Journal of Communications

ISSN 1796-2021

Volume 5, Number 7, July 2010

Special Issue: Advances in Wireless Communication - I

Guest Editors: Khan Mohammad Iftekharuddin, Mohammed Nazrul Islam, Mohammad Ataul Karim, and Mohammad Abdus Salam

Contents

Special Issue on Advances in Wireless Communication – I

Guest Editorial

Khan Mohammad Iftekharuddin
Department of Electrical & Computer Engineering, University of Memphis
206 Engineering Science Building, Memphis, Tennessee 38152, USA

Mohammed Nazrul Islam
Department of Security Systems, State University of New York, Farmingdale
2350 Broadhollow Road Farmingdale, New York 11735, USA

Mohammad Ataul Karim
Office of Research, Old Dominion University
4111 Monarch Way, Suite 203 Norfolk, Virginia 23508, USA

Mohammad Abdus Salam
Department of Computer Science, Southern University
Baton Rouge, Louisiana 70813, USA

Wireless communication systems and networks are becoming extremely popular because of cost effectiveness, mobility and flexibility and as such the demand for high-speed wireless communication services is growing rather rapidly. However, wireless systems do face a number of challenges including spectral efficiency, performance reliability, power requirement, QoS, security, and complexity and flexibility of the architecture. Both academia and industries are focusing on research in wireless systems and networks significantly, which motivated the publication of the current Special Issue in wireless communications.

While this Special Issue invited authors from around the world, we had also targeted pertinent papers that were presented at the IEEE International Conference on Computer and Information Technology (ICCIT 2009) held on December 2123, 2009, Dhaka, Bangladesh. We had received many submissions which were then peer-reviewed, and 15 papers were finally selected for publication to appear in issues of the journal. These authors represent academic and/or research institutions from Bangladesh, Canada, China, India, Iran, Japan, Jordan, Saudi Arabia, Thailand, and United Kingdom.

In the first paper of this Special Issue, Z. Li, J. Yang, and J. Yao propose threshold-aided opportunistic transmission strategies for use in multiuser multiple-input multiple-output downlink having limited feedback. They employ the primary threshold to reduce the feedback and the secondary threshold to guarantee the outage performance. The proposed strategies have been investigated through statistical analyses and Monte-Carlo simulation. The second paper by A. Mehbodniya, S. Aissa and J. Chitzadeh presents a vertical handoff algorithm in multitier (overlay) networks. The scheme employs pattern recognition technique using probabilistic neural network to determine the user location and decides on the handoff based on that information.

P. Wang, L. Zu, F. Liu and Y. Wang, in the third paper, develops a downlink resource allocation algorithm to maximize the system throughput employing the frequency selective fading phenomenon in orthogonal frequency division multiplexed network. Analytical and simulation results show that the algorithm enhances the system performance by assuring quality of service and guaranteeing the minimum reserved traffic rate for non-real-time services. In the fourth paper, S. Ahmed and M. Kawai present a code allocation table consisting of selected orthogonal binary user codes to reduce the peak to average power ratio in multicarrier code division multiple access systems. The authors show through analysis and simulation that the proposed scheme can perform better than Walsh-Hadamard codes in terms of both peak-to-average power ratio and bit error rate.

A scalable group key agreement is proposed next by Z. Li-Ping and W. Yi using the layer-cluster group model for a mobile ad hoc network. A multi-linear map is incorporated in the layer-cluster structure to meet the security demands of large mobile ad hoc networks as well as improve the system performance. This is followed by K. A. Darabkh who proposes a queuing model to investigate the impact of Fano decoding algorithm on the performance of a wireless network. The paper presents both analytical and simulation studies on the average number of packets residing in the system's buffer.

S. A. K. Tanoli, I. Khan and N. Rajatheva investigate next the performance of a cooperative network based on bit-interleaved coded modulation-iterative decoding over different fading channels, including Rayleigh, Nakagami-*m* and Rician. The bit error rate performance of the fading channels are evaluated and compared through simulation studies. The last paper of this Special Issue is authored by F. Haroon and K. M. Ahmed who propose the deployment of

doi:10.4304/jcm.5.7.509-510

IEEE 802.15.4a protocol in medium access control of an industrial field level communication network. The authors also introduce a reduced complexity adaptive SRake receiver to recover the weak impulse radio-time hopping ultra wideband signals in dense multipath propagation with strong noise.

The guest editors would like to express their sincere gratitude to the reviewers, who have finished their reviews in the shortest possible time and dedicated their valuable time to ensure the quality of this special issue. Finally, the guest editors would extend their sincere appreciation to the Associate Editor-in-Chief, Dr. Haohong Wang for providing them with this opportunity and facilitating preparation of an excellent journal special issue.

Editor Biographies

Khan Mohammad Iftekharuddin is an Associate Professor in the department of Electrical and Computer Engineering at the University of Memphis (U of M). He is also an associated faculty in the Institute for Intelligent Systems at U of M. Further, he holds a joint appointment with the joint graduate program in biomedical engineering at the U of M and University of Tennessee at Memphis. Prior to joining U of M, he was on the faculty of the departments of Computer Science and Electrical & Computer Engineering at North Dakota State University. His research interests include biomedical image analysis, sensor signal acquisition and modeling, digital, optical and multimedia signal and image processing, optical computing and interconnection, applications of artificial-neural inference techniques, automatic target recognition (ATR) and biologically inspired ATR. Dr. Iftekharuddin is the author of five book chapters and more than hundred refereed journal papers and conference proceedings. He is an associate editor for *Optical Engineering*, *International Journal of Imaging*, *Open Cybernetics and Systemic Journal* and *International Journal of Tomography and Statistics*, He has served as guest editor for five journal special issues. He is an elected fellow of SPIE, a senior member of IEEE, a member of IEEE CS and OSA. He obtained his B.Sc. degree from Bangladesh Institute of Technology in 1989. He received an M.S. and a Ph.D. both in electrical engineering from the University of Dayton in 1991 and 1995, respectively.

Mohammed Nazrul Islam is an Assistant Professor in the Department of Security Systems at the State University of New York at Farmingdale. He received his BS and MS in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology in 1991 and 1994, respectively, and his PhD from Muroran Institute of Technology, Japan in 1999. Prior to joining Farmingdale, he worked as a Research Scientist and Adjunct Assistant Professor at Old Dominion University. He also served respectively as an Associate Professor at Bangladesh University of Engineering and Technology, as a Postdoctoral Research Fellow at the University of South Alabama and as a Visiting Assistant Professor at the University of West Florida. He authored and coauthored more than 110 publications in refereed journals and conference proceedings. His research interests include optical communication, wireless communication, digital image processing and solid state devices. He is a Senior Member of IEEE and a Member of SPIE.

Mohammad Ataul Karim is Vice President for Research of Old Dominion University in Norfolk, Virginia. Previously, he served as dean of engineering at the City University of New York. His research areas include information processing, pattern recognition, computing, displays, and electro-optical devices and systems. Dr. Karim is author of 16 books, 7 book chapters, and over 350 articles. He is North American Editor of *Optics & Laser Technology* and an Associate Editor of the *IEEE Transactions on Education.* He has served as guest editor for over twenty journal special issues. Professor Karim is an elected fellow of the Institution of Electrical and Electronics Engineers (IEEE), Optical Society of America (OSA), Society of Photo-Instrumentation Engineers (SPIE), the Institute of Physics (InstP), the Institution of Engineering & Technology (IET), and Bangladesh Academy of Sciences. He received his BS in physics in 1976 from the University of Dacca, Bangladesh, and MS degrees in both physics and electrical engineering, and a Ph.D. in electrical engineering from the University of Alabama respectively in 1978, 1979, and 1981.

Mohammad Abdus Salam is an Associate Professor in the Department of Computer Science at Southern University, Baton Rouge, Louisiana. He received his BS degree in Electrical and Electronics Engineering from Bangladesh Institute of Technology, Rajshahi in 1991 and MS and Ph.D. degrees from Fukui University, Japan, respectively in 1998 and 2001. Prior to 2005, he worked as an adjunct faculty member of Mathematics and Computer Science at the City University of New York at York College, and as a postdoctoral fellow in the Department of Electrical and Computer Engineering at the University of South Alabama, Mobile, Alabama. He is a member of IEEE, IEICE (Japan), and AIAA. His research interests include wireless communication, error-control coding, and sensor networks.

Adaptive Opportunistic Transmission in MU-MIMO Downlink with Reduced Feedback

Zhao Li, Jiawei Yang, Junliang Yao
State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China
Email: {zli, jwyang}@xidian.edu.cn, jlyao@pcn.xidian.edu.cn

Abstract— **In this paper we proposed two threshold-aided adaptive opportunistic transmission strategies in multiuser multiple-input multiple-output (MU-MIMO) downlink with limited feedback. Threshold scheme in company with transmission mode adaptation are employed to handle feedback cost and improve system performance. In one time slot, each mobile station (MS) carries out channel estimation, selects appropriate transmission mode and feeds back adaptively. With feedback information the base station (BS) schedules one user out of candidate MSs and transmits to it. One of the proposed strategies employs single-level threshold. However, proper design of the threshold is difficult. Since a tight one would result in high outage probability whereas a loose one could not reduce feedback load effectively. Thus the other method employs double-level threshold. In this scheme, the primary threshold is used for feedback reduction, the secondary one is employed to guarantee the outage performance. In evaluating the strategies, statistical analysis and Monte-Carlo simulation are used. Results show that with properly designed thresholds, the proposed schemes can greatly reduce the feedback load, achieve high throughput and good outage performance.**

Index Terms— **Multiuser, MIMO, Downlink, Adaptive, Opportunistic, Threshold**

I. INTRODUCTION

MU-MIMO is a set of advanced MIMO technologies that exploit the availability of multiple independent mobile users in order to improve communication performance. When multiple users are simultaneously supported by BS two major issues are under investigation, multiuser interference (MUI) elimination (precoding method design) [1-3] and multiuser scheduling [4-7]. These two issues are always discussed jointly and need channel state information (CSI) feedback. In existing MU-MIMO works, MUI elimination always induces much feedback load and computational complexity.

In a multiuser environment with multiple independent wireless links, at any given point in time the probability that all of those links have low quality is small. This advantage is called multiuser diversity (MUD). Accordingly, some works focus on opportunistic beamforming (OB) [8-13], i.e. in each time slot only one user among multiple candidates is selected to exploit MUD. Moreover, when OB is under discussion MUI is no longer considered.

Obviously, all of the above techniques require BS to know the channel information of users, which is usually estimated at the MS and fed to BS through a finite-rate reverse link. Since feedback information is part of system overhead, a lot of works focus on feedback load reduction. Most of them employ threshold to eliminate MSs whose channel quality is below the given threshold [6-9, 12-13], i.e. users who are not eligible for transmission would keep silent. However, these works are usually discussed in a simplified system model, i.e. the receiver end is equipped with single antenna [8-9, 12] or only part of multiple antennas is utilized [6]. As a result the advantages of MIMO are not fully exploited. Moreover, those works always consider fixed transmission mode. If adaptive transmission [14] is employed, not only the communication performance can be improved, the system feedback cost could also dynamically vary along with channel status. In [13] both threshold scheme and transmission mode adaptation are adopted. However, threshold design and mode selection criterion are not elaborated. Moreover, better performance may be achieved by modifying the number and redesigning the value of thresholds.

In this paper we proposed two adaptive opportunistic transmission strategies in MU-MIMO downlink. General MIMO scenario is considered in which both BS and MS are equipped with multiple antennas. Threshold scheme in company with transmission mode adaptation are employed to handle feedback cost and improve system performance. In one time slot, each MS carries out transmission mode selection and adaptively feeds CSI. With feedback information, BS schedules one user out of multiple candidates and implements transmission.

Among the proposed strategies, the first one employs single-level threshold, and the criterion of transmission mode adaptation is deduced from this threshold. However, proper design of the threshold is difficult, since a tight one would result in high outage probability whereas a loose one could not reduce feedback load effectively. Thus a double-level threshold scheme is introduced subsequently. In this strategy two thresholds are employed. The primary one is used for feedback reduction, the secondary one is to guarantee low outage probability.

The rest of this paper is organized as follows. In Section II we describe the system model. In Section III the single-level threshold strategy is introduced. Section IV presents the double-level threshold scheme. In the above two sections we also give the design of threshold values and analyzed the system performance theoretically. Section V provides the simulation results. Finally, the

Manuscript received January 20, 2010; revised April 6, 2010; accepted May 12, 2010.

doi:10.4304/jcm.5.7.511-520

conclusion and future work are discussed in Section VI.

II. System Model

Consider a single cell MU-MIMO system with one BS and K MSs. BS has N_T transmit antennas and user-k is equipped with N_k receive antennas. Assume each MS undergoes statistically the same independent frequency non-selective fading. The channel matrix of user-k in time slot n is denoted by an $N_k \times N_T$ matrix $\mathbf{H}_k(n)$. Consider the Rayleigh block fading channel model. Make reasonable assumption that $N_T > N_k$, as MS always has rigorous size constraint compared with BS.

In a transmission cycle, BS should send training sequence to MSs, finish CSI collection and user selection before data transmission. These processes take a finite period of time called overhead slot. In downlink discussion we consider training sending as the primary overhead. Define T and τ as the length of the total downlink time slot and overhead slot respectively. Similar to [15-16], typical values of τ/T lie in the region of $5\% - 10\%$. For each user a low-rate error-free channel exists that conveys CSI back to the BS.

As we know, MIMO channel can be equivalent to a set of decoupled parallel subchannels by singular value decomposition (SVD). In slot n, assume that BS transmits to user-k and the number of substreams for user-k is $T_k(n)$, called mode, which satisfies $0 \le T_k(n) \le N_k$. When $T_k(n) = 1$, BS transmits to MS-k using beamforming (BF). When $T_k(n) > 1$, BS selects spatial multiplexing (SM). We employ a length-$T_k(n)$ vector $\mathbf{s}_k(n)$ to represent the data streams intended to user-k, satisfying $\varepsilon\{|s_{k,i}(n)|^2\} = 1$.

Apply SVD to $\mathbf{H}_k(n)$, we have $\mathbf{H}_k(n) = \mathbf{U}_k(n)\mathbf{\Lambda}_k(n)\mathbf{V}_k^H(n)$. $\mathbf{V}_k(n)$ is channel directional information (CDI) and $\mathbf{\Lambda}_k(n)$ is channel quality information (CQI) [17]. In this paper, we employ $\mathbf{V}_k(n)$ as the precoding matrix for user-k. $\mathbf{s}_k(n)$ is first transformed into a length-N_T symbol vector by multiplying $\mathbf{V}_k(n)$. Then the symbol vector is launched from transmit antenna array. The power allocation matrix at BS is denoted as a $T_k(n) \times T_k(n)$ diagonal matrix $\mathbf{Q}_k(n) = \text{diag}\left(q_{k,1}(n), \cdots, q_{k,T_k n}(n)\right)$. The power assigned to user-k satisfies total transmit power constraint $P_T = \text{tr}\left(\mathbf{Q}_k(n)\right) = \sum_{i=1}^{T_k(n)} q_{k,i}(n)$.

At the user side, the received signal vector for user-k is

$$\mathbf{r}_k(n) = \mathbf{H}_k(n)\mathbf{V}_k(n)\mathbf{Q}_k^{1/2}(n)\mathbf{s}_k(n) + \mathbf{n}_k(n) \qquad (1)$$

where $\mathbf{Q}_k^{1/2}(n)$ stands for the algebra square root operation of $\mathbf{Q}_k(n)$'s entries. $\mathbf{n}_k(n)$ is the noise vector whose elements are i.i.d. zero mean complex Gaussian random variable with variance N_0.

User-k generates an estimate $\bar{\mathbf{r}}_k(n)$ for $\mathbf{r}_k(n)$ by multiplying the conjugate transpose of left singular vector matrix $\mathbf{U}_k(n)$,

$$\bar{\mathbf{r}}_k(n) = \mathbf{\Lambda}_k(n)\mathbf{Q}_k^{1/2}(n)\mathbf{s}_k(n) + \mathbf{U}_k^H(n)\mathbf{n}_k(n) \qquad (2)$$

Water-flling power allocation [18] could achieve the optimal throughput performance. However, in many practical systems it puts a high demand on the linear range of transmit power amplifiers, which is extremely costly especially for multiple antenna systems. Thus in this paper, equal power allocation [14, 19] is used at BS. Moreover, we assume each MS has 2 antennas. The time index n is omitted for simplicity when discussing within a single time slot.

III. Single-Level Threshold Adaptive Transmission Strategy

A. Feedback strategy at MS

Recall that $N_T > N_k$ and each MS is equipped with 2 antennas, i.e. rank$(\mathbf{H}_k) = 2$ and $T_k \in \{1, 2\}$. Use R_k^{BF} and R_k^{SM} to denote the achievable rate of user-k in one time slot using BF and SM, respectively. They are given by (3) and (4), where B denotes the transmission bandwidth.

$$R_k^{BF} = B\log_2\left(1 + P_T\lambda_{k,1}^2/N_0\right) \qquad (3)$$

$$R_k^{SM} = B\sum_{i=1}^{T_k}\log_2\left(1 + \frac{P_T\lambda_{k,i}^2}{N_T N_0}\right) \qquad (4)$$

$\lambda_{k,1}$ and $\lambda_{k,i}$ stand for the maximum and the ith singular value of $\mathbf{H}_k$, respectively. If SM achieves higher capacity multiple subchannels are activated for transmission, otherwise BF is better and only the principal eigenmode is selected.

We employ single-level threshold Q_{pri}^{th} in the strategy. In order to implement adaptive transmission we introduce another threshold Q_{sec}^{th} derived from Q_{pri}^{th}, which will be discussed later. In one slot n, MSs decide locally whether they should attempt to access the channel and send feedback to the BS or not. In the negative case, they remain silent for that slot. The feedback strategy at MS is described in Fig. 1, where $\mathbf{v}_{k,1}$ is the principal right singular vector (PRSV) corresponding to $\lambda_{k,1}$.

B. Scheduling strategy at BS

On receiving feedback information, BS selects one user from the pool of eligible MSs [8]. We assume both CDI and CQI are available at BS and employ greedy scheduling strategy as described in Fig. 2, where L denotes the number of eligible users in one slot.

It is possible that for some time slots the BS cannot identify any eligible users to be considered for transmission. This happens in the case where no user has good enough channel gain higher than the preset threshold. Thus the number of feedback users is zero. This condition is defined as outage. In our discussion BS terminates transmission when outage occurs, i.e. the throughput of an outage slot is 0.

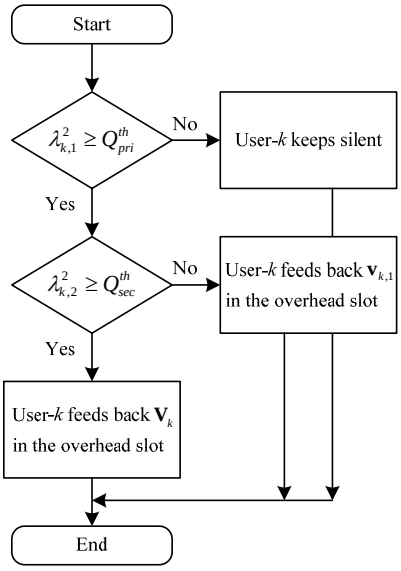

Figure 1. Feedback strategy at MS applying single-level threshold.

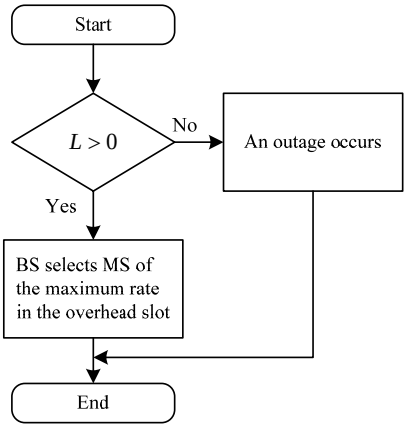

Figure 2. Scheduling strategy at BS applying single-level threshold.

C. Threshold design

Most existing works discussed the threshold design under simplified system model and resulted in the closed-form formulas [20]. As in general MIMO scenario the distribution of $\lambda_{k,i}^2$ is too complicated, we use numerical searching method to find Q_{pri}^{th}.

We assume each right singular vector (RSV) $\mathbf{v}_{k,i}$ is quantized into M bits, and antenna configuration is $N_T = 8$, $N_k = 2$. Fig. 3 and Fig. 4 illustrate the throughput (normalized by B and T) and the average CDI feedback load (normalized by M) under $\tau/T = 5\%$, $K = 10$, different Q_{pri}^{th} and SNR (Signal to noise ratio), respectively.

From Fig. 3 we can see that as Q_{pri}^{th} becomes larger than 10.5dB the throughput starts to decrease notably. In Fig. 4 the feedback load monotonically reduces with increasing Q_{pri}^{th}. Based on both figures it can be concluded that under $K = 10$, in order to maintain good system throughput performance as well as low feedback load, the appropriate value of primary threshold should be 10.5dB. The same method can be applied to determine Q_{pri}^{th} when

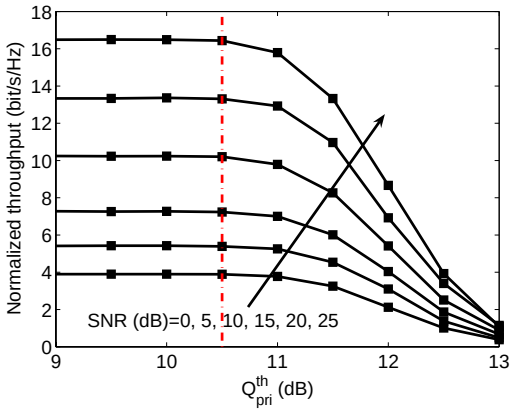

Figure 3. Normalized throughput per time slot under $\tau/T = 5\%$, $K = 10$, different Q_{pri}^{th} and SNR.

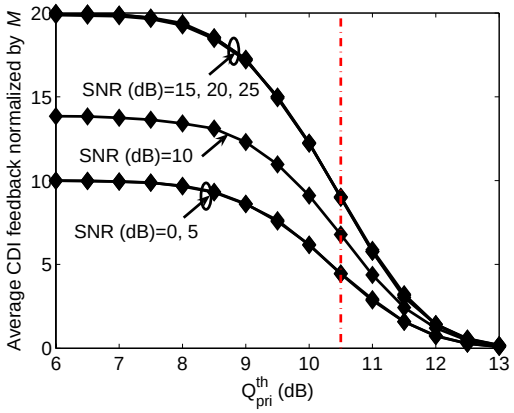

Figure 4. Average CDI feedback load per time slot normalized by M under $K = 10$, different Q_{pri}^{th} and SNR.

the number of users is taken other values. Q_{pri}^{th} under different K are presented in Table I.

In practical the threshold level can be set by the BS based on CSI collection within a certain period of time. Moreover, the threshold value could be adaptive to the communication environment to achieve better performance.

TABLE I.
Q_{pri}^{th} VALUES APPLYING SINGLE-LEVEL THRESHOLD

K	3	5	10	15	20	30	40
Q_{pri}^{th}(dB)	9.5	10.0	10.5	10.8	11.0	11.3	11.5

According to the algorithm in Fig. 1, we design Q_{sec}^{th} as follows. From (3) and (4) we have

$$R_k^{SM} \geq R_k^{BF} \Leftrightarrow \lambda_{k,2}^2 \geq \frac{(N_T - 1)\,\lambda_{k,1}^2}{1 + \gamma_0 \lambda_{k,1}^2 / N_T} \tag{5}$$

where $\gamma_0 = P_T/N_0$. The realtionship of Q_{pri}^{th} and Q_{sec}^{th} is given by (6),

$$Q_{sec}^{th} = \frac{(N_T - 1)\,Q_{pri}^{th}}{1 + \gamma_0 Q_{pri}^{th} / N_T} + \beta \tag{6}$$

β is the function of $\lambda_{k,1}$ and Q_{pri}^{th},

$$\beta = N_T \frac{\lambda_{k,1}^2 - Q_{pri}^{th}}{\left(1 + \gamma_0 Q_{pri}^{th}/N_T\right)^2} \tag{7}$$

The derivation of β is described in Appendix. Q_{sec}^{th} for each time slot could be readily obtained by using (6), (7) and values in Table I. Moreover, from (7) it can be seen that with fixed Q_{pri}^{th}, larger $\lambda_{k,1}$ would result in greater β as well as Q_{sec}^{th}.

D. Performance analysis

In this section we give the theoretical analysis of downlink throughput and system feedback load. Recall that L denotes the number of eligible MSs in one slot. According to the algorithm in Fig. 1, the throughput normalized by B and T is as follows,

$$Th = \begin{cases} \left(1 - \frac{\tau}{T}\right) \max\limits_{k=1,\cdots,L} \left(R_k^{BF}, R_k^{SM}\right), & L > 0 \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

The CDI feedback load (normalized by M) of L eligible users of which l users select BF in one slot, is given by (9).

$$F(L,l) = \begin{cases} 2ML - Ml, & L > 0 \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

Since one RSV corresponds to one singular value, CDI load varies in the same manner as CQI load does. Thus the later is not investigated in this paper.

Define the cumulative distribution function (CDF) of the ith eigenvalue $\zeta_{k,i}$ of $\mathbf{H}_k\mathbf{H}_k^H$ as $P_{\zeta_{k,i}}(u)$, $i \in \{1,2\}$. And $p_{\zeta_{k,i}}(u)$ represents the corresponding probability density function (PDF). For example, the probability that $\zeta_{k,i}$ is less than Q_{pri}^{th} is $P_{\zeta_{k,i}}(Q_{pri}^{th})$. Note that $\zeta_{k,i} = \lambda_{k,i}^2$. Based on the assumption that each user undergoes statistically the same independent fading, for any $k \in \{1,\cdots,K\}$ we have $P_{\zeta_{k,i}}(u) = P_{\zeta_i}(u)$ and $p_{\zeta_{k,i}}(u) = p_{\zeta_i}(u)$.

The probabilities that an arbitrary user is eligible for feeding information in an overhead slot is

$$P^{fd} = 1 - P_{\zeta_1}\left(Q_{pri}^{th}\right) \tag{10}$$

In Table II the possible states of MS are listed.

TABLE II.
POSSIBLE STATES OF MS APPLYING SINGLE-LEVEL THRESHOLD

State	Description
S^{BF}	Select BF and feed $\mathbf{v}_{k,1}$ in the overhead slot.
S^{SM}	Select SM and feed $\mathbf{V}_{k,1}$ in the overhead slot.
S_o	MS stays quiet.

(11)-(13) give the probabilities that an arbitrary MS is in one of the above states, respectively.

$$P^{BF} = Prob\left(S^{BF}\right) = P^{fd} P_{\zeta_2}\left(Q_{sec}^{th}\right) \tag{11}$$

$$P^{SM} = Prob\left(S^{SM}\right) = P^{fd}\left(1 - P_{\zeta_2}\left(Q_{sec}^{th}\right)\right) \tag{12}$$

$$P_o = Prob\left(S_o\right) = 1 - P^{fd} \tag{13}$$

From (11)-(13) we can easily have

$$P^{BF} + P^{SM} + P_o = 1 \tag{14}$$

Denote the number of eligible users in one slot as $\mathscr{N}$. We have

$$P_L = Prob\left(\mathscr{N} = L\right) = \begin{pmatrix} K \\ L \end{pmatrix} \left(P^{fd}\right)^L \left(1 - P^{fd}\right)^{K-L} \tag{15}$$

$$P_o^{sys} = Prob\left(\mathscr{N} = 0\right) = \left(1 - P^{fd}\right)^K \tag{16}$$

(15) denotes the probability that the number of eligible MSs in one slot is L. (16) indicates the system outage probability.

From (10)-(16), the expectation of throughput and CDI feedback load can be readily obtained,

$$\mathbb{E}\left(Th\right) = \left(1 - \left(P_{\zeta_1}\left(Q_{pri}^{th}\right)\right)^K\right)\left(1 - \frac{\tau}{T}\right) E\left(\Phi\right) \tag{17}$$

where

$$\Phi = \max_{k=1,\cdots,L} \left(R_k^{BF}, R_k^{SM}\right) \tag{19}$$

Note that Q_{sec}^{th} is function of $\lambda_{k,1}$, which is a random variable. In statistical analysis we use the expectation $\mathbb{E}(Q_{sec}^{th})$ instead of Q_{sec}^{th} to compute (11) and (12).

In order to calculate (10)-(18), the PDF of $\zeta_{k,i}$ should be obtained beforehand. This distribution is given in [20] as follows,

$$p_{\zeta_{(\cdot)}}(u) = \varphi \sum_{n=1}^{N_k} \sum_{m=1}^{N_k} (-1)^{n+m} u^{\alpha} \exp^{-u} \left|\mathbf{\Omega}_{(\cdot)}\right| \tag{20}$$

where $\alpha = n + m - 2 + N_T - N_k$. In this paper $p_{\zeta_{(\cdot)}} \in \{\zeta_1, \zeta_2\}$, $\zeta_1 > \zeta_2$. φ is a constant. $\mathbf{\Omega}_{(\cdot)}$ is an $(N_k - 1) \times (N_k - 1)$ square matrix. For room limitation the details of (20) are not elaborated.

