13], analysis on the influence caused by the charging machines [14-17], energy storage of EV, charging techniques[18,19], and so on. And many achievements have been obtained. In China the national grid has started the construction of charging station since 2009, which aims at fasting the promotion of EVs.

The aim of the promotion of EV and construction of charging stations is energy conservation and reduction of off-gas emission. Energy efficiency evaluation has been proved to be a way to evaluate the contribution of a device or a system to energy conservation. Therefore, it is necessary to build a proper and comprehensive energy efficiency evaluation system for charging station, which can ensure that EVs and charging stations play an important role in the worldwide energy conservation & emission reduction.

Hence in this paper a fuzzy model based on the fuzzy comprehensive evaluation and analytic hierarchy process is presented to evaluate the energy efficiency of charging station.

Fuzzy comprehensive evaluation method, mainly using of evaluation results of single factor related to the evaluation object, is to form the corresponding evaluation matrix, and to do fuzzy transformation using the weighting factor for determining the important degree of each factor, and the final evaluation results of the evaluation object will be obtained. Fuzzy evaluation set is determined by the use of the factor set, membership degree, weighting factor, and the best evaluation results will be obtained from the alternative set. An AHP hierarchy is a structured means of modeling the decision. It consists of an overall goal, a group of options or alternatives for reaching the goal, and a group of factors or criteria that relate the alternatives to the goal.

Firstly the evaluation system for the charging station is set up according to AHP, which is divided into three layers: goal layer, criterion layer and indices layer. The criterions and indices in the evaluation system are obtained by expert surveys based on Delphi method.

Then the method to establish the judgment matrix is

doi:10.4304/jcp.9.5.1185-1192

described, and 7 judgment matrices are established which are basic for the calculation of indices weights. And the process to calculate the weights of all indices in the evaluation indices system based on AHP is formulated, including the weights for the criterions in the second layer and indices in the third layer.

Finally the energy efficiency evaluation on the charging station is conducted according to the process of FCE, and the assessment rating can be obtained. The energy efficiency of a charging station in Chongqing is evaluated, assessment ratings for the total goal and six criterions are obtained, and according to which corresponding measures can be taken to improve the energy efficiency of EV charging station.

The remaining parts of the paper are arranged as follows: The evaluation system of energy efficiency is set up in Section II, and the process to evaluate the energy efficiency based on fuzzy comprehensive method is described in Section III, and the energy efficiency evaluation on a charging station in Chongqing is conducted in Section IV. Finally conclusions end the paper.

II. Setup of the Evaluation System

A. Delphi Method

Delphi method is a survey technique for achieving consensus among isolated anonymous participants with a controlled feedback of opinions. This technique is being increasingly used in many complex areas in which a consensus is to be reached. Some of these areas included the development of residential areas, theory and design application, and bridge condition rating and effects of improvements. Moreover, the Delphi method is a highly formalized method of communication that is designed the maximum amount of unbiased information from a panel of experts. Therefore, this method is adopted and used for obtaining a set of selection criteria for the selection of the procurement system.

Delphi method means asking a number of experts for advice on some questions, and then collecting the opinions of each adviser and distributing them to experts as reference materials. It is a method in which the experiences, knowledge, and presumptions of expert panelists on an issue or development process under study are collected in an interactive process, normally by interview or survey [20, 21]. As a data collection method, the Delphi can fall in the category of both a quantitative and qualitative study. It is useful when the phenomenon under study is complex or when the topic is somehow delicate – difficult to define, awkward to talk about, politically delicate, etc – or the number of members in the focus group is relatively small.

In this study the Argument Delphi method is used to set up the evaluation system for the energy efficiency of charging station, including the determination of criterions and evaluation indices. The Delphi method process was conducted mainly following the Argument Delphi method, and the whole process took about four months, the process is shown in Table I.

TABLE I.
PROCESS TO SET UP THE EVALUATION SYSTEM BASED ON ARGUMENT DELPHI METHOD

Phase	Purpose and content	Participants
First round: - Selection of the expert panel - Semi-structured interviews	- Identify the key issues in the evaluation and assist in formulating the topics - Find meaningful questions and future statements	15 in-depth interviews
Second round: - Questionnaire to the panelists - 43 future statements	- Evaluate the statements and argumentation for the evaluation indices	37 panelists representing - Specialists (13/13) - Generalists (12/14) - Industry (7/10) 32 responses
Third round: - Questionnaire to the panelists - 29 future statements	- Determine the evaluation indices	30 panelists representing - Technical experts (15/15) - Management experts (12/15) -27 responses

B. Energy Efficiency Evaluation System for Charging Station

According to the results obtained by the expert survey based on Argument Delphi method and the principle of AHP[21-25], the energy efficiency evaluation system for the charging station is set up as shown in Fig.1.

Analytic hierarchy process (AHP) is a systematic analysis method which was proposed by a professor in the University of Pittsburgh named Satai in the70 years of 20th century. It regards the evaluation subjects or problems as a system, and breaks down the problems into different elements according to the nature of question and the expected overall objective, and gathers those elements at different levels in accordance with the correlation and subordination among the elements, to form a multilevel analysis system which makes the problems organized and hierarchical. This research adopts AHP and makes pairwise comparison and forms a matrix to calculate the relative compared weight, and makes the consistency test of the matrix.

The evaluation system for the energy efficiency of charging station is composed of three layers. The top layer is the goal of the system - energy efficiency of the charging station (G0).

The second layer contains six criterions as follows:

B1-Departments & strategies, is the division of the department and strategies for management and energy conservation in the charging station. B1 contains 3

indices C1 to C3.

B2-Management & training of staff, incorporates 2 indices C4 and C5.

B3-Management of equipment, is related to 3 indices-C6 to C8.

B4-Efficiency of power supply system, is related with the efficiency of power supply equipment, such as transformers, switchgear. Four indices C9 to C12 are contained in B4.

B5-Efficiency of charging equipment, is comprised of three indices C13 to C15.

B6-Efficiency of monitoring system, is composed of three indices C16 to C18.

The third layer is composed of 18 indices, including: C1- Energy conservation department, C2- Energy conservation strategy, C3- Energy efficiency management, C4 - Energy conservation operation training, C5 - Energy conservation management training , C6 - Equipment depreciation degree, C7- Equipment technical rating, C8 - Equipment maintenance plan, C9 - Distribution transformers, C10 - Distribution switchgears, C11 - Electricity meters, C12 - Harmonic processing equipment, C13 – Rectifiers, C14 - DC charging machines, C15 - Billing equipment, C16 - Security monitoring system, C17 - Charging monitoring system, C18 - Intelligent charging monitoring system.

III. EVALUATION PROCESS FOR THE ENERGY EFFICIENCY OF CHARGING STATION

When the evaluation system for the energy efficiency of the charging station is established, the weights for the indices C1 to C18 in the third layer and criterions B1 to B6 should be calculated, and the evaluation based on fuzzy comprehensive evaluation can be carried out.

In this part, the establishment of the judgment matrix, method to calculate the weights of indices and evaluation process based on FCE will be described.

A. Establishment of Judgment Matrix for the Evaluation System

To calculate the weight of the indices in the evaluation system, firstly the judgment matrix $A=(a_{ij})_{n\times n}$ for each layer in the evaluation system should be set up. In AHP model, judgment matrix can be constructed by pair-wise comparisons between factors at the same level. The judgment matrix is the analysis basic of AHP and the weight of factors to top goal can be obtained from it. The value of element a_{ij} in the judgment matrix is determined by the relationship between indice *i* and indice *j*, the value set of a_{ij} is {1, 3, 5, 7, 9}, which represents different relationships between two components in the evaluation system, and the details for the value of a_{ij} is listed in Table II. A matrix of judgments A= (a_{ij}) is constructed with respect to a particular property the elements have in common. It is reciprocal, that is $a_{ij}=1/a_{ji}$, and $a_{ii}=1$.

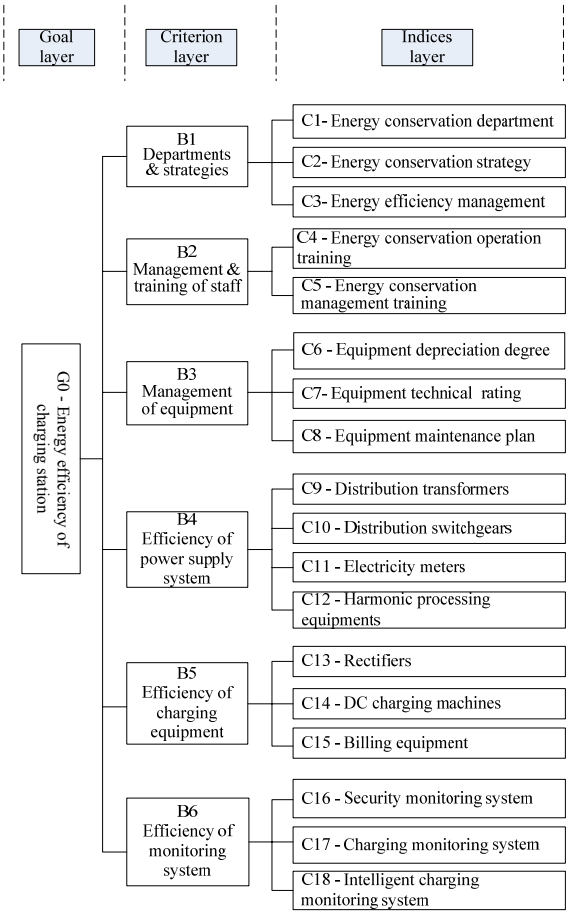


Figure 1. Structure of evaluation system based on AHP and Delphi method

TABLE II.
VALUE OF a_{ij} FOR DIFFERENT RELATION BETWEEN INDICE i AND INDICE j

a_{ij}	Relation between indice *i* and indice *j*
1	indice *i* and indice *j* are equally important for the objective
3	indice *i* is a little important than indice *j* for the objective
5	indice *i* is much important than indice *j* for the objective
7	indice *i* is much more important than indice *j* for the objective
9	indice *i* is extremely important than indice *j* for the objective
2,4,6,8	indice value between two corresponding status.

Again the expert surveys were carried out for the judgment matrix, in which 37 questionnaires were collected. Based on the results obtained in the expert surveys and rules shown in Table II, 7 judgment matrices A_0 to A_6 are established as shown in Table III.

Here A_0 is the judgment matrix for the goal G0 in the

evaluation system, A_1 is the judgment matrix for the criterion B1 in the evaluation system, A_2 is the judgment matrix for the criterion B2 in the evaluation system, A_3 is the judgment matrix for the criterion B3 in the evaluation system, A_4 is the judgment matrix for the criterion B4 in the evaluation system, A_5 is the judgment matrix for the criterion B5 in the evaluation system, A_6 is the judgment matrix for the criterion B6 in the evaluation system.

TABLE III.
JUDGMENT MATRICES

A_0						
G0	B1	B2	B3	B4	B5	B6
B1	1	3	3	1/3	1/5	1
B2	1/3	1	1	1/5	1/7	1/3
B3	1/3	1	1	1/5	1/7	1/3
B4	3	5	5	1	1/3	3
B5	5	7	7	3	1	5
B6	1	3	3	1/3	1/5	1

A_1				A_3			
B1	C1	C2	C3	B3	C6	C7	C8
C1	1	5	3	C6	1	1/7	1/3
C2	1/5	1	1/3	C7	7	1	3
C3	1/3	3	1	C8	3	1/3	1

A_2			A_4				
B2	C4	C5	B4	C9	C10	C11	C12
C4	1	1/5	C9	1	5	7	5
C5	5	1	C10	1/5	1	3	1
			C11	1/7	1/3	1	1/3
			C12	1/5	1	3	1

A_5				A_6			
B5	C13	C14	C15	B6	C16	C17	C18
C13	1	3	9	C16	1	3	9
C14	1/3	1	7	C17	1/3	1	5
C15	1/9	1/7	1	C18	1/9	1/5	1

B. Method to Compute the Weights of Indices

Because the evaluation is divided into three layers, thus the goal layer, criterion layer and indices layer, hence three weights of all the indices should be computed. The first one is the weights of the second layer B1 - B6 to the top layer G0, the second one is the weights of the third layer to the criterions in the second layer, such as weight of C1 - C3 to B1 and weight C4-C5 to B2, and the last one is the weights of indices in the third layer to the goal G0. The following is the process to calculate the weights.

1) Normalize the judgment matrix $A = (a_{ij})_{n\times n}$ according to equation (1),

$$\overline{a_{ij}} = \frac{a_{ij}}{\sum_{k=1}^{n} a_{kj}} \quad (i, j = 1, 2, ..., n) \tag{1}$$

2) Calculate the normalized weight $\overline{W_i}$ according to equation (2), which is the sum of the $\overline{a_{ij}}$ in one row of normalized A.

$$\overline{W_i} = \sum_{j=1}^{n} \overline{a_{ij}} \quad (j = 1, 2, ..., n) \tag{2}$$

3) With the calculated $\overline{W_i}$ the indicator weight W_i can be computed as shown in equation (3), and then the weight vector can be obtained.

$$W_i = \frac{\overline{W_i}}{\sum_{i=1}^{n} \overline{W_i}} \tag{3}$$

4) Compute the maximum characteristic root $\lambda_{\max}$ for the judgment matrix according to equation (4).

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^{n} \frac{(AW)_i}{W_i} \tag{4}$$

where A is the judgment matrix, W is the column vector for the weight, W_i is the ith component of the weight vector.

5) When the pair-wise comparisons are taken to construct judgment matrix, whose order is larger than two, there will be judgment errors as calculation process develops. Therefore, to ensure the accuracy of the method, the consistency check should be carried out, for which the checking factor CI should be computed as follows:

$$CI = \frac{\lambda_{\max} - n}{n-1}, CR = \frac{CI}{RI} \tag{5}$$

where CI is consistency checking factor, RI is average random consistency factor, CR is the consistency ratio, if $CR < 0.1$, the consistency of the weights is acceptable, otherwise corresponding matrix A should be adjusted, and the following progresses should be conducted [21]:

(1) Find the most inconsistent judgment in the matrix;

(2) Determine the range in which the judgment can be changed, then improve the inconsistency judgment;

(3) If the decision maker can change the judgment to a plausible value in that range, change the judgment; Otherwise use the second most inconsistent judgment. If no judgment is changed the decision is postponed until better criteria is obtained.

C. Principle of Fuzzy Comprehensive Evaluation

Fuzzy comprehensive evaluation is a decision making process that under the fuzzy environment, apply the fuzzy set theory, and make a comprehensive quantity evaluation on a system restrained from many uncertain factors.

When a subject is difficult to evaluate, a big system can be decomposed into small systems. These small systems are further decomposed into some elements. Because the importance of the smallest elements is easy to percept or to calculate, it is easy and accurate to evaluate on these elements. And when the smallest elements are weighed, the importance of the overall system is deduced. Thus the whole evaluation on the subject is scientific correspondingly. For example, someone's quality is decomposed into communication ability, learning capability, cultural level, operation ability, and so on. Then communication ability is decomposed into ability of communication with acquaintance, ability of communication with strangers, and so on; learning capability is decomposed into learning ability of new things, learning ability of common things, and so on. And then some experts are in this field invited to weight the subdivided abilities of this person to deduce this person's overall quality. Fuzzy comprehensive evaluation is suitable generally to evaluate and choose subjects with incomplete information. That is to say, when the subject evaluated is not well-informed, this method is usually adopted. Its prerequisite is that the evaluation indexes of subject investigated can be decomposable.

Fuzzy comprehensive evaluation has three advantages [27-30]. Firstly, it does not depend directly on a certain index, neither excessively on the absolute index. But comparative method can prevent from the inaccuracy of evaluation result resulted from unreasonable standard. Secondly, the important intensity of indexes is embodied by the weight, and the weight allows some certain discrepancy, but it will not change the final evaluation result. Technologically, it avoids the influence of accumulative error. Thirdly, the establishment of the membership function and the selection of operators establish connection among non-quantized indexes in index evaluation, which makes the evaluation result reflect well the whole characteristic and trend of the subject.

D. Process of Fuzzy Method for the Energy Efficiency Evaluation of Charging Station

In this study the evaluation on the energy efficiency of charging station is set up according to fuzzy comprehensive evaluation [26-29]. According to the basic principle of FCE, the main process is as follows:

1) Set up the domain for the factors affecting the evaluation objective $U=(u_1,u_2,...u_p)$, and for the evaluation system for the energy efficiency of charging station, U is composed of 18 indices in the third layer[30,31];

2) Establish the domain for the evaluation rating. No matter how many the levels of factors there are, there is only one evaluation rating. This evaluation rating is suitable for all factors, by which the evaluation standard is confirmed. This evaluation set is expressed by $V=(v_1,v_2,...v_n)$ and the corresponding membership function set $J=(J_1,J_2,...J_m)$. In this study the membership function set J is set as J= {5 (Excellent), 4 (Good), 3(Average), 2 (Qualified), 1 (Unqualified)}；

3) Calculate the weights vector $W=(w_1,w_2,...w_p)$;

4) Set up the fuzzy relation matrix $R=(r_{ij})_{m\times n}$, here r_{ij} is the membership between the ith element in U and jth element in V;

5) Calculate the composite operator K according to W and R as follows:

$$K=WR=\begin{bmatrix} w_1 \\ w_2 \\ \cdots \\ w_m \end{bmatrix}^T \cdot \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \cdots & \cdots & r_{ij} & \cdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \tag{6}$$

6) Compute the final evaluation score according to the weights and corresponding rating as follows:

$$G=\frac{\sum_{i=1}^{m} k_i J_i}{\sum_{i=1}^{m} k_i} \tag{7}$$

To make the final evaluation rating more intuitive, the evaluation rating is quantified as listed in Table IV.

Fig.2 is the flowchart of the fuzzy evaluation method for the energy efficiency of EVs charging station. The evaluation indices listed in Fig.1 comprise the influencing factor domain. The membership functions and evaluation indices weights calculation are calculated according to the results obtained by expert surveys.

TABLE IV.
QUANTIFIED EVALUATION RATING

Evaluation Rating	Excellent	Good	Average	Qualified	Unqualified
Score	[4.5 5]	[3.5 4.5]	[2.5 3.5]	[1.5 2.5]	[1 1.5]

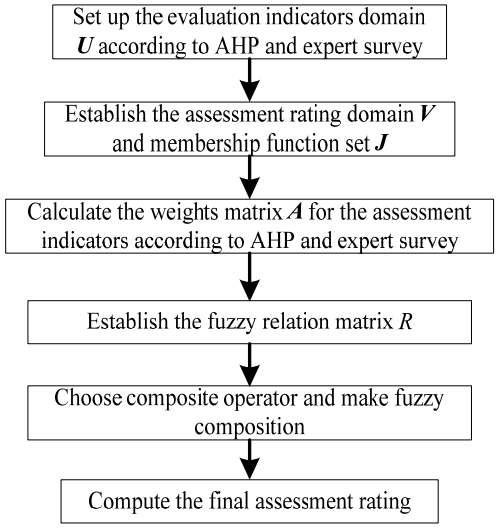


Figure 2. Flowchart of Fuzzy method for energy efficiency evaluation of electric vehicle charging station

IV. Energy Efficiency Evaluation on a Charging Station in Chongqing

To verify the method presented in the paper, an EV charging station in Chongqing was taken as an example. The main parameters of this charging station are as follows:

- Two power supply transformers: power capacity is 1600 kVA, voltage rating is 10/0.4kV;
- Outlets: two 10kV outlets and fifteen 0.4kV outlets;
- Reactive power compensation: four 200kVar capacitance compensation;
- Filters: four 300A active power filters;
- Charging machines: six high power DC charging machines.

A. Calculation of Weights Matrix

According to the structure of the charging station, evaluation system as shown in Fig.1 is established, and according to the judgment matrices shown in Table III, the weights of all components in the evaluation system are calculated.

According to the judgment matrix A_0, the weights of B1 to B6 in the second layer to G0 are 0.1052, 0.0442, 0.0442, 0.2382, 0.4628 and 0.1052 respectively, and the corresponding λ_{max} according to equation (4) is 6.1565, hence the consistency checking factor CI=0.0313, then CR=0.02524, which is far less than 0.1, hence the consistency is acceptable.

In the same way, according to the judgment matrices A1 to A6, the weighs of C1-C18 to B1-B6 can be calculated respectively. Finally the weights of 18 evaluation indices C1-C18 to the evaluation goal G0 can be obtained and the results for the weight calculation are listed in Table V.

TABLE V.
Calculated weights of the indices C1 to C18 for the six criterions B1 to B6 and the Evaluation Goal G0

criterion	weight to G0	indices	weight to criterion	weight to G0
B1	0.105	C1	0.637	0.067
		C2	0.105	0.011
		C3	0.258	0.027
B2	0.044	C4	0.167	0.007
		C5	0.833	0.037
B3	0.044	C6	0.088	0.004
		C7	0.669	0.029
		C8	0.243	0.011
B4	0.238	C9	0.635	0.151
		C10	0.151	0.036
		C11	0.062	0.015
		C12	0.151	0.036
B5	0.463	C13	0.655	0.303
		C14	0.290	0.134
		C15	0.055	0.025
B6	0.105	C16	0.672	0.071
		C17	0.265	0.028
		C18	0.063	0.007

B. Establishment of Fuzzy Relation Matrix

The fuzzy relation matrix is established according to the results obtained in the expert surveys, in which 27 related experts participated, and 23 responses are obtained. According to the data in Table V, the fuzzy relation matrix R1 for B1 can be obtained as follows:

$$R_1 = \begin{bmatrix} 0.146 & 0.375 & 0.287 & 0.157 & 0.035 \\ 0.058 & 0.632 & 0.145 & 0.165 & 0.000 \\ 0.324 & 0.467 & 0.185 & 0.024 & 0.000 \end{bmatrix} \quad (8)$$

According to its weight vector

$$W_1=[0.637 \quad 0.105 \quad 0.258] \quad (9)$$

The composite operator K_1 is

$$K_1 = W_1 \times R_1 = [0.1827 \quad 0.4257 \quad 0.2458 \quad 0.1235 \quad 0.0223],$$

The maximum element in K_1 is 0.4257, which is corresponding to "good" in the membership function set J, hence it's can be concluded that B1 belongs to "good" according to the maximum membership degree principle.

In the same way, according to the results obtained in the expert survey, the composite operators K_2 to K_6 for B2 to B6 can be obtained as follows:

K_2 = [0.1130 0.2690 0.2254 0.3401 0.0525],

hence B2 belongs to "Qualified";

K_3 = [0.5378 0.4041 0.0582 0.000 0.000],

hence B3 belongs to "Excellent";

K_4 = [0.4885 0.3048 0.1681 0.0251 0.0125],

hence B4 belongs to "Excellent;

K_5 = [0.1868 0.2923 0.3583 0.0970 0.0655],

hence B5 belongs to "Average";

K_6 = [0.1366 0.2397 0.3388 0.2121 0.0728],

hence B6 belongs to "Average".

According to the composite operator K_1 to K_6, the fuzzy relation matrix R_0 for G0 is

$$R_0 = \begin{bmatrix} 0.1827 & 0.4257 & 0.2458 & 0.1235 & 0.0223 \\ 0.1130 & 0.2690 & 0.2254 & 0.3401 & 0.0525 \\ 0.5378 & 0.4041 & 0.0582 & 0.0000 & 0.000 \\ 0.4885 & 0.3048 & 0.1681 & 0.0251 & 0.0125 \\ 0.1868 & 0.2923 & 0.3583 & 0.0970 & 0.0655 \\ 0.1366 & 0.2397 & 0.3388 & 0.2121 & 0.0728 \end{bmatrix} \quad (10)$$

Then the corresponding composite operator K_0 can be computed.

$K_0 = W_0 \times R_0$ = [0.2652 0.3073 0.2796 0.1011 0.0460]

And the final evaluation rating score G = 3.645 according to equation (7). Hence it can be concluded that the energy efficiency assessment result is "Good" according to the quantified evaluation rating relationship in Table IV.

G = 3.645 indicates that there should be some measures can be taken to increase the energy efficiency of the charging station, and the measures should be made according to the evaluation results of six criterions and weights of different indices.

According to the evaluation results for the charging station in Chongqing, for the six criterions in the second layer, B3 and B4 got the rating of "Excellent", B1 is "Good", B5 and B6 are "Average", B2 is only "Qualified", and the evaluation result is "Good", therefore the following suggestions are offered to improve the energy efficiency of the charging station.

- Improve the training and management of the staff, and establish the assessment mechanisms to improve the efficiency of the staff.
- Improve the efficiency of charging device, and make a more reasonable charging price.
- Improve the intelligence and efficiency of the monitoring system, which can improve the efficiency of B6.

V. CONCLUSIONS

A Fuzzy method for the energy efficiency evaluation of charging station is presented, which incorporates the setup of evaluation system based on AHP and expert surveys. The evaluation indices in the evaluation system are established based on Argument Delphi method, which is comprised of three layers, 6 criterions and 18 indices. Seven judgment matrices are established according to the results obtained in the expert surveys, then the weights of indices in the evaluation system are calculated. Expert surveys were carried out to obtain the fuzzy relation matrices. Energy efficiency evaluation for a charging station in Chongqing was conducted, the results indicate that the fuzzy method presented in the paper is effective to evaluate the energy efficiency of charging station, and according to the assessment result, measures can be taken to improve the energy efficiency.

REFERENCES

[1] M. Etezadi-Amoli, K. Choma, and J. Stefani, "Rapid-Charge Electric-Vehicle Stations," *IEEE Trans. on Power Delivery*, vol. 25, pp. 1883-1887, 2010.

[2] J. J. Jamian, H. Musa, M. W. Mustafa, H. Mokhlis, and S. S. Adamu, "Combined Voltage Stability Index for Charging Station Effect on Distribution Network," *International Review of Electrical Engineering-Iree*, vol. 6, pp. 3175-3184, 2011.

[3] D. D. Rasolomampionona, F. Maeght, P. Y. Cresson, and P. Favier, "Experimental solar-based charging station for electric vehicles," *Przeglad Elektrotechniczny*, vol. 87, pp. 58-62, 2011.

[4] P. Mohanty, N. Dasgupta, and A. Sharma, "Centralized solar lantern charging station under 'lighting a billion lives' campaign: a technological evolution," *Progress in Photovoltaics*, vol. 18, pp. 516-534, 2010.

[5] A. Chaurey and T. C. Kandpal, "Solar lanterns for domestic lighting in India: Viability of central charging station model," *Energy Policy*, vol. 37, pp. 4910-4918, 2009.

[6] Y. Ota, H. Taniguchi, T. Nakajima, K. M. Liyanage, J. Baba, and A. Yokoyama, "Autonomous Distributed V2G (Vehicle-to-Grid) Satisfying Scheduled Charging," *IEEE Trans. on Smart Grid*, vol. 3, pp. 559-564, 2012.

[7] T. Sousa, H. Morais, Z. Vale, P. Faria, and J. Soares, "Intelligent Energy Resource Management Considering Vehicle-to-Grid: A Simulated Annealing Approach," *IEEE Trans. on Smart Grid*, vol. 3, pp. 535-542, 2012.

[8] E. Sortomme and M. A. El-Sharkawi, "Optimal Scheduling of Vehicle-to-Grid Energy and Ancillary Services," *IEEE Trans. on Smart Grid*, vol. 3, pp. 351-359, 2012.

[9] M. Singh, P. Kumar, and I. Kar, "Implementation of Vehicle to Grid Infrastructure Using Fuzzy Logic Controller," *IEEE Trans. on Smart Grid*, vol. 3, pp. 565-577, 2012.

[10] W. Chenye, H. Mohsenian-Rad, and H. Jianwei, "Vehicle-to-Aggregator Interaction Game," *IEEE Trans. on Smart Grid*, vol. 3, pp. 434-442, 2012.

[11] A. T. Al-Awami and E. Sortomme, "Coordinating Vehicle-to-Grid Services With Energy Trading," Smart Grid, IEEE Transactions on, vol. 3, pp. 453-462, 2012.

[12] [G. Huaqun, W. Yongdong, B. Feng, C. Hongmei, and M. Maode, "UBAPV2G: A Unique Batch Authentication

Protocol for Vehicle-to-Grid Communications," *IEEE Trans. on Smart Gri*d, vol. 2, pp. 707-714, 2011.
[13] D. Dallinger, D. Krampe, and M. Wietschel, "Vehicle-to-Grid Regulation Reserves Based on a Dynamic Simulation of Mobility Behavior," *IEEE Trans. on Smart Gri*d, vol. 2, pp. 302-313, 2011.
[14] S. Wencong and C. Mo-Yuen, "Performance Evaluation of an EDA-Based Large-Scale Plug-In Hybrid Electric Vehicle Charging Algorithm," *IEEE Trans. on Smart Gri*d, vol. 3, pp. 308-315, 2012.
[15] S. Shahidinejad, S. Filizadeh, and E. Bibeau, "Profile of Charging Load on the Grid Due to Plug-in Vehicles," *IEEE Trans. on Smart Gri*d, vol. 3, pp. 135-141, 2012.
[16] R. J. Bessa, M. A. Matos, F. J. Soares, and J. A. P. Lopes, "Optimized Bidding of a EV Aggregation Agent in the Electricity Market," *IEEE Trans. on Smart Gri*d, vol. 3, pp. 443-452, 2012.
[17] A. Ashtari, E. Bibeau, S. Shahidinejad, and T. Molinski, "PEV Charging Profile Prediction and Analysis Based on Vehicle Usage Data," *IEEE Trans. on Smart Gri*d, vol. 3, pp. 341-350, 2012.
[18] Z. Amjadi and S. S. Williamson, "Prototype Design and Controller Implementation for a Battery-Ultracapacitor Hybrid Electric Vehicle Energy Storage System," *IEEE Trans. on Smart Gri*d, vol. 3, pp. 332-340, 2012.
[19] L. Fangcheng, L. Jinjun, Z. Bin, and Z. Haodong, "Fast charging system of electric vehicle (EV) based on hybrid energy storage system," *IEEE Conference and Exposition in Applied Power Electronics (APEC)*, 2, 2012, pp. 2115-2120.
[20] K. Dong-Min and O. K. Jin, "Design of Emergency Demand Response Program Using Analytic Hierarchy Process," *IEEE Trans. on Smart Grid*, vol. 3, pp. 635-644, 2012.
[21] W. Pedrycz and S. Mingli, "Analytic Hierarchy Process (AHP) in Group Decision Making and its Optimization With an Allocation of Information Granularity," *IEEE Trans. on Fuzzy Systems*, vol. 19, pp. 527-539, 2011.
[22] T. Hsin-Yi and H. Yu-Lun, "An Analytic Hierarchy Process-Based Risk Assessment Method for Wireless Networks," *IEEE Trans.s on Reliability*, vol. 60, pp. 801-816, 2011.
[23] H. Tanaka, S. Tsukao, D. Yamashita, T. Niimura, and R. Yokoyama, "Multiple Criteria Assessment of Substation Conditions by Pair-Wise Comparison of Analytic Hierarchy Process,"*IEEE Trans. on Power Delivery*, vol. 25, pp. 3017-3023, 2010.
[24] J. Jeonghwan, L. Rothrock, P. L. McDermott, and M. Barnes, "Using the Analytic Hierarchy Process to Examine Judgment Consistency in a Complex Multiattribute Task," *IEEE Trans. on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 40, pp. 1105-1115, 2010.
[25] L. Jie, M. Jun, Z. Guangquan, Z. Yijun, Z. Xianyi, and L. Koehl, "Theme-Based Comprehensive Evaluation in New Product Development Using Fuzzy Hierarchical Criteria Group Decision-Making Method," *IEEE Trans. on Industrial Electronics*, vol. 58, pp. 2236-2246, 2011.
[26] W. Yaonan, L. Chunsheng, and Z. Yi, "A Selection Model for Optimal Fuzzy Clustering Algorithm and Number of Clusters Based on Competitive Comprehensive Fuzzy Evaluation," IEEE Trans. on Fuzzy Systems, vol. 17, pp. 568-577, 2009.
[27] Jian Shu, Ming Hong, Linlan Liu, Yebin Chen, "A Water Quality Monitoring Method Based on Fuzzy Comprehensive Evaluation in Wireless Sensor Networks", *Journal of Network,* vol.7(1), pp.195-202, 2012.
[28] Hua Jiang, Junhu Ruan, "Fuzzy Evaluation on Network Security Based on the New Algorithm of Membership Degree Transformation—M(1,2,3)", *Journal of Network*, vol.4(5), pp. 324-331, 2009.
[29] Zhibin Liu, Shaomei Yang, "A Hybrid Intelligent Optimization Algorithm to Assess the NSS Based on FNN Trained by HPS", *Journal of Network*, vol5(9), pp. 1076-1083, 2010.
[30] Jianchang Lu, Leping Pei, "Security Evaluation of Power Network Information System Based on Analytic Network Process",*Journal of Network*, vol.8(4) , pp.866-873, 2013.
[31] Lu Guofu. Research on the Energy Efficiency Evaluation Index System for the Power Company [D], 2013 (in Chinese)

Hanwu Luo. He is within the school of electrical engineering, Wuhan University, his main research field focuses on the renewable resource planning. He is a Ph.D students now.

Jiangjun Ruan. Ph.D degree, professor. He is within the school of electrical engineering, Wuhan University, his main research fields focuses on the renewable resource planning, electromagnetic calculation.

Implementation of Multi-channel FIFO in One BlockRAM with Parallel Access to One Port

Zhipeng Gong[1,2], Tefang Chen[1]
1 School of Information Science and Engineering, Central South University, Changsha, China
2 College of Electrical and Information Engineering, Hunan Institute of engineering, Xiangtan, China
Email:{zpgong2000@126.com, ctfcyt@163.com}

Fumin Zou
School of Information Science and Engineering, Fujian University of Technology, Fuzhou, China
Email:fuminzou@ngi.fj.cn

Li Li, Yingxi Kang
College of Electrical and Information Engineering, Hunan Institute of engineering, Xiangtan, China
Email:{lili@hnie.edu.cn, yxkang@hnie.edu.cn}

***Abstract*—Because of flexibility of application and high cost performance, the low-and-middle-end FPGA has obtained an extensive market. As a fundamental memory structure, the FIFO memory is widely used in FPGA based project in various manners. But limited by the resources in chip and imperfection of development tools, the problem that the number of memory is insufficient while the overall capacity is enough often occurs in the implementation of multi-channel FIFO. This paper surveys various occasions of applications of multi-channel FIFO and put forward a method to achieve multi-channel FIFO in a single FPGA BlockRAM, which would support the parallel access to one port. The method may help to solve the problem mentioned above and improve the utilization of storage resources obviously. The steps of implementation and partial source code are present together with the detail analysis of simulation timing. Practical application indicates that the method is successful and effective.**

***Index Terms*—FPGA, FIFO, BlockRAM, Multi-channel**

I. INTRODUCTION

With the rapid development of FPGA technology, PFGA is widely used in the field of communication, medical instrument, consumer electronics, display, portable terminal, and so on[1][2][3]. Among various levels of FPGA, the low-and-middle-end FPGA has won a wide market due to its low cost, high performance, technical maturity and short design cycle [4]. As a fundamental storage structure in the design of system based on FPGA, FIFO is usually used as data buffer of digital signal processing system, communication bridge between network of different data rates and communication interface between modules in different clock domain [5][6]. Sometimes, in practice, it is needed to combine more than one FIFO into a new structure as multi-channel FIFO, according to rule of access to it, multi-channel FIFO generally can be classified into four categories as follows: 1) Serial Input and Serial Output FIFO (SISOFIFO), each channel of SISOFIFO is independent, no parallel operation at read port or write port is permitted, i.e., the data of the SISOFIFO is written in channel by channel and read out channel by channel. there is no constraint for access among channels; 2) Serial Input and Parallel Output FIFO (SIPOFIFO), just as its name implies, data of the SIPOFIFO is written in channel by channel but read out at the same time; 3) Parallel Input and Serial Output FIFO (PISOFIFO), the data of the PISOFIFO is written in at the same time but read out channel by channel, namely, the write port should be able to be accessed in parallel; 4) Parallel Input and Parallel Output FIFO (PIPOFIFO), the data of the PIPOFIFO is written in simultaneously and read out simultaneously, in other words, both ports of PIPOFIFO should be able to be accessed in parallel.

To implement FIFO in the FPGA based project, FPGA development tools often provide a relevant IP core generator [7]. With this generator, users can easily customize the FIFO step by step. But problem would occur when a FIFO is created by the IP core generator. Because each FIFO created by this means would consume a BlockRAM[8] in FPGA, even if the FIFO is a very small one. That is, once a BlockRAM is used, even a small part of it is used, the whole BlockRAM can't be used by others any more, and thus the remaining memory space is wasted. Generally, BlockRAM is a rare resource in low-and-middle-end FPGA, its number is limited, and it is often organized in uniform size. The capacity of a BlockRAM is always more than that of a FIFO need in many cases, e.g., in the case of multi-channel FIFO, the project need a lot of FIFO in number, but each in small size. Under this circumstance, we have to face the problem that there is enough memory in amount capacity but serious lack in number.

Currently, there are few literatures available for the solution to this problem. M.A.Khan proposed a method to implement a 5-channel SISOFIFO in literatures [9]. In

doi:10.4304/jcp.9.5.1193-1200

order to realize easily, the SISOFIFO is created based on distributed RAM, i.e., based on registers, not BlockRAM. Obviously, the distributed RAM in FPGA is used to realize the all kinds of logic on original purpose, and the amount of it is limited too. So this method is not suitable for the situation when there is a large amount of data to store. At the same time, this method can't make use of BlockRAM, so it would waste the memory resource.

With experience in previous FPGA projects, we have realized the design of PIPOFIFO of multi-channel FIFO and successfully applied it to data buffer of bus transceiver of Multifunction Vehicle Bus (MVB) and data cache of large LED display screen. The structure of PIPOFIFO is depicted in Figure 1. To realize the PIPOFIFO, a simple DPRAM is instantiated from a BlockRAM at first, its memory space is divided into multiple parts according to the number of the channels of PIPOFIFO (4 channels in Figure 1). The input data of all channels, from din_0 to din_3, are written in simultaneously under the uniform signal wr_enx, and the output data of all channels, from dout_0 to dout_3, are read out simultaneously under the uniform signal rd_enx. The label logics of all the channels are the same, so one set of them is enough. Actually the PIOPFIFO can be simply viewed as a binding of multiple normal FIFOs and act as one normal FIFO.

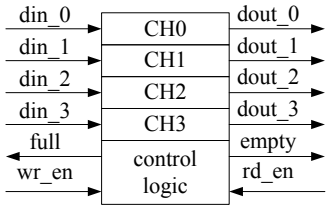


Figure 1. Structure of PIPOFIFO

In this paper, we propose a new solution for the remaining three kinds of multi-channel FIFO to implement them in one BlockRAM. As to SISOFIFO, the channels are independent and there are no timing constraints among them, so it is relatively easy to achieve. In the coming sections, we would concentrate our attention on the implementation of SIPOFIFO and PISIFIFO in one BlockRAM, it can significantly improve utilization of BlockRAM, reduce the cost of product and help to enhance market competitiveness.

II. Design And Structure

A. Basic Structure

From the analysis above, we have to implement multi-channel FIFO in one BlockRAM and provide write or read operation in parallel to some extent. It's obviously impossible to realize write or read operation at the same time for multiple FIFOs that built in one BlockRAM directly. In order to solve the problem of parallel access, it is necessary to place registers into the ports with function of parallel access as data buffer. The width of data buffer should be set in accord with the width of corresponding channel, and the depth is decided by the level of parallel access operation. The general structure of multi-channel FIFO is depicted in Figure 2. It is mainly composed of write control logic, DPRAM, read control logic and input/output buffering registers. The buffering registers are located at the parallel port, i.e., write port for PISOFIFO, and read port for SIPOFIFO. DPRAM can be instantiated from BlockRAM with IP core generator. Simple DPRAM is enough here because there is only one write port and so half of memory capacity can be saved by this means. Each channel have its own private memory space in the DPRAM, all the private memory spaces can't be overlapped. The read control logic and write control logic are relatively complex and there are differences in their structure between the scenes of serial access and parallel access, we'll describe this in detail later.

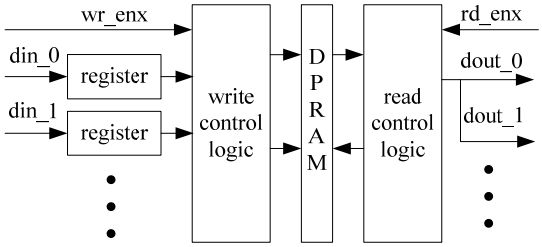


Figure 2. General structure of multi-channel FIFO

B. Parallel Write Control Logic

Parallel write control logic is designed to receive input data of all channels and write them into the DPRAM at corresponding area. The structure of write control logic in parallel is depicted in Figure 3. After receive a parallel write command, i.e. when wr_enx is active, the input data of all channels are registered. On detecting of data coming, the internal write control logic would activate the internal write signals for every channel in turn to fetch the data from registers and write to the DPRAM at corresponding memory area. Here is how it works:

Step 1. The control logic enters the idle state when system resets, all the labels are initialized at this time, e.g. p_write_ready, the label of write in parallel getting ready, is set to 1.

Step 2. When wr_enx is active, internal control logic accepts the input data of all channels and then stores them in corresponding buffering registers. p_write_ready is set to 0. A new external write command would not be accepted before p_write_ready is set to 1 again.

Step 3. Set write_in_process to 1, which means the control logic is enter the procedure of carrying data from buffering registers to DPRAM.

Step 4. Internal control logic generates the internal write command vector wr_eni, this causes a internal write operation for each channel, take channel 0 for example, the internal write operation goes as follows:

Step 5. Number of write channel is decided by ch_write, it is set to 0 firstly, i.e., the first channel is selected;

Step6. The write address of DPRAM wr_addr is a combination of ch_write and write_p_0, i.e., wr_addr={ch_write,write_p_0}, where write_p_0 means the write address pointer of channel 0.

Step 7. The LSB of the internal write commands vector wr_eni is set to 1, and the input data of channel 0 is written to corresponding memory area in DPRAM ;

Step 8. Write address pointer write_p_0 is increased by 1;

Step 9. Write channel number ch_write is increased by 1, so the next channel is selected;
Step 10. The remaining channels write data to DPRAM in the same way as channel 0 in accordance with the operation sequence from **step 5** to **step9** above;
Step 11. After the whole parallel write operation is completed, and if the FIFO is not full, p_write_ready is set to 1 to show the write control logic is ready for the next write operation;

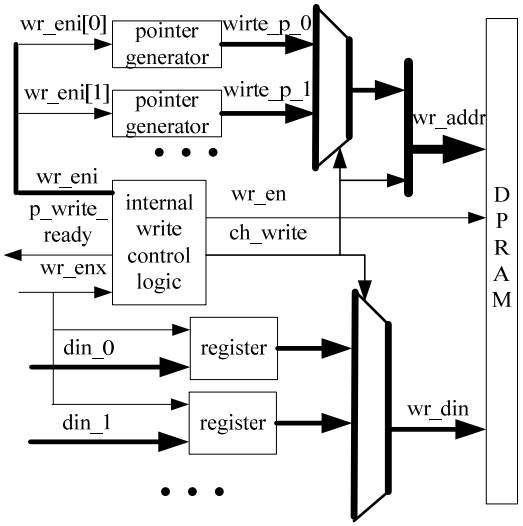


Figure 3. Structure of parallel write control logic

C. Serial Write Control Logic

The structure of serial write logic is depicted in Figure 4, the way it works is exactly the same as multiple independent FIFOs to write separately, it's not necessary to place buffering registers here that used in parallel case, and the labels for parallel write, e.g., p_write_ready, are not needed either. Each channel sets the full label ch_full according to the general rules of common FIFO as an independent FIFO. The write operation of each channel is carried out directly under the extern serial write command wr_enx. The number of write channel is produced by coding wr_enx. It is unnecessary to generate extra internal write command wr_eni, i.e., let wr_eni= wr_enx. Here is how it works:
Step 1. When system resets, the full label of each channel is set to 1, which means no channel is full. Each channel carry out the write operation independently, take channel 0 for example, it works as follows:
Step 2. When external write command wr_enx=0001, the write channel ch_write is coded as 0, i.e., the first channel is selected.
Step 3. The write address of DPRAM wr_addr is combination of ch_write and write_p_0, i.e., wr_addr = {ch_write,write_p_0}, where write_p_0 means the write address pointer of channel 0.
Step 4. The input data of channel 0 is written to corresponding memory area in DPRAM.
Step 5. Write address pointer write_p_0 is increased by 1.
Step 6. When another channel is selected, it works in the same way as channel 0 in accordance with the operation sequence from **step 2** to **step 5** above.

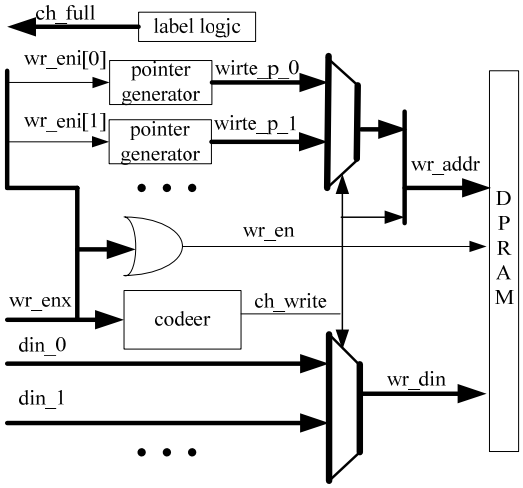


Figure 4. Structure of serial write logic

D. Parallel Read Control Logic

Parallel read control logic is designed to fetch data from proper area in DPRAM and sent them to the output port of corresponding channel. The structure of parallel read control logic is shown in Figure 5. When output buffering registers are free and every channel is not empty, the internal read control logic would give the read command to every channel in turn, thus, the data are read from the DPRAM and registered, and then the label of parallel read data getting ready, p_read_ready, is set to 1. When a active read request, rd_enx, is detected, parallel read control logic will output all the buffered data simultaneously. The following steps would show how it works:
Step 1. When system resets, the label of parallel read getting ready is cleared, i.e., p_read_ready is set to 0;
Step 2. When none of channel is empty, internal read control logic would begin to deliver the data of every channels from DPRAM to corresponding registers as follows;
Step 3. The read channel number ch_read is set to 0, i.e., the first channel is selected.
Step 4. The read address of channel 0 in DPRAM, rd_addr, is a combination of ch_read and read_p_0, i.e., rd_addr = {ch_read,read_p_0}, where read_p_0 means the read address pointer of channel 0;
Step 5. The data of channel 0 would be on the read bus and saved to the output buffering registers when the latch command of channel 0 is active.
Step 6. The LSB of internal read command vector rd_eni is set to 1, and read the address pointer of channel 0, read_p_0, is increased by 1;
Step 7. ch_read is increased by 1, i.e., the next channel is selected;
Setp 8. when another channel is selected, it works in the same way as channel 0 in accordance with the operation sequence from **step 4** to **step 7** above.
Step 9. The label of parallel read getting ready, p_read_ready, is set to 1;
Step 10. When external read request, rd_enx, is detected active, the data of all channels in buffering register will be output simultaneously;
Step 11. p_read_ready is set to 0;

Step 12. Turn to step2.

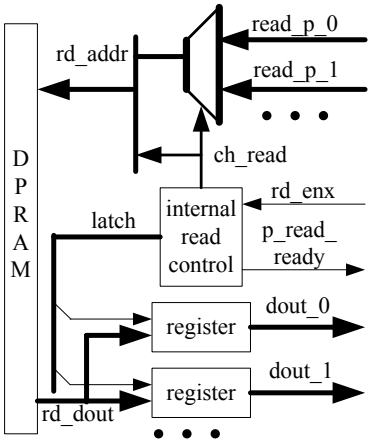


Figure 5. Structure of parallel read control logic

E. Serial Read Control Logic

Serial read control logic is designed to read data from multi-channel FIFO channel by channel. No output buffering registers are needed at this moment, and neither do the labels related to parallel read control logic. It is unnecessary to generate extra internal read command rd_eni any longer, external read command can be used as internal read command directly, i.e., let rd_eni= rd_enx. Each channel set the empty label ch_empty according to the general rules of common FIFO as an independent FIFO. The read operation of each channel is carried out at the external read requests directly. The code of read channel, ch_read, is decided by external read command rd_enx. How it works is similar to that of serial write logic.

F. Label Logic

Label logic is designed to generate all kinds of labels for multi-channel FIFO, including general labels as empty, full, and special labels for parallel access. The structure of label logic is shown in Figure 6. General labels can be produced by common rules of FIFO. Here are rules for special labels:

1) For PISOFIFO, p_write_ready is set to 1 when the input registers are empty and no channel of the PISOFIFO is full. Once the buffering data begin to write to DPRAM, p_write_ready is set to 0.

2) For SIPOFIFO, p_read_ready is set to 1 when data in registers of all channels are ready. Once an external read commands is detected active, p_read_ready is set to 0.

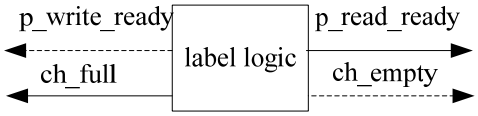


Figure 6. Structure of label logic

III. IMPLEMENTATION

A. Objedtive Platform

We constructed a XILINX FPGA development platform based on XC3S400A, which comes from the SPARTAN3A family, to implement the multi-channel FIFO. ISE13.2 XST [10] is employed as the synthesis tool. A 4-channel SIPOFIFO and a 4-channel PISOFIFO are realized on this platform. A 512 x 25 DPRAM is instantiated from a BlockRAM of 18Kb in size. The assignment of memory space to the 4 channels of FIFO is depicted in Figure 7.

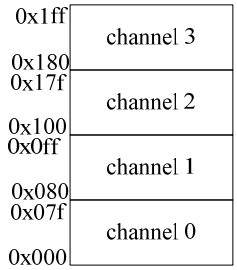


Figure 7. Assignment of memory space

We employ the Verilog Hardware Description Language (Verilog HDL) to implement SIPOFIFO and PISOFIFO, and part of relevant source code is presented in appendix A

B. Synthesis Result

The synthesis result of PISOFIFO and SIPOFIFO is listed in TABLE 1, including the resources they used and the maximum work frequency they can reach. From table 1, we can learn that each multi-channel FIFO needs only one BlockRAM and a few additional logic resources, the maximum frequency is up to 200Mhz. these features would enable the multi-channel FIFO to be integrated in most of projects based on FPGA and work well.

TABLE 1.
SYNTHESIS RESULT

FIFO type	Slices	LUTs	RAMB16BWEs	Frequency
PISOFIFO	226	271	1	193
SIPOFIFO	234	282	1	208

IV. SIMULATION AND EXPERIMENT

We employ modelsim6.5d [11] as the simulation tool and select XC3S400A as the objective chip, which comes from XILINX SPARTAN3A family of low-cost. The test bench runs on a personal computer with a core I3 2.3GHz processor and 2 GB of random access memory.

A. Simulation Of SIPOFIFO

As described earlier, a DPRAM of 512 words is divided into four parts, i.e., there are 128 words per part. Each channel takes up one part as shown in Figure 7. In test bench, we write 64 data, from 0x000000 to 0x00003f, to each channel in turn. The timing diagram of serial write is shown in Figure 8. For further explanation, we take the write operation of channel 0 for example. The data input to din_0 ranges from 0x000000 to 0x00003f, increased by 1 at a time. The write address of DPRAM, wr_addr, is composed of write channel number and write address pointer, i.e., wr_addr = {ch_ write, write_p_0}. As for channel 0, ch_write = 0. write_p_0 ranges from 0x00 to 0x3f. So the write address of DPRAM, wr_addr, ranges from 0x000 to 0x03f. The serial write operation of other channels is similar to that of channel 0.

After the write operation of all channels, the input data should be stored in DPRAM at proper area. The storage of the data of channel 0 and channel 1 in DPRAM is shown in Figure 9. According to the design, the memory space of channel 0 ranges form 0x000 to 0x07f and the memory space of channel 1 ranges form 0x080 to 0x0ff. The input data is the same for all channels, ranging from 0x000000 to 0x00003f. From Figure 9, we can discover that the storage of the data is consistent with the design and test bench.

The timing diagram of parallel read is shown in Figure 10. When the output data of every channel has been registered already, p_read_ready is activated to declare that the multi-channel FIFO is ready for parallel read operation. On receiving the external read request, i.e., when rd_enx is detected active, the parallel read control logic will output the data of all channels at the same time. Then the data is kept in the registers of test bench, from dout_reg_0 to dout_reg_3. Obviously, form Figure 10, we can learn that the data of all channels are the same as expected. It means that the parallel read operation is carried out successfully. p_read_ready is deactivated after a read operation, the parallel read control logic begins to read data form DPRAM to registers for the next read operation. For further description, we take channel 0 for example to illustrate the procedure of read operation from DPRAM to registers. When the output buffering registers are empty, it's time to prepare data for output again. Channel 0 is selected firstly to read data from its memory area in DPRAM to its registers. the read channel number, ch_read is set to 0, The read address of DPRAM, rd_addr, is composed of read channel number and read address pointer, i.e., rd_addr = {ch_ read, read_p_0}. When write_p_0=0x01, rd_addr must be 0x001, the data in DPRAM at 0x001would appear on the read bus and then would be latched in corresponding regisers as shown in Figure 10. The read operation of the remaining channels is similar to that of channel 0 and their corresponding addresses in DPRAM are 0x081, 0x101and 0x181 respectively.

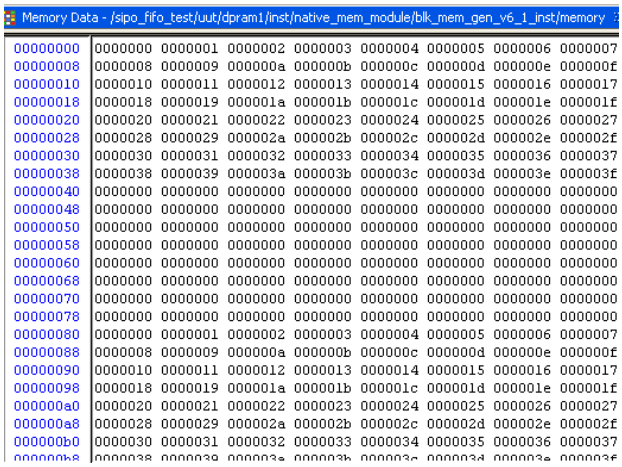


Figure 9. Storage of data in DPRAM

B. Simulation Of PISOFIFO

The timing diagram of parallel write is shown in Figure 11. When the input buffering registers are all empty, p_write_ready is activated to declare that the multi-channel FIFO is ready for parallel write operation. On receiving the external write request, i.e., when wr_enx is detected active, the parallel write control logic will accept the input data of all channels and keep them in the buffering registers simultaneously. During the time that followed, p_write_ready is deactivated and the parallel write control logic begins to write data to DPRAM form registers. For further description, we take channel 0 for example to illustrate the procedure of write operation to DPRAM from registers. When channel 0 is selected firstly to write data from its registers to its memory area in DPRAM, the write channel number, ch_write is set to 0. The write address of DPRAM, wr_addr, is composed of write channel number and write address pointer, i.e., wr_addr = {ch_ write, write _p_0}. When write _p_0=0x01, wr_addr must be 0x001, the data would be written into DPRAM at 0x001 when write enable signal, wr_eni[0], is set to 1, as shown in Figure 11. The write operation of the remaining channels is similar to that of channel 0 and their corresponding addresses in DPRAM are 0x081, 0x101and 0x181 respectively.

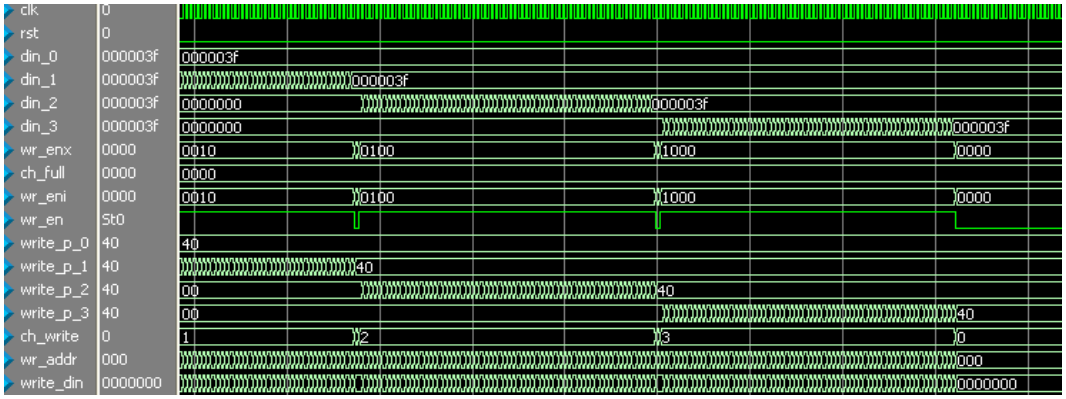


Figure 8. Timing diagram of serial write

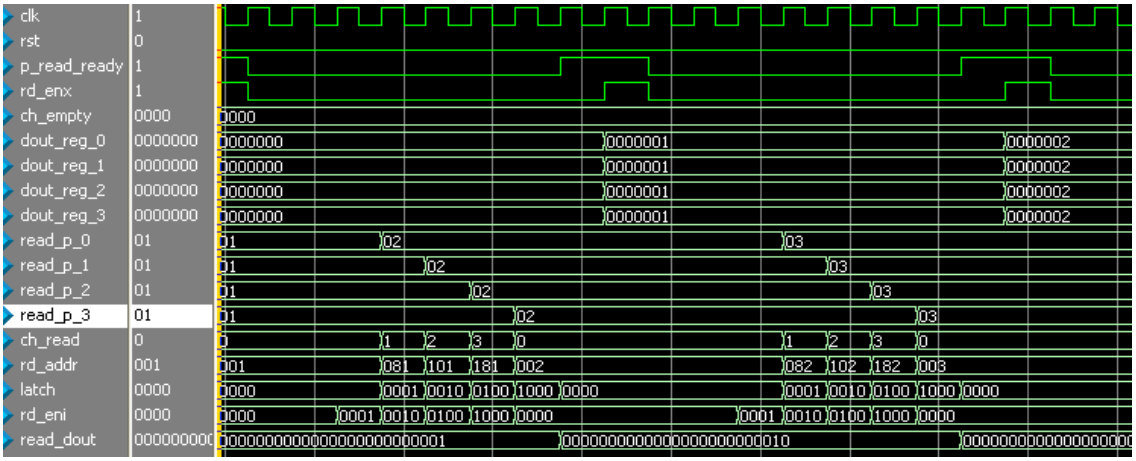


Figure 10. Timing diagram of parallel read

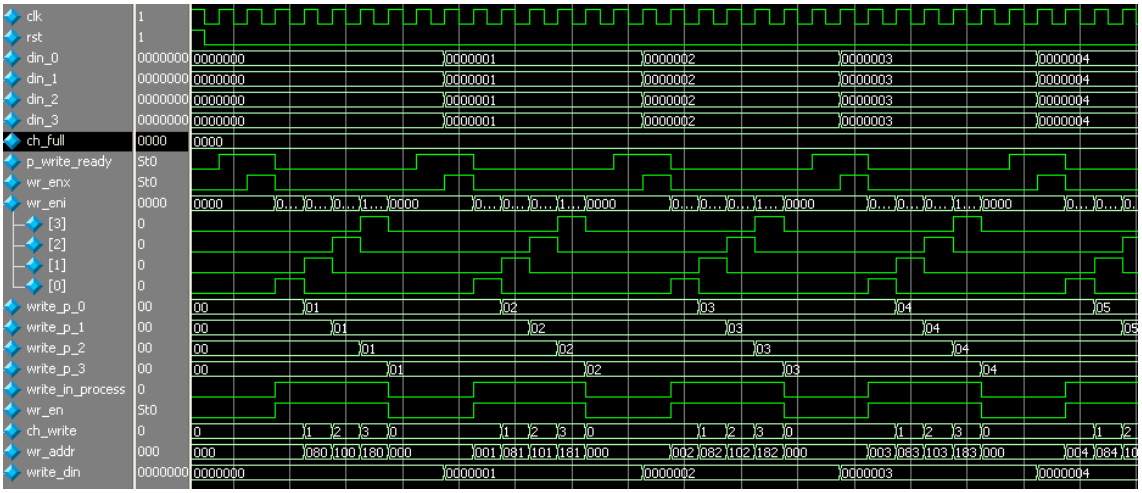


Figure 11. Timing diagram of parallel write

C. Conclusion Of Simulation

Through the simulation and analysis above, we can found that the timing waveform is consistent with the previous design of multi-channel FIFO, which is implemented with one BlockRAM and qualified to provide the parallel write or parallel read.

V. APPLICATION

In practice, the multi-channel FIFO is integrated in a FlexRay [12] communication controller (FCC) based on FPGA. In this project, we select the XC3S400A as the main chip. Its on-chip resources include 400k logic gates, 3584 logic slices, 56Kb distributed RAM, 360Kb BlockRAM, 4 DCM, and so on[13], the total amount of resources can meet the requirement of project. Analysis indicates that there are at least 32 FIFOs needed in the project. 16 of them are used as input data buffer with the structure of parallel write and serial read, the others are used as output data buffer with the structure of serial write and parallel read. If every FIFO is created by the IP core generator directly, these buffers would cost 32 pieces of BlockRAMs. But there are only 20 BlockRAMs in XC3S400A, and then resource of BlockRAM will be seriously insufficient. If we use XC3S1400A instead as the main chip, this situation would be relaxed slightly, because there are 32 BlockRAM in XC3S1400A, it is just enough for the data buffer. But the BlockRAM may be used somewhere else and it would be insufficient again. What's more, the substitute is much more expensive. Further analysis shows BlockRAMs in XC3S400A is perfectly adequate in total capacity but insufficient in number. We decide to employ the method proposed above to realize multi-channel FIFO in one BlockRAM. Study of project shows that the input port has a big data flow while the output port has a relatively less data flow. Hence, the capacity of FIFO of input port should be larger than that of the output port. At last, we decide to implement a 2-channel PISOFIFO in one BlockRAM for input port, each channel is 256 x 25 in size, and 4-channel SIPOFIFO in one BlockRAM for output port, and each channel is 128 x 25 in size. The structure of buffer for FlexRay communication controller is depicted in Figure 12. At last, 12 pieces of BlockRAMs is used for buffers

and 8 pieces of BlockRAM is remained for other purpose. The debug result shows that the system works in accordance with the design as expected.

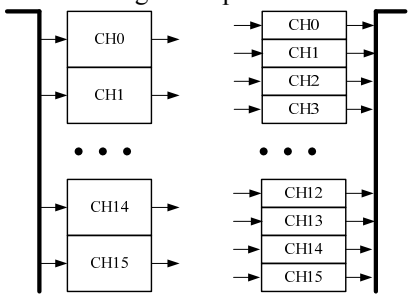


Figure 12. Structure of buffer for FCC

VI. CONCLUSION

This paper introduces a method of implementation of multi-channel FIFO with the utility of parallel access in a single piece of BlockRAM. This method can help to make full use of limited BlockRAM resources in FPGA and reduce the cost of terminal product, and it shows some practical value in project base on low-cost FPGA design. The simulation test and practical application indicate that this method is correct. The multi-channel FIFO created with it can run at a high frequency and is suitable for integration in most FPGA based projects.