IV. DOUBLE-LEVEL THRESHOLD ADAPTIVE TRANSMISSION STRATEGY

As for the single-level threshold transmission strategy, proper design of the threshold is difficult. Since a tight one would result in high outage probability whereas a loose one could not reduce feedback load effectively. Thus a double-level threshold scheme is proposed in this section.

A. Feedback strategy at MS

Similar to the algorithm in Fig. 1, we adopt two overhead slots here. In the first one, two tight thresholds $Q_{1,pri}^{th}$ and $Q_{1,sec}^{th}$ are employed to implement user and transmission mode selection, respectively. In the second overhead slot, loose thresholds $Q_{2,pri}^{th}$ and $Q_{2,sec}^{th}$ are used. As depicted in Section III, $Q_{m,sec}^{th}$ is derived from $Q_{m,pri}^{th}$ $(m = 1, 2)$. In one slot n, the feedback algorithm at MS is given in Fig. 5.

B. Scheduling strategy at BS

The scheduling algorithm at BS is given in Fig. 6, where L_1 and L_2 denote the number of eligible users in the first and second overhead slot, respectively. We assume that BS terminates transmission when outage takes place.

$$\mathbb{E}(F)=\sum_{L=1}^{K}\left\{\binom{K}{L}\left(1-P^{fd}\right)^{K-L}\sum_{l=0}^{L}\left[\binom{L}{l}\left(P^{BF}\right)^{l}\left(P^{SM}\right)^{L-l}(2ML-Ml)\right]\right\} \tag{18}$$

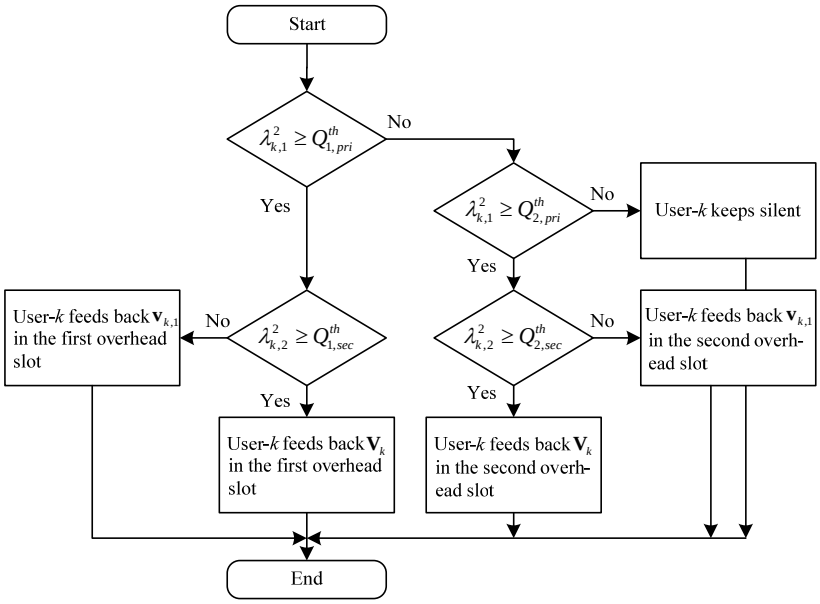

Figure 5. Feedback strategy at MS applying double-level threshold.

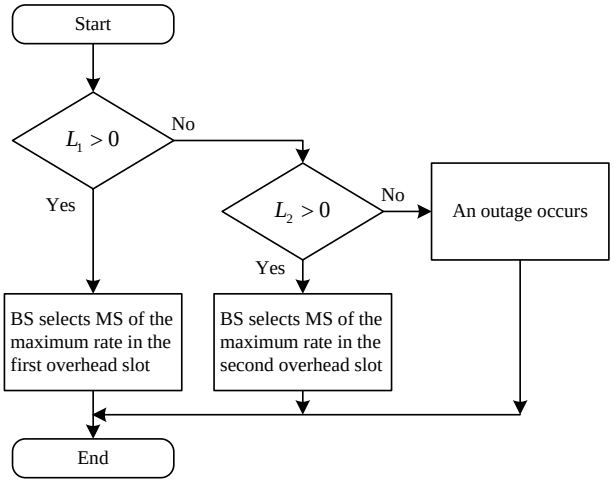

Figure 6. Scheduling strategy at BS applying double-level threshold.

C. Threshold design

We adopt the same idea as described in Section III. Accordingly, (6) and (7) can be rewritten in general form as follows,

$$Q_{m,sec}^{th}=\frac{\left(N_T-1\right)Q_{m,pri}^{th}}{1+\gamma_0 Q_{m,pri}^{th}/N_T}+\beta_m \tag{21}$$

where $m=1,2$. β_m is given by (22),

$$\beta_m=N_T\frac{\lambda_{k,1}^2-Q_{m,pri}^{th}}{\left(1+\gamma_0 Q_{m,pri}^{th}/N_T\right)^2} \tag{22}$$

Recall that $Q_{2,pri}^{th}$ is to guarantee the system outage performance. In order to make comparison to the single-level threshold strategy, $Q_{2,pri}^{th}$ is designed following the results in Table I. As a result the outage probabilities of both strategies would be the same.

Note that $Q_{1,pri}^{th}$ should be designed to effectively reduce the number of feedback users in the first overhead slot, thus it should be greater than $Q_{2,pri}^{th}$. Define the gap between these two thresholds as $\delta = Q_{1,pri}^{th} - Q_{2,pri}^{th}$. Fig. 7 illustrates the normalized average CDI feedback load per time slot under $K = 10$, different δ and SNR values.

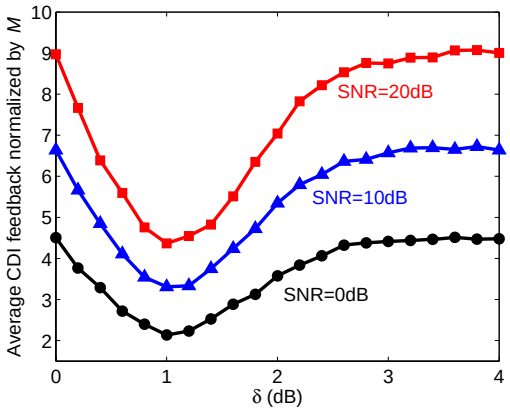

Figure 7. Average CDI feedback load per time slot normalized by M under $K = 10$, different δ and SNR.

As shown in the figure, when $\delta = 0$dB the feedback load of proposed scheme is the same with that of the single-level one, which can also be referred to Fig. 11. With increasing δ, $Q_{1,pri}^{th}$ becomes aggressive and results in decreasing number of eligible users. When $\delta < 1$dB,

$$Th = \begin{cases} \left(1-\frac{\tau}{T}\right) \max\limits_{k=1,\cdots,L_1} \left(R_k^{BF}, R_k^{SM}\right), & L_1 > 0 \\ \left(1-2\frac{\tau}{T}\right) \max\limits_{k=1,\cdots,L_2} \left(R_k^{BF}, R_k^{SM}\right), & L_1 = 0 \quad \text{and} \quad L_2 > 0 \\ 0, & \text{otherwise} \end{cases} \tag{23}$$

$$P_o = Prob(S_o) = P_{\zeta_1}\left(Q_{1,pri}^{th}\right)\left(1-\left(P_{\zeta_1}\left(Q_{1,pri}^{th}\right)\right)^{K-1}\right) + P_{\zeta_1}\left(Q_{2,pri}^{th}\right)\left(P_{\zeta_1}\left(Q_{1,pri}^{th}\right)\right)^{K-1} \tag{29}$$

$$\mathbb{E}\left(Th\right) = \left(1-\left(P_{\zeta_1}\left(Q_{1,pri}^{th}\right)\right)^{K}\right)\left(1-\frac{\tau}{T}\right)E\left(\Phi_1\right) + \left(P_{\zeta_1}\left(Q_{1,pri}^{th}\right)\right)^{K}\left(1-\left(1-P_2^{fd}\right)^{K}\right)\left(1-2\frac{\tau}{T}\right)E(\Phi_2) \tag{32}$$

$$\mathbb{E}(F) = \sum_{m=1}^{2}\sum_{L_m=1}^{K}\binom{K}{L_m}(1-P_m^{fd})^{K-L_m}\sum_{l_m=0}^{L_m}\left[\binom{L_m}{l_m}(P_m^{BF})^{l_m}(P_m^{SM})^{L_m-l_m}(2ML_m - Ml_m)\right] \tag{33}$$

although the outage probability increases, the probability that scheduling can be carried out in the first overhead slot is still high. Thus the feedback load reduces along with increasing δ (under $\delta <$ 1dB). When δ grows larger than 1dB, the probability that the second overhead slot is needed for scheduling becomes higher, consequently the feedback load (yielded in two overhead slots) starts to increase. As δ becomes too large, the outage probability in the first overhead slot approximates to 1 and both schemes result in the same cost. Based on the discussion above, there exists an optimal δ to minimize the feedback load. As shown in Fig. 7 the proper δ is 1dB under $K = 10$.

We use numerical searching method to find $Q_{1,pri}^{th}$, the results are given in Table III.

TABLE III.
$Q_{1,pri}^{th}$ VALUES APPLYING DOUBLE-LEVEL THRESHOLD

K	3	5	10	15	20	30	40
$Q_{1,pri}^{th}$(dB)	10.8	11.1	11.5	11.7	11.9	12.1	12.3

D. Performance analysis

In this section we give the theoretical analysis applying double-threshold scheduling strategy. Denote L_1 and L_2 as the number of eligible MSs in the first and second overhead slot, respectively. The throughput normalized by B and T is given by (23).

Note that when $L_1 = 0$ and $L_2 > 0$, due to an outage occurs in the first overhead slot the throughput of double-level threshold scheme is inferior to that of the single-level one, i.e. throughput loss is resulted.

The normalized CDI feedback load of L_m eligible users, of which l_m users select BF in the mth $(m = 1, 2)$ overhead slot is

$$F(L_m, l_m) = \begin{cases} 2ML_1 - Ml_1, & L_1 > 0 \\ 2ML_2 - Ml_2, & L_1 = 0 \quad \text{and} \quad L_2 > 0 \\ 0, & \text{otherwise} \end{cases} \tag{24}$$

The probabilities that an arbitrary user is eligible for feeding information in the first and second overhead slot are

$$P_1^{fd} = 1 - P_{\zeta_1}\left(Q_{1,pri}^{th}\right) \tag{25}$$

$$P_2^{fd} = \left(1-P_1^{fd}\right)^{K-1}\left(P_{\zeta_1}\left(Q_{1,pri}^{th}\right) - P_{\zeta_1}\left(Q_{2,pri}^{th}\right)\right) \tag{26}$$

Each MS has five states, listed in Table IV.

TABLE IV.
POSSIBLE STATES OF MS APPLYING DOUBLE-LEVEL THRESHOLD

State	Description
S_1^{BF}	Select BF and feed $\mathbf{v}_{k,1}$ in the overhead slot.
S_1^{SM}	Select SM and feed $\mathbf{V}_{k,1}$ in the overhead slot.
S_2^{BF}	Outage occurs in the first overhead slot. Select BF and feed $\mathbf{v}_{k,1}$ in the second overhead slot.
S_2^{SM}	Outage occurs in the first overhead slot. Select SM and feed $\mathbf{V}_{k,1}$ in the second overhead slot.
S_o	MS stays quiet.

(27)-(29) give the probabilities that an arbitrary MS is in one of the above states.

$$P_m^{BF} = Prob\left(S_m^{BF}\right) = P_m^{fd}P_{\zeta_2}\left(Q_{m,sec}^{th}\right) \tag{27}$$

$$P_m^{SM} = Prob\left(S_m^{SM}\right) = P_m^{fd}\left(1-P_{\zeta_2}\left(Q_{m,sec}^{th}\right)\right) \tag{28}$$

The first part in the RHS (Right hand side) of (29) denotes the probability that an arbitrary MS is ineligible in the first overhead slot and at least one of the other $K-1$ users is eligible for transmission. The second part indicates the probability that both the MS in the second overhead slot and the other $K-1$ users in the first overhead slot are ineligible for transmission.

From (27)-(29) we can easily have

$$\sum_{m=1}^{2}\left(P_m^{BF} + P_m^{SM}\right) + P_o = 1 \tag{30}$$

Denote the number of eligible users in the mth overhead slot for an arbitrary slot as $\mathcal{N}_m$ $(m = 1, 2)$. We have

$$\begin{aligned} P_{L_m} &= Prob(\mathcal{N}_m = L_m) \\ &= \binom{K}{L_m}\left(P_m^{fd}\right)^{L_m}\left(1-P_m^{fd}\right)^{K-L_m} \end{aligned} \tag{31}$$

From (27)-(29) and (31), the expectation of throughput and CDI feedback load can be readily obtained in (32) and (33). $\Phi_m = \max\limits_{k=1,\cdots,L_m}\left(R_k^{BF}, R_k^{SM}\right)$. Similar to the discussion in Section III, we also use $\mathbb{E}(Q_{m,sec}^{th})$ instead of $Q_{m,sec}^{th}$ in theoretical analysis.

V. Simulation Results

In this section, we use simulation results to illustrate the performance of the proposed scheduling strategies. Assume $K \in \{3, 10, 20, 30, 40\}$ and antenna configuration $N_T = 8$, $N_k = 2$. Six strategies are under investigation: $GS_{BF}woT$ (Greedy scheduling of BF users without threshold), $GS_{SM}woT$ (Greedy scheduling of SM users without threshold), $GS_{MA}woT$ (Greedy scheduling of mode-adaptive users without threshold), $GS_{MA}wSLT$ (Greedy scheduling of mode-adaptive users with single-level loose threshold, applying the values in Table I), $GS_{MA}wSTT$ (Greedy scheduling of mode-adaptive users with single-level tight threshold, applying the values in Table III) and $GS_{MA}wDT$ (Greedy scheduling of mode-adaptive users with double-level threshold). The first two schemes are similar to the conventional OB [8]. For the strategies without threshold we also consider one overhead slot overhead for BS to send training sequence.

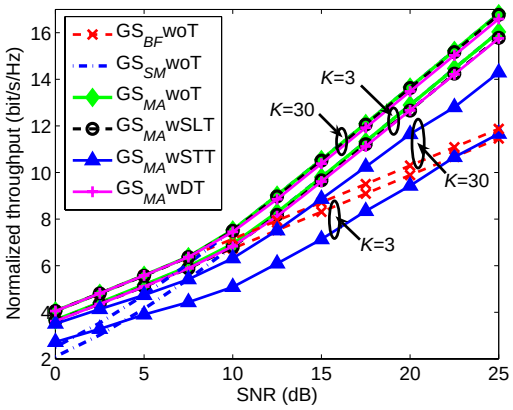

Figure 8. Normalized throughput per time slot with different strategies under $\tau/T = 5\%$, $K = 3, 30$ and different SNR.

Fig. 8 shows the normalized throughput per time slot for different schemes under $\tau/T = 5\%$, $K = 3, 30$ and different SNR. As shown in the figure, $GS_{BF}woT$ performs better at low SNR whereas $GS_{SM}woT$ achieves larger throughput at high SNR. Without threshold, every MS feeds CSI and the one of the maximum rate can be activated using greedy scheduling scheme. Thus $GS_{MA}woT$ is throughput optimal. With threshold, only eligible users feed information. The outage probability of a small K system is higher than that of a system with large K. Thus the throughputs of $GS_{MA}wSLT$, $GS_{MA}wSTT$ and $GS_{MA}wDT$ grow with increasing K. As figure shows, $GS_{MA}wSLT$ and $GS_{MA}wDT$ can achieve almost the same near-optimal throughput. This is because both schemes apply the same loose threshold values. Furthermore, since tight threshold results in high outage probability, the throughput of $GS_{MA}wSTT$ is poor compared with that of $GS_{MA}wSLT$ and $GS_{MA}wDT$.

From Fig. 8 it can be concluded that $GS_{MA}wSTT$ results in poor throughput performance. $GS_{MA}wSLT$ and $GS_{MA}wDT$ are of almost the same near-optimal throughput. As discussed in Section IV, double-level threshold strategy may incur throughout loss with respect to the single-level one. In order to illustrate the difference between $GS_{MA}wSLT$ and $GS_{MA}wDT$, Fig. 9 shows the relative throughput loss (RL_{Th}) of $GS_{MA}wDT$ with respect to $GS_{MA}wSLT$ under $\tau/T = 5\%$, different K and SNR. RL_{Th} is defined in (34),

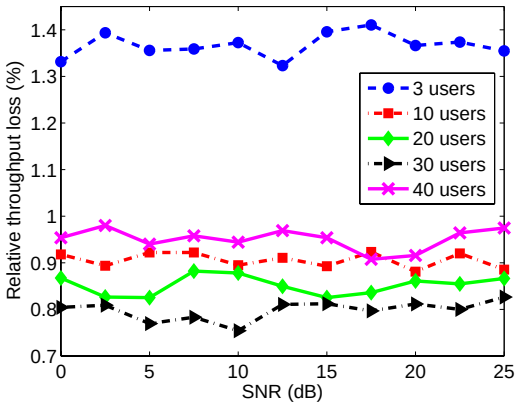

Figure 9. Relative throughput loss of $GS_{MA}wDT$ with respect to $GS_{MA}wSLT$ under $\tau/T = 5\%$, different K and SNR.

$$RL_{Th} = \frac{Th_{SLT} - Th_{DT}}{Th_{SLT}} \times 100\% \qquad (34)$$

where Th_{SLT} and Th_{DT} denote the throughput of $GS_{MA}wSLT$ and $GS_{MA}wDT$, respectively. Note that with properly designed $Q^{th}_{1,pri}$, BS starts data transmission immediately following the first overhead slot with high probability, thus both strategies are of almost the same throughput performance. As shown in the figure, RL_{Th} keeps low in the SNR region.

We may notice in Fig. 9 that the relative throughput loss for 40 users is less than for 3 users but greater than for 10, 20 and 30 users. This is due to the fact that the threshold values given in Table III for 40 users is a bit tighter, thus resulted in the higher probability that the second overhead slot is taken for scheduling which could have been used for data transmission. If the threshold is adjusted elaborately, the throughput loss would be reduced.

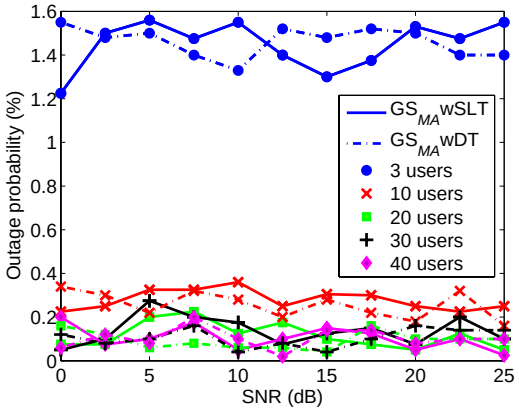

Figure 10. Outage probability of $GS_{MA}wSLT$ and $GS_{MA}wDT$ under different K and SNR.

Fig. 10 shows the outage probability of $GS_{MA}wSLT$ and $GS_{MA}wDT$ under different K and SNR values. The outage probability of $GS_{MA}wSTT$ ranges from 14% to 28% when K varies from 3 to 40. For space limitation these curves of poor performance are not plotted. As shown in Fig. 10, both schemes produce low outage probability in the SNR region. Their outage performance is statistically the same, which is consistent with the analysis in Section IV. As K increases, the outage probability approximates to zero. Since each MS undergoes independent fading, the probability that all users are in bad condition diminishes along with increasing K.

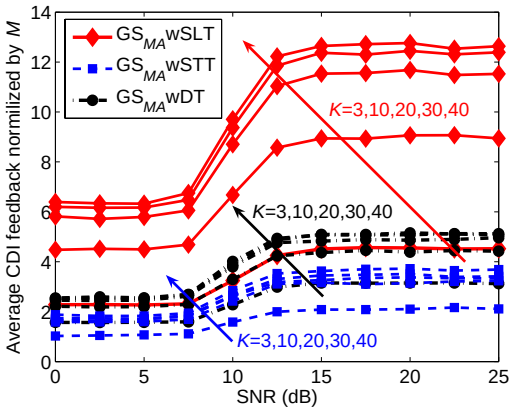

Figure 11. Average CDI feedback load per time slot normalized by M under different K and SNR.

Fig. 11 shows the average CDI feedback load per time slot normalized by M. The load of $GS_{BF}woT$, $GS_{SM}woT$ and $GS_{MA}woT$ are equal to K, $2K$ and varies from K to $2K$ along with increasing SNR, respectively [13]. For clarity, these curves are not plotted. As shown in the figure, BF is preferable at low SNR. SM becomes better at higher SNR, accordingly the load increases. When SNR$>$ 15dB, nearly all users select SM. With properly designed threshold, only limited number of eligible users feed information. Thus with fixed K the load of $GS_{MA}wSTT$, $GS_{MA}wSLT$ and $GS_{MA}wDT$ converges to an asymptotic value as SNR increases. As can be seen from the figure, applying $GS_{MA}wDT$ the feedback load can be significantly reduced compared with $GS_{MA}wSLT$. The load of $GS_{MA}wSTT$ is comparable with that of $GS_{MA}wDT$, yet the former is of poor outage performance.

Fig. 12 and Fig. 13 show the Normalized throughput per time slot and average CDI feedback load normalized by M using $GS_{MA}wDT$, respectively. Monte-Carlo simulation and statistical analysis are employed. It can be seen that the analytical results are consistent with the simulative one. When $GS_{MA}wSLT$ or $GS_{MA}wSTT$ is under investigation, the similar consistency can be obtained.

VI. Conclusion and Future Work

In wireless communications especially in multiuser scenarios, achieving good system performance with reduced cost is an interesting research issue. In this paper we proposed two adaptive opportunistic transmission strategies in MU-MIMO downlink with reduced feedback. Threshold policy in company with transmission mode adaptation are employed to handle feedback cost and improve system performance. Each MS carries out channel estimation, selects appropriate transmission mode and feeds back information adaptively. With feedback CSI, BS applies greedy scheduling method to activate the user of the maximum rate in each time slot. One of the proposed strategies employs single-level threshold ($GS_{MA}wSLT$). The other one employs double-level threshold ($GS_{MA}wDT$). Results show that with properly designed thresholds, both schemes can greatly reduce feedback load and achieve high throughput as well as low outage probability. Moreover, compared with $GS_{MA}wSLT$, $GS_{MA}wDT$ can further reduce the load and maintain the same outage performance at the cost of negligible throughput loss.

An interesting extension of this paper is to consider more generalized scenario where more users are scheduled from the pool of MS candidates. In this case, multiuser scheduling and MUI elimination should be taken into account. It can be expected that with properly

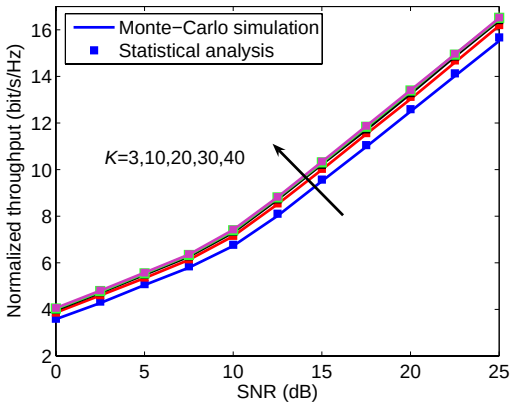

Figure 12. Normalized throughput per time slot under different K and SNR using Monte-Carlo simulation and statistical analysis.

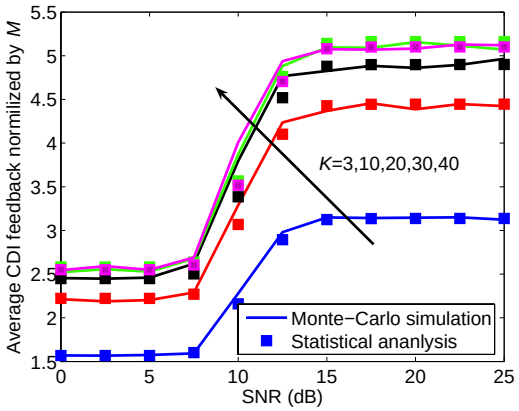

Figure 13. Average CDI feedback load per time slot normalized by M under different K and SNR using Monte-Carlo simulation and statistical analysis.

designed thresholds the modified strategy could also result in good performance. Moreover, this paper uses numerical searching in threshold design. This method is feasible in theoretical analysis. If the relation of threshold values and system parameters (such as antenna configuration, link statistics, target outage probability, and etc.) could be specified, the proposed strategies would be of more practical significance.

APPENDIX I
DERIVATION OF β

The threshold for the secondary eigenmode, Q_{sec}^{th} should be designed to ensure that secondary subchannel be appropriately activated, i.e. when $\lambda_{k,1}^2 \geq Q_{pri}^{th}$ and $\lambda_{k,2}^2 \geq Q_{sec}^{th}$ (5) should hold. β is derived as follows,

Under $\lambda_{k,1}^2 \geq Q_{pri}^{th}$ and $\lambda_{k,2}^2 \geq Q_{sec}^{th} = \frac{(N_T-1)Q_{pri}^{th}}{1+\gamma_0 Q_{pri}^{th}/N_T} + \beta$. If (5) does not hold, i.e. $R_k^{SM} < R_k^{BF} \Leftrightarrow \lambda_{k,2}^2 < \frac{(N_T-1)\lambda_{k,1}^2}{1+\gamma_0\lambda_{k,1}^2/N_T}$. We have

$$\frac{(N_T-1)\,Q_{pri}^{th}}{1+\gamma_0 Q_{pri}^{th}/N_T} + \beta < \frac{(N_T-1)\,\lambda_{k,1}^2}{1+\gamma_0\lambda_{k,1}^2/N_T} \tag{35}$$

and

$$\beta < \frac{(N_T-1)\left(\lambda_{k,1}^2 - Q_{pri}^{th}\right)}{\left(1+\gamma_0\lambda_{k,1}^2/N_T\right)\left(1+\gamma_0 Q_{pri}^{th}/N_T\right)} \tag{36}$$

Since $\lambda_{k,1}^2 \geq Q_{pri}^{th}$, the RHS of (36) is non-negative.

Based on the analysis above, it can be readily deduced that as long as β satisfies (37), (5) would hold under $\lambda_{k,1}^2 \geq Q_{pri}^{th}$ and $\lambda_{k,2}^2 \geq Q_{sec}^{th}$.

$$\begin{aligned}\beta &= \frac{N_T\left(\lambda_{k,1}^2 - Q_{pri}^{th}\right)}{\left(1+\gamma_0 Q_{pri}^{th}/N_T\right)^2} \\ &> \frac{(N_T-1)\left(\lambda_{k,1}^2 - Q_{pri}^{th}\right)}{\left(1+\gamma_0\lambda_{k,1}^2/N_T\right)\left(1+\gamma_0 Q_{pri}^{th}/N_T\right)}\end{aligned} \tag{37}$$

(7) is achieved.

ACKNOWLEDGMENT

This work was supported by National Science Fund for Distinguished Young Scholars (60725105), National Basic Research Program of China (973 Program) (2009CB320404), Program for Changjiang Scholars and Innovative Research Team in University (IRT0852), the National Nature Science Foundation of China (60902032), the 111 Project (B08038), the Key Project of Chinese Ministry of Education (107103), the Fundamental Research Funds for the Central Universities (JY10000901030).

REFERENCES

[1] M. Costa, "Writing on dirty paper," *IEEE Trans. on Inf. Theory*, vol. 29, no.3, pp. 439-441, 1983.

[2] Lai-U Choi, R.D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. on Wireless Commun.*, vol. 3, no. 1, pp. 20-24, 2004.

[3] T. Yoo, A.J. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Commun.*, vol. 24, no. 3, pp. 528-514, 2006.

[4] X. Zhang, J. Lee, H. Liu, "Low complexity multiuser MIMO scheduling with channel decomposition," in *Proc. of IEEE Intern. Conf. on Wireless Commun. and Networking*, Hong Kong, pp. 2452-2456, 2007.

[5] R. Chen, A.J. Goldsmith, R.W. Heath, Z. Shen, "Low-complexity user and antenna selection for multiuser MIMO systems with block diagonalization," in *Proc. of IEEE Intern. Conf. on Acoustics, Speech, and Sig. Process.*, Honolulu, pp. 613-616, 2007.

[6] W. Zhang, K.B. Letaief, "MIMO Broadcast Scheduling with Limited Feedback," *Proc. of IEEE Journal on Selected Areas in Commun.*, vol. 25, no.7, pp. 1457-1467, 2007.

[7] W. Xu, C. Zhao, Z. Ding, "Efficient User Scheduling under Low Rate Feedback for Correlated MIMO Broadcast Channels," in *IEEE Intern. Conf. on Commun.*, Beijing, pp. 3658-3663, 2008.

[8] D. Gesbert, M.-S. Alouini, "How much feedback is multiuser diversity really worth?" in *Proc. of IEEE Intern. Conf. on Commun.*, Paris, pp. 234-238, 2004.

[9] S. Sanayei, A. Nosratinia, "Exploiting multiuser diversity with only 1-bit feedback," in *Proc. of IEEE Wireless Commun. and Networking Conf.*, Sydney, pp. 978-983, 2005.

[10] P. Viswanath, D.N.C. Tse, R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. on Inf. Theory*, vol. 48, no.6, pp. 1277-1294, 2002.

[11] Il-Min Kim, Seung-Chul Hong, et al., "Opportunistic beamforming based on multiple weighting vectors," *IEEE Trans. on Wireless Commun.*, vol. 4, no.6, pp. 2683-2687, 2005.

[12] M. Nicolaou, A. Doufexi, S. Armour, "Reducing feedback requirements of the multiple weight opportunistic beamforming scheme via selective multiuser diversity," in *Proc. of IEEE Veh. Technol. Conf.*, Singapore, pp. 1-5, 2008.

[13] Z. Li, J. Yang, J. Li, J. Yao, "Opportunistic Transmission in MU-MIMO Downlink with Reduced Feedback and Mode Adaptation," in *Proc. of Intern. Conf. on Wireless VITAE*, Aalborg, pp. 379-383, 2009.

[14] A. Forenza, M.R. McKay, et al., "Adaptive MIMO Transmission for Exploiting the Capacity of Spatially Correlated Channels," *IEEE Trans. on Veh. Technol.*, vol. 56, no. 2, pp. 619-629, 2007.

[15] X. Tang, M.-S. Alouini, A.J. Goldsmith, "Effect of channel estimation error on M-QAM BER performance in Rayleigh fading," *IEEE Trans. on Commun.*, vol. 47, no. 12, pp. 1856-1864, 1999.

[16] J.S. Evans, "Optimal resource allocation for pilot symbol aided multiuser receivers in Rayleigh faded CDMA channels," *IEEE Trans. on Commun.*, vol. 50, no. 8, pp. 1316-1325, 2002.

[17] Y. Shao, J. Yuan, "A lower bound to the sum-rate of MIMO broadcast channels with limited-rate feedback," in *Proc. of IEEE Intern. Conf. on Commun.*, Beijing, pp. 3674-3678, 2008.

[18] D.P. Palomar, J.R. Fonollosa, "Practical Algorithms for a Family of Waterfilling Solutions," *IEEE Trans. on Sig. Process.*, vol. 53, no. 2, pp. 686-695, 2005.

[19] K. Zhang, Z. Niu, "Multiuser MIMO Downlink Transmission Over Time-Varying Channels," in *Proc. of IEEE Intern. Conf. on Commun.*, Glasgow, pp. 5514-5518, 2007.
[20] A. Zanella, M. Chiani, M.Z. Win, "A General Framework for the Distribution of the Eigenvalues of Wishart Matrices," in *Proc. of IEEE Intern. Conf. on Commun.*, Beijing, pp. 1271-1276, 2008.

Zhao Li was born in Xi'an, China. He received the B.S degree, M.S. degree and the Ph.D. degree in telecommunication engineering from Xidian University, Xi'an, China in 2003, 2006 and 2010, respectively. Since 2006, he has been with the State Key Laboratory of Integrated Service Networks at Xidian University. His research interests are in the areas of MIMO systems, multiuser communications and resource management in wireless networks.

Jiawei Yang was born in Jiangsu Province, China. He received the B.S. degree of remote sensing from Harbin Engineering University, Harbin, China in 1970, and the M.S. degree of telecommunication engineering from Xidian University, Xi'an, China in 1983. He was a visiting scholar at the University of Liverpool, Britain, from 1988 to 1990. He is a professor of telecommunication engineering at Xidian University. His research interests focus on wireless communications, characteristics of radio propagation and data transmission. Professor Yang is the senior member of China institute communications and Chinese institute of electronics.

Junliang Yao was born in Shanxi Province, China. He received the B.S. degree in telecommunication engineering from North China Electric Power University, Hebei Province, China in 2005. He is currently working towards the Pd.D. degree in the State Key Laboratory of Integrated Service Networks at Xidian University, Xi'an, China. His research interests are in the areas of signal processing and wireless communications.

A Location-Aware Vertical Handoff Algorithm for Hybrid Networks

Abolfazl Mehbodniya†, Sonia Aïssa†‡
† INRS-EMT, University of Quebec,
Montreal, QC, Canada.
‡ King Abdullah Univ. of Science and Technology (KAUST),
Jeddah, Kingdom of Saudi Arabia.
Email: mehbod@emt.inrs.ca, sonia.aissa@ieee.org

Jalil Chitizadeh
Ferdowsi University, Mashhad, Iran
Email: chitizad@ferdowsi.um.ac.ir

***Abstract*— One of the main objectives of wireless networking is to provide mobile users with a robust connection to different networks so that they can move freely between heterogeneous networks while running their computing applications with no interruption. Horizontal handoff, or generally speaking handoff, is a process which maintains a mobile user's active connection as it moves within a wireless network, whereas vertical handoff (VHO) refers to handover between different types of networks or different network layers. Optimizing VHO process is an important issue, required to reduce network signalling and mobile device power consumption as well as to improve network quality of service (QoS) and grade of service (GoS). In this paper, a VHO algorithm in multitier (overlay) networks is proposed. This algorithm uses pattern recognition to estimate user's position, and decides on the handoff based on this information. For the pattern recognition algorithm structure, the probabilistic neural network (PNN) which has considerable simplicity and efficiency over existing pattern classifiers is used. Further optimization is proposed to improve the performance of the PNN algorithm. Performance analysis and comparisons with the existing VHO algorithm are provided and demonstrate a significant improvement with the proposed algorithm. Furthermore, incorporating the proposed algorithm, a structure is proposed for VHO from the medium access control (MAC) layer point of view.**

***Index Terms*— Vertical handoff, Handoff, Mobile positioning, Overlay networks, 4G wireless networks, Heterogeneous networks.**

I. Introduction

In this section, we describe the handoff and vertical handoff (VHO) problems with their associated challenges, related work pertinent to this subject along with the paper's contribution.

A. Handoff and its Associated Problems

Handoff process is an important resource management module in most wireless networks. A new handoff process considered for the next generation of wireless networks is VHO. While the handoff process defines the transfer of an active mobile connection between cells within a single layer (also referred to as horizontal handoff), VHO defines this process between different layers [1]–[5]. Throughout this paper the term handoff is used mostly instead of horizontal handoff for writing simplification but it is important to mention that the two terms are analogous. Layered cell structure idea suggests using different cellular networks with different sizes and possibly different technologies, such that the user can still preserve its connection while moving out of the coverage of a lower layer, because it is still in the coverage of an upper layer. Another benefit of such networks is the capability of adapting network parameters to user's specifications, such as speed and demanded service. Using different types of networks (heterogeneous networks), their coordination, integration and VHO procedures is a significant challenge for beyond 3G and 4G wireless networks [6], [7].

Research on handoff is mainly categorized in two parts. In the first, handoff is addressed from the upper network layers point of view, which focus on data transfer and performance of network protocols, such as internet protocol (IP) and transmission control protocol (TCP) [8]–[10]. Works in the second category are done from the physical layer point of view, and mostly deal with radio propagation characteristics and handoff decision process. One significant characteristic of wireless systems is the signal variations caused by the movement of the mobile station (MS). The variations can cause some unnecessary handoffs on cell boundaries, a phenomenon referred to as ping-pong effect [2]. Such effect increases network traffic load which indirectly leads to an increase in the handoff blocking probability (HBP). The ping-pong effect also causes power loss at the transmitters and receivers. Crossover point is another metric used to compare different VHO algorithms, and is defined as the distance between the point where a handoff is made and the cell boundary. An ideal algorithm makes handoff on cells and layers boundaries. However, if the crossover distance is large, it can cause an increase in channel interference and a degradation in the

Manuscript received February 10, 2010; revised May 20, 2010; accepted June 13, 2010.
This work was partially supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

doi:10.4304/jcm.5.7.521-529

link Quality of Service (QoS). An idea for improving the precision of the handoff algorithm and the crossover point consists in exploiting the user position information. Different localization methods can be applied for subscriber-positioning in wireless networks. Localization methods can be classified in self-, remote- and indirect-positioning, according to the place where the measurements and evaluations take place [2]. In self-positioning systems, the handset itself performs the measurements and the necessary evaluations to finally estimate its own position. In remote-positioning systems, the fixed part of the network takes the measurements and estimates the terminal's position. Finally, in indirect-positioning systems the MS takes the measurements and transmits them to the base station (BS) for evaluation.

B. Related Work and Contribution

In [11], a policy-based VHO is proposed where different metrics are taken into consideration for the process of decision making in VHO such as user preferences and direction of the MS. Despite the distinctive feature of this work, the cost function does not consider signal variations and cellular boundaries, hence there is a high probability of ping-pong in the related algorithm. In [12], a VHO designed with fuzzy logic is proposed, which uses global positioning system (GPS) to obtain location information. In [13], a handoff scheme for multimedia networks has been proposed, however the work has been done just for flat networks and not for overlay networks. In [14], a VHO has been proposed considering channel distribution. This algorithm does not have a strategy for reducing the ping-pong effect and HBP. In [15], a location estimation technique based on pattern recognition is proposed to define optimum points for handoff. This algorithm is not designed for overlay networks and considers location estimation within cellular boundaries, which needs heavy calculations for advance selection of desired handoff locations. Another weak point in the latter algorithm is exploiting minimum distance criteria for pattern classification, and as it will be seen later detection error probability of this method is too high. In [16], the authors have proposed a received signal strength (RSS) classifier for path identification with application to handoff decision in multi-service networks. Their scenario only considers a restricted system consisting of two layers with only two cells. As a result, studying the effect of different network parameters on the pattern classification is not possible. In [18], a location estimation technique based on pattern recognition is proposed with probabilistic neural network (PNN) to be used with handoff algorithm similar to what is done in [15]. The algorithm considers only the traditional cellular structure with one layer and a simple scenario using only four BSs, and its handoff decision strategy is also different. Finally, [20] proposed an intelligent handoff algorithm using grey prediction and fuzzy decision. The grey prediction is used to predict the future RSS which is applied to a fuzzy decision system for handoff decision. This algorithm is designed for conventional single layer networks and needs significant computing resources.

In this paper, we propose an algorithm that improves the ping-pong effect and the crossover point. First, user location is estimated with a pattern recognition technique, and then location information is used in the VHO algorithm to deduce the right time and the right place at which the VHO should take place. The positioning technique proposed in this work is based on pattern recognition and we consider an indirect-positioning method. This technique is the most cost-effective technique and can be easily integrated into any wireless network process module [17]. Pattern recognition-based positioning can be a practical solution for underground and confined environments where using GPS is not possible. Our algorithm uses different RSS samples of BSs to construct the pattern classes. Since shadow fading is constant for a given path between a BS and a given point, pattern recognition techniques can be applied to mobile positioning algorithms. Due to the fact that the shadow fading is constant, the averaged signal samples of MSs, which travel along the same path will be similar and the averaged signal samples can be stored in nearby BS database. These samples are used later for positioning purposes using pattern recognition. Because of outstanding features of intelligent techniques, we use PNN as our pattern classifier similar to the work in [18]. Another contribution of this paper is the optimization process done for the pattern recognition part, which along with the proposed VHO algorithm shows significant performance improvement.

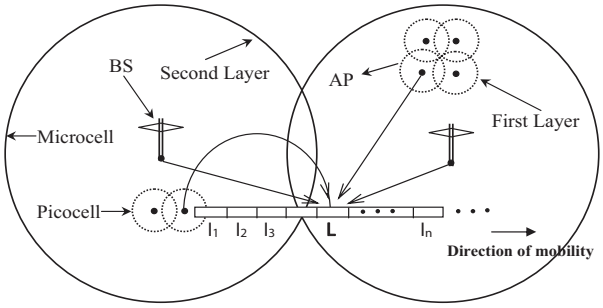

Fig. 1. The idea of using pattern recognition for vertical handoff.

The aforementioned process is categorized in the physical layer part of the network. As for the upper network layer point of view, we propose a structure for VHO between universal mobile telecommunication systems (UMTS) and wireless local area networks (WLAN), incorporating the location information in the signalling. The proposed algorithm can be generalized for any wireless system and is not necessarily restricted to UMTS and WLAN. As a matter of fact, this paper extends the approach in [19] which deals with physical layer based VHO. Moreover, we include the typical VHO structure proposal between UMTS and WLAN which discusses the corresponding procedures in network layers.

In the next section, the proposed idea for VHO in overlay networks is presented. In section III, further details on the PNN classifier are discussed. In section IV, the simulation environment and parameters are explained, and in section V performance analysis and simulation results are provided.

II. The Proposed Method

In this section, the proposed method for VHO is discussed from physical layer and upper layers point of view.

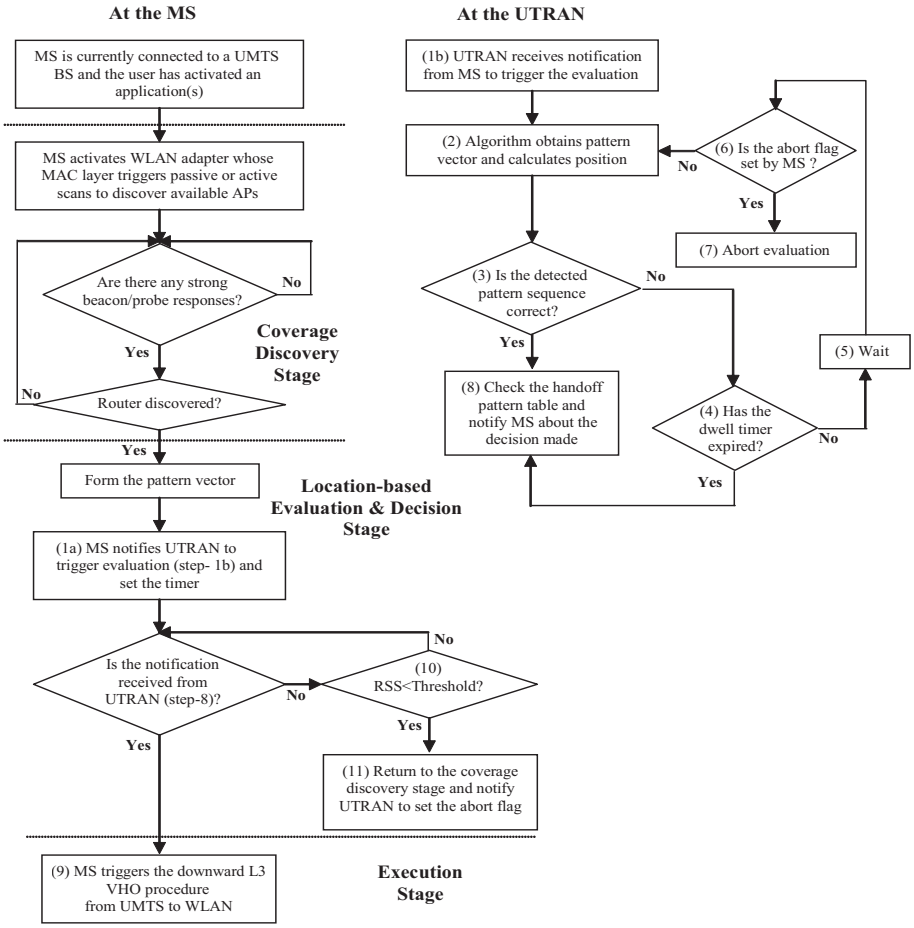

Fig. 2. The proposed algorithm for vertical handoff from UMTS to WLAN.

A. From Physical Layer Point of View

Fig. 1 illustrates the idea of using pattern classes for VHO in overlay networks. Let's assume that a given path (e.g., a road, a passage, a corridor or a mine gallery) within a wireless network coverage is divided into smaller blocks. Each block is called a pattern class, L, and is the representative of a small location area. As the MS travels along the specified path, it collects RSS samples for different cells in different layers with a specific sampling rate. During a training phase, a MS travels along a particular path and collects averaged signal samples from nearby BSs. Let $\xi^{ik} = [\xi_0^{ik}, \xi_1^{ik}, ..., , \xi_{N_{\max}-1}^{ik}]^T$ denote the averaged samples of ith BS in kth layer, B_{ik}. The total number of samples recorded from each BS is $N_{\max}$ and N_w is the block's length. Hence, the number of classes is given by:

$$C = \left\lfloor \frac{N_{\max}}{N_w} \right\rfloor, \tag{1}$$

where $\lfloor . \rfloor$ denotes the upper integer rounding operator. For each block, signal samples of different cells in different layers are arranged in vectors called pattern vectors, which are then used to train the PNN. Each pattern vector is identical to its block and is unique for that block in the area. After collecting all signal samples and forming the pattern vectors, the algorithm can run in online mode as MSs move along the path. When a complete pattern vector is formed, it will be classified within one of the existing pattern vectors. Knowing the pattern class number, we can decide on the handoff time so that to avoid unnecessary handoffs caused by uncertainty. But before that, we should somehow make sure about the correctness of the classification. For this purpose, we compare previously detected patterns with the current pattern number. If the difference is within a tolerable factor, T_M, the classified pattern is considered as correct and consequent decisions can be made:

$$\left\| \hat{l}_n - \hat{l}_{n-1} \right\| \leq \Delta l T_M, \tag{2}$$

where Δl is the distance between adjacent pattern vectors and $\hat{l}_n$ represents the n-th pattern vector.

B. From Upper Network Layers Point of View

The conventional method for discovering coverage and making the decision to trigger an upward or downward VHO, is primarily based on radio frequency (RF) measurements [3]. The basic approach consists of a MS currently connected to an upper layer and scanning for coverage by a lower layer. Upon discovering WLAN coverage through beacons sent by access points (APs) and receiving a router advertisement from the associated router, a downward VHO is activated. During

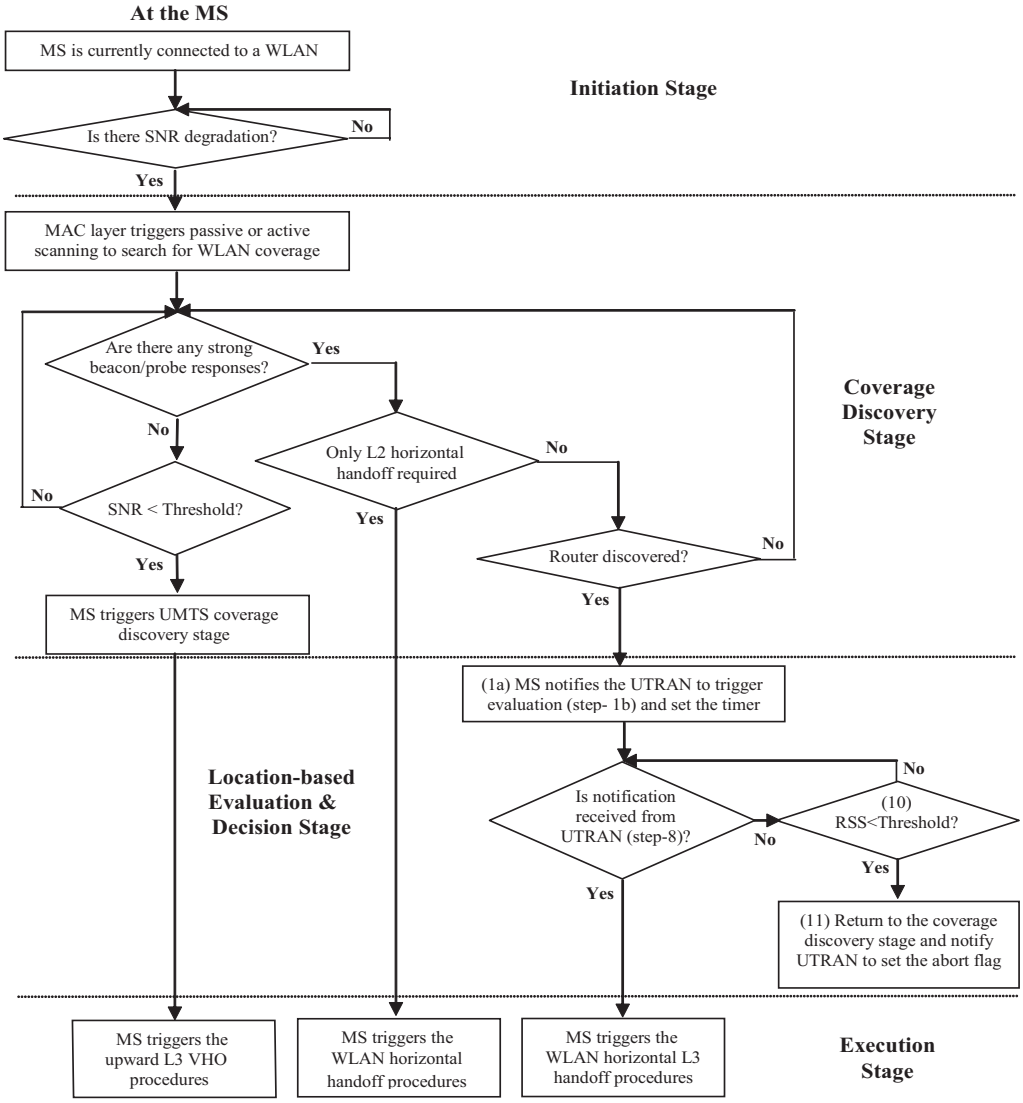

Fig. 3. The proposed algorithm for vertical handoff from WLAN to UMTS.

the MS's entrance into a WLAN cell, it remains connected to the current AP until a degradation in the signal-to-noise ratio (SNR) is sensed. The medium access control (MAC) layer is responsible for locating other APs. If another AP within the same hotspot is found, a layer 2 (L2) handoff (a handoff indicating that the packet header is processed up to data link layer) takes place without notifying the home network. However, if the AP belongs to another hotspot, then a Layer 3 (L3) handoff (a handoff indicating that the packet header is processed up to network layer) is needed. A horizontal L3 handoff to a visited WLAN network includes notifying the MS's home UMTS network, of the MS new location. On the other hand, if the discovery stage fails to locate other APs with strong beacons, an upward handoff is activated when the SNR value falls below a specific threshold. Upon such activation, the UMTS adapter uses RF measurements to discover and evaluate UMTS cells, and then triggers the L3 handoff once the discovery stage completes successfully. According to the relevant literature, there exist several schemes for evaluation of RF measurements to determine the best time to activate a handoff in homogeneous and heterogeneous networks [2].

Fig. 2 shows our proposed VHO algorithm from UMTS to WLAN. This algorithm is a general algorithm including upper layers point of view and is proposed to improve the performance of conventional algorithms. In Fig. 3, a similar algorithm for VHO but from WLAN to UMTS is proposed. These algorithms are a combination of the conventional algorithm in [21] and the above presented proposal for physical layer.

As illustrated in Fig. 2, the MS scans for coverage permanently using its WLAN network adaptor. If a strong signal is detected, the MS forms the pattern vector and notifies UMTS terrestrial radio access network (UTRAN) in step (1a), and then transfers the pattern vector to UTRAN. In step (1b), while UTRAN is notified, it estimates the location of the user (2), with the proposed algorithm, and checks for the correctness of

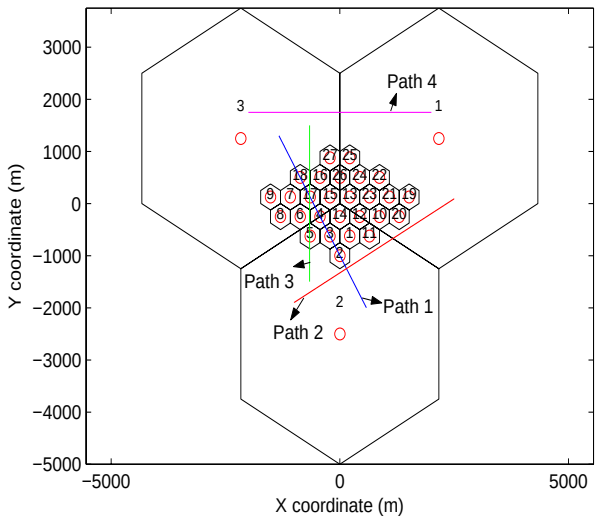

Fig. 4. The hierarchical cell layout in two-dimensional space.

the pattern sequence (3). If the detected sequence is correct, decisions will be made in step (8) and the MS is notified (9) to do a VHO of type L3 to WLAN. If there is no need for handoff, MS continues scanning for WLAN coverage as usual, and if RSS becomes less than a threshold, step (11), UTRAN will be notified to ignore the calculations. From UTRAN's point of view, in step (3), if the location estimation is not approved, a dwell timer will be used for handoff optimization and reduction of signal power fluctuations. The process is as follows: a constant value is set for this timer (1a), and when the algorithm enters UTRAN mode, the timer starts its count down. In step (4), the value of the timer is checked, if it reaches zero a handoff is established, otherwise the algorithm waits for few seconds. This interruption is considered for the purpose of reducing unnecessary calculations in UTRAN. After this interruption, step (11) at the MS side will be checked; if the abort flag was set, UTRAN mode calculations will be terminated, otherwise the algorithm returns back to UTRAN loop in step (2), and calculations and evaluations restart.

As for the Fig. 3, most parts of the algorithm are similar to downward VHO, the only difference is the utilization of the SNR for the detection of signal power degradation in WLAN cells.

III. Simulation Environment

In this section we discuss the structures and parameters considered for the PNN pattern classifier, radio propagation environment and VHO algorithm.

A. Pattern Classifier Structure

Minimum distance algorithm for pattern recognition does not provide enough accuracy. In order to achieve higher accuracy and lower execution time, neural networks are chosen for pattern classification. Neural networks have already been applied to various pattern recognition problems. Generally speaking, they can be regarded as techniques for nonlinear function approximation; pattern recognition can be regarded as a special case of function approximation where the function values form a discrete set. The output y_n of the n-th neuron in PNN structure is given by [18]:

$$y_n = e^{-\frac{\|X-W_n\|^2}{\sigma^2_{\text{PNN}}}}. \quad (3)$$

This activation function is a Normal function where $X = [x_1, x_2, ..., x_{N_P}]^T$ is the input vector to be classified or given to the network for training. In (3), $W_n = [w_{n,1}, w_{n,2}, ..., w_{n,N_P}]^T$ is the weight vector for n-th neuron in the network and σ^2_{PNN} is the smoothing parameter of the PNN. This parameter is in fact the bias value of the network. As it gets larger, the Normal function's slope gets smoother and several neuron's may respond to an input vector. In this case, the network acts like it is taking a weighted average between target vectors whose design input vectors are closest to the new input vector. On the other hand, if the value is small, only the neuron with the weight vector nearest to the input vector will respond and the network will function as a minimum distance algorithm. It is worth to mention that the main reason for choosing PNN for this algorithm among all other neural networks, is their high ability of classification, their very low computational complexity and their highest convergence time for this application. These features make PNN favorable to be implemented even at the mobile set processor unit for minimum delay real-time processing.

B. General Model Specifications

For our simulations, we consider a multi-layer cellular environment as shown in Fig. 4. It consists of two layers. The upper layer which is referred to as macrocell, has 3 cells and can be a network like UMTS. The lower layer, called microcell, has 27 cells and can be WLAN. Furthermore, it is assumed that the speed of the MS is known. Spatial sampling is used, which means that signals are sampled at constant distances. Four paths are considered as different scenarios. The user travels along each path and its position segment is estimated. This information can later be applied to the VHO algorithm to decide about the time and place of inter-layer or inter-cellular handoff. For simulations, we used *MATLAB* software and a set of functions called *RUNE* [22]. The time model in our simulator is a discrete-time step model. The model for the received signal (in dB) is given by:

$$P(d) = P_t + \alpha - 10\beta log(d) + f_i(d) + S(d), \quad (4)$$

where P_t is the effective transmit power, d is the distance to the transmitter, α is the path loss constant and β is the path loss exponent. In (4), $f_i(d)$ is the log-normal shadow fading factor and $S(d)$ denotes the Rayleigh fading parameter.

For each point in space, samples of shadow fading are the same and depend upon the following correlation function:

$$R_{ff}(\Delta) = E\{f(d+\Delta)f(d)\} = \sigma_s^2 e^{-\frac{|\Delta|}{d_0}}, \quad (5)$$

where σ_s is the standard deviation of shadow fading and d_0 is the correlation distance. Samples of this fading have a normal

TABLE I
MAIN PARAMETERS USED IN THE SIMULATIONS

Parameter	Microcell	Macrocell
Cell radius (m)	250	2500
Number of clusters	1	9
Attenuation at 1 meter distance, α (dB)	-40	-31
Thermal noise floor, N_0 (dBm)	-118	-118
Distance attenuation coefficient, β (dB)	3.3	4
Standard deviation of shadow fading, σ_s (dB)	7	6
Correlation distance, d_0 (m)	20	20
Sampling distance (m)	0.8	0.8

distribution. To model these samples in one-dimensional space, we use the following recursive function:

$$f_i(n) = x_i(n) + e^{-\frac{V.T}{d_0}} f_i(n-1), \quad (6)$$

where V is the average user velocity, T is the sampling period, d_0 is the correlation distance and x_i is a Gaussian random variable with zero mean and standard deviation given by:

$$\sigma = \sqrt{1 - e^{\frac{-2.V.T}{d_0}}}\ \sigma_s. \quad (7)$$

In a practical system, there is often a correlation between the lognormal shadow fading for the links between one MS and the BSs, modelled by assigning one random fading component related to each link, G_{ij}, and one fading component related to the MS, G_M. These are then added according to the expression [22]:

$$G = \sqrt{\rho}\ G_M + \sqrt{1-\rho}\ G_{ij}, \quad (8)$$

where ρ defines the effect of each component. In order to compensate the effect of Rayleigh fading, an averaging window is considered according to:

$$\overline{P}(n) = \frac{1}{n_w} \sum_{i=n-n_w-1}^{n} P(i), \quad (9)$$

where n_w is the averaging window size.