APPENDIX A IMPLEMENTTATION WITH VERILOG HDL

A.1 Implementtation Of SIPOFIFO

```
module sipo_fifo(
      … //description of signals
          );
//set internal read command
  always @ (posedge clk or posedge rst)
      if(rst) rd_en_inter_set<=1'b0;
      else if(rd_en_inter_set|rd_command_time)
     rd_en_inter_set<=1'b0;
      //no channel is empty and registers are free
      else if(~|ch_empty&~|reg_label)
rd_en_inter_set<=1'b1;
//labels for buffering registers.
always @ (posedge clk or posedge rst)
      if(rst) reg_label[0]<=1'b0;
      else if(rd_en_inter_shift[3])reg_label[0]<=1'b1;
      else if(rd_enx)reg_label[0]<=1'b0;
…
//latch data to registers.
always @ (posedge clk or posedge rst)
      if(rst)dout_0<=25'h0;
      else if(latch[0]) dout_0<=read_dout;
//generate the write address
wire[3:0] wr_eni=wr_enx;
always @( * )
      case(wr_eni)
            4'b0001: ch_write=2'b00;
            4'b0010: ch_write=2'b01;
            4'b0100: ch_write=2'b10;
            4'b1000: ch_write=2'b11;
      endcase
assign wr_addr={ch_write,
    ({7{wr_eni[0]}}& write_p_0[6:0])|
    ({7{wr_eni[1]}}& write_p_1[6:0])|
    ({7{wr_eni[2]}}& write_p_2[6:0])|
  ({7{wr_eni[3]}}& write_p_3[6:0])};
wire wr_en=|wr_eni;
//select the data to write to DPRAM
assign   write_din= ({25{wr_eni[0]}}& din_0)|
                    ({25{wr_eni[1]}}& din_1)|
                    ({25{wr_eni[2]}}& din_2)|
                    ({25{wr_eni[3]}}& din_3);
//label logic
always @ (posedge clk or posedge rst)
      if(rst) p_read_ready<=1'b0;
      else  p_read_ready<=&reg_label&~rd_enx;
// instance of DPRAM
dpram25_512 dpram1 (
  .clka(clk),
  .wea(wr_en),
  .addra(wr_addr),
  .dina(write_din),
  .clkb(clk),
  .rstb(rst),
  .addrb(rd_addr),
  .doutb(read_dout) );
…
endmodule
```

A.2 Implementtation Of PISOFIFO

```
module piso_fifo(
…//description of signals
      );
wire[3:0] rd_eni=rd_enx;
always @ (posedge clk or posedge rst)
      if(rst) p_write_ready<=1'b0;
      else  p_write_ready<=
       ~wr_enx&~write_in_process&~|ch_full;
        //not full, last write operation is over,
always @ (posedge clk or posedge rst)
      if(rst) write_in_process<=1'b0;
      else if(wr_enx&~write_in_process&~&ch_full)
            write_in_process<=1'b1;
      else if(wr_eni[3])write_in_process<=1'b0;
//read buffering registers
always @ (posedge clk or posedge rst)
      if(rst)
            begin
              din_reg_0<=25'b0;
              din_reg_1<=25'b0;
              din_reg_2<=25'b0;
              din_reg_3<=25'b0;
            end
      else if(wr_enx&p_write_ready)
            begin
                  din_reg_0<=din_0;
                  din_reg_1<=din_1;
                  din_reg_2<=din_2;
                  din_reg_3<=din_3;
            end
always @ (posedge clk or posedge rst)
```

```
        if(rst) wr_eni<=4'b0;
        else wr_eni<={wr_eni[2:0],wr_enx};
//write to DPRAM
assign wr_en=|wr_eni;
always @( * )
        case(wr_eni)
                4'b0001: write_din=din_reg_0;
                4'b0010: write_din=din_reg_1;
                4'b0100: write_din=din_reg_2;
                4'b1000: write_din=din_reg_3;
        endcase
//instance of DPRAM
dpram25_512 dpram1 (
    .clka(clk),
    .wea(wr_en),
    .addra(wr_addr),
    .dina(write_din),
    .clkb(clk),
    .rstb(rst),
    .addrb(rd_addr),
    .doutb(read_dout) );
…
endmodule
```

ACKNOWLEDGEMENT

The work in this paper is supported by The National Natural Science Foundation of China (60674003), Major projects of Fujian Province (2011HZ0002-1), Scientific Research Fund of Hunan Provincial Education Department (12C0638) , Research Fund of Hunan Institute of Engineering (2013GK3034), Provincial Natural Science Foundation of Hunan(13JJ9022) and Provincial Science & Technology plan project of Hunan(2013GK3029).

The authors also would like to thank the key laboratory for automotive electronics and electric drive of Fujian province for the support in testing work she provided.

REFERENCE

[1] Z.Lu, Y.Wu, "A Rotation-based Data Buffering Architecture for Convolution Filtering in a Field Programmable Gate Array", Journal of Computers, 8(6): pp.1411-1416, 2013.

[2] S.Kasap and K.Benkrid, "Parallel Processor Design and Implementation for Molecular Dynamics Simulations on a FPGA-Based Supercomputer", Journal of Computers, 7(6): pp.1312-1328, 2012.

[3] A.M.Fernandes, "HDL Based FPGA Interface Library for Data Acquisition and Multipurpose Real Time Algorithms", IEEE Transactions on Nuclear Science, vol.58, pp. 1526-1530, 2011.

[4] A.Afaneh, Y.He, "Implementation of accurate frame interleaved sampling in a low cost FPGA-based data acquisition system", International Conference on Intelligent Data Acquisition and Advanced Computing Systems, IEEE, pp.20-25, Sept. 2011.

[5] G..W.Zhong; H.B.Zheng, et al., "1024-point pipeline FFT processor with pointer FIFOs based on FPGA", International Conference on VLSI and System-on-Chip, IEEE, pp.122-125, Oct. 2011.

[6] H.S.Han, K.S.Stevens, "Clocked and asynchronous FIFO characterization and comparison", IEEE International Conference on Very Large Scale Integration Oct, IEEE, pp.101-108, 2009

[7] Xilinx Inc, LogiCORE IP FIFO Generator v8.2, http://www.xilinx.com, 2011.

[8] Xilinx Inc, LogiCORE IP Block Memory Generator v6.2, http://www.xilinx.com, 2011

[9] M.A.Khan, A.Q.Ansari, "n-Bit multiple read and write FIFO memory model for network-on-chip", World Congress on Information and Communication Technologies, IEEE, pp.1322-1327, Dec. 2011.

[10] Xilinx Inc, ISE Design Suite Software Manuals, http://www.xilinx.com, 2011.

[11] Mentor Graphics Corporation, ModelSim SE User's Manual, http://www.mentor.com, 2009.

[12] FlexRay consortium, FlexRay communication systems protocol specification, version 3.0.1. http: //www.flexray.com, 2010.

[13]Xilinx Inc, Spartan-3A FPGA Family, http://www.xilinx.com, 2009.

Zhipeng Gong Received the B.S. in physics from the Hunan Normal University, Changsha, China, in 1998, and the M.S. in transport information engineering and control from Central South University, Changsha, China in 2005. He is currently working toward the Ph.D. degree with the School of Information Science and Engineering, Central South University.

He is also currently a faculty member with the College of Electrical and Information Engineering, Hunan Institute of engineering, Xiangtan, China. His research interests include embedded system, FPGA and communication network in vehicle.

Fast Mode Decision and Encryption Policy in H.264/AVC Frame-skipping Transcoding

Xiaohong Zhang
School of Information Engineering /Jiangxi University of Science and Technology, Ganzhou, China
Email:xiaohongzh@263.net

Baolin Qiu
School of Information Engineering /Jiangxi University of Science and Technology, Ganzhou, China
Email:934917715@qq.com

***Abstract*—H.264/AVC adopts Rate-Distortion Optimization (RDO) techniques to search all the prediction modes supported by standard in turn, then determine the optimal encoding mode which has the minimum rate-distortion cost. Frame-skipping transcoding among heterogeneous networks is a way to lower bitrate transcoding, which can solve incompatibility issues of video stream transmission. Aiming to varied motion problem of the video sequences, this paper introduces a set of adaptive threshold for inter mode decision, which can make a more correct code judgment of motion characteristics for the macroblock, also can obtain more reliable video quality; Combing the mode information of the skipping frame which can not be recovered, a simple and practicable encryption policy is introduced here, which avoids the shortcomings of traditional encryption algorithm complexity. Experimental results show that the encoding time of the new algorithm is almost the same with that of the referenced algorithm, the peak value SNR increases steadily, so novel scheme guarantees nice video quality; Encryption policy achieves the desired encryption effect and efficiency, and broadens the applications of the video security technology.**

***Index Terms*— H.264/AVC; frame-skipping transcoding; adaptive threshold; inter mode decision; encryption**

I. INTRODUCTION

With the rapid development of media technique, people pay growing attention and concern about the video quality and safety. As the newest video coding standard, comparing with H.263, H.264/AVC developed by both ITU-T VCEG and ISO/IEC MPEG standards committees displays a number of new ideas and new techniques, such as H.264/AVC divides operation processing of encoding and network into two separate levels (layer): VCL (Video Coding Layer) and NAL (Network Abstraction Layer); flexible block transform in motion compensation; introducing 1/4 pixel precision motion compensation; loop de-blocking filtering technologies, and so on. H.264/AVC not only has tremendous progress and improvement in the video compression efficiency (under the same video quality, the compression ratio of H.264/AVC is twice as that of H.263), but also has been greatly improvement in the transmission, network compatibility [1]-[2]. Facing different network transmission limitations, original video stream is often required to decode, then undergo the second encoding or the other corresponding processing operations to meet vary requirements of terminal device, this process is defined as video transcoding. In H.264 standard, ME (Motion Estimate) and MC (Motion Compensation) are complicated and time-consuming [3]. The time they cost is more than 60% of the entire coding time [4]. It's the main factor that impacts the implementation and application of H.264 [5]. Obviously, in video transcoding, using the mode traversal search strategy will decrease transcoding efficiency, so that the time complexity increases dramatically. Frame-skipping transcoding aims at solving the limitation of data transmission rate and bandwidth between heterogeneous networks, which is a kind of transcoding converting the higher quality compression code stream into lower code rate compression stream. Fig. 1 is a simplified schematic diagram of transcoding:

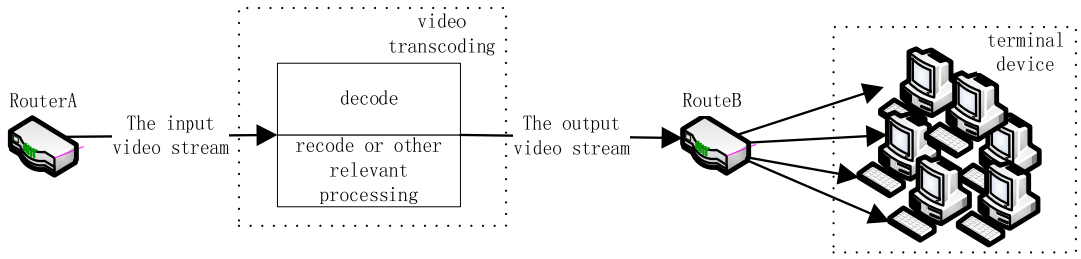


Figure 1. Schematic diagram of video transcoding

doi:10.4304/jcp.9.5.1201-1208

In order to avoid unnecessary mode search and compute in the frame-skipping transcoding process, as well as encrypts important frame, we can achieve accelerating encoding rate and access to safe video goals. In the field of skipping transcoding, many scholars and experts conduct a series of studies and improvement about relevant algorithms [6]-[8], for example, according to the macroblock motion features to avoid unnecessary mode search in order to accelerate transcoding speed [7]; selectively using the screened out mode to reduce the number of encode modes in order to accelerate transcoding efficiency [9]; combined the time-domain correlation of macroblock to reduce the number of prediction modes in order to accelerate transcoding speed [10] and so on. To ensure the security of video, a series of achievements have obtained, such as detecting and extracting the interesting regions, then integrate with H.264 encryption algorithms reasonably [11]; a selective encryption algorithm is proposed to protect video data which can performed in both types of CABAC and CAVLC [12] .

The key of frame-skipping transcoding is using the original macroblock coding information as reasonable as possible. Combined with encryption policy, it can not only speed up the second video encoding but also ensure the video quality and safety. Reference [7] introduces a pair of fixed threshold values to define and classify the movement of macroblock in the frame, and adopts the corresponding coding mode selection algorithm for the different motion characteristics of the macroblock. Therefore, the algorithm can eliminate some unnecessary mode search and computer, which plays a significant role on the enhancement for the execution efficiency of frame-skipping transcoding. However, the motion characteristics of video sequences are quite unstable in the real world. If we use the same fixed threshold criteria to measure and define the video feature when the motion characteristics of video sequence vary considerably large, it is not accurate. At the end, the inaccurate judgment of movement characteristics will result into the decrease of video quality. This paper introduces adaptive transcoding thresholds for fast mode selection and encryption policy. According to the current movement of macroblock, it is flexible to change the value of adaptive thresholds in the process of program execution which is used to define the motion characteristics of macroblock. Therefore, the judgment that what modes of macroblock need to be searched and calculated again is more accurate. Then, encrypting the key frame of video sequences with the encryption policy introduced here will get quality assured video sequences ultimately. Experimental results show that: on the premise that ensures the efficiency of frame-skipping transcoding, the algorithm introduced here achieves higher peak signal to noise ratio and better encryption effect, obtains better applicability.

II. H.264/AVC INTER-MODE DECISION

In the inter frame coding of H.264/AVC, it can store reconstructed image of the current coding frame in order to provide reference frame for the inter frame coding. The coding unit of inter frame coding is 16×16 macroblock. Relevant regulations of H.264 show that there are seven segmentation methods for inter frame coding of macroblock. Macroblock size of 16×16 can be divided into 16×16, 16×8, 8×16 or 8×8, segmentation method of 8×8 will proceed subdivision into sub-macroblock. Every sub-macroblock will be divided into 8×8, 8×4, 4×8 and 4×4, then conducts motion estimation and rate-distortion cost calculation to determine the optimum encoding mode of the current macroblock, respectively. Table I gives the corresponding relationship between the inter coding mode and segmentation method in baseline profile of H.264. For example: horizontal size is 16, vertical size is 8, corresponding to the segmentation method is 16×8. “2 sizes/mode 2”, in the corresponding cell means that the macroblock is divided into two sub-macroblocks whose size is 16×8 and corresponding coding mode is mode 2.

Procedures for inter mode decision of macroblock in H.264/AVC standard are as follows:

Step 1:According to segmentation method of 16×16,16×8,8×16 in the macroblock level,conduct motion estimation and calculate the rate-distortion cost for the current macroblock to be encoded in turn;

Step 2: As to the 8×8 segmentation method followed, conduct motion estimation, calculate rate-distortion cost for the four sub-macroblock by the way of 8×8,8×4,4×8 and 4×4 in turn,then select the mode with the minimal value of rate-distortion cost as the best encoding mode of current 8×8 sub-blocks ;

Step 3: Add the minimum rate-distortion value of the four 8×8 sub-blocks calculated respectively in step 2 up,use the sum as rate-distortion cost of the current macroblock in 8×8 segmentation method;

Step 4: Calculate motion vector and rate-distortion value of the mode 0 (Direct mode) ;

Step 5: Calculation rate-distortion cost of mode 9 (I4MB, intra mode) and mode 10 (I16MB, intra mode) in turn;

Step 6: Select the mode with minimum rate-distortion cost from step 1,3,4,5 as the best inter encoding mode of the current macroblock.

TABLE I.
CORRESPONDENCE TABLE OF SEGMENTATION METHOD AND CODING MODE

Vertical / Horizontal	16	8	4
16	1 size /mode 1 or I16MB	2 sizes /mode 2	not have
8	2 sizes /mode 3	4 sizes /mode 4	8 sizes/mode 5
4	not have	8 sizes /mode 6	16 sizes /mode 7 or I4MB

The way that H.264 searches and calculates rate-distortion cost for all supported modes will result in a substantial increase in execution time and lead to the decline of coding efficiency. Therefore, the search

strategy used in H.264 standard is inadvisable in the frame-skipping transcoding. Otherwise, it will limit the implementation of real-time.

III. MODE DECISION OF FRAME-SKIPPING TRANSCODING

A. The Frame-Skipping Framework Based On H.264/AVC

In the original compressed stream, there is much helpful macroblock information can be used in frame-skipping transcoding, such as mode information, the pixel residuals and motion vector information and so on. Appropriate using of such information will play a significant role in promoting the transcoding speed. Fig. 2 is map that shows the CODEC (enCOder/DECoder pair) framework of frame-skipping transcoding based on H.264

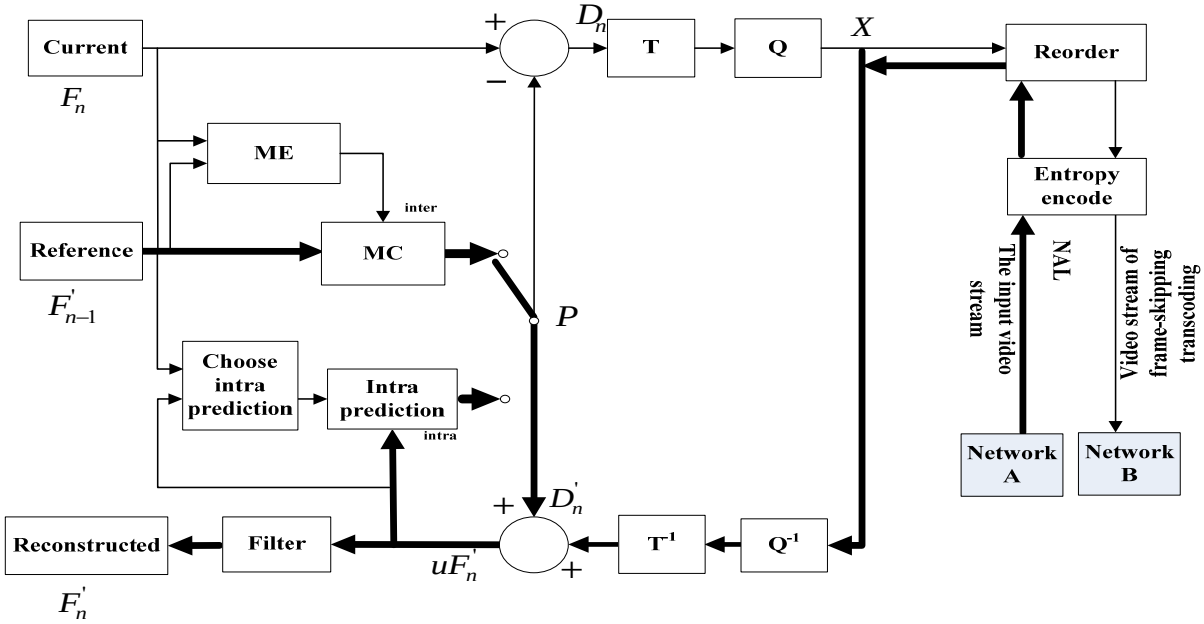


Figure 2. The CODEC framework of frame-skipping based on H.264.

The figure contains two part processes of the encoder and the decoder, the thin line part shows process of encoder, the thick line part shows process of decoder.

B. Fast Mode Decision Of Frame-Skipping Transcoding

Introduce a set of variable thresholds to achieve real-time adaptive capacity of the motion feature for current video sequence so that the algorithm can distinguish the actual motion characteristics more precisely for macroblocks of different video sequences. Thus, achieve more appropriate and accurate classification of the current macroblock, which leads to the result of better video quality.

Generally, it's scarcely possible that the adjacent video images make sudden change, so the adjacent image frames have a very strong correlation. During frame-skipping transcoding process, with this principle, the encoding mode of the current macroblock $curr\_mode\,[n][k]$ (n is the number of frame, k is the number of macroblock) makes reference to the mode of macroblock k in frame n ($mode[n][k]$) and the mode of the dominant macroblock k in frame n-1 ($mode[n\text{-}1][k]$). Select the mode with smaller ratio-distortion cost as the inter encoding mode of current macroblock by calculating and comparing the cost. The macroblock of the skipped frame n-1 whose area covered by the current macroblock is the largest among the four covered macroblock is defined as the dominant macroblock.

Equation (1) is proposed by reference [7], defines a variable that is used to measure and characterize motion characteristic of the current macroblock. The value of the variable is the sum of the two absolute values of original motion vector's horizontal and vertical components.

$$MM_{MB} = |MV[0]| + |MV[1]|. \quad (1)$$

$MV[0]$ is the horizontal component of original motion vector, $MV[1]$ is the vertical component of original motion vector.

As MM_{MB} reflects the motion characteristic of the current macroblock, the variable also shows the degree of relevant for $mode[n\text{-}1][k]$ and $mode[n\text{-}1][k]$ in the two adjacent frames.

(1) When the value of MM_{MB} is small, indicate that the time-domain characteristics of current macroblock are relatively stable, the two adjacent frames have a strong correlation. Use $mode[n][k]$ ($sub\_mode[n][k]$, sub-macroblock mode) as $curr\_mode\,[n][k]$ directly;

(2) When the value of MM_{MB} is not small enough, indicate that the time-domain characteristics of current macroblock are not stable, the two adjacent frames

haven't a strong correlation. The algorithm should computer RD cost of $mode[n][k]$ ($sub\_mode[n-1][k]$) and $mode[n-1][k]$ ($sub\_mode[n-1][k]$) to determine $curr\_mode[n][k]$.

(3) When the motion of video sequences is too intense, means that the time-domain characteristics of current macroblock are quite unsteady, so algorithm needs to search all modes supported in H.264 to determine the best encoding mode of the current macroblock.

C. Fast Mode Decision With Adaptive Threshold

In order to obtain more guaranteed quality of video sequences, this paper introduces a simple and linear calculation method to calculate the changeable threshold. The threshold interval is defined as [low_value , $high\_value$]. Combined with this set of threshold, the algorithm classifies the motion characteristic of current macroblock.

$$\begin{cases} low\_value = a-k*ratio \\ high\_value = b-k*ratio \end{cases} . \quad (2)$$

$ratio$ in the equation is discrete, its value is natural number. Comparing PSNR Y (PSNR of Y component) values of the two adjacent frames, when a continuous decline in the value of PSNR Y occurs, then the value of $ratio$ is 1; when continuous decline in the value of PSNR Y occurs twice, then the value of $ratio$ is 2 and the value of $ratio$ on the other situation can be obtain in the same manner; if PSNR Y doesn't decline, then the value of $ratio$ is 0. k in the equation reflects the adaptation speed that the threshold adapts the movement of the current video sequence. Therefore, it will bring inaccurate reflection to the judgment of motion characteristic of macroblock when the value of k is assigned too large or too small. According to experimental statistics of video sequences, assign 5 to k. Introduce variable $deta$, its value is determined by the decreased PSNR Y value of the two adjacent frames. $deta$ characterizes the intensity of movement. If the decreased PSNR Y value of the two adjacent frames is small, then $deta = 0$; If the decreased PSNR Y value of the two adjacent frames is large, then, $deta = 1$.

From (2) we know that the adaptive threshold value is determined ultimately by basis value (i.e. a and b) and the fine turning value(i.e. Minus $k*ratio$). Reasonable assigning of basis value can adapt the video motion feature faster, and the fine turning can make the threshold adapt the motion feature of video sequence more accurately. The more PSNR decreases, the worse the adaptability of threshold to the video, i.e., threshold is too large or too small. Combed with algorithm analysis, on the one hand, when the threshold value is too large, it will result in that most of the values of MM_{MB} are smaller than the threshold, so the current macroblock will continue to follow original mode to re-encode. While the original encoding mode is not necessarily the best encoding mode in frame-skipping transcoding, which will result in that the quality of video declines i.e. PSNR drops. On the other hand, the coding efficiency will decline if the threshold is too small. Therefore, under the premise of guaranteed coding efficiency, in order to avoid mischoosing the inaccurate mode by mistake and obtain more accurate judgment of motion characteristics, the paper assigns smaller number to basis value appropriately. For the sake of faster and flexible adaptation of video movement, design two levels for basis value a and b.

If $deta = 0$, then $a = 65$ and $b = 85$;

If $deta = 1$, then $a = 35$ and $b = 50$.

The initial values of a and b are determined by experimental statistics of a series of different video sequences with varied motion characteristics.

In the process of frame-skipping transcoding, compare MM_{MB} reflected the motion characteristic with low_value and $high\_value$. If $MM_{MB} \le low\_value$, means that motion characteristic of macroblock is stable; otherwise, it's unstable, algorithm needs to calculate some rate-distortion values to finally determine $curr\_mode[n][k]$ .The following is specific algorithm steps:

Step 1: Calculate the value of (1), and determine the value of $ratio$ and $deta$ by comparing the PSNR Y of adjacent frames;

Step 2: According to the change of PSNR, determine the value of a and b;

Step 3: Combined with $ratio$, a and b, work out the current value of low_value and $high\_value$;

Step 4: Compare MM_{MB} with low_value. If more than half of macroblocks (sub-macroblocks) satisfy $MM_{MB} \le low\_value$, the motion characteristic of macroblock is slow and stable. Then, use $mode[n][k]$ ($sub\_mode[n][k]$) as the optimal encoding mode of the current macroblock. Otherwise, go to Step 5;

Step 5: Compare MM_{MB} with $high\_value$. If more than half of macroblocks (sub-macroblocks) satisfy $MM_{MB} > high\_value$, the motion characteristics of macroblock is violent movement. Determine the optimal encoding mode of the current macroblock by searching all inter mode supported in H.264. Otherwise, go to Step 6;

Step 6: Otherwise, satisfy $low\_value \le MM_{MB} < high\_value$. In this case, determine the optimal encoding mode of the current macroblock by calculating RDcost of $mode[n][k]$ ($sub\_mode[n][k]$) and $mode[n-1][k]$ ($sub\_mode[n][k]$). Use the mode with smaller RDcost as the optimal mode of current macroblock;

The following flow chart shows the inter encoding mode decision of frame-skipping transcoding with adaptive threshold:

Since the algorithm avoids motion estimation and calculation of some unnecessary modes, the execution time of program is reduced, which promotes the frame-skipping transcoding speed.

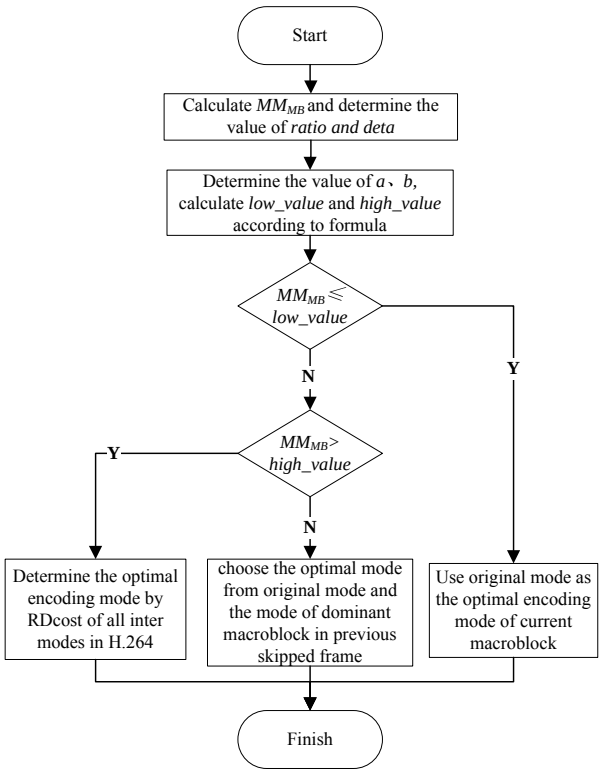


Figure 3. The flow chart of inter encoding mode decision.

IV. ENCRYPTION POLICY OF FRAME-SKIPPING TRANSCODING IN H.264/AVC

Nowadays, with the rapid development of multimedia technology and popularization of network, multimedia information, people's life, and daily work are inextricably linked. Since video data of frame-skipping transcoding relates to many security issues like its own restrictions of network (such as network congestion, bandwidth limitations, etc.) and human factors (such as information theft, data corruption during network transmission, etc.), also relates to users' security requirements of video data [13]. Therefore, introduce secure encryption algorithms to protect frame-skipping transcoding data. And use the mode information of macroblock in skipped frame which comes from frame-skipping transcoding to process the encryption key.

A. Traditional Video Encryption

As to the research of video coding encryption algorithm, many achievements have been acquired now. The security of encryption algorithm is closely related to the encryption positions which even directly affect the efficiency and practicability of encryption algorithm. According to the designer's consideration of the encryption algorithm, the selection of encryption position is flexible, such as some header information of NAL (Network Abstraction Layer) data packets, MVD (motion vector difference), some key syntax elements, DCT. Generally, the main encryption way is data scrambling or critical data extraction.

B. The Encryption Policy By Removing The First Frame Of Bit Stream

Inspired by the unrecoverable information of skipped frame in the process of frame-skipping transcoding, propose an encryption policy to destroy some vital information reasonably when conduct encoding, which results in that the decoder client without the authorization can not correctly decode the bit stream of video by the way of statistical analysis or other means. Therefore, the video data can achieve a high level of security.

In the inter-frame coding process of H.264 standard, the reference image of the current slice which has encoded before conducts motion compensation to obtain the prediction PRED (i.e. indicated by P in Fig. 3). The PRED is subtracted the current block, then get a residual or difference macroblock D_n. After transformation and quantization of D_n, a set of quantized transform coefficients X is obtained. Take these coefficients to re-order as well as entropy. Together with some side information (such as the prediction mode of macroblock, motion vector, etc) which is indispensable for decoding macroblock, achieve the compressed bit stream. Then, the bit stream will be passed to NAL to store or transmit to the receiver.

Combined with the brief statement of H.264 encoding process in the previous section, obviously, I frame is the foundation for the decoding of subsequent P-frame, so I frame is crucial. Although the video format is unbroken, residuals is undamaged, motion vector is correct, the decoder can not decode the original video sequence without PRED information. Therefore, this encryption algorithm that remove the whole compressed bitstream unit of I frame which is used as reference frame during encoding video is proposed and implemented. Undoubtedly, it is serious damage for the compressed bitstream of original video sequence. The whole encrypted compressed bitstream of P frames which is lack of the whole compressed bitstream unit of I frame are stored and transmit to the client-side. Therefore, the subsequent P frame can never be decoded correctly without supporting information of reference frame (i.e. I frame) if the compressed bitstream of video sequence is obtained by an unauthorized intruder. Then, the other frame which uses this subsequent P frame as reference frame can never continue to be decoded, which will trigger the chain reaction. The reaction leads to a result that the whole compressed bitstream of video can not be decoded, which guarantees high security but easy to do.

This encryption algorithm uses NALU (NAL unit) of I-frame compressed data which is removed completely as the encryption key. The key should be transmitted to legitimate users through secure channel. The decryption algorithm can decrypt the encrypted video sequence with the encrypted compressed bitstream and the key which is received by legitimate user. The following Fig. 4 shows the main process chart of encryption policy:

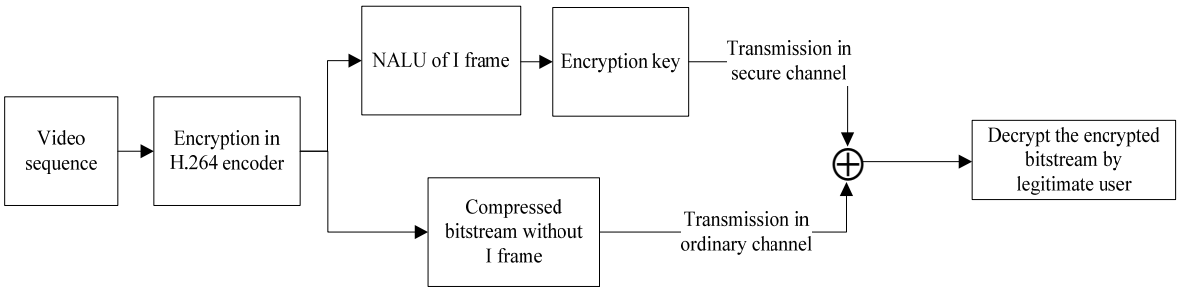


Figure 4. The process chart of encryption policy.

The encryption algorithm only needs to conduct some code integration for the encoder in JM of H.264 coding standard. The program has strong portability. In addition, it is simple and practicable, which makes it easy to promote practical applications; the encryption algorithm needn't keep secret, the security of video bitstream only depends on the key; the brute force attack on the encryption algorithm is unrealistic, which makes the security of encryption compressed bitstream is relatively high.

C. The Key Processing

During frame-skipping transcoding, use the mode data of macroblock in skipped frame to encrypt coding mode of macroblock in the I frame which is used as the key to scrambling encryption. The encrypted mode *mode_encryption* is calculated by (3):

$$mode\_encryption = best\_mode + rand + mode\_skip . \quad (3)$$

In (3) *best_mode* is the optimal encoding mode of macroblock before encryption. *rand* is an artificially generated random integer within limits which can be changed by man. *mode_skip* comes from the non-zero encoding mode of macroblock in the skipped frame.

It is arbitrary nature of *rand* and unrecoverable nature of *mode_skip* that make the concealment and security of *best_mode* stronger, which produce one more security assurance. The operation of protection to the key only involves addition and subtraction, which will just cause a little computational overhead. Therefore, appropriate protection to the key isn't making an unnecessary move. On the contrary, higher security of service for the user is obtained at a quite low computational cost. At first, free the operation of protection for the key after legitimate user receives all the encrypted bitstream. Then, with the key, the encrypted compressed bitstream can be decrypted correctly. The correct video sequence will be obtained at last.

V. Experimental Results and Analysis

The experimental platform is the baseline profile of H.264 in JM8.6, the operating system is XP system, use PC with 2.09GHz dual-core intelprocessor and 1.99GB memory. In order to verify the effectiveness of the proposed algorithm, experiment uses five test sequences with different motion characteristics. The size of video sequence is 176×144. The test sequences are akiyo_qcif.yuv, mobile_qcif.yuv, foreman_qcif.yuv , highway_qcif.yuv and carphone_qcif.yuv respectively. Each sequence contains 200 frames. The first frame is the I frame, the rest of frames are P frames.

A. Experimental Results

Experimental results are composed of three part data. The first part comes from frame-skipping transcoding with standard mode search strategy of H.264, the second part comes from frame-skipping transcoding with fast mode decision algorithm proposed in reference [7], the last part comes from frame-skipping transcoding with the fast mode decision used adaptive threshold which is proposed by this paper. The paper mainly compared PSNR Y and time of motion estimation (Time). The specific results are shown in Table 2

The encrypted video bitstream without any information of I frame can't get reference frame. The illegal user can never decrypt the encrypted bitstream as well as seen the decrypted results without correct key. Use the right key to decrypt the encrypted bitstream and select one frame randomly (select unified the 55th frame) to be displayed in Fig. 5:

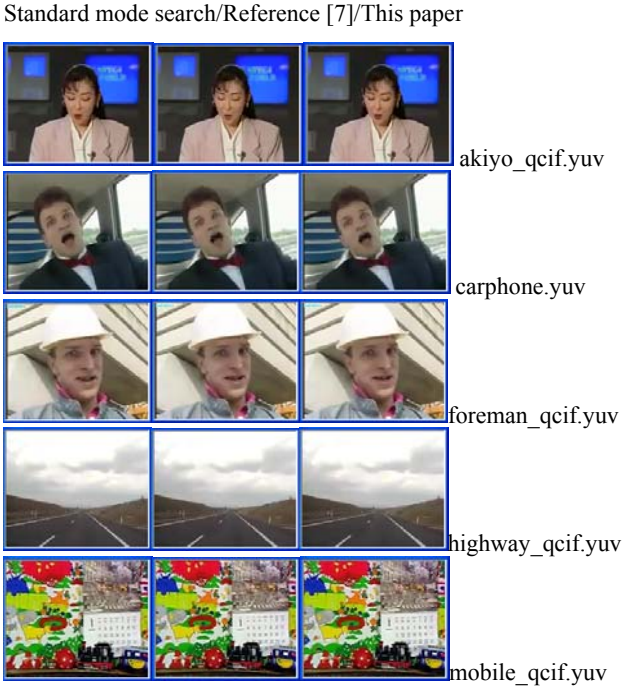


Figure 5. The 55th frame after decrypting correctly

B Analysis Of Experimental Results

Data obtained from the above experiments indicated that:

(1) Synthesize the motion estimation time that the former 200 frames of the five test sequences. Although, standard mode search pattern obtained the higher peak signal to noise ratio, the total time spent on motion estimation increases, because the standard results in the increasing of complexity. The average time that algorithm proposed by this paper is 84.8654s less. The average time spent on motion estimation is reduced by about 35.25%;

(2) Compared with the reference [7], the PSNR of video sequence is improved steadily in this paper while the time of motion estimation are basically maintain the same time, which achieves more guaranteed quality of video service under the premise of good efficiency in process of video encoding;

(3) From the point of visual sense, image quality obtained by this paper are almost the same as that obtained by standard mode search pattern, and the other frames of the test sequences have similar experimental effects;

(4) The encrypted video bitstream encrypted by this paper can not be decrypted correctly without the decrypting key. The decrypted video sequences are completely correct when the correct key is obtained, which is related to the way of encryption directly.

So you can see that the fast mode decision and encryption policy proposed by this paper can achieve better effect of frame-skipping transcoding under the premise of guaranteed execution efficiency. Also, obtain more reliable video quality and more secure video sequence.

VI. CONCLUSIONS

The rational use of information of macroblock in original compressed bitstream can greatly promote the efficiency of frame-skipping transcoding. According to real-time movement of different video sequences, the adaptive threshold for frame-skipping transcoding proposed by this paper can use motion vector in the original compressed bitstream more accurately and reasonably. Experiments show that the algorithm proposed makes the video quality more reliable and more guaranteed under the premise of high efficiency of transcoding, as well as the encryption policy is simple and easy to do while it makes the video service keep a high security.

ACKNOWLEDGEMENTS

This work is jointly supported by the National Natural Science Foundation of China (No. 11062002, 61363076), Natural Science Foundation of Jiangxi Province (No. 20122BAB211019) and Scientific Research Key Plan Projects of Jiangxi Education Department (Nos. GJJ13435).

REFERENCES

[1] M. M. Jeong. “A new low-complexity integer distortion estimation method for H.264/AVC encoder,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 2, 2010, pp. 207-212.

[2] C. Chiang, W. Pan, and C. Hwang. “Fast H.264 encoding based on statistical learning,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 9, 2011, pp. 1304-1315.

[3] X. Su, P. Bai, Y. Wu, and Y. Feng. "An Improved Cross Diamond Motion Search Algorithm Based on H. 264," *Journal of Computers*, vol. 6, no. 12, 2011, pp. 2603-2606.

[4] Z. Shen, Z. Li, and T. Pan. “Motion estimation based on the partition and evaluation of the search area in H.264/AVC,” *Journal of Image and Graphics*, vol. 15, no. 2, 2010, pp. 242-246.

[5] G. Yuan, P. Liu, and K. Jia. "A Fast Motion Estimation Algorithm Based on Motion Vector Distribution Prediction," *Journal of Software*, vol. 8, no. 11, 2013, pp. 2863-2870.

[6] Z. Lv, K. Jia, and W. Siu. “Fast inter mode decision and motion estimation in H.264 frame-skipping transcending,” *Application Research of Computers*, vol. 26, no. 12, 2009, pp. 4814-4817.

[7] Q. Tang, and Panos Nasiopoulos. “Efficient motion re-estimation with rate-distortion optimization for MPEG-2 to H.264/AVC transcoding,” *IEEE Transactions* on *Circuits and Systems for Video Technology*, vol. 20, no. 2, 2010, pp. 262-274.

[8] T. Shanableh, E. Peixoto, and E. Izquierdo, "MPEG-2 to HEVC Video Transcoding With Content-Based Modeling," IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no.7, 2013, pp. 1191-1196.

[9] X. Jin. “Low complexity H.264/AVC encoding and transcoding,” *Shanghai Jiao Tong University*, 2011.

[10] Z. Lv, K. Jia, and W. Siu. “Fast inter coding in frame-skipping transcoding based on SVM,” *Systems Engineering and Electronics*, vol. 34, no. 6, 2012, pp. 1266-1271.

[11] J. Yu, Q. Liu, and Y. He. “H.264 video encryption algorithm based on region of interest,” *CHINESE JOURNAL OF COMPUTERS*, vol. 33, no. 5, 2010, pp. 945-953.

[12] Z. Shahid, C. Marc, and P. William. "Fast Protection of H. 264/AVC by Selective Encryption of CAVLC and CABAC for I and P frames," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 5, 2011, pp. 565-576.

[13] T. Stutz, and A. Uhl. "A survey of h. 264 avc/svc encryption," IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 3, 2012, pp. 325-339.

Xiaohong Zhang, PH.D, professor, main researches are nonlinear dynamics theory, cellular neural networks, and information security.

Baolin Qiu is currently a master student in Jiangxi University of Science and Technology, China, His research interests include video compression coding.

Crowd Density Estimation based on Improved Harris & OPTICS Algorithm

Cheng Xu[1] , Hong Bao[1] * , Lulu Zhang[1] , Ning He[2]
Beijing Key Laboratory of Information Service Engineer,Beijing Union University, Beijing, China[1]
Information Technology College, Beijing Union University, Beijing, China[2]

***Abstract*—In this paper, we propose a method to estimate crowd density using improved Harris and Optics Algorithms. We pre-processed the raw images at first and the corner features of the crowd were detected by the improved Harris algorithm, then the formed density point data were used to analyze the corner characters of crowd density by the optics density clustering theory. This theory is related to the distribution of the feature points where the crowd density is estimated by the machine learning algorithm.We used a standard database PETS2009 to do the experiments in this paper and the self-shooting datasets to illustrate the effectiveness of our method. The proposed approach has been tested on a number of image sequences. The results show that our approach is superior to other methods including the original Harris algorithm. Our method improves the efficiency of estimation and has a significant impact on preventing the accidents on crowd area with high density.**

***Index Terms*—density clustering, Harris algorithm, OPTICS algorithm, crowd density**

I. INTRODUCTION

With the wide application of surveillance cameras, intelligent video surveillance technology leads to more and more researcher opportunities. It is an important technique to monitor the crowd in the video, and can be used in the public place for crowd management, auxiliary public service design, intelligent security monitoring and auxiliary case detection, virtual people, etc.

People are often in a state of motion and a moving populace is at risk of committing some accidents such as crowd stampede. The various physical information of crowd movement have gradually become the core issue in intelligent monitoring systems. Cameras are visible at traffic intersections, stadiums, stations and other places. Video processing technology on cameras is useful in personal monitoring and it can enhance the efficiency of management by improving the analysis technologies of crowd density. This can also permit real-time monitoring with a mass monitoring device and can analyze the scenario of crowded intelligent video content [1]. This paper is based on the improved Harris corner detection to detect the crowd characteristics in an image plane. Firstly, we used the Optics algorithm to do the density clustering analysis; then we utilized the reach ability graph for analysis, and then used machine learning to judge the situation of the contemporary crowd density . The results were divided into high and low density crowd.

The crowd monitoring theory mainly focuses on crowd density and statistics as in [2]. Related research methods on crowd density analysis can be divided into two ways which are based on pixels and texture statistics. The pixel method involves a complex calculation. It is suitable only for low crowd densities. V. Verona and A. N. Marana extracted crowd characteristics using gray level co-occurrence matrix, wavelet packet decomposition and other analysis methods [3]. Wavelet packet decomposition provides an effective way to decipher crowd images via multi-scale analysis. Texture analysis method has been used to solve the misclassification due to overlapping in high density crowd to a certain extent

Numerous international universities and research institutes regard the video monitoring system technology as an important research field [4, 7 and 8]. In 1995, the British EPSRC organization also studied crowd density estimation and the potential dangerous situations thereof. In 1997, the Defense Advanced Research Projects Agency (DARPA) has set up a major VSAM video surveillance project. W4 monitoring system has been used in 2000, and achieved the goals of detecting and tracking people who are in their homes, car parks, banks, etc. In 2003 European successively introduced PRISMATICA and ADVISOR system. In 2005, ISCAPS was established by more than a dozen European communications companies and academic organization; it aimed at realizing automatic intelligent monitoring of crowded areas. In 2013, the large-scale separation of pedestrians and portrait retrieval intelligent recognition system have obtained a breakthrough in China. It can carry on the processing of large-scale portrait; identify the key crowd automatically, and therefore provide safety for the public [5,18,19].

Davies proposed image processing methods by statistical pixels [2]. It can estimate crowd density by background subtraction and crowd pixel calculation. This method proved to be simple and effective. Background

Manuscript received 9 2013; revised 2, 2014;
This project was supported by the Project of Construction of Innovative Teams and Teacher Career Development for Universities and Colleges under Beijing Municipality (CIT&TCD20130513), and the National Natural Science Foundation of China (Grant No. 61271370 and grant No.61271369), and the project No. 201411417SJ053
*Corresponding author： Hong Bao;

doi:10.4304/jcp.9.5.1209-1217

extraction and the real-time realization of crowd density estimation are the most important parts. In some larger monitoring places, the big workload scenario training is often too broad to use pixel statistical method directly for crowd density estimation. It is beyond practical significance [6].

A. Albiol, et. al. proposed a statistical method to study the changing characteristics of a dynamic angular point [9]. The method estimated the behavior of the crowd by the changing rules of dynamic angular point. The advantage of this approach is not affected by illumination changes; but if people gather for a long time,it will cause a static target angular point which will then be deleted. In addition, if there is a small crowd of people appearing a few meters around the target which likewise become individual dynamic angular points, this will cause a large effect on the dynamic angular point area and result in a miscalculation. Moreover, this method only gives the change in the regularity of crowd gathering. There are no details for the identification scheme.

D. Conte, et.al. [10] improves the method of A. Albiol, et.al.[9]. It presents a kind of machine learning method to analyze the monitoring group by considering the various characteristics of different population densities. This method was able to solve certain problems which the direct nor the indirect detection methods can completely solve. This method has good robustness; it is based on a training, hence it is not necessary to set clear prior knowledge. It can be directly used to evaluate the relationship between angular point and density in a dynamic populace. The experimental results showed that this method improved the precision and maintains the robustness of the indirect method as in [9].

D. Conte, et.al.was able to monitor the crowd in efficiently and in real-time [11]. Their system exhibited high precision and high speed in finding the target, and hence can be used for real-time monitoring applications. Its main methods are based on extracting the scale invariant corner features according to the relationship of the camera and the actual population proportion. It divides bottom-up level area to associate angular point characteristics and horizontal zoning for a moving continuity situation. It makes each region obtain the same stereo effect from the same target .The method uses a simple training program with a good real-time performance and robustness. Such program can automatically evaluate the required parameters of the system. It is suitable for people counting applications and crowd behavior analysis, and can also be used to identify the segmented regions in different levels.

There are lots of density clustering algorithms based on image feature extraction; for example, OPTICS (Ordering Pointers to Identify the Clustering Structure) algorithm is a kind of automatic interaction density clustering algorithm which overcomes the strong dependency of parameters like DBSCAN (Density-Based Spatial Clustering of Applications with Noise). OPTICS Algorithm is a relatively advanced clustering method. It starts from arbitrary data objects then expands toward denser regions as far as possible. It finally gets the reachability-distance and expansion order of each data object. In order to express the expansion order more intuitively, each object data of the expansion will be drawn in 2D coordinates according to order, where the longitudinal axis shows the reachability- distance, and the horizontal axis representing the expansion order sequence (also called clustering sequence). With these, the reachability graph can be obtained. Each tag in the reachability graph represents a cluster. B. Rochau, et.al. [7] and others use OPTICS to observe a massive cluster. Torrecilla et.al. recognized the plankton in a remote sensing image that provided a potential tool to monitor the spatial and temporal variations of marine biodiversity [8].

In this paper, we propose an improved Harris corner detection method to detect crowd feature corners, then we used the coordinates matrix of those feature corners to do clustering analysis. Finally, we can get the information of crowd density by using a reachability graph for machine learning.

The paper is organized as follows: in Section II, we introduce the basic theories of crowd monitoring, the latest methods about crowd density determination, and then we compare the two methods presented in different papers that both use corner detection to analyze crowd density. In Section III, we introduce the improved Harris algorithm and the cluster analysis method of crowd feature corners as well as present the experimental results In Section IV, we summarize the whole work of this paper and present our conclusion.

II Crowd Density Estimation based on Density Clustering

The whole crowd density estimation flow chart based on density clustering algorithm is shown in Fig. 1. Our method utilizes the improved Harris algorithm for image feature extraction, OPTICS algorithm for density clustering, a reachability graph, and density clustering analysis method by machine learning for crowd density evaluation.

First of all, video data were obtained from the scene, then preliminary image processing was done such as increasing the contrast, reducing image noise, and getting the image gray level in each frame. The main process consists of improving Harris algorithm to calculate all the features of corners in the image and to save feature points into a matrix form. Intermediate process such as reachability graph is used by the OPTICS algorithm to generate clustering characteristic vectors by feature points analysis. In advanced processing, we used Support Vector Machines (SVM) to learn the data characteristics which is a means to train our system.. Finally, we quantified the crowd density and we divided the results into superior or high and low crowd densities. The overall framework is shown in Fig. 1.

A. Harris Algorithm and Its Improved Algorithm

There are different kinds of quantifiable characteristics in a crowd image, such as texture, pixels, etc.. Angular point characteristics detection is an effective method to

distinguish the crowd stage with additional features. Harris algorithm was presented in 1988 [10]. Now the image corner detection method has two types, namely: extraction of the image edge which is in the form of chain code where the maximum curvature point at the angular point is extracted; the other utilizes the curvature distribution in a gray-scale image, from which the maximum curvature point is also obtained and assigned as the angular point [12]. Harris corner algorithm belongs to the second type. Comparing with the former type, the calculation and the difficulty of Harris algorithm is significantly reduced.

For an image, the angular point is related to the curvature characteristics of the autocorrelation function. Autocorrelation function describes the changes of local gray scale image .The autocorrelation function is given by:

$$E(x,y)=\sum_{u,v} w_{u,v}\,|\,I_{x+u,y+v}-I_{u,v}\,|^2=[x \quad y]M\begin{bmatrix}x\\y\end{bmatrix} \quad (1)$$

The matrix M for approximate Hessian matrix, is:

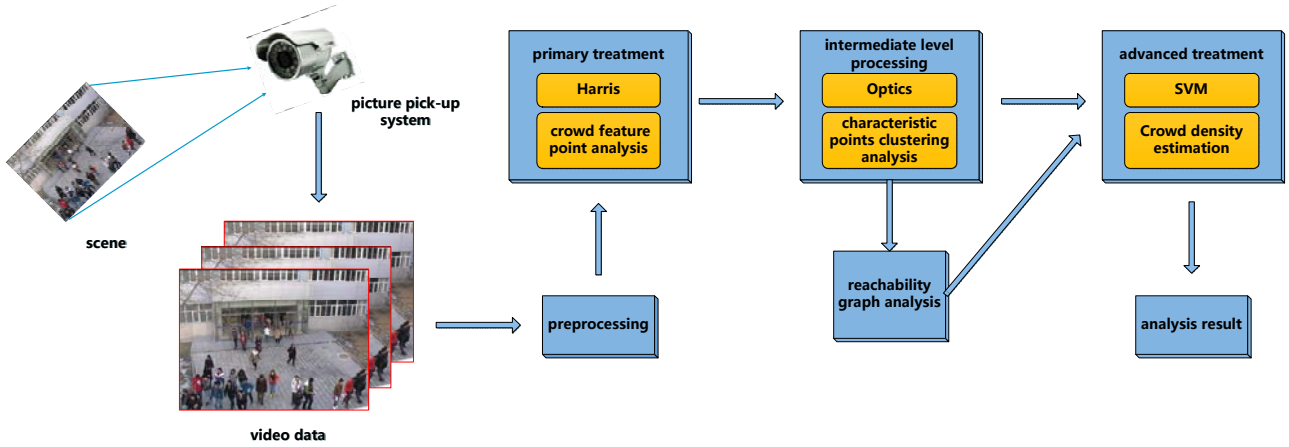


Figure 1. The whole crowd density estimation based on density clustering algorithm

（1）

$$M=\begin{bmatrix}A(x,y) & C(x,y)\\ C(x,y) & B(x,y)\end{bmatrix} \quad (2)$$

where:

$$A(x,y)=I_X{}^2\otimes G(x,y) \quad (3)$$

$$B(x,y)=I_Y{}^2\otimes G(x,y) \quad (4)$$

$$C(x,y)=I_{XY}\otimes G(x,y) \quad (5)$$

Harris algorithm R (x, y) is defined as:

$$R(x,y)=\det(M)-k*(trace(M))^2 \quad (6)$$

Harris corner detection involves a simple matrix with the first derivative calculation, and it is based on the threshold to extract the local so-called ‘interest points’.

$$\det(M)=AB-C^2 \quad (7)$$

$$trace(M)=(A+B) \quad (8)$$

The parameter k is usually between 0.04 and 0.06. When $R(x, y)$ is more than a given threshold, it will treat that point as the angle of the image point.

Using Harris algorithm for corner detection will have some main drawbacks. First, this does not have scale invariance features. Second, this is a pixel-based algorithm for extracting angular points. Finally, its testing time is too long [13].

Harris algorithm is based on Gaussian convolution function; it uses linear differential equation to extract feature points per frame. The corner detection is based on the gradient changes in the gray image. Harris corners are often those with higher changes in gray value as well as in the gradient. The similarity concept for image pixels is described by the difference between the gray values of a center pixel to 8 neighboring pixels in a detection window as shown in Fig. 2. As shown in this figure, it is obvious that the gray value of the center pixel is less than that of its neighboring pixels.

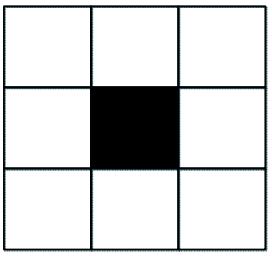


Figure 2. Detection window showing a center pixel of different gray value as compared to its surrounding 8 pixels

If the absolute values of the difference between the center pixel *I(i,j)* and the neighboring values is in the range of *t* (where we preset *t=20*), we can consider that such pixel is similar to the center pixel. At the same time, the similar pixel number of *I(i,j)* will increase in similarity.

After the 8 neighboring pixels of the center pixel *I(i,j)* are operated, we can get the similarity number *m* of the center pixel with the neighboring pixels using (9).

$$m = \sum R(i+x, j+y) \quad (-1 \le x \le 1, -1 \le y \le 1, \text{且} x \ne 0, y \ne 0) \tag{9}$$

where:

$$R(i+x, j+y) = \begin{cases} 1, \Delta(i+x, j+y) \le t \\ 0, \Delta(i+x, j+y) > t \end{cases} \tag{10}$$

According to the value of m, if *m=8*, this indicates that the 8 neighboring pixels are the similarity pixels around the center pixel; if *m=0*, this indicates that there is no neighboring pixels similar to the center pixel. This center pixel is then regarded as a "noise point" or an "isolated point". If *m* follows the range *1≤ m ≤ 7*, then this pixel can be regarded as a candidate pixel for a corner point. Finally, we use CRF to judge whether the center pixel is a corner and likewise an angular point.

Figure 3. The improved Harris algorithm used in a crowd image

This process not only increases the number of corners and their detection accuracy, but most importantly reduces the runtime thereby improving Harris algorithm. A comparison between the runtimes of the improved and original Harris algorithm is shown inFig. 4 for a series of 100 randomly selected images from the 2009 IEEE Performance Evaluation of Tracking and Surveillance (PETS2009) pedestrian dataset. We utilized MatLab 2012a to perform the experiment.

The original and the improved Harris algorithms are adopted to detect feature points. For every image runtime, it can be seen that the original Harris algorithm takes an average of about 7s for each image, while the improved method takes only about 3s each. This shows that our method is computationally faster .

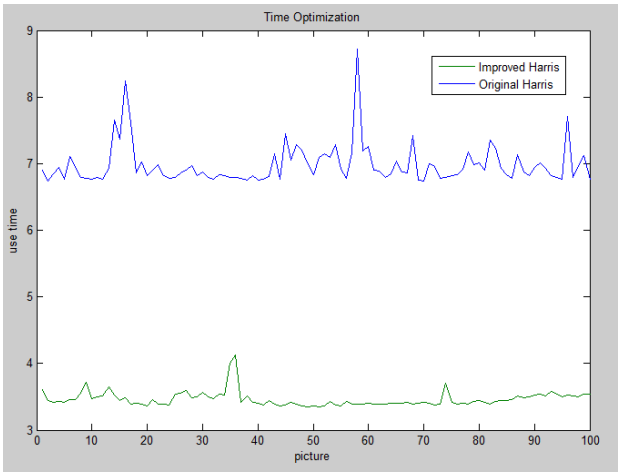


Figure 4. Original and improved Harris algorithms' corresponding runtimes

B. Feature Point Density Clustering Analysis

In this part, we treated the feature points as objects to do clustering analysis after image feature extraction. We then combined the density with the clustering analysis results, as presented in Fig. 5. The OPTICS algorithm based on density just sorts object datasets and outputs an ordered list of them in the form of clusters. As shown in Fig. 5, cluster-ordering cannot only distinguish the three groups *A*, *B*, and *C* but can also extract the sub-groups *C1*, *C2*, and *C3*. Compared with the traditional clustering algorithms, OPTICS algorithm is insensitive to the input parameters [14].

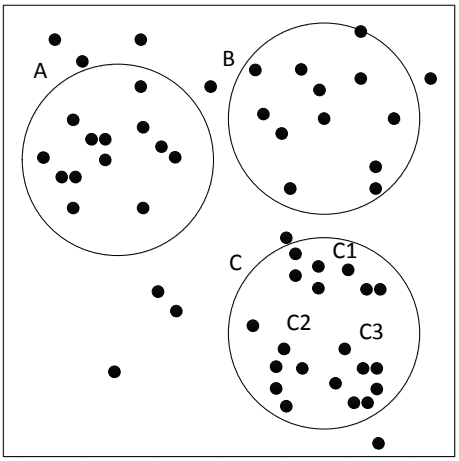


Figure 5. Different density clustering algorithm

The core distance and reachability-distance of OPTICS calculation formula is described as follows: [15]

Object *p* is the core distance:

$$\text{core\_}dis\tan ce_{\varepsilon,MinPts}(p) = \begin{cases} undefined, & \text{if } |N_\varepsilon(p)| < MinPts \\ \text{distance to the MinPtsh point}, & \text{otherwise} \end{cases} \tag{11}$$

Reachability-distance:

$$\text{reachability\_}dis\tan ce_{\varepsilon,MinPts}(p,o)= \quad (12)$$
$$\begin{cases} undefined, & \text{if } |N_{\varepsilon}(o)| < MinPts \\ \max(\text{core\_}dis\tan ce_{\varepsilon,MinPts}(o), dis\tan ce_{\varepsilon,MinPts}(p,o)), & \text{otherwise} \end{cases}$$

The reachability – distance between objects q and p refers to the larger Euclidean core distance If p is not the core object, then the reachability-distance between p and q is meaningless.

To predict the population in the video through feature points in the crowd, we applied the clustering algorithm by constructing the crowd eigenvector. We then used the SVM model which was previously trained by the crowd density method presented in [16,17] to assess the current population density .

III Experiment and Analysis

We used PETS2009 pedestrian datasets and the Beijing Union University campus crowd density datasets. The PETS2009 datasets are single camera shot using a fixed down-inclined video. Some of the aims of using these datasets are to estimate the population's density and quantity, to do target tracking in the crowd, and to perform crowd behavior analysis.

The test results are shown in Figs. 6to 9. Among them, *A1 – A4*, and *B1 – B4* are from the self-shooting data; while *C1 – C4*, and *D1 – D4* are from PETS2009 data. Figs. 6 and 8 were obtained using different methods under low density; whereas Figs. 7 and 9 were obtained under high density.

In Figs. 6 – 9, *A1*, *B1*, *C1*, and *D1* are the test crowd images; *A2*, *B2*, *C2*, and *D2* are the processing results of the improved Harris algorithm;*A3*, *B3*, *C3*, and *D3* are the processing results of angular point coordinate vector analysis; and *A4*, *B4*, *C4*, and *D4* are the processing results of OPTICS clustering showing the reachability graphs. .

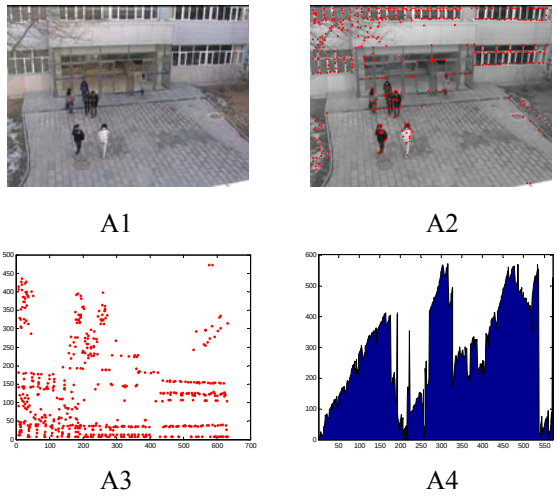


Figure 6. Self-shooting datum with low crowd density (A1) and the analysis results from the improved Harris algorithm (A2), angular point coordinate vector (A3), and OPTICS clustering (A4)

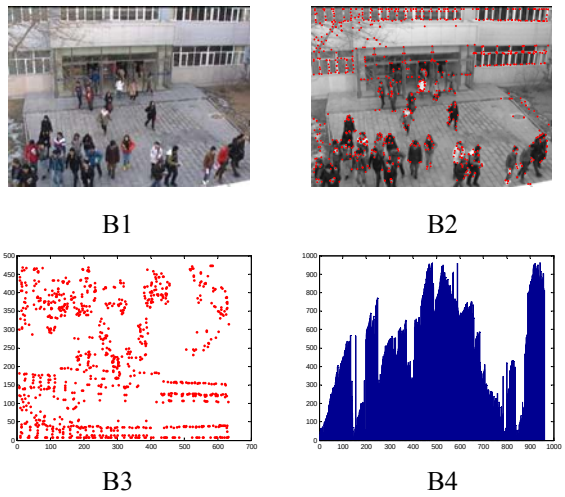


Figure 7. Self-shooting datum with high crowd density (B1) and the analysis results from the improved Harris algorithm (B2), angular point coordinate vector (B3), and OPTICS clustering (B4)

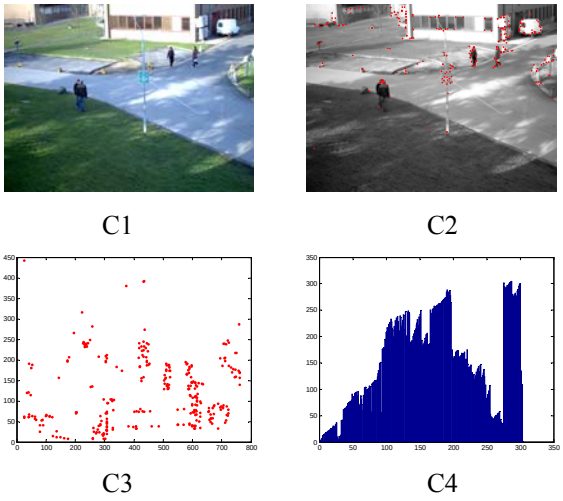


Figure 8. PETS2009 datum with low crowd density (C1) and the analysis results from the improved Harris algorithm (C2), angular point coordinate vector (C3), and OPTICS clustering (C4)

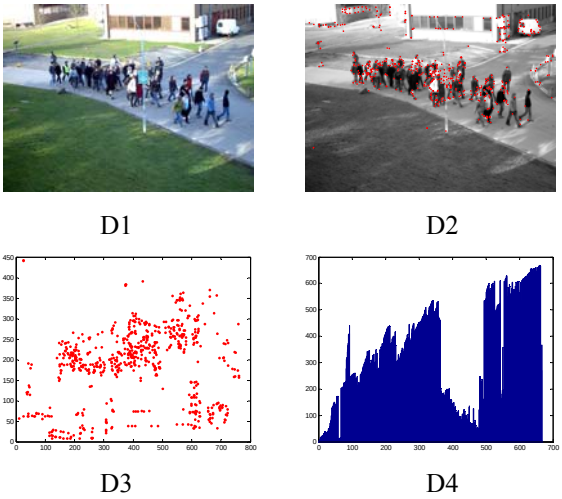


Figure 9. PETS2009 datum with high crowd density (D1) and the analysis results from the improved Harris algorithm (D2), angular point coordinate vector (D3), and OPTICS clustering (D4)

It can be seen from the reachability graphs of Figs. 6 - 9 that high density crowd clustering has obvious feature points whose number exceeds that of the low crowd density.

The four data groups: *A*, *B*, *C*, and *D* were taken in an ideal environment. Here the camera parameters are fixed, and they are almost without obstructions. But if the camera is far from the crowd, it will be subject to several obstructions or interferences such as buildings, trees, etc.. The effect of these are shown in Figs. 10 and 11. It is almost impossible to infer the people's density condition in the angular point coordinate vectors of *E3* and *F3*; but after using OPTICS particularly in *E4*, the reachability graph in *F4* revealed the gathering of a large crowd.

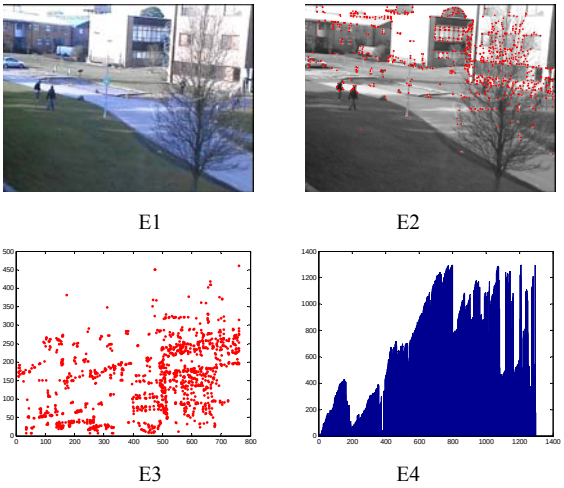


Figure 10. PETS2009 datum with low crowd density with obstacles (E1) and the analysis results from the improved Harris algorithm (E2), angular point coordinate vector (E3), and OPTICS clustering (E4)

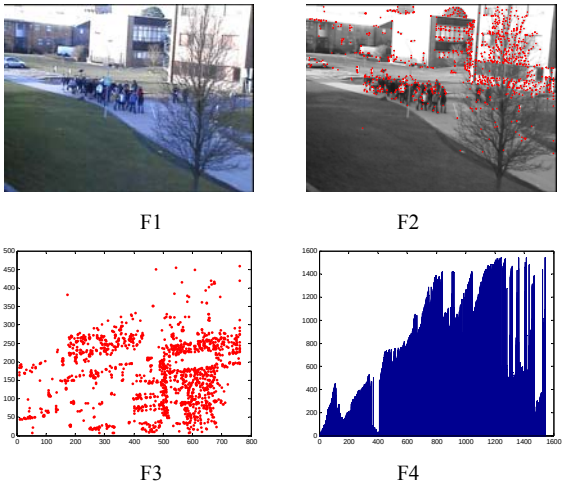


Figure 11. PETS2009 datum with high crowd density with obstacles (F1) and the analysis results from the improved Harris algorithm (F2), angular point coordinate vector (F3), and OPTICS clustering (F4)

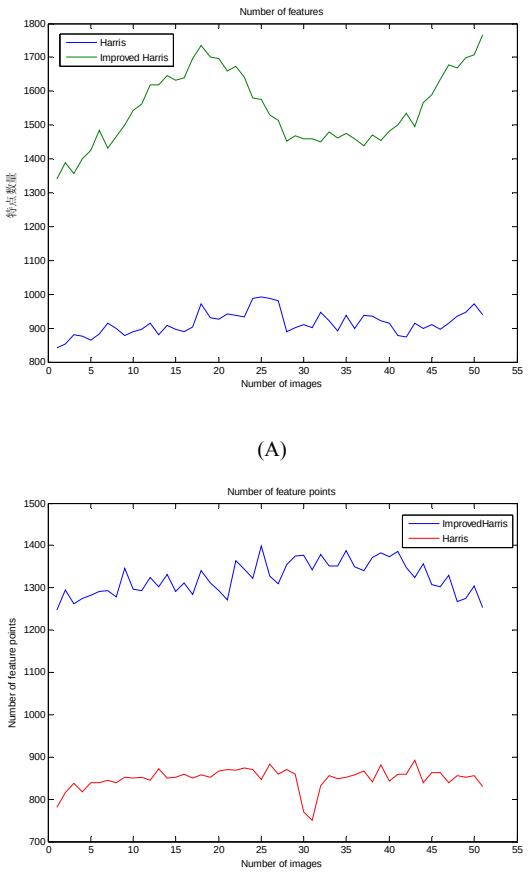


Figure 12. Different density population number of feature points more

We selected 50 pictures of low and high crowd densities from the PTES2009 standard datasets to do the experiments. We found that the characteristics of the improved Harris corner algorithm is significantly better than the original algorithm as shown in Fig. 12, where low crowd density is seen on the left and the high density

on the right side. It can be shown that due to the position of the cameras, the improvement of feature point detection is larger in a low crowd density.

Harris algorithm's speed is significantly improved at around 1.7 times faster than the original . We freely chose low and high crowd density images from PTES2009 datasets and video datasets to compare the performance of the different methods as shown in Table II.

From Table II, we can conclude that when the crowd is far away from the camera, the clustering effect is not apparent, and that the processing result of angular point method is not good. Obstacles within the camera's periphery, as in the case of a non-ideal environment,cause severe detection errors The detection accuracy of the above method will be high if the distance is moderate.. Furthermore, a moderate situation has an environment that has no buildings nor clutter interferences. To compare the different kinds of methods mentioned in this paper, we did the experiment using the PETS2009 dataset. The results are shown in Tables I and II. The improved algorithm proved to be faster than the original algorithm as shown by the shorter runtime in table I; and it has a relatively higher feature point extraction capability as shown in Table II. . Through the experimental results of crowd density estimation in Table III , we can see that the improved algorithm is feasible.

TABLE I.
COMPARISON OF THE RUNTIMES OF THE ORIGINAL AND IMPROVED HARRIS ALGORITHMS

Test Data	Crowd Density	Original Algorithm Run time (average)	Improved Algorithm Runtime (average)	Runtime Difference (average)
PETS 2009 benchmark data set 1 (200images)	low density	6.8s	3.1s	3.7s
PETS 2009 benchmark data set 2 (200images)	high density	7.2s	3.6s	3.6s
Capturing video data set 1 (2000images)	low density	6.5s	2.7s	3.8s
Capturing video data set 2 (2000images)	high density	6.9s	3.3s	3.6s

From Table III, we can see that the overall estimation accuracy is generally higher for the improved method than the original. It can be seen that this has a higher accuracy rate for the different testing sample density, number of images and forecast number. . Based from these results we can therefore say that the proposed method can be used to estimate the crowd density stage of the rest of the video sequence in the PETS2009 database .

The results in this section infer the following:

(1) The use of a fixed camera to get the images yields good test results. The method also works well without too many corners, obstacles and clutter interferences. However, its prediction error on benchmark datasets with obstacle situations is larger, because the image data characteristics are less and that the overall accuracy of the crowd density estimation is very ideal.

(2) The proposed method is not sensitive to the lighting condition of the environment. As long as there is no overexposure, illumination does not affect its results.

(3) The position of the camera is very important. It will impact and cause some errors in the detection results.

TABLE II.
COMPARISON OF THE EFFECTIVE FEATURE POINTS OF THE ORIGINAL AND IMPROVED HARRIS ALGORITHMS

Test Data	Original Algorithm Effective Feature Points (average)	Improved Algorithm Feature points (average)	Effective feature point difference (average)
PETS 2009 benchmark data sets (S1_L1 200 images)	79%	82%	3%
PETS 2009 benchmark data sets （S1_L2 200 images）	78%	83%	5%
PETS 2009 benchmark data sets （S1_L3 200 images）	76%	80%	4%
PETS 2009 benchmark data sets （S2_L1 200 images）	79%	80%	1%
PETS 2009 benchmark data sets （S2_L2 200 images）	78%	82%	4%
PETS 2009 benchmark data sets （S2_L3 200 images）	78%	80%	2%
PETS 2009 benchmark data sets （S3_MF 200 images）	80%	81%	1%
PETS 2009 benchmark data sets （S3_MF 200 images）	79%	82%	3%

TABLE III.
CROWD DENSITY ESTIMATION RESULTS

Data sources	Testing samples	Number of images	Forecast number	Original Harris Algorithm Feature Points		Improved Harris Algorithm Feature Points		
				Correct identification number	Accuracy rate	Correct identification number	Accuracy rate	Accuracy rate Difference
Capturing video data set 1	low density	10000	2500	2450	98%	2455	98.2%	0.2%
	high density	10000	2500	2403	96%	2451	98%	2%
Capturing video data set 1	low density	10000	2500	2427	97%	2432	97.3%	0.3%
	high density	10000	2500	2398	96%	2419	97%	1%
PETS 2009 benchmark data set 1	low density	750	150	142	94.7%	145	96.7%	2%
	high density	310	90	85	94.4%	87	96.7%	2.3%
PETS 2009 benchmark data set 2	low density	250	50	47	94%	48	96%	2%
	high density	630	190	175	92.1%	180	94.7%	2.6%

III CONCLUSION

With the development of modern social economy and the speeding up of urbanization, the density of people in a strategic place at a given time has increased greatly over the years. This however can lead to some untoward accidents that need to be prevented, hence there is a need to determine the crowd density and do the necessary management. We have used the medium density clustering method toextract feature points on a crowd video. We have also done data mining in a fixed background scenario for density estimation. The improved Harris algorithm predicts the crowd density by the number of the detected feature points. This further has a reduced runtime and a lower prediction error as compared to the other two methods that have been described also in this paper. This inference is based from the test results that utilize both the PETS2009 dataset and the images in the self-shooting data. The experimental results showed that the overall accuracy of our proposed method is satisfactory, and it is applicable in predicting different population densities most especially for extremely populated areas.

Further testing of our algorithm in a different environmental weather condition shall be one of the foci of our future research. We shall also aim at improving the robustness of our method and further increase its speed.

ACKNOWLEDGMENT

This research was supported by the Project of Construction of Innovative Teams and Teacher Career Development for Universities and Colleges under Beijing Municipality (CIT&TCD20130513), and the National Natural Science Foundation of China (Grant nos.. 61271370 and61271369), and the project no. 201411417SJ053. We thank PETS for sharing its source images, and the reviewers for their insightful comments. We also would like to thank associate professor Xu Guangmeis and He Ning whose guidance and help have been critical in the writing phases of this paper.

REFERENCES

[1] D. Kim, Crowd Density Estimation Using Multi-class Adaboost. Advanced Video and Signal-Based Surveillance (AVSS), 2012 IEEE Ninth International Conference. Page(s):447 - 451

[2] W. Bo, H. Bao, and S. Yang, Crowd Density Estimation Based on Texture Feature Extraction, Journal of Multimedia, 2013, 8(4): 331-337.

[3] V. Verona, and A. N. Marana, Wavelet packet analysis for crowd density estimationl Prnc. the listed International Symposia on Applied Information, Innsbmck,Austria: Aeta Press, 2001: 535—540.

[4] J. Junior, J.C.S., S.R. Musse, and C.R. Jung, Crowd Analysis Using Computer Vision Techniques, IEEE Signal Processing Magazine, 2010,27(5):66–77.

[5] G.T. Bei S.Y. Kwak,, and H.R. Byun, . Motion pattern analysis using partial trajectories for abnormal movement detection in crowded scenes, Electronics Letters, 2013, 49(3): 186-187.

[6] Y. L. Hou, and G.K.H. Pang, People counting and human detection in a challenging situation, IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems & Humans, 2011, 41(1):24-33.

[7] B. Rochau, W. Brandner, and A. Stolte, A benchmark for multiconjugated adaptive optics: VLT-MAD observations of the young massive cluster Trumpler14, Monthly notices of the Royal Astronomical Society, 2011, 418(2): 949-959.

[8] E.Torrecilla, D. Stramski, and R.A. Reynolds, Cluster analysis of hyperspectral optical data for discriminating phytoplankton pigment assemblages in the open ocean, Remote Sensing of Environment, 2011, 115(10): 2578-2593.

[9] A. Albiol, M.J.Silla,and J. M. Mossi, Video Analysis using Corner Motion Statistics, Proceedings of the IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, 2009: 31-38.

[10] D. Conte., P. Foggia, G. Percannella, F. Tufano, and M. Vento, A method for counting people in crowded scenes, Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2010 pp. 225 – 232.

[11] D. Conte, P. Foggia, G. Percannella, and M. Vento, Counting moving persons in crowded scenes, Machine

Vision and Applications, July 2013, Volume 24, Issue 5, pp 1029-1042.
[12] J. Wright, A. Yang, S. Sastry S, and Y. Ma, Robust face recognition via sparse representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2):210−227.
[13] K. Andrea,and T. Sziranyi, Harris function based active contour external force for image segmentation, Pattern recognition letters, 2012, 33(9): 1180-1187.
[14] B. Zhou, F. Zhang, and L. Peng, Higher - order SVD analysis for crowd density estimation,. Computer vision and image understanding, 2012, 116(9): 1014-1021.
[15] X.F.Wang, and D.S. Huang, A Novel Density-Based Clustering Framework by Using Level Set Method, IEEE Transactions on Knowledge and Data Engineering, 2009, 21(11):1515-1531.
[16] I. Ali, and M. Dailey, Multiple human tracking in high-density crowds, Image and vision computing, 2012, 30(12): 966-977.
[17] Z. Liu, and Z. Wei, Image Classification Optimization Algorithm based on SVM, Journal of Multimedia, 2013, 8(5): 496-502.
[18] J.Li, and W. Gong, Real Time Pedestrian Tracking using Thermal Infrared Imagery, Journal of Computers 5.10 (2010): 1606-1613.
[19] W. Li, Y. Lin, B. Fu, M. Sun and W. Wu, Cascade Classifier Using Combination of Histograms of Oriented Gradients for Rapid Pedestrian Detection, Journal of Software 8.1 (2013): 71-77.
[20] PETS 2009 Eleventh IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. Retrieved from: http://www.cvg.rdg.ac.uk/PETS2009/ Retrieved on: February 24, 2014.

Optimal Resource Allocation Scheme for Satisfying the Data Rate Requirement in Hybrid Network of D2D-Cellular

Wenwen Liu, Yang Yang, Tao Peng, Wenbo Wang
Wireless Signal Processing and Network Lab, Key Laboratory of Universal Wireless Communication Ministry of Education, Beijing University of Posts & Telecommunications
Email: {wenwl67, buptyy1015}@gmail.com

***Abstract*—In this paper, an optimal wireless resources allocation scheme is put forward to improve the performance for both cellular and Device-to-Device (D2D) users in a hybrid network. The resource unit and power allocation can be modeled into a sum-rate optimization problem with the constraint of power limitation and rate ratio. The rate ratio is used to measure the data rate requirements from both cellular and D2D users. By using the Lagrange Multiplier Method, we derived out the optimal formula about allocating the power and resource unit for both cellular and D2D users. According to the optimal formula, the relationship between the transmission power and resource unit allocation is mutually conditional. Then the resource allocation scheme algorithm was given. And the power allocation can be resolved by the water-filling algorithm. Simulation results show that the proposed scheme improves the sum-throughput significantly as well as the fairness.**

***Index Terms*—hybrid network, Device-to-Device (D2D) communication, resource allocation, power control, fairness**

I. INTRODUCTION

In a traditional cellular network, the communication is always forwarded by the central node (e.g. Base Station, BS). The communication under the centralized control has benefit in resource allocation and interference management, but it would lead to low wireless resource utilization. In order to resolve this problem, a new hybrid wireless network is proposed. This hybrid network is built up by placing a certain number of D2D (Device-to-Device) terminal users in the traditional wireless communication network (e.g. LTE/LTE-A cellular wireless network). The internal communication between D2D terminal users can be implemented by using the cognitive radio technology. There are two communication models for the UE (User Equipment) in the hybrid network: (1) the cellular mode, in which UE is connected to another UE via the BS; (2) the D2D model, in which source UE can communicate with the destination UE directly. Studies have proved that: by putting the D2D communication model into the traditional cellular network, the data rate and capacity of the network can be increased, while the transmit power is reduced, and then it is easier to balance the load and coverage of the network [1-3].

At present, the mobile operators and equipment manufacturers are trying to explore a way to introduce the D2D function in cellular networks, especially the LTE-A system. In the 3GPP RAN#52 meeting and SA1#55 meeting which are held in 2011 June and August, Qualcomm Corp has already put forward proposals about doing some D2D researches in the LTE-A system. The literature [4] presented the conception of integrating the D2D communication in the LTE-A system. The literature [5] introduces a FlashLinQ technology, which can make a device to detect other devices nearby and then communicates with them directly via the air interface. Different from the normal D2D technology, FlashLinQ can be used in the authorized frequency, which is more efficient. Introducing the D2D function in cellular networks can bring two main benefits. First, the D2D approach can improve the spectral efficiency of wireless transmission. When the two terminals are close, the channel quality of the direct link is always better than when the two terminals connect to base station respectively. And the direct communication only takes up one slot of radio resources, otherwise two slots are required to transmit data indirectly through the base station. Second, by using the D2D function, it would be much more convenient for a mobile device to find out other devices around it. In the traditional cellular network, if a mobile device wants to do this job, it should locate the other device through the network operator, search the information in the operator's database, and then get the result sent back from the operator. In this way, it does not only have a slow speed, but also has a poor scalability, which cannot support a large number of requests like this. If the D2D function is implemented, the mobile device can detect each other directly by transmitting and monitoring the broadcast packet, which makes it much more fast and flexible.

Resource scheduling is a key problem in the MAC layer of wireless network. It directly affects the performance of the whole wireless network system. The resource allocation depends on many factors: network

Manuscript received December 18, 2013;.
Corresponding author: Wenwen Liu (wenwl67@gmail.com) .

doi:10.4304/jcp.9.5.1218-1225

architecture, power consumption, the allocation of spectrum resource blocks and radio interference, which are all mutually dependent and interacting.

Recent studies have shown that there are two important aspects in allocating the network resources: spectrum resources scheduling and power control. For spectrum resource scheduling, there are two theoretical models: (1) packet-based model [6,7], the scheduling granularity is based on the executing frequency of scheduling algorithm; (2) flow-based model [8-10], the scheduling granularity is based on the frequency of service data arrival and processing. For the power control, it has two algorithms: (1) the centralized control algorithm in [11-13] have shown that it can get a global optimal solution, but it requires the global information and longer processing time, which means it can't satisfy the real wireless environment; (2) the distributed algorithm has shown in [14,15] that it can be fast convergence to a optimal solution based on the local partial information, which is suitable for the real environment.

In this paper, we assume the independent communication between the D2D users is implemented by using the uplink resource in the cellular wireless network which is controlled by the base station. For the uplink resource allocation, there are some issues in the hybrid wireless network:

a)In cellular model, such as the traditional cellular network of OFDM (Orthogonal Frequency Division Multiplexing), the UEs use the orthogonal uplink spectrum resources to transfer data, so there is no interference between them when using the uplink spectrum resource.
In hybrid network, the D2D UEs can reuse the uplink spectrum resource to communicate. It would have interference when the D2D UEs are trying to communicate with cellular UEs by using the same spectrum resource, so it is under an interference-limited constraint.

b)In hybrid network, the mutual interference between UEs would lead to a complex wireless link environment. Once a wireless resource is allocated to a new UE, the status of wireless channel would be changed. So the channel gain would be a time-varying parameter, which is different from the traditional cellular model.

c)In traditional cellular model, in order to satisfy the data rate requirement, the pre-set rate ratio of each UE is used to satisfy the proportional fairness criterion.
In hybrid network, the D2D UEs is the secondary users comparing with cellular UEs according to the theory of traditional cognitive radio technology. Indeed, the D2D UEs still have to pay for reusing the cellular resource, why they can't deserve the corresponding service? So the weight value of data rate of D2D UEs is taken into account in this paper.

d)In traditional cellular model, the control central (BS) knows the information of uplink resource scheduling in advance.
In hybrid network, the uplink resource scheduling information of cellular network should be used in the communication of D2D users, which is also controlled by the cellular BS in order to avoid the interference from the cellular UE.

According to the above analysis, the cellular and D2D transmission can be formed as a sum-rate optimization problem. The target is to schedule the frequency resources allocation and transmission power for cellular and D2D users, and then gain the maximum data sum-rate. We first prove that the optimization problem is convex and then solve it by constructing Lagrange function and KKT conditions. Finally we can get the conclusion that: (1) the relationship between the transmission power and resource unit allocation is mutually conditional; (2) the power allocation can be resolved by the water-filling algorithm. Based on the former conclusion, we propose a hybrid scheduling algorithm for the dynamic cellular and D2D accessing process and do some comprehensive analyzes on it. The system simulation results can prove the above conclusions in hybrid network. The contribution of this paper is as follows: (1) the scheduling algorithm of spectrum resource allocation and power control can be used for both cellular and D2D users; (2) transferring the interference-limited system into the noise-limited system can make the sum-rate optimization problem to be solved, it also conforms the idea of making the maximum use of the noise tolerance of cellular users; (3) the weight value of D2D user data rate is set up to ensure fairness in the D2D users.

The rest of this paper is organized as follows. Section II describes the scenario description and system model. In section III, the optimal uplink resource allocation for the D2D and cellular users are derived and proved in details, and then uplink resource scheduling algorithm is provided based on the former conclusion. The system simulation results of the proposed scheme in section III are analyzed in section IV. Finally the conclusions are summarized in section V.

II. Scenario Description and System Model

In this section, we present the scenario description of the hybrid wireless network, which defines the model of hybrid network and optimization constraints.

A. Scenario Description

The basic scenario is a hybrid network system, which contains a cellular system and D2D pairs. As shown in Fig. 1, there are some users need to be served in the hybrid network, and they are all able to transmit data in both cellular mode and D2D mode. We make the following definition and assumption:

a)BS could receive the CSI (Channel Signal Information) immediately from the mobile terminals which include cellular and D2D terminal users;

b)By sensing on the spectrum, the D2D terminal could detect several resource blocks (RBs) which are idle to be used, and then report the number of RBs to the BS;

c)The spectrum resource allocation and transmission power of D2D user is also controlled by BS as the cellular users;

d)D2D user can reuse the uplink resource of cellular network, and we assume that only one D2D pair is allowed to reuse the spectrum resource of one cellular user to communicate in the same time slot.

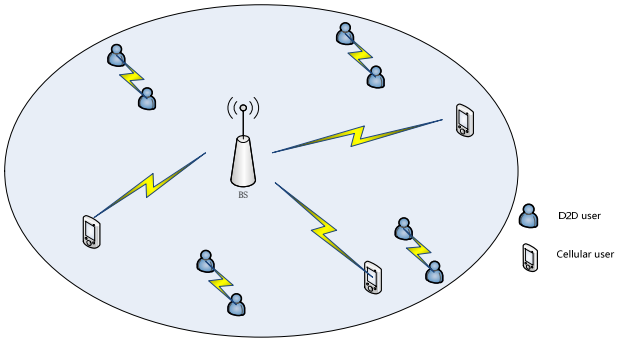


Figure1. Hybrid network

B. System Model

We assume that: in one cell, there are M activation cellular users, N activation D2D pairs, K RBs can be used to communication. The channel gain and transmission power of m-th cellular user and n-th D2D pair on the k-th RB is denoted as $h_{mk}, p_{mk}, h_{nk}, p_{nk}$, and the density of noise power is N_0, the spectrum bandwidth of each RB is ω. The sum of available rate that of the m-th cellular user is $R_{cm}=\sum_{k=1,2,...,K}\rho_{mk}\gamma_{mk}$, while the sum of data rate of the n-th D2D pair is $R_{dn}=\sum_{k=1,2,...,K}\rho_{nk}\gamma_{nk}$, where ρ_{mk} and ρ_{nk} is the indicating factor of the m-th cellular user and the n-th D2D pair use the k-th RB to communicate respectively, the value of ρ_{mk} and ρ_{nk} is 0 or 1, γ_{mk} and γ_{nk} is the data rate of the m-th cellular user and n-th D2D pair use the k-th RB to communicate respectively as follows:

$$\gamma_{mk}=\omega\log_2(1+\frac{p_{mk}|h_{mk}|^2}{N_0\omega+\sum_{n=1}^{N}\rho_{nk}p_{nk}|h_{nk}|^2}) \quad (1)$$

$$\gamma_{nk}=\omega\log_2(1+\frac{p_{nk}|h_{nk}|^2}{N_0\omega+\sum_{m=1}^{M}\rho_{mk}p_{mk}|h_{mk}|^2}) \quad (2)$$

Let $I_1=N_0\omega+\sum_{n=1}^{N}\rho_{nk}p_{nk}|h_{nk}|^2$ denote the sum of background noise and the interference from D2D user n to cellular user m which are using the same RB k to communication at the same time, and let $I_2=N_0\omega+\sum_{m=1}^{M}\rho_{mk}p_{mk}|h_{mk}|^2$ denote the sum of background noise and the interference from cellular user m to D2D user n which are using the same RB k to communication at the same time. So the SNR of the cellular user m which uses RB k to communication is $SINR_{mk}=\frac{p_{mk}|h_{mk}|^2}{I_1}$, and the SNR of the D2D user n which uses RB k to communication is $SINR_{nk}=\frac{p_{nk}|h_{nk}|^2}{I_2}$. The maximum cellular transmission power is set as $p_{\max 1}$, and the D2D transmission power is set as $p_{\max 2}$.

Base on the above analysis, the maximum total rate of hybrid network and the optimization constraints are modeled as follows:

$$\begin{aligned}
&\max_{p_{mk},\rho_{mk},p_{nk},\rho_{nk}}(\sum_{m=1}^{M}R_{cm}+\sum_{n=1}^{N}R_{dn})\\
&=\max_{p_{mk},\rho_{mk},p_{nk},\rho_{nk}}[\sum_{m=1}^{M}\sum_{k=1}^{K}\rho_{mk}\omega\log_2(1+\frac{p_{mk}|h_{mk}|^2}{N_0\omega+\sum_{n=1}^{N}\rho_{nk}p_{nk}|h_{nk}|^2})+\sum_{n=1}^{N}\sum_{k=1}^{K}\rho_{nk}\omega\log_2(1+\frac{p_{nk}|h_{nk}|^2}{N_0\omega+\sum_{m=1}^{M}\rho_{mk}p_{mk}|h_{mk}|^2})]\\
&=\max_{p_{mk},\rho_{mk},p_{nk},\rho_{nk}}[\sum_{m=1}^{M}\sum_{k=1}^{K}\rho_{mk}\omega\log_2(1+\frac{p_{mk}|h_{mk}|^2}{I_1})+\sum_{n=1}^{N}\sum_{k=1}^{K}\rho_{nk}\omega\log_2(1+\frac{p_{nk}|h_{nk}|^2}{I_2})]\\
&s.t.\begin{cases}
\sum_{k=1}^{K}p_{mk}\le p_{\max 1},\\
\sum_{k=1}^{K}p_{nk}\le p_{\max 2},\\
\sum_{m=1}^{M}\rho_{mk}=1,\\
\sum_{n=1}^{N}\rho_{nk}=1,\\
p_{mk},\rho_{mk},p_{nk},\rho_{nk}\ge 0,\\
\{R_{c1}:R_{c2}:,\ldots,:R_{cM}:R_{d1}:R_{d2}:,\ldots,:R_{dN}\}=C\{\theta_{c1}:\theta_{c2}:,\ldots,:\theta_{cM}:\theta_{d1}:\theta_{d2},\ldots,:\theta_{dN}\}=C\{\theta_i\}.
\end{cases}
\end{aligned} \quad (3)$$

In this formulation, the first and second constraints denote the maximum power of cellular UEs and D2D UEs should be less than $p_{\max 1}$ and $p_{\max 2}$. The third and fourth constrains denote the cellular UEs use the resources orthogonally, and the D2D UEs reuse the resources orthogonally with each other. The fifth condition is obvious.

The last expression stands for the proportional fairness criterion, where $\{\theta_i\}=\{\theta_{c1},\theta_{c2},\ldots,\theta_{cM},\theta_{d1}\theta_{d2},\ldots,\theta_{dN}\}$, $i=c1,c2,\ldots,cM,d1,d2,\ldots,dM$ is the predefined proportional fairness factor which reflects the weight value of the cellular and D2D users to guarantee the fairness of hybrid network and QoS of the D2D users.

III. Optimal Resource Selection and Power Allocation

In this section, we derive the relationship of the transmission power and spectrum resource allocation of the hybrid network. In particular, we turn the interference-limited system into the noise-limited system to solve the optimization problem, which based on the several researches [16-18] on the D2D communication reuse the resource of cellular network. We can conclude that the results for RB allocation per time slot and transmission power of D2D users and cellular users are mutual restraint. After this, we give the specific scheduling algorithm.

A. Problem Formulation

From [17,18], the outstanding advantage of hybrid network is to improve overall network performance. Even though it scarifies the performance of cellular network, it greatly increases the wireless resource utilization. To maximize the hybrid network load and resource utilization, we suppose that the SNR of all the communication links can just meet to the customer satisfaction, which means we suppose the system model is based on the maximum noise. So we turn the interference-limited system into the noise-limited system, and the formula (3) turns to:

$$\max_{p_{mk},\rho_{mk},p_{nk},\rho_{nk}} (\sum_{m=1}^{M} R_{cm}+\sum_{n=1}^{N} R_{dn})$$

$$= \max_{p_{mk},\rho_{mk},p_{nk},\rho_{nk}} [\sum_{m=1}^{M}\sum_{k=1}^{K} \rho_{mk}\omega\log_2(1+\frac{p_{mk}|h_{mk}|^2}{I_{1\max}})+\sum_{n=1}^{N}\sum_{k=1}^{K} \rho_{nk}\omega\log_2(1+\frac{p_{nk}|h_{nk}|^2}{I_{2\max}})] \quad (4)$$

where the $I_{1\max}$ and $I_{2\max}$ is the maximum noise withstand the limiting conditions, which are defined as constant.

Obviously, it is neither a linear nor a standard convex problem. By [19], we can transform the problem to a convex problem and it can be expressed as

$$\min_{p_{mk},\rho_{mk},p_{nk},\rho_{nk}} -\left[\sum_{m=1}^{M}\sum_{k=1}^{K} \rho_{mk}\omega\log_2(1+\frac{p_{mk}|h_{mk}|^2}{I_{1\max}})+\sum_{n=1}^{N}\sum_{k=1}^{K} \rho_{nk}\omega\log_2(1+\frac{p_{nk}|h_{nk}|^2}{I_{2\max}})\right] \quad (5)$$

According to the Lagrangian Arithmetic, the objective function (5) and the constrictions in function (3) can be transformed into a Lagrange function as

$$L=\left[\sum_{k=1}^{K}\sum_{m=1}^{M} -\rho_{mk}\omega\log_2(1+\frac{p_{mk}|h_{mk}|^2}{I_{1\max}})-\sum_{k=1}^{K}\sum_{n=1}^{N} \rho_{nk}\omega\log_2(1+\frac{p_{nk}|h_{nk}|^2}{I_{2\max}})\right]+\sum_{m=1}^{M}\lambda_m(\sum_{k=1}^{K} p_{mk}-p_{\max 1})$$
$$+\sum_{n=1}^{N}\mu_n(\sum_{k=1}^{K} p_{nk}-p_{\max 2})+\sum_{k=1}^{K}\eta_k(\sum_{m=1}^{M}\rho_{mk}-1)+\sum_{k=1}^{K}\vartheta_k(\sum_{n=1}^{N}\rho_{nk}-1)+\sum_{m=1}^{M}\sum_{k=1}^{K}\phi_{mk}(-p_{mk})+\sum_{n=1}^{N}\sum_{k=1}^{K}\gamma_{nk}(-p_{nk})$$
$$+\sum_{m=1}^{M}\sum_{k=1}^{K}\xi_{mk}(-\rho_{mk})+\sum_{n=1}^{N}\sum_{k=1}^{K}\delta_{nk}(-\rho_{nk})+\sum_{i=1}^{M+N} D_i(\frac{R_i}{\theta_i}-C) \quad (6)$$

Where λ_m, μ_n, η_k, ϑ_k, ϕ_{mk}, γ_{nk}, ξ_{mk}, δ_{nk} and D_i is the Lagrange multiplier factor, C is a constant and defined as $C=\frac{R_i}{\theta_i}=\frac{R_{c1}}{\theta_{c1}}=\frac{R_{c2}}{\theta_{c2}}=\ldots=\frac{R_{cM}}{\theta_{cM}}=\frac{R_{d1}}{\theta_{d1}}=\frac{R_{d2}}{\theta_{d2}}\ldots=\frac{R_{dN}}{\theta_{dN}}$.

We derivative the p_{mk}, p_{nk}, ρ_{mk} and ρ_{nk} separately as

$$\frac{\partial L}{\partial p_{mk}}=-\rho_{mk}\frac{\omega}{\ln 2}\frac{|h_{mk}|^2}{I_{1\max}+p_{mk}|h_{mk}|^2}+\lambda_m-\phi_{mk}+\frac{D_i}{\theta_i}\frac{\omega}{\ln 2}\frac{|h_{mk}|^2}{I_{1\max}+p_{mk}|h_{mk}|^2} \quad (7)$$

$$\frac{\partial L}{\partial p_{nk}}=-\rho_{nk}\frac{\omega}{\ln 2}\frac{|h_{nk}|^2}{I_{2\max}+p_{nk}|h_{nk}|^2}+\mu_n-\gamma_{nk}+\frac{D_i}{\theta_i}\frac{\omega}{\ln 2}\frac{|h_{nk}|^2}{I_{2\max}+p_{nk}|h_{nk}|^2} \quad (8)$$

$$\frac{\partial L}{\partial \rho_{mk}}=-\omega\log_2(1+\frac{p_{mk}|h_{mk}|^2}{I_{1\max}})+\eta_k-\xi_{mk}+\frac{D_i}{\theta_i}\omega\log_2(1+\frac{p_{mk}|h_{mk}|^2}{I_{1\max}}) \quad (9)$$

$$\frac{\partial L}{\partial \rho_{nk}}=-\omega\log_2(1+\frac{p_{nk}|h_{nk}|^2}{I_{2\max}})+\vartheta_k-\delta_{nk}+\frac{D_i}{\theta_i}\omega\log_2(1+\frac{p_{nk}|h_{nk}|^2}{I_{2\max}}) \quad (10)$$

By using KKT constriction, we have $\lambda_m,\ \mu_n,\ \eta_k,\ \vartheta_k,\ \phi_{mk},\ \gamma_{nk},\ \xi_{mk},\ \delta_{nk},\ D_i \ge 0$, $p_{mk}\cdot\phi_{mk}=0$, $\lambda_m(\sum_{k=1}^{K} p_{mk}-p_{\max 1})=0$, $p_{nk}\cdot\gamma_{nk}=0$, $\mu_n(\sum_{k=1}^{K} p_{nk}-p_{\max 2})=0$, $\rho_{mk}\cdot\xi_{mk}=0$, $\rho_{nk}\cdot\delta_{nk}=0$, $\eta_k(\sum_{m=1}^{M}\rho_{mk}-1)=0$, $\vartheta_k(\sum_{n=1}^{N}\rho_{nk}-1)=0$. And then, we have analysis and discussion as follows:

(1) When $p_{mk}\neq 0$, the equation $\phi_{mk}=0$ holds,

$$\therefore \frac{\rho_{mk}}{\ln 2}(1-\frac{D_i}{\theta_i})\frac{\omega|h_{mk}|^2}{I_{1\max}+p_{mk}|h_{mk}|^2}\le\lambda_m,$$ and then the optimal RB selection k_m^* for cellular UE m is

$$k_m^*=\arg\max_{k=1,2,\ldots,K}\frac{1}{\ln 2}(1-\frac{D_i}{\theta_i})\frac{\omega|h_{mk}|^2}{I_{1\max}+p_{mk}|h_{mk}|^2}; \quad (11)$$

(2) When $p_{nk}\neq 0$, the equation $\gamma_{nk}=0$ holds,

$$\therefore \frac{\rho_{nk}}{\ln 2}(1-\frac{D_i}{\theta_i})\frac{\omega|h_{nk}|^2}{I_{2\max}+p_{nk}|h_{nk}|^2}\le\mu_n,$$ and then the optimal RB selection k_n^* for D2D UE n is

$$k_n^*=\arg\max_{k=1,2,\ldots,K}\frac{1}{\ln 2}(1-\frac{D_i}{\theta_i})\frac{\omega|h_{nk}|^2}{I_{2\max}+p_{nk}|h_{nk}|^2}; \quad (12)$$

(3) When $\rho_{mk}\neq 0$, the equation $\xi_{mk}=0$ holds,

$$\therefore (1-\frac{D_i}{\theta_i})\omega\log_2(1+\frac{p_{mk}|h_{mk}|^2}{I_{1\max}})\le\theta_k,$$ and then the optimal cellular UE allocation m_k^* for using RB k is

$$m_k^*=\arg\max_{m=1,2,\ldots,M}(1-\frac{D_i}{\theta_i})\omega\log_2(1+\frac{p_{mk}|h_{mk}|^2}{I_{1\max}})=\arg\max_{m=1,2,\ldots,M}(1-\frac{D_i}{\theta_i})r_{mk} \quad (13)$$

(4) When $\rho_{nk}\neq 0$, the equation $\delta_{nk}=0$ holds,

$$\therefore (1-\frac{D_i}{\theta_i})\omega\log_2(1+\frac{p_{nk}|h_{nk}|^2}{I_{2\max}})\le\vartheta_k,$$ and then the optimal D2D UE allocation n_k^* for using RB k is

$$n_k^*=\arg\max_{n=1,2,\ldots,N}(1-\frac{D_i}{\theta_i})\omega\log_2(1+\frac{p_{nk}|h_{nk}|^2}{I_{2\max}})=\arg\max_{n=1,2,\ldots,N}(1-\frac{D_i}{\theta_i})r_{nk} \quad (14)$$

Based on the above theoretical derivation, we can draw two conclusions as follows:

Theorem 1: The two problems: the allocation of the resource block for each UE (including cellular UE and D2D UE) and the optimal power controlling for UE using

the appropriate resource block to communicate, are mutually restraint.

Proof: From formula (13) and (14), the m_k^* is a function of p_{mk} and the n_k^* is a function of p_{nk}, which means for the k-th resource block, choosing which user (concluding cellular and D2D user) using the k-th resource block to communicate can be determined by the cellular users and the D2D users' power allocation.

From formula (11) and (12), the k_m^* is a function of p_{mk} and the k_n^* is a function of p_{nk}, which means for the m-th cellular UE and n-th D2D UE, choosing which resource block to communicate that can be determined by the power allocation of these UEs.

In particular, from the above formula $\frac{\rho_{mk}}{\ln 2}(1-\frac{D_i}{\theta_i})\frac{\omega|h_{mk}|^2}{I_{1\max}+p_{mk}|h_{mk}|^2} \le \lambda_m$ with equality, the optimal power p_{mk} of m-th cellular UE using the k-th RB to communicate is

$$p_{mk}=\frac{\left(\frac{1}{\ln 2}-\frac{r_{mk}}{\theta_{mk}}\right)\omega}{\lambda_m}-\frac{I_{\max 1}}{|h_{mk}|^2} \tag{15}$$

And also from the above formula $\frac{\rho_{nk}}{\ln 2}(1-\frac{D_i}{\theta_i})\frac{\omega|h_{nk}|^2}{I_{2\max}+p_{nk}|h_{nk}|^2} \le \mu_n$ with equality, we have the optimal power p_{nk} of n-th D2D UE using the k-th RB to communicate is

$$p_{nk}=\frac{\left(\frac{1}{\ln 2}-\frac{r_{nk}}{\theta_{nk}}\right)\omega}{\mu_n}-\frac{I_{\max 2}}{|h_{nk}|^2} \tag{16}$$

Thus we have shown that in order to determine the value of p_{mk} and p_{nk}, the resource allocation is also required as a prerequisite. □

Note that the formula (15) and (16) satisfy the water-filling algorithm form ($p=[\mu-\frac{1}{SNR}]_+$) in [18-20]. So we have the following corollary.

Corollary 1: In the hybrid wireless network, the water-filling power allocation algorithm can be used to realize the power allocation of cellular UEs and D2D UEs.

B. Algorithm of RB Selection and Power Allocation

We assume that the setting of resource blocks for selecting the whole cell is $\Omega=\{1,2,...,K\}$, the resource blocks setting of cellular UEs using is presented as $\Omega_m \subseteq \Omega$, and the resource blocks setting which D2D UEs have sensed for using is $\Omega_n \subseteq \Omega$.

The proposed algorithm is as follows:

Step 1: Initialized the resource blocks allocation for cellular and D2D users.

For cellular user m is assigned with the resource block which has the highest channel gain in the hybrid wireless network by base station; the resource block can be assigned to the user that with highest channel gain when the resource block is chosen by several users at the same time. Then add these block number to Ω_m.

For D2D user n, add the sensed resource blocks could be used to Ω_n.

Step 2: Calculate the rate and priority for each user.

For cellular user m, load the whole transmit power on the RB allocated in Ω_m, then calculate the rate R_{mk} and priority $Y_{mk}=\frac{R_{mk}}{\theta_{mk}}$.

For D2D user n, load the transmit power on the RB sensed in Ω_n, and then calculate the rate R_{nk} and priority $Y_{nk}=\frac{R_{nk}}{\theta_{nk}}$.

Step 3: The iterative process.

According to

$$x^*=\arg\min_{m=1,2,...,M;n=1,2,...,N}\{Y_{mk},Y_{nk}\}=\arg\min_{m=1,2,...,M;N=1,2,...,N}\{\frac{R_{mk}}{\theta_{mk}},\frac{R_{nk}}{\theta_{nk}}\}$$

, find the right terminal which has the highest priority.

If the right terminal x^* is a cellular user, m^*, $m=x$, $\Omega_m \neq \varnothing$:

(1) Choose the right RB k_{m^*} with highest channel gain to the cellular UE m^* according to $k_{m^*}=\arg\max_{k=1,2,...,K}|h_{mk}|$;

(2) Calculate the data rate $R_{m^*}=\sum_{k\in\Omega_m}\log_2(1+\frac{p_{m^*k}|h_{m^*k}|^2}{N_0\omega+\sum_{n=1}^{N}\rho_{nk}p_{nk}|h_{nk}|^2})$, where the p_{m^*k} is the transmit power of the m^*-th cellular user by using the water-filling algorithm;

(3) Update the priority of m^*-th cellular user Y_{m^*}, $Y_{m^*}=\frac{R_{m^*}}{\theta_{m^*}}$;

(4) Update the channel gain of the whole hybrid network, Ω_m, and Ω_n.

TABLE I.
KEY PARAMETERS IN THE SYSTEM SIMULATIONT

Simulation Parameters	Value
Max transmitting power of cellular UE	24dBm
Max transmitting power of D2D UE	15dBm
Spectral density of noise power	173dBm/Hz
Cell Length	500m
Number of D2D UE	30(variable)
Number of cellular UE	30
Pathloss factor	4
Frequency of carrier	2GHz
Radius of D2D pairs	20m(variable)
Threshold of lowest transmitting rate R_{th}	2Mbps(variable)

If the right terminal x^* is a D2D pair n^*, $n=x$, $\Omega_n \neq \varnothing$:

(1) Choose the right RB k_{n^*} with the highest channel gain to the D2D pair n^* according to $k_{n^*}=\arg\max_{k=1,2,..,K}|h_{nk}|$;

(2) Calculate the data rate $R_{n^*}=\sum_{k\in\Omega_n}\log_2(1+\frac{p_{n^*k}|h_{n^*k}|^2}{N_0\omega+\sum_{m=1}^{M}\rho_{mk}p_{mk}|h_{mk}|^2})$, where the p_{n^*k} is the transmit power of the n^*-th D2D pair by using the water-filling algorithm;

(3) Update the priority of n^*-th D2D pair Υ_{n^*}, $\Upsilon_{n^*}=\frac{R_{n^*}}{\theta_{n^*}}$;

(4) Update the channel gain of the whole hybrid network, Ω_m, and Ω_n.

IV. PERFORMANCE ANALYSIS

In this section, we present system simulation results to illustrate the performance of the proposed resource and power allocation scheme. The simulated system network is contained in a 500m×500m area where BS is located in the center. Cellular UEs and D2D UE transmitters are randomly placed in this area according to uniform distribution, and the corresponding D2D receivers are randomly placed around D2D transmitters with the radius of 40m (the factor is variable). We set θ=1,2,3 for the three grades of the priority of data rate in the hybrid network. The channel gain is $h=d^{-\alpha}f$, where α is pathloss factor, d is the distance between the interference source and the terminal received the interference, f is independent unit-mean exponential random variable for modeling the frequency-selective fading channels. Key parameters are shown in TABLE I.

By randomly operate the above scenario simulation for 2000 times, we get the average number of iteration as shown in Figure 2. So the convergence of the proposed algorithm in Section III is guaranteed. And with the higher lowest transmitting rate R_{th}, the convergence of sum throughput is better.

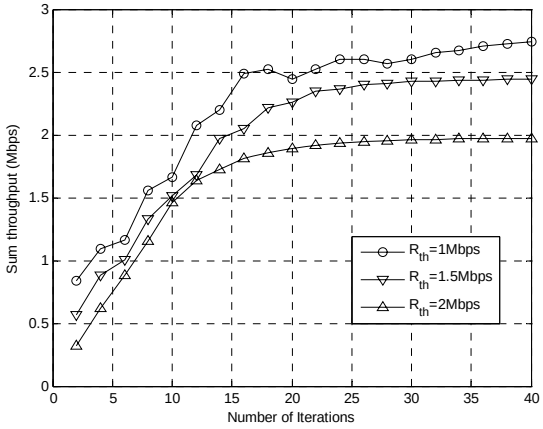


Figure2. Convergence performance

Fig.3 shows the relationship between the sum-throughput and the radius of the D2D pair. We can see that with the increasing of radius between D2D pair, both sum-throughputs of the proposed and traditional algorithm are decreased. It is because that the radius increment requires increasing the D2D transition power, which may bring interference to cellular users. We can still find out that the sum-throughput of proposed algorithm is always higher than the traditional algorithm. Comparing with the traditional algorithm, the decline of proposed algorithm is slower which means the minimum throughput constraint of D2D pairs has effect on the sum-throughput. Because in our scheduling algorithm, some D2D pair with low throughput should be allowed to occupy more spectrum resource blocks in order to maintain the fairness.

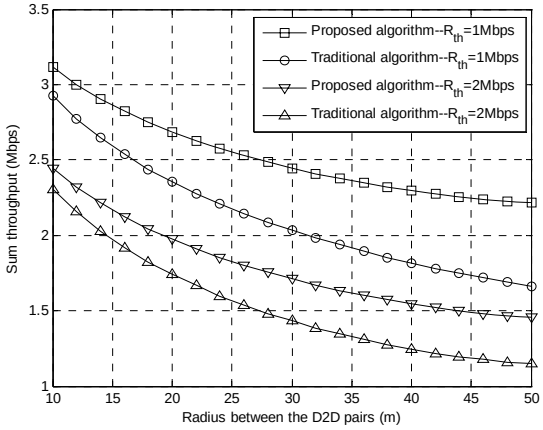


Figure3. Relationship between sum throughput and radius of D2D pairs

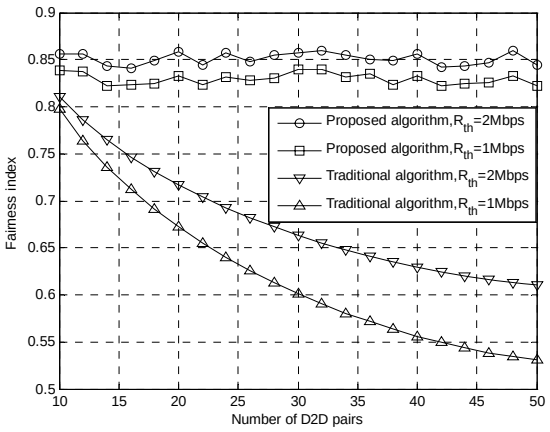


Figure4. Fairness performance

In Fig.4, the fairness of our proposed algorithm is compare with the algorithm presented in [20]. From [20], the Jain's fairness index is defined as: $f(T_1,T_2,...,T_N)=\frac{(\sum_{i=1}^{N}T_i)^2}{N\sum_{i=1}^{N}T_i^2}$, where T is the throughput of each terminal user, N is the number of terminal user. And the Jain's fairness index is used to evaluate the fairness of the algorithms. From Fig.4, the fairness index of the proposed algorithm is better than the traditional algorithm because it can satisfy the data rate requirement of the D2D pairs. And the fair index of the traditional algorithm decreases with the increasing of the number of D2D pairs increase in the hybrid network. Meanwhile the fairness index of the proposed algorithm maintains a relative stable value.

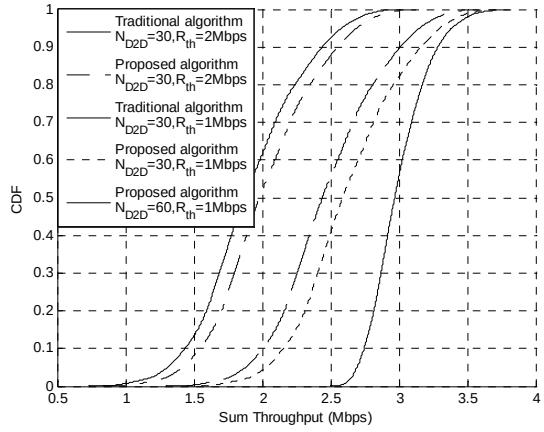


Figure5. CDF of sum-throughput for different number and R_{th} of D2D pairs

Fig.5 illustrates the cumulative Distribution Function (CDF) of sum-throughput for hybrid network. The parameters are shown in TABLE 1, the number of D2D is 30, 60 and the threshold of lowest transmitting rate is 1Mbps and 2Mbps. From fig.5, the leftmost curve is denote the traditional algorithm with the number of D2D pairs is 30, the R_{th} is 2Mbps, the rightmost curve is denote the proposed algorithm with the number of D2D pairs is 60, the R_{th} is 1Mbps, and the sum-throughput of two algorithm scheme increase with the number of D2D pair. We can see the proposed algorithm scheme in this paper is better than the traditional algorithm because of the steep slope. And the more number of D2D pair the superior performance presents.

V. CONCLUSIONS

In this paper, we have given the optimal resources allocation and power control for the hybrid network which is built up by placing a certain number of D2D (Device-to-Device) terminal users in a traditional wireless communication network. The problem is formed as a sum-rate optimization with rate ratio constraint for both cellular and D2D users. First we drive out the optimal wireless resources allocation scheme for achieving the maximum the data rate of hybrid network including the cellular and D2D users. Then we concluded that: the relationship of the transmission power and spectrum resource allocation of the hybrid network is mutually conditional; and the water-filling power allocation algorithm can be used to implement the power allocation of cellular UEs and D2D UEs. And then we proposed the specific scheduling algorithm. The simulation results verify the correctness of our conclusions.

ACKNOWLEDGMENT

This work is supported by National Science and Techn ology Major Projects under grant 2012ZX03003011 and 2012ZX03003007; and in part by the National Key Basic Research Program of China (973 Program) under grant n o.2012CB316005; the Joint Funds of NSFC-Guangdong under Grant U1035001.

REFERENCES

[1] Daquan Feng ; Lu Lu ; Yi Yuan-Wu ; Li, G.Y. ; Gang Feng ; Shaoqian Li. Device-to-Device Communications Underlaying Cellular Networks[j]. Communications IEEE Transactions on Volume: 61 , Issue: 8 2013 , Page(s): 3541 - 3551

[2] Doppler, K. ; Chia-Hao Yu ; Ribeiro, C.B. ; Janis, P.. Mode Selection for Device-To-Device Communication Underlaying an LTE-Advanced Network. Wireless Communications and Networking Conference (WCNC), 2010 IEEE Page(s): 1 - 6.

[3] Rodziewicz, Marcin. Network coding aided Device-to-Device communication. European Wireless, 2012. EW. 18th European Wireless Conference. Page(s): 1 – 5.

[4] K. Doppler, M. Rinne, C. Wijting, C.B. Ribeiro, and K. Hugl, Device-to-device communications as an underlay to LTE-Advanced Networks, IEEE Commun. Mag., vol. 47, no. 12, pp. 42-49, Dec. 2009.

[5] M. Scott Corson, R. Laroia, J. Li, V. Park, T. Richardson, and G. Tsirtsis, Toward proximity-aware internetworking, IEEE Wireless Mag., vol. 17, no. 6, pp.26-33, Dec. 2010.

[6] Yun Niu, Liji Wu, Xiangmin Zhang. An IPSec Accelerator Design for a 10Gbps In-Line Security Network Processor.[J] IEEE Journal of Computers, Vol 8, No 2 (2013), 319-325, Feb 2013.
[7] Lyu-Han Chen ; Wu, E.H.-K. ; Ming-I Hsieh ; Jorng-Tzong Horng ; Gen-Huey Chen . Credit-based low latency packet scheduling algorithm for real-time applications . Communication, Networks and Satellite (ComNetSat), 2012 Page(s): 15 - 19.
[8] Bin Jiao, Qunxian Chen, Shaobin Yan. A Cooperative Co-evolution PSO for Flow Shop Scheduling Problem with Uncertainty.[J] IEEE Journal of Computers, Vol 6, No 9 (2011), 1955-1961, Sep 2011.
[9] Kizhakkethottam, Jubilant J ; Karthik, S ; Chandra, S S Vinod. Novel Fair Share Flow scheduling algorithm for ATM network. Communications, Computers and Signal Processing (PACRIM), 2013 Page(s): 217 - 222.
[10] Po-Kai Huang ; Xiaojun Lin ; Chih-Chun Wang. A Low-Complexity Congestion Control and Scheduling Algorithm for Multihop Wireless Networks With Order-Optimal Per-Flow Delay. Networking, IEEE/ACM Transactions on Volume: 21 , Issue: 2. 2013 , Page(s): 495 – 508.
[11] Seppala, J. ; Koskela, T. ; Tao Chen ; Hakola, S.. Network controlled Device-to-Device (D2D) and cluster multicast concept for LTE and LTE-A networks. Wireless Communications and Networking Conference (WCNC), 2011 IEEE, Page(s): 986 – 991.
[12] Di Zhang ; Xuewen Liao ; Jianguo Deng ; Wei Wang. Interference coordination mechanism for device-to-device communication in uplink period underlaying cellular networks. Wireless Communications & Signal Processing (WCSP), 2012 Page(s): 1 – 5.
[13] Kaufman, Brett ; Lilleberg, Jorma ; Aazhang, Behnaam. Spectrum Sharing Scheme Between Cellular Users and Ad-hoc Device-to-Device Users. Wireless Communications, IEEE Transactions on Volume: 12 , Issue: 3 2013 , Page(s): 1038 – 1049.
[14] Pei Liu ; Chunjing Hu ; Tao Peng ; Wenbo Wang. Distributed cooperative admission and power control for Device-to-Device links with QoS protection in cognitive heterogeneous network. Communications and Networking in China (CHINACOM), 2011 2 7th, Page(s): 712 - 716.
[15] Xiaoyue Zhu, Si Wen, Gen Cao, Xin Zhang. QoS-based resource allocation scheme for Device-to-Device (D2D) radio underlaying cellular networks. Telecommunications (ICT), 2012 19th International Conference on 23-25 April.
[16] Tao Peng, Qianxi Lu, Haiming Wang, Shaoyi Xu. Interference avoidance mechanisms in the hybrid cellular and device-to-device systems. 2009 IEEE 20th International Symposium on 13-16 Sept. 2009, 617 – 621.
[17] Wei Wang, Wang Wenbo, Lu Qianxi, Peng Tao. An Uplink Resource Allocation Scheme for OFDMA-Based Cognitive Radio Networks. International Journal of Communication Systems Volume 22, Issue 5, pages 603–623,May 2009.
[18] Mung Chiang, "geometric programming for communication systems", Princeton University, USA.
[19] Dongyu Wang, Xiaoxiang Wang, Yuan Zhao, "An interference Coordination Scheme for Device-to-Device Multicast in Cellular Networks," VTC 2012,Sept.2012.
[20] Yu Wei. A Dual Decomposition Approach to the Sum Power Gaussian Vector Multiple Access Channel Sum Capacity Problem. The John Hopkins University : 2003 Conferenceon Information Sciences and Systems, March 12-14, 2003.
[21] Yingjie Xia, Mingzhe Zhu. Power-aware Small World Topology in Ad Hoc Networks.[J] IEEE Journal of Computers. Vol 7, No 1(2012), 244-251.
[22] Yu, W., Ginis, G. and Cioffi, J. M. Distributed multiuser power control for digital subscriber lines. IEEE Journal on Selected Areas in Communication. Volume: 20, Issue: 5, June 2002, pp 1105-1115.

Wenwen Liu (wenwl67@gmail.com**)** received her M.Sc. degree in Communications and Information Engineering from Nanjing University of Posts & Telecommunications, China, in 2009. Currently, she pursues studies toward a Ph.D. in Communications and Information Engineering at Beijing University of Posts & Telecommunications, China. Her research interests include the integration of device-to-device communication, hybrid network, cognitive radio, and radio resource management.

Yang Yang (buptyy1015@gmail.com) received his M.Sc. degree in Communications and Information Engineering from Beijing University of Posts & Telecommunications, China, in 2011. Currently, he pursues studies toward a Ph.D. in Communications and Information Engineering at Beijing University of Posts & Telecommunications, China. His research interests include the integration of device-to-device communication cognitive radio, radio resource management, and hybrid network capacity.

Tao Peng (pengtao@bupt.edu.cn) received his BS, MS and PhD degree from School of Telecommunications Engineering, Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1999, 2002 and 2010. He is now the associate professor at BUPT, working at wireless Signal Processing and Network (WSPN) Lab. His research interests include Long Term Evolution (LTE) system, Cognitive Radio networks and so on.

Wenbo Wang (wbwang@bupt.edu.cn) received his BS, MS and PhD degree from School of Telecommunications Engineering, Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1986, 1989 and 1992. He is now a professor at BUPT, in charge of Wireless Signal Processing and Network (WSPN) Lab. The main research of WSPN lies in key technology and theory of Wireless Communication, Signal Processing, Wide-Band Access, and so on.

Polynomial Smooth Twin Support Vector Machines Based on Invasive Weed Optimization Algorithm

Shifei Ding[1, 2], Huajuan Huang[1, 2], Junzhao Yu[1, 2], Fulin Wu[1, 2]
School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China, 221116
Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, 100190
Email: dingsf@cumt.edu.cn

***Abstract*—Smoothing functions can transform the unsmooth twin support vector machines (TWSVM) into smooth ones, and thus better classification results can be obtained. It has been one of the key problems to seek a better smoothing function in this field for a long time. In this paper, a novel version for smooth TWSVM, termed polynomial smooth twin support vector machines (PSTWSVM), is proposed. In PSTWSVM, using the series expansion, a new class of polynomial smoothing is proposed, and then their important properties are discussed. It is shown that the approximation accuracy and smoothness rank of polynomial functions can be as high as required. Subsequently, the polynomial functions are used to convert the original constrained quadratic programming problems of TWSVM into unconstrained minimization problems, and then are solved by the well-known Newton-Armijo algorithm. Meanwhile, in order to find the suitable parameters of PSTWSVM, Invasive Weed Optimization (IWO) algorithm is used to optimize the proposed algorithm. Then we propose an algorithm called polynomial smooth twin support vector machines based on invasive weed optimization algorithm (PSTWSVM-IWO). Finally, the effectiveness of the proposed method is demonstrated via experiments on synthetic and UCI benchmark datasets.**

***Index Terms*—Polynomial function, Newton-Armijo, Invasive weed optimization algorithm, Parameter optimization, Twin support vector machines**

I. INTRODUCTION

Support vector machine (SVM) proposed by Vapnik and co-worker [1] is a computationally powerful kernel-based tool for binary data classification and regression. Because the theory of SVM is based on the idea of structural risk minimization principle, SVM has successfully solved the high dimensionality and local minimum problems. Therefore, compared with other machine learning methods, such as artificial neural network [2-3], SVM owns better generalization ability. Within a few years after its introduction SVM has played excellent performance in many real-world predictive data mining applications such as text categorization [4], time series prediction [5], pattern recognition [6] and image processing [7], etc.

Although SVM owns better generalization ability compared with many other machine learning methods, however, its computational complexity in training stage is too expensive. To address this problem, so far, many improved algorithms have been presented, such as chunking algorithm [8], decomposition algorithm [9] and sequential minimal optimization (SMO) [10], etc. However, these algorithms are too complex. On the other hand, many researchers have proposed some deformation algorithms based on standard SVM. For example, in 2006, Mangasarian et al. [11] proposed a nonparallel plane classifier for binary data classification, named generalized eigenvalue proximal support vector machine (GEPSVM). The essence of GEPSVM is to look for two nonparallel planes, so that data points of each class are proximal to one of them. GEPSVM has good learning speed, but its classification accuracy is low. In 2007, Jayadeva et al. [12] proposed a new machine learning method called twin support vector machine (TWSVM) for the binary classification in the spirit of GEPSVM. TWSVM would generate two non-parallel planes, such that each plane is closer to one of the two classes and is as far as possible from the other. In TWSVM, a pair of smaller sized quadratic programming problems (QPPs) are solved, instead of solving a single large one in SVM, makes the computational speed of TWSVM approximately 4 times faster than the traditional SVM. Because of its excellent performance, TWSVM has been applied to many areas such as speaker recognition [13], medical detection [14], etc.

Similar to SVM, TWSVM solves its QPPs in the dual space. However, this solving method will be affected by time and memory constraints when dealing with the large datasets, which would make the learning speed of TWSVM low. In order to address this problem, in 2008, M. Arun Kumar et al. [15] used the sigmoid function to approach the objective function of TWSVM and then proposed smooth twin support vector machines

This work is supported by the National Natural Science Foundation of China (Nos.61379101), the National Key Basic Research Program of China (No. 2013CB329502), and the Natural Science Foundation of Jiangsu Province (No.BK20130209).

Corresponding author, Shifei Ding, E-mail: dingsf@cumt.edu.cn.

doi:10.4304/jcp.9.5.1226-1233

(STWSVM). STWSVM directly solved QPPs in the original space instead of the dual space. Experimental results showed that STWSVM could make the classifier faster to compute in the classification phase than TWSVM. However, because of the low approximation ability of the sigmoid function, the classification accuracy of STWSVM was unsatisfactory. In order to further improve the classification performance of STWSVM, looking for a new smooth function with better approximation ability is the key problem.