Table I indicates the parameters used in the simulations.

C. VHO Algorithm Specifications

Although PNN classifier has better performance than minimum distance classifiers, it can further be improved. We propose three methods for this purpose. The first technique sorts input data (training or test data) in an ascending or descending order before they are fed to the PNN network. In this way, random effects in the propagation will be compensated to some extent. The second technique consists of multiplying the test data by a magnitude factor (MF) before classification because training data sets, which are in fact averaged signal samples, have a small variation in scale and this magnitude factor will compensate this variation. The third technique uses the same averaging window size for training data or test data. The value we have chosen for the neural network bias is 10, so that the network takes into consideration the effect of all nearby neurons. The averaging window size is 10 in all simulation

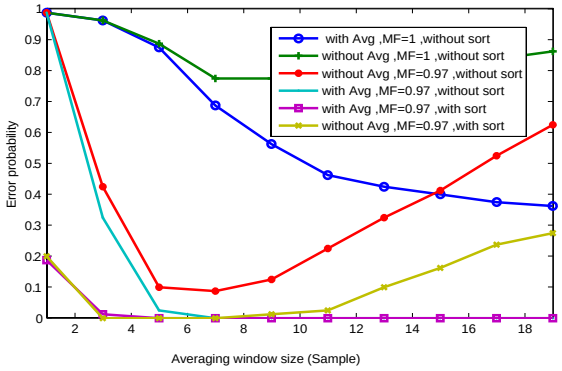

Fig. 5. DEP diagram for three proposed optimizations.

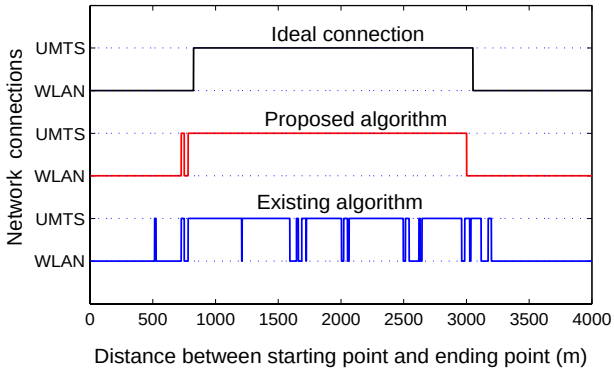

Fig. 6. Network connections for Path 1.

scenarios. The tolerance factor (TM) is set to 1, which means that only if the recent classified pattern completely matches the previous pattern, it will be valid.

In all scenarios, results of the proposed algorithm have been compared with the existing VHO algorithm. By the term existing algorithm, we refer to current available hysteresis based handoff and VHO algorithms, such as those in UMTS and Global System for Mobile Communications (GSM) standard [7], where a handoff is made by comparing different RSS values from different BSs. In the existing algorithm, if MS is in the lower layer and the lower layer signal becomes weaker than a threshold, handoff is made to the upper layer, and if the MS is in the upper layers and a stronger signal is detected in lower layer, handoff is made to lower layer. These two thresholds are both set to -45dB, which is the approximate carrier power at cell boundary in our simulation environment. For cells inside a layer, handoff is made to the BS which has the strongest signal plus a hysteresis value. This handoff is called horizontal handoff, and hysteresis is set to 15dB for both layers.

In the simulation environment, first the variables are defined, then PNN is trained with signal samples collected in the specified path. In online mode, at each iteration, the user location is updated and the new RSS matrix is calculated. If there are enough samples to construct the pattern vector, the latter is made and classified. If the tolerance factor is also

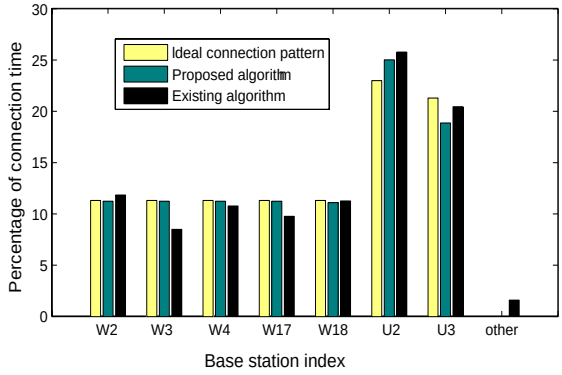

Fig. 7. Percentage of connection times for each BS in Path 1.

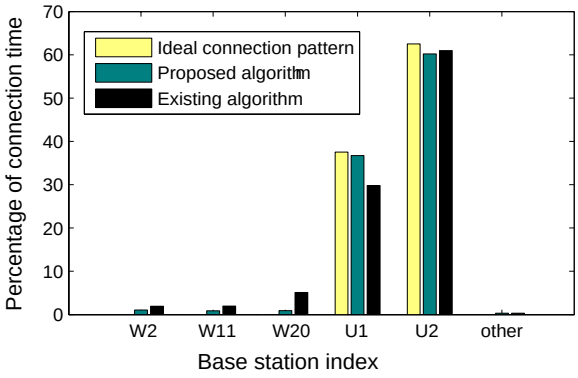

Fig. 9. Percentage of connection time for each BS in Path 2.

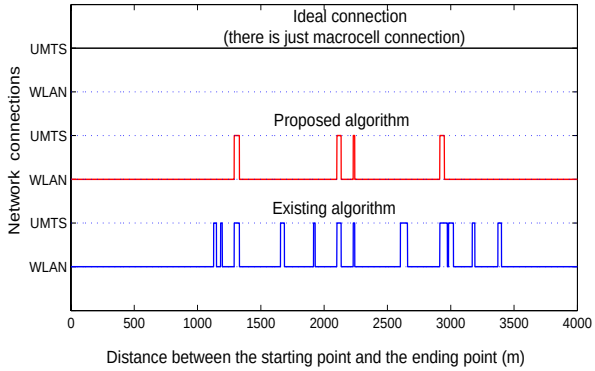

Fig. 8. Network connections for Path 2.

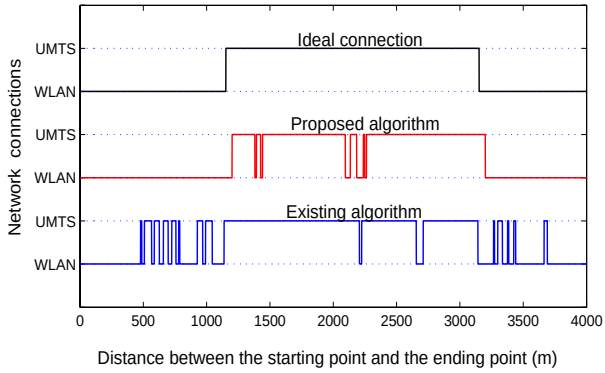

Fig. 10. Network connections for Path 3.

correct, the algorithm enters into the proposed VHO algorithm mode and decides about handoff according to cell boundaries defined with path blocks. If there are not enough samples or that the classified pattern is not correct, the program continues with the existing VHO algorithm.

IV. Simulation Results

A. Evaluation of the Proposed VHO Algorithm

We present simulation results pertaining to the optimization methods for the pattern classification proposed in section III.C along with the algorithm proposed in section II.A, within the simulation environment elaborated in section III.

Fig. 5 shows the effect of using different combinations of proposed optimization techniques on pattern detection error probability (DEP) versus averaging window size. Totally, 6 combinations have been considered and the pattern vector sequence is [WLAN (2 4 18)]. It is observed that the proposed optimization tools have a great effect on reducing the error probability, and their simultaneous use reduces the error probability to zero for averaging windows higher than 3.

We evaluate the performance of our VHO algorithm for four paths illustrated in Fig. 4. In Path 1, the emphasis is on maximizing WLAN or lower layer usage in order to maximize the received bandwidth and minimize the connection costs. For this path, we used signals for WLANs number 2, 4 and 18 to make the pattern vector (in some figures we have replaced WLAN with W and UMTS with U for more clarity). To reduce channel interference and increase QoS, handoff points are defined on cells (horizontal handoff) and layers (vertical handoff) boundaries.

In Fig. 6, networks connection for Path 1 is shown. The first plot shows the ideal connection for each of 4000 points in the path. The second plot shows the performance of the proposed algorithm, and the third one shows the performance of the existing algorithm. It is obvious that in the proposed algorithm unnecessary handoffs and ping-pong effect are reduced significantly. Similarly, in the proposed algorithm, crossover distance is much less than the existing algorithm and almost identical to the ideal connection pattern.

Fig. 7 shows the percentage of connection times for each BS in Path 1. For ideal planned connection scheme, there are only connections to WLANs 2, 3, 4, 17 and 18 (with the same connection time for each) and UMTS cells number 2 and 3. Comparing bars in Fig. 7, we observe that the proposed algorithm connection time is much similar to the ideal case and that it does not transfer unwanted connections to other BSs. Here we define a metric called, WLAN usage factor. It shows the percentage of time, during which a MS is connected to WLAN (for Path 1, a higher value is desirable). For the proposed algorithm, this factor is 56.1% and for the existing algorithm it is 53.4%, so we have 2.7% improvement. Unnecessary handoffs for Path 1 are 30 times for the existing

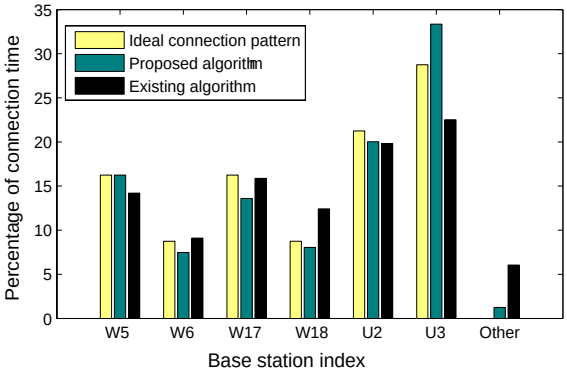

Fig. 11. Percentage of connection time for each BS in Path 3.

algorithm and 3 times only for the proposed algorithm; this also shows 90% reduction in ping-pong effect. If plots in Fig. 6 are compared, we note that the layer crossover distance is reduced 4 times.

Figs. 8 to 11 show similar graphs for Path 2 and Path 3. In Path 2, for ideal connection, no VHOs are considered because the status will not be stable. WLAN 11 and UMTSs 1 and 2 are used to build the pattern vector in Path 2. For this path, there was 66% reduction in horizontal handoffs and 70% reduction in ping-pong VHOs. Path 3 is similar to Path 1, however it is interesting for studying the effect of horizontal handoffs between WLANs. Cells considered for forming the patter vector are WLAN 2, 17 and UMTS 3 for this path and reduction values for horizontal handoff and VHO for this path are 100% and 40%, respectively. Path 4 is interesting to take just horizontal handoffs into consideration, which results in 76% reduction in ping-pong effect. Because Path 4 is far enough from WLANs, the number of VHOs is zero in this case.

V. Conclusion

A new location-based vertical handoff algorithm (VHO) has been proposed from physical layer point of view for wireless heterogeneous networks. This algorithm uses signal samples of different wireless technologies in the coverage area to construct a patter vector. Later, this vector is used for localizing the MS position by means of a probabilistic neural network. Location information is then applied to the VHO, for more exact decisions on the time to trigger the handoff procedure. Moreover, we proposed procedures for upward and downward vertical handoff between WLAN and UMTS from higher network layers point of view. In our simulations, we studied several paths in a hierarchial structure of UMTS and WLAN cells. For maximizing the user received bandwidth, the best strategy that can be defined is maximizing WLAN usage. Minimizing ping-pong handoffs between layers was another goal of this work. Our algorithm shows significant improvement over existing techniques, especially that we have designed it for handoff procedures in heterogeneous networks.

References

[1] E. Stevens-Navarro, Y. Lin and V. W. S. Wong, "An MDP-based vertical handoff decision algorithm for heterogeneous wireless networks," *IEEE Trans. on Vehicular Technology*, vol. 57, no. 2, pp.1243–1254 , Mar. 2008.

[2] I. Stojmenovic, Handbook of wireless networks and mobile computing, Wiley, 2002.

[3] K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J. P. Makela, R. Pichna and J. Vallstron, "Handoff in hybrid mobile data networks," *IEEE Personal Commun.*, vol. 7, no. 2, pp. 34–47, Apr. 2000.

[4] P. Bellavista, M. Cinque, D. Cotroneo, L. Foschini, "Self-adaptive handoff management for mobile streaming continuity," *IEEE Trans. on network and service management*, vol. 6, no. 2, pp. 80–94, June 2009.

[5] A. Hasib and A. O. Fapojuwo, "A mobility model for heterogeneous wireless networks," in *Proc. IEEE Radio and Wireless Symp.*, Jan. 2008, pp. 815–818 .

[6] M. R. Kibria, A. Jamalipour and V. Mirchandani, "A location aware three-step vertical handoff scheme for 4G/B3G networks," in *Proc. Global Telecommun. Conf. (Globecom'05)*, vol. 5, Nov. 2005, pp. 2752–2756.

[7] J. Korhonen, Introduction to 3G Mobile Communications, Artech House Publications, 2003.

[8] janam W. T. Chen, J. C. Liu and H. K. Huang, "An adaptive scheme for vertical handoff in wireless overlay networks," in *Proc. 10th International Conf. on Parallel and Distributed Systems*, July 2004, pp. 541–548.

[9] C. W. Lee, L. M. Chen, M. C. Chen and Y. S. Sun, "A framework of handoffs in wireless overlay networks based on mobile IPv6," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 11, pp. 2118–2128, Nov. 2005.

[10] R. Chakravorty, P. Vidales, K. Subramanian, I. Pratt, J. Crowcroft, "Performance issues with vertical handovers - experiences from GPRS cellular and WLAN hot-spots integration," in *Proc. IEEE Annual Conf. on Pervasive Computing and Commun.*, Mar. 2004, pp. 155–164.

[11] F. Zhu and J. McNair, "Optimizations for vertical handoff decision algorithms," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC'04)*, vol. 2, Mar. 2004, pp. 867–872.

[12] M. Ylianttila, J. Makela and K. Pahlavan, "Geolocation information and inter- technology handoff," in *Proc. IEEE International Conf. on Commun. (ICC'00)*, vol. 3, June 2000, pp. 1573–1577.

[13] F. Hu and N. K. Sharma, "priority determinded multiclass handoff scheme with guaranteed mobile QoS in wireless multimedia networks," *IEEE Trans. on Vehicular Technology*, vol. 53, no. 1, pp. 118–135, Jan. 2004.

[14] Y. H. Wang, H. M. Huang and C. P. Hsu, "Handoff strategy for Multi-tier IP-based wireless network," in *Proc. 17th International Conf. on Advanced Information Networking and Applications*, Mar. 2003, pp. 790–793.

[15] K. D. Wong and D. C. Cox, "A pattern recognition system for handoff algorithms," *IEEE Journal on Selected Areas in Commun.*, vol. 18, no. 7, pp. 1301–1312, July 2000.

[16] J. Makela, M. Ylianttila and K. Pahlavan, "Handoff decision in multi-service networks," in *Proc. IEEE International Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC'00)*, vol. 1, Sep. 2000, pp. 655–659.

[17] S. Kyriazakos, D. Drakoulis and G. Karetsos, "Optimization of the handover algorithm based on the position of the mobile terminal," in *Proc. Symp. on Commun. and Vehicular Technology*, Oct. 2000, pp. 155–159.

[18] R. Narasimhan and D. C. Cox, "A handoff algorithm for wireless systems using pattern recognition," in *Proc. IEEE International Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC'98)*, Sep. 1998, pp. 335–339.

[19] A. Mehbodniya and J. Chitizadeh, "An intelligent vertical handoff algorithm for next generation wireless networks," in *Proc. IFIP International Conf. on Wireless and Optical Commun. Networks, 2005 (WOCN'05)*, Mar. 2005, pp. 244–249.

[20] C. H. Lee and C. J. Yu, "An intelligent handoff algorithm for wireless communication systems using grey prediction and fuzzy decision system," in *Proc. IEEE International Conf. on Networking, Sensing and Control*, vol. 1, Mar. 2004, pp. 541–546.

[21] 3GPP TS 25.331, 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Radio Resource Control (RRC); Protocol Specification, V3.11.0, 2002.

[22] J. Zander and S.L. Kim, Radio resource management for wireless networks, Artech House, 2001.

Abolfazl Mehbodniya received his Bachelor's degree and his Master's degree in electrical engineering from Ferdowsi University of Mashhad, Iran in 2002 and 2005, respectively. He is now working towards his Ph.D. degree at the National Institute of Scientific Research-Energy, Materials, and Telecommunications (INRS-EMT), University of Quebec, Montreal, QC, Canada.

His research interests are in wireless communications, ultra wideband technology, radio resource management and ad hoc networks.

Sonia Aïssa received her Ph.D. degree in Electrical and Computer Engineering from McGill University, Montreal, QC, Canada, in 1998. Since then, she has been with the National Institute of Scientific Research-Energy, Materials, and Telecommunications (INRS-EMT), University of Quebec, Montreal, QC, Canada, where she is a Professor.

¿From 1996 to 1997, she was a Researcher with the Department of Electronics and Communications of Kyoto University, Kyoto, Japan, and with the Wireless Systems Laboratories of NTT, Kanagawa, Japan. ¿From 1998 to 2000, she was a Research Associate at INRS-EMT, Montreal. From 2000 to 2002, while she was an Assistant Professor, she was a Principal Investigator in the major program of personal and mobile communications of the Canadian Institute for Telecommunications Research (CITR), leading research in resource management for code division multiple access systems. From 2004 to 2007, she was an Adjunct Professor with Concordia University, Montreal. In 2006, she was Visiting Invited Professor with the Graduate School of Informatics, Kyoto University, Japan. Her research interests lie in the area of wireless and mobile communications, and include radio resource management, performance evaluation, design and analysis of multiple antenna (MIMO) systems, and cross-layer design and optimization, with a focus on cellular, ad hoc, and cognitive radio networks.

Dr. Aïssa was the Founding Chair of the Montreal Chapter IEEE Women in Engineering Society in 2004-2007, a Technical Program Cochair for the Wireless Communications Symposium (WCS) of the 2006 IEEE International Conference on Communications (ICC 2006), and PHY/MAC Program Chair for the 2007 IEEE Wireless Communications and Networking Conference (WCNC 2007). She was also the Technical Program Leading Chair for the WCS of the IEEE ICC 2009, and is currently serving as Cochair for the WCS of the IEEE ICC 2011. She has served as a Guest Editor of the *EURASIP journal on Wireless Communications and Networking* in 2006, and as Associate Editor of the IEEE WIRELESS COMMUNICATIONS MAGAZINE in 2006-2010. She is currently an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and of the IEEE COMMUNICATIONS MAGAZINE, and Associate Editor of the *Wiley Security and Communication Networks Journal*. Awards and distinctions to her credit include the Quebec Government FQRNT Strategic Fellowship for Professors-Researchers in 2001-2006; the INRS-EMT Performance Award in 2004 for outstanding achievements in research, teaching and service; the IEEE Communications Society Certificate of Appreciation in 2006 and 2009; and the Technical Community Service Award from the FQRNT Center for Advanced Systems and Technologies in Communications (SYTACom) in 2007. She is also co-recipient of Best Paper Awards from IEEE ISCC 2009 and IEEE WCNC 2010; and recipient of NSERC Discovery Accelerator Supplement Award.

Jalil Chitizadeh received the BSc degree in Electrical and Electronic Engineering from Napier University , Scotland in 1980, the MSc degree in communication systems from Essex University, England in 1985. He received the PhD degree and DIC degree in Communication Engineering from Imperial College of London University in 1989. He worked in Iran Telecommunications Research Center (ITRC)from 1981 to 1983. He is currently an associate professor and a project leader at Electrical Engineering department and Computer Communication Center of Ferdowsi University of Mashad, Iran. His research interests are in Cellular wireless networks, Mobile ad-hoc networks and high speed networks.

A Radio Resource Allocation Algorithm for QoS Provision in PMP-based Systems

Ping Wang
Broadband Wireless communications and Multimedia laboratory, Key Laboratory of Embedded System and Service Computing supported by Ministry of Education, Tongji University, Shanghai, China.
Email: pwang@tongji.edu.cn

Lijun Zu, Fuqiang Liu and Yiling Wang
Broadband Wireless communications and Multimedia laboratory, Key Laboratory of Embedded System and Service Computing supported by Ministry of Education, Tongji University, Shanghai, China.
Email: {ninhaozulijun, liufuqiang, zagigi}@163.com

***Abstract*—Based on the studies on downlink resource allocation in point to multi-point (PMP) mode in the 802.16e systems, an efficient downlink resource allocation algorithm with low complexity is proposed to maximize the system throughput, which takes the advantage of frequency-selective fading property in orthogonal frequency division multiple (OFDM) networks. The proposed algorithm not only provides the traffic rate as high as possible, but also satisfies the minimum reserved traffic rate for non-real-time services. The simulation shows that the proposed algorithm performs better in terms of assuring individual QoS and offering fairness among users at the cost of slight degradation in throughput.**

***Index Terms*—radio resource management, downlink resource allocation, QoS, PMP, 802.16e**

I. INTRODUCTION

It is well known that one of the most challenging technical problems for future wireless systems is to provide various services to meet the requirement of an increasing number of consumers while ensuring individual quality of service (QoS). Owing to the scarce frequency spectrum, resource allocation has been considered as one of the key technologies for increasing utilization rate of the limited power and spectrum in future wireless networks. According to the IEEE 802.16e standard, the medium access control (MAC) layer supports both point to multi-point (PMP) mode and mesh mode [1, 2], between which the former is the primary topology structure. The downlink transmission operates on a PMP basis, and data for MSs need to be transferred by a central BS via a wireless link. The PMP mode provides a comparatively high traffic rate, since network resources can be shared among users. As a result, broadband wireless access systems widely adopted the PMP mode in the last few years.

The MAC layer in IEEE 802.16e network uses the conception of service flow, which provides unidirectional packet transmission [1]. A service flow is characterized by a set of quality of service (QoS) parameters such as latency, jitter, and throughput assurances. Owing to the low traffic rate and high error rate in a wireless link, as well as user's mobility, a QoS-guaranteeing resource allocation algorithm becomes one of the key techniques in broadband wireless network. Consequently, to support various applications under limited radio resources and time-varying channel, a dynamic resource allocation, which can achieve both higher system spectral efficiency and better QoS, has been identified as one of the imperative tasks in wireless communication since a few years ago.

Letaief and Zhang provide an overview of recent research on dynamic resource allocation, especially for MIMO and orthogonal frequency division multiple (OFDM) systems [3]. They propose an algorithm to maximize data rate for a given power budget with a target BER. The proposed scheme in [4] adjusts the power to meet the predefined delay requirement, which is one of the decisive factors to guarantee end-to-end delay. Ref. [5] proposes an efficient subcarrier and power allocation algorithm, which formulates necessary conditions of downlink scheduling for proportional fairness in orthogonal frequency division multiple access (OFDMA) systems. However, since it considers power allocation after subcarriers' allocation for assuring individual QoS and subsequent proportional fairness (PF) allocation, it adds a lot of complexity. As we know, in a scenario with high SNR, the performance of allocating power evenly to each subcarrier is almost equivalent to that of the classic optimal power allocation algorithm, at the same time the complexity is reduced greatly. Ref. [6] utilizes a user and connection based scheme to improve transmission efficiency and guarantee quality of service. It schedules the user with the highest priority first and the allocation starts from the maximum deviation channel, with the result that the QoS requirement of the user with low priority may not be met. The target QoS of resource allocation in [7] corresponds to a minimum user data rate, a target bit-error rate and a maximum BER-outage

Manuscript received January 31, 2010; revised April 18, 2010; accepted May 11, 2010.

doi:10.4304/jcm.5.7.530-536

probability. Ref. [8] distributes subcarriers and bits among users based on their different quality of service requirements and traffic type. Though they provide an efficient resource allocation for the users with different traffic class, the system throughput may reduce greatly.

In order to meet the requirements of non-real-time polling service (nrtPS) in IEEE802.16e network, a downlink resource allocation algorithm is proposed to maximize the system throughput, which takes the advantage of frequency-selective fading property of OFDM networks. What's more, the proposed Delta algorithm combines the key QoS parameters of nrtPS and assures the required minimum data rate for each user. Under the circumstances of using non-real-time polling service in PMP-based OFDM networks, such items as subcarrier/bit allocation statistics, throughput and fairness index are compared and analyzed by using four different resource allocation algorithms. The four algorithms are Delta, MT [3], fixed modulation and code scheme (FMCS) algorithm based on 16 quadrature amplitude modulation (QAM) and maximum throughput confined by minimum traffic rate (MTMR) [9]. The simulation shows that though the proposed Delta algorithm sacrifices few throughput, it comes close to the best MT algorithm in terms of system throughput and the best FMCS algorithm in terms of fairness index. And Delta algorithm meets QoS demand of the minimum reserved traffic rate for all users at the same time. Besides, its complexity is reduced greatly, which makes it appropriate for application in non-real-time business.

The remainder of this paper is organized as follows. Section 2 briefly describes the basic consideration of an adaptive radio resource allocation algorithm based on OFDM system. In view of QoS characteristics of nrtPS defined in 802.16e protocol, a relevant adaptive allocation algorithm in the multi-carrier system is discussed and a downlink resource allocation algorithm to maximize the system throughput is introduced in section 3. Section 4 presents simulation result for the proposed Delta and other three algorithms. Finally, the paper is concluded in section 5.

II. Basic Considerations

Power is one of the important resources in wireless communication system. The system capacity directly depends on BS's transmitting power. In OFDM system, a scheduler needs to provide the power allocation of each subcarrier as well as subcarrier assignment. The classic optimal power allocation algorithm is called Water-filling theorem [10], of which the basic idea is allocating to the channels of good quality as more power as possible, the ones of bad quality relatively less. However, according to information theory, it is easily proved that allocating the power to each subcarrier on average can dramatically reduce the complexity of the resource allocation just at the cost of slight degradation in system capacity.

Assume that there are J independent parallel Gaussian channels, of which noise power are respectively $N_1 \ldots N_J$. According to Shannon's formula, the capacity of J channels is

$$C = \underset{P_i}{Max} \frac{1}{2} \sum_{j=1}^{J} \log_2 (1 + \frac{P_j}{N_j}) \tag{1}$$

where P_j denotes the signal power of channel J, and satisfies the equation of power limitation $\sum_{j=1}^{J} P_j = P$.

By the method of Lagrange multipliers, we can get the power allocation method which can maximize the system capacity:

$$P_j = (\gamma - N_j)^+ \tag{2}$$

where $(x)^+$ represents positive function, γ constant is to be determined by $\sum_{j=1}^{J} (\gamma - N_j)^+ = P$. Thus the whole capacity of parallel Gaussian channels equals to

$$C = \frac{1}{2} \sum_{j=1}^{J} \log_2 (1 + \frac{(\gamma - N_j)^+}{N_j}) \tag{3}$$

From (2), the less the noise power of each channel is, the more power each channel can obtain and vice versa.

In order to simplify the power allocation scheme, we assume that the power is high enough. Hence, $\sum_{j=1}^{J} (\gamma - N_j)^+ = P$ retrogresses to:

$$\sum_{j=1}^{J} (\gamma - N_j) = P \tag{4}$$

Then it is easy to get the equation:

$$\gamma = \frac{1}{J} (P + \sum_{j=1}^{J} N_j) \approx \frac{P}{J} \tag{5}$$

Substituting (5) into (3) yields the channel capacity:

$$C \approx \frac{1}{2} \sum_{j=1}^{J} \log_2 (\frac{P}{J \cdot N_j}) \tag{6}$$

From (6), the channel capacities obtained from average power allocation algorithm and water-filling power allocation algorithm are roughly equivalent at high SNRs. As a result, in order to simplify the algorithm, it is generally regarded that the channel capacity with the average allocation algorithm approximately equals to the best allocation result at high SNRs. In OFDM system, we can allocate the power to each subcarrier on average, and then consider the assignment of carriers and bits together, which can reduce the complexity of adaptive allocation algorithm to a large extent.