In this paper, using the series expansion, a new class of polynomial smoothing is proposed. We have proved that the proposed smoothing functions have better smooth performance and their approximation accuracy can be as high as required. Subsequently, the polynomial functions are used to convert the original constrained quadratic programming problems of TWSVM into unconstrained minimization problems, and then are solved by the well-known Newton-Armijo algorithm. Based on the above idea, a novel version for smooth TWSVM, termed polynomial smooth twin support vector machines (PSTWSVM), is proposed in this paper. Besides, in order to overcome PSTWSVM parameters selection problem, we use Invasive Weed Optimization (IWO) algorithm [16] which has fast global searching ability to select PSTWSVM parameters, so that we would obtain the optimal parameters combination. Finally, the experimental results show the effectiveness and stability of the proposed method.

The paper is organized as follows: In section 2, we propose the PSTWSVM model and prove its global convergence. In section 3, PSTWSVM-IWO is detailed introduced and analyzed. Computational comparisons on synthetic and UCI datasets are done in section 4 and section 5 gives concluding remarks.

II. Polynomial Smooth Twin Support Vector Machines

A. Twin Support Vector Machines

Consider a binary classification problem of classifying m_1 data points belonging to class +1 and m_2 data points belonging to class -1. Then let matrix A in $R^{m_1 \times n}$ represent the data points of class +1 while matrix B in $R^{m_2 \times n}$ represent the data points of class -1. Two nonparallel hyper-planes of the linear TSVMs can be expressed as follows.

$$x^T w_1 + b_1 = 0 \quad \text{and} \quad x^T w_2 + b_2 = 0 \tag{1}$$

The target of TSVMs is to generate the above two nonparallel hyper-planes in the n-dimensional real space R^n, such that each plane is closer to one of the two classes and is as far as possible from the other. A new sample point is assigned to class +1 or -1 depending upon its proximity to the two nonparallel hyper-planes. The linear classifiers are obtained by solving the following optimization problems.

$$\min_{w^{(1)},b^{(1)},\xi^{(2)}} \frac{1}{2}\left\|Aw^{(1)} + e_1 b^{(1)}\right\|^2 + c_1 e_2^T \xi^{(2)}$$
$$s.t. \quad -(Bw^{(1)} + e_2 b^{(1)}) \geq e_2 - \xi^{(2)},$$
$$\xi^{(2)} \geq 0. \tag{2}$$

$$\min_{w^{(2)},b^{(2)},\xi^{(1)}} \frac{1}{2}\left\|Bw^{(2)} + e_2 b^{(2)}\right\|^2 + c_2 e_1^T \xi^{(1)}$$
$$s.t. \quad (Aw^{(2)} + e_1 b^{(2)}) \geq e_1 - \xi^{(1)},$$
$$\xi^{(1)} \geq 0. \tag{3}$$

where c_1 and c_2 are penalty parameters, $\xi^{(1)}$ and $\xi^{(2)}$ are slack vectors, e_1 and e_2 are vectors of ones of appropriate dimensions.

In TWSVM, generally, we solve the QPPs in the dual space. However, this solving method will be affected by time and memory constraints when dealing with the big datasets. In order to improve the computational speed, the TWSVM model represented by (2) and (3) would be transformed into two unconstrained non-smooth optimization problems by using the plus function.

According to the KKT theorem, we can get

$$\xi^{(2)} = \max\{0, e_2 + (Bw^{(1)} + e_2 b^{(1)})\} \tag{4}$$

$$\xi^{(1)} = \max\{0, e_1 - (Aw^{(2)} + e_1 b^{(2)})\} \tag{5}$$

The optimization problems (2) and (3) can be rewritten as

$$\min \frac{1}{2}\left\|Aw^{(1)} + e_1 b^{(1)}\right\|^2 + c_1 e_2^T \max\{0, (e_2 + Bw^{(1)} + e_2 b^{(1)})\} \tag{6}$$

$$\min \frac{1}{2}\left\|Bw^{(2)} + e_2 b^{(2)}\right\|^2 + c_2 e_1^T \max\{0, (e_1 - Aw^{(2)} + e_1 b^{(2)})\} \tag{7}$$

Let $(x_1)_+ = \max\{0, (e_2 + Bw^{(1)} + e_2 b^{(1)})\}$,

$(x_2)_+ = \max\{0, (e_1 - Aw^{(2)} + e_1 b^{(2)})\}$,

where $(x_1)_+$ and $(x_2)_+$ are the plus functions. Apparently, the objective functions of the unconstrained optimization problems (6) and (7) are convex and non-smooth.

Theorem 1 The unconstrained TWSVM model can be represented as (6) and (7) and the model is continuous but non-smooth.

Theorem 1 shows that (6) and (7) are non-smooth, so we can't use the gradient optimization method such as the Newton-Armijo method to solve (6) and (7). In order to address this problem, we will use the polynomial smooth function to approach (6) and (7).

B. The Polynomial Smooth Function

Weierstrass Theorem [17] Set arbitrary continuous function $f(x)$, $x \in [m, n]$, existing polynomial $P_n(x)$ makes $\lim_{n\to\infty} \max_{m \leq x \leq n} |f(x) - P_n(x)| = 0$.

Weierstrass's theorem shows that any continuous real-valued function in closed interval can be arbitrarily approached by the polynomial function. From theorem 1 we can know that the plus function is a continuous function, so we can use the polynomial function to approach it. In this paper, we will give the common formula of the polynomial smooth function by transforming it to an equivalent infinite series.

Lemma 1 [18] Two expansion of $m = \frac{1}{2}$ can be expressed as

$$\sqrt{1+x}=1+\frac{1}{2}x-\frac{1}{2\cdot 4}x^2+\frac{1\cdot 3}{2\cdot 4\cdot 6}x^3-\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6\cdot 8}x^4+\cdots=$$
$$1+\frac{1}{2}x-\sum_{n=2}^{\infty}\frac{(2n-3)!!}{(2n)!!}(-x)^n,\ -1\le x\le 1 \quad (8)$$

Theorem 2 The plus function x_+ can be transformed to an equivalent infinite series in $[-\frac{1}{k},\frac{1}{k}]$ as follows.

$$x_+=\frac{1}{2k}\left(\frac{1+k^2x^2}{2}-\sum_{n=2}^{\infty}\frac{(2n-3)!!}{(2n)!!}(1-k^2x^2)^n\right)+\frac{x}{2} \quad (9)$$

Proof According to the definition of x_+, we can get

$$x_+=\max(0,x)=\frac{|x|+x}{2}=\frac{|kx|}{2k}+\frac{x}{2}=\frac{1}{2k}\sqrt{1+(k^2x^2)-1}+\frac{x}{2} \quad (10)$$

According to lemma 1 and (10), x_+ can be rewritten as

$$x_+=\frac{1}{2k}\left(\frac{1+k^2x^2}{2}-\sum_{n=2}^{\infty}\frac{(2n-3)!!}{(2n)!!}(1-k^2x^2)^n\right)+\frac{x}{2} \quad (11)$$

End.

Theorem 3 The polynomial approximation function for x_+ in $[-\frac{1}{k},\frac{1}{k}]$ is

$$P_n(x,k)=\begin{cases}x, & x\ge\frac{1}{k}\\ \frac{1}{2k}\left(\frac{1+k^2x^2}{2}-\sum_{l=2}^{n}\frac{(2l-3)!!}{(2l)!!}(1-k^2x^2)^l\right)+\frac{x}{2}, & |x|<\frac{1}{k},k>0\\ 0, & x\le-\frac{1}{k},\end{cases} \quad (12)$$

where n is a positive integer. The approximation image of the plus function by the polynomial function when $k=10$, $n=1,2$ is shown as figure 1. From Figure 1, we can see that the approximation accuracy of $P_n(x,k)$ will be higher with n larger.

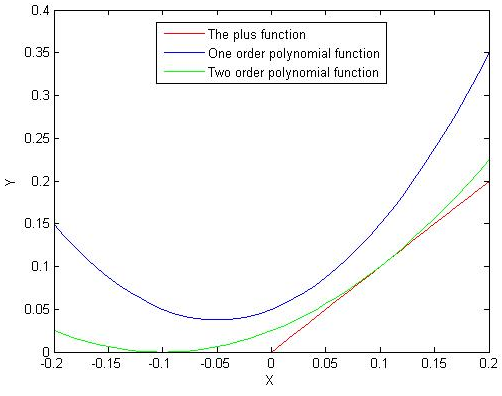


Figure 1. The approximation image of the plus function

Theorem 4 $P_n(x,k)$ is defined as (12), it has some characteristics as follows.

(1) $P_n(x,k)$ has n-order smoothness about x.

(2) $\lim_{n\to\infty}\max(P_n(x,k)-x_+)=0$.

Proof (1) If $P_n(x,k)$ has n-order smoothness about x, it must meet the following conditions.

$$P_n(\frac{1}{k},k)=\frac{1}{k},\quad P_n(-\frac{1}{k},k)=0;$$

$$\nabla P_n(\frac{1}{k},k)=1,\quad \nabla P_n(-\frac{1}{k},k)=0;$$

$$\nabla^n P_n(\frac{1}{k},k)=0,\quad \nabla^n P_n(-\frac{1}{k},k)=0,\quad n\ge 2$$

According to (12), it can be got

$$P_n(\frac{1}{k},k)=\frac{1}{k},\quad P_n(-\frac{1}{k},k)=0$$

We find the partial derivative of x, it can have when $n\ge 1$,

$$\nabla P_n(x,k)=\begin{cases}1, & x\ge\frac{1}{k}\\ \frac{kx}{2}\left(1+\sum_{l=2}^{n}\frac{(2l-3)!!}{(2l-2)!!}(1-k^2x^2)^{l-1}\right)+\frac{1}{2}, & |x|<\frac{1}{k},k>0\\ 0, & x\le-\frac{1}{k}\end{cases}$$

when $n\ge 2$,

$$\nabla^2 P_n(x,k)=\begin{cases}0, & x\ge\frac{1}{k}\\ \frac{k}{2}\left(1+\sum_{l=2}^{n}\frac{(2l-3)!!}{(2l-2)!!}(1-k^2x^2)^{l-1}\right)-\\ \frac{k^3x^2}{2}\sum_{l=2}^{n}\frac{(2l-3)!!}{(2l-4)!!}(1-k^2x^2)^{l-2}, & |x|<\frac{1}{k},k>0\\ 0, & x\le-\frac{1}{k}\end{cases}$$

$\nabla P_n(x,k)$, $\nabla^2 P_n(x,k)$ and $\nabla^n P_n(x,k)$ ($n>2$) are existence and continuation in $x=\pm\frac{1}{k}$. So $P_n(x,k)$ has n-order smoothness about x.

(2) According to weierstrass's Theorem, it can be got easily

$\lim_{n\to\infty}\max(P_n(x,k)-x_+)=0$.

End.

Theorem 4 shows that the polynomial smooth function transformed to an equivalent infinite series can achieve arbitrary precision to approach the plus function when n is large enough.

C. The Optimal Smoothing Factor

There is a parameter k called smoothing factor in (12). We give the formula of optimal smoothing factor as follows.

Theorem 5 Give arbitrary precision E, if the smooth function $P_n(x,k)$ meets the condition $|P_n(x,k)-x|\le E$ when it approaches to x_+, the smoothing factor k is called the optimal smoothing factor and is denoted as $k_{opt}(n,E)$.

Because the error of $P_n(x,k)$ approaching to x_+ is maximum in $x=0$, we can get $k_{opt}(n,E)$ when it meets the condition $P_n(x,k)-x_+ \le E$ in $x=0$.

Therefore, if $x=0$, calculate (12), we can get

$$k_{opt}(n,E) \ge \frac{\frac{1}{2}-\sum_{l=2}^{n}\frac{(2l-3)!!}{(2l)!!}}{2E} \quad (13)$$

D. PSTWSVM Algorithm

Because $P_n(x,k)$ has n -order smoothness when $n \ge 2$, Newton-Armijo optimization algorithm can be used to solve the following unconstrained optimization problems.

$$\min \frac{1}{2}\left\|Aw^{(1)}+e_1b^{(1)}\right\|^2 + c_1e_2^T P((e_2+Bw^{(1)}+e_2b^{(1)}),k) \quad (14)$$

$$\min \frac{1}{2}\left\|Bw^{(2)}+e_2b^{(2)}\right\|^2 + c_2e_1^T P((e_1+Aw^{(2)}+e_1b^{(2)}),k) \quad (15)$$

Algorithm 1 PSTWSVM based on the Newton-Armijo method

Input: Give the initial value $(w^0,b^0)\in R^{n+1},\eta$, let the iteration number $i=0$, the order of polynomial function n, the arbitrary precision E;

Output: The optimal value of the objective function.

Step1: calculate $P_n(x,k)$ and $g^i=\nabla P_n(x,k)$.

Step2: If $\left\|g^i\right\|\le\eta$, select $(w^*,b^*)=(w^i,b^i)$, then terminate programs. Otherwise according to $\nabla^2 P_n(x,k)d^i=-g^i$, calculate the down direction d^i.

Step3: take $\delta\in(0,\frac{1}{2})$, $\lambda_i=\max\{1,\frac{1}{2},\frac{1}{4},\cdots\}$, let $P_n(x,k)-P_n((w^i,b^i)+\lambda_i d^i,k)\ge-\delta\lambda_i g^i d^i$, then let $(w^{i+1},b^{i+1})=(w^i,b^i)+\lambda_i d^i$.

Step4: Let $i\leftarrow i+1$, turn to Step2.

E. The Nonlinear PSTWSM

If the previous conclusions are extended to nonlinear smooth PSTWSVM, it can be used to deal with the nonlinear problem.

In order to obtain the nonlinear classifiers we consider the following kernel generated surfaces

$$K(x^T,C^T)u_1+b_1=0,\ K(x^T,C^T)u_2+b_2=0, \quad (16)$$

where $C^T=[A\quad B]^T$, $(u_{(i)},b_{(i)})\in(R^m\times R)$ $(i=1,2)$ and K is an chosen kernel. The nonlinear TWSVM are obtained by solving the following optimization problems.

$$\min_{w^{(1)},b^{(1)},\xi^{(2)}} \frac{1}{2}\left\|K(A,C^T)w^{(1)}+e_1b^{(1)}\right\|^2 + c_1e_2^T\xi^{(2)}$$
$$s.t. \quad -(K(B,C^T)w^{(1)}+e_2b^{(1)})\ge e_2-\xi^{(2)},$$
$$\xi^{(2)}\ge 0. \quad (17)$$

$$\min_{w^{(1)},b^{(1)},\xi^{(2)}} \frac{1}{2}\left\|K(B,C^T)w^{(2)}+e_2b^{(1)}\right\|^2 + c_2e_1^T\xi^{(1)}$$
$$s.t. \quad (K(A,C^T)w^{(2)}+e_1b^{(2)})\ge e_1-\xi^{(1)},$$
$$\xi^{(1)}\ge 0. \quad (18)$$

Introducing the plus function, (17) and (18) can be transformed into the following optimization problems without constraint.

$$\min \frac{1}{2}\left\|K(A,C^T)w^{(1)}+e_1b^{(1)}\right\|^2 + c_1e_2^T(e_2+K(B,C^T)w^{(1)}+e_2b^{(1)})_+ \quad (19)$$

$$\min \frac{1}{2}\left\|K(B,C^T)w^{(2)}+e_2b^{(2)}\right\|^2 + c_2e_1^T(e_1-K(A,C^T)w^{(2)}+e_1b^{(2)})_+ \quad (20)$$

We can get the nonlinear PSTSVMs-NA model using the polynomial smooth function.

$$\min \frac{1}{2}\left\|K(A,C^T)w^{(1)}+e_1b^{(1)}\right\|^2 + c_1e_2^T P((e_2+K(B,C^T)w^{(1)}+e_2b^{(1)}),k) \quad (21)$$

$$\min \frac{1}{2}\left\|K(B,C^T)w^{(2)}+e_2b^{(2)}\right\|^2 + c_2e_1^T P((e_1-K(A,C^T)w^{(2)}+e_1b^{(2)}),k) \quad (22)$$

The previous conclusions and theorems are also applicable to the nonlinear PSTWSVM model.

III. PSTWSVM BASED ON INVASIVE WEED OPTIMIZATION ALGORITHM

A. Analysis the Penalty Parameters of PSTWSVM

The role of penalty parameters c_1 and c_2 is to adjust the ratio between the confidence range with the experience risk in the defining feature, so that the generalization ability of PSTWSVM can achieve the best state. The values of c_1 and c_2 smaller expresses the punishment on empirical error smaller. Do it this way, the complexity of PSTWSVM is smaller, but its fault tolerant ability is worse. The values of c_1 and c_2 are greater, the data fitting degree is higher, but its generalization capacity will be reduced. From the above analysis, we can know that the parameters selection is very important for PSTWSVM.

After the above analysis, in this paper, Invasive Weed Optimization (IWO) algorithm which has fast global searching ability is used to select the PSTWSVM parameters and the mixed kernel parameters.

B. Invasive Weed Optimization

In 2006, a novel stochastic optimization model, invasive weed optimization (IWO) algorithm [16], was proposed by Mehrabian and Lucas, which is inspired from a common phenomenon in agriculture: colonization of invasive weeds. Not only it has the robustness, but also it is easy to understand and program. So far, it has been applied in many engineering fields [19-20].

In the classical IWO, weeds represent the feasible solutions of problems and population is the set of all weeds. A finite number of weeds are being dispread over the search area. Every weed produces new weeds depending on its fitness. The generated weeds are randomly distributed over the search space by normally distributed random numbers with a mean equal to zero. This process continues until maximum number of weeds is reached. Only the weeds with better fitness can survive and produce seed, others are being eliminated. The process continues until maximum iterations are reached or hopefully the weed with best fitness is closest to optimal solution.

The process is addressed in details as follows:

Step 1: *Initialize a population*

A population of initial solutions is being dispread over the *D* dimensional search space with random positions.

Step 2: *Reproduction*

The higher the weed's fitness is, the more seeds it produces. The formula of weeds producing seeds is

$$weed_n = \frac{f - f_{\min}}{f_{\max} - f_{\min}}(s_{\max} - s_{\min}) + s_{\min} \qquad (23)$$

where, f is the current weed's fitness. f_{max} and f_{min} respectively represent the maximum and the least fitness of the current population. s_{max} and s_{min} respectively represent the maximum and the least value of a weed.

Step 3: *Spatial dispersal*

The generated seeds are randomly distributed over the D dimensional search space by normally distributed random numbers with a mean equal to zero, but with a varying variance. This ensures that seeds will be randomly distributed so that they abide near to the parent plant. However, standard deviation (σ) of the random function will be reduced from a previously defined initial value (σ_{init}) to a final value (σ_{final}) in every generation. In simulations, a nonlinear alteration has shown satisfactory performance, given as follows

$$\sigma_{cur} = \frac{(iter_{\max} - iter)^n}{(iter_{\max})^n}(\sigma_{init} - \sigma_{final}) + \sigma_{final} \qquad (24)$$

Where, $iter_{max}$ is the maximum number of iterations, σ_{cur} is the standard deviation at the present time step and n is the nonlinear modulation index. Generally, n is set to 3.

Step 4: *Competitive exclusion*

After passing some iteration, the number of weeds in a colony will reach its maximum (P_MAX) by fast reproduction. At this time, each weed is allowed to produce seeds. The produced seeds are then allowed to spread over the search area. When all seeds have found their position in the search area, they are ranked together with their parents (as a colony of weeds). Next, weeds with lower fitness are eliminated to reach the maximum allowable population in a colony. In this way, weeds and seeds are ranked together and the ones with better fitness survive and are allowed to replicate. The population control mechanism also is applied to their offspring to the end of a given run, realizing competitive exclusion.

C. The Algorithm steps of PSTWSVM-IWO

The accuracy in the sense of CV is used for the fitness of IWO. So the algorithm steps of PSTWSVM-IWO are as follows:

Step1: Select the training dataset and the testing dataset.

Step2: Preprocessing the dataset.

Step3: Constructe the PSTWSVM model.

Step4: Select the optimal parameters using IWO algorithm.

Step5: Train the PSTWSVM model using the optimal parameters.

Step6: Predict the testing dataset.

Step7: Output the classification accuracy.

IV. EXPERIMENT RESULTS AND ANALYSIS

In order to verify the efficiency of PSTWSVM and PSTWSVM-IWO, we conduct two experiments. In the first experiment, in order to show the advantage of PSTWSVM, we make experiments on several benchmark datasets using four algorithms, that is, GEPSVM, TWSVM, STWSVM and PSTWSVM. In the second experiment, we make experiment on NDC dataset to compare PSTWSVM with PSTWSVM-IWO. The dual QPPs arising in TWSVM are solved using mosek optimization toolbox for MATLAB [21] which implements fast interior point based algorithms. Classification accuracy of each algorithm is measured by standard tenfold cross-validation methodology.

A. The First Experiment

In this experiment, we make experiments on several benchmark datasets using four algorithms, that is, GEPSVM, TWSVM, STWSVM and PSTWSVM. The optimal parameters of these algorithms are searched from $\{2^i \mid i = -6, -4, -2, 0, 1, 2, 4, 6\}$ using the grid search algorithm. In PSTWSVM, the parameter of Newton-Armijo method is set $\varepsilon_1 = 1.0E-3$, the approximation accuracy of smooth function is set $\varepsilon_2 = 1.0E-3$. For the nonlinear case, we only consider the Gaussian kernel function. The optimal value of Gaussian kernel parameter is selected over the range $\{2^i \mid i = -6, -4, -2, 0, 1, 2, 4, 6\}$. The order of polynomial is set $n = 5$.

TABLE1.
COMPARISON FOR LINEAR KERNEL

Dataset	PSTWSVM	STWSVM	TWSVM	GEPSVM
Hepatitis	78.05 ± 4.31	77.39 ± 2.15	78.08 ± 2.16	77.28 ± 2.78
Housing	86.21 ± 2.39	84.42 ± 3.87	85.42 ± 4.53	74.81 ± 2.85
Wdbc	96.10 ± 6.32	94.89 ± 4.31	96.22 ± 6.67	92.81 ± 2.54
Glass6	96.52 ± 4.56	95.70 ± 6.05	96.55 ± 2.40	96.21 ± 2.72
Votes	95.50 ± 1.23	94.96 ± 4.24	95.85 ± 2.24	95.63 ± 2.74

TABLE 2.
COMPARISON FOR GAUSSIAN KERNEL

Dataset	PSTWSVM	STWSVM	TWSVM	GEPSVM
Australian	88.23 ± 3.14	86.83 ± 3.24	85.56 ± 2.17	85.24 ± 2.09
Breast-cancer	71.16 ± 1.28	69.38 ± 1.12	69.42 ± 2.32	67.78 ± 1.56
Heart	84.54 ± 4.32	82.89 ± 4.56	82.24 ± 3.59	80.13 ± 3.42
Pima	82.12 ± 3.07	79.76 ± 3.05	78.52 ± 2.48	76.56 ± 2.16
Votes	97.26 ± 2.35	95.14 ± 3.27	95.09 ± 2.56	93.45 ± 2.12
Sonar	91.47 ± 2.24	89.67 ± 3.12	88.99 ± 4.67	85.96 ± 2.23
CMC	77.56 ± 3.25	75.58 ± 3.85	69.54 ± 2.09	66.52 ± 2.16

TABLE 3.
DESCRIPTION OF NDC DATASETS

Dataset	# Training data	# Test data	# Feature
NDC-500	500	50	32
NDC-700	700	70	32
NDC-900	900	90	32
NDC-1k	1000	100	32
NDC-2k	2000	200	32
NDC-3k	3000	300	32
NDC-4k	4000	400	32
NDC-5k	5000	500	32
NDC-10k	10,000	1000	32
NDC-1l	100,000	10,000	32
NDC-3l	300,000	30,000	32
NDC-5l	500,000	50,000	32
NDC-1m	1,000,000	100,000	32

TABLE 4.
COMPARION FOR LINEAR KERNEL

Dataset	PSTWSVM-IWO	PSTWSVM	TWSVM
	Train (%)	Train (%)	Train (%)
	Test (%)	Test (%)	Test (%)
	Time (s)	Time (s)	Time (s)
	81.07	80.05	79.93
NDC-3k	79.69	77.64	77.66
	3.2412	9.1545	27.08
	81.08	79.89	79.80
NDC-4k	74.92	73.78	73.75
	3.5605	10.0665	60.94
	79.87	78.33	79.15
NDC-5k	82.24	80.26	80.23
	4.0734	11.0761	114.24
	87.23	86.48	86.45
NDC-10k	88.56	87.32	87.38
	4.1178	15.1239	1092.07
	85.59	84.35	-
NDC-1l	87.75	86.28	-
	4.996	16.014	-
	81.43	78.75	-
NDC-3l	79.34	75.78	-
	5.899	18.103	-
	79.54	78.26	-
NDC-5l	82.01	79.17	-
	6.1312	18.3505	-

"-" We stop experiment as computing time was very high

Table 1 shows the comparison of classification accuracy for PSTWSVM with GEPSVM, TWSVM and STWSVM for linear kernel on five UCI datasets. Table 2 shows the comparison of classification performance for nonlinear extensions of PSTWSVM with GEPSVM, TWSVM and STWSVM. Table 1 and table 2 show that the accuracy performance of PSTWSVM is better than STWSVM. Therefore, we can know the approximation ability of polynomial function is better than the sigmoid function.

TABLE5.
COMPARION FOR GAUSSIAN KERNEL

Dataset	PSTWSVM-IWO	PSTWSVM	TWSVM
	Train（%）	Train (%)	Train (%)
	Test（%）	Test (%)	Test (%)
	Time (s)	Time (s)	Time (s)
	100.00	100.00	100.00
NDC-500	80.25	80.15	79.18
	0.4654	0.5632	0.789
	99.88	99.54	99.27
NDC-700	85.23	83.17	84.29
	1.3009	4.3123	1.7322
	99.76	99.56	99.56
NDC-900	82.85	81.36	80.58
	1.5017	2.5354	3.4675
	98.83	98.47	98.85
NDC-1k	85.32	84.12	83.85
	1.5029	3.5198	4.1176
	100.00	100.00	99.68
NDC-2k	88.45	88.24	88.27
	11.09	21.25	25.8958
	100.00	100.00	99.53
NDC-3k	91.67	90.12	90.45
	65.299	78.609	85.445

B. The Second Experiment

Similar to SVM and TWSVM, the learning performance and generalization ability of PSTWSVM is very dependent on its parameters selection. In the above experiment, we used the grid search algorithm to find the parameters values, which is a commonly used method. However, it will lead to low efficiency when dealing with big data. In this paper, we try to use invasive weed optimization (IWO) algorithm to optimize PSTWSVM and propose an algorithm called PSTWSVM-IWO. In this section, we will conduct experiment on NDC datasets which are generated by David Musicant' NDC Data Generator [22] to test the ability of our algorithm for dealing with big data. The parameters of IWO are as follows: $D=5$, $P\_MAX=30$, $s_{\max}=5$, $s_{\min}=1$, $n=3$, $\sigma_{init}=[1,0.1,1,1,1]$, $\sigma_{final}=[0.1,0.1,0.1,0.1,0.1]$. In IWO algorithm, the accuracy in the sense of CV is used for the fitness of IWO. Therefore, the fitness value is closer to 100, the obtained parameters is closer to the optimal value. Table 3 gives a description of NDC datasets. Table 4 shows the comparison of computing time and accuracy for three algorithms with linear kernel. On the other hand, table 5 shows the comparison of classification performance for these algorithms with Gaussian kernel. Figure 2~3 are the fitness curves of IWO searching the optimal parameters for dealing with NDC-500, NDC-700, respectively.

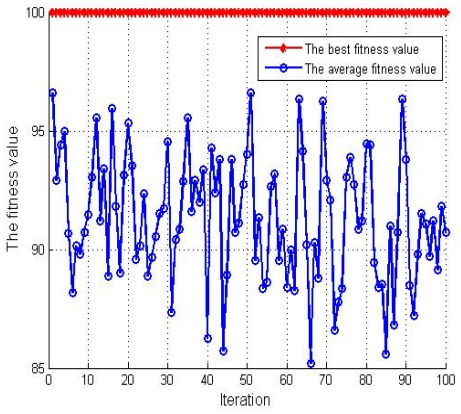


Figure 2. The fitness curves of IWO searching the optimal parameters for dealing with NDC-500

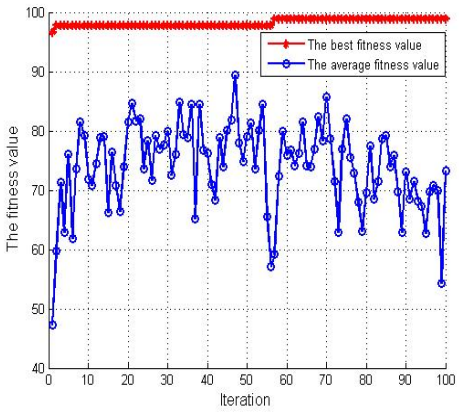


Figure 3. The fitness curves of IWO searching the optimal parameters for dealing with NDC-700

From table 4, we can see that in view of the high computing time, TWSVM can't work when the training samples reach 100000. However, PSTWSM-IWO and PSTWSVM can get reasonable accuracy in the relatively short time when the training samples reach 500000, which indicates that PSTWSM-IWO and PSTWSVM have the advantage on dealing with big data comparing with TWSVM. Furthermore, from table 4 and 5, we also see that the classification accuracy of PSTWSVM-IWO and the computing time are better than PSTWSVM. Therefore, PSTWSVM-IWO is suitable for dealing with big data. Figure 2 and figure 3 shows that the optimization ability of IWO is very strong.

V. Conclusion And Furture Work

In order to improve the performance of STWSVM, seeking a better smoothing function is the key problem. In this paper, a novel version for smooth TWSVM, called polynomial smooth twin support vector machines (PSTWSVM), is proposed. Firstly, using the series expansion, a new class of polynomial smoothing is proposed, and then we prove their important properties. Subsequently, the polynomial functions are adopted to convert the original constrained QPPs of TWSVM into unconstrained minimization problems, and then are solved by the well-known Newton-Armijo algorithm. The experiments show that the proposed algorithm can obtain better classification than STWSVM. In view of the good optimization ability of Invasive Weed Optimization (IWO) algorithm, it is used to optimize PSTWSVM in this paper. And then we propose an algorithm called polynomial smooth twin support vector machines based on invasive weed optimization algorithm (PSTWSVM-IWO). We enhance our algorithm to deal with big data, the results indicate that PSTWSVM-IWO is a good method to deal with large datasets.

References

[1] L. Ding and F. Yu. "A classification algorithm for network traffic based on improved support vector machine", Journal of Computers, 2013, vol. 8, no.4, pp.1090-1096.

[2] Y.N. Zhang and S.F. Ding. "An algorithm research for prediction of extreme learning machines based on rough sets", Journal of Computers, 2013, vol. 8, no.5, pp.1335-1342.

[3] L.S. Yin and Y.G. He. "Adaptive chaotic prediction algorithm of RBF neural network filtering model based on phase space reconstruction", Journal of Computers, 2013, vol. 8, no.6, pp.1449-1455.

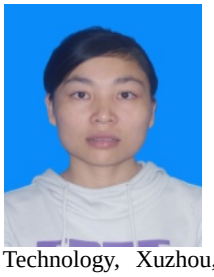
[4] D. Morariu and R. Cretulescu. "Improving a SVM Meta-classifier for Text Documents by using Naïve Bayes", International Journal of Computers Communications & Control, 2010, vol. 5, no.3, pp.351-361.

[5] Z.Y. Chen and Z.P. Zhi. "Distributed customer behavior prediction using multiplex data: A collaborative MK-SVM approach", Knowledge-Based Systems, 2012, vol.35, pp.111-119.

[6] F.Q. Shi and J. Xu. "Emotional cellular-based multi-class fuzzy support vector machines on product's KANSEI extraction", Applied Mathematics & Information Sciences, 2012, vol.6, no.1, pp. 41-49.

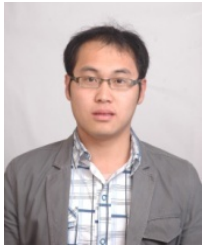
[7] C.S. Lo and C. M. Wang. “Support vector machine for breast MR image classification”, Computers & Mathematics with Applications, 2012, vol.64, no.5, pp.1153-1162.
[8] C. Cortes and V.N. Vapnik, “Support vector networks”, Machine Learning, 1995, vol.20, pp.273-297.
[9] E. Osuna and R. F. Girosi. “An improved training algorithm for support vector machines”, Proceedings of the 1997 IEEE Workshop on Neural Networks for Signal Processing. New York: IEEE Press, 1997, pp.276-285.
[10] J.C. Platt. “Using analytic QP and sparseness to speed training of support vector machines”, In M. Kearns, S. Solla and D.Cohn, Advances in Neural Information Processing Systems 11. Cambridge, MA: MIT Press,1999, pp. 557-563.
[11] O. L. Mangasarian. “Multisurface proximal support vector machine classification via generalized eigenvalues”, IEEE Transactions on Pattern Aalysis and Machine Intelligence, 2006, vol.28, no.1, pp. 69-74.
[12] Jayadeva, K. Reshma and S. Chandra, “Twin support vector machines for pattern classification”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, vol. 29, no.5, pp. 905-910.
[13] H. H. Cong, C. F. Yang and X. R. Pu. “Efficient Speaker Recognition based on Multi-class Twin Support Vector Machines and GMMs”, 2008 IEEE Conference on Robotics, Automation and Mechatronics, 2008, pp.348-352.
[14] X. S. Zhang and X. B. Gao and Y. Wang. “Twin Support Vector Machine For MCs Detection”, Journal of Electronics (China), 2009, vol.26, no.3, pp.318-325.
[15] M. Arun Kumar and M. Gopal. “Application of smoothing technique on twin support vector machines”, Pattern Recognition Letters, 2008, vol.28, pp.1842-1848.
[16] A.R. Mehrabian, C. Lucas, “A novel numerical optimization algorithm inspired from weed colonization”, Ecological Informatics, 2006, vol. 1, no.4, pp:355-366.
[17] Y. Q. Liu, S. Y. Liu and M. T. Gu. “Self-training Polynomial Support Smooth Semi-supervised Support Vetor Machines”, Journal of System Simulation, 2009, vol.21, no.18, pp.5740-5743.
[18] R. Bin and L. L. Cheng, “Polynomial smoothing support vector regression”, Control Theory & Applications, 2011, vol.28, no.2, pp.261-265.
[19] G.R. Gourab, D. Swagatam, C. Prithwish and N.S. Ponnuthurai, “Design of Non-Uniform Circular Antenna Arrays Using a Modified Invasive Weed Optimization Algorithm”, IEEE Transactions on antennas and propagation, vol.59, no.1, pp.110-118.
[20] F.M. Monavar and N. Komjani, “Bandwidth enhancement of microstrip patch antenna using jerusalem cross-shaped frequency selective surfaces by invasive weed optimization approach”, Progress In Electromagnetics Research, vol.121,pp.103-120.
[21] Mosek, http://www.mosek.com, 2007.
[22] O. L. Mangasarian and D. R, “Musicant, Lagrangian support vector machines”, J. Machine Learn. Res. 2001, vol.1, pp.161-177.

Shifei Ding received his bachelor's degree and master's degree from Qufu Normal University in 1987 and 1998 respectively. He received his Ph.D degree from Shandong University of Science and Technology in 2004. He received postdoctoral degree from Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences in 2006. And now, he works in China University of Mining and Technology as a professor and Ph.D supervisor. His research interests include intelligent information processing, pattern recognition, machine learning, data mining, and granular computing et al. He has published 3 books, and more than 150 research papers in journals and international conferences.



Huajuan Huang, born in 1984. Received her B.Sc.degree and M.Sc. degree in applied computer Technology from Guangxi University for Nationalities, Guangxi, China, in 2006 and 2009 respectively. Since 2011, she has been a Ph.D. degree candidate in applied computer Technology from the China University of Mining and Technology, Xuzhou, China. Her current research interests include data mining, parrern recognition and computational intelligence.

Junzhao Yu is currently a graduate student now studying at School of Computer Science and Technology, China University of Mining and Technology. He received his B.Sc. degree in computer science from China University of Mining and Technology in 2011. His research interests include pattern recognition, machine learning, and twin support vector machines et al.



Fulin Wu is currently a graduate student now studying in School of Computer Science and Technology, China University of Mining and Technology, and his supervisor is Prof. Shifei Ding. He received his B.Sc. degree in computer science from China University of Mining and Technology in 2012. His research interests include cloud computing, feature selection, pattern recognition, machine learning et al.

Principal Component Analysis Based Network Traffic Classification

Ruoyu Yan
College of Computer and Information Engineering, Henan University of Economics and Law, Zhengzhou, China
Email: hn_yry@163.com

Ran Liu
Software Institute, North China University of Water Resources and Electric Power, Zhengzhou, China
Email : liuran@ncwu.edu.cn

***Abstract*—At present traffic classification is widely concerned in many research fields such as network security, traffic scheduling and traffic accounting. How to identify network traffic fast and accurately is a very meaningful thing. But most machine learning based methods have a lower speed and efficiency, and can not guarantee their stability and usability. For this reason a Principal Component Analysis (PCA) based method is proposed in the paper. At first the method use Fast Correlation-Based Filter (FCBF) algorithm to filter training data set to obtain suitable flow attributes. Then these flow attributes are processed by PCA to build feature subspace for each flow class. After that a nearest neighbor rule is used to accurately identify flow class of testing traffic sample. In the end some experiments on public data sets are done to compare performance with some existing methods. The experimental results show that the PCA based method has higher accuracy, stability and faster speed than Naive Bayes (NB) estimation method and Naive Bayes Kernel (NBK) estimation method.**

***Index Terms*—traffic classification, Principal Component Analysis, network flow, Naïve Bayes, machine learning**

I. INTRODUCTION

With the increasing of Internet scale network traffic classification is more and more important in network security, traffic scheduling and traffic accounting etc [1-3]. Due to the emerging new network applications and application layer load encryption, the traditional traffic classification methods, such as port number match [4-6] and packet payload features analysis [7-8] can not well meet the needs of various network managements. In order to overcome the deficiencies of the two kinds of methods, researchers begin to study how to use machine learning method to classify traffic fast and accurately. At present there are many machine learning methods to classify network traffic [9-19]. In order to classify network traffic efficiently most of them have to solve two problems: one is how to select suitable traffic attributes set; the other is how to select suitable machine learning algorithm to build classification model. To extract suitable traffic attributes network flow is a common used object. The so called network flow is a unidirectional stream of packets with five tuples: destination IP, source IP, destination port, source port and layer 3 protocol type. After traffic attributes vector is determined, from the point of view of machine learning, the traffic classification problem can be described as: when the collection of network flow classes $T = \{T_1, T_2 ..., T_k\}$ and the collection of network flows $X = \{X_1, ..., X_n\}$ belonging to some of known classes are given, how to use machine learning method to process traffic features vector to construct traffic classification model $f : X \to T$ and then use the model to identify unknown class of traffic.

In [8] authors propose a traffic classification method based on network flow attributes. It mainly analyzes application layer load and application traffic class is identified by attributes extracted from application layer. Although the method is effective, analysis of application layer load not only consumes computation, but also may potentially cause user privacy disputes. Besides when the application layer load or attribute fields are encrypted, the method is usually useless. In [9] authors apply one-class SVMs to traffic classification and present a simple optimization algorithm for each set of SVM working parameters. In [10] authors apply SVM method to payload-based traffic classification. Although it is more accurate to classify traffic by searching application attributes in payload content, obtaining the attributes manually is very time consuming. In [11] authors then introduce Naive Bayes (NB) method based on probability model. The method requires that network flow attributes for classification must be conditional independence and follow a Gaussian distribution. But the original flow attributes hardly satisfy the above-mentioned conditions and overall accuracy of the method is only about 65%. In order to overcome the negative effects of conditional independence assumption and Gaussian distribution assumption, in [12] authors use FCBF algorithm and Kernel Estimation (KE) algorithm to improve on primitive Naïve Bayes method. Their experimental results have shown that the classification accuracy can increase to 90% or more.

doi:10.4304/jcp.9.5.1234-1240

In addition, in [13] authors firstly introduce K Nearest Neighbor (K-NN) and Linear Discriminant Analysis (LDA) to solve flow classification problem. But when K-NN method is used to process testing sample, the similarity between testing sample and training sample must be calculated one by one. This would not only lead to a larger processing overhead, but also the classification performance is easily affected by noisy data. In [14] authors put forward a method which combines Gaussian mixture model and spectral clustering to process network flow attributes. Their experiments show that the method has accuracy over 90%. To achieve this goal the quality of flow attributes must be guaranteed by the arrival sequence of data packets. But the dynamic routing and network congestion stop data packets arriving in order in real network environment [15]. So the stability and usability of the method can not be guaranteed. In [16] authors propose a traffic classification method based on Rough Set Theory and Genetic Algorithm. Due to limited conditions they only do some small-scale experiments on 2254 flow records and don't give a further comparison with the existing methods.

To sum up, most of these methods mentioned above are complicated, time consuming and inefficient, and especially the stability and usability can not be guaranteed. Although the Naïve Bayes method proposed by Moore et al. [12] has higher overall accuracy and is easily implemented, the method can not guarantee the stability of the classification result, which is verified in the following experiments. For this reason a Principal Component Analysis (PCA) based method is proposed in the paper. The method also use FCBF to filter training data set to obtain suitable flow attributes. Then PCA is applied to the processed training data set to build feature subspace for each flow class. Finally a nearest neighbor rule is used to accurately judge which class the testing traffic sample belongs to.

II. Traffic Classification Method Based on PCA

A. The Proposed Traffic Classification Model

The detection algorithm model based on PCA is shown in Fig.1. Note that testing data set and training data set are already preprocessed by FCBF method which is introduced in [5].

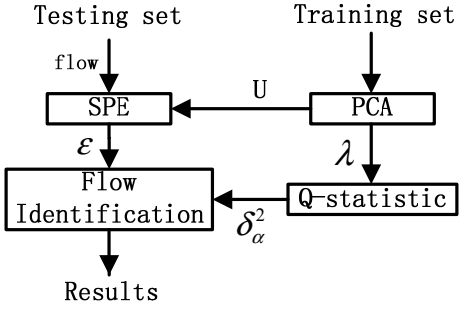


Figure 1. Flow classification model

PCA module: Principal Component Analysis is applied to each flow class of training data set, then for each flow class, the feature profile, namely the set of eigenvectors (U) and the set of eigenvalues (λ), is obtained. The subspace spanned by the eigenvectors is then regarded as the eigenspace of each class of flow data. Detailed steps to calculate U and λ are given in the next.

SPE module: Given a new data vector y to represent a testing flow, we project it onto a *k*-dimensional subspace U which represents a class of flow behavior and its reconstruction onto the subspace is written as $\hat{y} = UU^T y$. As PCA seeks a projection that best represents the original data in a least-square sense, the squared prediction error (SPE) is used in the experiments to measure the distance between original vector and its reconstruction vector: $\varepsilon = \left\| y - \hat{y} \right\|^2$. Then the SPE is a good metric to decide which class the testing flow belongs to.

Q-statistic module: This module applies a statistical test for the residual vector $\tilde{y} = y - \hat{y}$ to get a threshold namely δ_α^2 of SPE at the $1-\alpha$ confidence level. Detailed steps to compute δ_α^2 are given in the next.

Flow identification module: This module identifies flow class according to the feature profile of each flow class. Detailed steps to classify flow are given in the next.

B. Using PCA to Obtain Eigenvectors and Eigenvalues

Principal Component Analysis is one of the most widely used dimensionality reduction techniques to analyze and compress data. It transforms lots of variables into less uncorrelated variables by finding a few orthogonal linear combinations of the original variables with the largest variance. In the transformation, the first principal component is the linear combination of the original variables with the largest variance; the second principal component is the linear combination of the original variables with the second largest variance and orthogonal to the first principal component, and so on. The first several principal components always contribute most of the variance in the original data set, so that the rest can be discarded with minimal loss of the variance for dimension reduction of the data [20]. The transformation works as follows.

Step 1) Given a set of observations X_1 , X_2 ,..., , X_n in the training data set, where each observation is represented by a vector of length *p*, the training data set is represented by a matrix $X_{n\times p}$

$$X=\begin{bmatrix} x_{11} & x_{12}\cdots x_{1p} \\ x_{21} & x_{22}\cdots x_{2p} \\ \cdots & \cdots \cdots \\ x_{n1} & x_{n2}\cdots x_{np} \end{bmatrix} = \begin{bmatrix} X_1 & X_2\cdots X_p \end{bmatrix} \quad (1)$$

Step 2) Calculate the mean observation:

$$\mu = \frac{1}{p}\sum_{i=1}^{p} X_i \quad (2)$$

Step 3) Calculate the deviation from the mean:

$$\Phi_i = X_i - \mu \quad ,1 \le i \le n \quad (3)$$

Step 4) Calculate the sample covariance matrix of the training data set:

$$C = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)(X_i - \mu)^T = \frac{1}{n}\sum_{i=1}^{n}\Phi_i\Phi_i^T = \frac{1}{n}AA^T \quad (4),$$

where $A = [\Phi_1, \Phi_2, ..., \Phi_n]$

Step 5) Suppose $(\lambda_1, u_1), (\lambda_2, u_2), ..., (\lambda_p, u_p)$ are p eigenvalue-eigenvector pairs of the sample covariance matrix C, which can be computed by the Singular Value Decomposition (SVD) [21]. Then the k eigenvectors having the largest eigenvalues are chosen. It implies that k is the inherent dimensionality of the subspace governing the signal. The dimensionality of the subspace k can be determined by

$$\frac{\sum_{i=1}^{k}\lambda_i}{\sum_{i=1}^{p}\lambda_i} \ge \alpha \quad (5)$$

where α is the ratio of the amount of information in the subspace to the total amount of information in the original space. We arrange the set of eigenvectors (principal components) $(u_1, u_2, ..., u_k)$ as columns of a matrix $U = (u_1, u_2, ..., u_k)$ of size $p \times k$ where k denotes the number of normal axes and arrange the set of eigenvalues $(\lambda_1, \lambda_2, ..., \lambda_p)$ as columns of a vector $\lambda = (\lambda_1, \lambda_2, ..., \lambda_p)$.

C. Using Q-statistic to Obtain Threshold

Q-statistic is a useful statistical test for the residual vector $\tilde{y}$ to calculate threshold δ_α^2 for the SPE at the $1-\alpha$, which is developed by Jackson and Mudholkar and is given in [22] as:

$$\delta_\alpha^2 = \phi_1\left[\frac{c_\alpha\sqrt{2\phi_2 h_0^2}}{\phi_1} + 1 + \frac{\phi_2 h_0(h_0 - 1)}{\phi_1^2}\right]^{\frac{1}{h_0}} \quad (6)$$

Where $\phi_i = \sum_{j=r+1}^{m}\lambda_j^i \quad i = 1,2,3$, and $h_0 = 1 - 2\phi_1\phi_2/(3\phi_2^2)$. λ_j is eigenvalue, and c_α is the $1-\alpha$ percentile in a standard normal distribution. Jackson and Mudholkar's result holds regardless of how many principal components are retained in the normal subspace.

Note that in this setting, the $1-\alpha$ confidence limit corresponds to a false alarm rate of α, if the assumptions under which this result is derived are satisfied. An important property of this approach is that it does not depend on the mean traffic amount in the network. Thus, one can apply the same test on networks with different sizes and utilization levels.

D. Flow Identification Method

There are four steps to determine the class of testing flow:

Step 1) Use PCA module to profile behavior of each flow class based on training flow data. Suppose there are p classes of flow behaviors. Given n vectors X_1 , X_2 ,..., , X_n representing the observations of a class of flow data as an example, the average vector μ and each mean-adjusted vector can be computed by (2) and (3). p eigenvalue-eigenvector pairs $(\lambda_1, u_1), (\lambda_2, u_2), ..., (\lambda_p, u_p)$ of the sample covariance matrix C of the data set are then calculated. The number of principal eigenvectors $u_1, u_2, ..., u_k (k << p)$, which are used to represent the distribution of the original data, is determined by (5). Any training data vector belonging to a certain flow class can be approximately represented by a linear combination of k eigenvectors so that the dimensionality of the data is reduced, hopefully without sacrificing valuable information. The subspace spanned by the eigenvectors is then regarded as the eigenspace of the flow class.

Step 2) For each flow class i, use SPE module to calculate a SPE ε_i between the testing data vector and its reconstruction onto subspace of the flow class.

Step 3) For each SPE ε_i, use *Q-statistic* module to get its threshold $\theta_i = \delta_\alpha^2$ at the $1-\alpha$ confidence level.

Step 4) Identify the testing flow as a known flow class or a new flow class. For the testing vector to be identified, find the minimum ε_i. If the minimum ε_i is below its predefined threshold θ_i, the vector is then identified as flow class i. Otherwise it is identified as a new flow class.

III. Comparison and Analysis

A. Measurement Data Used

Data sets provided in [12] are used in the paper. Theses data sets are collected from a network which hosts several Biology-related facilities. These facilities have about 1,000 users connected to the Internet via a full-duplex Gigabit Ethernet link. Traffic was monitored to generate traffic-set for a full 24 hour period and for both link directions. Because the raw traffic is so huge, traffic sampling is used by Moore et al.. In order to construct the sets of flows, the day trace was split into ten blocks of approximately 1680 seconds each and the start of each

sample was selected randomly (uniformly distributed over the whole day trace). The fundamental object classified is a traffic-flow which is represented as a flow of one or more packets between a given pair of hosts. The flow is defined by five tuples consisting of the IP address of the pair of hosts, the protocol type and the port numbers used by the two hosts. In the case of TCP, a flow has a finite duration defined by the semantics of the TCP protocol. Here training and testing sets consist only of TCP and are made-up of semantically 2 complete TCP connections.

There are 377,526 flow samples altogether in data sets, which are classified into ten classes. Each flow has 249 attributes parameterizing its behavior, among which the last attribute marks its class. In the paper we only use five classes of flows as experimental data, as these flows account for about 98.8% of the entire flows. The five flow classes include: WWW, MAIL, BULK, DATABASE, SERVICE. The detailed application name, flow number and proportion of each class are shown in Table I. It can be seen that WWW and MAIL flow class has much larger flow number than BULK, DATABASE and SERVICE flow class, so WWW and MAIL flow class are called big flow class, and BULK, DATABASE and SERVICE flow class are called small flow class.

TABLE I.
INFORMATION ABOUT EACH FLOW CLASS

Flow class	Application	Flow number	Percentage (%)
WWW	Http,Https	328,091	86.91
MAIL	Imap, Pop2/3, Smtp	28,567	7.567
BULK	Ftp	11,539	3.056
DATABASE	Postgres, Sqlnet, Oracle	2,648	0.701
SERVICE	X11, Dns, Ident, Ldap	2,099	0.556

B. Evaluation Criteria

In this paper two metrics are used to assess the performance of different classification methods. In particular, refinements to those methods will be assessed on the basis of the two evaluation criteria.

Accuracy: The accuracy is the raw count of flows which are correctly classified divided by the total number of flows. This metric can be used to describe the classification accuracy for the whole system and can also provide an accuracy of classification on a per-class basis.

Trust: This is a per-class measure and it is an indication of how much the classification can be trusted. In other words, this measure is a probability that a flow that has been classified into some class, is in fact from this class. The trust value is higher for a flow class, the other class of flow samples is lesser classified as the class of flow by classification model.

Accuracy and trust of a particular flow class reflect the ability of the model to classify the flow class.

C. Experimental Results Under Stratified Sampling

At first, in order to compare the classification stability of PCA and Naïve Bayes method, the data set collected by Moore (called Moore_Set) is equally divided into two data sub-set. They are Set_1 and Set_2. In Set_1 and Set_2, the proportion of each class of flow sample is consistent with that in Moore_Set. For each flow class, 0.1% of the flow samples are taken from Set_1 to compose training set. In Moore_Set, each flow samples includes 249 network flow attributes, in which there are many redundancy attributes and irrelevant attributes. Due to the lower classification accuracy and heavier computation load of the classification model caused by theses attributes, FCBF method is used to filter training set. Then after NB, NBK and PCA method are applied to the preprocessed training set, for the three machine learning algorithms, flow classification model can be learned respectively. At last, for performance verification, each classification model is applied to testing set Set_2. For the purpose of analyzing the sensitivity of classification model to training data size, training data set is constructed respectively with the sampling scale gradually increasing to 0.5%, 1%, 5%, 10%, 50%. For each sale of training data set, the classification experiment process described before is repeated ten times. The final experimental results are shown in Fig. 2.

In Fig. 2, X-axis is Logarithmic coordinates, which represents the number of training flow sample. Y-axis represents overall accuracy of classification model. It can be shown that NB method has significantly poorer classification results. That is because NB method directly uses Gaussian distribution hypothesis which can not effectively fit the distribution of network flow attributes. Different from NB method NBK method and PCA method are able to maintain higher classification accuracy. But the classification accuracy of the two methods vibrates slightly with the increase of training data. This is mainly because FCBF algorithm filters flow attributes according to local information of the training data set, namely FCBF algorithm selects proper flow attributes under local optimality, which leads to the unstability of classification results. Besides, on the one hand NBK method needs the attribute filtering mechanism to satisfy conditional independence assumption, on the other hand the local optimality of NBK method can lead to the unstability of classification results. Therefore how to avoid local optimality and optimize flow attribute selection is still a need for further in-depth study.

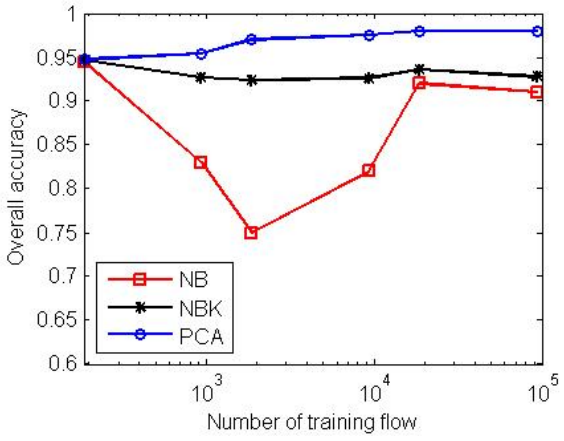


Figure 2. Overall accuracy with stratified sampling

Although Fig. 2 has shown the changes of overall accuracy of classification methods with the increasing of training data, in order to compare the classification accuracy and trust of each flow class, Table II and Table III show the classification results under stratified sampling when the training data size is 93236.

From Table II and Table III, it can be seen: when comparing the accuracy and trust of small flow class, such as SERVICE and BULK flow class, PCA method is significantly better than NB and NBK methods. This is because NB and NBK methods are all dependent on sample priori probability. That means bigger training set can improve the classification performance. However PCA method is independent on priori probability distribution, which can efficiently avoid this situation. In summery, PCA method is better than NB and NBK methods under stratified sampling when measuring from these performance indicators.

TABLE II.
ACCURACY OF ALL CLASSES WITH STRATIFIED SAMPLING

Method	WWW	MAIL	DB	BULK	SERV
NB (%)	97.95	87.52	1.35	22.50	8.24
NBK (%)	98.83	90.64	8.76	14.33	0.01
PCA (%)	99.27	96.38	91.42	89.10	65.17

TABLE III.
TRUST OF ALL CLASSES WITH STRATIFIED SAMPLING

Method	WWW	MAIL	DB	BULK	SERV
NB (%)	94.15	81.68	20.10	65.33	1.05
NBK (%)	94.60	76.82	48.70	95.68	9.21
PCA (%)	96.53	97.47	93.46	91.85	66.27

D. Experimental Results Under Uniform Sampling

In order to analyze the dependence degree of Naive Bayes to sample priori probability, and to further study the classification stability of PCA and Naïve Bayes method, some experiments are done under uniform sampling. Firstly, for each flow class 100 flow samples are randomly selected from Set_1 to compose a training set. That means in training set each class of flow numbers are equal. Likewise FCBF method is used to filter training data set. Then after NB, NBK and PCA method are used to process training data set respectively, flow classification model of each method can be built. Finally, for performance verification, each classification model is applied to testing data set Set_2. Again for each different flow sample number, such as 300, 500, 700 and 900, training data set is constructed as mentioned before. For each training data set, the experiments are repeated also ten times. The average experimental results are shown in Fig. 3.

In Fig. 3, X-axis represents the sample number of each flow class. Y-axis represents overall accuracy of classification model. It can be shown that, among the three methods, with the increasing of training data only the overall accuracy of PCA method improves in a relatively stable manner. However the overall accuracy of NB and NBK method does not increase, on the contrary decreases with the increasing of training data. This is not only due to the vibration caused by local optimality of FCBF, but the more important factor is the larger flow distribution differences between testing data set and training data set. Because NB method and NBK method, which are all based on Bayes' theorem, assume that a priori probability remains unchanged. When this assumption is unsatisfied, NB method and NBK method become invalid.

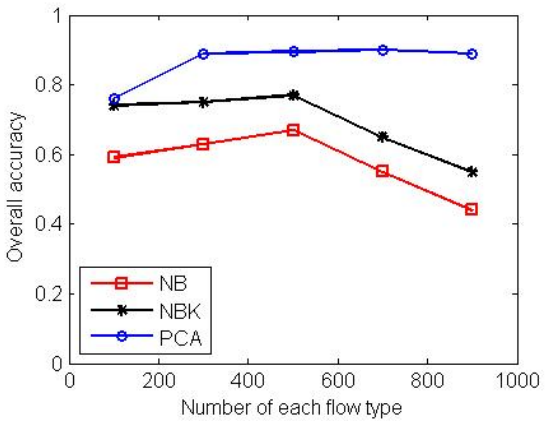


Figure 3. Overall accuracy with uniform sampling

For further comparison and analysis, Table IV and Table V show classification accuracy and trust for each flow class under uniform sampling when the training data composes 900 flow samples for each flow class.

TABLE IV.
ACCURACY OF ALL CLASSES WITH UNIFORM SAMPLING

Method	WWW	MAIL	DB	BULK	SERV
NB (%)	43.94	81.85	37.94	16.45	87.75
NBK (%)	55.16	86.70	64.68	12.73	95.36
PCA(%)	89.07	92.82	97.50	90.70	96.28

TABLE V.
TRUST OF ALL CLASSES WITH UNIFORM SAMPLING

Method	WWW	MAIL	DB	BULK	SERV
NB (%)	97.28	84.23	12.60	27.74	15.74
NBK (%)	97.53	87.45	15.82	67.38	27.36
PCA (%)	98.15	97.21	83.16	54.68	60.17

It can be shown that, under uniform sampling, the accuracy of identifying WWW flow all decreases compared to that under stratified sampling when the three classification models process flow sample respectively. Among them, the accuracy of PCA method identifying WWW flow decreases about 10%. This is because the number of WWW flow in training data set is only 900, which is much smaller compared to 200000 WWW flows in whole data set. Namely the relative deficiency of sampling causes the decrease of accuracy. However the accuracy of NB and NBK method identifying WWW flow decreases about 50%. This significant decrease is

not only because under-sampling losses many information, but also because uniform sampling losses priori probability information. In summery, PCA method is better than NB and NBK methods under uniform sampling when measuring from accuracy and trust.

E. Classification Speed Comparison

In the paper, FCBF algorithm is used to select suitable flow attributes before each flow classification experiment, but each time the selected attributes are different owing to the difference of training data sets. So it is hard to measure the speed of classification method. For this reason a group of fixed flow attributes should be selected. Hence FCBF algorithm is run on whole Moore_Set data to assure the selected attributes typical and seven flow attributes are obtained. Then 10% of Moore_Set data are randomly selected as training data set, and the rest of Moores_Set data belong to testing data set. Afterwards NB, NBK and PCA methods are run to process training data set respectively to obtain flow classification model. Finally for performance verification, each classification model is applied to testing data set. Above experiment is repeated ten times. The experiments results are shown in Table VI. The testing time is average identification time for one testing sample. Comparing the experimental results of the three methods, it is clear that PCA method not only has a shorter training time, but also has an absolute advantage in testing speed. The main reason is that when PCA method classifies flow samples, it only computes SPE value and executes a simple threshold comparison. However for NB and NBK methods, they need compute flow sample probability of belonging to each flow class, which is relatively complicated. In practical application, a good classification method should have fast speed to classify flow and need not build classification model frequently in large scale network. Compared to NB and NBK methods, PCA is clearly suitable for massive network traffic.

TABLE VI.
AVERAGE TRAINING TIME AND TESTING TIME OF 3 METHODS

Time	NB	NBK	PCA
Training Time (S)	0.43	0.46	0.33
Testing Time (μs)	51.23	134.71	4.13

IV . CONCLUSIONS

This paper proposes a method based on PCA and flow concept to classify network traffic. FCBF algorithm firstly processes training data set to select suitable flow attributes. Next PCA algorithm analyzes these flow attributes in training data set and builds feature subspace for each flow class. When classifying testing flow sample, SPE and Q-statistic algorithm compare the distance between testing flow sample and each feature subspace to judge flow class. Compared to NB and NBK method, experiments under stratified sampling and uniform sampling demonstrate that the method has higher accuracy, trust and faster speed. Furthermore the method has better classification stability under the changing in the training data set size. For the purpose of improving on network traffic online identification, in the future some research work should be done to collect only a smaller portion of data packets in one network flow to extract flow attribute. This will identify flow class in advance greatly.

ACKNOWLEDGMENT

This work is supported mainly by National Natural Science Foundation of China (61101211, 61202285), Foundation of He'nan Educational Committee (12A520005, 13B520901), Research Programme of He'nan Fundamental and Advanced Technology of China (132300410337) and The Heng Yang Joint Funds of The Hunan Provincial Natural Sciences Foundation of China (11JJ9010).

REFERENCES

[1] A. Sperotto, G. Schffrath, and R. Sadre, *et al.,* "An overview of IP flow-based intrusion detection," *IEEE Communications Surveys and Tutorials*, vol. 12, pp. 1-14, 2010.

[2] R. Y. YAN and Q. H. ZHENG, "Using Cross Entropy to Detect and Classify Network Anomalous Traffic," *Journal of XI'AN Jiao Tong University*, vol. 44, no. 6, pp. 10-15, 2010.

[3] Z, X. SUN and J. LIN, "Research of intelligent rule-base based on multilayer intrusion detection," *Journal of Computers*, vol. 4, no. 6, pp. 453-460, 2009.

[4] C. Fraleigh, S. Moon, and B. Lyles, et al., "Packet-level traffic measurements from the sprint IP backbone," *IEEE Network*, vol. 17, no. 6, pp. 6–16, 2003.

[5] R. Smith, C. Estan, and S. Jha *et al.*, "Deflating the big bang: fast and scalable deep packet inspection with extended finite automata," *in Proceedings of the ACM SIGCOMM 2008*, Seattle, USA, August 2008, pp. 207-218.

[6] A. Madhukar and C. Williamson, "A longitudinal study of P2P traffic classification," *in Proceedings of the* 14th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, 2006. MASCOTS, pp.179-188.

[7] S. Sen , O. Spatscheck, and D. Wang, "Accurate, scalable in network identification of P2P traffic using application signatures," *in Proceedings of the 13th international conference on World Wide Web*, New York, USA, 2004, pp. 512-521.

[8] A. W. Moore and K. Papagiannaki, "Toward the accurate identification of network applications", *In Proceedings of the Passive and Active Measurement Workshop*, Boston, USA, 2005, pp.41-54.

[9] Este, F. Gringoli, and L. Salgarelli, "Support Vector Machines for TCP Traffic Classification," *Computer Networks*, vol. 53, no. 14, pp. 2476-2490, 2009.

[10] Finamore, M. Mellia, M. Meo, and D. Rossi, "KISS: Stochastic Packet Inspection Classifier for UDP Traffic," *IEEE/ACM Transactions on Networking*, vol. 18, no. 5, pp.1505-1515, 2010.

[11] D. Zuev and A.W. Moore, "Traffic classification using a statistical approach," *In Proceedings of the Passive and Active Measurement Workshop*, Boston, USA, 2005, pp.321-324.

[12] A.W. Moore and D. Zuev, "Internet traffic classification using Bayesian analysis techniques," *In Proceedings of the*

2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, New York, USA, 2005, pp.50-60.
[13] M. Roughan, S. Sen, and O. Spatscheck, et al., "Class-of-service mapping for QoS: A statistical signature-based approach to IP traffic classification," *In Proceedings of ACM SIGCOMM Internet Measurement Conference 2004*, New York, USA, 2004, pp.135-148.
[14] L. Bernaille, R. Teixeira and K. Salamatian, "Early application identification," *In Proceedings of Conference on Future Networking Technologies 2006*, New York, USA, 2006.
[15] V. Paxson, "Measurements and analysis of end-to-end Internet dynamics", *Ph.D. Thesis*, Berkeley, California: University of California, 1997.
[16] N. Li, Z.L. Chen and G. Zhou, "Network Traffic Classification Using Rough Set Theory and Genetic Algorithm," *In Proceedings of ICIC 2006*, 2006, pp.945~950.
[17] R.O. Duda, P.E. Hart and D.G. Stork, "Pattern Classification," Beijing : China Machine Press, 2nd edition, 2004.
[18] G.H. Golub and C.F. Van Loan, "Matrix Computation," Baltimore : Johns Hopkins Univ. Press, 1996.
[19] J. E. Jackson and G. S. Mudholkar, "Control Procedures for Residuals Associated with Principal Component Analysis", *Technometrics*, 1979.
[20] J. TAN, X. S. CHEN and M. DU, "An internet traffic identification approach based on GA and PSO-SVM," *Journal of Computers*, vol. 7, no. 1, pp.19-29, 2012.
[21] L. DING, F. YU, S. PENG and C. XU, "A classification algorithm for network traffic based on improved support vector machine," *Journal of Computers*, vol. 8, no. 4, pp.1090-1096, 2013.
[22] Y. WANG and S. Z. YU, "Supervised Learning Real-time Traffic Classifiers," Journal of Networks, vol. 4, no. 7, pp. 622-629, 2009.

Ruoyu Yan was born in Shaoyang, China, on April, 1974. He received a M.S. degree from Beijing Jiaotong University in computer science, China, in 2004 and received a Ph.D. from Xi'an Jiaotong University in computer science, China, in 2010. Now he is a lecturer in Henan University of Economics and Law in China. His research interests include computer network, network security and database system.

Ran Liu was born in Xingyang city,Henan, P.R. China in December 23, 1979. He received master's degree in engineering from North China University of Water Resources and Electric Power, Zhengzhou, P.R. China. He is lecturer of computer science, and the director of multimedia technology department, the Software Institute, North China University of Water Resources and Electric Power, Zhengzhou, P.R. China. He has published more than 10 papers in Journal of Science and Technology. His research interests include data mining ,distributed database, and multimedia technology.

Study on Multi-document Summarization Based on Text Segmentation

Meng Wang
Assoc. Prof., Computer College, Guangxi University of Technology, CHINA
Email:mwang007@163.com

Xinlai Tang
Assoc. Prof., Lushan College, Guangxi University of Technology, CHINA
Email: lz_txl@163.com

Xiaorong Wang
Assoc. Prof., Computer College, Guangxi University of Technology, CHINA
Email:ccnuxnxs@163.com

***Abstract*—This paper introduces a novel approach of automatic multi-document summarization based on text segmentation. The approach acquires concepts with a How-Net oriented tool and calculates the importance degree of sentences by means of employing the improved DotPlotting model and establishing a sentence-based vector space model (VSM). A conclusion is made according to the importance degree and similarity of sentences. Experimental results show that the performance on ROUGE of the approach put forward hereby is effective and significant.**

***Index Terms*—Text segmentation, summarization, How-Net**

I. Introduction

Today electric text messages of various kinds have emerged in great numbers, with pages available on the Internet almost doubled every year in numbers [1]. For instance, in January 2012, the number of hosts advertised in DNS is 888,239,420 [2]. To help people cope with the ever-increasing text documents, advanced technologies facilitating text summarization have been developed. Automatic text summarization aims to automatically produce a short and well-organized summary of a single or multiple documents. As a fundamental and effective tool for document understanding and organization, the multi-document summarization enables better information services by creating concise and informative reports for a large collection of documents [3,4].

Recently, more and more models have successfully been applied to summarization. Jean-Yves Delort adopts an unsupervised probabilistic approach to model the novelty in a document collection and applies it to the generation of update summaries [5]. James Gung uses temporal information to improve abstracted summarization [6]. Kristian Woodsend adopts a method where such individual aspects are learned separately from data but optimized jointly by employing an integer linear program to abstract summary [7]. Seonggi Ryang presents a new approach to automatic text summarization based on Reinforcement Learning, which models the construction of a summary within the framework of reinforcement learning and attempts to optimize the given score function with the given feature representation of a summary [8].

Till now, there are two types of summarizations, one of which is the abstractive summarization and the other is the extractive one. Extractive summary consists of selecting important sentences and paragraphs etc. from the original document and combines them into a shorter form based on statistical and linguistic features of sentences. The abstractive summarization aims to represent main concepts and ideas of a document by paraphrasing the source document in a clear natural language. Most of the recent works have concentrated on the extraction summarization method where there are two main techniques for feature extract, namely, sentence-based and keyword-based text summarization [9]. The former identifies the most salient sentences in a document while the latter summarizes documents by topics. Each of the approaches is featured in a set of keywords.

In this paper, a special Chinese automatic summarization method is proposed on the basis of Text segmentation. The method consists of three main parts: 1) feature finding: using concepts as minimal semantic unit rather than words, and using HowNet as a tool to obtain concepts in the text. 2) Text segmentation: using an improved DotPlotting method to segment texts. The method not only gives consideration to the defects of traditional DotPlotting, but also improves the speed of text segmentation by using a concept matrix. 3) Automatic summary. According to the segmentation results, the system can obtain the summary of the text on the basis of similarity. Summarization evaluation metric ROUGE motivated by the MT evaluation metric is used. Experimental results indicate clear superiority of the proposed method over the traditional ones in the proposed evaluation scheme.

The rest of this paper is organized as follows: Section 2 describes how to obtain concepts by using HowNet;

doi:10.4304/jcp.9.5.1241-1246

Section 3 demonstrates text segmentation based on the improved DotPlotting method; Section 4 introduces the way to abstract summary; Section 5 presents some experiments and their numerical results.

II. CONCEPT-OBTAINED

A. Introduction to HowNet

How-Net is a knowledge database which has been released recently on the Internet [2]. In How-Net, the concepts expressed in Chinese or English are described and the relations between concepts and the attributes of concepts are revealed. This knowledge database is used as the resource of evaluating the sememe, for it can offer some useful information The format (which is defined as HowNet tool) can be described as follows:

NO. = serial number
W_X = word
G_X = part of speech
E_X = example of word
DEF = definition of word

Examples of lemma in HowNet can be represented as follows:

NO.=005987
W_X= blow up
E_X= the plane will blow up， Boat blowed up
G_X=V
DEF={FormChange| shape change:StateFin={Out Of Order| shatter}}

A lemma in HowNet presents part of speech and definition of a word. In the definition of words, the basic sememe ({FormChange|形变(shape change)}) and the related sememe ({StateFin={OutOfOrder|坏掉(shatter)}) are defined respectively. The former reflects the meanings of a word, while the latter represent the frame feature about a word. Both of the sememes can help to obtain the word concept in the text.

B. Concept Acquisition Based On HowNet

From the structure of How-Net, one learns that the DEF item expresses the meaning of words very well. Words with the same DEF item are regarded as sememes for they share the same word meaning. Words composed of a set of single elements are different words with the same concept. Two problems are processed when the actual concepts are being acquired. The first problem is: the obtained principle is the same DEF item in the process of obtaining a word concept for polysemous words. The actual sememe item of polysemous word cannot be distinguished, which then influences the accuracy of concept acquisition. The second problem is: the distinction of DEF is too strict, thus some related information will be probably missed if alignment-search depends entirely on DEF item.

We can firstly solve the selection problem of DEF item of polysemous words; the word concept is obtained by using improved DEF item. We use the ICTCLAS platform of ICT (Institute of Computing Technology, Chinese Academy of Sciences) to conduct words segmentation and part-of-speech tagging for the document. Some words will be deleted, such as prepositions, numerals and function words which have little influence on text summarization. Some key words will be extracted, such as nouns and adjective. The text with segmented and part-of-speech tagging will be obtained. There are two cases for selecting DEF item of polysemous words. One is that the part-of-speech of some polysemous words is varied in different contexts. The DEF items of these polysemous words can be determined by tagged part-of-speech. Another case is that the same part-of-speech in different DEF items for polysemous words, but different part-of-speech of words will pair up different words in different contexts. Just taking two words (NO. is 005987 and NO. is 005990) in HowNet for instance. The probable meaning of the first word usually adopts a grammar form of N+V, while the probable meaning of the second word usually adopts a grammar form of V+N. Different contexts will have different grammatical forms, and therefore DEF item of this type of polysemy can be determined in this view.

The detailed process is as follows:

1) The DEF item is redefined. The DEF item is extended to the union of contained basic sememe and relation sememe of this word. If the meaning of abstract sememe in How-Net is too large and broad, the abstract sememe will be filtered, such as “attribute”, “event” and “entity”.

2) The document model before concept acquisition is established by sentences. The document model is expressed as $S_j(W_1,W_2,\cdots W_n)$(the document contains the j sentence, with each sentence containing n words).

3) We scan the sentence where the vector space model is established. We assume that the scanning sentence is currently the j^{th} sentence.

4) We scan the word W_i of the sentence, and find the corresponding DEF item of the word. At the same time we scan the sentence to search whether some words have the same meaning as the sememe of the DEF item. If the search result is a negative one, we will tag the concept of the word W_i, and scan the next word W_i +1 of the sentence, and proceed to Step (4). When all the words of the sentence are scanned, we will scan the next sentence and proceed to Step (3). If the search result is a positive one, we will proceed Step (5).

5) The word W_k is extracted and then the corresponding DEF item of the word W_k is found. If the DEF item sememe word of the word W_k does not contain the word W_i, the word W_i’s concept and the word W_k’s concept will be tagged with the word W_i’s DEF item. If the DEF item sememe word of the word W_k contains the word W_i, we will compare the sememe distance of the two words in the DEF item. The DEF item of the word which is closer to the basic sememe will be selected as the concept of the two words. Then we scan the next word W_i+1, and proceed to Step (4). When all the words of the sentence are scanned, we will scan the next sentence, and proceed to Step (3).

When all the steps are completed, concepts of all words are contained. The word concept contained by the above method solves the digestion problem of the

polysemous words. Meanwhile, the words which have the same relationship in the same context are treated as a concept. This can ensure the orthogonal relation of each conceptual element in concept vector space model based on concept, and help generate high quality text summarization.

This paper selects the "H7N9 Bird Flu" as the topic, and downloads 100 documents from http://news.qq.com/zt2013/H7N9/ as the test corpus. Statistics concepts and words use concept-based method and word-based method respectively. The results show that compared with word frequency statistics method, the number of concept has reduced greatly by concept statistical algorithms. Only consider document collection frequency greater than 2 times words and concepts, the document include 2,869 words and 1,789 concepts. Moreover, 1,789 concepts include 3,345 words. This illustrates that more of the words could be included in less of the concept when the concept of statistical method is used. This cannot miss the word frequency statistics appearing in the small amount of articles and express is an important word concepts.

Figure 1 shows the comparison between the numbers of words and concepts using word vector space model and concept vector space model under three different themes. The number of concepts is significantly less than that of words.

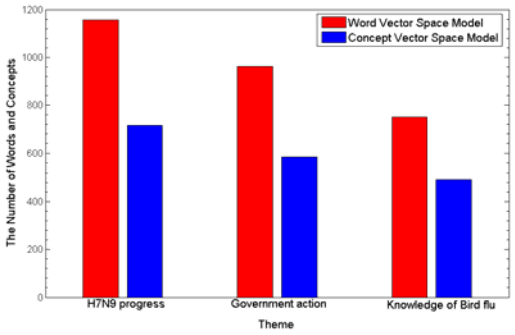


Figure 1. Number of words and concepts

III. Text Segmentation Method Based on Concept Density

A. The Traditional Doptlotting Model

The DotPlotting[10] is a famous model in the field of text segmentation. This method is based on the vocabulary degree of polymerization and the image analysis technology. It identifies the semantic paragraph border by point diagrams reflecting overall distribution of document vocabularies. If one word appears at x and y position repeatedly in the document, the word will be marked by a point in four coordinate points (x, x),(x, y),(y, x),(y, y) in the point diagram. Then, all vocabularies of the document will be marked in the point diagram. A symmetrical two-dimensional point diagram will be formed based on this method. The two-dimensional point diagram can clearly reflect sub-topic distribution of one document and measure theme consistency by establishing density evaluation function.

The density function of Dotplotting is expressed as:

$$f_D = \sum_{j=2}^{|P|} \frac{V_{P_{j-1},P_j} \cdot V_{P_j,n}}{(P_j - P_{j-1})(n - P_j)} \qquad (1)$$

Here n represents the length of the document. P_j represents the position of the jth semantic paragraph boundary. |P| represents the semantic paragraph number of the document. $V_{x,y}$ represents the word frequency vector of text fragment from the x^{th} word to the y^{th} word.

In the traditional DotPlotting model, if we only use the vocabulary as a basic semantic unit, the two-dimensional point diagram will have a lot of coefficient matrix, which will be unable to accurately abstract the border of some semantic paragraphs when density function evaluation is adopted. In the formula (1), every single density is $\frac{V_{P_{j-1},P_j} \cdot V_{P_j,n}}{(P_j - P_{j-i})(n - P_j)}$ .The density of each position P_j is to calculate the vocabulary similarity from its previous semantic paragraph to its back in all texts. So the density of each position P_j is determined based on its previous semantic paragraph border and the end position n of the document. That causes an asymmetry density function, resulting in completely different text segmentations between scanning the document from front to back and from back to front. Sincewe do text segmentation by evaluating density function in a symmetric two-dimensional point diagram, we must solve the problem of density function asymmetry for the traditional DotPlotting model.

B. Improved Dotplotting Model

From the analysis and research for the traditional DotPlotting model, this paper will use concepts of the second part instead of words to create symmetric two-dimensional point diagram. At the same time, density function is improved to solve the problem of density function asymmetry for the traditional DotPlotting model.

$$f_D^{'} = \sum_{j=2}^{|P|} \frac{V_{P_{j-1},P_k} \cdot V_{P_j,n}}{(P_j - P_{j-1})(n - P_j)} + \sum_{j=1}^{|P-1|} \frac{V_{0,P_j} \cdot V_{P_j,P_{j+1}}}{P_j(P_{J+1} - P_j)} \qquad (2)$$

The second part of formula (2) is "Backward" density which aims to solve the density function of symmetry. By modifying the formula (1), the "Backward" density of P_j is determined by next semantic paragraph boundary P_j+1 and the start position 0 of the document. This density function can get the same density function value whenever the document is scanned from front to back or from back to front.

C. Text Segmentation Algorithm

The semantic paragraph boundary determination method of the DotPlotting model is: If B is the established semantic boundary set, the remaining boundaries are candidate semantic boundaries; the remaining boundary set is the candidate boundary of the next round which is composed of the candidate boundary set C. For each candidate boundary i of C, P=B∪ {i}, we calculate the overall density by P division recording to the formula (2). We will select the overall density of the smallest candidate boundary as the next best semantic

paragraph boundary, and combine it with the set B. The specific description of the algorithm is as follows:

(1) For a given document W, we have to pretreat it. We acquire word concept according to the concept acquisition method of the second part, establish a two-dimensional point concept diagram, and determine the semantic paragraph partition number K.

(2) Initialize the semantic boundary set B as an empty set; each paragraph is a boundary which is seen as a candidate segmentation point. We establish a candidate boundary set C based on the candidate segmentation point and we use S to record the best segmentation variable.

(3) We repeat operations (4)-(5) from segmentation paragraph 1 to segmentation paragraph k.

(4) For each boundary candidate point i of the set C, P=B ∪ {i}, we calculate the overall density d by P division according to the formula (2). If dmin is greater than d, dmin =d. We will record S=i.

(5)The boundary S will be a target boundary added to set B. At the same time, S will be deleted from the candidate boundary set C.

Semantic paragraph boundaries are successively added in this algorithm. The end of natural paragraph in the document is set as candidate semantic paragraph segmentation point. We check each candidate boundary when selecting new semantic paragraph boundaries. We try to add each candidate boundary to the boundary set B and form the new boundary set P. We evaluate segmented mode composed of the boundaries from the new boundary set by density function. The candidate boundary which has the minimum value of density function is selected as a segmentation boundary and it is added to segmentation boundary set until the number of boundary is equal to K.

IV. Automatic Summary based on Text Segmentation

For those original documents, the system should exclude those useless words, such as prepositions, empty words and numerals etc during pretreatment, and only some important nouns and adjectives are treated. In this section, the proposed method will be introduced in details. The process of abstracting Summary by text segmentation is displayed in Figure 1.

(1) Calculate Importance of Concepts

We apply TF*IDF to assign weight to the individual concept and the importance of each concept is defined as follows:

$$Wd_{it} = TFd_{it} * \log \frac{N}{Nd_t} \qquad (3)$$

Wd_{it} in formula (3) is TF*IDF of concept t in the i-th document. TFd_{it} in formula (3) denotes concept frequency of t in the i-th document. N is the number of documents and Nd_t is the number of documents where t occurs.

(2) Calculate Importance of Sentences

After the CVSM S_j (C_1, W_{1j}; C_2, W_{2j}; C_n, W_{nj}) of all sentences in the text are established, the importance of each sentence is defined as follow:

$$W(S_j) = \lambda \frac{\sum_{i=1}^{n} F_{ij} \times w_i(d_t)}{M} \qquad (4)$$

Wherein $W_i(d_t)$ is the importance of C_i, F_{ij} is the frequency of appearance of C_i in sentence S_j, M is all the words that sentence S_j contains; λ is the correct factor when the sentence is at the beginning or ending of paragraphs. It is 1.5 in this system.

(3) Compute Similarity of Sentence

In order to avoid overlap sentences in summary, we work out the cross-sentence word overlap according to the following formula:

$$R_s = 2 * \frac{(\#\text{overlapping words})}{(\#\text{words insentence1} + \#\text{words in sentence2})} \qquad (5)$$

The system sets 0.7 as threshold. If R_s exceeds the value, we deem that each pair has the same semantic, and select higher sentence value as summary to remove the following score sentence.

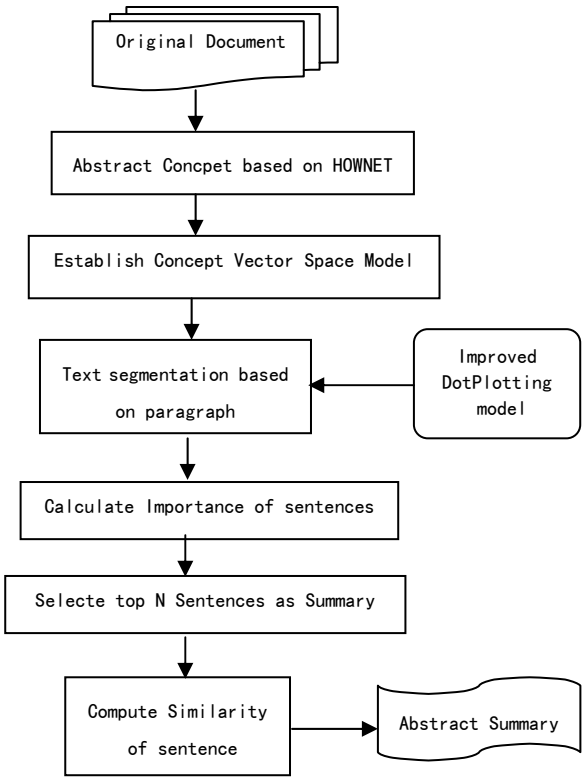


Figure 2. The process of Abstracting Summary

V. System Evaluations

A. Performance Evaluations

Generally, summaries can be evaluated by using intrinsic and extrinsic measures. While intrinsic methods attempt to measure the quality of summary by using human evaluation thereof, extrinsic methods measure the same through a task-based performance measuring such information retrieval oriented task. We adopt the former

to evaluate the quality of summarization by defining the following parameters for evaluation.

This system uses intrinsic evaluation method to verify the algorithm of this paper. Traditional intrinsic evaluation indexes mainly include recall rate, accurate rate and F-Score. At present, the intrinsic evaluation method is generally automatic summarization evaluation method ROUGE [11,12] as proposed by Lin Chin-Yew et al. This method has been gradually adopted in DUC automatic summarization evaluation since 2006, but the testing data of DUC is in English. However, when we conduct automatic summarization evaluation for Chinese texts, the according corpus must be established. After that, we use the ROUGE method to evaluate the text automatic summarization.

(1) We use three parameters which are recall, precision and F_measure to evaluate the summarization system. Recall refers to the ratio of accurate recognition by system; precision refers to the ratio of exact recognition. The formula: recall $R = N_{hm}/N_h$, precision $P = N_{hm}/N_m$, N_{hm} is the number of sentences abstracted by the summarization system and experts simultaneously, N_h is the number of sentences abstracted by experts and Nm is the number of sentences abstracted by the summarization system, $F\_Score = \frac{2 \times P \times R}{P + R}$.

(2) ROUGE stands for Recall-Oriented Understudy for Gisting Evaluation. There are five different ROUGE measures: ROUGE-N, ROUGE-L, ROUGE-W, ROUGE-S and Rouge-Su. Three of them have been used in the Document Understanding Conference (DUC) 2004, namely, Rouge-N, Rouge-S and Rouge-Su. ROUGE-N is an n-gram recall between a candidate summary and a set of reference summaries. ROUGE-S is Skip-Bigram Co-Occurrence Statistics, but Rouge-SU is an Extension of ROUGE-S which resolves the problem when ROUGE-S does not give any credit to a candidate sentence if the sentence does not have any word pair co-occurring with its references. This system will use Rouge-2 and Rouge-SU4 as evaluation criteria.

B. Evaluated Summary System

1 Coverage Baseline: choose the first sentence in first document, then choose the first sentence in the second document, and choose the first sentence in the n-th document; select the second sentence in the first document, then choose the second sentence in the second document…, until the summary is long enough. (Method 1)

2 Centroid-based summaries: this system is proposed by Dragomir R. Radev in Centroid-based summarization of multiple documents. (Method 2)

3 Text Segmentation based summary (TSS): the author describes the system. (Method 3)

C. Evluation Result And Analysis

Summary evaluation is a very important aspect for text summarization. Our evaluations on the three proposed summarization methods have been conducted based on a database of China's National Linguistics Work Committee which covers 200 articles covering economics, newspaper and literacy aspects.

We select three independent human evaluators which are employed to conduct manual summarization on the 200 documents contained in the evaluation database to obtain an objective summary. Each evaluator was requested to select exactly five sentences which he/she deems the most important for summarizing every document. Because of the disparities in the evaluators' sentence selections, 5 to 15 sentences in each document can be selected by at least one of the evaluators. Evaluation of recall, precision and F_measure parameter of each method are shown in Figures3-5.

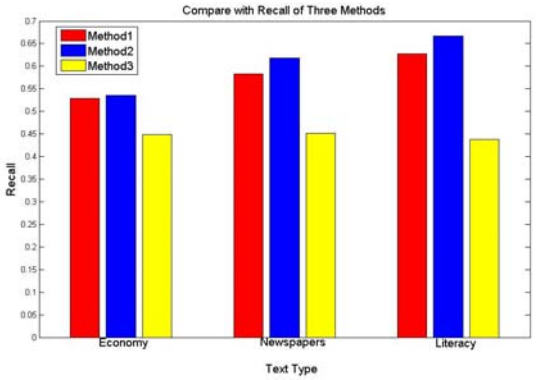


Figure 3. Result of Recall

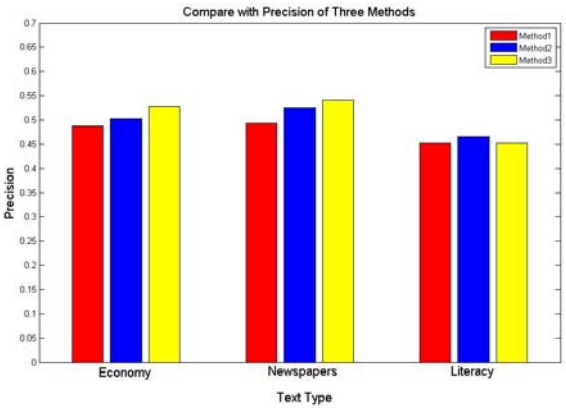


Figure 4. Result of Precision

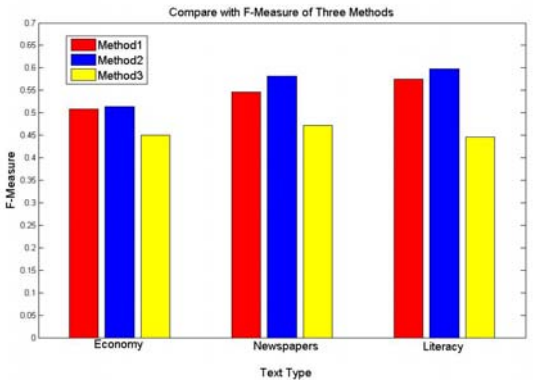


Figure 5. Rresult of F-Measure

From Figures 3-5, it distinctly shows that the economic type achieves a comparatively ideal result, while the newspaper and literature types shall be improved. The reason is that emphasis for texts of economy is explicit and the structure of those texts follows a regular pattern.

On the contrary, texts of newspapers and literature have no such features.

As is shown in Table 1, on this data set, the score of our system (TSS) is close to the mean score of all the participating systems (DUC2006) on ROUGE-2 and ROUGE-SU4. The mean scores of ROUGE-2 and ROUGE-SU4 are 0.0736 and 0.1288 respectively in DUC2006. TTS is slightly lower than the mean scores, but TTS generates a summary in Chinese instead of English. Different languages have different characteristics. The syntactic structure in Chinese is more complexly than that of English. In addition, most of participating systems (DUC2006) have adopted language tool, external corpora and knowledge database to help them understand the content of documents. This system uses basically statistical linguistics technology independent of any external resources, thus it isfaster and more independent.

TABLE 1.
RESULTS OF ROUGE-2 AND ROUGE-SU4

System Type	Rouge-2	Rouge-SU4
Method 1	0.0662	0.1112
Method 2	0.0691	0.1189
Method 3	0.0735	0.1281

From Tables 1-2, we can learn that TTS is more effective and efficient in performance than the other two systems, which means that using statistical linguistics technology to process documents can distinctly improve the quality of summary in a cost-effective manner.

ACKNOWLEDGMENT

This work is supported in part by a grant from Guangxi Natural Science Foundation (NO. 2013GXNSFBA019280) and Educational Commission of Guangxi Province of China (2013YB180).

REFERENCES

[1] D. R. Radev and W. Fan, “Automatic summarization of search engine hit lists”, Proceedings of the ACL Workshop on Recent Advances in Natural Language Processing and Information Retrieval, Hong Kong, 2000, pp. 99
[2] ISC “ISC Internet Domain Survey”, Available http://ftp.isc.org/www/survey/reports/current/
[3] Zhou Xi, Wang Li, Pan Fuping , Dong Bin and Yan Yonghong. Automatic Scoring for English Spoken Question and Answer [J]. International Journal of Advancements in Computing Technology, 2013, vol 5(3), pp 448-455.
[4] Zhu Junwu, Jiang Yi, Li Bin and Sun, Maosheng. Ontology-based Automatic Summarization of Web Document [J]. International Journal of Advancements in Computing Technology, 2012, vol 4(14), pp 298-306.
[5] Enrique Alfonesca, and Jean-Yves Delort. A topic-model Based Approach for Update Summarization. Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics ECAL 2012, Avignon, France, pp. 214–223.
[6] James Gung and Jugal Kalita. Summarization of Historical Articles Using Temporal Event Clustering. Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 631–635.
[7] Seonggi Ryang and Takeshi Abekawa. Framework of Automatic Text Summarization Using Reinforcement Learning. In Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (July 2012), pp. 256-265.
[8] Kristian Woodsend and Mirella Lapata. Multiple Aspect Summarization Using Integer Linear Programming. In Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (July 2012), pp. 233-243.
[9] Seniz Demir,Sandra Carberry and Kathleen F. McCoy. Summarizing Information Graphics Textually. Computational Linguistics (2012), pp. 527-574
[10] Reynar JC. Topic Segmentation: Algorithms and Applications [D], University of Pennsylvania, USA, 1998.
[11] Lin C Y. ROUGE: A Package for Automatic Evaluation of Summaries // Proc of ACL Workshop on Text Summarization. Barcelona,Spain, 2004: 74 – 81
[12] Lin C Y, Hovy E. Automatic Evaluation of Summaries Using N-gram Co-occurrence Statistics / / Proc of the North American Chapter of the Association for Computational Linguistics on Human Language Technology. Edmonton, Canada, 2003: 71 - 78

Meng Wang received the master’s degree in Computer Science and Technology from Central China Normal University, in 2005. He is a Ph.d. candidate of Wuhan University of Technology and Associate Professor at Guangxi University of Technology. His interests are in Natural Language Processing and Feature Selection.

Xinlai Tang received the master’s degree in Computer Science and Technology in Huazhong University of Science and Technology, in 2006. He is a Ph.d. candidate of Wuhan University of Technology and Associate Professor at Guangxi University of Technology. He is interested in Natural Language Processing.

Xiaorong Wang received the master’s degree in Computer Science and Technology from Central China Normal University, in 2005. She is a Ph.d. candidate of Wuhan University of Technology and Associate Professor at Guangxi University of Technology. Her interests are in Natural Language Processing.

Research on an Edge Detection Algorithm of Remote Sensing Image Based on Wavelet Enhancement and Morphology

Yu Xiong
Faculty of Land Resource Engineering, Kunming University of Science and Technology, Kunming China
Email:xiong.yu@foxmail.com

Jun Li, Xiaoqing Zuo and Zhenting Chen
Faculty of Land Resource Engineering, Kunming University of Science and Technology, Kunming China
Faculty of Computer Information, Kunming Metallurgy College, Kunming, China
Email: 1165978835@qq.com, zxq@163.com and cztgis@qq.com

Abstract—According to the basic features of high resolution remote sensing images, a novel edge detection algorithm based on wavelet enhancement and mathematical morphology is proposed. First, the remote sensing image is decomposed by a wavelet transform to get the low frequency part and high frequency part. Then a Laplacian sharpening operation is doing on the low frequency part to enhance the edge information. At the same time, the edge information of high frequency part is enhanced by means of linear enhancement after being denoised with threshold method. Last, the edge of image which has been enhanced is detected by using mathematical morphology. Experimental results indicate that our method can achieves better image processing effect than traditional method, has strong ability of eliminating noise and keeping clear image edge. It is more suitable for high resolution remote sensing image in edge detection and extraction.

***Index Terms*—high resolution remote sense image; edge detection; wavelet enhancement; morphology**

I. INTRODUCTION

With the rapid development of remote sensing technology, the amount of high resolution remote sensing data has a sharp increase. High resolution remote sensing image has the advantages of high spatial resolution, rich texture information and clear geometry of objects on earth [3, 4]. However, it is exactly the reason why the information of high resolution remote sensing image cannot be extracted automatically and effectively. As the edge focuses on the most information of an image, how to extract the edges effectively is extremely important for image-recognition [1,2].

Corresponding author: Zuo Xiaoqing, zxq@163.com

Edge detection technology can not only keep the structural information of boundary, but also can reduce the amount of data in image processing and simplify image-processing analysis process. There are many researches on edge detection at home and abroad, most of them are the algorithms based on operators and mathematical morphology, such as Sobel operator [5], Robert operator, Log Operator and Canny operator [6], they have the advantages of small calculating amount and fast speed. But they are seriously affected by noise, the accuracy and anti-noise ability can not be fused well. Besides, they are less efficient for detecting very low contrast edges. Some algorithms involving wavelet computation[2] might yield better results, but can be computationally intensive, and may not be suitable for low cost implementations. Hence there is a need to develop edge detection algorithm tackling effectively varieties of images.

For the problems proposed above, a novel edge detection algorithm based on wavelet enhancement and mathematical morphology is put forward. First, the remote sensing image is decomposed by wavelet transform to get the low frequency part and high frequency part. Then a Laplacian sharpening operation is doing on the low frequency part to enhance the edge information. At the same time, the edge information of high frequency part is enhanced by means of linear enhancement after being denoised with threshold method. Then we do an inverse wavelet transform operation on the low and high frequency part that are enhanced to get the enhancement image. Last, the edge of high remote sensing image is detected by using the mathematical morphology. Experimental results indicate that the new edge detection achieves better image processing effect than traditional method, has strong ability of eliminating noise as well as keeping clear image edge. It is more suitable for high resolution remote sensing image in edge detection and extraction.

doi:10.4304/jcp.9.5.1247-1252

II. BASIC PRINCIPLE

A Basic Theory of Wavelet Transform

When digital images are to be viewed or processed at multiple resolutions, wavelet transform is a mathematical tool of choice. The definition of a wavelet function is: assume $\psi(t)$ is a square and integrable function, thus $\psi(t) \in L^2(R)$. If the Fourier transform of the function $\hat{\psi}(w)$ satisfies the following conditions[8]:

$$C_\psi = \int_R \frac{|\psi(t)|}{w} dw < \infty \tag{1}$$

Then we can call that $\psi(t)$ is a basic wavelet. The definition of wavelet transform of function $f(x)$ is

$$\psi f(a,b) = (C_\psi |a|)^{-1/2} \int_{-\infty}^{+\infty} f(x) \psi \left[\frac{x-b}{a} \right] dx \tag{2}$$

The inverse wavelet transform is

$$f(x) = (C_\psi)^{-1/2} \int\int_{-\infty}^{+\infty} a^{-2} \psi f(a,b) (|a|^{1/2} \times \psi \left[\frac{x-b}{a} \right]) da db \tag{3}$$

For an image which size is $N \times N$, the wavelet decomposition process of it is shown as follows[1,2,7]:

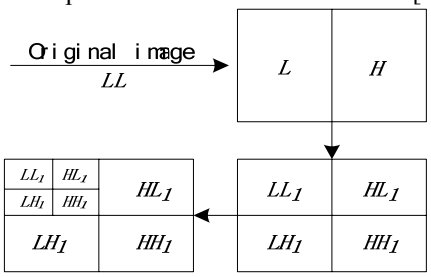


Fig.1 image wavelet decomposition schematic diagram

The image features that are described by 4 sub-bands LL,HL,LH,HH are shown as below:

(1) LL sub-band focuses on the original image of main energy and it mainly reflects low frequency part of the image.

(2) HL sub-band mainly reflects high frequency part of the horizontal direction of an image. As the edges of an image are focused on the high frequency part, HL sub-band has preserved the horizontal direction edge information well.

(3) LH sub-band mainly reflects high frequency part of the vertical direction of an image and it has preserved the vertical direction edge information well.

(4) HH sub-band mainly reflects high frequency part of the diagonal line direction of an image and it has preserved the diagonal line direction edge information well.

Fig.2 is a remote sensing image and Fig.3 is the result of doing a wavelet multi-resolution decomposition operation on Fig.2. Where the wavelet base type is 'coif3' and decomposition progression is '1'. Fig 3.(a),(b),(c),(d) are the low frequency information, high frequency information of horizontal, vertical, diagonal line direction of the original image respectively.

Fig.2 A remote sensing image

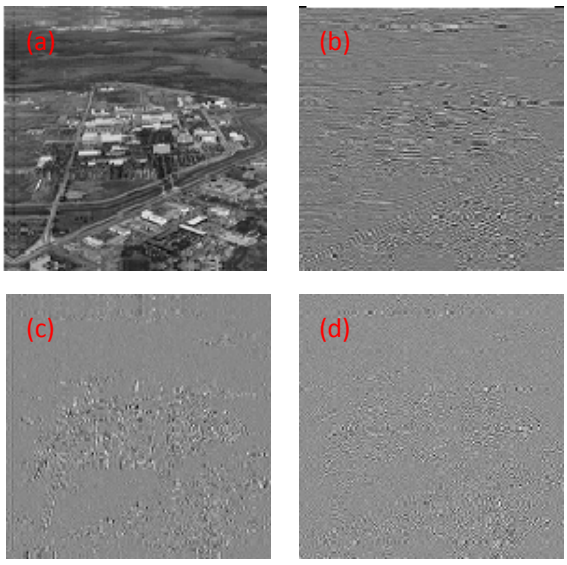


Fig.3 Wavelet decomposition result

B Theory of Laplacian Sharpening

As wavelet transform will do a low-passing filtering operation on the low frequency part of an image, the edge of low frequency part is blurred. As a result, the low frequency part will be sharpened by using Laplacian algorithm to enhance the edge information. It is better for the following edge detection work[9].

In the following part, we will introduce the definition of Laplacian operator. The Laplacian of an image $f(x,y)$, denoted $\nabla^2 f(x,y)$, is defined as

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \tag{4}$$

Commonly used digital approximations of the second derivatives are

$$\frac{\partial^2 f}{\partial x^2} = f(x+1,y) + f(x-1,y) - 2f(x,y) \tag{5}$$

and

$$\frac{\partial^2 f}{\partial y^2} = f(x,y+1) + f(x,y-1) - 2f(x,y) \tag{6}$$

so that

$$\nabla^2 f = [f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1)] - 4f(x,y) \tag{7}$$

Both derivatives sometimes are defined with the signs opposite to those shown here, resulting in masks that are the negatives of the preceding two masks. Enhancement using the Laplacian is based on the equation

$$g(x,y) = f(x,y) + c[\nabla^2 f(x,y)] \tag{8}$$

where $f(x,y)$ is the input image, $g(x,y)$ is the enhanced image, and c is 1 if the center coefficient of the mask is positive, or -1 if it is negative[1]. Because the Laplacian is a derivative operator, it sharpens the image[1,2]. In this paper, the Laplacian mask that we choose is [0, 1, 0; 1,-8, 1; 0, 1, 0].

C Theory of Wavelet Threshold Denoising

Noise and the edge information both focus on the high frequency part of an image, but we want to preserve the edge information when filtering the noise. For this, wavelet threshold denoising method is introduced to filter the noise exists in the high frequency part. Assume W and Wr are the values of the high frequency sub-band coefficients. W and Wr are before filtering and after filtering respectively. So we can define wavelet soft threshold function is [8]

$$\hat{w}_{j,k} = \begin{cases} \mathrm{sgn}(w_{j,k})(|w_{j,k}| - T) & |w_{j,k}| \geq T \\ 0 & |w_{j,k}| < T \end{cases} \tag{9}$$

where $\mathrm{sgn}(.)$ expresses symbolic function.

The threshold we choose is VisuShrink threshold. Method of VisuShrink was proposed in 1994 and it is also called the general threshold. The calculating formula for threshold T is [8]

$$T = \sigma_n \sqrt{2 \ln N} \tag{10}$$

Where, σ_n is the standard deviation of noise and N is the length or size of the signal. σ_n can be estimated by means of median estimation method that is put forward by Donoho and Johnstone[9]

$$\hat{\sigma}_n = \frac{Median(|Y(i,j)|}{0.6745} \tag{11}$$

D Edge Detection Based on Morphology

Mathematical morphology is a tool of extracting the components of an image and the components are very useful to express and describe the edge and shape of areas of an image. The gray-scale dilation of f by structuring element b, denoted $f \oplus b$, is defined as[10,11]

$$(f \oplus b)(x,y) = \max\{f(x-x', y-y') + b(x',y') \,|\, (x',y') \in D_b\} \tag{12}$$

Where D_b is the domain of b, and $f(x,y)$ is assumed to equal $-\infty$ outside the domain of f.

Also, the gray-scale erosion of f by structuring element b, is defined as

$$(f \Theta b)(x,y) = \min\{f(x-x', y-y') + b(x',y') \,|\, (x',y') \in D_b\} \tag{13}$$

Where D_b is the domain of b, and $f(x,y)$ is assumed to equal ∞ outside the domain of f.

Thus we can define the following morphology edge detection operators

$$Grad_1 = f \oplus b - f \tag{14}$$

$$Grad_2 = f - f \Theta b \tag{15}$$

The dilation type operator $Grad_1$ is mainly extracted the outside edge of an image. The realization process is first doing a dilation operation on the image and then getting the edge information by subtracting the original image from the image that is dilated. Similarly, the erosion type operator $Grad_2$ is mainly extracted the inside edge of an image. The realization process is first doing an erosion operation on the image and then getting the edge information by subtracting the image that is dilated from t he original image.

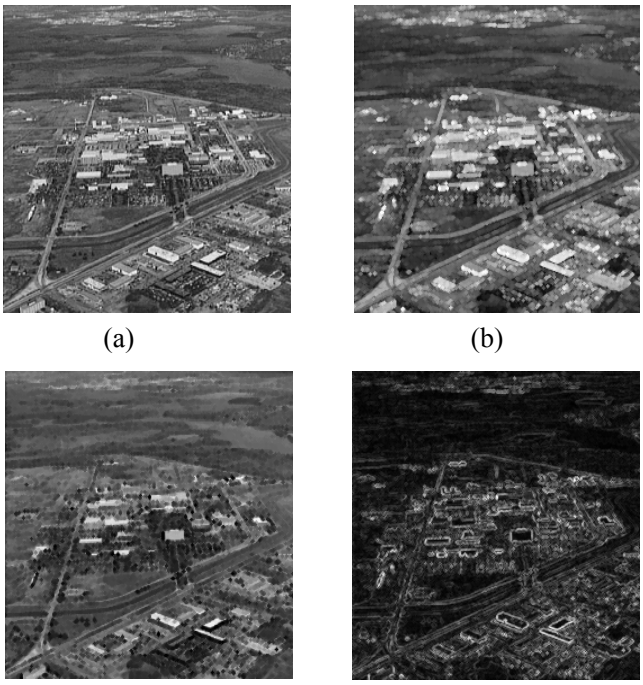
Dilation and erosion can be combined to achieve a variety of effects. For instance, subtracting an eroded image from its dilated version produces a "morphological gradient" [1], which is a measure of local gray-level variation in the image. It can be expressed as

$$Grad_3 = f \oplus b - f \Theta b \tag{16}$$

Dilation and erosion type operator $Grad_3$ is a combination of dilation operator and erosion operator. It computes the difference between maximum gray intensity and minimum gray intensity of a specific area. It has the advantages of simple calculation and geometry meaning

definitude.

Fig.4 shows the results of dilation and erosion and morphological gradient. Fig.4(b) is dilating Fig.4(a) using a flat 3*3 structuring element([0 1 0; 1 1 1; 0 1 0]). Fig. 4(c) is the erosion result. Fig.4 (d) is the morphological gradient of the image in Fig.4 (a).



(a) (b)

(c) (d)

Fig.4 morphological experiment results

III. OUR METHOD

In order to get a better edge image of a high-resolution remote sensing image, according to the theory above, the paper proposed a novel edge detection algorithm based on wavelet enhancement and morphology. It mainly consists of six procedures.

1. The remote sensing image is wavelet transformed to get low and high frequency coefficients of the image. The wavelet base type is 'coif3' and decomposition progression is 3.

2. Enhance the low frequency coefficients with a Laplacian filter. The Laplacian filter is 3*3 with a -8center.

3. For the high frequency coefficients, firstly, we use wavelet threshold de-noising method that is introduced in 2.3 to remove noise. The threshold function we choose is Eq.(10) and threshold is Eq.(11).

4. After denoised process, a linear enhancement operation will be done on high frequency coefficients to enhance the edge information of the high part.

5. Compute the inverse transform with the low frequency coefficients and high frequency coefficients that are processed by step 2 and 4 respectively. Then we can get the image that is wavelet enhanced [12,13].

6. Extract the edge image from the wavelet enhancement image using dilation and erosion type operator $Grad_3$. The flat structuring element is 3*3 ([0 1 0; 1 1 1; 0 1 0]).

Fig.5 shows a flow chart representing the algorithm above.

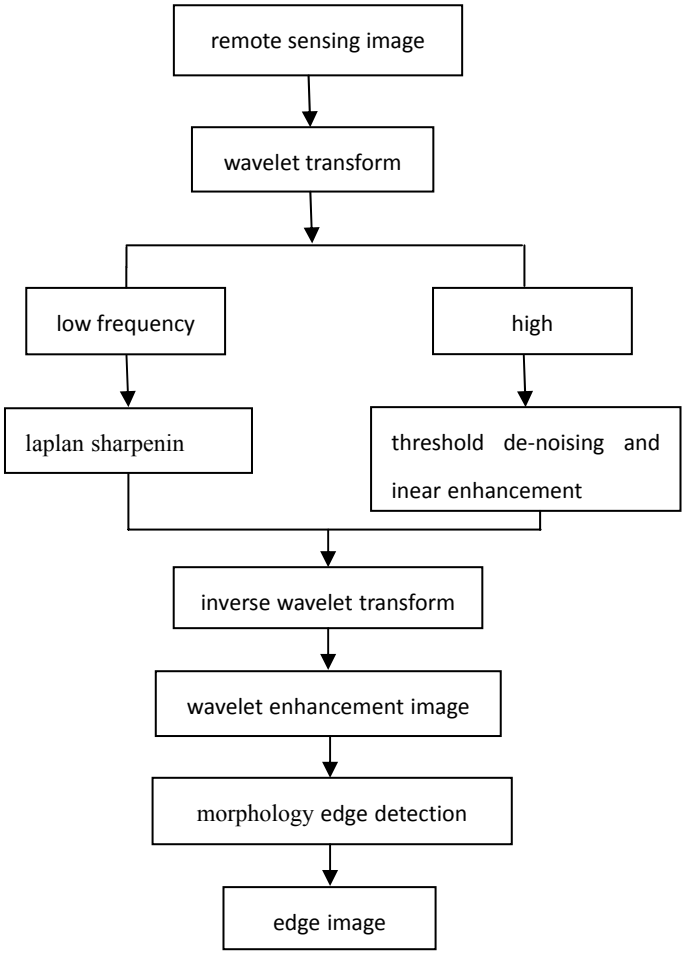
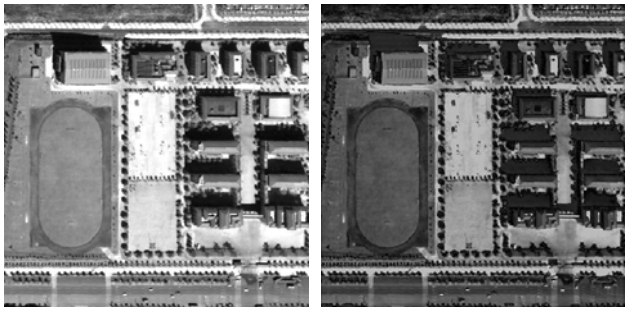


Fig.5 flow chart of our method

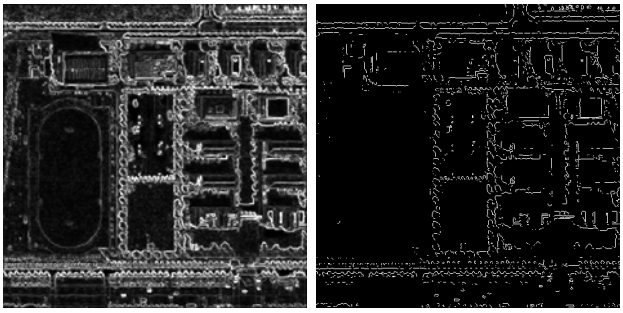
IV. EXPERIMENT VERIFICATION

A. Verification of Performance of Our Method

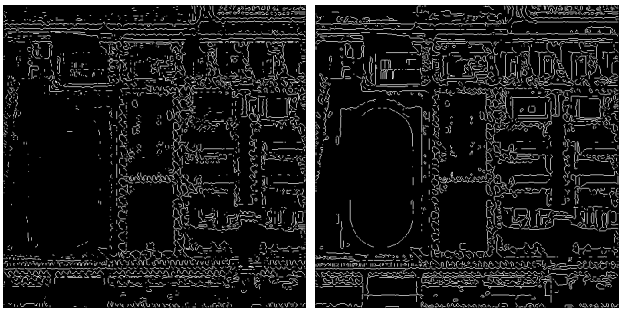
To verify the performance of our method, we choose a Quick-bird multispectral image with a resolution of 0.61 m. The experimental area contains various features and the size is 456*456 pixels. Fig.6 (a) is the original remote sensing image, Fig.6(b) is the wavelet enhancement image. Fig.6(c) is the result of using our method. Fig.6 (d),(e),(f) are the results of using traditional Sobel operator, Log operator and Canny operator respectively.



(a) (b)



(c) (d)

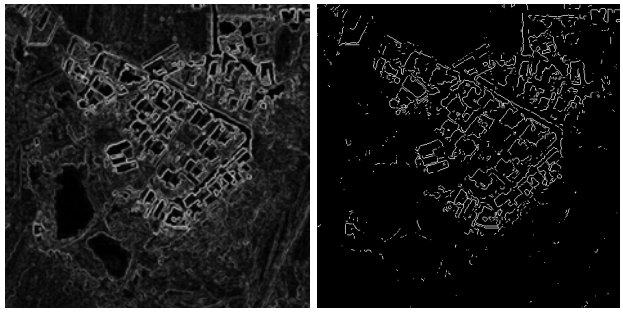


(e) (f)

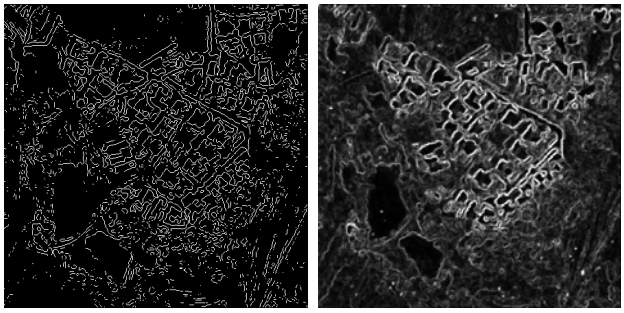
Fig.6 Experimental results

To further certify the performance of our method, we choose another Quick-bird multispectral experimental area with the resolution of 0.8m. Similarly, Fig.7(a) is the original remote sensing image, Fig.7(b) is the wavelet enhancement image. Fig.7(c) is the result of using our method. Fig.7(d),(e),(f) are the results of using traditional Sobel operator, Log operator and Canny operator respectively.

(a) (b)



(c) (d)



(e) (f)

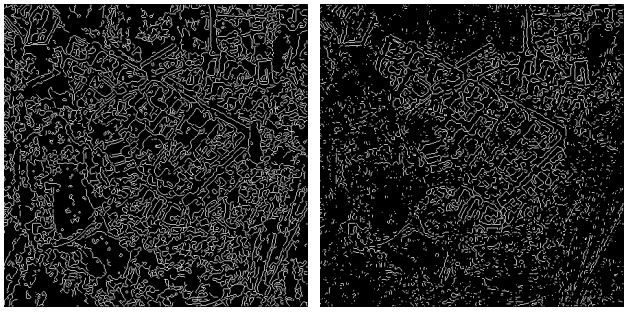
Fig.7 Experimental results

From the two experimental results we can see that the detection results (Fig.6 (d), Fig.7 (d)) of using Sobel operator have less integrated edges and a high rate of miss detection. Compared with Fig.6(d) and Fig.7(d), the results of using Log operator(Fig.6(e) and Fig.7(e)) are better, but there still exist some miss detection. Although Fig.6 (f) and Fig.7 (f) have good integrity, there exist some false detection. The results of using our method (Fig.6(c) and Fig.7(c)) have the best detection performance and least rate of miss detection. By contrast, they can detect some weak edge information. The method proposed by the paper has the highest location rate, the best edge connection degree and the clearest details of edges.

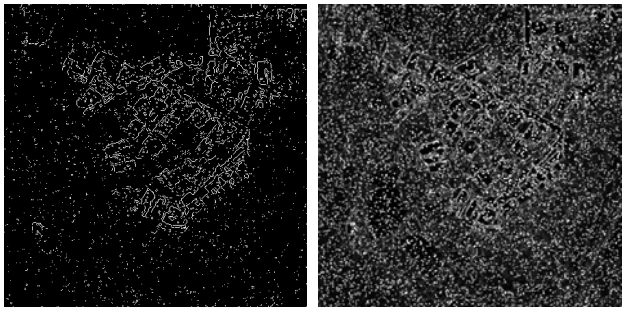
B. Verification of Anti-noise Property of Our Method

To verify the anti-noise property of our method, we add salt-and-pepper noise with probability 0.05[15]. Fig.8 shows the results of the anti-noise property of our algorithm and some traditional algorithm. Fig.8 (a) is the image corrupted by salt-and-pepper noise with density 0.05. Fig.8 (b) is the result of extracting edge image from Fig.8 (a) by using our method. Fig.8(c), (d), (e) are the results of using Sobel, Log and Canny operator. Fig.8 (f) is the result of using morphological edge detection operator $Grad_3$.

(a) (b)



(c) (d)



(e) (f)

Fig.8 Experimental results

From the experimental results, we can clearly see that the results by using the traditional algorithms have been influenced by the noise by different degrees and have plenty of false detection, especially the method by using morphological edge detection operator $Grad_3$ (Fig.8 (f)). The results above show that our method has a good anti-noise property.

C. Objective Evaluation of Our Method

The results above are all some subjective evaluation. The following we will give the objective evaluation of our method. The peak signal-to-noise ratio (PSNR) is usually used as a major quality assessment in the image processing system. Assume the size of an image is M*N, then PSNR can be expressed as[14]

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} [f'(i,j) - f(i,j)]^2$$

$$PSNR = 10 \lg(\frac{255^2}{MSE}) \tag{17}$$

where $f'(i,j)$ is the image after processing and $f(i,j)$ is the original image.

The PSNR of different algorithms are shown in Table 1 as below:

TABLE 1.

PSNR OF DIFFERENT METHODS

Experimental figure	Our method(dB)	Sobel(dB)	Log(dB)	Canny(dB)
Fig.6(a)	7.4936	5.0117	5.0211	5.023
Fig.7(a)	7.3737	5.1591	5.1623	5.1631
Fig.8(a)	7.5016	5.0214	5.0247	5.0266

From Table 1 we can learn that the PSNR value of our method is much higher than that of any other algorithms. It further proves our method has a good performance.

V .CONCLUSION

According to the basic features of high resolution remote sensing images, a novel edge detection algorithm based on wavelet enhancement and mathematical morphology is presented. And some experiments were taken to validate it. The experimental results show that our method can achieve the better edge detection result and has a good anti-noise property. It is more suitable for high resolution remote sensing image edge detection. However, our method is only applicable to gray image, how to develop an algorithm that is available to color image is the focus of our future work.

REFERENCES

[1] C.Gonzalez, E.Woods, L. Eddins, "Digital image processing using Matlab", Publishing House of Electronics Industry. Beijing, 2011, pp.70-80. (in Chinese)

[2] D Marr, E Hildreth., "Theory of edge detection", In: proc Roy Soc, London, 1980, pp.187-217.

[3] A. Panetta, J. Wharton, S. Agaian, "Logarithmic Edge Detection with Applications", Journal of Computers, 2008,9(3), pp.11-19

[4] LI Hui,XIAO Peng-Feng,FENG Xue-Zhi,LIN Jin-Tang, "Edge detection of high-resolution imagery by integrating spectral and scale characteristics", Journal of Infrared and Millimeter Waves, 2012, 31(5),pp.469-474(in Chinese)

[5] Tyagi, Y, Runtambekar, T.A, Sexena, Preeti, Tanwani, Sanjay, "A Hybrid Approach to Edge Detection Using Ant Colony Optimization and Fuzzy Logic", International Journal of Hybrid Information Technology, 2012, 5(1),pp.37-46.

[6] ZHU Chang-sheng,GUAN Jian,ZHOU Wei,HE Dong-liang, "Parallel Feature Extraction for SAR Images Based on Edge Detector", Opto-Electronic Engineering, 2010, 37(9),pp.44-50

[7] JIA Cheng-li, KUANG Gang-yao, "An Improved Edge Detection Algorithm for SAR Images", Journal of Electronics & Information Technology, 2007, 29(2), pp.379-382.

[8] Donoho DL and Johnstone IM, "Adapting to unknown smoothness via wavelet shrinkage", J. of American Statististic Assocciation, 1995,90(432),pp.1200-1224

[9] D.L.Donoho, I.M.Johnstone, "Ideal spatial adaptation by wavelet shrinkage", Biometrika, 1994,81(3),pp.425-455.

[10] Yuanni Wang, Fei Ge, "Adaptive order morphology edge detection", Journal of Computers, 2012,4(7),pp.846-850

[11] Xun liu, Zhisheng You, "An Image Edge Detection Method Based on Multi-Scale Morphology", Opto-Electronic Engineering, 2003, 30(3),pp.56-58.(in Chinese)

[12] GAO Li,LING Xiao-ming, "Color Edge Detection Based on Mathematical Morphology in HSI Space", Opto-Electronic Engineering, 2010, 37(4),pp.125-129.

[13] Hongjiu Tao, Jian Liu, Jinwen Tian, "Remote sense image edge detection based on wavelet transform and mathematical morphology", Infrared and Laser Engineering, 2002, 31(2),pp. 154-157.(in Chinese)

[14] S.D.Chen, and K.Z.Liu, "Color image edge features and performance evaluation of face detection," Journal of Software, 2005,8(5), pp.727-731.

[15] Deng, Shaojianqg, Tian Yuan, Hu Xipeng, "Application of new advanced CNN structure with adaptive thresholds to color edge detection", Communications in Nonlinear Science and Numerical Simulation, 2012, 17(4), pp.1637-1648.

Modeling of Stripper Temperature based on Improved T-S Fuzzy Neural Network

Shuzhi Gao
School of Information and Engineering, Shenyang University of Chemical Technology, Shenyang, China
Email: szg6868@126.com

Yihao Zhang
School of Information and Engineering, Shenyang University of Chemical Technology, Shenyang, China
Email: yihaozhang310@163.com

Xianwen Gao
School of Information Science and Engineering, Northeastern University, Shenyang, China
Email: gaoxianwen@ise.neu.edu.cn

***Abstract*—In the Polyvinyl Chloride (PVC) industry, proper control of stripper temperature is directly related to product quality of PVC resin. Considering multivariable, strong coupling, nonlinear and time-varying characteristics of the temperature control system for PVC stripper the current modeling method is difficult to obtain a relatively accurate mathematical model. Then, this paper studies the stripper temperature modeling method based on improved T-S fuzzy neural network, and proposes new nearest neighbor clustering fuzzy rules. In order to improve the learning performance, hybrid learning algorithm based on T-S fuzzy neural networks is developed. As for non-linear layer parameters, conjugate gradient algorithm is applied, while recursive least squares algorithm is adopted to handle the linear parameters. Simulation results demonstrate the effectiveness and accuracy of the proposed modeling method given.**

***Index Terms*— Stripper, T-S fuzzy neural network, New nearest neighbor clustering algorithm, least square method modeling**

I. INTRODUCTION

PVC (Polyvinyl Chloride) resin, a kind of general-purpose plastic, is a resultant from polymerization reaction of vinyl chloride monomer (VCM) [1]. Due to toxicity of vinyl chloride monomer in PVC resin products, the concentration of VCM residual must be properly controlled within a certain range. For this purpose, stripping process is normally developed to strip such harzardous elements. It is also recognized that stripping process is a typical and complex industrial process, along with high nonlinearity and time-varying characteristics. Hence, it is critical to accurately model the stripping process and to manipulate the process with advanced control method, so that good quality of PVC resin, low production cost and satisfactory environmental protection could be achieved.

On the other hand, PVC stripping process is a multivariable, strong coupling and nonlinear time-varying industry process, while temperature is one of the most important impact factors of stripping efficiency. However, it is difficult to analytically model the temperature variation of the stripping operation[2][3], while few research efforts on this topic have been reported in the literature. It is considered that T-S fuzzy model is essentially a nonlinear methodology, which could approximate any nonlinear system precisely, and is able to identify the process with high modeling accuracy. Meanwhile, as characterized by neural networks and fuzzy logic system, it can well describe the complex dynamic characteristics of the multi-variable system by a few rules. Currently, modeling based on T-S fuzzy system has been successfully applied to real industries, such as thermal process of gas temperature[4], distributed coke oven gas collector pressure[5] and system sizing reactor temperature[6].

Therefore, this paper will study temperature modeling method of PVC stripper based on T-S fuzzy model, and will present an improved T-S fuzzy model with established novel fuzzy rules for new and nearest neighbor clustering algorithm. Firstly, the importance of stripper temperature modeling is analyzed. Secondly, main input parameters of the system are obtained by analyzing various parameters that affect the stripper temperature. Next, the structure of T-S fuzzy neural network is described, and the fuzzy rules of the improved nearest neighbor clustering algorithm are determined. Finally simulation results demonstrating feasibility of the proposed modeling method are given.

II. PROBLEM FORMULATION

Manuscript received June 11, 2013; revised June28, 2013;
Corresponding author: Gao Xianwen
This work was supported by the Key Program of National Natural Science Foundation of China (61034005).

doi:10.4304/jcp.9.5.1253-1258

As discussed, RMNs are extended from Markov networks to the relational setting. We review Markov networks first, and then RMNs. The definitions and equations mentioned in this section are mainly taken from.

The theoretical and practical basis established by Berens for describing the VCM diffusion through a PVC matrix in suspension and emulsion resins successfully treats desorption phenomena at low VCM concentrations as a diffusive fickean process associated with the resin grain primary particles diameter. Desorption kinetics follow a Fickean model:

$$\frac{dC}{dt} = -k(C - C_{eq}) \tag{1}$$

Where k is the desorption constant, h^{-1}, C is the average VCM concentration in the resin, weight fraction, C_{eq} is C for resin at equilibrium with VCM existing the gas phase outside the resin particles.

Berens found that the diffusivity of VCM in PVC is a function of temperature following an Arrhenius-like relationship and is independent of molecular weight at least in the commercial resin range. The desorption constant can be calculated from Berens data and primary particle diameter measurements, and follows the same kind of relationship to temperature.

$$k = k_0 e^{-(\frac{E_a}{RT})} \tag{2}$$

Where E_a is the activation energy, cal/mol, R is ideal gas constant, cal/mol K, T is the temperature, C° .

The column trays are considered as perfectly mixed tanks and no effect of residence time distribution is taken in account so the tray-to-tray VCM in resin concentration changes are given by

$$\frac{C_{i+1} - C_{eqi}}{C_i - C_{eqi}} = e^{(-kt_r)} \tag{3}$$

$$t_R = \frac{\pi D^2}{4}\frac{F_A h_L}{Q} \tag{4}$$

Where t_R is the residence time, h, D is column diameter, m, F_A is the fraction of the column section not occupied by downcomer, h_L is liquid height on tray, m, Q is the tray volumetric liquid/solid feed, m^3 / h.