III. Resource Allocation Algorithm Based On Maximum Throughput

A. Algorithm Description

Generally speaking, the target of an algorithm in multi-carrier system is maximum throughput, or best fairness index, or tradeoff between them. This paper focuses on maximizing system throughput and guaranteeing the minimum reserved traffic rate of non-real-time polling service, and only downlink OFDM system is considered. Assume that the number of users served by a BS in a cell

is K, and the number of subcarriers available is N. At transmitting terminal, data from K users will be mapped to subcarriers and bits, i.e., different subcarriers with corresponding bits will be allocated to different users. An adaptive algorithm can allocate resource for each subcarrier at transmitting terminal on condition that channel state is always given. We define $\rho_{k,n}$ as the occupying identifier of a subcarrier, whose value is 1 when the nth subcarrier is occupied by the kth user, otherwise $\rho_{k,n}$ equals 0. Besides, $c_{k,n}$ signifies the number of bits available for the kth user on the nth subcarrier.

After inverse fast Fourier transform and adding cyclic prefix, signal is transmitted through the downlink. Generally, when the duration of cyclic prefix is larger than maximum multipath delay spread, inter signal interference can be eliminated and then every subcarrier is supposed to be experiencing flat fading in the channel. It is assumed that channel state information is acquired through dedicated control channel. With cyclic prefix removed and fast Fourier transform, users can obtain their own data information from relevant subcarrier, according to subcarrier allocation and modulation information.

When a system has a target bit error rate (BER), there exists a relationship between BER and $c_{k,n}$ as follows:

$$c_{k,n} = f(BER, p_{k,n}, h_{k,n}) \tag{7}$$

where $p_{k,n}$ is the transmitting power needed when allocating the nth subcarrier to the kth user, $h_{k,n}$ is the magnitude of channel gain, and function f indicates a mapping relationship.

For every OFDM subcarrier, there is:

$$T = \sum_{k=1}^{K}\sum_{n=1}^{N} \rho_{k,n} c_{k,n} \tag{8}$$

For user k, the number of bits transmitted on this subcarrier is

$$R_k = \sum_{n=1}^{N} \rho_{k,n} c_{k,n} \tag{9}$$

Based on the analysis mentioned above, the problem that the proposed algorithm needs to solve can be concluded as:

$$V = \underset{c_{k,n},\rho_{k,n}}{Max} \sum_{k=1}^{K}\sum_{n=1}^{N} \rho_{k,n} c_{k,n} \tag{10}$$

while it should also submit to the following restraint:

$$\sum_{n=1}^{N} \rho_{k,n} c_{k,n} \geq r_k \quad \forall k \tag{11}$$

$$\text{If } \rho_{k,n} = 1, \text{ then } \rho_{k',n} = 0 \quad \forall k \neq k' \tag{12}$$

$c_{k,n}$ can be obtained from (10), and the minimum traffic rate for every user is confined in (11). From (12), we can see that a subcarrier can be occupied by at most one user. The target of the proposed algorithm is to obtain proper $\rho_{k,n}$ and $c_{k,n}$. Since $c_{k,n}$ is an integer and has non-linear relationship with $\rho_{k,n}$, the resolution of (10) belongs to non-linear problem. Here, $\rho_{k,n}$ is supposed to be the average value of transmitting power, i.e., power is distributed evenly, as a result that calculation complexity is reduced and sub-optimal solution is achieved at the same time.

When $\rho_{k,n}$ is fixed, $c_{k,n}$ can be solved from (7). Firstly, we ignore the restraint of (11), and take only the target of maximum throughput into consideration. According to (10), subcarriers are only allocated to users who can provide maximum throughput to the system. Then, bit allocation is adjusted according to the minimum reserved rate of each user. The subcarriers occupied by QoS satisfied users needs to be allocated to those unsatisfied ones till QoS requirements of all users are satisfied. However, modulation order of some subcarriers must be set lower to satisfy QoS requirements of all users during the re-adjustment processing of subcarriers. What's more, adjustment of subcarriers in this algorithm is supposed to obey the following rule, that is, always choosing the one that least affects the system throughput.

$$Delta_{k,n} = \frac{c_{k_n^*,n} - c_{k,n}}{c_{k,n}} \quad \forall n \tag{13}$$

where $Delta_{k,n}$ is the criteria of judgment when adjusting bits, and $c_{k_n^*,n}$ indicates the traffic rate available when subcarrier n is initially allocated to user k^*. The flow chart of this algorithm is illustrated in figure 1.

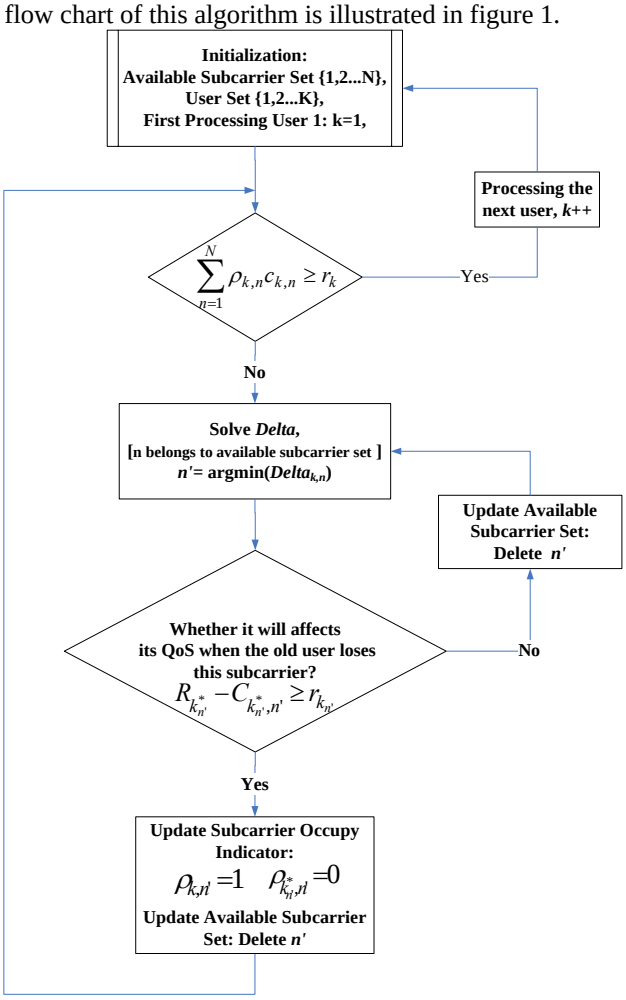

Figure 1. Flow chart of the Delta algorithm

B. Description of Comparison Algorithms

Here the three traditional resource allocation algorithms used for comparison with the proposed algorithm are introduced.

Solution 1: Every subcarrier utilizes fixed modulation mode, that is, 16QAM. Ideally, each subcarrier carries 4 bits, thus an OFDM signal carries 4*N* bits all together. We call this FMCS algorithm.

Solution 2: Adaptive resource allocation targeted at maximum throughput, which has the same result as initial allocation of the proposed algorithm, is reputed as MT algorithm [3].

Solution 3: The allocation algorithm targeted at maximum throughput, token involved and restrained by minimum rate, with precedence judgment criteria is simplified, is called as MTMR algorithm [9]. It can be represented as follows: at time slot *t*, allocating subcarrier *k* to user k_n^*.

$$k_n^* = \arg\max_k (T_k^{n-1} \cdot c_{k,n}(t)) \quad n = 1,2,...N \qquad (14)$$

Every time a subcarrier is allocated, users' demand of minimum rate and the number of tokens should be updated. For example, when subcarrier *n* is allocated, the number of tokens for user *k* is updated as follows:

$$T_k(t) = \max\{0, T_k(t-1) + r_k - c_k(t)\} \qquad (15)$$

where $c_k(t) = \sum_{n=1}^{N} c_{k,n}(t)$ represents the number of bits which are already allocated to user *k* at time slot *t*, $T_k(t)$ represents the service demand of the user, and $c_{k,n}(t)$ is the real transmission capability of subcarrier *n* allocated to user *k*.

Ⅳ. Simulation And Analysis

In order to evaluate the performance of the proposed allocation algorithm, the simulation scenario mainly focuses on non-real-time polling service (nrtPS). The proposed Delta algorithm will be compared with other three algorithms (i.e., FMCS, MT and MTMR) in fairness index among users, allocation of subcarriers and bits, and system throughput.

A. Simulation Environment

TABLE I.
PARAMETERS IN SIMULATION

Parameter	Value
The number of users (K)	6 ~ 22 Users
The number of subcarriers (N)	256
Modulation mode allowed	BPSK / QPSK / 16QAM / 64QAM
Target BER (BER)	10e-4
Service model	nrtPS
QoS property: Min traffic rate (bits/symbol)	Uniform distribution over the interval (65, 90)

This simulation scenario contains 256 subcarriers for data transmission and supports four modulation techniques including BPSK, QPSK, 16QAM and 64QAM, where the highest modulation order is six. In simulation, the radius of a cell is 1000m, the maximal transmitting power of a BS is 43dbm and the system target BER is set to 10^{-4}. For the service model, assume that each user has only one nrtPS connection, of whose packet the size ranges from 65bit/symbol to 90bit/symbol. Table 1 shows the parameters in detail.

B. Results and Discussion

At first the fairness index among users of the four algorithms (Delta, MT, FMCS and MTMR) is analyzed. Here the Min/Max fairness index is used to measure the fairness of different algorithms. Its definition is as follows:

$$F = \frac{\min\{s_i\}}{\max\{s_i\}} \qquad (16)$$

where F represents the fairness index, S_i denotes the throughput achieved by user *i*.

Figure 2 illustrates the relationship between Min/Max fairness index and the number of users by using the four different algorithms.

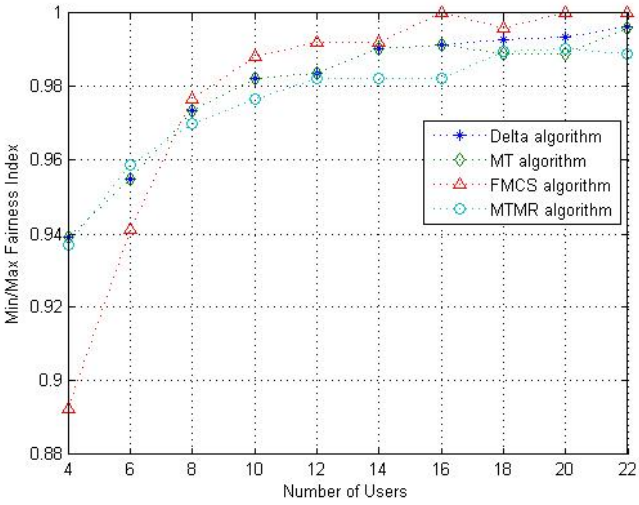

Figure 2. Relationship between Min/Max fairness index and the number of users

In simulation, the number of users changes from 4 to 22. As shown in figure 2, with the number of users in a cell increasing, the Min/Max fairness index of the four algorithms all trends to rising. Especially, for FMCS algorithm, when the user number goes up to a certain value (*K*=16 in this simulation), the fairness index reaches 1. This is because not only it uses a fixed modulation, but also its resource allocation is similar to the polling of packet scheduling algorithm. Thus the fairness index is supposed to be the highest. The fairness of the other three ascends steadily. Comparatively, the fairness index of MTMR algorithm is slightly lower than Delta and MT algorithms, by 1% under the worst circumstances (*K*=14-16). The proposed algorithm, Delta algorithm, is rather equivalent to MT algorithm in system fairness, but when the user number is rising (*K*>=18), Delta algorithm shows superior performance.

Through several simulations, the statistics indicate that Delta algorithm and MT algorithm perform approximately the same in scenarios when users are relatively few. This is because every user in this scenario can be met QoS demand of minimum traffic rate, that is to say, there are always enough subcarriers to allocate and thus the adjustment of bits' allocation for QoS provision is unnecessary. However, Delta algorithm reaches a better result for bit allocation when the number of users increases to a certain amount. Figure 3 illustrates the number of bits per symbol allocated to each user in a scenario of 10 users. The results of Delta algorithm and MT algorithm hardly differ from each other. What's more, both of them meet QoS demand of minimum traffic rate. Theoretically, if QoS demand of each user differs, FMCS algorithm, which focuses on fairness index among users, will ignore the restraint of QoS provision. Therefore, when adopting FMCS algorithm, the traffic rates of user 4 to user 9 are a little bit lower than the required minimum traffic rate. The statistics acquired by using MTMR algorithm fluctuate largely. For example, the traffic rates of user 1 and user 4 are comparatively high, while those of user 3 and 10 are too low to satisfy the requirement of minimum traffic rate. The reasons why the result of subcarriers' allocation differs so much are mutual influence during the allocation of adjacent subcarriers, as well as the update of the number of tokens.

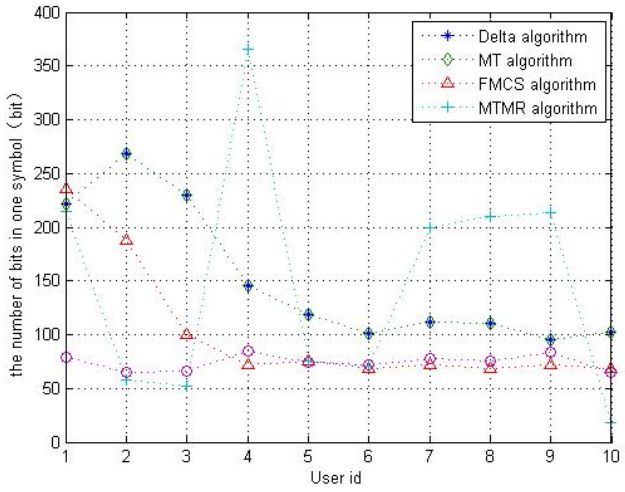

Figure 3. Number of bits per signal allocated to users in 10 users' scenario

Figure 4 shows the number of bits per symbol allocated to each user in 15 users' scenario. The result of MTMR algorithm still fluctuates largely. When the user number is increasing, there are some users experiencing "hungry" by using whether MTMR, FMCS, or MT algorithm, whose target is to maximize the system throughput. While Delta algorithm, an improvement of MT algorithm, allocates fewer bits to certain users than MT algorithm, it satisfies QoS requirements of minimum traffic rate of all users. In addition, Delta algorithm has generally achieved as high traffic rate as possible, this is because Delta algorithm is designed to adjust allocated subcarrier appropriately to meet the requirements of the users who have not obtained the minimum reserved traffic rate on the premise of maximizing the system throughput.

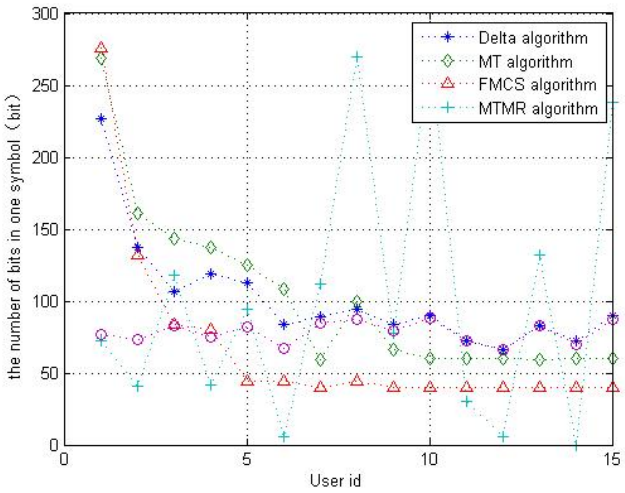

Figure 4. Number of bits per signal allocated to users in 15 users' scenario

Figure 5 shows system throughput in scenario of 10 users. As FMCS algorithm requires special modulation mode, it gets a relatively small throughput, approximately 1020 bit/symbol. Figure 5 also proves the conclusion mentioned before, that the performances of Delta algorithm and MT algorithm are quite the same when the number of users in the scenario is small. However, the throughput of MTMR is slightly smaller than that of Delta algorithm.

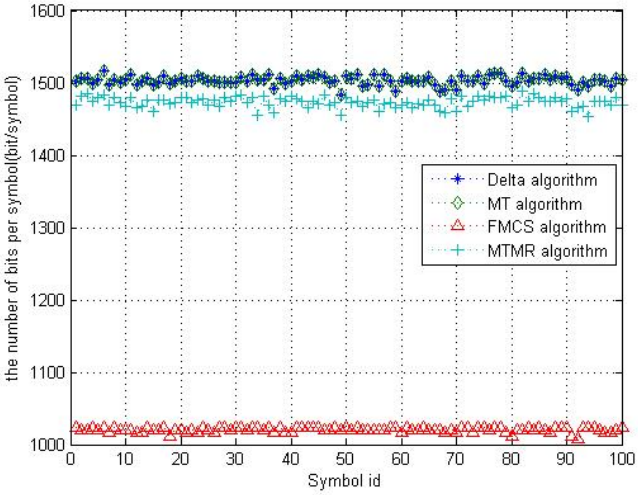

Figure 5. System throughput in 10 users' scenario

Figure 6 indicates system throughput in scenario of 15 users. The tendency is nearly the same as that of 10 users' scenario. Note that the performances of Delta algorithm and MT algorithm are both better than that of MTMR algorithm.

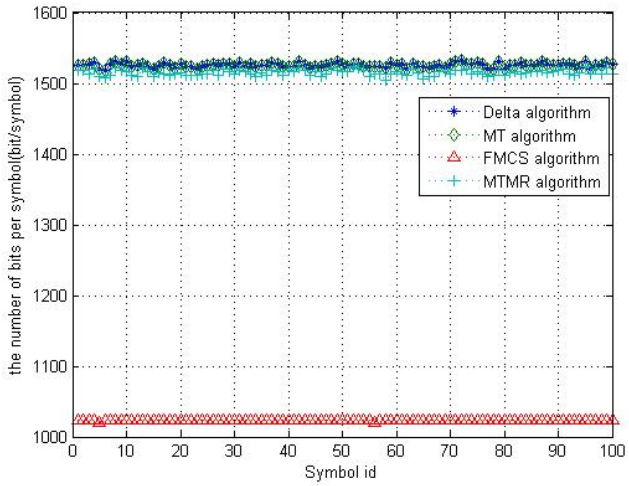

Figure 6. System throughput in 15 users' scenario

V. CONCLUSION

Based on MT algorithm, this paper proposes a new resource allocation algorithm with low complexity, which not only provides the maximum traffic rate as much as possible, but also guarantees the minimum reserved traffic rate for non-real-time services. Through simulation, the proposed Delta algorithm is compared with MT, FMCS and MTMR algorithms in fairness index among users, allocation of subcarriers and bits, and system throughput. The simulation indicates that Delta algorithm, as a promotion of MT algorithm, achieves QoS provision at the cost of slight degradation in throughput, at the same time does well in fairness index among users too. Particularly, in scenario with a relatively large number of users (above 10 in this simulation), Delta algorithm exceeds over the other three in comparison.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under grant (No. 70631002), the National Science and Technology Major Project of China under Grant (No. 2009ZX03003-005 and No. 2008BAH30B10) and program for Young Excellent Talents in Tongji University (No. 0800219099).

REFERENCES

[1] IEEE Std. 802.16-2004. "IEEE Standard for Local and Metropolitan Area Networks-Part 16: Air Interface for Fixed Broadband Wireless Access Systems," Oct. 2004.

[2] IEEE Std. 802.16e-2005. "IEEE Standard for Local and Metropolitan Area Networks-Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems," Feb. 2006.

[3] K. B. Letaief, Y. J. Zhang, "Dynamic multiuser resource allocation and adaptation for wireless systems," *IEEE Transactions on Wireless Communications*, Vol. 13(4), pp. 38-47, Aug. 2006.

[4] G. X. Xu, Y. Ji, J. H. Zhang and P. Zhang, "Adaptive OFDMA subcarrier assignment for QoS guaranteed services," *IEEE 61st Vehicular Technology Conference (VTC 2005-Spring)*, Vol. 3, pp. 1817-1820, May 2005.

[5] T. D. Nguyen and Y. Han, "A Proportional Fairness Algorithm with QoS Provision in Downlink OFDMA Systems," *IEEE Communications Letters*, Vol. 10(11), pp. 760-762, Nov. 2006.

[6] X. Zhu, J. Huo, X. Xu, C. Xu, W. Ding, "QoS-Guaranteed Scheduling and Resource Allocation Algorithm for IEEE 802.16 OFDMA System," *IEEE International Conference on Communications (ICC08)*, pp. 3463-3468, May 2008.

[7] A. Alsawah and I. Fijalkow, "Practical Radio Link Resource Allocation for Fair QoS-Provision on OFDMA Downlink with Partial Channel-State Information," *EURASIP Journal on Advances in Signal Processing*, Vol. 2009, pp. 1-16, 2009.

[8] Y. Lu, T. Luo, C. Yin and G. Yue, "Adaptive radio resource allocation for multiple traffic OFDMA broadband wireless access system," *The Journal of China Universities of Posts and Telecommunications*, Vol. 13(4), pp. 1-6, December 2006.

[9] G. X. Xu, "A study on resource allocation strategy in the next generation of mobile communications," *Doctor thesis of Beijing University of Posts and Telecommunications*, pp. 57-86, 2004.

[10] S. T. Chung and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: A unified view," *IEEE Transactions on Communications*, Vol. 49, pp. 1561-1571, Sept. 2001.

Ping Wang, born in China, 1978-2-28. He graduated from the department of computer science and engineering at Shanghai Jiaotong University, China and received Ph. D. degree in 2007. His major field of study is wireless communication.
He joined the college of electronic and information engineering at Tongji University in 2007 and now is a lecturer. He has published several papers in Computer Communications, Wireless Personal Communications, IEICE Transactions on Information and Systems, IEICE Transactions on Communications, etc. His current and previous interests include routing algorithms and resource management in wireless networks, vehicular ad hoc network and video transcoding.

Lijun Zu, born in China, 1986-8-5. He graduated from the department of information and communication engineering at Tongji University and received B.S. degree in 2008. His major field of study is wireless communication. Now he is is a graduate in the department of information and communication engineering at Tongji University. His main research interests are in mobility modeling and radio resource management for the next generation mobile communications.

Fuqiang Liu, born in China, 1963-3-7. He graduated from the department of automation at China University of Mining and received Ph. D. degree in 1996. His major field of study is signal processing. Now he is a professor in the department of information and communication engineering at Tongji University. His main research interests are in routing algorithms in wireless broadband access and image manipulation.

Yiling Wang, born in China, 1984-7-14. She graduated from the department of information and communication engineering at Tongji University and received B.S. degree in 2006. His major field of study is signal processing. Now she is a graduate in the department of information and communication engineering at Tongji University. Her main research interests are in cross-layer resource allocation in OFDM systems.

Adaptive Selection of Spreading Code Subsets from Orthogonal Binary Code Sets for Reduced PAPR in MC-CDMA Systems

Sabbir Ahmed† and Makoto Kawai*
Graduate School of Science and Engineering, Ritsumeikan University, Kusatsu City, Japan
Email†: gr045054@ed.ritsumei.ac.jp
Email*: kawai@is.ritsumei.ac.jp

***Abstract*— One of the major concerns with multicarrier CDMA (MC-CDMA) systems is the high peak to average power ratio (PAPR) that can lead to degraded transmission power efficiency. Based on the well known fact that suitable allocation of user spreading codes can be used as a PAPR reduction tool, we in this contribution, investigate the PAPR property of a recently proposed spreading code set called "Orthogonal Binary User (OBU) Codes". Considering different levels of active user densities, we present code allocation table consisting of selected OBU code combinations yielding low PAPR. On the basis of analytical and simulation results, we show that the presented code allocation table is capable of making the system perform better than the well known Walsh-Hadamard codes from both PAPR and BER perspectives.**

***Index Terms*— PAPR, MC-CDMA, Spreading codes, Orthogonal binary user codes, Walsh-Hadamard codes, BER.**

I. INTRODUCTION

Recently, the demand for high-speed wireless multimedia services is growing very rapidly and hence different advanced multiple access technologies are drawing significant attention. Multicarrier code division multiple access or MC-CDMA is one such multiple access method which is an amalgamation of orthogonal frequency division multiplexing (OFDM) and code division multiple access (CDMA) techniques. Considering its capability of offering the combined features of OFDM and CDMA, MC-CDMA appears to be a strong candidate as a multiple access method for future generation wireless communication systems.

But one problem of implementing multicarrier based systems is the large value of peak to average power ratio (PAPR). Since transmitter power amplifiers are often operated near the saturation region, occurrence of high peaks in power envelop of the transmitted signal can lead to severe BER degradation due to non-linear amplification. On the other hand, if amplifiers are operated in the linear region their power efficiency is degraded. Hence, it is desirable that the transmitted signal possesses reduced peaks and in order to achieve this objective, researchers have suggested methods like signal clipping, selected mapping, partial transmit sequences etc. [1-3].

In this context, MC-CDMA systems offer one additional degree of freedom over OFDM systems, i.e., the choice of spreading codes. It has been shown before that the transmitted signal amplitude of MC-CDMA systems is closely related to the collective non-periodic auto-correlation and cross-correlation values of the underlying spreading codes and hence the selection process of spreading codes itself can be used as a tool to characterize the value of PAPR [4, 5].

A survey of related literature reveals that investigations concerning the PAPR issues of MC-CDMA systems, especially with relation to spreading codes, are being carried out mainly from two perspectives. Firstly, studying comparative crest factor (CF) or PAPR of different orthogonal and non-orthogonal spreading code sequences like Walsh-Hadamard (WH), Golay Complementary, Orthogonal Gold and Zadoff-Chu [4-7] and exploring adaptive usage of codes belonging to different code family [7]. And secondly, searching for low PAPR producing codes by investigating different code allocation strategies through the selection of different combinations of codes from a specific code set [5, 8, 9]. All these schemes show peak reduction capability but necessitate some trade off also. For example, applying the concept of adaptive usage of WH and Golay Complementary codes, [7] had proposed a scheme that proved efficient from peak reduction perspective but required transmission of significant amount of side information causing possible negative effect on the system throughput. Again, based on the observation that WH codes show lowest PAPR for systems working at full load capacity, recently [13] proposed a scheme where a system is made to work always at full load by artificially introducing data

Manuscript received December 30, 2009; revised April 7, 2010; accepted May 11, 2010. Part of this paper was presented at the 18th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '07).

doi:10.4304/jcm.5.7.537-544

symbols from inactive users. This system shows acceptable PAPR reduction, but it is also obvious that it will suffer from high multi-access interference (MAI) and will require high transmission power. Another recent work explores the advantage of different combinations of WH codes through cyclic shift of code sequences for every transmitted symbol [14]. But the introduced mechanism only looks for certain spreading code combinations ignoring a quite large number of other possible combinations. Apart from the computational complexity, it also requires transmission of side information on every occasion a user symbol changes. Thus for higher order modulation the system complexity will be very high. On the other hand, [5, 9] proposed WH sequence based code allocation table for reduced PAPR where creating the code allocation table is a one time activity which is performed before the system goes into operation and thus the requirement of side information is much less compared to [7, 14].

In light of the issues discussed so far, we were motivated to explore the effects of new spreading codes in reducing the PAPR problem of MC-CDMA; a potential future multiple access systems capable of very high data rates. In this endeavour, here in this study our objective is to consider a recently proposed orthogonal set of binary code sequences called orthogonal binary user (OBU) codes [10]. Re-organizing the structure of our previous work [15] based on the OBU code set, here at first we demonstrate the construction of code allocation table as was done in [5, 9]. Then we propose a different variant of code allocation strategy with less computational complexity yet near about same performance. We analyze both single and multiuser scenario, compare the peak property of OBU code with that of WH codes both analytically and with simulation. Finally, we show that codes from our constructed code allocation table perform better compared to corresponding WH codes from both PAPR and BER perspectives.

II. System Description

The transmitter block diagram of a downlink MC-CDMA system is shown in Fig. 1 [7, 9]. Here, $d^{(k)} = \left[d_1^{(k)},\ d_2^{(k)},...,\ d_M^{(k)}\right]$ denotes M modulated data symbols of the kth user, $k = 1, 2,...., K$. These serial data symbols are at first converted into M parallel symbols. As a result, the symbol rate is reduced by a factor of M. After this serial-to-parallel conversion, each symbol is frequency domain spread by a user spreading code $C^{(k)} = \left[c_1^{(k)},\ c_2^{(k)},...,\ c_L^{(k)}\right]$ where L denotes the length of the spreading code. It means every symbol on each of the M parallel paths is copied L times and then multiplied by each chip of the spreading code simultaneously. In the next stage, frequency domain spread data symbols from all the other users are added on an element-by-element basis. The summed spread symbols from all the users are then fed into a frequency interleaver. The function of the interleaver is to achieve frequency diversity by placing

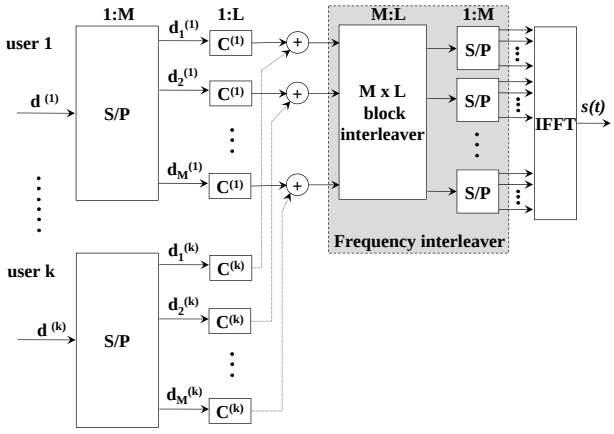

Figure 1. MC-CDMA Transmitter.

chip-symbol elements that correspond to the same symbol on subcarriers that are a distance M apart. After frequency interleaving, the symbol elements are fed into an MxL point IFFT block for OFDM operation. The output samples are then converted back into serial data to generate the complex baseband signal $s(t)$. This signal for an MC-CDMA symbol $0 \le t \le Ts$, can be expressed as,

$$s(t) = \sum_{m=1}^{M} \sum_{l=1}^{L} \sum_{k=1}^{K} d_m^{(k)} c_l^{(k)} e^{j2\pi\{M(l-1)+(m-1)\}t/T_s} \qquad (1)$$

The definition of PAPR for an MC-CDMA symbol is given by (2).

$$PAPR = \frac{\max_{0 \le t \le T_s} |s(t)|^2}{E\left[|s(t)|^2\right]} \qquad (2)$$

In this connection, another parameter of interest is the crest factor (CF) whose relationship with PAPR is expressed by (3).

$$CF = \sqrt{PAPR} \qquad (3)$$

This parameter has been used in some of the previous related studies [4-6] and we in this study also consider CF as the measure of variation of the transmission signal's envelop.