It is known that exposure to high temperatures, especially above 100°C, causes coloring of the resin due to polymer degradation with HCl evolution. This is a major difficulty in removing VCM from PVC as high color is quality non-conformity. Arlman (1954) presented a correlation for HCl evolution from PVC in the absence of oxygen.

$$E_{HCL} = \frac{b}{M_n} \tag{5}$$

Where E_{HCL} is HCl evolution rate, μmol HCl / gPVC-h, $b = 7 \times 10^5$ μmol HCl / gPVC-h (at 182°C), M_n = PVC molecular weight

Arlman found for this process an activation energy of 138 KJ/mol, so E_{HCL} can be put as a function of temperature and polymer molecular weight:

$$E_{HCL} = \frac{k_0}{M_n}\exp(\frac{-E_a}{RT}) \tag{6}$$

Here $k_0 = 1 \times 10^{10}$ μmol HCl / gPVC-h, E_a = 33 000 cal/mol.

Associating the HCl evolution to resin damage, a convenient multiple of the amount of HCl evolved during stripping can be used as damage index for comparison between different operations modes. Moreover, based on previous analysis, it can be seen that temperature plays a key role in stripping process, because it affects both VCM desorption and resin damage. Therefore, proper temperature control is becoming crucial for PVC stripping, and an accurate modeling of stripping process is desired.

III. T-S FNN-BASED MODELING

A. Data Preprocessing

During the process of PVC stripping, due to the accuracy of measurement instruments, methods employed and human factors, various errors can not be avoided, which will affect process modeling and control quality. In this paper, gross error detection, measurement error test, filtering and other methods are used in data preprocessing of process variables. Then, normalized method is adopted to eliminate the different magnitude of the model structure and parameters in the process variable data, as follows:

$$a_{ji} = \frac{a_{jio} - a_{j\min}}{a_{j\max} - a_{j\min}}, \tag{7}$$

Among them a_{ji}, a_{jio}, $a_{j\max}$, $a_{j\min}$ are for normalized value, actual value, maximum value, minimum value of each group respectively. The method guarantees the data are between [0, 1].

B. Structure of T-S fuzzy Neural Network

The method of T-S fuzzy modeling has received increasing attention in recent years [7]-[9]. Takagi and Sugeno proposed a T-S model in 1985, called the Sugeno fuzzy model. The post-condition of fuzzy rules is a linear combination of input variables. Without loss of generality, multi-input and multi-output system can be viewed as several multi-input and single-output MISO [10].

Set input vector $X = [x_1, x_2 \cdots, x_r]^T$, every component of x_i is a fuzzy linguistic variable, and variable value of their language is:

$$T(x_i) = \{A_i^1, A_i^2, \cdots, A_i^{m_i}\}, \tag{8}$$

$A_i^j(j=1,2,3\cdots,m_i)$ is the j linguistic variable value of x_i. It is defined as a fuzzy set on the domain of U_x. The corresponding membership function is $\mu_{A_i^j}(x_i)(i=1,2,\cdots,r;\ j=1,2,3\cdots,m_i)$.

In the T-S model, the form of j fuzzy rule is as follows:

$$IF\ x_1\ is\ A_1^j, x_2\ is\ A_2^j,\cdots,x_r\ is\ A_r^j$$

$$THEN\quad y^j = p_0^j + p_1^j x_1 + \cdots p_r^j x_r$$

$$(j=1,2\cdots,m; m=\prod_{i=1}^{m} m_i) \tag{9}$$

According to Takagi-Sugeno fuzzy inference characteristics, it can be combined with a neural network for the construction of adaptive learning ability of neuro-fuzzy system[11-12]. Fig. 1 is built by the former pieces and back pieces of the network, and the former pieces of the network are used to match the antecedent of fuzzy rules, while the later is used to generate fuzzy rules.

(1)The antecedent network

The antecedent network consists of four layers; the first layer is the input layer, and its role is to transmit the input values to the next layer. The node number is $N_1 = r$.

Each node of the second layer represents a linguistic variable, and its role is to calculate the components of the input membership function.

Each node of the third layer represents a fuzzy rule, and its role is to match the antecedent of fuzzy rules to calculate the fitness of each rule.

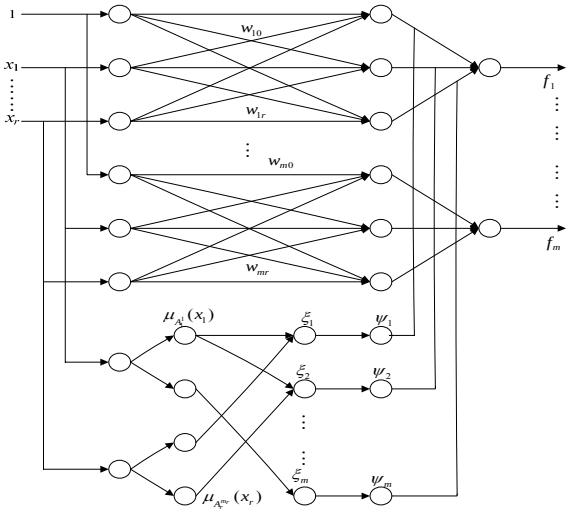


Figure 1 Structure diagram of T-S fuzzy neural network

The node number of the forth layer is the same as that of the third layer, and is for normalization, that is

$$\Psi_l = \frac{\xi_l}{\sum_{k=1}^{m}\xi_k} = \frac{\exp\left[-\sum_{i=1}^{r}\frac{(x_i - c_{ij})^2}{2\sigma_{ij}^2}\right]}{\sum_{k=1}^{m}\exp\left[-\sum_{i=1}^{r}\frac{(x_i - c_{ij})^2}{2\sigma_{ij}^2}\right]} \tag{10}$$

(2) The back network

The fifth layer is fuzzy decision-making $f_l = W_l * \Psi_l$, $W_l = w_{l0} + w_{l1}x_1 + \cdots w_{lr}x_r$, $l = 1,2\cdots m$.

The sixth floor is defuzzification layer. The number of its nodes corresponds to number of output variables.

C. Determination of Rules

The nearest neighbor clustering method is used to determine the rules in this paper. Firstly, the first data is taken as the first group of cluster center. Second, if the distance of a data from the cluster center is less than a preset value, then put the data into this group. In other words, the group of cluster centers and this data should be the closest. Otherwise, put it to a new group of cluster centers [13].

A modification is carried out based on this clustering algorithm. Each cluster center is compared with each other. When the distance between two clustering centers is less than the expected value, these two clustering centers are merged into one. So the clustering center is set up the average value and clustering number is equal to the original number minus one.

With the nearest neighbor clustering, steps to design the fuzzy system are listed as follows:

Step1: Select an appropriate width r of Gaussian function, define vector $A(l)$ for storing the sum of all output vectors, and counter $B(l)$ for statistics number of various sample types, where l is the number of categories.

Step2: From an input-output data $(x_0^1,\ y_0^1)$, set x_0^1 as a cluster center x_0^1, make $C_1 = x_0^1$, $A(1) = y_0^1$, $B(1) = 1$, set the radius r.

Step3: Considering the second sample data $(x_0^2,\ y_0^2)$, calculate $|x_0^2 - c_1|$ the distance from x_0^2 to the cluster center C_1. If $|x_0^2 - c_1| \le r$, then C_1 is the nearest neighbor clustering of x_0^2, set $A(1) = y_0^1 + y_0^2$, $B(1) = 2$; If $|x_0^2 - C_1| > r$, then take x_0^2 as a new cluster center, set $C_2 = x_0^2$, $A(2) = y_0^2$, $B(2) = 1$.

Step4: Suppose k pairs of input-output data $(x_0^k,\ y_0^k)$, (k=2, 3,…), and the existing cluster centers are x_0^1, x_0^2,…, the M clusters of the x_0^M.The distance from x_0^k to the M clusters centers $|x_0^k - x_0^l|$, $l = 1,2,3\cdots,M$ is calculated. Set the smallest distance is $|x_0^k - x_0^{l_k}|$, $(l_k \in l)$, that means $x_0^{l_k}$ is the nearest neighbor clustering of x_0^k.

If $|x_0^k - x_0^{l_k}| > r$, then take x_0^k as a new cluster center, set $C_{M+1} = x_0^k$, $M = M+1$, $A(M) = y_0^k$, $B(M) = 1$; If $|x_0^k - x_0^{l_k}| \le r$, then $A(j) = A(j) + y_0^k$, $B(j) = B(j) + 1$.

Step5: When C_1, C_2 …… C_{M+1} as cluster center, compare them one by one, If $|C_i - C_j| < \varepsilon$,which

means the distance between C_i and C_j is small enough, and they are close in width also. Then merge C_i and C_j, get $C = \frac{C_i + C_j}{2}$, numbers of cluster center minus 1.

Step6: x_0^k do not create a new cluster, then according to k pairs of input-output data $(x_0^j,\ y_0^j)$ (j = 1, 2 ..., k), the fuzzy system is designed as follows:

$$f_k(x) = \frac{\sum_{i=1}^{M} A(k)\exp(-\frac{|x - x_0^l|^2}{\sigma})}{\sum_{i=1}^{M} B(k)\exp(-\frac{|x - x_0^l|^2}{\sigma})} \tag{11}$$

If a new cluster is created, then it has the fuzzy system:

$$f_k(x) = \frac{\sum_{i=1}^{M+1} A(k)\exp(-\frac{|x - x_0^l|^2}{\sigma})}{\sum_{i=1}^{M+1} B(k)\exp(-\frac{|x - x_0^l|^2}{\sigma})} \tag{12}$$

Step7: Set k=k+1, then back to step 4.

D. Hybrid Learning Algorithm

This study is based on T-S fuzzy neural network, and introduces Gaussian function as fuzzy membership function. The parameters to be identified of the Gaussian membership function are the central value, width value and the link weight function coefficients of the Gaussian membership function. The first two parameters are the premise parameters of fuzzy rules, which are non-linear parameters. The weight function coefficients corresponding to the conclusion of fuzzy inference rules are linear parameters. According to the study of the structural characteristics of the fuzzy neural network, a hybrid learning algorithm could be proposed. The conjugate gradient algorithm can be used for the non-linear Layer parameters. While recursive least squares algorithm is used for the linear parameters, which significantly improve the network learning efficiency [14-15].

The calculation steps with conjugate gradient algorithm to optimize the parameters are as follows:

(1) Set initial point $M^{(0)}$ with accuracy $\varepsilon > 0$, D is the number of elements of vector M;

(2) Calculate $g_0 = \nabla E(M^{(0)})$, $k = 0$; Set $S^{(k)} = -g_k = -\nabla E(M^{(K)})$, which is a k-time search direction vector, the gradient direction g of each parameter of M is

$$\frac{\partial E}{\partial c_{ij}} = (d_i - y_i)\frac{\partial f}{\partial \xi_l}\frac{\partial \xi_l}{\partial c_{ij}} \tag{13}$$

$$\frac{\partial E}{\partial \sigma_{ij}} = (d_i - y_i)\frac{\partial f}{\partial \xi_l}\frac{\partial \xi_l}{\partial \sigma_{ij}} \tag{14}$$

(3)Set the learning rate of the center and width of M as η_1, η_2 respectively, or $\eta = \{\eta_1, \eta_2\}$, In order to avoid the inappropriate choice of learning rate causing system instability, this paper adopts an adaptive learning rate method. In the direction of a new vector $M^{(k+1)}$ is $M^{(k+1)} = M^{(k)} + \eta s^{(k)}$. When $\Delta E > 0$, $\eta = \eta\varphi$, $\varphi > 1$; When $\Delta E < 0$, $\eta = \eta\beta$, $\beta < 1$; φ, β are constants. The corresponding gradient is $g_{k+1} = \nabla E(M^{(k+1)})$

(4) If $E < \varepsilon$, terminate the iteration, otherwise transfer to (5);

(5) If $k < D - 1$, calculate conjugate direction.

$$s^{(k+1)} = -g_{k+1} + \frac{\|g_{k+1}\|^2}{\|g_k\|^2} s^{(k)} \tag{15}$$

Set $k = k + 1$, back to (3); if $k = D - 1$, set $M^{(0)} = M^{(D)}$,back to (2).

Conclusive parameters, which are the connection weights, are calculated with recursive least squares method.

Specific steps of hybrid learning algorithm are:

(1) Set the initial value: $k = 0$, set the precision of error $\varepsilon > 0$ and the initial learning rate η etc.;

(2) For the sample input p, via the network prior transmission consequent parameters W can be obtained by the least squares method, and then the corresponding network output could be obtained, then find out error E;

(3)Through back-propagation of error E, the gradient vector of the parameters c_{ij} and σ_{ij}, set $k = k + 1$;

(4) Calculating the value of c_{ij} and σ_{ij} with the conjugate gradient method;

(5)The next moment of error E could be obtained by prior dissemination based on revised parameter;

(6) If E reaches the error accuracy, at this time C, σ and W are the final values, learning terminates; otherwise, repeat step (2) to (6) until it meets the ending conditions.

IV. SIMULATION

The structure of stripper temperature fuzzy neural network modeling method is shown in fig. 2.

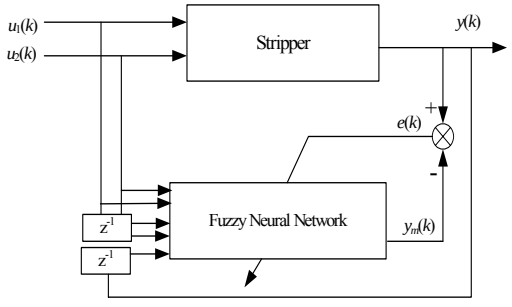


Figure 2 Modeling sketch of stripper temperature based on fuzzy neural network

In the figure above, u_1 is the steam flow; u_2 is slurry flow; y is the actual output value of temperature of

stripper top; y_m is output value of fuzzy neural network model; z-1 is the delay factor; k is the sampling time.

First, a dynamic model of stripper temperature can be built based on 100 groups of data collected from the site, and Fig. 3 shows the modeling results of stripper temperature, compared with that with BP, as showed in Fig. 4.

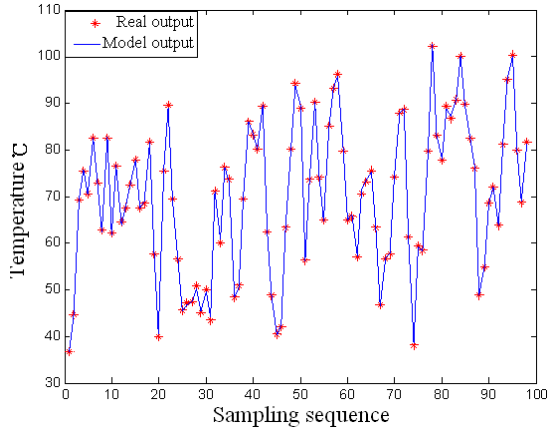


Figure 3 Modeling of stripper temperature based on FNN

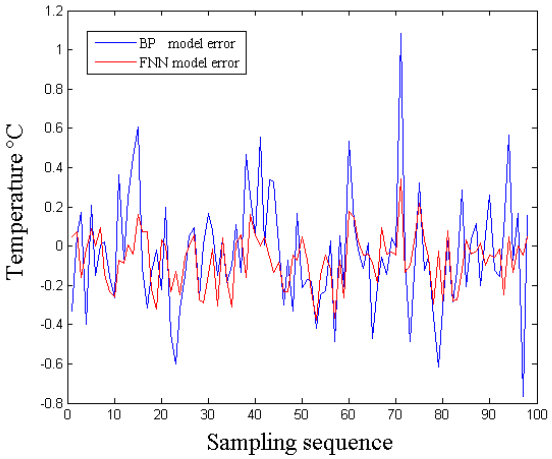


Figure 4 Modeling of stripper temperature based on FNN

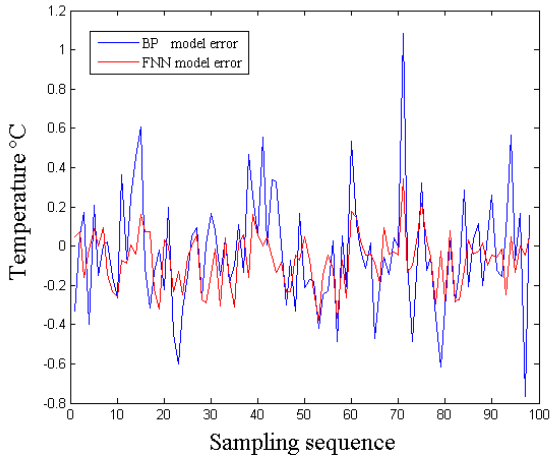


Figure 5 The modeling error

Fig. 5 show the modeling errors with both methods, and it can be seen that the variation with FNN is within the range of -0.2-0.2, which is much less than that with BP modeling method, mostly between -0.8—0.8. It means that accurate stripper temperature model could be obtained with a hybrid learning algorithm of fuzzy neural network. Comparison of performance of BP model error and FNN model error is showed in Tab. Ⅰ.

TABLE I.
COMPARISON OF BP AND FNN MODEL ERROR

	BP model error	FNN model error
min	-0.76429	-0.376
max	1.085714	0.344
varriance	0.083195	0.019344

The generalization test results of another 100 sets of data are shown in Fig. 6, and generalization error curve is shown in Fig. 7. Clearly, this fuzzy neural network used in the generalization results can meet the requirements.

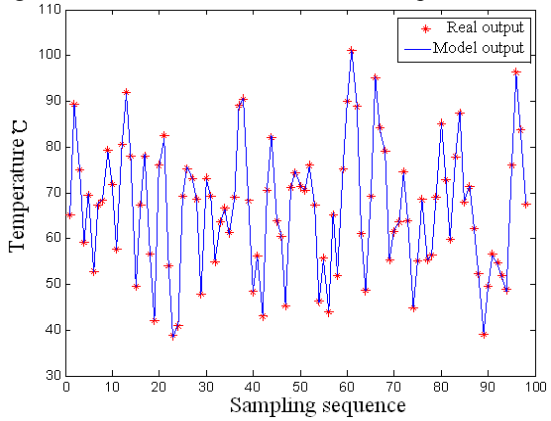


Figure 6 The generalization curve of FNN about stripper temperature model

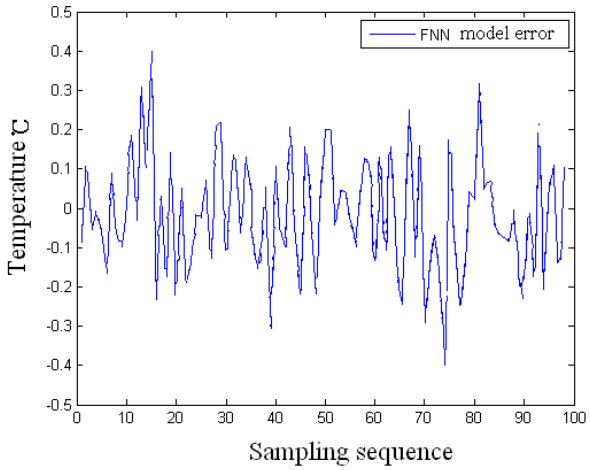


Figure 7 The generalization error curve of the model based on FNN

V. CONCLUSION

In this paper, the stripper temperature control model of the system is built up based on improved T-S fuzzy neural network, considering the time-varying, nonlinear and strong coupling problems. New nearest neighbor clustering method is put forward and used to determine fuzzy rules. At the same time, a hybrid learning algorithm is introduced for training neural networks, the non-linear conjugate gradient algorithm is used to solve the problem of layer parameters, while recursive least squares algorithm is introduced to solve the problem of the linear parameters. In this way, the identification accuracy of the model is improved. Simulation results show that higher

approximation accuracy of stripper temperature control system could be obtained with the method based on the T-S fuzzy neural network.

REFERENCES

[1] Hui Zheng-gang,The application of stripping technology in the production of PVC[J], Polyvinyl Chloride, Chemical Industry Press, Huludao, 2007,7(2):8-12.

[2] Azeem, MF; Ahmad, N; Hanmandlu, M. Fuzzy modeling of fluidized catalytic cracking unit [J], Fuzzy modeling of fluidized catalytic cracking unit 2007, 7(1): 298-324.

[3] Chang, H; Hou, WC. Optimization of membrane gas separation systems using genetic algorithm [J]. CHEMICAL ENGINEERING SCIENC, 2006, E (16). 5355-5368.

[4] Guo Ying, Lv Jian-hong, Wu Bo,e.t. Fuzzy modeling Based on TS model and Its Application for Thermal Process[J], Journal of System Simulation,2010,22(1):210-215.

[5] Qin Bin,Wu min,Wang xin,Yang Chun Hua,MAS-based distributed decoupling control for the pressure of gas collectors of coke[J], Control Theory & Applications,2006,23(2):961-966.

[6] Tan Yu-ling ,Chen Yuan, Intell igent Control Based on TS Model to the Temperature for Adhesive Preparation Processing[J], Computer Measurement & Control ,2010,18(2):351-353.

[7] He PL, Hou YX. An asymmetric robust learning algorithm of fuzzy clustering neural networks. Journal of Computer Research and Development, 2001, 38(3):296−301.

[8] Mastorocostas, PA; Hilas, CS. A block-diagonal recurrent fuzzy neural network for system identification [J]. NEURAL COMPUTING & APPLICATIONS, 2009, 18 (7): 707-717.

[9] Hong, YST; White, PA. Hydrological modeling using a dynamic neuro-fuzzy system with on-line and local learning algorithm [J], ADVANCES IN WATER RESOURCES, 2009, 32(1): 110-119.

[10] Li Guo-yong, Intelligent Control and MATLAB Implementation [M], Beijing:Electronics Industry Press, 2005,132-140.

[11] Yuan Yu-Hao, Zhang Guang-Ming, Research on T-Fuzzy Descriptor Systems:A review[J], ACTA AUTOMATICA SINICA, 2010, 36(7): 901-911.

[12] Ding Hai-Shan, Mao Jian-Qin, Lin Yan, Indirect Adaptive Fuzzy Control Based on Fuzzy Tree Model [J],ACTA AUTOMATICA SINICA, 2008, 34(6): 676-683.

[13] Shen HB, Wang ST, Wu XJ. Fuzzy kernel clustering with outliers. Journal of Software, 2004, 15(7):1021−1029.

[14] Wang Xue-miao, Fuzzy Neural Network Optimization and Its Application [D], Da Lian, 2007.

[15] Wang Hong-yuan and Shi Guo-dong, Artificial Neural Network Technology and its application[M], Beijing: China Petrochemical Press , 2002,97-106.

Shuzhi Gao received BSc degree from Shenyang University of Chemical Technology in 1989 and M. Sc. degree from Northeastern University in 2012. She is currently a professor in Shenyang University of Chemical Technology and PhD supervisor. Her main research interest is modeling of complex industry process and intelligent control.
E-mail; szg6868@126.com

Yihao Zhang received BSc degree from Henan University in 2010.He is a graduate in Shenyang University of Chemical Technology from 2011. His main research interest is modeling of complex industry process and intelligent control.
E-mail;yihaozhang310@163.com

Xianwen Gao received BSc degree from Shenyang University of Chemical Technology in 1978 and M. Sc. degree from Northeastern University in 1993. In 1998, he received PhD degree in control theory and control engineering from Northeastern University. He is currently a professor in Northeastern University. His main research interests are modeling of complex industry process and intelligent control, etc.
E-mail: gaoxianwen@ise.neu.edu.cn

Enteromorpha Prolifera Detection with MODIS Image Using Semi-supervised Clustering

Shunyao Wu[a,b], Fengjing Shao[a,b], Ying Wang[b], Rencheng Sun[b], Jinlong Wang[c]
[a] College of Automation Engineering, Qingdao University, Qingdao 266071, China
Email: shunyaowu@gmail.com
sfj@qdu.edu.cn
[b] College of Information Engineering, Qingdao University, Qingdao 266071, China
Email: yingwanglucky@gmail.com
qdsunstar@163.com
[c] School of Computer Engineering, Qingdao Technological University, Qingdao 266033, China
Email: wangjinglong@gmail.com

***Abstract*— In recent years, enteromorpha prolifera detection has received increasing attention. Supervised learning with remote sensing images can achieve satisfactory performances for green tide monitoring. However, data distributions between images obviously differ, and it would be too costly to label a massive amount of images for enteromorpha prolifera detection. Thus, this paper focuses on detecting enteromorpha prolifera using not only limited labelled data, but also a large amount of unlabelled data. We propose an effective semi-supervised clustering framework for enteromorpha prolifera detection, which can reduce the labelling cost and alleviate the overfitting problem. Experimental results prove the effectiveness and potential of our approach, with almost a 15% increase from baseline. In addition, the proposed approach can provide quantitative assessments for band data of moderate resolution imaging spectroradiometer (MODIS) images, and several often ignored bands, such as bands 5, 6, and 7, are shown to be useful for enteromorpha prolifera detection.**

***Index Terms*— enteromorpha detection, semi-supervised clustering, remote sensing images**

I. Introduction

Enteromorpha prolifera is a type of green algae that is widely distributed along the coastal areas of China [1]. Its natural reproductive capacity and environmental adaptability are particularly strong, and it may provoke a green tide under suitable conditions. In recent years, frequent occurrences of green tide have resulted in great losses for offshore tourism and aquaculture. Therefore, the rapid and accurate tracking and detection of green tides are urgent issues that must be solved.

Recently, various kinds of data were utilised for enteromorpha prolifera detection, such as sonar data [2], aerial images [3], and remote sensing images [4]. With the rapid development of remote sensing technology and data sharing, remote sensing images, such as moderate resolution imaging spectroradiometer (MODIS), have become the main data source to monitor green tides [4]–[7]. For example, Xing et al. [5] adopted a dynamic threshold strategy to detect enteromorpha prolifera based on multi-temporal and multi-source remote sensing images. Gu et al. [6] incorporated three modes of data, namely satellite optical, satellite microwave, and aerospace remote sensing, and used a decision tree algorithm to determine whether a sample was enteromorpha prolifera. Shi et al. [7] employed the fuzzy c-means method with images from the HJ-1A/1B satellites charge-coupled device sensor to complete the classification task for enteromorpha prolifera detection.

It is certain that enteromorpha prolifera detection with supervised learning can effectively work. Generally, training samples are obtained from one image or a collection of images through artificial visual interpretation. The training set is then used to understand a model for identifying unrecognised or uncertain (some islands appear as enteromorpha prolifera in remote sensing images) enteromorpha prolifera in the same image or collection. However, there exist two major problems with this approach. Firstly, remote sensing data nowadays increase at an exponential rate, and with such a massive amount of images, human labelling is too costly and unsuitable for the rapid automation of monitoring green tide. Secondly, due to a multitude of reasons, such as weather and solar flares, data distributions between images obviously differ, making it a challenge to detect enteromorpha prolifera through supervised learning with insufficient training samples. Table I shows the classification accuracy of supervised methods on five images. E1-E5 stands for five remote sensing images. "Combine" denotes the combination of the five images. Enteromorpha prolifera detection is treated as a binary classification to determine whether a sample has enteromorpha prolifera or not, we test supervised learning methods on E1-E5

Manuscript received August 10, 2011; revised January 2, 2012; accepted April 16, 2012.

This work was partially supported by the State Key Program of National Natural Science of China (No. 91130035), the National Natural Science Foundation of China (No. 60974085), the National Public Benefit Research Foundation (No. 200905030), the National Science Foundation of Shandong Province (No. ZR2012FZ003), the National Science Foundation of Shandong Province (No. ZR2012FQ017) and the Science and Technology Planning Project of Qingdao (No. 12-1-4-4-(8)-jch).

doi:10.4304/jcp.9.5.1259-1265

TABLE I.
CLASSIFICATION ACCURACY.

Dataset	One rule	Decision tree	Support Vector Machine
E1	100%	100%	93.75%
E2	90%	91.5%	90%
E3	84.75%	84.75%	81%
E4	95.25%	99.5%	97.5%
E5	95%	92%	94.5%
Combine	75.44%	95.56%	93.2%

and "Combine". As shown in Table I, the supervised methods can achieve a good performance. Nevertheless, when setting E1 as the training set and other images as the test sets, the method is ineffective, with the accuracy rate falling to 50%. The reason for this is because the training set is insufficient for supervised learning, which brings about the overfitting problem. There thus exists a trade-off between labelling costs and classification accuracy, while it is also complicated to determine the amount of training images and the sampling strategy needed for supervised learning.

To overcome these problems, this paper introduces semi-supervised learning for enteromorpha prolifera detection. Differing from supervised learning, semi-supervised learning [8]–[10] can efficiently make use of a small amount of labelled data (or prior knowledge) and a large number of unlabelled data. And semi-supervised learning has been widely applied to many fields, such as text mining [11] and bioinformatics [12]. Thus, we propose a novel framework for enteromorpha prolifera detection with semi-supervised clustering based on metric learning. On the one hand, the supplement size of the labelled data for semi-supervised learning is much smaller than for supervised learning, but the semi-supervised approach can effectively utilise unlabelled data to improve the classification accuracy. On the other hand, this method can alleviate the overfitting problem using partition samples according to their similarities as opposed to the learning classification model. In addition, attribute weights can be obtained through metric learning, which is useful for identifying the critical band for green tide detection.

The rest of the paper is organized as follows. In section II, we propose a semi-supervised learning framework for enteromorpha prolifera detection. In section III, we compare the performance of our approach with *k*-means on nine MODIS images, and apply the proposed approach to green tide monitoring. Finally, we conclude the paper and discuss some future work in Section IV.

II. SEMI-SUPERVISED LEARNING FRAMEWORK FOR ENTEROMORPHA PROLIFERA DETECTION

Traditional unsupervised clustering automatically groups similar objects, while separating different objects without providing a training set. Nevertheless, this strategy has many disadvantages, such as poor performance, unreadable clustering results, and limited ability to satisfy user requirements. To solve these problems, semi-supervised clustering has emerged as a solution, as it incorporates a small amount of prior knowledge so as to guide the clustering procedure and obtain better performances [13] [14].

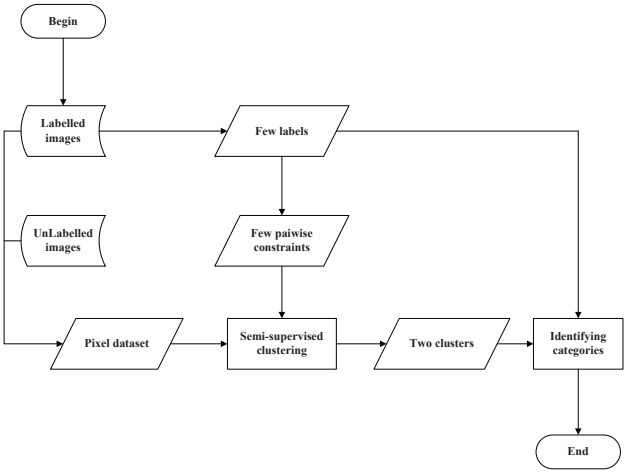


Figure 1. Main flow of our framework.

This paper aims to use semi-supervised clustering with limited prior knowledge to detect enteromorpha prolifera in a large amount of unlabelled images and thus provide a promising strategy for the automation of green tide monitoring. Figure 1 demonstrates the main flow of our framework. Firstly, we pre-processed remote sensing images and extracted enteromorpha prolifera and non-enteromorpha prolifera pixels to obtain datasets. Subsequently, we generated pairwise constraints based on labels and utilised semi-supervised clustering incorporating pairwise constraints to partition the datasets into two clusters. Finally, we identified the categories of the two clusters according to the labels, so as to detect enteromorpha prolifera.

A. Data pre-processing

MODIS images have three different resolutions: 1km, 500m, and 250m. The satellite transits China twice every morning, with the monitoring areas covering the coastal waters of Qingdao, the Yellow Sea, and the East China Sea. Thus, it is conducive for accessing important enteromorpha prolifera information, such as the covered area, distribution, and drift trends.

We selected MODIS HKM images from the American National Aeronautics and Space Administration [1] as our data source, and each pixel in the images had seven bands of information. As many factors, such as attitude, altitude, and speed of aircrafts, may cause geometric distortions, we implemented more detailed pre-processing for the remote sensing images, such as geometric correction. Further, we extracted enteromorpha prolifera and non-enteromorpha prolifera samples from a small number of labelled images. The band information between the land and ocean was obviously different. Thus, it is feasible

[1] http://rapidfire.sci.gsfc.nasa.gov/realtime

to distinguish the land and ocean using coastline extraction [15] or according to latitude and longitude data. Finally, we select pixels from the normal ocean as non-enteromorpha prolifera samples, and obtain enteromorpha prolifera samples from the sea area with the green tide.

B. Semi-supervised clustering based on metric learning

Given a set of pixels $\mathcal{P} = \{\mathbf{p}_1, ..., \mathbf{p}_\mathcal{P}\}$, a must-link constraint set $\mathcal{S}$, and a cannot-link set $\mathcal{D}$, the objective of clustering was to effectively partition $\mathcal{P}$ into $k(k = 2)$ clusters using prior knowledge.

1) Pairwise constraints: Pairwise constraints is a popular type of instance-level knowledge in semi-supervised learning, and have become a common form to express user requirements. For instance, in the global positioning system (GPS) for intelligent navigation application, it is reasonable to determine whether the two GPS points of one car are on the same lane using trace contiguity and maximum separation [16]. Moreover, compared with labels, pairwise constraints are more consistent with the clustering objective as they focus on the differences between objects.

Pairwise constraints consists of must-link constraint set $\mathcal{S}$ and cannot-link constraint set $\mathcal{D}$.

- If $\mathbf{p}_i$ and $\mathbf{p}_j$ are in the same cluster, $(\mathbf{p}_i, \mathbf{p}_j)$ belongs to $\mathcal{S}$.
- If $\mathbf{p}_i$ and $\mathbf{p}_j$ are in different clusters, $(\mathbf{p}_i, \mathbf{p}_j)$ belongs to $\mathcal{D}$.

Given the limited amount of labels, we obtained the pairwise constraints based on whether the labels of two pixels were the same or not. Further, we generated a larger amount of prior knowledge according to the properties of pairwise constraints. Provided that cluster number was 2, if $(\mathbf{p}_i, \mathbf{p}_j) \in \mathcal{S}$ and $(\mathbf{p}_i, \mathbf{p}_l) \in \mathcal{D}$, $(\mathbf{p}_i, \mathbf{p}_j)$ belonged to $\mathcal{D}$; if $(\mathbf{p}_i, \mathbf{p}_j) \in \mathcal{S}$ and $(\mathbf{p}_i, \mathbf{p}_m) \in \mathcal{S}$, $(\mathbf{p}_j, \mathbf{p}_m)$ belonged to $\mathcal{S}$; if $(\mathbf{p}_i, \mathbf{p}_l) \in \mathcal{D}$ and $(\mathbf{p}_l, \mathbf{p}_o) \in \mathcal{D}$, $(\mathbf{p}_j, \mathbf{p}_o)$ belonged to $\mathcal{S}$.

2) Learning new metric with pairwise constraints: As an important semi-supervised clustering strategy, metric learning can learn new metric by solving optimisation, which aims to satisfy prior knowledge as much as possible. Xing et al. [13] proposed an effective convex optimisation to learn new metrics with pairwise constraints, which ensures the distances between objects belong to the must-link constraint set $\mathcal{S}$ as small as possible. As Xing′s method is both effective and efficient, our optimisation is constructed in accordance with this method.

$$min_{\mathbf{w}} \sum_{(\mathbf{p}_i, \mathbf{p}_j) \in \mathcal{S}} D_{\mathbf{w}}(\mathbf{p}_i, \mathbf{p}_j) + \lambda \mathbf{w}^T \mathbf{w}$$

s.t.

$$\sum_{(\mathbf{p}_i, \mathbf{p}_j) \in \mathcal{D}} D_{\mathbf{w}}(\mathbf{p}_i, \mathbf{p}_j) \geq 1$$

$$\mathbf{w} \succeq 0 \qquad (1)$$

The first term of the objective function aims to reduce the differences between pixels belonging to the must-link constraint set $\mathcal{S}$; the second term is the regulation term which ensures the consistent of band weights; the constraint means the differences between pixels belonging to the cannot-link constraint set $\mathcal{D}$ should be large enough.

$D_{\mathbf{w}}(.,.) = \sum_{j=1}^{d} w_j d_\phi(.,.)$, where $d_\phi(.,.)$ corresponds to the Bregman divergences; d stands for the dimension number of the dataset, and $d = 7$, as the MODIS image used in this paper has seven bands. The Bregman divergences [17] can be applied to many useful distances, such as Itakura-Saito, Mahalanobis, Squared Euclidean, and so on. Thus, this paper utilises the Bregman divergences as the metric of our framework so as to improve its expansibility.

On the condition that $\phi : \mathcal{S} \to \mathcal{R}$ is a strictly convex function defined on a convex set $\mathcal{S} \subseteq \mathcal{R}^d$ to ensure that ϕ is differentiable on $ri(\mathcal{S})$, the Bregman divergences d_ϕ are defined as follows:

$$d_\phi(\mathbf{p}_i, \mathbf{p}_j) = \phi(\mathbf{p}_i) - \phi(\mathbf{p}_j) - < \mathbf{p}_i - \mathbf{p}_j, \phi(\mathbf{p}_j) > \quad (2)$$

where $\nabla\phi$ is the gradient vector of ϕ.

Given the different function ϕ, the Bregman divergences can be transformed into different types of distances. For example, given $\phi(x) = x^2$, $d_\phi(x, y) = (x - y)^2$, and $D_{\mathbf{w}}(.,.)$ becomes a parameterised square Euclidean distance, $D_{\mathbf{w}}(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{d} w_j (x_j - y_j)^2$. In this paper, we utilize the parameterised square Euclidean distance when constructing the optimization with pairwise constraints.

3) Algorithm procedure: Generally, *k*-means clustering contains two steps: step 1, assigning each pixel to the nearest cluster; step 2, re-estimating the cluster representations. The two steps are iterated until convergence is obtained.

In this paper, we firstly employed the large-scale optimisation software, MOSEK, to solve the optimisation. After solving the optimisation, the obtained metric was used to assign the cluster index for step 1. Cluster representations were then obtained by computing the arithmetic mean of clusters for step 2. The procedure of the semi-supervised clustering was as follows:

4) Identifying categories of clusters: As an unsupervised method, clustering cannot determine the categories of objects. To identify categories of clusters, this paper counted the number of labelled enteromorpha prolifera samples in each cluster and then assigned the cluster with a greater number of enteromorpha prolifera samples as the enteromorpha prolifera class, with the other cluster being normal ocean class.

III. Results and Discussion

In this section, we validate the effectiveness of our approach with several experimental results.

A. Datasets and experimental setting

We downloaded nine MODIS images from the National Aeronautics and Space Administration site and named

Algorithm 1: Procedure of semi-supervised clustering

Input: Pixel dataset $\mathcal{P}$, must-link set $\mathcal{S}$, cannot-link set $\mathcal{D}$.

Output: Two Clusters obtained with semi-supervised clustering.

Procedure Begin

Step 1: Integrate must-link set $\mathcal{S}$ and cannot-link set $\mathcal{D}$ by constructing the optimisation according to formula (1);

Step 2: Solve the optimisation problem to obtain the weight vector of bands $\mathbf{w}$;

Step 3: Scale all the pixels based on the band weight vector $\mathbf{w}$;

Step 4: Divide the scaled dataset into two clusters using the *k*-means algorithm.

Procedure End

them E1 to E9. In the experiments, we extracted the training set from E1, and then respectively combined E1 and the other images to obtain eight datasets. When providing $2m$ (m is a positive integer) labelled pixels, we generated m must-link constraints and $1.2m$ cannot-link constraints for the semi-supervised clustering [2]. In addition, we set d for the parameter λ in the objective function of the optimisation (1).

B. Evaluation criteria

Two common clustering external indexes, NMI and Purity, were used to evaluate the clustering results and Accuracy was adopted to evaluate the accuracy of enteromorpha prolifera detection.

Normalised mutual information is one type of measure based on information entropy and is widely adopted [14] [19].

$$NMI(\mathcal{C},\mathcal{B}) = \frac{I(\mathcal{C},\mathcal{B})}{\sqrt{H(\mathcal{C})H(\mathcal{B})}} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{k} n_{ij} log\frac{n\cdot n_{ij}}{n_i\cdot n_j}}{\sqrt{\sum_{i=1}^{k} n_i log\frac{n_i}{n}\sum_{j=1}^{k} n_j log\frac{n_j}{n}}} \quad (3)$$

Here, H presents the entropy and I computes the mutual information. $\mathcal{C}$ presents the clustering results after applying our approach to partition $\mathcal{P}$, while $\mathcal{B}$ denotes the pre-specified structure. n represents the number of pixels in the dataset. The number of items in $\mathcal{C}$ and $\mathcal{B}$ are both k. We use n_j to express the object number in the ith cluster and n_j in the jth cluster; n_{ij} thus denotes the item number included in ith and jth cluster.

Purity built one-to-one correspondences between the clusters and classes, and measured the proximity of the cluster assignments to the pre-specified structure.

$$Purity(\mathcal{C},\mathcal{B}) = \frac{max(\sum_{i=1}^{k} n_{i,Map(i)})}{n} \quad (4)$$

[2]The ratio between must-link and cannot-link constraints was set according to [18]

TABLE II.
CLUSTERING ACCURACY UNDER PURITY/NMI INDEX.

Dataset	*k*-means	our approach
E2	0.6167/0.1677	0.8333/0.4791
E3	0.5000/0	0.7481/0.5000
E4	0.7086/0.2029	0.7818/0.3013
E5	0.7481/0.3285	0.8525/0.5757
E6	0.8249/0.4751	0.8937/0.6040
E7	0.7536/0.3507	0.7692/0.3735
E8	0.6182/0.1481	0.7481/0.3283
E9	0.5307/0.2419	0.8237/0.6750
mean	0.6626/0.2394	0.8063/0.4796

Here, $\mathcal{C}$ stands for the cluster indexes, $\mathcal{B}$ for the underlying class labels, and i for the cluster index. $Map(i)$ is the class label corresponding to the cluster index i, and $n_{i,Map(i)}$ is the number of pixels not only belonging to cluster i, but also class $Map(i)$.

Accuracy is a simple index used to calculate the consistency of results with the underlying class labels for classification.

$$Accuracy(\mathcal{C},\mathcal{B}) = \frac{\sum_{i=1}^{k} n_i}{n} \quad (5)$$

Here, $\mathcal{C}$ stands for classification results while $\mathcal{B}$ presents the underlying class labels. i represents the category index and n_i the number of correct instances for category i.

C. Comparison with k-means

Tables II and III provide the empirical results of our approaches in comparison to *k*-means clustering. Each result was generated using 20 trials. We provided 10% labelled data and generated 5% must-link constraints and 6% cannot-link constraints for each dataset. As shown in Tables II and III, the clustering quality and classification accuracy of our approach were better than the *k*-means on the whole datasets. Table IV shows the t-test results comparing our approach to *k*-means. This index indicates that the improvement observed with our proposed method was statistically obvious.

The experimental results also demonstrated the effectiveness of category identification, as the accuracy was mostly in line with the purity with the exception of the E4 dataset. This was because the clustering did not function as expected, and all of the pixels in E4 were partitioned into the same cluster. The reason was that this image contained heavy fog, which had a negative influence on sampling.

Through the metric learning, band information could be organised according to the weights calculated, and we find, band 2 $>$ band 5 $>$ band 6 $>$ band 4 $>$ band 1 $>$ band 7 $>$ band 3. Obviously, band 2 is the most important of all. It is consistent with the traditional methods in marine remote sensing [4]. In addition, bands 5, 6, and 7 were also important and useful for enteromorpha prolifera detection, even though they are often neglected by the marine experts.

TABLE III.
CLASSIFICATION ACCURACY.

Dataset	*k*-means	our approach
E2	0.6167	0.8333
E3	0.5000	0.7481
E4	0.5101	0.5792
E5	0.7481	0.8525
E6	0.8249	0.8937
E7	0.7536	0.7692
E8	0.6051	0.7481
E9	0.5307	0.8237
mean	0.6361	0.7810

TABLE IV.
T-TEST OF OUR APPROACH VS. *k*-MEANS

Dataset	NMI	Purity	Accuracy
E2	2.2828e-006	2.2828e-006	2.2828e-006
E3	0	7.9678e-299	7.9678e-299
E4	9.7101e-006	1.8137e-006	0.1923
E5	3.7652e-005	3.7572e-005	3.7572e-005
E6	0.0022	0.0060	0.0060
E7	8.9731e-027	6.5897e-027	6.5897e-027
E8	2.2834e-006	2.2828e-006	2.2828e-006
E9	0.0014	9.7640e-004	9.7640e-004

D. Classification performance versus labels

As shown in Figures 2 and 3, we demonstrated the performance versus the number of labels for the E5 and E9 datasets. With an increasing number of labels, we provided more pairwise constraints as stated in the experiment setting. It was clear that our approach was superior to the *k*-means method. However, the performance of our approach was deduced with the increase in labels. There are two possible reasons for this. On the one hand, the incoherence between must-link and cannot-link constraints may decrease the clustering quality, especially in cases of large amounts of prior knowledge [20]. On the other hand, it may cause the overfitting problem to learn with a large number of labels extracted from E1.

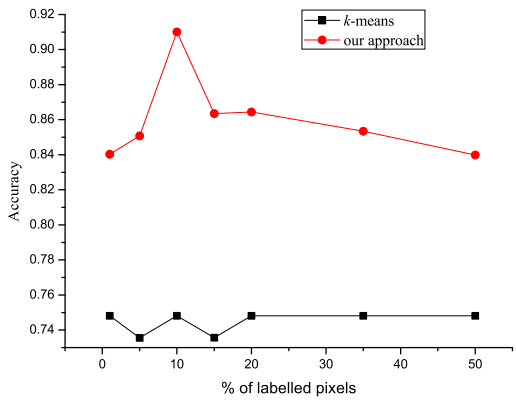


Figure 2. Performance versus labels for the E5 dataset.

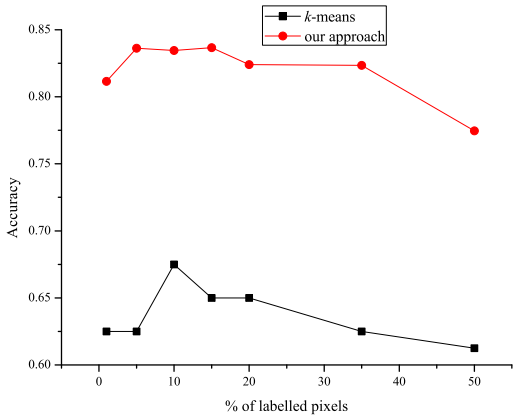


Figure 3. Performance versus labels for the E9 dataset.

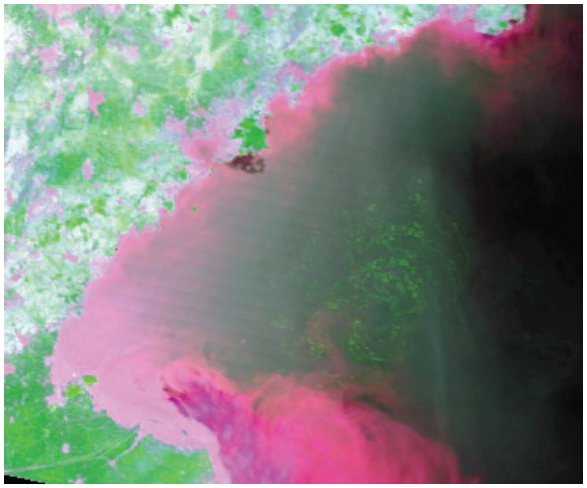


Figure 4. Artificial visual interpretation of the E5 image.

E. Green tide monitoring application

Finally, we applied the proposed approach to green tide monitoring. During the period from May 14th to July 17th in 2008, greed tide occurred in the sea area of Qingdao, China. Thus, we select two MODIS images on May 31st and June 29th, E5 and E9 images, to test our approach.

Figures 4 and 6 are the artificial visual interpretation of E5 and E9 images. Figures 5 and 7 demonstrate the detection effect on the E5 and E9 images, respectively. According to the previous analysis, we select 10% labelled pixels in E1 images for semi-supervised clustering. As shown in Figures 5 and 7, Our approach was thus successful in effectively identifying green tide according to the real disaster area marked by the artificial visual interpretation. In particular, due to the thick curtain of fog in the E9 image, its performance was not as great as the E5 image.

IV. CONCLUSIONS

This paper proposes an efficient semi-supervised clustering framework for enteromorpha prolifera detection, which provides a promising strategy incorporating limited

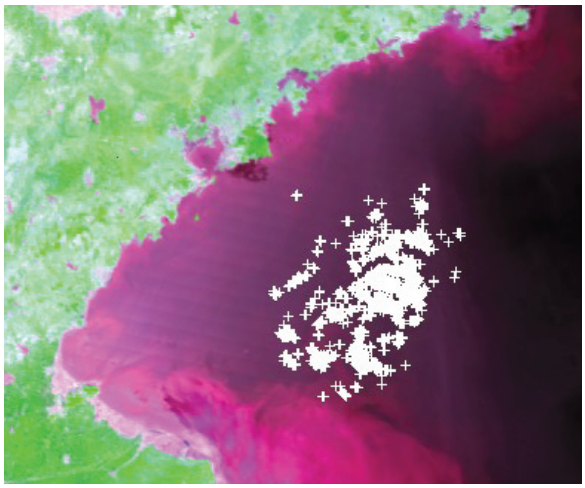


Figure 5. Green tide detection for the E5 image.

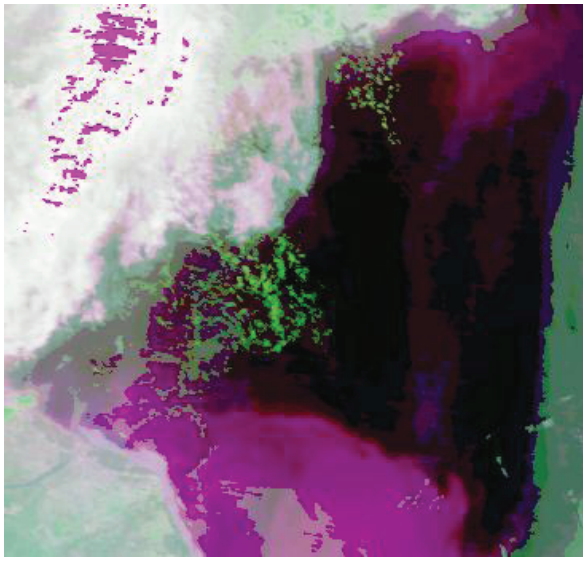


Figure 6. Artificial visual interpretation of the E9 image.

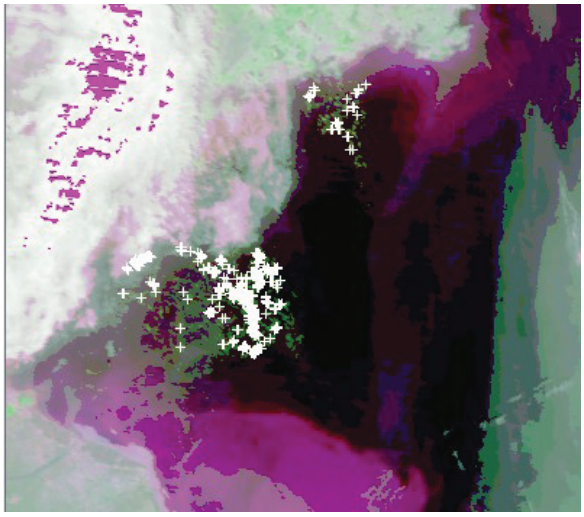


Figure 7. Green tide detection for the E9 image.

labelled data and a large amount of unlabelled data for the automation of green tide monitoring. With few labels, we generated pairwise constraints for semi-supervised clustering and then adopted metric learning to incorporate pairwise constraints, so as to partition the pixel set into two clusters. Finally, we identified the corresponding categories of clusters for enteromorpha prolifera detection by counting the amount of enteromorpha prolifera samples. Experimental results demonstrate the effectiveness of our approach, which is promising for using the weights of band information obtained based on metric learning to monitor green tide.

In the future, we aim to design a more systematic method by combing multi-source remote sensing images and implementing an effective toolkit for the automation of green tide monitoring. Other work involves applying and comparing the performance of metric learning strategy with other semi-supervised learning algorithms for enteromorpha prolifera detection. Active learning will also be introduced into the semi-supervised clustering procedure.

ACKNOWLEDGMENT

The authors are grateful to the anonymous referees for their valuable comments and suggestions to improve the presentation of this paper.

REFERENCES

[1] Z. Liang, X. Lin, M. Ma, J. Zhang, X. Yan, and T. Liu, "A preliminary study of the enteromorpha prolifera drift gathering causing the green tide phenomenon [j]," *Periodical of Ocean University of China*, vol. 4, pp. 601–604, 2008.

[2] L. Zhu, G. Yao, and X. Zhang, "Researching in enteromorpha disaster monitoring based on sonar detection," *Science and Technology Information*, pp. 187–188, 2009.

[3] X. Dong, J. Dong, and L. Qu, "Enteromorpha detection in aerial images using support vector machines," in *IEEE Youth Conference on Information, Computing and Telecommunication*, 2009, pp. 299–302.

[4] C. Hu, "A novel ocean color index to detect floating algae in the global oceans," *Remote Sensing of Environment*, vol. 113, no. 10, pp. 2118–2129, 2009.

[5] Q. Xing, X. Zheng, P. Shi, J. Hao, D. Yu, S. Liang, D. Liu, and Y. Zhang, "Monitoring green tide in the yellow sea and the east china sea using multi-temporal and multi-source remote sensing images," *Spectroscopy and Spectral Analysis*, vol. 31, no. 6, pp. 1644–1647, 2011.

[6] X. Gu, X. Chen, Q. Yin, Z. Li, H. Xu, Y. Shao, and Z. Li, "Stereoscopic remote sensing used in monitoring enteromorpha prolifra disaster in chinese yellow sea," *Spectroscopy and Spectral Analysis*, vol. 31, no. 6, pp. 1627–1632, 2011.

[7] Y. Shi, L. Shi, M. Xa, N. Yang, and J. Ding, "The application of hj-1a/1b's ccd data to entermorpha prolifera monitoring over the yelow sea and east sea," *Remote Sensing Information*, vol. 2, pp. 47–50, 2012.

[8] X. Zhu and A. B. Goldberg, "Introduction to semi-supervised learning," *Synthesis lectures on artificial intelligence and machine learning*, vol. 3, no. 1, pp. 1–130, 2009.

[9] S. Xu, "Semi-supervised tensor graph-optimized linear discriminant analysis for two-dimensional face recognition," *Journal of Software*, vol. 8, no. 8, pp. 2047–2051, 2013.

[10] Z. Huang and Y. Chen, "A prototype patterns selection algorithm based on semi-supervised learning," *Journal of Software*, vol. 8, no. 8, pp. 1984–1990, 2013.

[11] K. Nigam, A. McCallum, and T. Mitchell, "Semi-supervised text classification using em," *Semi-Supervised Learning*, pp. 33–56, 2006.

[12] T.-P. Nguyen and T.-B. Ho, "Detecting disease genes based on semi-supervised learning and protein–protein interaction networks," *Artificial Intelligence in Medicine*, vol. 54, no. 1, pp. 63–71, 2012.

[13] E. P. Xing, M. I. Jordan, S. Russell, and A. Ng, "Distance metric learning with application to clustering with side-information," in *Advances in neural information processing systems*, 2002, pp. 505–512.

[14] S. Basu, M. Bilenko, and R. J. Mooney, "A probabilistic framework for semi-supervised clustering," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004, pp. 59–68.

[15] S. Dellepiane, R. De Laurentiis, and F. Giordano, "Coastline extraction from sar images and a method for the evaluation of the coastline precision," *Pattern Recognition Letters*, vol. 25, no. 13, pp. 1461–1470, 2004.

[16] K. Wagstaff, C. Cardie, S. Rogers, and S. Schrödl, "Constrained k-means clustering with background knowledge," in *ICML*, vol. 1, 2001, pp. 577–584.

[17] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh, "Clustering with bregman divergences," *The Journal of Machine Learning Research*, vol. 6, pp. 1705–1749, 2005.

[18] M. Halkidi, D. Gunopulos, M. Vazirgiannis, N. Kumar, and C. Domeniconi, "A clustering framework based on subjective and objective validity criteria," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 1, no. 4, p. 4, 2008.

[19] X. Hu, X. Zhang, C. Lu, E. K. Park, and X. Zhou, "Exploiting wikipedia as external knowledge for document clustering," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009, pp. 389–396.

[20] I. Davidson, K. Wagstaff, and S. Basu, "Measuring constraint-set utility for partitional clustering algorithms," *Knowledge Discovery in Databases: PKDD 2006*, pp. 115–126, 2006.

Shunyao Wu received the B.S. and M.S. degree in computer Science and technology and computer software and theory from Qingdao Technological University, China, in 2008 and 2011, respectively. He is currently working towards his Ph.D. degree in systems theory at Qingdao University, China. His current research interest includes data mining, machine learning, complex network and bioinformatics.

Fengjing Shao received her B.S. degree in computer science from Shandong University, China in 1982 and her Ph.D. degree in information engineering from Osaka University, Japan in 1991. She is a professor in College of Information Engineering, Qingdao University, China. Her research interests include complex network, network complexity and data mining.

Ying Wang received her B.S. degree in computer science from Qingdao University, China in 2011. She is currently a master student and working towards her M.S. degree at Qingdao University, China. Her current research interest includes data mining and remote image detection.

Rencheng Sun received his B.S. degree in applied mathematics from Qingdao University, China in 2002, his M.S. degree in computer software and theory from Qingdao University, China in 2005, and his Ph.D. degree in systems theory at Qingdao University, China in 2010. He is currently an associated professor in College of Information Engineering, Qingdao University. His research interests include complex network and data mining.

Jinlong Wang received his B.S. and Ph.D. degree in computer Science and technology from Zhejiang University, China, in 2002 and 2007, respectively. He is currently a professor in School of Computer Engineering, Qingdao Technological University. His research interests include data mining, machine learning and artificial intelligence.

A Fast Potential Fault Regions Locating Method Used in Inspecting Freight Cars

Zongxiao Zhu
Huazhong University of Science and Technology, Wuhan, China
South-Central University for Nationalities, Wuhan, China
Email: zhuzongxiao@gmail.com

Guoyou Wang
Huazhong University of Science and Technology, Wuhan, China
Email: gywang@mail.hust.edu.cn

***Abstract*—The trouble of freight car detection system (TFDS) is a popular application in Chinese railway today. In this paper, the discrete-point sampling model is further developed to locate potential fault regions in the photos taken by the TFDS. The discrete-point sampling model not only contains the image's region boundary information and region information, but also reflects the transition from region to boundary. The most salient component's contours in samples are drawn by hand and recorded as data templates used for matching in test images. Experimental results show that by components' classification, the method based on this model can classify different types of freight cars' parts universally and locate the potential fault regions more accurately and quickly than regional gray matching or edge matching. The results of anti-noise testing in laboratory and more than two years daily operation at several inspecting stations show that our method has a strong ability to survive with nonlinear deformations, and has a good extensibility to be used with different parts, which meet application demands for the full-automatic inspection system.**

***Index Terms*— TFDS, discrete-point sampling model, components' classification, potential fault regions locating**

I. Introduction

Every day in China, thousands and tens of thousands of freight trains are running on more than 90,000 kilometers railway. The trouble of freight car detection system [1]-[2] （TFDS, official name） is a popular application demanded by the Chinese department of railway now. A TFDS takes 53 photos each freight car with five cameras from sides and bottom as showed in Fig. 1, which means each freight train may have 2,000-3,000 photos according to the number of its boxcars. Four experienced workers will look into all these photos and make judgments whether there are troubles or not in 10 minutes. Our goal is using computer to do the same job with the same accuracy and same speed as those experienced workers do, and change the human-control model to the computer combining with human-control model until the final realization of the complete computer-control model.

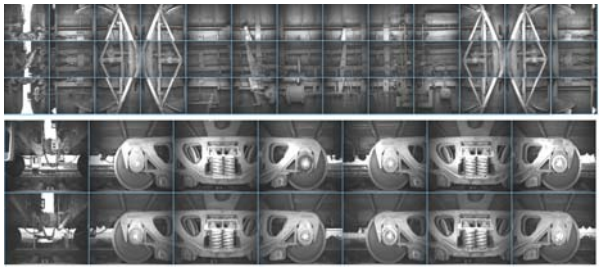


Figure 1. Fifty-three photos from Chinese trouble of freight car detection system.

The key problem in TFDS is how to classify components and locate potential fault regions as quickly as possible. Since China's railway system has been developing rapidly, different kinds of freight cars are running at the same time with a variety of components. Both new and old components, even with the same function, may be designed in different years with several types. Different types of components may have different fault regions in different quantities. The computer fault detection algorithm should first automatically identify the component's type in the freight car, then locate potential fault regions, and detect if there is fault or not in these potential fault regions.

The challenge also comes from the difficulty of components modeling. In TFDS, the status of freight car and photography conditions vary significantly at different inspecting stations such as illumination changes, different weather conditions, varying speed and camera vibration, etc. As a result, the photos of one component type may be quite different in details despite overall similarity. It is impossible to gain identical photos. In this case, if the common regional threshold segmentation or edge extraction algorithms are adopted, irregular and random changes will occur in the regions or edges. So it is very difficult to establish a unified mathematical model representing this type of component.

Manuscript received March 15, 2013.
This work was supported by Wuhan Huamu Science and Technology Company.
Guoyou Wang is the corresponding author.

doi:10.4304/jcp.9.5.1266-1273

Before knowing about TFDS, the railway workers who look over those photos have already been checking out and repairing the faults in freight cars for more than five years. They are quite familiar with those regions with potential faults in various types of components, and they can classify the type of components and locate the region with potential faults as soon as they see the photo. In order to ensure the computer to reach the same detecting ability, many methods have been tried, including line matching[3], edge matching [4], corner point matching[5], SIFT[6] and its variant features [7] matching. However, the classification capacity, locating time and accuracy of these methods are no match for those experienced railway workers. This situation has not been changed until a method based on the discrete-point sampling model [8]-[9] (DPSM) is developed. This model takes advantage of both methods based on object region and boundary, obtaining the boundary information from a yin discrete-point sampling map and regional information from a yang discrete-point sampling map. Yin and yang sampling maps reflect the transitional features from object region to object boundary in the image. Summarized from a large number of images, the transitional features are combined for component structure modeling and component type classification. Better results have been achieved in classifying the component types and locating potential fault regions. Currently, this model has been successfully applied in several freight car stations with TFDS.

This paper is organized as follows. Section 2 briefly introduces the developed discrete-point sampling model. Section 3 presents how the discrete-point sampling model be used to locate the potential fault regions in a TFDS. In Section IV, we describe and discuss the results of our experiments. Section V offers our conclusion.

II. THEORY

A. Discrete-point Sampling

The visual basis for application of the discrete-point is primarily established upon the spatial characteristics of vision, including the cumulative effect in space and the limitation of visual acuity [10]. Visual acuity of human eyes has a certain limit. For instance, based on the optical diffraction principle, a single light spot in a scene will no longer be a single spot when focused on the retina, but a pattern composed of a central bright disk and a series of surrounding dark and bright rings. The width of a fine line on the retina is always larger than 0.0087rad, no matter how fine the original test line is. Under the best light conditions, the best eyes of human beings can only distinguish grids composed of lines with width corresponding to 0.0097-0.011rad on the retina. Considering all these visual features, the major structural information in an image can be conveyed by a series of discrete-point in the image. Human brain can easily combine these discrete points and obtain meaningful information. Different amounts of discrete-point can convey different quantities of information and present the details of the image at different levels, which can be seen

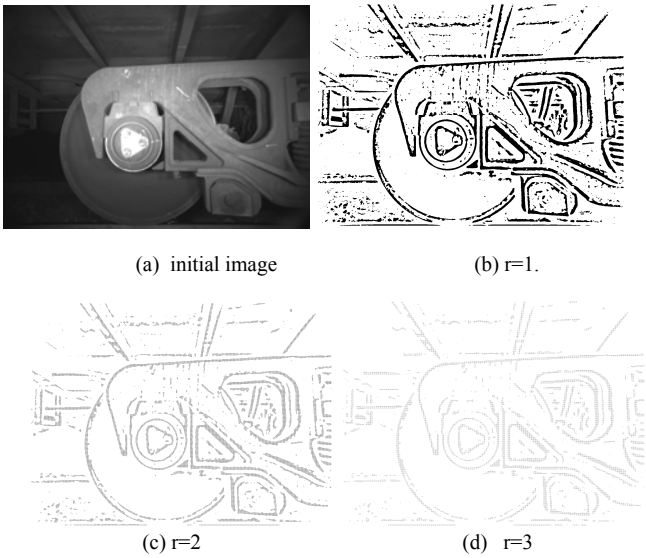


Figure 2. Initial image(1400×1024) and images composed of discrete-point with different sampling radii.