III. Spreading Codes

As mentioned before the focus of this study is the OBU code set. For comparison, we chose WH code set since they have been studied extensively in the past. For example references [5-7] showed that Golay complementary codes show better PAPR performance compared to other orthogonal codes at low active user density but WH is a better choice at medium or heavy user densities. Moreover, with respect to employing WH codes in multiuser environment, significant improvement in PAPR has been reported through methodical allocation of user code other than selecting them in a straight forward manner [5, 8, 9, 14].

A. Walsh-Hadamard Codes (WH Codes)

Walsh-Hadamard code is generated from a Hadamard matrix whose rows form an orthogonal set of codes. The codes sequences in this code set are the individual rows of a Hadamard matrix. Hadamard matrices are square matrices whose entries are either +1 or -1 and whose rows and columns are mutually orthogonal. If *N* is a non-negative power of 2, the *NxN* Hadamard matrix, is defined recursively as follows.

$$H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & H_{-n} \end{bmatrix}, \quad H_2 = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix} \quad (4)$$

B. Orthogonal Binary User Codes

Reference [10] proposed these codes and reported that they were constructed through performing search operations in the binary sample space that consisted of zero mean and linear phase codes only. Later they were expanded for non-linear phase code also [11]. In contrast to WH codes, the OBU codes do not necessitate the condition that there can be only one code in the set for a given number of zero crossings. Again, for WH codes, decimal values of all *n*-bit codes are multiples of either $2^{(n/2)} + 1$ or $2^{(n/2)} - 1$. OBU codes do not impose this restriction. Rather, efforts were made to reduce the number of codes that fulfill this criterion in order to avoid codes that are common to WH codes. Following all these principles, the authors of [10] have shown that for a given code length, more than one OBU code sets can be formulated. In this context, one may remember that for a given code length there exists only one set of WH codes. Since shorter code length implies smaller binary sample space, there exist a considerable number of short OBU codes that are common with same length WH codes. But this number reduces as the code length becomes higher. Table I shows an example of a typical 16 bit OBU code set. For the sake of convenient representation, the codes are shown as decimal numbers. All the occurrences of "-1" bits in a code sequence is at first replaced by "0" bits and then the decimal equivalent of the code sequence is calculated and shown on the table.

TABLE I.
EXAMPLE OF 16 BIT ORTHOGONAL BINARY USER CODE SET

Code index	Decimal notation
1	65535
2	383
3	3727
4	39321
5	12979
6	42405
7	50115
8	15683
9	21717
10	43605
11	52275
12	23333
13	61455
14	26393
15	26857
16	38505

III. ANALYTICAL STUDY

In the context of PAPR, one important property of spreading codes is its aperiodic auto-correlation value. If we consider the *L*-bit long *x*th code $C^{(x)} = \left[c_1^{(x)}, c_2^{(x)}, ..., c_L^{(x)}\right]$ from a code set having *P* distinct code sequences, the aperiodic auto-correlation $AC_n^{(x)}$ is given by,

$$AC_n^{(x)} = \sum_{i=0}^{L-n-1} c_i^{(x)} [c_{i+n}^{(x)}]^* \quad (5)$$

where [ ]* refers to the conjugate operation.

Figure 2 depicts the comparative collective aperiodic auto-correlation of 16 bit OBU and WH codes. It shows that OBU codes in general exhibits lower values of aperiodic auto-correlation than that of WH codes.

In MC-CDMA system, since the power spectrum of spreading code is related to the transmitted signal amplitudes, PAPR can be estimated by analyzing the correlation properties of those codes. Based on this principle, [9] proposed an algorithm for finding code combinations resulting in least PAPR. Following that algorithm, the comparative minimum normalized peaks of WH and OBU codes are shown in Fig. 3. It is evident that some selected OBU code combinations possess better aperiodic auto-correlation values than corresponding WH codes combinations and hence show lower peak values for all most all the user cases.

However, in multiuser environment, where multiple codes are used simultaneously, apart from aperiodic auto-correlation, a message dependent parameter called the collective aperiodic cross-correlation has significant effect on PAPR [5]. This parameter is defined by,

$$CC_n = \sum_{x=0}^{P-1} \sum_{y=0, y\neq x}^{P-1} b^{(x)} [b^{(y)}]^* CC_n^{(x),(y)} \quad (6)$$

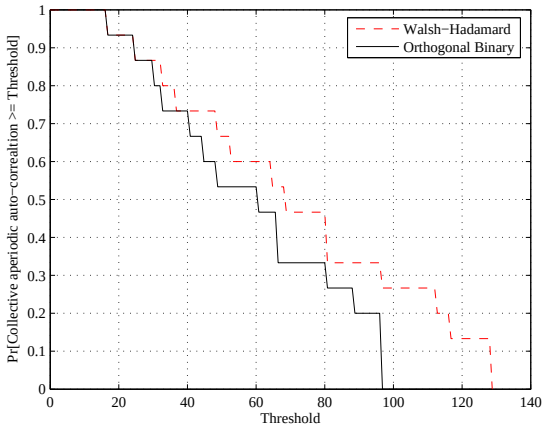

Figure 2: Comparative aperiodic auto-correlation.

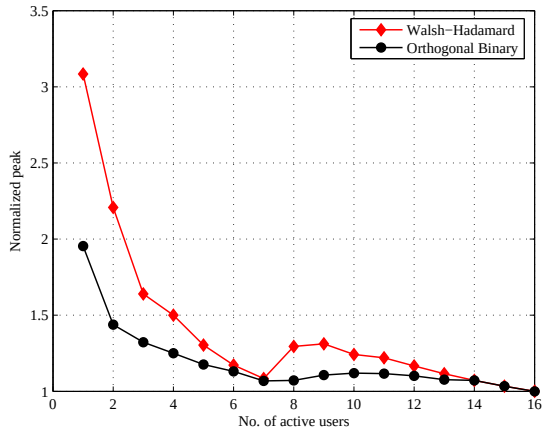

Figure 3: Minimum normalized peak for different active user level.

However, in multiuser environment, where multiple codes are used simultaneously, apart from aperiodic auto-correlation, a message dependent parameter called the collective aperiodic cross-correlation has significant effect on PAPR [5]. This parameter is defined by,

$$CC_n = \sum_{x=0}^{P-1} \sum_{y=0, y\neq x}^{P-1} b^{(x)}[b^{(y)}]^* CC_n^{(x),(y)} \qquad (6)$$

where, $b^{(x)}$ and $b^{(y)}$ are message symbols from the xth and the yth user respectively and $CC_n^{(x),(y)}$ is the aperiodic cross-correlation between xth and yth codes given by,

$$CC_n^{(x),(y)} = \sum_{i=0}^{L-n-1} c_i^{(x)}[c_{i+n}^{(y)}]^* \qquad (7)$$

In our study, we consider 16QAM as the modulation method where the number of possible distinct message symbols is large compared to BPSK or QPSK. Thus in addition to the just presented analytical result, a comprehensive simulation is performed.

IV. Spreading Code Allocation Methodology

Considering the architecture of Fig. 1 and our previous discussion with respect to spreading codes, the symbols just prior to interleaving can be represented as,

$$\begin{bmatrix} Q_1 \\ Q_2 \\ . \\ Q_M \end{bmatrix} = \begin{bmatrix} d_1^1 & d_1^2 & \ldots & d_1^K \\ d_2^1 & d_2^2 & \ldots & d_2^K \\ \ldots & .. & \ldots & .. \\ d_M^1 & d_M^2 & \ldots & d_M^K \end{bmatrix} x \begin{bmatrix} c_1^{(1)} & c_2^{(1)} & \ldots & c_K^{(1)} \\ c_1^{(2)} & c_2^{(2)} & \ldots & c_k^{(2)} \\ \ldots & \ldots & \ldots & \ldots \\ c_1^{(K)} & c_2^{(K)} & \ldots & c_K^{(K)} \end{bmatrix} \qquad (8)$$

where, we have considered the length of spreading codes equal to that of total number of maximum possible users. Here, the spreading codes have been allocated on a plain sequential manner, i.e., user index and code index is same. But it is possible to check other combinations of code allocation as discussed below resulting in different values of PAPR.

If we consider a code set consisting of 16 unique orthogonal code sequences and 8 active users in the system, we find that there can be as many as $C_8^{16} = 12870$ unique different possible ways of allocating codes to the users. And since the collective aperiodic auto and cross correlation values of these combinations may vary, so may the CF produced by them. As a result, finding the best code combinations, i.e., the combinations that produce lowest CF, is of interest. But, recalling the fact that the collective aperiodic cross correlation is dependent on message symbols [5], a perfect estimation of lowest CF producing code combination requires consideration of all possible message combinations along with all possible code combinations. For example with 16QAM modulation, if n number of users each send 1 symbol, 16^n no. of different symbol combinations is possible. So, the number of possible code sequence-symbol combination comes to C_n^{16} x 16^n. The second term increases exponentially with increasing n. But, since higher user density results in less PAPR [5-7], relatively low values of n is of prime concern.

We define two alternative methods of the spreading code allocation strategies. We name the first method as best code allocation and the second one as the optimum code allocation method.

A The Best Code Allocation Approach

In the first method, users are assigned specific spreading codes from all the available spreading codes in such a manner so that the resultant crest factor (CF) becomes minimum. This approach considers each user case discretely and formulates code combinations yielding least CF for those cases. For example, if the number of active users is k and the total number available spreading codes is N, $\left[C^{(1)}, C^{(2)},..., C^{(k)}\right]$ is chosen from $\left[C^{(1)}, C^{(2)},..., C^{(N)}\right]$ so that the CF of the transmit signal is minimum. When choosing the best spreading combination, the effect of user data is also taken into account by considering random combinations of input symbols from all users.

The method is illustrated in Fig. 4. Here, as soon as the state of the user case changes, the entire code set is searched to find the new code combination producing lowest CF. For example, if the number of users change

1. Start with user case n = 1.

2. Select a code combination having n number of code sequences from total available N codes. within the N length code set .

3. Generate random 16 QAM data frames for all n users

4. Apply those code sequences to perform MC-CDMA operations.

5. Calculate CF from the time domain signal.

6. Repeat steps 3 to 5 for a predefined number of times and calculate average CF.

7. Repeat steps 2 to 6 for $C_n^N - 1$ times for all other combinations.

8. Record the lowest CF and corresponding code combination.

9. Increment n by 1 and repeat steps 2 to 8 for all other possible values of n.

Figure 4. Principle of the best code allocation approach.

1. Start with user case n = 1.

2. Select a single code from total available N-(n-1) codes within the N length code set

3. Generate random 16 QAM data frames for all n users

4. Apply those code sequences to perform MC-CDMA operations.

5. Calculate CF from the time domain signal.

6. Repeat steps 3 to 5 a predefined number of times and calculate average CF.

7. Repeat steps 2 to 6 for $C_n^{N-(n-1)} - 1$ times for all other combinations.

8. Record the lowest CF and corresponding code combination.

9. Increment n by 1 and repeat steps 2 to 8 for all other possible values of n.

Figure 5. Principle of the optimum code allocation approach.

from 4 to 5, $C_5^{16} = 4368$ possible code combinations are searched to find the best one.

A drawback of this method is that on the arrival of a new user, besides allocating code to this user, it necessitates that all users who were in the system before the change of user case occurred also get their spreading codes reassigned since the new code combination may not preserve the previously allocated codes. This may result in an additional overhead from switching and resource allocation perspective and thus may increase the system complexity to a higher extent.

B The Optimum Code Allocation Approach

In contrast to the first approach, this approach, considers change of user cases as a continuous process. Here, codes allocated to users on a particular user case are marked as used and new users are assigned codes only from those codes that have not been allocated yet. Returning to the previous example, when the number of users changes from 4 to 5, already allocated user codes are not included in the search space for finding the code to be allocated to user number 5. As a result, only $C_1^{12} = 12$ code combinations are now searched to find the best combination. This approach has the drawback of showing possible inferior CF performance compared to the previous approach as the search space is limited. But computational complexity for building the code allocation table is reduced and also no additional overhead related to dynamically changing already allocated user codes is present as was the case with the other approach. The scenario is depicted in Fig. 5. We call it optimum code allocation approach since it optimizes system complexity at the cost of possible higher crest factor.

In general for both the approaches, the base station transmits the code allocation table to a new user along with the user index. Then, for the best allocation approach, a notification is sent to all existing users whenever a new user arrives. But in the second approach, the existing users do not need any notification in case of new arrivals.

V. SIMULATION MODEL AND PERFORMANCE EVALUATION

TABLE II.
SIMULATION PARAMETERS

Modulation scheme	16 QAM
Spreading code length	16
No. of symbols per frame	4
No. of subcarriers	64
Non-linear amplifier model	SSPA
Level of non-linearity, r	3

Our simulation was based on the transmitter model shown in Fig. 1 In addition, for analyzing bit error rate, we considered a solid state power amplifier (SSPA) model [12], an AWGN channel and a simple correlator based receiver architecture. The salient simulation parameters are listed in Table II. And the non-linear characteristic of the amplifier is given by (9).

$$F[x] = \frac{x}{\left[1 + \left(x/A\right)^{2r}\right]^{1/2r}} \tag{9}$$

where, x is the amplitude of the input signal, A is the saturated output level and r is the parameter that decides the level of non-linearity. For fixing amplifier operating point we considered the parameter output back-off (OBO) which is given by.

$$OBO = 10\log_{10}\left(\frac{P_{sat\ out}}{P_{out}}\right) \tag{10}$$

where, $Psat_{out}$ is the saturation power referred to the output and $Pout$ is the output power.

A. Single User Scenario

In the context of uplink communication, since the signals from individual users are amplified by different amplifiers, every spreading code sequence of a particular code set is examined individually. Our simulation result shown in Fig. 6 illustrates the comparative CF values of WH and OBU codes. Here, CF values for all the 16 distinct sequences available from 16 bit long WH and OB code set are plotted and compared on an individual basis. The results show that 50% of OBU codes generate a lower CF than corresponding WH codes and 93.75% of OBU codes produce a lower or equal CF than corresponding WH codes. Comparison was also done with the help of complementary cumulative distribution function (CDF) of CF considering all the code sequences. This is depicted in Fig. 7 and it clearly shows the superior CF properties of OBU codes over WH codes.

B. Multiuser Scenario

In a synchronized downlink communication environment, the base station transmits a signal that constitutes of summed signals of multiple users spread by different spreading codes. Hence the CF property of the transmitted signal, as mentioned before, is determined by the collective properties of individual spreading codes.

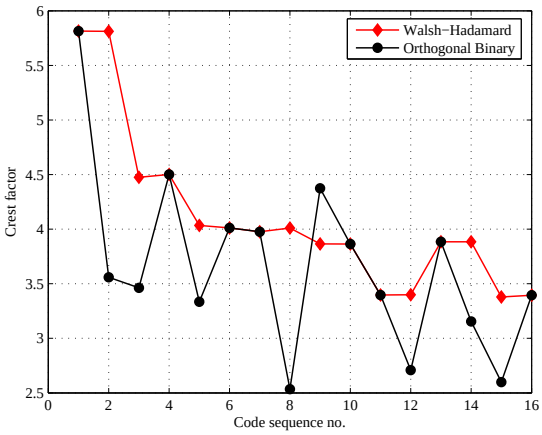

Figure 6. Crest factor of 16 bit long codes.

TABLE III.
THE CODE ALLOCATION TABLE

No. of active users	Best combinations	Optimum combinations
1	0x0100	0x0100
2	0x0102	0x0102
3	0x0112	0x0112
4	0x4806	0x0912
5	0x6901	0x4912
6	0x6816	0x6912
7	0x6916	0x6916
8	0x16E9	0xE916
9	0x16EB	0xE917
10	0x96EB	0xE997
11	0xD7E3	0xE99F
12	0xBFE9	0xE9DF
13	0xF7F9	0xEBDF
14	0xDFDF	0xFBDF
15	0xFFFD	0xFFDF
16	0xFFFF	0xFFFF

We constructed the code allocation table for OBU codes considering a multiuser system that can support a maximum of 16 simultaneous users. For each possible user cases, i.e., 1 to 16, we simulated the system considering 4 random 16QAM symbols per frame. Thus one MC-CDMA symbol consisted of a minimum of 4 to a maximum of 16x4=64 QAM user symbols spread by 16 bit long spreading codes. We put more emphasis on the spreading codes by examining their all possible combinations and for each of this combination the user symbol effect is included by considering 100 frames random data.

Table III shows the code combination values resulting from the two allocation approaches. Here, the combination values are represented in hexadecimal format [5] which is explained using an example. Let us take the particular user case of a total of 5 active users in the system. The best code combination value for this case reads $(6901)_{16}$. When this hexadecimal number is converted into its binary equivalent, it becomes $(0110\ 1001\ 0000\ 0001)_2$. This binary string has the bit value of "1" in its 2nd, 3rd, 5th, 8th and 16th position from the MSB and it implies that the combination of codes having the indices of 2, 3, 5, 8 and 16 (refer to Table I) produces the lowest CF for all possible combinations.

For the same user case, the optimum combination displays the values of $(4912)_{16}$, i.e., $(0100\ 1001\ 0001\ 0010)_2$ which means codes having the indices of 2, 5, 8, 12 and 15 represent the code combination with the lowest CF value resulting from the optimum allocation approach. Again, if we move to the next user case and check the optimum combination value, we find that it shows $(6912)_{16}$, which when converted to binary is $(0110\ 1001\ 0001\ 0010)_2$. By comparing this value with that of the previous user case we find that only the code with an index of 3 has been added without any change to the already allocated code indices. But for the best allocation approach, comparing the same two user cases it is found that there are changes associated with more than one bit positions in the binary strings representing the code combinations. These two observations confirm the difference between the two code allocation approaches that was mentioned earlier.

Figure 8 illustrates the comparative results of different code allocation approach for OBU codes. As is evident from this figure, sequential allocation shows very poor CF values at low user densities but its performance gradually improves elsewhere. By sequential allocation we refer to allocation of codes in the same order as found on the code set table.

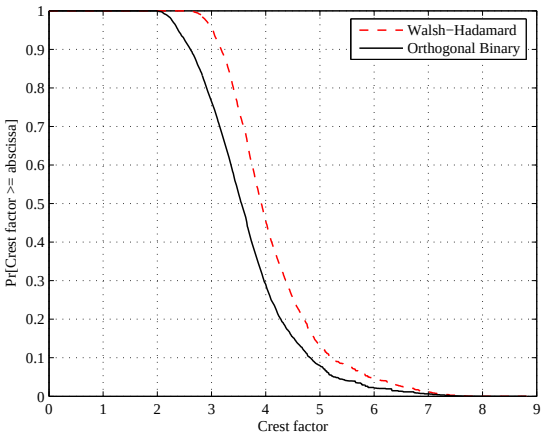

Figure 7. Comparative CDF of crest factors.

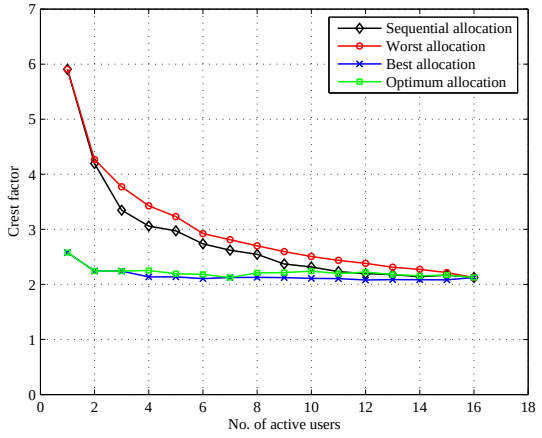

Figure 8. Crest factor of OBU codes for different allocation method.

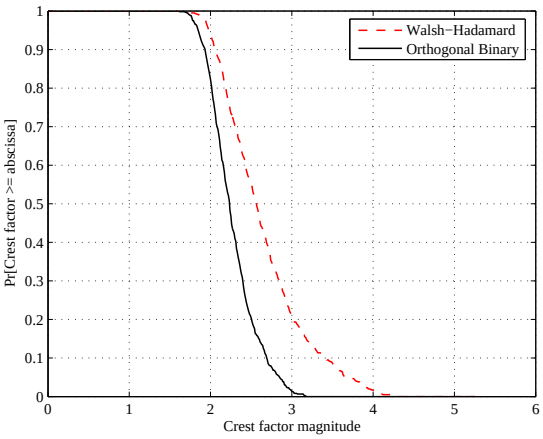

Figure 9. Comparative crest factor at low user density.

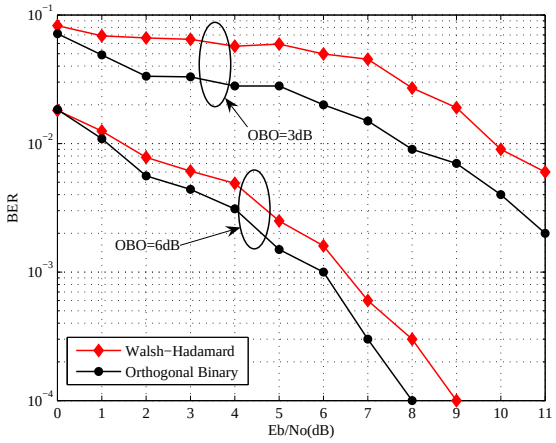

Figure 10. Comparative bit error rate

That means user *k* is assigned the code sequence having the index *k* and so on. CF plots for the other three cases found in this illustration are not sequential, i.e., user index and code index are necessarily not the same. In this respect, the worst allocation corresponds to those code combinations which yield maximum CF. This finding came up when different combinations were being tested for CF comparison. It is understandable that these are the least desired combinations. The unintended presence of these code combinations can be eliminated by avoiding random code allocation.

Now, the last two graphs of Fig. 8 are of special interest as they correspond to best and optimum code allocation approaches. Compared to sequential allocation approach, both of them show an improvement in performance. Specifically in the region of low user density their superiority over sequential approach is noticeable. It is interesting to notice that optimum combination approach demands less computational and implementational overhead compared to best allocation approach and yet demonstrates almost the same performance.

Best code combinations formulated from WH codes have been reported before [5, 9]. Since the problem of WH codes is to generate higher peaks at low user density, we decided to perform a relative CF comparison of best code combinations of OBU and WH codes focusing on relatively lower number of total active users. For this simulation, we chose a user density of <=37.5% and compared the complementary CDF of CF of those two best code combinations as shown in Fig. 9. From this figure, we find that best combinations of OBU codes perform better than the corresponding WH best code combinations.

Finally, Fig. 10 shows the comparative BER depicting the effect of power amplifier non-linearity on the peak produced by the two best combination spreading codes. The BER simulation was carried out by concentrating on the region of low user density by selecting the case of 4 active users. Two different level of amplifier OBO, i.e., 3dB and 6dB were investigated, and for both the cases the results show that the best combinations of OBU codes are less affected by amplifier non-linearity than the corresponding WH best code combinations. From these results, it is obvious that selected code subsets from the OBU code set exhibit better performance than WH code subset.

VI. Conclusions

Multicarrier systems often possess high peaks in their time domain signal that may lead to non-linear amplification and there by cause severe BER degradation. CF is a parameter that quantifies the peaks in a signal and with the help of this parameter, we analyzed the comparative performances of WH codes and OBU codes in the context of an MC-CDMA system that uses a non-linear amplifier and works under the influence of an AWGN channel. We found that in a single user scenario the OBU codes produce lower CF values than WH codes.

And for a multiuser environment, when spreading codes are applied in a sequential manner, the OBU codes show decreasing CF with increasing user density. Being, motivated by the approaches of [5, 9], we constructed code allocation tables by selecting low CF producing code subsets. In this context, we also introduced the optimum allocation approach that minimizes system complexity yet achieves good performance. Our analytical and simulation results showed that selected code combinations from the OBU code set exhibit better performance than corresponding WH codes. On the basis of our study results and the fact that OBU code and non-linear phase Walsh like orthogonal codes give better performance than WH codes in conventional single carrier CDMA systems [10, 11], we affirm that for multiuser MC-CDMA systems, use of specific OBU code combinations as spreading codes can lead to considerable reduction in CF and thus result in improved system performance.

The code allocation table only considers increase in user cases. In case of decrease in the number of users some codes will be released. It is important to consider this scenario also in order to have a more efficient and practical code allocation procedure. We consider it as a scope of future work.

REFERENCES

[1] X. Li and L. J. Cimini, "Effects of clipping and filtering on the performance of OFDM", *IEEE Communications Letters*, vol. 2, no. 5, pp. 131-133, May 1998.

[2] R. W. Bauml, R. F. H. Fischer and J. B. Huber, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping", *IEE Electronics Letters*, vol. 32, no. 22, pp. 2056-2057, October 1996.

[3] L. J. Cimini and N. R. Sollenberger, "Peak-to-average power ratio reduction of an OFDM signal using partial transmit sequences", *IEEE Communications Letters*, vol. 4, no. 3, pp. 86-88, March 2000.

[4] B. J. Choi, E. L. Kuan and L. Hanzo, "Crest factor study of MC-CDMA and OFDM", *Proc. IEEE VTC 1999-Fall*, vol. 1, pp. 233-237, 1999.

[5] L. Hanzo, M. Munster, B. J. Choi and T. Keller, "Advanced peak factor reduction techniques, in OFDM and MC-CDMA for broadband Multi-user Communications, WLANs and broadcasting", *WILEY-IEEE press*, pp. 288-306, 2003.

[6] S. Nobilet, J-F. Helard and D. Mottier, "Spreading sequences for uplink and downlink MC-CDMA systems: PAPR and MAI minimization", *ETT*, vol. 13, no. 5, pp. 465-474, 2002.

[7] H. Ochiai and H. Imai, "OFDM-CDMA with peak power reduction based on the spreading sequences", *Proc. IEEE ICC '98*, pp. 1299-1303, June 1998.

[8] N. Hathi, I. Darwazeh and J. O'Reilly, "Peak-to-average power ratio performance comparison of different spreading code allocation strategies for MC-CDMA and MC-DS-CDMA", *Electronics Letters*, vol. 32, no. 20, pp. 1219-1220, September 2002.

[9] K. Kang, K. Choi and S. K. Shin, "Reduced search for optimum code sets to reduce PAPR in MC-CDMA System", *5th International Symposium on Wireless Personal Multimedia Communications*, vol. 1, pp. 135-139, October 2002.

[10] R. Poluri and A. N. Akansu, "New orthogonal binary user codes for multiuser spread spectrum communications", *13th European Signal Processing Conference, Antalya, Turkey*, September 4-8 2005.

[11] A. N. Akansu and R. Poluri, "Walsh-Like Nonlinear Phase Orthogonal Codes for Direct Sequence CDMA Communications", *IEEE Transactions On Signal Processing,* Vol. 55, No. 7,pp. 3800-3806, July 2007.

[12] C. Rapp, "Effects of HPA-nonlinearity on a 4-DPSK/OFDM signal for a digital sound broadcasting system," *Proc. Second European Conference on Satellite Communications*, Liege, Belgium, pp. 179-184, October 1991.

[13] L. A. P. Hernández and M. G. Otero, "User Reservation Approach for Peak-to-Average Power Ratio Reduction in MC-CDMA Systems", *IEEE 69th Vehicular Technology. Conferenc*e, Barcelona, Spain, April 26-29, 2009.

[14] L. Yang and E. Alsusa, "Dynamic Code-allocation based PAPR Reduction Techniques fro MC-CDMA Systems", *2007 IEEE Wireless Communications and Networking Conference*, Hongkong, March 11-15, 2007.

[15] S. Ahmed, T. Noguchi and M. Kawai, "Selection of Spreading Codes for Reduced PAPR in MC-CDMA Systems," *18th IEEE Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07)*, Athens, Greece, September 2007.