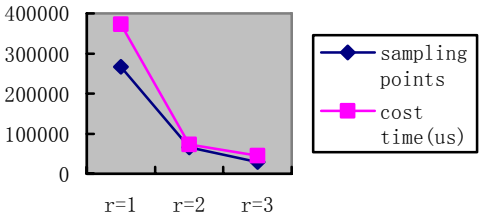


Figure 3. Relationship among sampling radius, sampling points and sampling time.

clearly in Fig. 2. It can be seen that as the sampling radius increases, the detailed information of the image is gradually decreasing, but it is still quite easy for human eyes to locate the wheel. If the current task is wheel locating, Fig. 3 (d) with r=3 can provide sufficient information with 25674 yin pixels, which account for only 2% of the total number of pixels in the original image. Fig. 3 gives the relationship among sampling radius, sampling points and sampling time in Fig.2. As shown in the figure, when sampling radius increases from 1 to 2, the number of sampling points reduces by 75.2%, and sampling time shortens by 80.3%. Taking into account the convenience of grouping, we use r=2 as default sampling radius in our applications.

B. Sampling Model

In this paper, we define yin points are pixels located in the region with uneven brightness, while yang points are pixels located in the region with even brightness. This involves how we perceive brightness contrasts. Our sampling model based on the perception of brightness contrast takes the following two factors into account:

1) The human visual system mainly perceives the changes of brightness rather than brightness itself. The

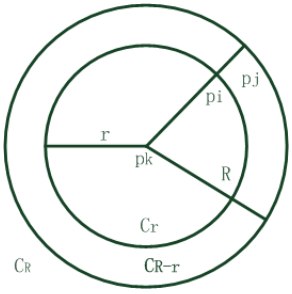


Figure 4. The diagram of discrete-point sampling model.

psychological brightness of a surface is generally determined by the relationship between the brightness of itself and its surroundings. The sampling model should contain such relationship.

2) There are certain limits for human eyes to perceive the change of brightness. Firstly, the brightness difference between two images must be no less than $\triangle B_{min}$, which is the minimum brightness difference perceivable for human eyes. And however the ambient brightness B changes, the ratio $\triangle B_{min}$ /B remains the same and equals to a constant. This constant should be regarded as a threshold in the sampling model to tell whether the discrete point is in a region with even brightness or not.

The second factor is based on the Weber-Fechner Law[11]. The minimum brightness difference perceivable for human eyes $\xi = \triangle B_{min}/B$, is called contrast sensitivity threshold or Weber-Fechner Ratio. Generally, for human eyes, $\xi = 0.005 \sim 0.02$. In case of extremely high or low brightness, ξ can rise up to 0.05. When watching TV, ξ may be larger due to the effects of stray lights. All these can be taken as a reference for the value selection of ξ in the sampling model.

Comprehensively considering the above two factors, a sampling model for discrete points is further developed in this paper, as shown in Fig. 4. In this model, C_r and C_R represent the inner circle and outer circle of concentric circles, and their radiuses are r and R respectively. $R \geq r+1$. C_{R-r} is a concentric ring. pi is the No. i pixel in C_r, and gi is the brightness value of p_i. p_j is the No. j pixel in C_{R-r}, and g_j is the brightness value of p_j. r is chosen as the sampling radius for discrete-point, which means the distance between each sampling point and its 4-neighborhood sampling points is r. The values of r and R are related to the type and size of the object to be detected.

According to the Weber-Fechner Law, the difference between mean brightness of ring region C_{R-r} and mean brightness of inner circular region C_r is the brightness difference in the neighborhood of the center point. This brightness difference divided by the mean brightness of inner circular region C_r is $U(p_k)$, the relative brightness uniformity in the neighborhood of p_k . So the sampling operator for p_k is developed as follows:

$$U(p_k) = \frac{\left| \frac{1}{C_{R-r}} \sum_{p_j \in C_{R-r}} g_j - \frac{1}{C_r} \sum_{p_i \in C_r} g_i \right|}{\frac{1}{C_r} \sum_{p_i \in C_r} g_i} \quad (1)$$

Next, Weber-Fechner Ratio ξ is introduced as a threshold to evaluate the brightness uniformity of p_k's neighborhood, Sampling operator for yin discrete-point:

$$T(p_k) = \begin{cases} 1 & U(p_k) > \xi \\ 0 & U(p_k) \leq \xi \end{cases} \quad (2)$$

Sampling operator for yang discrete-point:

$$T(p_k) = \begin{cases} 1 & U(p_k) < \xi \\ 0 & U(p_k) \geq \xi \end{cases} \quad (3)$$

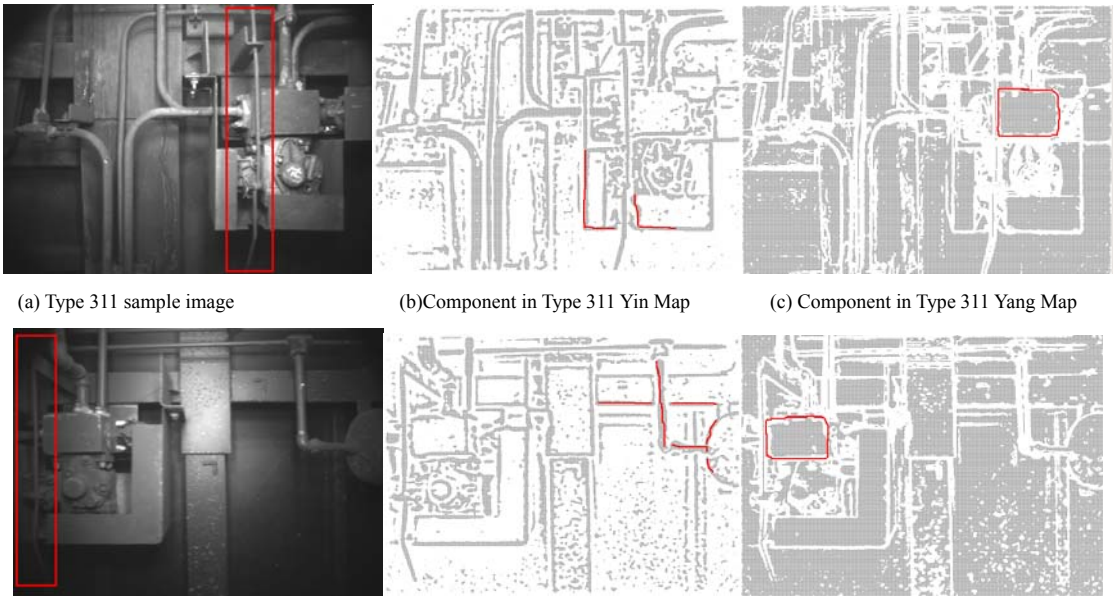
In applications, the sampling map (yin or yang), sampling radius r and the threshold ξ should be appropriately selected according to the type and size of objects to be detected, and noise level in the image. If the object is a plane, yang sampling map should be choose and Formula (1), (3) should be used for sampling the original image. If the object is a line or a point, yin sampling map should be choose and Formula (1), (2) should be used for sampling the original image. Sampling radius r should be determined by the size of objects. Generally speaking, the larger the size of object to be detected is, the larger the sampling radius is, and the common value of r is 2 in our applications. The threshold ξ to evaluate the brightness uniformity in the neighborhood is affected by the overall noise level of the image. Higher noise level requires larger ξ to restrain noise interference; otherwise, the value of ξ can be smaller, and the common value of ξ is 0.05 in our applications. In practice, there is no need to select the optimal Weber-Fechner Ratio ξ for the human eyes. If the entire photo is too dark to identify the details, ξ can be decreased to increase detailed information, which is contrary to physiological property of human eyes. If the photo has too many details and contains a lot of noise, ξ can be increased to reduce unnecessary information.

III. Application

Based on the DPSM, a new technique is invented to automatically classify the components of freight cars and locate the regions with potential faults. This technique is efficient, accurate, simple and general. This technique can judge the structural type of the component in the image and locate the regions with potential faults in 200~300 ms. Meanwhile, it is applicable to all parts of the freight car, which lays a solid foundation for the further research on automatic fault detection of freight cars. The procedures of this technique are as follows:

Step 1: Select learning samples for all types of components in a certain part of freight cars, and at least one learning sample for each type of component.

Step 2: After careful observation of the yin sampling maps and yang sampling maps from learning samples, some components in these learning samples are selected and their contours are drawn up by hand in their sampling maps. At least one component's contours should be drawn up for each learning sample. The contours are not necessarily continuous or closed. The standard for drawing up the contours of the component is to clearly distinguish the current learning sample from other



(a) Type 311 sample image (b)Component in Type 311 Yin Map (c) Component in Type 311 Yang Map

(d) Type 312 sample image (e)Component in Type 312 Yin Map (f) Component in Type 312 Yang Map

Figure 5. Contours of the component in Part n_3_1 drawn by hand in yin sampling maps.

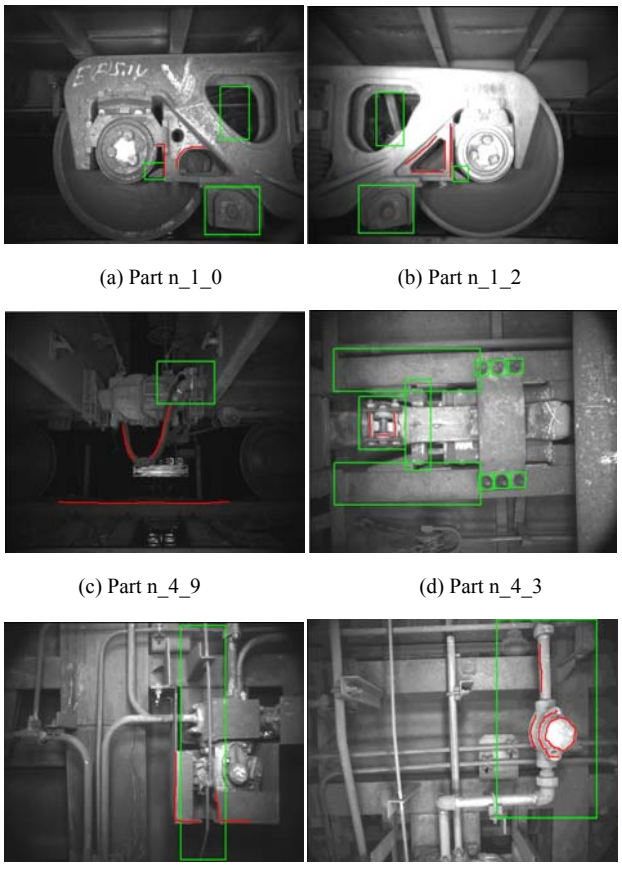
learning samples within the same part of freight cars. Fig. 5 shows the contours of the component in Part n_3_1 drawn by hand in yin –yang sampling maps.

Step 3: Store the information about the hand-drawn contours of all components from all learning samples within the same part of freight car in the format of *.ink data file, and prepare the ink data file for this part of freight car. The information stored in *.ink data file mainly includes: yin and yang location sign of each component in each sample, each component's code, the code of the part that the components belong to, the code of structural type that the part belongs to, number of fault regions, serial number of learning sample in the part, coordinates of a component locating point in its learning sample, search radius of component locating point, type of component locating point, key length for component locating and coordinate sequence of component contour points, etc.

Step 4: Select an *.ink data file for a test image according to which part of freight car it shows, and then calculate the yin or yang sampling map of test image using formula 2 or 3 according to the yin and yang location sign of every component in this *.ink data file. Yin and yang sampling maps might be calculated only once for each test image.

Step 5: In the yin or yang sampling map of test image, take the coordinates of component locating point in its learning sample as the search center, search the component locating point in the test image within the search radius of the component locating point according to the type of component locating point and key length for component locating. Only those points that can meet the key length for component locating and correspond to the type of component locating point can be regarded as the potential component locating points in the test image.

Step 6: Load the coordinate sequence of component contour points onto all the potential component locating points of the test image, i.e. map the coordinate sequence of component contour points in the corresponding learning sample into yin or yang sampling map of test image for matching. Suppose N_m is the actual number of matching points, N_{ink} is the number of points in the coordinate sequence of component contour points. If this mapping receives the "response" of yin or yang sampling map, which means there are yin or yang sampling points appearing in the mapping region, then these points will be



(a) Part n_1_0 (b) Part n_1_2

(c) Part n_4_9 (d) Part n_4_3

(e) Part n_3_1 (f)Part n_3_4

Figure 6. Classification and locating results in different parts of TFDS.

counted as matching points. The total number of matching points N_m divided by the number of points in the coordinate sequence of component contour points N_{ink} is the match ratio of the test image to the contour of this component, as shown in Formula 6:

$$\lambda = \frac{N_m}{N_{ink}} \tag{6}$$

Step 7: Rank the order of the match ratios calculated according to all the coordinate sequences of component contour points. The component type with highest match ratio is most likely to be the component type existing in the test image. Then the part's structural type in the test image can be determined according to the code of the part that this component type belongs to and the code of structural type that the part belongs to.

Step 8: Locating of region with potential faults. After the proper classification of structural type in a certain part of freight car, it becomes easier to locate the regions with potential faults according to its structural type. First, these regions are marked in the above learning samples, and information about the type and location of faults is also recorded in the ink data file that contains the information about components in learning samples. After the component in the test image is located, the location of its regions with potential faults can be confirmed as well. Fig. 6 shows the location of regions with potential faults in different part s.

When preparing the ink files, attention should be paid to the following aspects and the accuracy of classification may be effectively improved:

1. Selecting component types with salient differences;

Based on the visual saliency of human eyes, component types with significant differences should be selected, and the conditions for locating points should be set reasonably, so as to achieve an initial acceptable result of classification and locating.

2. Selecting typical sample image for each type of component;

When selecting learning samples for drawing component contour, pay attention to the representativeness of these sample images, and use the least sample images to ensure the search ranges of locating points can cover all the possible regions and the initial conditions of locating points can be met by most test images except for special ones.

3. Appropriately drawing component with auxiliary structural difference

If there are lots of errors appear when classify two groups of structural types , keep the component's contours with significant difference, and draw contours of another component with auxiliary structural difference, so as to improve the accuracy of classification.

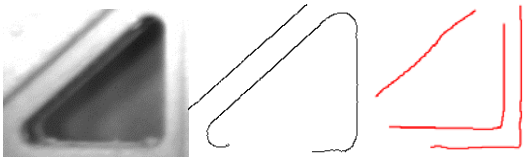
IV. Results and Discussion

In this section, we will compare our method with several classical template matching methods, and we will test its performances in classifying and locating, addressing noise interfering, and performing extensible learning. We will use four experiments, including matching comparison, locating test, noise test and extensibility test. Our test environment utilizes an Intel Core i5-2540 2.60GHz CPU , 4 GB RAM, and VC++2010.

TABLE I
MATCHING COMPARISON RESULTS WITH 743 IMAGES

Methods	Accuracy(%)	Average-time (ms)	Mean	Variance	Size (Bytes)
Gray matching 1	99.46	314.64	0.9753	0.000502	200×160
Gray matching 2	99.73	306.96	0.9568	0.001380	200×160
Gray matching 3	100.00	306.11	0.9925	0.000035	200×160
Edge matching	100.00	6955.88	0.9820	0.000005	190×20
Out method	100.00	154.22	0.9300	0.004109	827×4+3

(a) Initial image (1400×1024); (b) edge map; (c) yin DPSM;



(d) gray template; (e) edge template (f) hand-sketch template.

Figure 7. Three different template matching methods with their templates.

A. Matching Comparison

A total of 744 images with good quality in a component type I of the wheel part as shown in Fig. 3 (a) are selected from a TFDS station. One image is randomly chosen to be the learning sample, as Fig. 7 (a) shows; the chosen image is used to generate an edge map (Fig. 7 (b)) and a yin DPSM (Fig. 7 (c)). The triangular region in the middle of the image is regarded as the salient component; it is marked as component type I of a wheel part, which is used to generate a gray template (Fig. 7 (d)), an edge template (Fig. 7 (e)) and a hand-sketch template (Fig. 7 (f)). All of these three templates are used for matching and locating in the whole range (1400×1024) of the remaining 743 images. All of the best matching areas are cut out as sub-images, to be checked by manual inspection. If there is a whole triangular region, such as Fig. 7 (d) in a sub-image, this matching is correct. If there

TABLE II
REGIONS LOCATING RESULT

Type	151	152	153	154	155	total
Photos' number	773	229	38	33	77	1150
Correct rate(%)	96.9	97.4	97.4	100	96.1	97.04
Time(ms)	231	245	245	284	257	239

is no such triangular region or part of it, then this matching is incorrect.

Table I records, for the three templates, the correct rate, the average time cost, the matching rate's mean and variance and the template's size. The gray template matching uses the cvMatchTemplate function in OpenCV, where schemes 1, 2, and 3 correspond to SQDIFF_NORMED, CCOEFF_NORMED, and CCORR_NORMED, respectively. The template-matching methods that are based on the gray correlation are easily affected by illumination changes or occlusions. Edge matching based on distance transformations [12]-[13] can overcome such effects, while it is computationally expensive and its real-time performance is still not satisfactory. In the method that we proposed in this paper, it costs 90 ms on average to generate a yin DPSM of a whole image, and 65 ms for single hand-sketch template matching. Thus, with half the time of the gray template and 10% of its size, our method not only makes the same matching and positioning result as the gray matching scheme 3 and edge matching do, but also prepares well for subsequent classifications and locating.

B. Test on the Locating of Potential Fault Regions

In this test, the DPSM is applied to locate potential fault regions in wheel-part images from TFDS. 1,150 wheel-part images of the freight car are randomly selected from the total images taken in one day by a test station. These images are used as test samples, including 5 different structural types. Each image is in the original size of 1400×1024. The locating technique based on DPSM is adopted to locate the regions with potential faults in all the wheel images. The accuracy and efficiency of this technique is tested, and the test result is recorded according to the type of the wheel, as shown in TABLE II.

The traditional methods of matching locating by gray-level template or edge template cannot predict the structural types of wheel part, and thereby fail to select the correct template in advance. The locating technique based on DPSM proposed in this paper can simultaneously accomplish such three tasks as detection, locating and classification and reach high efficiency. According to TABLE II., locating accuracy of all five wheel types has reached over 95%, and the average locating accuracy of all the images reaches 97.04%. The average time for one image is equivalent to the time cost by matching with a single gray-level template. Therefore, both classification accuracy and locating efficiency of the technique in this paper can meet the requirements of automatic fault detection of freight cars.

TABLE III
CLASSIFICATION RESULT WITH RANDOM NOISE

SNR	151 (773)	152 (229)	153 (38)	154 (33)	155 (77)	total (1150)
∞	98.06	99.56	100	100	96.10	98.35
25.70	97.67	99.56	97.37	100	97.40	98.09
19.95	98.71	98.69	97.37	96.97	97.40	98.52
17.88	97.93	97.82	78.95	93.94	90.91	96.70
15.02	97.67	93.01	84.21	78.79	84.42	94.87
13.59	92.11	79.91	55.26	45.45	70.13	85.65
13.29	79.56	68.99	28.95	21.21	58.44	72.70

TABLE IV
CLASSIFICATION RESULT WITH RANDOM NOISE

SNR	151 (773)	152 (229)	153 (38)	154 (33)	155 (77)	total (1150)
21.43	99.61	97.82	97.37	100	94.81	98.87
18.50	98.06	60.70	73.68	9.09	20.78	82.09
16.80	97.41	37.12	65.79	3.03	14.29	76.09

TABLE V
RESULTS OF COMPONENT CLASSIFICATIONS IN FIVE PARTS

Parts	n_1_0	n_1_2	n_4_3	n_3_1	n_3_4
primitive types	5	5	5	6	8
no primitive proportion %	0	0	2.6	12.6	11.8
Correct rate %	99.46	99.39	96.21	92.64	91.35

C. Noise Test

After testing millions of TFDS images, it is confirmed that the rapid classification and locating technique based on the DPSM has strong anti- noise capability. In order to analyze the anti- noise capability of image classification and locating method based on DPSM, salt and pepper noise of different signal-to-noise ratios are added to 1150 tested images used in test (a). The classification and locating method based on the DPSM is adopted to classify the images, and experimental results are shown in Table III and Table IV.

According to the experimental result, the classification and locating method based on a DPSM has strong resistance to both random noise and salt-pepper noise. It should be noted that, when the signal-to-noise ratio reduces to a certain degree, Weber-Fechner Ratio ξ can be properly increased to reduce the detailed information and further improve the anti-interference capability of detection system.

D. Extensibility Test

With training, anyone who has experience with a computer can master our hand-sketch shape matching method that is based on a DPSM and can apply it to images of different parts from TFDS. If a new type of component occurs, then we must draw a new hand-sketch template for it, which can be accomplished flexibly by inspectors in each of the TFDS stations according to the actual situations of their recent passing freight cars. Our method has been successfully applied to more than 10

parts in several TFDS stations for more than two years. In the following five-part test, we extract 2000 images for each part randomly from a TFDS, and Fig. 6 shows their classification and locating results. Table IV shows the experimental results of component classifications with mainstream types in five parts of TFDS based on our method. The mainstream types of components refer to the common component types in this part, for which the hand-sketch has been drawn and recorded. Because non-mainstream types refer to some component types that occasionally appear in this part and that differ greatly from the mainstream components, their hand-sketches are not drawn. During classification, those non-mainstream types should be distinguished as a special category, i.e., the category that could not be recognized. However, because of the irregular and wide changes, the non-mainstream types are likely to be detected as mainstream types of components, which lead to errors in the classification results. The experimental results in table IV show that, for parts with small structural changes and without non-mainstream types of components, the accuracy of this classification method can reach over 99.0%; while for parts with complex structure and large proportions of non-mainstream types of components, the accuracy of this classification method decreases to 90%. In this case, the threshold for some false fault alarms should be appropriately decreased, and manual judgments are needed.

V. CONCLUSIONS

The DPSM reflects the transient process from object region to object edge and provides a platform for storing, representing, searching and detecting object structures. To detect the images in a certain part of TFDS is actually to give an overall description of all the image samples in this part. This description not only includes edge information and regional information, but also includes the relative positions of edges and regions and allows the mutual conversion between edges and regions, which may be caused by position deviation, light changes or covered spots. Nevertheless, as long as the component of freight car exists, the transient process from component region to component edge will exist undoubtedly. This transient process is not reflected by a certain pixel or edge, but distributed in a certain region that is not necessarily continuous. In other words, as long as there is a certain region in the yin sampling map or yang sampling map that can contain the contour of a certain component of the freight car, it can be deduced that the original image probably contains this component. Moreover, whenever the components of freight cars are upgraded, the ink data files can be replaced to update the classification and locating algorithm, so this algorithm can serve as an open and standard tool for TFDS. In fact, compiling ink data files is storing knowledge features, the ink files is corresponding to the storage space of knowledge features of freight car components, and mapping the component data in ink files to sampling maps is searching for knowledge features. All these steps are completed based on yin-yang sampling maps. Yin-yang sampling maps can store the knowledge about the point, line or plane objects. Edge graphs and threshold graphs cannot properly demonstrate the knowledge features of the point, line or plane objects in complex and changing backgrounds. DPSM can solve this problem to some extent under certain conditions, provide a platform for object detection to effectively extract knowledge features, and search knowledge features at a very fast speed only in a small storage space. In many cases, accompanied by discrete point grouping algorithm and computing geometric algorithms, the yin-yang sampling maps can replace edge maps or threshold maps as a more efficient and robust model for object detection in a large quantity of images.

ACKNOWLEDGMENT

This work was supported in part by the Natural Science Foundation of China under Grant 60975021.

REFERENCES

[1] Z.H. Liu, D.Y. Xiao and Y.M. Chen, “Displacement fault detection of bearing weight saddle in TFDS based on hough transform and symmetry validation,” in *9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pp. 1404-1408, Chongqing, China, May. 2012.

[2] X.D. Yang, L.J. Ye and J.B. Yuan, “Research of Computer Vision Fault Recognition Algorithm of Center Plate Bolts of Train,” in *1st International Conference on Instrumentation, Measurement, Computer, Communication and Control*, pp.978-981, Beijing, China, Oct. 2011.

[3] Wang, Yanxia, Yan Ma, and Qixin Chen. "A Method of Line Matching Based on Feature Points." *Journal of Software* 7.7 (2012): 1539-1545.

[4] H.Z. Wang and J. Oliensis, “Rigid Shape Matching by Segmentation Averaging,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 4, pp. 619-635, Apr. 2010.

[5] Yang, Huihua, et al. "An Efficient Vehicle Model Recognition Method."*Journal of Software* 8.8 (2013): 1952-1959.

[6] Gao, Tao, et al. "Feature particles tracking for moving objects." *Journal of Multimedia* 7.6 (2012): 408-414.

[7] Zeng, Lin, et al. "A Self-adaptive and Real-time Panoramic Video Mosaicing System." *Journal of Computers* 7.1 (2012): 218-225.

[8] Zhu, Zongxiao, Guoyou Wang, and Jianguo Liu. "Object detection based on multiscale discrete-point sampling and grouping." *Sixth International Symposium on Multispectral Image Processing and Pattern Recognition*. International Society for Optics and Photonics, YiChang, China,2009.

[9] Zhu, Z., et al., “Fast and Robust 2D-Shape Extraction Using Discrete-Point Sampling and Centerline Grouping in Complex Images”. *Image Processing, IEEE Transactions on*, 2013. 22(12): p. 4762-4774.

[10] Ali, Mohamed Ather and Klyne, M.A., “Vision in Vertebrates”. New York: Plenum Press. 1985, pp. 28.

[11] Jianhong Shen, “On the foundations of vision modeling I. Weber's law and Weberized TV restoration,” Physica D: Nonlinear Phenomena, vol.175, no.3-4, pp. 241-251,2003.

[12] R.G. Caves, P.J. Harley and Q. Shuan, “Matching map features to synthetic aperture radar (SAR) images using template matching,”, *IEEE Trans. Geosci. Remote Sensing*, vol. 30, no.4, pp.680-685, Jul. 1992.

[13] J. Kang and G.Yang, “Fast morphological pyramid matching algorithm based on the Hausdorff distance,” in *2011th IEEE International Conference on Cyber*

Technology in Automation, Control, and Intelligent Systems, IEEE, pp.288-292, KunMing, China, Mar. 2011.

Zongxiao Zhu received the BS degree and the MS degree in electrical and electronic engineering from Xi'an Jiaotong University, Xi'an, China in 2000 and 2003, respectively. He is currently pursuing the Ph.D. degree in control science and engineering at the Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science & Technology, Wuhan China.

From 2003 to 2004, he was an engineer with SuZhou Shihlin Electric&Engineering Co., where he was in charge of designing small power converters (on the market in 2004). Since 2004, he has been a faculty of College of computer Science, South-Central University for Nationality (SCUN), Wuhan, China. In 2007, he founded the Information Processing Laboratory for Minority Language (IPLML) in SCUN and began to manage a multidisciplinary research team aiming at using information technology to salvage, protect and broadcast endangered minority cultures. His research interests include image processing, object detection, and endangered minority culture's protection with information technology.

Mr. Zhu is a member ACM and a member of Chinese computer federation (CCF).

Guoyou Wang received BS degree in Electronic Engineering and the MS degree in pattern recognition and intelligent system from Huazhong University of Science and Technology, Wuhan, China in 1988 and 1992, respectively. He is currently a professor with the Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science & Technology, Wuhan China. His research interests include image processing, image compression, pattern recognition, artificial intelligence, and machine learning.

A Sub-1V High-PSRR Piecewise-Linear Bandgap Reference

Qianneng Zhou
College of Electronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China
Email: zhouqn@cqupt.edu.cn

Qi Li, Hongjuan Li, Jinzhao Lin, Yu Pang, Guoquan Li, Lu Deng
Chongqing University of Posts and Telecommunications, Chongqing, China
Email: {lihj1, pangyu}@cqupt.edu.cn

***Abstract*—A piecewise-linear bandagp reference (BGR) with high power supply rejection ratio (PSRR) and low temperature coefficient is designed for analogue and mixed signal systems in this paper. By adopting LDO regulator, the designed high PSRR piecewise-linear BGR achieves well performances and has a simple architecture. Simulation results show that the PSRR of the designed piecewise-linear BGR with LDO regulator at 10Hz, 1kHz and 100kHz achieves, respectively, -110.42dB, -109.18dB and -64.51dB. Compared to the designed BGR without LDO regulator, the PSRR of the designed high PSRR piecewise-linear BGR with LDO regulator has the improvements of about 35dB, 36.9dB and 29.28dB at 10Hz, 1kHz and 100kHz respectively. The designed piecewise-linear BGR with LDO regulator generates an output voltage of 0.68V with 1.65ppm/ºC temperature coefficient in the range from -50 ºC to 125 ºC. The deviation of the output voltage is within 98.23μV when the power supply voltage V_{DD} changes from 1.2V to 7V.**

***Index Terms*—bandgap reference (BGR), piecewise-linear, LDO regulator, power supply rejection ratio (PSRR)**

I. INTRODUCTION

Bandgap reference (BGR) is a very important block for many analogue and mixed signal electronic devices such as data converters, DC-DC converters, DRAMs, linear regulators and so on. The BGR should be independent of fluctuations of power supply voltage and temperature, and also be implemented without modification of fabrication process. In modern CMOS technology, the output voltage of BGR is usually a weighted sum of the forward-bias emitter-base voltage of diode-connected parasitic vertical PNP transistor and the thermal voltage [1]. Conventional BGR has a large temperature coefficient over the whole temperature range and cannot meet the requirements of high precision circuits. To improve the temperature performance of BGR, many temperature compensation techniques have been reported, such as correlated double sampling technique [2], curvature correction [3-6] and so on. Recently, demands for low-voltage BGR circuits have increased enormously because they are widely used in portable electronic applications. Simultaneously, the most significant noise injected to the output of BGR circuit is the supply noise regarding to the other noise. On the other hand, a high power supply rejection ratio (PSRR) BGR is desired to achieve high performance analogue and mixed signal systems, particularly in wireless communications. Therefore, a BGR structure, which has high PSRR over broad frequency range, should be chosen to reject the supply noise coupled from the high-speed digital circuit on the chip.

In the recent past, many approaches have be developed to improve the PSRR of BGR, such as supply independent current source technique [7], pre-regulator technique [8-11], subtractor technique [12], pseudo floating voltage source technique [13], cascode current-mode technique [14], self-cascode current mirror technique [15], low dropout (LDO) regulator technique [16], voltage follower technique with PMOS as input transistor [17] and so on. In general, these reported BGR with PSRR enhancement have achieved well performance. However, to further improve the performance of BGR, the high PSRR BGR structure must still be analyzed and discussed.

In this paper, a high PSRR CMOS BGR with less than 1V output voltage is designed and analyzed. Employing an improved piecewise-linear temperature compensation technique and a LDO regulator, the designed BGR achieves a high PSRR performance over a wide frequency range and a well temperature characteristic over a wide temperature range. The analysis and design of piecewise-linear BGR without LDO regulator will be given in Section II. Section III will discuss the high PSRR piecewise-linear BGR with LDO regulator. Simulation results are described in Section IV. Finally, conclusions are given in Section V.

II. ANALYSIS AND DESIGN OF PIECEWISE-LINEAR BGR

The designed BGR is shown in Fig.1, which consists of a start-up circuit and a core circuit of BGR. There are two possible equilibrium points in the core circuit of BGR, so a start-up circuit is necessary. M_{s1}~M_{s5} form the start-up circuits, as shown in Fig.1 (b). The core circuit of piecewise-linear BGR will provide a sub-1V bandgap

doi:10.4304/jcp.9.5.1274-1281

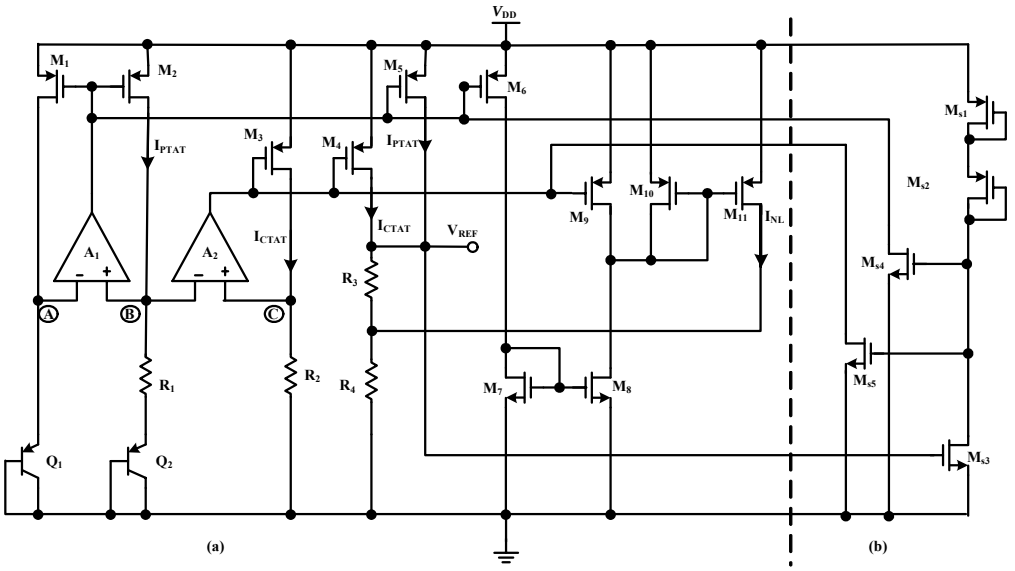


Figure 1. Designed BGR (a) core circuit of BGR, (b) start-up circuit

voltage V_{REF} and has a good temperature characteristic. At the same time, all MOS transistors adopt the long channel transistor so that the channel-length modulation effect is negligibly small in this paper.

For convenience, it is assumed that W_j, L_j and I_j are, respectively, channel-width, channel-length and drain current of transistor M_j in this paper, here j=1, 2, 3…. The core circuit of piecewise-linear BGR consists of transistors M_1~M_{11}, bipolar transistors Q_1 and Q_2, resistors R_1~R_4, and amplifiers A_1 and A_2. Amplifiers A_1 and A_2 are entirely the same, and their dc gain A_d has that A_d>>1. Amplifier A_1 forces the voltages V_A and V_B be equal, and amplifier A_2 forces the voltages V_B and V_C be equal, i.e. V_A=V_B=V_C=V_{EB1}. Here, V_A, V_B, V_C and V_{EB1} are, respectively, the voltages of node A, node B, node C and emitter-base voltage of Q_1. Transistors M_1, M_2 and M_5 are entirely the same, and Q_2 has an emitter area that is m times that of Q_1. So, the drain currents of transistors M_2 and M_5 can be derived as

$$I_2 = I_5 = \frac{kT}{q}\frac{1}{R_1}\ln m \tag{1}$$

where, k is Boltzmann's constant, q is electronic charge, and T is absolute temperature. It is concluded that currents I_2 and I_5 are proportional to the absolute temperature T. Amplifier A_2 force voltages V_B and V_C be equal, and transistors M_3 and M_4 are entirely the same. Therefore, the drain currents of M_3 and M_4, i.e. I_3 and I_4, can be derived as

$$I_3 = I_4 = \frac{V_{EB1}}{R_2} \tag{2}$$

V_{EB1} has a negative temperature coefficient, so it is concluded that currents I_3 and I_4 have a negative temperature coefficient. Transistors M_2 and M_6 have the same aspect ratio, and M_3 and M_9 have also the same aspect ratio. Therefore, it is concluded that I_6=I_2 and I_3=I_9.

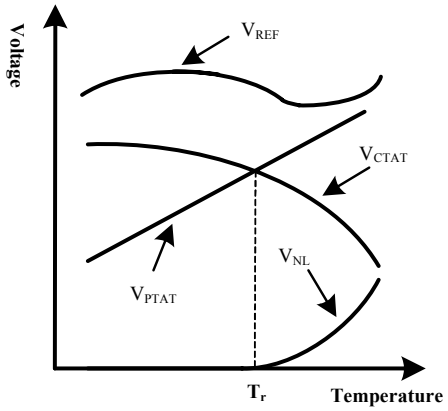


Figure 2. Operation of designed BGR

Transistorss M_7 and M_8 form the current-mirror pair, and transistor M_8 has the aspect ratio that is α times that of transistor M_7. By optimizing the value of parameter α, it can be achieved that I_9=I_8=αI_2 under the condition of room temperature Tr. Therefore, the following expression can be derived as

$$\begin{cases} I_8 = \alpha\dfrac{kT}{q}\dfrac{1}{R_1}\ln m < I_9 = \dfrac{V_{EB1}}{R_2}, when \quad T < T_r \\ I_8 = \alpha\dfrac{kT}{q}\dfrac{1}{R_1}\ln m = I_9 = \dfrac{V_{EB1}}{R_2}, when \quad T = T_r \\ I_8 = \alpha\dfrac{kT}{q}\dfrac{1}{R_1}\ln m > I_9 = \dfrac{V_{EB1}}{R_2}, when \quad T > T_r \end{cases} \tag{3}$$

M_{10} and M_{11} form current-mirror pair and have entirely the same aspect ratio. Therefore, the drain current I_{NL} of transistor M_{11} can be derived as

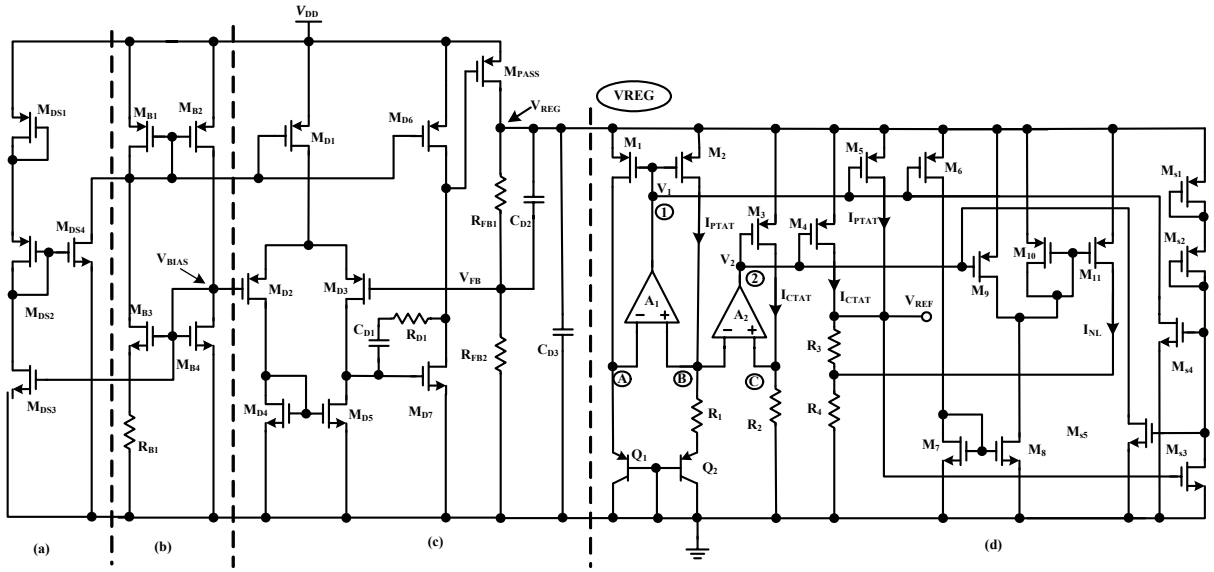


Figure 3. Designed high PSRR BGR with LDO regulator (a) start-up circuit, (b) supply-independent bias circuit, (c) LDO regulator, (d) core circuit of piecewise-linear BGR

$$\begin{cases} I_{NL} = 0, & when \quad T \le T_r \\ I_{NL} = \alpha \dfrac{kT}{q}\dfrac{1}{R_1}\ln m - \dfrac{V_{EB1}}{R_3}, & when \quad T > T_r \end{cases} \tag{4}$$

So, the output voltage V_{REF} of piecewise-linear BGR can be derived as

$$\begin{aligned} V_{REF} &= (R_3 + R_4)\left(\frac{kT}{q}\frac{1}{R_1}\ln m + \frac{V_{EB1}}{R_2}\right) + R_4 I_{NL} \\ &= V_{PTAT} + V_{CTAT} + V_{NL} \end{aligned} \tag{5}$$

$$V_{PTAT} = (R_3 + R_4)\frac{kT}{q}\frac{1}{R_1}\ln m \tag{6}$$

$$V_{CTAT} = (R_3 + R_4)\frac{V_{EB1}}{R_2} \tag{7}$$

$$V_{NL} = R_4 I_{NL} \tag{8}$$

According to (1) ~ (8), it is concluded that V_{PTAT}, V_{CTAT} and V_{NL} are, respectively, the voltages with positive-temperature coefficient, negative-temperature coefficient and piecewise temperature coefficient, and their relations are shown in Fig.2. Equation (5) shows that the piecewise-linear BGR can achieve a low temperature coefficient bandgap reference voltage V_{REF} by optimizing resistors R_1~R_4 and parameter α in theory. However, the operating supply voltage of piecewise-linear BGR, which is shown in Fig.1, is the power supply voltage V_{DD}, and it cannot achieve high PSRR over a wide frequency range. Therefore, the piecewise-linear BGR cannot be effectively applied to analogue and mixed signal systems that have the requirement of high PSRR. Therefore, to further improve PSRR of BGR, a high PSRR piecewise-linear BGR is designed by adopting LDO regulator in this paper, and will be analyzed in Section III.

III. Analysis and Design of High PSRR Piecewise-Linear BGR

Fig.3 shows the designed high PSRR piecewise-linear BGR, and all MOS transistors adopt the long channel transistor so that the channel-length modulation effect is negligibly small. The designed high PSRR piecewise-linear BGR consist of a start-up circuit, a supply-independent bias circuit, a LDO regulator and a core circuit of piecewise-linear BGR. The core circuit of piecewise-linear BGR is entirely the same as that designed in Section II, but whose operating supply voltage is the output voltage V_{REG} of LDO regulator instead of power supply voltage V_{DD}. Therefore, the designed high PSRR piecewise-linear BGR with LDO regulator can achieve an output voltage V_{REF} with low temperature coefficient and high PSRR. The supply-independent bias circuit produces supply-independent bias voltages, and will be discussed in Section III.A. Because there are two possible equilibrium points in the supply-independent bias circuit, a start-up circuit is necessary. M_{DS1}~M_{DS4} form the start-up circuits, as shown in Fig.3 (a). The function of LDO regulator is to produce an internally regulated voltage V_{REG} that is the operating supply voltage of core circuit of piecewise-linear BGR instead of power supply voltage V_{DD}. The analysis and design of LDO regulator will be given in Section III.B.

A. Supply-Independent Bias Circuit

As shown in Fig.3 (b), the supply-independent bias circuit consists of M_{B1}~M_{B4} and R_{B1} [18]. MOS transistors M_{B1}~M_{B4} operate in the saturation region, and M_{B1} and M_{B2} are entirely the same. The channel lengths of M_{B3} and M_{B4} are the same, but M_{B3} has a channel width that is N times that of M_{B4}. Therefore, the drain currents I_{B3} and I_{B4} can be expressed as

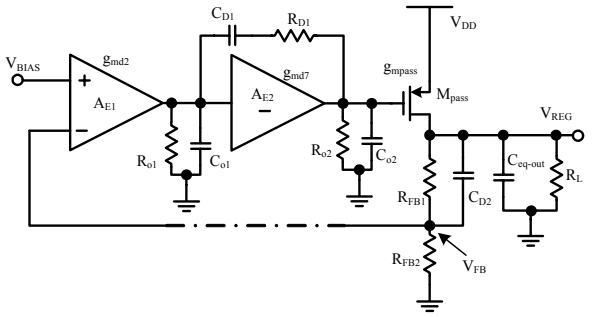


Figure 4. Topologic architecture of LDO regulator

$$I_{B3} = I_{B4} = \frac{2}{\mu_n C_{ox}(W_{B4}/L_{B4})} \frac{1}{R_{B1}^2}\left(1-\frac{1}{\sqrt{N}}\right)^2 \quad (9)$$

where, μ_n is the mobility of an electron, C_{ox} is the gate oxide capacitance per unit area. Equation (9) indicates that the bias-current I_{B4} is independent of the power supply voltage V_{DD}. Therefore, the gate-voltage V_{BIAS} of MOS transistor M_{B4} can be derived as

$$V_{BIAS} = \frac{2}{\mu_n C_{ox}(W_{B4}/L_{B4})} \frac{1}{R_{B1}}\left(1-\frac{1}{\sqrt{N}}\right)+V_{THN} \quad (10)$$

where, V_{THN} is the threshold voltage of NMOS transistor. Equation (10) indicates that the bias voltage V_{BIAS} is also independent of the power supply voltage V_{DD}.

B. Analysis and Design of LDO Regulator

The designed LDO regulator is shown in Fig.3 (c), which consists of error amplifier, a PMOS power transistor M_{pass}, and a feedback resistive network. The feedback resistive network consists of capacitor C_{D2}, and resistors R_{FB1} and R_{FB2}. The error amplifier consists of MOS transistors M_{D1}~M_{D7}, resistor R_{D1} and capacitance C_{D1}. The error amplifier compares the reference voltage V_{BIAS}, which is provided by the supply-independent bias circuit, with the feedback voltage V_{FB} that is provided by the feedback resistors R_{FB1} and R_{FB2}, and generates an error voltage signal which is fed into the gate of power transistor M_{pass} to change its over-drive. The over-drive adjusts the drain current of M_{pass} and forces the output voltage V_{REG} of LDO regulator to be corrected to the proper level. The error amplifier and power transistor M_{pass} form a negative feedback system, which is equivalent to a three-stage amplifier negative feedback system. Therefore, the open-loop stable of LDO regulator is critical issue.

To analyze the stability of LDO regulator, the open-loop transfer function of LDO regulator should be analyzed and discussed. For convenience, the topologic architecture of LDO regulator is shown in Fig.4. The first-stage error amplifier A_{E1} consists of MOS transistors M_{D1}~M_{D5}, whose equivalent input transconductance is written as g_{md2}. The second-stage amplifier A_{E2} consists of MOS transistors M_{D6} and M_{D7}, whose equivalent input transconductance is written as g_{md7}. R_{o1} and R_{o2} are the output resistance of amplifier A_{E1} and A_{E2} respectively, and C_{o1} and C_{o2} are the parasitic capacitances at the output of A_{E1} and at the gate of M_{pass} respectively. g_{mpass} is the equivalent transconductance of M_{pass}, and R_L is the equivalent load resistance. In Fig.4, $C_{eq\text{-}out}=C_{D3}+C_L$, here C_L is the equivalent capacitance of internal power line.

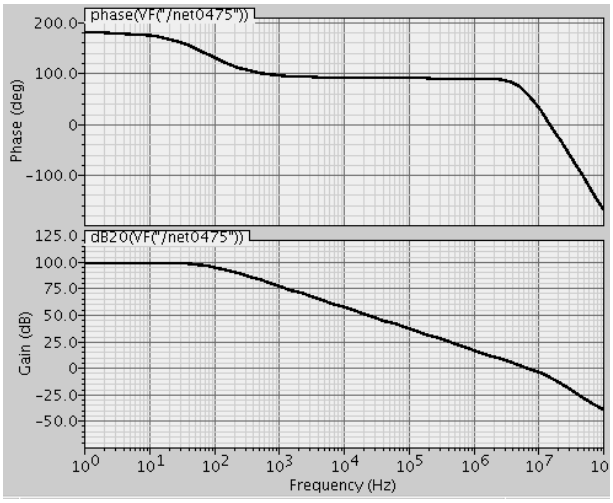


Figure 5. Open-loop frequency response of LDO regulator

To simplify the transfer function without losing accuracy with the goal of providing a clearer insight into the designed LDO regulator structure, it is assumed that capacitors C_{D1}, C_{D2} and $C_{eq\text{-}out}$ are much greater than the parasitic capacitance C_{o1} and C_{o2}, and the gains of each stage is much greater than one, i.e. $g_{md2}R_{o1}$, $g_{md7}R_{o2}$ and $g_{mpass}R_L>>1$. On the other hand, the feedback resistance R_{FB2} is much greater than the load resistance R_L, and R_{o1} and R_{o2} are greater than R_{D1}. Therefore, the loop transfer function of the designed LDO regulator can be approximated to

$$T_{loop}(s) = \frac{T_0(1+s/z_1)(1+s/z_f)}{(1+\frac{s}{p_{-3dB}})(1+\frac{s}{p_1})(1+\frac{s}{p_2}+\frac{s^2}{p_2 p_3})(1+\frac{s}{p_f})} \quad (11)$$

$$T_0 = g_{md2}R_{o1}g_{md7}R_{o2}g_{mpass}R_L\frac{R_{FB2}}{R_{FB2}+R_{FB1}} \quad (12)$$

$$p_{-3dB} = \frac{1}{g_{md7}R_{o1}R_{o2}C_{D1}} \quad (13)$$

$$p_1 = \frac{1}{C_{eq-out}R_L} \quad (14)$$

$$p_2 = \frac{g_{m2}}{C_{o1}+C_{o2}} \quad (15)$$

$$p_3 = \frac{1}{R_{D1}(C_{o1}//C_{o2})} \quad (16)$$

$$p_f = \frac{1}{C_{D2}(R_{FB1} // R_{FB2})} \quad (17)$$

$$z_1 = \frac{1}{C_{D1}(R_{D1} - 1/g_{md7})} \quad (18)$$

$$z_f = \frac{1}{R_{FB1}C_{D2}} \quad (19)$$

C_{o1} and C_{o2} are the lumped capacitance, so the non-dominant poles p_2 and p_3 will shift to a higher frequency than the unity-gain frequency (UGF). To cancel the effect of non-dominant poles in the designed LDO regulators, the zero z_f should be lower than poles p_1 and p_f, so R_{FB2} should be much smaller than R_{FB1}, i.e. $z_f << p_f$. That is to say, the effect of the pole p_1 can be cancelled by z_f. To ensure z_1 be left-plane zero, R_{D1} should be selected much larger than $1/g_{md7}$. At the same time, since C_{D1} and C_{D2} are the compensation capacitor, it is practical to take the assumption of $z_1 < p_f$ by optimizing resistors R_{D1} and R_{FB1}, and compensation capacitors C_{D1} and C_{D2}. From the above discussion, the LDO regulator will be stable because it is similar to a single pole system. Fig.5 shows the simulated open-loop frequency response of the designed LDO regulator. Simulation results show that the phase margin is about 61°, which is sufficient to ensure the loop stability of LDO regulator.

C. Analysis of PSRR

To improve the PSRR of piecewise-linear BGR, a LDO regulator is adopted in this paper, as shown in Fig.3. The operating supply voltage of core circuit of piecewise-linear BGR is the output voltage V_{REG} instead of power supply voltage V_{DD}. Therefore, the PSRR of piecewise-linear BGR with LDO regulator will be improved. Under the condition of low frequency, the PSRR can be quantitatively analyzed as follows.

For convenience, it is assumed that g_{mj} and i_{mj} are, respectively, the transconducatnce and the small-signal drain current of transistor M_j, here j=1, 2, 3…. Assumed that power supply voltage has an incremental variation v_{dd}, the incremental current i_{b1} of M_{B1} can be derived as

$$i_{b1} = \frac{v_{dd}}{\frac{1}{g_{mb1}} + g_{mb3}R_{B1}r_{dsb3}} \quad (20)$$

where, r_{dsb3} is the source-drain resistance of M_{B3}. M_{B1} and M_{B2} form current mirror pair, and they are entirely the same. Therefore, the gate-source variation v_{bias} of M_{B4} can be derived as

$$v_{bias} = \frac{v_{dd}}{g_{mb4}(\frac{1}{g_{mb1}} + g_{mb3}R_{B1}r_{dsb3})} \quad (21)$$

Assumed that v_{reg} is the output voltage variation of LDO regulator, and the feedback voltage variation v_{fb} can be derived as

$$v_{fb} = \frac{v_{reg}}{R_{FB1} + R_{FB2}} R_{FB2} \quad (22)$$

As shown in Fig.3, the error amplifier of LDO regulator is made up of MOS transistors M_{D1}~M_{D7}, resistor R_{D1} and compensation capacitor C_{D1}. To simple the analysis, it is assumed that the dc gain A_v of error amplifier is far greater than 1, i.e. $A_v >> 1$. Neglected the effect of drain current variation of M_{D1} and M_{D6}, the gate voltage variation v_{gpass} of power transistor M_{PASS} can be derived as

$$v_{gpass} = \frac{A_v R_{FB2} v_{reg}}{R_{FB1} + R_{FB2}} - \frac{A_v v_{dd}}{g_{mb4}(\frac{1}{g_{mb1}} + g_{mb3}R_{B1}r_{dsb3})} \quad (23)$$

According to (20) ~ (23) and the Kirchhoff current law (KCL) at the output node VREG of the LDO regulator, it is derived as

$$\frac{v_{reg}}{v_{dd}} = g_{mpass}R_{eq-L}\frac{1 + \frac{A_v}{g_{mb4}(\frac{1}{g_{mb1}} + g_{mb3}R_{B1}r_{dsb3})}}{1 + g_{mpass}R_{eq-L}\frac{A_v R_{FB2}}{R_{FB1} + R_{FB2}}} \quad (24)$$

where, R_{eq-L} is the equivalent resistance seen from node VREG to ground. In the similar way, it is assumed that v_a, v_b and v_1 are, respectively, the voltage variations at node A, node B and node 1. So, v_a and v_b can be derived as

$$v_a = g_{m1}(v_{reg} - v_1)r_a \quad (25)$$

$$v_b = g_{m2}(v_{reg} - v_1)r_b \quad (26)$$

where, r_a and r_b are the resistance seen from node A and node B to ground respectively. Amplifier A_1 and A_2 are entirely the same, and whose dc gain A_d is far greater than 1, i.e. $A_d >> 1$. In Fig.3, the voltage variation v_1 at node 1 has that $v_1 = A_d \times (v_b - v_a)$. According to (25) and (26), it is derived as

$$v_1 = \frac{A_d g_{m1}\beta v_{reg}}{1 + A_d g_{m1}\beta} \quad (27)$$

where, $\beta = r_b - r_a$. MOS transistors M_1, M_2, M_5 and M_6 are entirely the same, and it is concluded that $g_{m1} = g_{m2} = g_{m5} = g_{m6}$. Therefore, it is derived as

$$i_{1,2,5,6} = g_{m1}\frac{1}{1 + A_d g_{m1}\beta}v_{reg} \quad (28)$$

The voltage variation v_2 at node 2 has that $v_2 = A_d \times (v_c - v_b)$, here v_c is voltage variation at node C. MOS transistors M_3, M_4 and M_9 are entirely the same, and it is concluded that $g_{m3} = g_{m4} = g_{m9}$. In the similar way, it is derived as

$$i_{3,4,9} = g_{m3}\frac{1 + A_d g_{m1}\beta + A_d g_{m1}r_b}{(1 + A_d g_{m1}\beta)(1 + A_d g_{m4}R_2)}v_{reg} \quad (29)$$

Transistors M_7 and M_8 form the current mirror pair, and transistor M_8 has the aspect ratio that is α times that

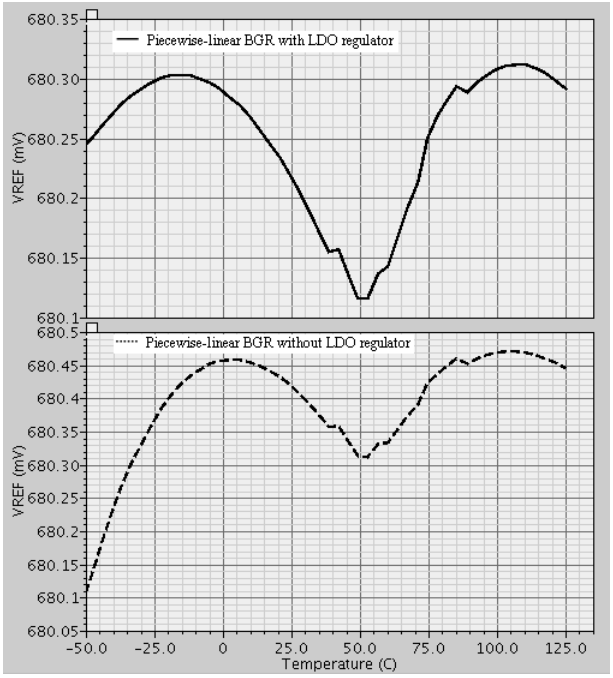


Figure 6. Output voltage V_{REF} of piecewise-linear BGR with- and without- LDO regulator as a function of temperature T

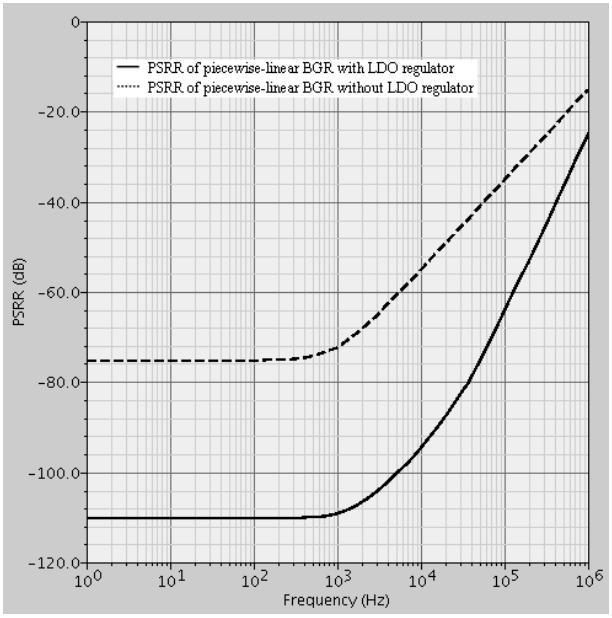


Figure 7. Simulated PSRR of piecewise-linear BGR with- and without- LDO regulator

of transistor M_7. Transistors M_{10} and M_{11} are entirely the same, so it is derived as

$$i_{11} = \alpha g_{m1} \frac{1}{1+A_d g_{m1}\beta} v_{reg} - g_{m3} \frac{1+A_d g_{m1}\beta + A_d g_{m1} r_b}{(1+A_d g_{m1}\beta)(1+A_d g_{m4} R_2)} v_{\text{reg}} \tag{30}$$

To simple the analysis, it is assumed that $A_d g_{m1}\beta >> 1$ and $A_d g_{m4} R_2 >> 1$. According to (28) ~ (30) and KCL at the output of BGR, it is derived as

$$\frac{v_{ref}}{v_{\text{reg}}} \approx R_3 \frac{\beta + r_b}{\beta A_d R_2} + \frac{R_3 + (1+\alpha)R_4}{A_d \beta} \tag{31}$$

where, v_{ref} is the output voltage variation of piecewise-linear BGR with LDO regulator. Therefore, the PSRR of the designed piecewise-linear BGR with LDO regulator can be written as

$$PSSR_{dB} = 20\lg\left|\frac{v_{ref}}{v_{vdd}}\right| = 20\lg\left|\frac{v_{ref}}{v_{reg}} \times \frac{v_{reg}}{v_{dd}}\right| = 20\lg\left|\frac{v_{ref}}{v_{reg}}\right| + 20\lg\left|\frac{v_{reg}}{v_{dd}}\right| \tag{32}$$

According to (24), (31) and (32), it is concluded that the PSRR of the designed piecewise-linear BGR will be improved significantly by adopting the LDO regulator.

IV. Simulation Results

To confirm the circuit of the designed piecewise-linear BGR in this paper, the piecewise-linear BGRs with- and without- LDO regulator are designed and simulated in SMIC 0.18μm CMOS process with 1.35-V power supply voltage.

Fig.6 shows the output voltage V_{REF} of piecewise-linear BGR with- and without- LDO regulator as a function of temperature T with 1.35-V power supply voltage. Simulation results show that the temperature coefficient of the piecewise-linear BGR without LDO regulator is 2.89ppm/ºC when temperature varying from -50 ºC to 125 ºC. And, the output voltage V_{REF} of piecewise-linear BGR with LDO regulator has only the temperature coefficient of 1.65ppm/ºC.

The simulated PSRR of piecewise-linear BGR with- and without- LDO regulator is shown Fig.7. The piecewise-linear BGR with LDO regulator at 10Hz, 100Hz, 1kHz, 10kHz and 100kHz achieves , respectively, -110.42dB, -110.41dB, -109.18dB, -94.65dB and -64.51dB. And the piecewise-linear BGR without LDO regulator at 10Hz, 100Hz, 1kHz, 10kHz and 100kHz achieves, respectively, -75.35dB, -75.31dB, -72.28dB, -55.19dB and -35.23dB. Compared to the piecewise-linear BGR without LDO regulator, the designed high PSRR piecewise-linear BGR with LDO regulator has an improvement of PSRR with about 35dB, 36.9dB and 29.28dB at 10Hz, 1kHz and 100kHz respectively. Therefore, the PSRR improvement is achieved by adopting LDO regulator.