Sabbir Ahmed received his B.E. degree in Electrical and Electronics from Bangladesh University of Engineering and Technology (BUET), Bangladesh and M.E. degree in Information Science and System Engineering from Ritsumeikan University, Japan in 1999 and 2007 respectively. He is currently enrolled as a PhD student in the same school from where he did his M.E. His research interests include mobile and wireless communication, multicarrier systems, multiple access methods and non-linear effects. He received the Best Student Paper Award at the 25th AIAA International Communication Satellite System Conference (ICSSC '07) held in Seoul, South Korea, 2007.

Makoto Kawai was born in 1949. He received B.S., M.S. and Dr. Eng. Degrees in electrical engineering from Kyoto University, Kyoto, Japan, in 1972, 1974 and 1987, respectively. During 1974 and 1999 he worked mainly for research and development of wireless communication systems at Nippon Telegraph and Telephone Corporation (NTT) and also worked at Advanced Telecommunications Research Institute International (ATR) from 1996 to 1999. He was an associate professor at Graduate School of Informatics, Kyoto University from 1999 to 2003. He is now a professor of Graduate School of Science and Engineering and College of Information Science and Engineering, Ritsumeikan University, Shiga, Japan. He is a member of IEEE, AIAA, IEICE, IPSJ and IEEJ.

A New Group Key Agreement Protocol for Ad Hoc Networks

Zhang Li-Ping
Lab. of information security,
College of Computer Science and Technology,
China University of Geosciences,
Wuhan, China
Carolyn321@163.com

Wang Yi
Department of information engineering,
Wuhan Police Vocation College,
Wuhan, China

***Abstract*—A mobile ad hoc network is a collection of autonomous nodes that communicate with each other by forming a multi-top wireless network. Different from conventional wireless networks, the resource of the nodes in ad hoc networks is limited and there may be tens of thousands of low-power energy constrained nodes in ad hoc networks. As such, the costs of the nodes resource and the network size should be taken into consideration when constructing a group key agreement protocol in the ad hoc networks. In this paper, an efficient and scalable group key agreement protocol based on layer-cluster group model for mobile ad hoc networks was proposed. In this protocol, multi-linear map is employed on layer-cluster structure to establish and allocate group key. So that it can not only meet security demands of larger mobile ad hoc networks but also improve executing performance.**

***Index Terms*—ad hoc networks, layer-cluster, group key agreement, multi-linear map**

I. INTRODUCTION

Wireless ad hoc networks are becoming progressively popular as they have the ability to form "on the fly" and can dynamically handle the joining or leaving of nodes in the network. However, the use of wireless links gives chances to attacks ranging from passive interception, replaying, and data interpolation, denial of service and identity forgery. In addition, wireless ad hoc networks usually operate in a wide open space and their topologies change frequently, so that the nodes are prone to be compromised. Because of these attacks, security measures should be adopted to protect the ad hoc communications.

Most security requirements, such as privacy authenticity and integrity, can be addressed by building upon a solid key management framework [1]. A secure group key agreement is the prerequisite for the security of these primitives, and thus essential to achieving secure infrastructure in ad hoc networks. However, the larger size of the group and the dynamic character of group changes pose a challenge on group key management research for wireless ad hoc networks.

Nodes in wireless ad hoc networks are usually low power devices that run on batter power and become unusable after failure or energy depletion. As a result, there is a need to employ energy-efficient group key agreement protocol in order to increase the overall network longevity.

Furthermore, given the potentially large number of mobile devices, scalability becomes another critical issue. The scalability problem can be solved by partitioning the communicating devices into subgroups, with a leader in each subgroup, and further organizing the subgroups into hierarchies [2].

In this paper we propose a new group key agreement protocol *LCML* , aimed at addressing a lightweight and fast solution in ad hoc networks. In protocol *LCML* , the network is partitioned into several clusters to construct *h* layers. On this layer-cluster model, multi-linear map is employed to establish group key which can not only meet security demands of mobile ad hoc networks but also reduce the communication costs.

The rest of this paper is organized as follows. In section II, we discuss related works on group key agreement protocols for ad hoc networks. Section III presents our key agreement protocol. In section IV, the security of the proposed protocol is discussed. We discuss the performance in section V, and conclude the paper in section VI.

II. RELATED WORKS

Recently, the area of group key management over

Manuscript received January 31, 2010; revised April 18, 2010; accepted May 11, 2010.

The Project was Supported by the Special Fund for Basic Scientific Research of Central Colleges, China University of Geosciences （Wuhan）No. CUGL090246

doi:10.4304/jcm.5.7.545-550

wireless ad hoc networks has received a significant amount of attention in literature. Since the foundational Diffie-Hellman (DH) protocol [3], several other protocols have been proposed for the group case. The first group key agreement protocol known as ING protocol was proposed by Ingemarsson et al. [4] in 1982. Following their work, Steiner et al. [5, 6, 7] proposed a family of protocols known as Group Diffie-Hellman (GDH.1, GDH.2 and GDH.3). In these protocols, the last group member servers as a controller and performs most of the computation on behalf of other group members in the group, therefore it needs more energy compared with other group members. Due to the limitation of the nodes energy the GDH protocol is inappropriate to the ad hoc networks. Kim et al. extended the work of a tree-based key agreement scheme by Perrig [8] to design a Tree-Based Group Diffe-Hellman (TGDH) protocol in [9]. Compared with GDH, it scales down the number of exponentiations and received messages required by the last group member to avoid excessive computational and communication costs required by one node. But TGDH protocol still requires each group member to perform large modular exponentiations and transmit/receive large messages. So the TGDH protocol is also inadequate for ad hoc networks. Kim et al. also proposed another tree-based key agreement scheme named as STR [10], which is quite similar to TGDH. In 2005, an efficient GKA protocol for low-power mobile devices was proposed by Cho et al [11]. However, this protocol requires a special member U_n to perform high computation on behalf of other members in the group. In the same year, Teo et al. [12] proposed an energy-efficient and scalable group key agreement scheme named as C-H protocol, which claimed that it is adapted to the large ad hoc networks. Although the C-H protocol logarithmically scales down the number of exponentiations, it increases the communication costs, compared to the GDH protocol and TGDH protocol. Based on their work, Zhang et al. [13] proposed a new protocol CH-ECC. In this protocol, the elliptic curve cryptosystem is employed by circular hierarchical group model to establish group key. So that it scales down the costs of communication. However the scalability problem is not taken into account in this protocol.

In order to solve the scalability problem, Jason H. et al. proposed a scalable key management and clustering scheme for ad hoc networks [2]. In this protocol, the communicating devices are divided into subgroups, with a leader in each group, and then organizing the subgroups into hierarchies. On this hierarchic structure, Diffie-Hellman protocol is used to establish group key. While this is one of the most recognized energy-efficient clustering protocols, its performance can be further enhanced.

Dan Boneh and Alice Silverberg studied some questions in linear algebra and cryptography and then presented several applications of multi-linear forms to cryptography [14].

Now, we give a definition of a d multi-linear map. Let G_1 be a cyclic additive group of prime order p and G_2 be a cyclic multiplicative group of same order p. We assume that the discrete logarithm problems (DLP) in both G_1 and G_2 are intractable. A map $e: G_1^{\ d} \to G_2$ is a d multi-linear map if it satisfies the following properties [14]:

1. Multi-linear: For $\forall a_1,...,a_d \in Z_p^*$ and $\forall P_1,...,P_d \in G_1^*$, $e(a_1P_{1,...,}a_dP_d) = e(P_{1,...,}P_d)^{a_1 \cdots a_d}$

2. Non-degenerate: if $P \in G_1$ is a generator of G_1 then $e(P_{,...,}P)$ is a generator of G_2;

3. Computable: There exists an efficient algorithm to compute $e(P_{1,...,}P_d)$ for $P_1,...,P_d \in G_1^*$.

Based on their work [14], some group key management protocols were proposed [15, 16]. A common advantage of those protocols is that the one-round multi-party key exchange can be easily performed. In addition the security of those protocols always based on the Decisional Multi-linear Diffie-Hellman problem and Decisional Multi-linear Diffie-Hellman Assumption.

Definition1. The Decisional Multi-linear Diffie-Hellman (DMDH) problem is given $(P, a_1P, a_2P, ..., a_{d+1}P)$ and $z \in G_2$, to decide whether $z = e(P,P,...,P)^{a_1a_2...a_{d+1}}$ or not.

Definition2. Decisional Multi-linear Diffie-Hellman Assumption claims that for any polynomial time algorithm T and any $d > 1$, the advantage $DMDH_{T,d}(t)$ of T in solving the Decisional Multi-linear Diffie-Hellman problem is negligible, where $DMDH_{T,d}(t)$ is the probability that T can distinguish $e(P,P,...,P)^{a_1a_2...a_{d+1}}$ from $z \in G_2$.

Although Dan Boneh and Alice Silverberg point out those multi-linear maps is hard to build we believe that this issue can be solved by new techniques soon.

III. LAYER-CLUSTER KEY AGREEMENT PROTOCOL

A. Notation and Terminology

We use the following notation throughout the rest of this paper:

h: total number of layers in the group model;

L_i; ith layer for $i \in [0,\ldots,h\text{-}1]$ in the group model;

n: group size i.e. the total number of the nodes in the group model;

n_1: total number of subgroups when the group size is n;

tsg_{L_i} : total number of subgroups at layer L_i;

$SG_j^{(L_i)}$: jth subgroup at layer L_i ($j \in [0,\ldots, tsg_{L_i}\text{-}1]$);

$U_{SG_j}^{(L_i)}$: subgroup controller of the jth subgroup at layer L_i;

ub_{L_i} : the upper bound of the size of subgroup at layer L_i;

lb_{L_i} : the lower bound of the size of subgroup at layer L_i;

$t_{SG_j^{(L_i)}}$: total number of subgroup members in jth subgroup at layer L_i, $lb_{L_i} \le t_{SG_j^{(L_i)}} \le ub_{L_i}$;

$U_{(j,k)}^{(L_i)}$: kth member of L_i and it in subgroup $SG_j^{(L_i)}$ ($k \in [0,\ldots, \sum_{j=0}^{tsg_{L_i}-1} t_{SG_j^{(L_i)}} -1]$);

$\{m\}_e$: a symmetric key encryption scheme;

B. Description of layer-cluster group model

In order to secure group communication for a large ad hoc network containing n users, the proposed protocol (*LCML*) adopt a layer-cluster group model as shown in Fig.1.

Denote the highest layer as L_0 while the lowest layer as L_{h-1}. In the layer-cluster group model each layer L_i ($i \in [0,\ldots,h-1]$) consists of subgroups denoted as $SG_j^{(L_i)}$ ($j \in [0,\ldots, tsg_{L_i} -1]$) and each subgroup $SG_j^{(L_i)}$ have some subgroup members denoted as $U_{(j,k)}^{(L_i)}$, in which k represents the position of the subgroup member at the layer L_i. The size of subgroup $SG_j^{(L_i)}$ is restricted by a lower and an upper bound. Each layer has one lower and upper bound which will be used across all the subgroups in that layer. And each layer can has different a pair of bound. Denote the minimum lb_{L_i} among all lb_{L_i} ($i \in [0,\ldots,h-1]$) as $lb_{\min}$ and the maximal ub_{L_i} among all ub_{L_i} ($i \in [0,\ldots,h-1]$) as $ub_{\max}$ in layer-cluster group model. Further the subgroups in each layer should be disjoint.

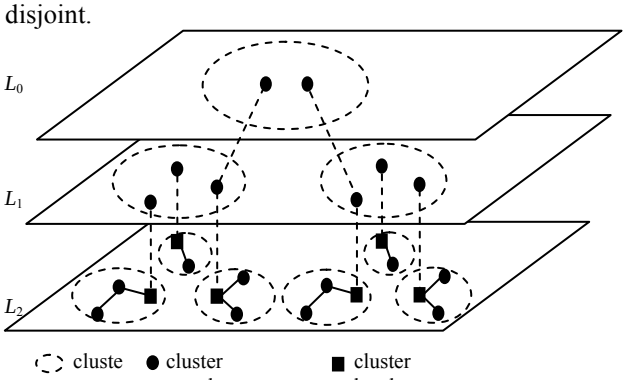

Fig.1. An illustration of the layer-cluster group model with h=3

In the layer-cluster group model a cluster is represented by a subgroup and a cluster member is represented by a subgroup member. In each subgroup all the subgroup members are arranged in a ring and let the subgroup member which represents the cluster head be the first member. Each subgroup $SG_j^{(L_i)}$ is managed by a subgroup controller $U_{SG_j}^{(L_i)}$ who is also the first member of that subgroup, i.e. $U_{(j,\ \sum_{s=0}^{j-1} t_{SG_s^{(L_i)}}+1)}^{(L_i)} = U_{SG_j}^{(L_i)}$ ($j \in [1,\ldots,\ tsg_{L_i} -1]$). The subgroup controller of all the subgroups in layer L_i except the highest layer $L_i \neq L_0$ join the layer L_{i-1}. So the subgroup members $U_{(j,k)}^{(L_i)}$ in each layer except the lowest layer $L_i \neq L_{h-1}$ are also subgroup controllers $U_{SG_k}^{(L_{i+1})}$ of subgroup $SG_k^{(L_{i+1})}$ at the next layer L_{i+1}, i.e. $U_{(j,k)}^{(L_i)} = U_{SG_k}^{(L_{i+1})}$.

C. Group key agreement protocol based on layer-cluster group model

In this section, we propose a new group key agreement protocol (*LCML*) based on layer-cluster group model for ad hoc networks. This protocol comprises three phases as follows:

*Phase*1 : the proposed protocol *LCML* starts at the lowest layer L_{h-1}. The process of subgroup key agreement in subgroup $SG_0^{(L_{h-1})}$ at the lowest layer L_{h-1} is described in details as follows:

1. Every subgroup member $U_{(0,k)}^{(L_{h-1})}$ of subgroup $SG_0^{(L_{h-1})}$ chooses an integer $r_{(0,k)} \in Z_p^*$ randomly as its private key.

2. Every subgroup member $U_{(0,k)}^{(L_{h-1})}$ computes its public key $r_{(0,k)}P$ and broadcast it to subgroup $SG_0^{(L_{h-1})}$.

3. After subgroup member $U_{(0,k)}^{(L_{h-1})}$ obtain all public keys of other subgroup members in $SG_0^{(L_{h-1})}$ it can compute subgroup key $K_{SG_0^{(L_{h-1})}}$ as follows:

$$K_{SG_0^{(L_{h-1})}} = e(r_{(0,0)}P,\ldots,r_{(0,k-1)}P, r_{(0,k+1)}P,\ldots,r_{(0,t_{SG_0^{(L_{h-1})}}-1)}P)^{r_{(0,k)}}$$

since

$$\begin{aligned} K_{SG_0^{(L_{h-1})}} &= e(r_{(0,1)}P, r_{(0,2)}P\ldots,r_{(0,t_{SG_0^{(L_{h-1})}}-1)}P)^{r_{(0,0)}} \\ &= e(r_{(0,0)}P, r_{(0,2)}P\ldots,r_{(0,t_{SG_0^{(L_{h-1})}}-1)}P)^{r_{(0,1)}} \\ &= e(r_{(0,0)}P,\ldots,r_{(0,k-1)}P, r_{(0,k+1)}P,\ldots,r_{(0,t_{SG_0^{(L_{h-1})}}-1)}P)^{r_{(0,k)}} \\ &= \ldots = e(P,P,\ldots,P)^{r_{(0,0)}r_{(0,1)}\cdots r_{(0,t_{SG_0^{(L_{h-1})}}-1)}} \end{aligned}$$

According to similar methods mentioned above, other subgroups at the lowest layer L_{h-1} can obtain their subgroup keys respectively.

*Phase*2 : Each subgroup member $U_{(j,k)}^{(L_m)}$ of subgroup $SG_j^{(L_m)}$ at the layer L_m ($m \in [h-2,\ldots,0]$) will run the subgroup key agreement protocol similar to Phase1 to obtain its subgroup key $K_{SG_j^{(L_m)}}$. Because the subgroup member $U_{(j,k)}^{(L_m)}$ is also the subgroup controller $U_{SG_k}^{(L_{m+1})}$ of subgroup $SG_k^{(L_{m+1})}$ at the next lower layer L_{m+1}, so each subgroup member $U_{(j,k)}^{(L_m)}$ possesses the subgroup key $K_{SG_k^{(L_{m+1})}}$ of the subgroup $SG_k^{(L_{m+1})}$ at the layer L_{m+1}. Therefore in the phase2 each subgroup

member $U_{(j,k)}^{(L_m)}$ will use the hash value of the subgroup key $K_{SG_k^{(L_{m+1})}}$ as its private key to compute the subgroup key $K_{SG_j^{(L_m)}}$, instead of choosing a new integer randomly. This phase continues until all subgroup members $U_{(0,k)}^{(L_0)}$ in the subgroup $SG_0^{(L_0)}$ obtain the final group key $K = K_{SG_0^{(L_0)}}$.

*Phase*3 : Each subgroup member $U_{(0,k)}^{(L_0)}$ at the highest layer L_0 encrypts the final group key $K = K_{SG_0^{(L_0)}}$ using the subgroup key $K_{SG_k^{(L_1)}}$ of subgroup $SG_k^{(L_1)}$ and broadcast $\{K\}_{K_{SG_k^{(L_1)}}}$ to its respective subgroup $SG_k^{(L_1)}$ at the layer L_1.

At the layer L_m for $m \in [1,\ldots, h\text{-}2]$ each subgroup $U_{(j,k)}^{(L_m)}$ first decrypts the encrypted message received from its subgroup controller $U_{SG_j}^{(L_m)}$ who belongs to a subgroup in the layer $L_{m\text{-}1}$ and concatenated its subgroup key $K_{SG_j^{(L_m)}}$ follow the decrypted message. Because the subgroup member $U_{(j,k)}^{(L_m)}$ is also the subgroup controller $U_{SG_k}^{(L_{m+1})}$ of subgroup $SG_k^{(L_{m+1})}$ at the next lower layer L_{m+1}, so each member $U_{(j,k)}^{(L_m)}$ can encrypts the message using the subgroup key $K_{SG_k^{(L_{m+1})}}$ and broadcast the encrypted message to its respective subgroup $SG_k^{(L_{m+1})}$ at the layer L_{m+1}. This process will end when all subgroup members $U_{(j,k)}^{(L_{h-1})}$ at the lowest layer $L_{h\text{-}1}$ have obtained the final group key and the corresponding subgroup keys by decrypting the encrypted message received from its subgroup controller $U_{(j,k)}^{(L_{h-1})}$ who belongs to a subgroup in the layer $L_{h\text{-}2}$.

D. Re-Keying Operations

1. Member joins. When a new node U_{n+1} wants to join the group and there existing a subgroup $SG_j^{(L_{h-1})}$ contains less than $ub_{L_{h-1}}$ subgroup members at the lowest layer $L_{h\text{-}1}$, then the node U_{n+1} join this subgroup $SG_j^{(L_{h-1})}$. This subgroup will run the subgroup key agreement protocol to get the new subgroup key. And the corresponding subgroups at the layer L_m ($m \in [h\text{-}2,\ldots,0]$) above this subgroup will also run the subgroup key agreement protocol to update their corresponding subgroup key. After updating of group key K, all the new subgroup keys and the new group key K will be broadcasted down the layers to corresponding subgroup members securely using symmetric key cryptography.

If all the subgroup $SG_j^{(L_{h-1})}$ at the lowest layer $L_{h\text{-}1}$ contains $ub_{L_{h-1}}$ subgroup members, then construct the layer-cluster group model again and run the *LCML* protocol to establish and allocate new group key.

2. Member leaves. Let $U_{(j,k)}^{(L_{h-1})}$ be a subgroup member who wants to leave the subgroup $SG_j^{(L_{h-1})}$. In this subgroup, other subgroup members $U_{(j,t)}^{(L_{h-1})}$ ($t \neq k$), after receiving the leaving requirement from subgroup member $U_{(j,k)}^{(L_{h-1})}$, will delete the information of subgroup member $U_{(j,k)}^{(L_{h-1})}$ and run the subgroup key agreement protocol again to refresh the subgroup key $K_{SG_j^{(L_{h-1})}}$. Moreover, all the corresponding subgroup keys above this subgroup will be updated. And then all new keys will be broadcasted down to corresponding subgroup member securely.

IV. SECURITY ANALYSIS

The security of protocol *LCML* is based on decisional multi-linear Diffie-Hellman assumption and the security of the symmetric key encryption scheme.

In the subgroup key agreement protocol, every subgroup member broadcast its public key to the subgroup. So every subgroup member can obtain other subgroup member's public keys in the subgroup to compute the subgroup key by using DMDH assumption. Obviously the security of subgroup key agreement is based on DMDH assumption. Assume that the adversary want to get the subgroup key, he need to extract the subgroup member's private key from its public key in which is equivalent to solving an instance of discrete logarithm problem. Obviously the adversary can not obtain any private key of subgroup member then it can not obtain the subgroup key.

In the group key agreement process the subgroup member $U_{(j,k)}^{(L_m)}$ of subgroup $SG_j^{(L_m)}$ at L_m ($m \in [h\text{-}2,\ldots,0]$) uses the hash value of the subgroup key $K_{SG_k^{(L_{m+1})}}$ as its private key to run the subgroup key agreement protocol. This process will end when the subgroup at the highest layer L_0 has computed its subgroup key K. Obviously, based on the security of subgroup key agreement protocol an adversary will not be able to obtain the subgroup key $K_{SG_k^{(L_{m+1})}}$ and he will not be able to get the subgroup key $K_{SG_j^{(L_m)}}$ at L_m too.

In the protocol *LCML*, the final group K and the respective corresponding subgroup keys are encrypted and broadcasted down the layers to corresponding subgroup members using symmetric key cryptography. If the symmetric key encryption scheme is secure against chosen ciphertext attacks, then the adversary will not be able to obtain the group key unless he is able to successfully break the secure encryption scheme.

Theorem1. Protocol *LCML* provides forward secure and backward secure.

Proof. *forward secure*: Let A be an active adversary who has been a member of some subgroup during some previous time period. Now assume the adversary A tries to read the subgroup traffic after he has left. A has

with it the old group key and a series of corresponding subgroup keys. However, he can not read the subgroup traffic, since the protocol updates group key and all corresponding subgroup keys that A previously knows. So the adversary A can not read the subgroup traffic after he has left unless he join the subgroup again which provide the forward secure.

backward secure: In *LCML* protocol, when A joins a subgroup, this subgroup and all the corresponding subgroup above this subgroup will update their subgroup keys so the adversary A cannot derive any previous subgroup key and previous group key before he join the subgroup. Then the adversary A can not read the previous subgroup traffic before he joins the subgroup since he does not know any previous subgroup keys and the group key. According to the analysis above, the protocol *LCML* provides the backward secure.

V. COMPLEXITY ANALYSIS

We compared the computational overhead and communication costs of our proposed protocol with TGDH, GDH and C-H [12]group key agreement protocols. In Table 1, the computational overhead refers to the number of modular exponentiations and the number of DMDH operations required to compute the final group key and the communication cost is represented by the number of messages transmitted and received. Furthermore, as mentioned in [17, 18], compared with the computational overhead of symmetric key cryptography, the computational overhead of modular exponentiations are several orders of magnitude higher. So we neglect the computational complexity of symmetric key encryption/decryption as compared to modular exponentiat -ions. In table 1, the notation c refers to the number of members in each subgroup in protocol C-H and the notation h is presented the number of layers or the height of the tree. Furthermore, the users refer to the all subgroup members in each subgroup across the lowest layer $L_{h\text{-}1}$.

For TGDH protocol, it requires each user to perform $2h$ modular exponentiations, send and receive h messages respectively. In the TGDH protocol the height of the tree is $h = \log_2 n$ however the number of layers in the *LCML* protocol is $h \le \log_{lb_{\min}} n_1$. For example, for a group size n=2^{20}=1048576, the height of the TGDH tree is $h = \log_2 2^{20} = 20$ while the number of layers in *LCML* protocol is $h \le \log_8 2^{15} = 5$ with n_1= 32768, $lb_{\min}$ =8 and $ub_{\max} > 32$. Compared with TGDH protocol, the proposed protocol *LCML* reduces the computation cost and the messages need to send.

As shown in Table 1, the C-H protocol requires each user to perform three modular exponentiations, transmit two messages and receives c+2 messages. A subgroup member in the $SG_j^{(L_m)}$ ($m \in [1,\ldots,h\text{-}2]$) has to compute 3($h$-$m$) modular exponentiations, transmit 3h-3m-1 messages and receive (h-m)(c+1)+1 messages respectively. While a subgroup member in the $SG_0^{(L_0)}$ requires to perform 3h modular exponentiations, send 3h-1 messages and receive h(c+1) messages. Compared with TGDH, the C-H protocol scales down the number of exponentiations and the transmitted messages, but increases the number of received messages. In our proposed protocol, each user requires to perform one modular exponentiation and one DMDH operation, transmit one message and receive less than $ub_{L_{h-1}}$ messages. A subgroup member in the $SG_j^{(L_m)}$ ($m \in [1,\ldots,h\text{-}2]$) has to compute h-m modular exponentiations and h-m DMDH operations, transmit 2(h-m)-1 messages and receive less than $ub_{\max}(h-m)$ messages respectively. While a subgroup member in the $SG_0^{(L_0)}$ requires to perform h modular exponentiations and h DMDH operations, send 2h-1 messages and receive less than $(ub_{\max})h-1$ messages. Obviously, our proposed protocol *LCML* scales down the communication costs as compared to the C-H protocol.

TABLE 1

COMPUTATION AND COMMUNICATION COST

		Exp.	DM DH	Messages sent	Messages to be received
TGDH		2h	0	h	h
GDH	U_1-$U_{n\text{-}2}$	3	0	2	3
	$U_{n\text{-}1}$	2	0	1	2
	U_n	n	0	1	n-1
C-H	$SG_0^{(L_0)}$	3h	0	3h-1	h(c+1)
	$SG_j^{(L_m)}$ $m \in [1,\cdots,h\text{-}2]$	3(h-m)	0	3(h-m)-1	(h-m)(c+1)+1
	users	3	0	2	c+2
LCML	$SG_0^{(L_0)}$	h	h	2h-1	$\le (ub_{\max})h-1$
	$SG_j^{(L_m)}$ $m \in [1,\cdots,h\text{-}2]$	h-m	h-m	2(h-m)-1	$\le ub_{\max}(h-m)$
	users	1	1	1	$\le ub_{L_{h-1}}$

VI. CONCLUSION

In ad hoc networks, secure group key agreement protocols play a key role. They are one of the most crucial technologies for ad hoc networks. However, most existing group key agreement protocols require either centralized key servers or expensive public key operations, which make them unsuitable for ad hoc networks. In this paper, we proposed a new group key agreement protocol based on DMDH assumption and layer-cluster group model. Compared with TGDH, GDH and C-H group key agreement protocols, the proposed protocol *LCML* improve the executing performance. So it is more suitable for ad hoc networks.

ACKNOWLEDGMENT

Thanks for the helpful discussion about the proof of security from PhD Candidate XuPeng and PhD MingHui Zheng.

REFERENCES

[1] G.Ateniese, M.Steiner and G.Tsudik. New multi-party authentication services and key agreement protocols. IEEE Journal on Selected Areas in Communications, 2000, 18(4):628-640

[2] Jason H. Li, Renato Levy, Miao Yu. A Scalable Key Management and Clustering Scheme for Ad Hoc Networks. In: INFOSCALE'06, 2006, 1-10

[3] W.Diffie, M.Hellman. New directions in cryptography. IEEE Transactions on Information Theory, 1976, 22: 644-654

[4] I.Ingemarsson, D.T.Tang and C.K.Wong, A conference key distribution system. IEEE Transactions on Information Theory, 1982, 28(5): 714-720

[5] Steiner M, Tsudik G., Waidner M. Differ-Hellman key distribution extended to group communication. In: Usenix Conference on Computer and Communications Security, ACM Press, 1996, 31-37

[6] Steiner M, Tsudik G, Waidner M. DLIQUES: A New Approach to Group Key Agreement. In: Proceeding of the 18th International Conference on Distributed Computing Systems, 1998, 380-387

[7] Steiner M, Tsudik G, Waidner M. Key Agreement in Dynamic Peer Groups. IEEE Transactions on Parallel and Distributed Systems, 2000, 11(8): 769-780

[8] A.Perrig. Efficient collaborative key management protocols for secure autonomous group communication. International Workshop on Cryptographic Techniques and Electronic Commerce, 1999, 192-202

[9] Kim Y, Perring A, Tsudik G. Tree-Based Group Key Agreement. ACM Transaction on Information and System Security, 2004, 7(1): 60-96

[10] Kim Y, Perring A, Tsudik G. Group Key Agreement Efficient in Communication. IEEE Transaction on Computers, 2004, 53(7): 905-921

[11] S. Cho, J. Nam, S.Kim and D. Won. An Efficient Dynamic Group Key Agreement for Low-Power Mobile Devices. ICCSA'2005. 2005,3480:498-507

[12] Joseph Chee Ming Teo, Chik How Tan. Energy-Efficient and Scalable Group Key Agreement for Large Ad Hoc Networks. PE-WASUN'05, 2005, 114-121

[13] Zhang Li-Ping, Cui Guo-Hua, Yu Zhi-Gang. An Efficient Group Key Agreement Protocol for Ad Hoc Network. WICOM08, October, Dalian, China, 2008

[14] D.Boneh and A.Silverberg. Applications of multi-linear forms to cryptography, Contemporary .Mathematics, 2003, 324:71-90.