Fig.8 shows the simulated line-regulations of piecewise-linear BGR with- and without- LDO regulator. When power supply voltage V_{DD} varies from 1.2V to 7V, the output voltage deviation of piecewise-linear BGR without LDO regulator is 3.83mV, but the output voltage deviation of piecewise-linear BGR with LDO regulator is only 98.23μV. Compared to the piecewise-linear BGR without LDO regulator, the piecewise-linear BGR with

TABLE I.
PERFORMANCE SUMMARY OF BANDGAP REFERENCE

		Reported BGR				Designed BGR in this paper	
		[4]	[5]	[16]	P[17]	without LDO regulator	with LDO regulator
Process		0.5μm BiCMOS	0.13μm CMOS	0.25μm BiCMOS	0.09μm CMOS	0.18μm CMOS	0.18μm CMOS
Power supply voltage (V)		1.6	0.9	2.7	2.7	1.35	1.35
Output voltage		1.285 V	0.615 V	1.26 V	213.982 mV	0.68 V	0.68 V
Temperature coefficient (ppm/ ºC)		7.2	6.5	276.67	6.071	2.89	1.65
Temperature range (ºC)		-40 ~ 100	-50 ~ 150	0 ~ 100	-20 ~ 120	-50 ~ 125	-50 ~ 125
PSRR@ 25 ºC	10Hz	-70 dB			-82.7 dB	-75.35 dB	-110.42 dB
	100Hz			-82 dB		-75.31 dB	-110.41 dB
	1kHz	-69 dB		-79 dB		-72.28 dB	-109.18 dB
	10kHz			-64 dB		-55.19 dB	-94.65 dB
	100kHz	-35 dB		-48 dB		-35.23 dB	-64.51 dB

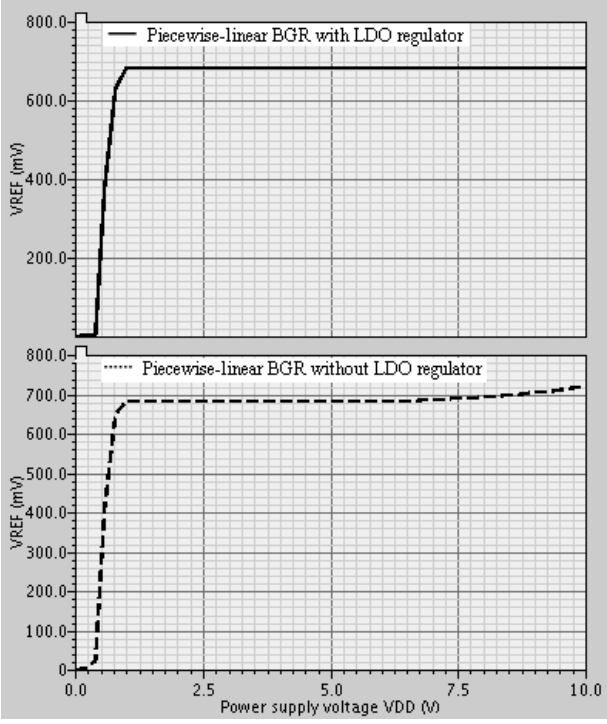


Figure 8. Simulated line-regulation of piecewise-linear BGR with- and without- LDO regulator

LDO regulator has a well line regulation. Finally, to provide an evaluation on the designed high PSRR BGR with LDO regulator in this paper, comparison of some reported BGR is shown in Table I. As shown in Table I, the designed piecewise-linear BGR with LDO regulator has a well performance.

V. CONCLUSIONS

A CMOS high PSRR piecewise-linear BGR, which has an output below 1V, has been designed and analyzed in this paper. Compared to piecewise-linear BGR without LDO regulator, the designed high PSRR piecewise-linear BGR achieves high PSRR performance by adopting LDO regulator. Simulation results shows that the designed high PSRR piecewise-linear BGR with LDO regulator provides an output voltage with excellent stability, a low temperature coefficient, and high PSRR performance. It is well suited for analogue and mixed signal systems.

ACKNOWLEDGMENT

This work was supported in part by a grant from National Science Foundation of China (Grant No. 61102075, and 61301124), Natural Science Foundation Project of CQ CSTC (Grant No. CSTCJJA40011 and CSTC2010BB2412), Science and Technology Research Project of Chongqing Education Commission (Grant No. KJ120503, KJ120507, and KJ120533), Dr. Start Fund of Chongqing University of Posts and Telecommunications (Grant No.A2010-09), 2013 University Innovation Team Construction Plan Funding Project of Chongqing, and Special Project of Internet of Things from Ministry of Industry and Information Technology, Chongqing Development Plan of Innovative Young Talents (Grant No. cstc2013kjrc-qnrc0126).

REFERENCES

[1] G. A. Rincon-Mora, P. E. Allen, "A 1.1-V current-mode and piecewise linear curvature corrected bandgap references," *IEEE J. Solid-State Circuits*. vol.33, p. 1551-1554, October 1998.

[2] P. Y. Chen, S. J. Chang, C. M. Huang, and J. F. Lin, "A 1-V bandgap reference without on-chip resistors," *IEEE Asia Pacific Conference on Circuits and Systems*. pp. 160-163, Kaohsiung, December 2012.

[3] P. B. Basyurt, D. Y. Aksin, "Design of a curvature-corrected bandgap reference with 7.5ppm/C temperature coefficient in 0.35μm CMOS process," *IEEE International Symposium on Circuits and Systems*. pp. 3142-3145, Seoul, May 2012.

[4] Z. K. Zhou, Y. Shi, Z. Huang, P. S. Zhu, Y. Q. Ma, Y. C. Wang, Z. Chen, X. Ming, and B. Zhang, "A 1.6-V 25-mA 5-ppm/°C curvature-compensated bandgap reference," *IEEE Trans. Circuits Syst. I, Reg. Papers*. vol. 59, pp. 677-684, April 2012.

[5] C. M. Andreou, S. Koudounas, and J. Georgiou, "A novel wide-temperature-range, 3.9ppm/ºC CMOS bandgap reference circuit," *IEEE J. Solid-State Circuits*. Vol. 47, pp. 574-581, February 2012.

[6] S. Sano, Y. Takahashi, M. Horiguchi, and M. Ota, "A sub-1V 3.9mW bandgap reference with a 3σ inaccuracy of ±0.34% from -50°C to +150°C using piecewise-linear-current curvature compensation," *International Symposium on VLSI Circuits*. pp. 22-23, Honolulu, June 2012.
[7] S. Mechrmanesh, M. B. Vahidfar, and H. A. Aslanzadeh, M. Atarodi. "A 1-volt, high PSRR, CMOS bandgap voltage reference," *IEEE International symposium on circuits and systems*. vol. 1, pp. I-381-I-384, Bangkok, May 2003.
[8] Y. Hu, and M. Saean, "A 900mV 25μW high PSRR CMOS voltage reference dedicated to implantable micro-devices," *IEEE International symposium on circuits and systems*. vol. 1, pp. I-373-I-376, Bangkok, May 2003.
[9] D. Xiao, W.M. Li, X.F. Zhu, and X.D. Fu, "A curvature-compensated bandgap reference with improved PSRR," *6th International conference on ASIC*. vol.2, pp. 548-551, Shanghai, October 2005.
[10] Z.H. Ning, L.N. He, Y. Wang, and Y.L. Shao, "A novel high PSR voltage reference with secondary temperature compensation," *International conference on electrical and control engineering*. Vol.3, pp. 3200-3203, Wuhan, May 2010.
[11] X.Z. Knang, Z.W. Tang, "A novel high PSRR bandgap over a wide frequency range," *International conference on solid-state and integrated circuit technology*. pp. 418-420, Beijing, October 2010.
[12] J. Yu, Y.F. Zhao, Z.M. Wang, and T.L. Zhang, "A curvature-compensated bandgap reference with high PSR," *International conference granular computing*. pp. 752-755, Hangzhou, August 2008.
[13] H.Y. Zhang, P.K. Chan, and M.T. Tan, "A high PSR voltage reference for dc-to-dc converter applications," *International symposium on circuits and systems*. vol. 1, pp. 816-819, Taipei, May 2009.
[14] A. Dey, and T.K. Bhattacharyya, "A CMOS bandgap reference with high PSRR and improved temperature stability for system-on-chip applications," *International conference of electron devices and solid-state circuits*. pp. 1-2, Tianjin, November 2011.
[15] T.L. Cao, Y. Han, X.P. Liu, H. Luo, and H. Zhang, "A 0.9-V high-PSRR bandgap with self-cascode current mirror," *International conference on circuits and systems*. vol.1, pp. 267-271, Kuala Lumpur, October 2012.
[16] L. Lei, L. Lukas, A. Aytac, S. Sebstian, W. Ralf, and H. Stefan, "A low power bandgap voltage reference circuit with PSRR enhancement," *8th International conference on Ph.D. research in microelectronics and electronics*. pp. 213-216, Aachen, June 2012.
[17] K.P. Francisco, J.A. Hora, "Very low bandgap voltage reference with high PSRR enhancement stage implemented in 90nm CMOS process Technology for LDO application," *International conference on electronics design, systems and applications*. pp. 216-220, Malaysia, November 2012.
[18] B. Razavi, *Design of Analog CMOS Integrated Circuits*, McGraw-Hill, 2001.

Qianneng Zhou received the B.S. degree, the M.Sc degree, and Ph.D degree, all in microelectronic, from Harbin Institute of Technology (HIT), Harbin, China, in 1998, 2002 and 2009, respectively. Since 2009, he has been an associate professor with Chongqing University of Posts and Telecommunications (CQUPT). His current interests include dc-dc converters, amplifiers, LDO regulator and LED drivers.

Qi Li received the B.S. degree in microelectronic from Chongqing University of Posts and Telecommunications (CQUPT), Chongqing, China, 2012. She is currently working toward the M.Sc degree at CQUPT. Her research interests include dc-dc converters, and LDO regulator.

Hongjuan Li received the M.Sc degree in School of Automation Engineering from Northeast Dianli Universisty, Jilin, China, in 2006. She is currently a lecturer with Chongqing University of Posts and Telecommunications, Chongqing (CQUPT), China. Her research interests include Embedded System, and RFID system.

Jinzhao Lin received Ph.D degree in 2001 from Chongqing University. Since 2003 he has been a professor with Chongqing University of Posts and Telecommunications (CQUPT). His current interests include wireless communications and ASIC design.

Yu Pang received Ph.D degree in 2010 from McGill University. Since 2012 he has been an associate professor with Chongqing University of Posts and Telecommunications (CQUPT). His current interests include wireless communications and ASIC design.

Guoquan Li received his Ph.D. degree in College of Communication Engineering from Chongqing University (CQU), Chongqing, China in 2012. He is currently a lecturer with Chongqing University of Posts and Telecommunications, Chongqing (CQUPT), China. His research interests include digital baseband signal processing, multi-user MIMO system, coding theory and technology.

Lu Deng received the B.S. degree in wireless communication from Chongqing University of Posts and Telecommunications (CQUPT) in 2011. She is currently working toward the M.Sc degree at CQUPT. Her research interest is digital signal processing.

Chaotic Cuckoo Search Algorithm for High-dimensional Functions

Aijia Ouyang[a,b]
a. College of Computer, Hunan Science & Technology Economy Trade Vocation College, Hengyang, China
b. College of Information Science and Engineering, Hunan University, Changsha, China
Email: ouyangaijia@163.com

Guo Pan[a,b,*]
a. Logistics Information Dept., Hunan Vocational College of Modern Logistics, Changsha, China
b. College of Information Science and Engineering, Hunan University, Changsha, China
Email: 690849909@qq.com

Guangxue Yue[a,b] and Jiayi Du[b]
a. College of Mathematics, Physics and Information Engineering, Jiaxing University, Jiaxing, China
b. College of Information Science and Engineering, Hunan University, Changsha, China
Email: {guangxueyue, dujiayi}@163.com

***Abstract*—Given that the basic cuckoo search algorithm is vulnerable to local optimum and unsatisfactory calculation precision, chaotic operator (CO), employed as local algorism to optimize elitist individuals in the population, can effectively enhance the properties of cuckoo search (CS) algorithm. Tested by 15 benchmark high-dimensional functions, the experiment result indicates that chaotic cuckoo search (CCS) algorithm can increase the calculation precision and step up the convergent speed with improved robustness, which can be applied to other engineering optimization problems.**

***Index Terms*—chaos, cuckoo search, lévy flight, high-dimensional function, optimization**

I. Introduction

Swarm intelligence algorithm is an efficient method to solve the global optimization problems. It mainly includes particle swarm optimization (PSO) [1], artificial immune (AI) [2], genetic algorithm (GA) [3], differential evolution (DE) [4], invasive weed optimization (IWO)[5] and so on.

After having studied the reproductive behavior of cuckoos and flying properties of lévy, Yang from Cambridge University puts forward cuckoo search algorithm, the properties of which were tested by many functions [6]. It reveals that the algorithm exceeds particle swarm algorithm and genetic algorithm with more powerful global searching ability and solving ability of multi-objective problems, fewer selected parameter as well as better search path [7]. It has been widely used in scientific research and industry. Multi-objective cuckoo search algorithm is applied to solve engineering design optimization in Literature [8] and Jiles-Atherton vector hysteresis parameters estimation problem in Literature [9] respectively. Directed cuckoo search algorithm successfully deals with the sheet nesting problem in Literature [10]. Cuckoo search algorithm handles distributed generation allocation problem so as to reduce power consumption in Literature [11] and image segmentation problem in Literature [12] respectively.

Chaos is an essential characteristic of nonlinear system with a series of particular features such as randomness, ergodicity and regularity, the discovery of which has profound effects on scientific development [13]. As an effective mechanism to avoid being in local optimum, chaos has been introduced to evolutionary computation, providing a new research field and application method for it [14, 15]. Most researches involved, however, only replace the random sequence in mutation operator with the chaos sequence to indicate that chaotic mutation is an effective realization of mutation operator with real number code evolutionary algorithm . Although clear to understand, simple to implement and easy to adapt itself, these algorithms still have some problems as chaos is not made full use of, such as neglecting the regularity of chaos and little use of prior knowledge to improve the local search ability of the algorithm [16].

In this paper, chaos optimization method is applied to make cuckoo search algorithm free from local optimum. It is discovered that rather than in a complicated mess, chaos enjoys fine internal structure. As a singular attractor, chaos attracts systematic movement and confines it to a specific range. Not following a particular orbit, particles of the system wander the chaotic area with freedom, so its future status is unpredictable. Even a slight difference of the original condition will cause immense change. Therefore, two aspects are being studied at present. First, chaos is not expected, and systematic movement is controlled to make chaos

* Corresponding author, Guo Pan, E-mail: panguo@hnu.edu.cn

doi:10.4304/jcp.9.5.1282-1290

unlikely to appear. Second, chaos is regarded as something beneficial, which uses less energy to make greater benefits. Meanwhile, characterized with ergodicity, chaos movement is ergodic with its own regularity in a certain range. As is seen in Literature [17] and Literature [18], using chaos variables to obtain optimization search can undoubtedly get rid of local optimum. A hybrid optimization algorithm is proposed in this paper by combining chaos optimization algorithm with cuckoo search algorithm, which can enhance the optimization efficiency by avoiding the disadvantage of all the statuses are ergodic.

The paper consists of five sections, the second describes the basic cuckoo search algorithm and chaos operator, the third gives the process of the chaotic cuckoo search algorithm, the fourth deals with the experimental results and analysis, and the last is the conclusion.

II. The Basic Algorithm

A. Cuckoo Search Algorithm

1) Reproductive behavior of cuckoos

According to entomologists' observation, some cuckoos breed their offspring by brood parasitism, the way certain birds lay eggs in other birds' nest and have them hatched and raised by other birds. Cuckoos are often alone. They seek hosts with similar hatching and breeding period, feeding habit and egg shape and color, mainly passerine birds. When the host is away from the nest, the cuckoo will quickly lay an egg in it, one every time. Mistaking the egg with its own because of the obscure impression, the host is then in charge of hatching it. The cuckoo always removes one or all the eggs of the host to force it to relay eggs. The newly-born cuckoo tends to push out the host's own baby bird so as to enjoy exclusive breeding.

2) Lévy flight

Viswanathan and some other researchers proved that the albatross adopts the pattern of Lévy flight for foraging. By using satellite positioning system they discovered their flying intervals follow power-law distribution, explained their findings through the space distribution property of invariable food scale on the sea surface and published some symbolic papers [19-21]. After having researched the foraging path of bees [22] and drosophilas [23], Reynolds discovered the appearing occurrence of the straight line portion in their flight path corresponds with scale-free inverse square of lévy distribution in that they both take on lévy flight properties. When the target positions are distributed randomly and sparsely, lévy flight is the ideal search strategy to M independent explorers [24]. In addition, lévy flight is found in creatures such as spider monkeys, gray seals and reindeer as well in human being's behavior [25,26]. Lévy flight is a sort of random walk. The step size meets a heavy-tailed stable distribution. Short-distance exploration and occasional long-distance walk interphase in it. Lévy flight used in intelligent optimization algorithm can enlarge search area, increase population diversity and jump out of local optimal point easily [27].

3) Features of the algorithm

Cuckoo algorithm is evolved form cuckoo's reproductive behavior and lévy flight mechanism. Cuckoos have special reproductive behavior in the nature. They lay eggs in the nest of other birds on which they rely to hatch the eggs. If the host bird spots the secret, severe collision will take place. If the host finds that the egg is not its own, it will remove the egg or rebuild a nest to breed its own offspring. Therefore, the cuckoo inevitably lays eggs the instant the host lays its own in the nest [28]. Once the cuckoo egg is kept in the nest, the host will hatch both the eggs at the same time. In addition, the cuckoo egg always enjoys the priority to be hatched. The newly-born cuckoo baby has an intuition to push out other birds' eggs, and in this way, it enjoys the exclusive breeding [29, 30]. Moreover, the flight of many animals and insects follow the flight properties [23]. Inspired by the two phenomena, Xin-She Yang as well as some other researchers present cuckoo search algorithm and make the following three ideal assumption,

a) Each cuckoo lays only one egg in a certain nest at random.

b) The cuckoo egg in the host nest with high quality will be hatched and then reproduce cuckoos of the next generation.

c) n standing for the number of the host nests used by the cuckoo is determined. The probability that the cuckoo egg is spotted by the host is $pa \in [0,1]$.

Given the three presumptions, the nest-seeking path and position updating formula can be expressed as

$$x_{ij}^{m+1} = x_{ij}^{m} + \alpha \times \text{Levy}(\lambda) \quad (1)$$

where x_{ij}^{m} and x_{ij}^{m+1} stand for the position of dimension j(j=1, 2, …, d) of the nest i (i=1, 2, …, n) in the generation of m and m+1, and lévy(λ) is the skipping path of the flight search at random. The distance and direction of the path is undetermined, and sometimes the skipping path lévy(λ) is considerably long. To make it applied in CS algorithm successfully, Literature [6] defines an adaptive amount α, which is a constant more than zero. The value of it varies, and generally, α=0.01.

Lévy distribution was put forward by French mathematician Lévy in the 1930s. It is believed that the relationship between the continuous skipping path of lévy flight and the time t follows *l*évy distribution. Many other scholars conduct research on it and try to explain randomly phenomenon in the nature with it, such as Brownian Movement and random walk. Yang obtained probability density function in the forms of power through simplified distribution function and Fourier transform,

$$\text{levy} \sim u = t^{-\lambda};\ 1<\lambda<3 \quad (2)$$

Here, λ is the power coefficient. Formula 2 is a probability distribution with heavy tail. Although it describes the randomly wandering process of the cuckoo essentially, it fails to describe the distribution mathematically with simplicity and easy programming to realize the algorithm. Yang, therefore, adopted formula of lévy flight skipping path put forward by Mantegna in 1992 and realized the algorithm [31].

$$s = \frac{\mu}{|\nu|^{1/\beta}} \quad (3)$$

In Formula 3, *s* is the lévy flight skipping path lévy(λ) . The relation between Parameter β and λ in Formula 2 is λ=1+β. The value of β is 0<β<2. Make β=1.5[7] in CS algorithm. Parameter μ and ν are random numbers of normal distribution, following the normal distribution expressed in Formula 4. The value of corresponding Standard deviation of normal distribution σ_μ and σ_υ is reflected in formula 5.

$$\begin{cases} \mu \sim N(0, \sigma_\mu^2) \\ \nu \sim N(0, \sigma_\nu^2) \end{cases} \quad (4)$$

$$\begin{cases} \sigma_\mu = \{\frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]2^{(\beta-1)/2}\beta}\} \\ \sigma_\nu = 1 \end{cases} \quad (5)$$

Make *Step*=α×lévy(λ), and *Step* is the path in which the cuckoo starts form the old nest position x_{ij}^m and searches for new nest position x_{ij}^{m+1} in the solution space by Formula 1. As lévy(λ) depends on the two normal distribution random number μ and ν, which can be various, passive and negative, the path length and direction of the cuckoo following lévy flight search mechanism varies randomly, which is apt to jump form one region to another, making the global optimization ability of CS algorithm extremely powerful. CS algorithm learns from the reproductive behavior of the cuckoo, defines the probability that the cuckoo egg will be spotted by the host as *pa*=0.25[6]. The cuckoo egg with worse adaptive ability will be ruled out, while the one with better adaptation will be hatched, making the newly-born cuckoos consists of excellent individuals. Thus CS algorithm enjoys great contraction ability. In actual optimization problems, the position of the nest x_{ij} stands for effective value space of all the variables, and the fitness of the nest represents the corresponding objective function when the value of the variable varies. Details of CS algorithm go to Literature [6] and [7].

B. Chaos Operator

Chaos is a unique movement pattern of nonlinear system with particular features of sensitivity to the initial value, randomness and ergodicity. If the features are made full use of, the optimization solution of CS algorithm will be promoted.

Chaotic search generates by iteration chaos sequence through certain particular format. Extend the numerical range of the chaos variables to the value range of the optimization variables through the form of carrier wave.

The math procedure of chaotic search is as follows:

If some individual pauses, *d*-dimension is generated and it randomly initializes vectors $y_0=[y_{0,1}, y_{0,2}, \ldots, y_{0,D}]'$, $y_{0,d} \in [0,1]$. There lie slight differences in the values. As the iteration initial value, Vector y_0 starts the chaos sequence iteration according to Logistic equation.

$$y_{n+1,d} = \mu y_{n,d}(1 - y_{n,d}) \quad (6)$$

Thus iteration sequence $y_{n,d}$ is obtained. In the formula, n=0, 1,…, $N_{\max}$; d=0, 1,…, D.

The formula can generate many fields around the local optimal solution by iteration. Through the form of carrier wave, and according to Eq.(7),

$$y'_{n,d} = x_{i,d} + R_{i,d}(2y_{n,d} - 1) \quad (7)$$

Chaos iteration variable $y_{n,d}$ is transformed into optimization variable $y'_{n,d}$,namely, extending the value of chaos variable $y_{n,d}$ to a region where the current position of the individual $x_{i,d}$ is made to be the center and $R_{i,d\ as}$ to be the radius. $R_{i,d}$ is the chaotic search radius and $y'_{n,d}$ is determined by the initialized range of the function variable $x_{i,d.}$ The value range is as follows,

$$y'_{n,d} \in [x_{i,d} - R_{i,d}, x_{i,d} + R_{i,d}] \quad (8)$$

Compute the adaptive value of the function $f(y'_n)$, and update the historical optimal adaptive value f^* in the chaos iteration process and the historical optimal position x_i^* . If f^* is superior to F_i , position x_i^* and velocity v_i^* are used to replace the original position and velocity of the individual. Here,

$$v_i^* = \frac{x_i^* - x_i}{\|x_i^* - x_i\|} \quad (9)$$

III. CHAOTIC CUCKOO SEARCH ALGORITHM

Although chaotic search can avoid being caught in local minimum because of its ergodicity, pure chaotic search can obtain good solution only through huge iteration step numbers and it is sensitive to initial solution in particular. Therefore, a two-stage chaotic CS algorithm is put forward by combining CS algorithm with the above chaotic search, in which CS algorithm is used to lead global search and CO leads local search according to the result of CS algorithm. In order to maintain population diversity and strengthen the dispersion of the search, the algorithm keeps some superior individuals, dynamically contracts search range in view of the best position of the population, and replaces the worse nest position with the one generated in the contract region randomly. The steps of chaotic cuckoo search algorithm can be described as follows,

Step 1. The objective function is $f(X)$, and $X=(x_1, \ldots, x_d)^T$. Initialize the population, generate randomly n initialized positions $X_i (i = 1, 2, \ldots, n)$, and set up parameters of the algorithm.

Step2. Compute the objective function of every bird nest, and record the current optimum solution.

Step 3. Maintain the optimal nest position of previous generation, and update other nest positions according to the position updating formula (1).

Step4. Compare the current bird nest position with that of the previous nest. If better, it is made to be the current optimal position.

Step 5. R stands for the possibility that the nest host will recognize the cuckoo egg. Compare it with probability P_a. If $R>P_a$, change the nest position randomly and obtain a set of new nest positions.

Step 6. Maintain 20% nest positions with the optimal properties in the population.

Step 7. Conduct CO search to the optimal nest position in the population, and update the new nest position.

Step 8. If the stopping criterion of the algorithm is met, output the optimal nest position, otherwise, continue the following steps.

Step 9. Contract search region according to formula (10) and (11)

$$x_{ij}^{\min} = \max\{x_{ij}^{\min}, x_{g,j} - r(x_{ij}^{\max} - x_{ij}^{\min})\} \quad (10)$$

$$x_{ij}^{\max} = \min\{x_{ij}^{\max}, x_{g,j} + r(x_{ij}^{\max} - x_{ij}^{\min})\} \quad (11)$$

Step 10. Generate the rest 80% nest positions of the population randomly in the contracted space, and evaluate it. If the stopping criterion is not met, return to Step 2.

Step 11. Output the global optimal position.

IV. Experiments and Analyses

A. Benchmark Functions

Fifteen widely used test functions are chosen from [32], and the proposed algorithm in this paper, genetic algorithm (GA), differential evolution (DE), particle swarm optimization (PSO) and cuckoo search (CS) are executed for them. These functions are shown in the equations as follows.

Note that functions F1-F9 have many local minima so that they are challenging enough for performance evaluation. For example, F7 has $n!$ local minima and both F8 and F9 have $2n$ local minima, where n=100.

$$F_1 = \sum_{i=1}^{d} -x_i \sin(\sqrt{|x_i|}) \qquad F_2 = \sum_{i=1}^{d} (x_i^2 - 10\cos(2\pi x_i) + 10) \qquad F_3 = -20\exp(-0.2\sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i^2}) - \exp(\frac{1}{d}\sqrt{\sum_{i=1}^{d}\cos(2\pi x_i)}) + 20 + \exp(1)$$

$$F_4 = \frac{1}{4000}\sum_{i=1}^{d} x_i^2 - \prod_{i=1}^{d}\cos(\frac{x_i}{\sqrt{i}}) + 1 \qquad F_5 = \frac{\pi}{d}\left\{10\sin^2(\pi y_1) + \sum_{i=1}^{d-1}(y_i - 1)^2 \bullet [1 + 10\sin^2(\pi y_{i+1})] + (y_d - 1)^2\right\} + \sum_{i=1}^{d} u(x_i, 10, 100, 4)$$

Where, $y_i = 1 + \frac{1}{4}(x_i + 1)$, $u(x_i, a, k, m) = \begin{cases} k(x_i - a)^m, & x_i > a \\ 0, & -a \le x_i \le a \\ k(-x_i - a)^m, & x_i < -a \end{cases}$

$$F_6 = \frac{1}{10}\left\{\sin^2(3\pi x_1) + \sum_{i=1}^{d-1}(x_i - 1)^2 \bullet [1 + \sin^2(3\pi x_{i+1})] + (x_d - 1)^2[1 + \sin^2(2\pi x_d)]\right\} + \sum_{i=1}^{d} u(x_i, 5, 100, 4) \qquad F_7 = -\sum_{i=1}^{d}\sin(x_i)\sin^{20}(\frac{i \times x_i^2}{\pi})$$

$$F_8 = \sum_{i=1}^{d}\left[\sum_{j=1}^{d}(\chi_{ij}\sin\omega_j + \psi_{ij}\cos\omega_j) - \sum_{j=1}^{d}(\chi_{ij}\sin x_j + \psi_{ij}\cos x_j)\right]^2 \qquad F_9 = \frac{1}{d}\sum_{i=1}^{d}(x_i^4 - 16x_i^2 + 5x_i) \qquad F_{10} = \sum_{i=1}^{d-1}\left[100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2\right]$$

$$F_{11} = \sum_{i=1}^{d} x_i^2 \qquad F_{12} = \sum_{i=1}^{d} x_i^4 + \text{rand}[0,1) \qquad F_{13} = \sum_{i=1}^{d}|x_i| + \prod_{i=1}^{d}|x_i| \qquad F_{14} = \sum_{i=1}^{d}\left(\sum_{j=1}^{i} x_j\right)^2 \qquad F_{15} = \max\{|x_i|, i = 1, 2, \cdots, d\}$$

B. Experimental Results and Discussion

TABLE I
COMPARISON OF RESULTS BY FIVE METHODS FOR F1-F5 ON 30 DIMENSION FUNCTION

Function	Algorithm	Best	Mean	Worst	Std.	Convergence
F1	GA	-9.8969E+03	-9.6331E+03	-9.3970E+03	2.0421E+02	195
	DE	-1.2569E+04	-1.2569E+04	-1.2569E+04	**0.0000E+00**	84
	PSO	-3.4249E+03	-3.2998E+03	-3.0110E+03	1.7331E+02	123
	CS	**-9.9789E+03**	**-9.6541E+03**	**-8.8397E+03**	4.5917E+02	141
	CCS	-1.2560E+04	-1.2034E+04	-1.1785E+04	3.2312E+02	**78**
F2	GA	2.0011E+01	2.2439E+01	2.4610E+01	1.8786E+00	203
	DE	9.8814E+00	1.0520E+01	1.2718E+01	1.2151E+00	99
	PSO	4.3767E+01	5.1612E+01	5.3535E+01	4.2250E+00	141
	CS	1.3770E-01	1.1153E+00	2.3152E+00	8.9050E-01	170
	CCS	**2.6545E-05**	**5.1887E-05**	**8.2446E-05**	**2.2871E-05**	**91**
F3	GA	8.7461E+00	1.0836E+01	1.1443E+01	1.1550E+00	232
	DE	1.9000E-06	2.0200E-06	2.0700E-06	7.1300E-08	121
	PSO	4.4440E-01	4.9860E-01	5.2340E-01	3.2988E-02	173
	CS	6.9286E-05	2.3578E-04	4.5698E-03	2.0830E-03	189
	CCS	**2.1316E-14**	**5.8776E-14**	**8.8284E-14**	**2.5200E-14**	**108**
F4	GA	1.7719E+01	2.8074E+01	2.9334E+01	5.2037E+00	217
	DE	3.1600E-12	3.7300E-12	1.4800E-11	5.3600E-12	117
	PSO	1.1796E+01	1.3323E+01	1.4862E+01	1.2518E+00	158
	CS	2.5547E-06	9.8740E-05	1.2354E-05	4.3200E-05	169
	CCS	**1.0842E-15**	**3.5478E-15**	**7.8462E-15**	**2.3212E-15**	**95**
F5	GA	2.1000E+05	2.8900E+05	3.5100E+05	5.7702E+04	194
	DE	**5.5000E-13**	**1.1100E-12**	**3.3900E-12**	1.2300E-12	101
	PSO	1.0470E-01	2.0890E-01	4.1660E-01	1.2965E-01	129
	CS	1.3614E-05	8.2345E-04	2.5401E-03	1.0530E-03	134
	CCS	1.0604E-10	7.2014E-10	5.5561E-09	**2.1047E-09**	**78**

TABLE II
COMPARISON OF RESULTS BY FIVE METHODS FOR F6-F10 ON 30 DIMENSION FUNCTION

Function	Algorithm	Best	Mean	Worst	Std.	Convergence
F6	GA	2.9100E+06	3.8900E+06	8.4900E+06	2.4326E+06	185
	DE	2.3565E-04	4.5620E-04	7.8899E-04	2.2700E-04	103
	PSO	1.5000E-02	2.1500E-02	2.2500E-02	3.3250E-03	114
	CS	8.8321E+04	5.4668E+04	9.1122E+03	3.2458E+04	151
	CCS	**7.3200E-12**	**8.2900E-12**	**9.9100E-12**	**1.0014E-12**	**76**
F7	GA	-2.8577E+01	-2.8286E+01	-2.7861E+01	2.9404E-01	158
	DE	-2.7408E+01	-2.7128E+01	-2.7076E+01	1.4558E-01	81
	PSO	-1.8290E+01	-1.6178E+01	-1.3520E+01	1.9517E+00	95
	CS	-2.1698E+01	-2.1161E+01	-2.0667E+01	4.2108E-01	130
	CCS	**-3.8541E+01**	**-3.8454E+01**	**-3.8322E+01**	**9.0007E-02**	**55**
F8	GA	8.3358E+03	2.8004E+04	6.5488E+04	2.3707E+04	425
	DE	3.1663E+01	5.0399E+01	6.7792E+01	1.4753E+01	**257**
	PSO	6.3445E+03	2.6571E+04	6.2989E+04	2.3438E+04	329
	CS	2.1047E+03	7.6873E+03	9.0806E+03	3.0142E+03	399
	CCS	**3.2170E-05**	**4.9867E-05**	**6.1243E-05**	**1.1008E-05**	277
F9	GA	-7.6387E+01	-7.6334E+01	-7.4475E+01	8.8931E-01	201
	DE	-7.8332E+01	-7.8332E+01	-7.8332E+01	0.0000E+00	108
	PSO	-6.8735E+01	-6.8674E+01	-6.5024E+01	1.7352E+00	105
	CS	-6.3215E+01	-6.1024E+01	-5.8014E+01	2.1321E+00	147
	CCS	**-7.2565E+01**	**-7.1902E+01**	**-7.1285E+01**	**4.5621E-01**	**66**
F10	GA	3.8969E+03	4.1243E+03	5.1371E+03	5.3910E+02	210
	DE	2.6995E+01	3.6461E+01	4.4843E+01	7.2911E+00	105
	PSO	4.1134E+01	6.2174E+01	1.1393E+02	3.0588E+01	95
	CS	1.2134E+02	1.0288E+02	9.8232E+01	9.9774E+00	124
	CCS	**2.3360E-02**	**3.1365E-02**	**3.8378E-02**	**1.9887E-01**	**71**

TABLE III
COMPARISON OF RESULTS BY FIVE METHODS FOR F11-F15 ON 30 DIMENSION FUNCTION

Function	Algorithm	Best	Mean	Worst	Std.	Convergence
F11	GA	1.0979E+03	2.3505E+03	2.9588E+03	7.7475E+02	156
	DE	4.3189E-11	5.4503E-11	6.3732E-11	8.4000E-12	82
	PSO	6.0821E-02	8.5370E-02	1.9144E-01	5.6681E-02	91
	CS	2.3147E-04	2.8510E-03	5.6201E-02	2.5789E-02	123
	CCS	**9.3257E-12**	**6.3479E-12**	**4.4323E-12**	**1.7886E-12**	**54**
F12	GA	2.8304E-01	3.1541E-01	9.4890E-01	3.0655E-01	174
	DE	2.6124E-01	4.7580E-01	6.5037E-01	1.5914E-01	95
	PSO	4.8274E-02	4.3897E-01	5.6509E-01	2.2001E-01	94
	CS	4.2357E-01	6.5478E-01	8.1479E-01	1.6059E-01	130
	CCS	**2.0361E-02**	**5.2014E-02**	**8.5433E-02**	**2.1314E-02**	**70**
F13	GA	9.3551E+00	1.1939E+01	1.2270E+01	1.3032E+00	204
	DE	2.6667E-07	2.9055E-07	3.9517E-07	5.5800E-08	124
	PSO	1.1879E+00	1.3636E+00	1.5306E+00	1.3992E-01	112
	CS	1.0287E-07	6.3144E-08	8.7293E-08	1.6300E-08	150
	CCS	**3.2154E-15**	**6.1981E-15**	**8.7557E-15**	**1.8769E-15**	**96**
F14	GA	3.5832E+04	3.7522E+04	4.2208E+04	2.6971E+03	304
	DE	1.2238E+04	1.6988E+04	1.7988E+04	2.5083E+03	182
	PSO	8.5820E-01	1.3556E+00	1.7450E+00	3.6293E-01	241
	CS	4.6965E+00	5.4284E+00	9.9811E+00	2.3378E+00	289
	CCS	**5.2112E-12**	**8.9876E-12**	**4.5231E-11**	**1.2365E-11**	**111**
F15	GA	2.7267E+01	2.7349E+01	3.3072E+01	2.7174E+00	155
	DE	3.4494E+00	3.7802E+00	3.9989E+00	2.2588E-01	104
	PSO	2.2730E-01	2.5260E-01	2.6940E-01	1.7304E-02	140
	CS	2.1935E+00	3.6503E+00	7.3986E-01	1.1882E+00	125
	CCS	**1.2387E-11**	**4.2551E-11**	**7.5884E-11**	**2.3017E-11**	**84**

To facilitate the experiments, we used the g++ to program some cpp files for implementing the five algorithms on a personal computer with a 32-bit ubuntu 10.04 operating system, a 4GB of RAM, and a 3.10GHz-core(TM) i5-based processor.

The parameter setting of GA, PSO, DE, CS and CCS are as follows. The maximum generations and population size of the five algorithms are the same: the maximum generation G_{max}=1000, population size N_p=100. It can ensure the fairness of competition for each algorithm. For GA, the crossover rate c=0.95, the mutation factor m=0.1; for DE, the crossover rate CR=0.1, the mutation factor F=0.5; for PSO, constriction factor c_1=1.49, c_2=1.49, the maximum flying velocity of particles V_{min}=-0.5, the minimum flying velocity of particles V_{max}=0.5; the maximum inertia weight factor of population ω_{max}=0.9, the minimum inertia weight factor of population ω_{min}=0.4; for CS and CCS, discovery probability of alien egg p_a=0.004, route length β=1.34.

TABLE IV
COMPARISON OF RESULTS BY FIVE METHODS FOR F1-F5 ON 100 DIMENSION FUNCTION

Function	Algorithm	Best	Mean	Worst	Std.	Convergence
F1	GA	-3.1046E+04	-2.9297E+04	-2.8968E+04	9.1198E+02	401
	DE	-2.2140E+04	-2.2072E+04	-2.1287E+04	3.8708E+02	178
	PSO	-6.1125E+03	-5.4758E+03	-4.8518E+03	5.1469E+02	250
	CS	-2.1989E+04	-2.1465E+04	-2.1439E+04	2.5337E+02	291
	CCS	**-2.9596E+04**	**-2.9587E+04**	**-2.9458E+04**	**5.9668E+01**	**147**
F2	GA	1.5654E+02	1.6064E+02	1.6378E+02	2.9641E+00	410
	DE	4.6918E+02	4.8911E+02	5.0298E+02	1.3874E+01	201
	PSO	4.3103E+02	4.4606E+02	4.7216E+02	1.6991E+01	299
	CS	1.7136E+02	1.7231E+02	1.7447E+02	**1.1210E+00**	351
	CCS	**1.2357E+02**	**1.2826E+02**	**1.3420E+02**	4.3519E+00	**197**
F3	GA	1.1801E+01	1.2974E+01	1.3001E+01	5.5943E-01	480
	DE	5.4024E-01	5.7761E-01	5.7917E-01	1.7995E-02	265
	PSO	2.5611E+00	3.1076E+00	3.1999E+00	2.8191E-01	384
	CS	2.8904E+00	3.1235E+00	3.4471E+00	2.2827E-01	427
	CCS	**2.3665E-02**	**3.7889E-02**	**5.2017E-02**	**1.0505E-02**	**201**
F4	GA	7.8163E+01	9.9526E+01	1.5024E+02	3.0226E+01	456
	DE	9.2770E-01	9.3300E-01	9.6640E-01	1.7131E-02	240
	PSO	7.5466E+01	7.8627E+01	8.0636E+01	2.1281E+00	321
	CS	1.4695E-01	2.3126E-01	8.2708E-01	3.0271E-01	364
	CCS	**8.5654E-03**	**5.2551E-03**	**4.6886E-03**	**1.1125E-03**	**214**
F5	GA	6.6758E+05	7.0339E+05	9.0341E+05	1.0377E+05	405
	DE	1.6189E+00	1.9531E+00	1.9859E+00	1.6582E-01	225
	PSO	1.5607E+00	1.9689E+00	2.3562E+00	3.2480E-01	281
	CS	2.8715E-01	7.9651E-01	1.2521E+00	3.9414E-01	257
	CCS	**1.2556E-04**	**4.2017E-04**	**8.0221E-04**	**2.2243E-04**	**186**

TABLE V
COMPARISON OF RESULTS BY FIVE METHODS FOR F6-F10 ON 100 DIMENSION FUNCTION

Function	Algorithm	Best	Mean	Worst	Std.	Convergence
F6	GA	1.2882E+07	1.4777E+07	1.5765E+07	1.1962E+06	421
	DE	7.8299E+00	8.5995E+00	8.9695E+00	4.7468E-01	213
	PSO	6.3090E-01	6.4300E-01	7.2470E-01	4.1660E-02	249
	CS	3.7961E+01	8.9404E+01	1.1219E+02	3.1048E+01	321
	CCS	**2.3224E-02**	**4.2517E-02**	**6.1204E-02**	**1.1228E-02**	**178**
F7	GA	-8.4439E+01	-8.2837E+01	-8.1177E+01	1.3318E+00	388
	DE	-5.0084E+01	-4.9558E+01	-4.9109E+01	**3.3663E-01**	178
	PSO	-2.7925E+01	-2.7906E+01	-2.6478E+01	6.7769E-01	226
	CS	-4.1175E+01	-3.8993E+01	-3.7071E+01	1.6766E+00	279
	CCS	**-9.3654E+01**	**-9.1201E+01**	**-9.0879E+01**	1.2392E+00	**112**
F8	GA	9.8754E+05	1.5426E+06	1.9083E+06	3.7854E+05	801
	DE	3.4724E+05	5.7849E+05	6.2045E+05	1.2013E+05	**405**
	PSO	9.9542E+05	1.7632E+06	2.1474E+06	4.7891E+05	666
	CS	1.9769E+06	1.9970E+06	2.0367E+06	2.4847E+04	804
	CCS	**3.2145E-01**	**5.6887E-01**	**8.3214E-01**	**1.7426E-01**	421
F9	GA	-7.4533E+01	-7.4454E+01	-7.4237E+01	1.2514E-01	423
	DE	-7.0789E+01	-7.0170E+01	-6.9612E+01	4.8072E-01	199
	PSO	-6.7447E+01	-6.6435E+01	-6.3643E+01	1.6087E+00	200
	CS	-6.3032E+01	-6.2786E+01	-6.2715E+01	1.3583E-01	269
	CCS	**-7.7920E+01**	**-7.7915E+01**	**-7.7911E+01**	**2.0108E-03**	**125**
F10	GA	1.2483E+04	1.2555E+04	1.7918E+04	2.5453E+03	431
	DE	7.5285E+02	8.3170E+02	9.7427E+02	9.1634E+01	228
	PSO	8.8102E+02	1.3039E+03	1.3593E+03	2.1361E+02	205
	CS	5.5349E+02	6.2408E+02	9.0421E+02	1.5146E+02	258
	CCS	**4.2365E-01**	**4.5142E-01**	**4.8997E-01**	**2.1106E-02**	**157**

The experiment is divided into two parts, part 1 is for solving the 30 dimension functions. The experimental result is displayed from Table I to Table III. CCS is better than GA, DE, PSO and CS in terms of precision except for F5, for F5, the precision of DE is the highest. the standard deviation of CCS is smaller than the four methods except for F1, for F1, the standard deviation of DE is the smallest. It shows that the robustness of CCS is strong. The convergent speed is higher than the four methods except for F8, for F8, the convergence of DE is the highest. The second part is for solving the 100 dimension functions. The experimental result is displayed from Table IV to Table VI. CCS shows that its super performance in precision, robustness and convergence. CCS is the most precise algorithm among the five methods. CCS is smaller than GA, DE, PSO and CS in terms of standard deviation except for F2 and F7, for F2 and F7, CS and DE is the smallest respectively. The convergence of CCS is faster than the four methods except for F8 and F14, for F8 and F14, DE is the fastest in convergence.

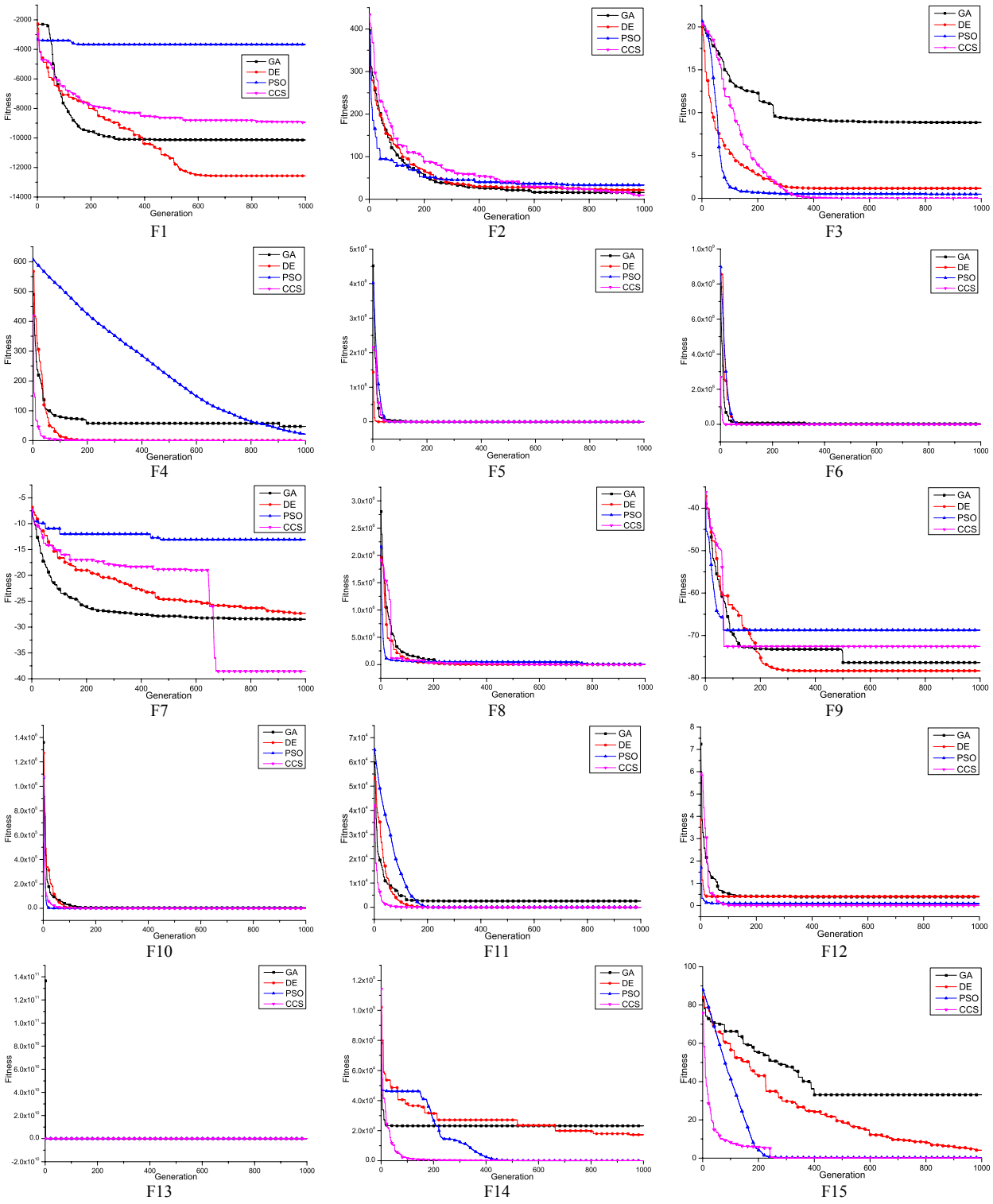


Figure 1. Convergence comparison of four algorihtms for benchmarks

The comparisons of convergent curves of four algorithms (GA, DE, PSO, CCS) for solving the benchmarks of high dimensional functions from F1 to F15 are listed in Figure 1. In subfigure F1, the convergence of DE is the fastest, but the fitness value deviates the theoretical optimum. The convergent curves of CCS for F2, F5, F9, F10 and F15 are rounded to the nearest the theoretical optimum. The convergent speeds of CCS for F3, F4, F6, F7, F11, F12, F13 and F14 are the fastest among the four algorithms. In subfigure F8, the convergent speed of DE is the fastest, the convergent speed of CCS is only faster than GA, the experimental results in the paper show that CCS has no advantage in convergence for F8. In subfigure F5, the fitness value of DE is the smallest. In Table I and Table V, the standard deviation of DE is the smallest and the robustness of DE is stronger than the other four methods for F1 and F7 respectively.

TABLE VI
COMPARISON OF RESULTS BY FIVE METHODS FOR F11-F15 ON 100 DIMENSION FUNCTION

Function	Algorithm	Best	Mean	Worst	Std.	Convergence
F11	GA	1.1314E+04	1.2385E+04	1.8144E+04	2.9993E+03	294
	DE	3.4943E+00	4.0955E+00	4.7194E+00	5.0017E-01	178
	PSO	3.5648E+00	3.6542E+00	3.6737E+00	4.7413E-02	201
	CS	3.2145E+03	5.6541E+03	9.8987E+03	2.7618E+03	261
	CCS	**1.8218E-01**	**3.0205E-01**	**3.6683E-01**	**3.2598E-02**	**101**
F12	GA	6.3992E-01	7.0922E-01	7.3981E-01	4.1788E-02	365
	DE	2.0997E-01	5.2664E-01	9.0501E-01	2.8412E-01	204
	PSO	6.9893E-01	9.9771E-01	1.5638E+00	3.5866E-01	210
	CS	6.1301E-01	7.4751E-01	5.1749E-01	9.4354E-02	248
	CCS	**1.2587E-01**	**1.3014E-01**	**1.4201E-01**	**4.5689E-03**	**142**
F13	GA	4.7663E+00	5.6044E+00	6.1119E+00	5.5484E-01	427
	DE	7.7861E-01	8.6395E-01	8.7092E-01	4.1969E-02	199
	PSO	1.3387E+01	1.5045E+01	1.7128E+01	1.5305E+00	247
	CS	2.5647E+01	3.8775E+01	5.6817E+01	1.2778E+01	270
	CCS	**4.5139E-02**	**6.1567E-02**	**8.9302E-02**	**1.4884E-02**	**173**
F14	GA	2.9544E+05	3.0252E+05	3.2549E+05	1.2827E+04	688
	DE	2.8302E+05	2.8791E+05	3.1719E+05	1.5088E+04	**294**
	PSO	8.6048E+02	8.7005E+02	1.4564E+03	2.7869E+02	541
	CS	3.7669E+03	3.8340E+03	4.4435E+03	3.0437E+02	600
	CCS	**4.2635E-06**	**4.3650E-06**	**4.4198E-06**	**3.0208E-08**	301
F15	GA	5.7815E+01	6.2661E+01	6.5075E+01	3.0188E+00	327
	DE	6.5758E+01	6.7462E+01	6.8746E+01	1.2239E+00	187
	PSO	3.2187E+00	3.5279E+00	3.8999E+00	2.7849E-01	301
	CS	1.2293E+01	1.3073E+01	1.4176E+01	7.7249E-01	214
	CCS	**2.3667E-04**	**2.7881E-04**	**2.8710E-04**	**9.2080E-06**	**184**

V. CONCLUSIONS

Cuckoo search algorithm is a novel intelligence algorithm, it has been successfully applied to scientific research and industrial technology. A chaotic cuckoo search algorithm is presented in the paper, which is employed to solve the high-dimensional functions. 15 Benchmarks of high-dimensional functions are used to verify the performance of CCS. The simulation results show that CCS have competitive advantages over GA, DE, PSO and CS in terms of computation precision, robustness, convergent speed.

Many real-world optimization problems involve a high dimension function of decision variables. For example, in shape optimization, a large number of shape design variables are often used to represent complex shapes, such as turbine blades, aircraft wings, and heat exchangers, etc. CCS can be applied to solve this kind of problems.

ACKNOWLEDGMENT

This work is partially supported by the National Natural Science Foundation of China (No.61202109), A Project Supported by Scientific Research Fund of Hunan Provincial Education Department (Grant No. 08D092, No.13C333), A Project Supported by Scientific Research Fund of Hunan Provincial Statistical Bureau (2008C36), A Project Supported by the Science and Technology Research Foundation of Hunan Province (No.2008GK2019 and No. 2013GK3082). A Project supported by the Natural Science Foundation of Zhejiang Province, China (No.LY12F02019), A Project supported by the public technology applied research project of Zhejiang province (No.2011C23130) Project supported by the Science and Technology Foundation of Jiaxing City, China (No.2012AY1027).

REFERENCES

[1] X. Qiu, X. Qiu and F. Liao, "A Fractal Evolutionary Particle Swarm Optimizer," *Journal of Computers*, vol.8, no.5, pp.1303-1308, 2013.

[2] L. Ye, Y. Xing, and W. Xiang, "An Artificial Immune Classification Algorithm based on Particle Swarm Optimization," *Journal of Computers*, vol.8, no.3, pp.772-778, 2013.

[3] T. Chen and G. Zhou, "Modeling Production scheduling problem and its solution by genetic algorithm," *Journal of Computers*, vol.8, no.8, pp.2126-2133, 2013.

[4] F. Gu, and H. Liu, "An Adaptive Multiobjective Differential Evolution Algorithm," *Journal of Computers*, vol.8, no.2, pp.294-301, 2013.

[5] L. Xie, A. Ouyang, L. Liu, M. He, X. Peng and X. Zhou, "Chaotic Hybrid Invasive Weed Optimization for Machinery Optimizing," *Journal of Computers*, vol.8, no.8, pp.2093-2100, 2013.

[6] X.-S. Yang and S. Deb, "Cuckoo Search via Lévy flights," *2009 World Congress on Nature & Biologically Inspired Computing*, pp.210-214, 2009.

[7] X.-S. Yang and S. Deb, "Engineering optimisation by cuckoo search," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol.1, no.4, pp.330-343, 2010.

[8] X.-S. Yang and S. Deb, "Multiobjective cuckoo search for design optimization," *Computers & Operations Research*, vol.40, no.6, pp.1616-1624, 2013.

[9] L.S. Coelho, F.A. Guerra, N.J. Batistela and J.V. Leite, "Multiobjective cuckoo search algorithm based on Duffing's oscillator applied to Jiles-Atherton vector hysteresis parameters estimation," *IEEE Transactions on Magnetics*, vol.49, no.5, pp.1745-1748, 2013.

[10] A. Elkeran, "A new approach for sheet nesting problem using guided cuckoo search and pairwise clustering," *European Journal of Operational Research*, vol.231, no.3, pp.757-769, 2013.

[11] Z. Moravej and A. Akhlaghi, "A novel approach based on cuckoo search for DG allocation in distribution network," *International Journal of Electrical Power & Energy Systems*, vol.44, no.1, pp.672-679, 2013.

[12] S. Agrawal, R. Panda, S. Bhuyan and B.K. Panigrahi, "Tsallis entropy based optimal multilevel thresholding using cuckoo search algorithm," *Swarm and Evolutionary Computation*, vol.11, pp.16-30, 2013.

[13] R. Vali, S.M. Berber, S.-K. Nguang, "Analysis of Chaos-Based Code Tracking Using Chaotic Correlation Statistics," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol.59, no.4, pp.796-805, 2012.

[14] G.-C. Liao and T.-P. Tsao, "Application of a fuzzy neural network combined with a chaos genetic algorithm and simulated annealing to short-term load forecasting," *IEEE Transactions on Evolutionary Computation*, vol.10, no.3, pp.330-340, 2006.

[15] M. Shen, W.-N. Chen, J. Zhang, H.S.-H. Chung and O.Kaynak, "Optimal selection of parameters for nonuniform embedding of chaotic time series using ant colony optimization," *IEEE Transactions on Cybernetics*, vol.43, no.2, pp.790-802, 2013.

[16] R. Caponetto, L. Fortuna, S. Fazzino, M.G. Xibilia, "Chaotic sequences to improve the performance of evolutionary algorithms," *IEEE Transactions on Evolutionary Computation*, vol.7, no.3, pp.289-304, 2003.

[17] L.D.S. Coelho, T.C. Bora, L. Lebensztajn, "A Chaotic Approach of Differential Evolution Optimization Applied to Loudspeaker Design Problem," *IEEE Transactions on Magnetics*, vol.48, no.2, pp.751-754, 2012.

[18] L.D.S. Coelho, V.C. Mariani, "Combining of chaotic differential evolution and quadratic programming for economic dispatch optimization with valve-point effect," *IEEE Transactions on power systems*, vol.21, no.2, pp.989-996, 2006.

[19] G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince and H. E. Stanley, "lévy flights search patterns of wandering albatrosses," *Nature*, vol. 381, pp. 413-415, 1996.

[20] G.M. Viswanathan, V. Afanasyev, S.V. Buldyrev, S. Havlin, M.G.E da Luz, E.P Raposo, and H.E. Stanley, "Lévy flights in random searches," *Physica A: Statistical Mechanics and its Applications*, Vol.282, no.1–2, pp.1-12, 2000.

[21] G.M. Viswanathan, V. Afanasyev, S. V Buldyrev, S. Havlin, M.G.E da Luz, E.P Raposo, and H.E. Stanley, "Lévy flights search patterns of biological organisms," *Physica A: Statistical Mechanics and its Applications*, Vol.295, no.1-2, pp.85-88, 2001.

[22] A.M. Reynolds, "Cooperative random Lévy flight searches and the flight patterns of honeybees," *Physics Letters A*, Vol.354, no.5-6, pp.384-388, 2006.

[23] A.M. Reynolds and M.A. Frye, "Free-flight odor tracking indrosophila is consistent with an optimal intermittent scale-free search," *PLoS ONE*, vol.2, no.4, pp.1-9, 2007.

[24] A.M. Raynolds, A.D. Smith, R. Menzel, U. Greggers, D.R. Reynolds and J.R. Riley, "Displaced honey bees performoptimal scal-free search flights," *Ecology*, vol.88, no.8, pp.1955-1961, 2007.

[25] A.L. Sellreier and M. Grove, "Ranging patterns of hamadryas baboons: Randomwalk analyses," *Animal Behavior*,vol.80, no.l, pp.75-87, 2010.

[26] G. Ramos-Fernandez, J.L. Mateos, O. Miramontes, G. Cocho, H. Larralde and B. Ayala-Orozco, "Lévy walk patterns in the foraging movements of spider monkeys," *Behavioral Ecology and Sociobioblogy*, vol.55, no.3, pp.1743-1750, 2004.

[27] M.F. Shlesinger, "Search research," *Nature*, vol.443, pp.281-282, 2006.

[28] R.B. Payne, M.D. Sorenson and K. Klitz, "The cuckoo," Oxford: Oxford University Press, 2005.

[29] C.T. Brown, L.S. Liebovitch and R. Glendon, "Lévy Flights in Dobe Ju/'hoansi Foraging Patterns," *Human Ecology*, vol.35, no.1, pp.129-138, 2007.

[30] I. Pavlyukevich, "Lévy flights, non-local search and simulated annealing," *Journal of Computational Physics*, Vol.226, no.2, pp.1830-1844, 2007.

[31] R.N. Mantegna, "Fast, accurate algorithm for numerical simulation of Lévy stable stochastic processes," *Physical Review E*, vol.49, no.5, pp. 4677-4683, 1994.

[32] Y. Wang and C. Dang, "An Evolutionary Algorithm for Global Optimization Based on Level-Set Evolution and Latin Squares," *IEEE Transactions on Evolutionary Computation*, vol.11, no.5, pp.579-595, 2007.

Aijia Ouyang received the M.E. degree in computer science and technology from Guangxi University for Nationalities, Nanning, China, in 2010. He is currently a Ph.D. candidate of Hunan University, Changsha, China. His research interests include parallel computing and artificial intelligence. He has published research articles in international conference and journals of parallel computing.

Guo Pan is an associate professor in the Logistics Information Department, Huna n Vocational college of Modern Logistic s, Changsha, Hunan Province, China. Sh e received the B.E. degree in the Depart ment of Computer from National Univer sity of Defense Technology, Changsha, China, in 2001, and the M.E. degree in th e Department of Computer Science and Technology from Hunan University, in 2 007, respectively. She is currently a Ph.D. candidate in the Colle ge in the College of Information Science and Engineering at Hu nan University, Changsha, China. Her research interests include parallel computing, big data, DNA computing and intelligence algorithms.

Guangxue Yue is a Professor in the Department of Computer Science and Technology at Jiaxing University, Jiaxing, China. He received the M.E. degree, the Ph.D. degree in the Department of Computer Science and Technology from Hunan University, Changsha, China, in 2006 and 2012, respectively. His research interests include computer network, distributed system, QoS of streaming media. He has published more than 70 international journals/conference papers on computer network, distributed system and information security.

Jiayi Du received his MSc and BSc in computer science from Hunan University, China, in 2004 and 2010. He is currently a PhD candidate in Hunan University, China. His research interest includes modeling and scheduling for parallel and distributed computing systems, embedded system computing, cloud computing, parallel system reliability, and parallel algorithms.

Call for Papers and Special Issues

Aims and Scope.

Journal of Computers (JCP, ISSN 1796-203X) is a scholarly peer-reviewed international scientific journal published monthly for researchers, developers, technical managers, and educators in the computer field. It provide a high profile, leading edge forum for academic researchers, industrial professionals, engineers, consultants, managers, educators and policy makers working in the field to contribute and disseminate innovative new work on all the areas of computers.

JCP invites original, previously unpublished, research, survey and tutorial papers, plus case studies and short research notes, on both applied and theoretical aspects of computers. These areas include, but are not limited to, the following:

- Computer Organizations and Architectures
- Operating Systems, Software Systems, and Communication Protocols
- Real-time Systems, Embedded Systems, and Distributed Systems
- Digital Devices, Computer Components, and Interconnection Networks
- Specification, Design, Prototyping, and Testing Methods and Tools
- Artificial Intelligence, Algorithms, Computational Science
- Performance, Fault Tolerance, Reliability, Security, and Testability
- Case Studies and Experimental and Theoretical Evaluations
- New and Important Applications and Trends

Special Issue Guidelines

Special issues feature specifically aimed and targeted topics of interest contributed by authors responding to a particular Call for Papers or by invitation, edited by guest editor(s). We encourage you to submit proposals for creating special issues in areas that are of interest to the Journal. Preference will be given to proposals that cover some unique aspect of the technology and ones that include subjects that are timely and useful to the readers of the Journal. A Special Issue is typically made of 10 to 15 papers, with each paper 8 to 12 pages of length.

The following information should be included as part of the proposal:

- Proposed title for the Special Issue
- Description of the topic area to be focused upon and justification
- Review process for the selection and rejection of papers.
- Name, contact, position, affiliation, and biography of the Guest Editor(s)
- List of potential reviewers
- Potential authors to the issue
- Tentative time-table for the call for papers and reviews

If a proposal is accepted, the guest editor will be responsible for:

- Preparing the “Call for Papers” to be included on the Journal’s Web site.
- Distribution of the Call for Papers broadly to various mailing lists and sites.
- Getting submissions, arranging review process, making decisions, and carrying out all correspondence with the authors. Authors should be informed the Instructions for Authors.
- Providing us the completed and approved final versions of the papers formatted in the Journal’s style, together with all authors’ contact information.
- Writing a one- or two-page introductory editorial to be published in the Special Issue.

Special Issue for a Conference/Workshop

A special issue for a Conference/Workshop is usually released in association with the committee members of the Conference/Workshop like general chairs and/or program chairs who are appointed as the Guest Editors of the Special Issue. Special Issue for a Conference/Workshop is typically made of 10 to 15 papers, with each paper 8 to 12 pages of length.

Guest Editors are involved in the following steps in guest-editing a Special Issue based on a Conference/Workshop:

- Selecting a Title for the Special Issue, e.g. “Special Issue: Selected Best Papers of XYZ Conference”.
- Sending us a formal “Letter of Intent” for the Special Issue.
- Creating a “Call for Papers” for the Special Issue, posting it on the conference web site, and publicizing it to the conference attendees. Information about the Journal and Academy Publisher can be included in the Call for Papers.
- Establishing criteria for paper selection/rejections. The papers can be nominated based on multiple criteria, e.g. rank in review process plus the evaluation from the Session Chairs and the feedback from the Conference attendees.
- Selecting and inviting submissions, arranging review process, making decisions, and carrying out all correspondence with the authors. Authors should be informed the Author Instructions. Usually, the Proceedings manuscripts should be expanded and enhanced.
- Providing us the completed and approved final versions of the papers formatted in the Journal’s style, together with all authors’ contact information.
- Writing a one- or two-page introductory editorial to be published in the Special Issue.

More information is available on the web site at http://www.academypublisher.com/jcp/.

(Contents Continued from Back Cover)