[15] Wei Wang, Jiangfeng Ma and SangJae Moon. Efficient Group Key Management for Dynamic Peer Networks. MSN2005, 2005, 3796:753-762.

[16] Zhang Li-Ping, Cui Guo-Hua et al. Group Key Agreement Protocol Based on Circular Hierarchical for Ad Hoc Network. Computer Science, 2008,35(10):61-64

[17] Carman D.W., Kruss P.S., Matt B.J. Constraints and approaches for distributed sensor network security. NAI Labs Technical Report, 2000.

[18] Trappe W, Wang Y, Liu K.J. Resource-aware conference key establishment for heterogeneous networks. IEEE/ACM Transactions on Networking, 2005, 13(1): 134-146.

Zhang Li-Ping, born in 1978, Ph.D. Her research interests include public key cryptography, provable security of cryptographic protocol and cryptography theory and practice, Ad hoc networks.
Email:carolyn321@163.com

Wang Yi, born in 1975, Master. Her research interests include information security, algorithm analysis and numeric analysis.

Queuing Analysis and Simulation of Wireless Access and End Point Systems using Fano Decoding

Khalid A. Darabkh
The University of Jordan/Computer Engineering Department, Amman, Jordan
Email: k.darabkeh@ju.edu.jo

***Abstract*— Wireless networks are a growing technology due to its ability to receive data in areas where it is very hard to plug-in using wires. TCP Reno assumes in his congestion algorithms that the packet loss is the major cause of network congestion. In wireless networks, this is not correct because having a high bit error rate leads also to a packet loss. Link layer approach is one of the most efficient proposed solutions to maintain TCP over wireless networks. For example, having hybrid ARQ type 1 with Fano decoding, which is an error correction technique, is very appropriate and is of concern in wireless networks due to its capability of offering decoding steps, which are dependent to the channel state. In this paper, we propose a novel queuing model to see the effect of employing Fano decoding on the buffer of wireless access or end points since it is a very effective network parameter and cannot be neglected. Our queuing model is concerned not only about those departed packets after being decoded using Fano algorithm, but also the way packets arrive to the wireless access or end point systems. An analytical study has been conducted to derive a general form expression for the average number of packets residing in the system's buffer. On the other hand, a simulation study using programming language has been performed to validate our analytical results.**

***Index Terms*— Expected queue size, Fano decoding, partial decoding, queuing analysis and simulation, TCP performance.**

I. INTRODUCTION

Wireless networks have become very popular nowadays for a lot of reasons, most importantly the flexibility to set up a network where it is difficult using wired networks. TCP (transmission control protocol) is the main protocol to deliver the service with completeness, no errors, and fast delivery enhancements. It was originally designed for wired networks. TCP Reno is the typical congestion control mechanism that is implemented by end hosts. TCP congestion schemes limit the connection throughput by interpreting the loss of the packet as an indication of congestion [1], [2]. This is correct for wired networks where the transmission media is noiseless, but in wireless networks the media is noisy by nature [3]. Even low BER (bit-error-rate) may generate packet loss as a result of DUP ACKs (duplicate acknowledgments) and timeout. Consequently, the sender, which is not wireless-aware, shrinks its congestion window due to a congestion mark understanding. Hence, multiple corrupted packets may decrease the sending rate of the TCP sender dramatically.

The following are the major approaches that have been done to improve TCP over wireless networks [4]: Split mode, Snoop protocol, and link layer. In the split mode, there are two separate TCP streams. In this approach, a certain packet may be acknowledged by the access point without reaching its destination. This violates the semantic of TCP as being an end-to-end scheme [2]. This further leads to extra overhead for wireless access points by adding a new layer (i.e., transport layer) to manage that. In snooping protocol, the idea of ELN (explicit loss notification) has been introduced [5]. This is done by activating an ELN flag in the TCP header to distinguish DUP ACKs received by sender. It is set when these DUP ACKs come as a result of a noisy environment. In this case, the congestion algorithms will not be invoked and the congestion window will be kept large. It works efficiently only when the error rate is low [6]. Moreover, it has a limitation of possibly false notification, especially when there are multiple packets lost in a window [6]. In the last approach, a set of algorithms have been associated with ARQ (automatic repeat request) (i.e., hybrid ARQ of type 1) to enhance the TCP performance over wireless networks through the MAC (media access control) sub-layer of wireless access and end points and most importantly the convolutional decoding algorithms [4]. These decoding algorithms are considered error detection and correction algorithms and are of interest in wireless networks much more than the CRC (cyclic redundancy check) error detection technique. In this paper, we investigate the last approach from a queuing point of view. In other words, we study the effect of employing decoding algorithms on the buffer of access or end points since it is a very important network parameter and may affect the network performance badly if it is not taken into consideration.

Convolutional codes are a category of channel coding that are described by adding extra bits to the original data for bit flipping prevention over a noisy channel [7]. A convolutional coder includes L-stage shift register and x codeword blocks modulo-2 adders [8]. Hence, it has a constraint length of L. There are two important decoding (error detection and correction) algorithms for convolutional codes, the maximum-likehood decoding

Manuscript received January 31, 2010; revised April 18, 2010; accepted May 11, 2010.

doi:10.4304/jcm.5.7.551-561

(Viterbi's algorithm) and sequential decoding [8]. These algorithms increase the ratio of good packets to the total number of received packets. Moreover, they may decrease overall message latency since correcting corrupted packets helps to reduce the number of retransmissions and accordingly network congestion. Consequently, the TCP performance over wireless networks is improved. Convolutional coding with Viterbi decoding is very popular and is considered the predominant FEC (forward error correction) technique that is widely employed in satellite and mobile communications [8], [9]. It is characterized by affording a fixed decoding time. Therefore, it is unable to provide faster decoding for a received codeword sequence that contains fewer errors.

Fano decoding was developed by Fano [9] and came as an efficiency improvement to a sequential decoding algorithm that was developed by Wozencraft [9]. Fano algorithm has advantages over Viterbi decoding algorithm because it is operated at a variable decoding rate and has computational and storage requirements that grow linearly as a function of the constraint length rather than exponentially as the case of Viterbi decoding [10]. Moreover, it has been proven that under high SNR (signal-to-noise ratio), the Fano decoding consumes less power than the Viterbi algorithm [11]. Actually, the complexity of Fano algorithm becomes dependent on the channel state (noisy or pure) [10], [12], and [13]. The brief description of this algorithm is as follows [10]: Once the codeword sequence is received by the decoder, a comparison is made with the codeword allowed by the decoder according to the encoder state diagram. Each codeword sequence is divided into groups where each group consists of *m* digits. Fano algorithm works in a way similar to the code tree. It chooses the path for which the sequence group has the shortest hamming distance (i.e., difference in bits between encoder output and the received sequence group is minimum). This process is repeated until groups end. If a lot of errors appear through this process then it becomes impossible to have a match for the received sequence. This is an indication that the wrong path is chosen. Thus, back and forth steps have to be done to avoid getting an accumulatively high number of errors.

It is well known that every intermediate hop has a buffer to absorb the variable packet processing rate. In this paper, we propose a new queuing model that describes packet arrivals and departures of a wireless system uses Fano decoding. The maximum decoding time has to be variable since it depends on how high the percentage of errors is. In our queuing model, the decoder is described as having not only an upper variable decoding limit (*T*), but also a lower variable decoding limit (*K*). The minimum decoding time of Fano decoding is also variable (i.e., not constant) since it depends on how low the percentage of errors is. We consider in our queuing model that the channel state is also variable. Having a Fano decoder with upper and lower operating decoding limits that are variable and adaptive with the channel state is very efficient and suitable for noisy media such as wireless networks. Furthermore, it gives generality for our system in terms of queuing theory. Our major aim of this queuing model is to find a closed form expression of the average number of packets waiting in the system's buffer. The system is assumed to be of infinite capacity. Hence, it becomes very necessary to find the average number of packets in the buffer waiting to be served so that it can be selected as a value to be operated on. It is clear that the value of system capacity (buffer size) cannot be chosen randomly since choosing too large or too small may degrade the performance of the system significantly. The average packet waiting time increases when the system capacity gets large. Thus, many retransmissions may occur due to a timeout limit from the sender side although there is no packet loss. On the other hand, many retransmissions may occur due to packets being discarded because of small system capacity. Hence, choosing a suitable buffer size may also improve TCP performance over wireless networks. Our queuing model works with not only Fano decoding, but also any other decoding algorithm of variable complexity such as Low-density parity-check and Turbo decoders. We also provide an expression of the maximum possible packet arriving probability in order to keep the decoder's buffer stable, which then is verified through the results of our general form expression of the average buffer size. Moreover, we simulate our proposed queuing model using Matlab programming to validate our analytical observations and results. To the best of our knowledge, our queuing model, analysis, simulation, and results are completely new and there is no similar work to ours.

II. ANAYTICAL STUDY

In our analytical study, we provides details about the system model, system assumptions and probability state transitions, system steady-state transition probabilities, and system analysis that are very helpful to recognize our proposed general form expression for the average queue size.

A. System Model

To model the system (decoder and queue), a discrete-time Markov chain is used. A Markov chain is a stochastic process [14] (bunch of random variables) with a very limited memory [14]. However, in our model, the time is portioned into equal time slots where at least a packet is allowed to arrive in any arbitrary time slot. Hence, Bernoulli process is the best to describe arriving packets. The arriving probability arrives with probability (λ) and does not with probability $(1-\lambda)$. The decoder can start decoding a packet after its arrival immediately on the succeeding slot if there are no packets waiting in the queue. Since Fano algorithm offers variable decoding time which is dependent to the channel state, Pareto distribution, which is a heavy-tailed distribution, is considered to be the best to describe the decoding time [15]-[17]. In this distribution, there is a parameter called (β). When the value of this parameter gets high, the PDF

(probability density function) of Pareto distribution goes to zero early. From the decoder side, this means dispatching low decoding time. On the other hand, when (β) gets low, the PDF goes to zero late. Considering decoder side, this implies dispatching high decoding time. Hence, we can take advantage of that by connecting (β) to the SNR of the channel [12], [18]. Thus, when SNR is high, this indicates we have low noise power. Therefore, low decoding time is required. We will lose the system's generality if this lower decoding time is considered to be fixed as assumed in [19]. However, when SNR is low, this implies that the noise power is high. Thus, high decoding time is needed. However, once a packet reaches that upper bound decoding limit and needs more, it is considered to be partially decoded and lost which is the same as assumption used in [19]-[21]. One way of recovering those packets is to resend them during the slots follow immediately getting partial decoding [20], [22].

B. System Assumptions and Probability State Transitions

The states of this discrete-time Markov model are represented by (*n, j*) as employed in [19], [23], where *n* describes the number of packets in the system including the packet being decoded, and *j* stands for the number of slots elapsed in decoding the packet currently in service. We assume two major notations to understand the probability state transitions shown in Fig. 1. The probability that the decoding process is completed in *j+1* slots is referred to c_{j+1} and the probability that the decoder finishes decoding in *j+1* slots given that decoding takes more than *j* slots is notated as μ_j. Thus, the conditional probability is:

$$\mu_j = \frac{c_{j+1}}{(1-F_j)}. \tag{1}$$

Where F_j refers to the distribution of decoding time and it can be represented as follows:

$$F_j = 1 - P_F(jT_r), \quad 1 \le j \le T, \tag{2}$$

Where T_r is the slot duration and assumed to be 1, *T* is the maximum decoding slots allowed for decoding. Hence,

$$P_F(j) = \Pr\{t > j\} = \left(\frac{j}{2}\right)^{-\beta}. \tag{3}$$

Note that the decoding time of Fano decoders has the Pareto distribution where β is called the Pareto parameter and it is a function of SNR [1], [20] of the wireless channel which means that if SNR gets low, then the wireless channel gets worse. However, 2 represents the minimum decoding time needed by the decoder to decode a packet.

$$F_j = \sum_{i=2}^{j} c_i, \quad j \ge 2. \tag{4}$$

To describe the behavior of Fano decoding using queuing theory with high level of generality, the Pareto parameter, packet arriving probability, maximum decoding limit, and minimum decoding limit should be variable. We have already introduced elementary analytical results for the average buffer size but for the special case where at least four decoding time slots are required to finish decoding [24]. In this work, we start first by assuming that packets need at least two decoding time slots to finish decoding (i.e., minimum decoding limit) as shown in (1) and (3). Consequently, we become able to generalize our close form expression for the average buffer size to be a function of upper variable decoding limit (*T*), channel condition (β), packet arriving probability (λ) in addition to lower variable decoding limit (*K*), which is presented in subsection *E*. We present in this work to the extent level details of how to reach that general form. It is important to state that the number of state probabilities increases when the upper variable decoding limit (*T*) increases (as shown in Fig. 1) while it is not the case for lower variable decoding limit since it has just a relation with μ_j. Therefore, the step transitions always get different when that lower limit is changed.

According to our initial assumption saying that any packet needs at least two decoding slots,

$$F_0 = 0. \tag{5}$$

$$F_1 = 0. \tag{6}$$

It can clearly be shown that,

$$1 - F_j = \prod_{i=2}^{j} (1 - \mu_{i-1}), \quad j \ge 2. \tag{7}$$

And the probability c_j is given by

$$c_j = \begin{cases} F_j - F_{j-1}, & 2 \le j \le T, \\ 0, & Otherwise. \end{cases} \tag{8}$$

C. System Steady-state Transition Equations

All states found in Fig. 1 are neither transient nor null recurrent. Thus, stationary (limiting, equilibrium, or steady-state) probabilities exist [14]. Let $p_{n,j}$ is the probability that the decoder has *n* packets in its queue including the packet being decoded which is in the j^{th} slot of decoding. Accordingly, the steady-state transition equations are given below.

$$p_{0,0} = (1-\lambda)p_{0,0} + \sum_{j=1}^{T-1} \mu_j (1-\lambda) p_{1,j}, \tag{9}$$

$$p_{0,j} = 0, \tag{10}$$

$$p_{1,0} = \sum_{j=1}^{T-1} \mu_j \left[\lambda p_{1,j} + (1-\lambda) p_{2,j}\right] + \lambda p_{0,0}, \tag{11}$$

$$p_{1,1} = (1-\lambda) p_{1,0}, \tag{12}$$

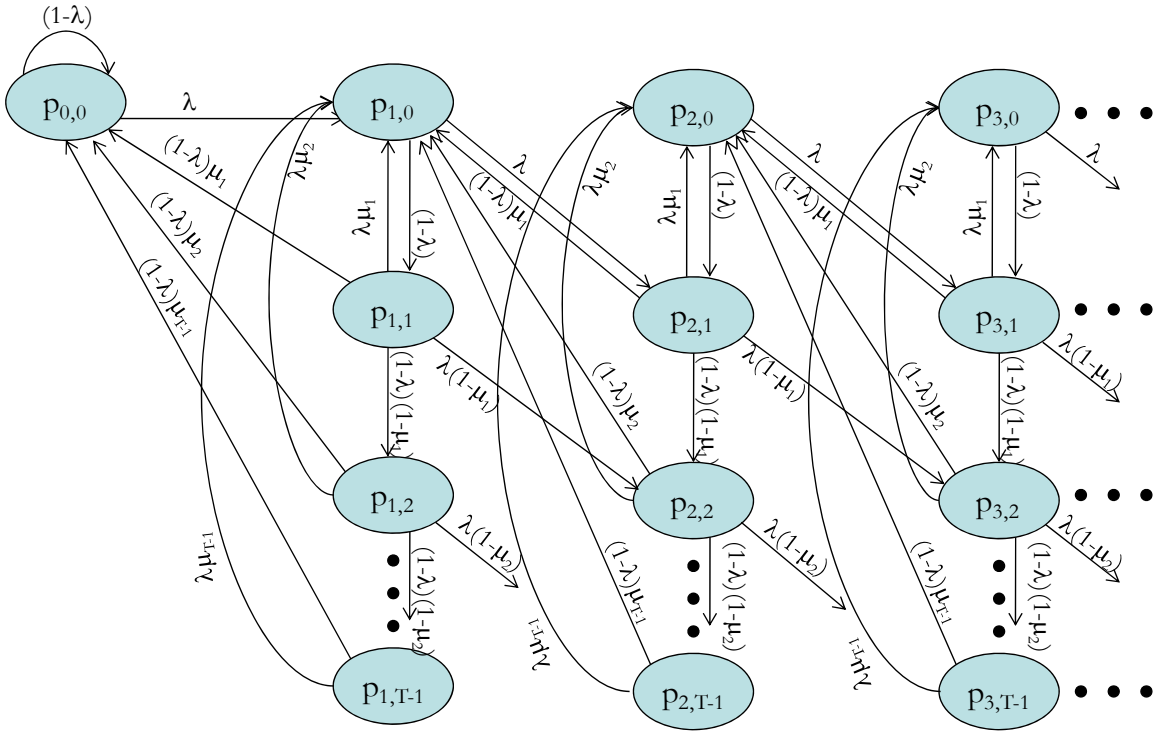

Figure 1: Probability state transitions of decoder's queue with maximum and minimum decoding limits of T and 2 respectively. The summation of outgoing links of all states is equal to one.

$$p_{1,j} = (1-\lambda)(1-\mu_{j-1})p_{1,j-1}, \quad 2 \le j \le T-1, \tag{13}$$

$$p_{n,0} = \sum_{j=1}^{T-1} \mu_j \left[\lambda p_{n,j} + (1-\lambda)p_{n+1,j}\right], \quad 2 \le n, \tag{14}$$

$$p_{n,1} = \lambda p_{n-1,0} + (1-\lambda)p_{n,0}, \quad 2 \le n, \tag{15}$$

$$p_{n,j} = (1-\mu_{j-1})\left[(1-\lambda)p_{n,j-1} + \lambda p_{n-1,j-1}\right], \quad 2 \le n, \quad 2 \le j \le T-1, \tag{16}$$

Taking into consideration that the conditional probability to finish decoding in T slots given that at least required $T-1$ slots, which refers to μ_{T-1}, is equal to one.

D. System Analysis

According to Fig. 1, dependency between states is visible. To solve the problem of discovering intelligent way to find the average number of packet in the queue, the following generating functions are defined:

$$P(z) = p_{0,0} + P_0(z) + P_1(z) + \sum_{j=2}^{T-1} P_j(z), \quad 2 \le j \le T-1. \tag{17}$$

Where,

$$P_i(z) = \sum_{n=1}^{\infty} p_{n,i} z^n, \quad 0 \le i \le T-1. \tag{18}$$

and $P(z)$ is the generating function that denotes for the number of packets in the queue waiting to get served. It is important to refer to the well known fact that $P(1) = 1$ [20], [25]. Thus, we can evaluate the average number of packets in the queue by taking the derivative of $P(z)$ at $z = 1$ [19], [20], [25], and [26]. However, referring to (18), we can write

$$P_0(z) = p_{1,0} z + \sum_{n=2}^{\infty} p_{n,0} z^n. \tag{19}$$

Substituting (11) and (14) for $p_{1,0}$ and $p_{n,0}$ into (19), we get

$$P_0(z) = \sum_{n=1}^{\infty} \sum_{j=1}^{T-1} \mu_j \left[\lambda p_{n,j} + (1-\lambda)p_{n+1,j}\right] z^n + \lambda z p_{0,0}. \tag{20}$$

After rewriting, we get

$$P_0(z) = \sum_{j=1}^{T-1} \mu_j \left[\lambda \sum_{n=1}^{\infty} p_{n,j} + (1-\lambda) \sum_{n=1}^{\infty} p_{n+1,j}\right] z^n + \lambda z p_{0,0}. \tag{21}$$

The rightmost summation can be written as

$$\sum_{n=1}^{\infty} p_{n+1,j} z^n = \frac{1}{z}\left(P_j(z) - p_{1,j} z\right). \tag{22}$$

Assuming,

$$f(z) = 1 - \lambda + z\lambda. \tag{23}$$

After finding the term $\sum_{j=1}^{T-1} \mu_j (1-\lambda) p_{1,j}$, found at (9), in terms of $p_{0,0}$ and substituting it along with (22) and (23) into (21), we find

$$P_0(z) = \frac{f(z)}{z} \sum_{j=1}^{T-1} \mu_j P_j(z) + \lambda(z-1) p_{0,0}. \tag{24}$$

Applying the fact that $\mu_{T-1} = 1$,

$$P_0(z)=\frac{f(z)}{z}\left[P_{T-1}(z)+\sum_{j=1}^{T-2}\mu_j P_j(z)\right]+\lambda(z-1)p_{0,0}. \quad (25)$$

For $P_j(z)$where j=1 we have,

$$P_1(z)=p_{1,1}z+\sum_{n=2}^{\infty}p_{n,1}z^n. \quad (26)$$

Substituting (12) and (15) into (26) yields

$$P_1(z)=(1-\lambda)p_{1,0}z+\sum_{n=2}^{\infty}\left[\lambda p_{n-1,0}+(1-\lambda)p_{n,0}\right]z^n. \quad (27)$$

We can rewrite (27) as follows

$$P_1(z)=\lambda\sum_{n=2}^{\infty}p_{n-1,0}z^n+(1-\lambda)\sum_{n=1}^{\infty}p_{n,0}z^n. \quad (28)$$

The leftmost summation can be simplified as

$$\sum_{n=2}^{\infty}p_{n-1,0}z^n=z\sum_{n=2}^{\infty}p_{n-1,0}z^{n-1}=zP_0(z). \quad (29)$$

Hence,

$$P_1(z)=f(z)P_0(z). \quad (30)$$

Deriving an expression for j>1,

$$P_j(z)=p_{1,j}z+\sum_{n=2}^{\infty}p_{n,j}z^n. \quad (31)$$

Substituting (13) and (16) into (31) gives

$$\begin{aligned}P_j(z)&=(1-\lambda)(1-\mu_{j-1})p_{1,j-1}z\\&+(1-\mu_{j-1})\sum_{n=2}^{\infty}\left[(1-\lambda)p_{n,j-1}+\lambda p_{n-1,j-1}\right]z^n.\end{aligned} \quad (32)$$

After rearranging, we get

$$\begin{aligned}P_j(z)&=(1-\lambda)(1-\mu_{j-1})\sum_{n=1}^{\infty}p_{n,j-1}z^n\\&+\lambda(1-\mu_{j-1})\sum_{n=2}^{\infty}p_{n-1,j-1}z^n.\end{aligned} \quad (33)$$

Utilizing the same principle used in (22) and (29), we get

$$P_j(z)=(1-\mu_{j-1})f(z)P_{j-1}(z). \quad (34)$$

Applying (7), we recursively get

$$P_j(z)=(1-F_j)f(z)^jP_0(z). \quad (35)$$

Substituting (35) into (24) and applying (1) yields

$$P_0(z)=\frac{q(z)}{z}P_0(z)+\lambda(z-1)p_{0,0}. \quad (36)$$

Where,

$$q(z)=\sum_{j=1}^{T-2}c_{j+1}f(z)^{j+1}+(1-F_{T-1})f(z)^T. \quad (37)$$

Finally, we get

$$P_0(z)=\frac{\lambda z(z-1)}{z-q(z)}p_{0,0}=y(z)p_{0,0}. \quad (38)$$

After substituting (30), (35), and (38) into (17), we get

$$P(z)=\left[1+y(z)+f(z)y(z)+\sum_{j=2}^{T-1}(1-F_j)f(z)^jy(z)\right]p_{0,0}. \quad (39)$$

We can get the value of $p_{0,0}$ after applying the fact that $P(1)=1$ on (17). Hence, we get

$$p_{0,0}=\frac{1}{\left[1+y(1)+f(1)y(1)+\sum_{j=2}^{T-1}(1-F_j)f(1)^jy(1)\right]}, \quad (40)$$

Where,

$$y(1)=\lim_{z\to1}\frac{\lambda z(z-1)}{z-q(z)}, \quad (41)$$

Substituting z=1 into (23) and (37), and considering (4), we get

$$q(1)=1, \quad (42)$$

Hence,

$$y(1)=\lim_{z\to1}\frac{(2z-1)\lambda}{1-\lambda\left(\sum_{j=1}^{T-2}(j+1)c_{j+1}f(z)^j+T(1-F_{T-1})f(z)^{T-1}\right)}, \quad (43)$$

After applying L'Hospital's law just once, we conclude

$$y(1)=\frac{\lambda}{1-\lambda\tau}, \quad (44)$$

Where,

$$\tau=\sum_{j=1}^{T-2}(j+1)c_{j+1}+T(1-F_{T-1}), \quad (45)$$

Thus, after substituting (44) into (40), we finally conclude

$$p_{0,0}=\frac{1}{\left[1+\frac{\lambda(2+\eta)}{1-\lambda\tau}\right]}, \quad (46)$$

Where,

$$\eta=\sum_{j=2}^{T-1}(1-F_j). \quad (47)$$

The average number of packets in the decoder's queue is found by taking the derivative of (39) and then substituting for $z=1$. Hence, we get

$$P'(1)=\left[\begin{aligned}&y'(1)+f'(1)y(1)+f(1)y'(1)\\&+\sum_{j=2}^{T-1}(1-F_j)\left[jf(1)^{j-1}f'(1)y(1)+f(1)^jy'(1)\right]\end{aligned}\right]p_{0,0}, \quad (48)$$

After taking the derivative of $y(z)$ found at (38), we get

$$y'(1)=\lim_{z\to1}\lambda\frac{[z-q(z)](2z-1)-[1-q'(z)]z(z-1)}{[z-q(z)]^2}, \quad (49)$$

By using L'Hospital's law twice, we get

$$y'(1)=\lambda\left(\frac{q''(1)+2[1-q'(1)]}{2[1-q'(1)]^2}\right), \quad (50)$$

Taking the derivative of (37) and then substituting for z=1 when using (23)

$$q'(1)=\lambda\left(\sum_{j=1}^{T-2}c_{j+1}(j+1)+T(1-F_{T-1})\right), \quad (51)$$

Also for the second derivative when applying further the fact that $f''(z)=0$

$$q''(1)=\lambda^2\left(\sum_{j=1}^{T-2}(j+1)jc_{j+1}+T(T-1)(1-F_{T-1})\right), \quad (52)$$

Hence, we finally get

$$y'(1)=\lambda\left(\frac{\lambda^2\bar{\tau}+2[1-\lambda\tau]}{2[1-\lambda\tau]^2}\right), \quad (53)$$

Where,

$$\bar{\tau}=\sum_{j=1}^{T-2}j(j+1)c_{j+1}+T(T-1)(1-F_{T-1}). \quad (54)$$

Substituting (44), (46), and (53) into (48) leads to the final expression of the average number of packets in the buffer

$$P'(1)=\frac{(1-\lambda\tau)\left[y'(1)(2+\eta)+\lambda y(1)\left(1+\bar{\eta}\right)\right]}{1+\lambda(2+\eta-\tau)}, \quad (55)$$

Where,

$$\bar{\eta}=\sum_{j=2}^{T-1}j(1-F_j). \quad (56)$$

E. Average Buffer Size with Variable Lower Bound Decoding Limit

Due to the nature of wireless links of being uncontrollable and noisy (i.e., high BER) because of many different conditions such as multipath interference, urban obstacles, mobility of wireless end points, large moving objects, as well as weather conditions, it becomes interesting to derive further a new general form of average number of packets in the buffer of a decoder that is also bounded by variable K which denotes to minimum decoding limit. With refereeing to our analysis in subsection *D* (where *K*=2) and our previous analysis [24] (for *K*=4), we conclude that (39) will be modified to include any value of *K* as follows,

$$P(z)=\left[\begin{array}{l}1+y(z)+f(z)y(z)+\ldots+f(z)^{K-1}y(z)\\+\sum_{j=K}^{T-1}(1-F_j)f(z)^j y(z)\end{array}\right]p_{0,0}. \quad (57)$$

The term $(1-F_j)$ will be just associated with $P_j(z)$ when $j\geq K$ (for details, see how to derive (35)). When applying the fact that $P(1)=1$, we can get after updating (40) according to our previous modified equation (i.e., (57)) the following,

$$p_{0,0}=\frac{1}{\left[1+\frac{\lambda(K+\eta)}{1-\lambda\tau}\right]}, \quad (58)$$

Where,

$$\eta=\sum_{j=k}^{T-1}(1-F_j), \quad (59)$$

Finding $P'(1)$ from (57), we get the general expression for the average buffer size with complete variable parameters,

$$P'(1)=\left[y'(1)(K+\eta)+\lambda y(1)\left(K(K-1)-\sum_{i=1}^{K-1}i+\bar{\eta}\right)\right]p_{0,0}, \quad (60)$$

Where,

$$\tau=\sum_{j=1}^{T-K}(j+K-1)c_{j+K-1}+T(1-F_{T-1}), \quad (61)$$

$$y'(1)=\lambda\left(\frac{\lambda^2\bar{\tau}+2[1-\lambda\tau]}{2[1-\lambda\tau]^2}\right), \quad (62)$$

$$\bar{\tau}=\sum_{j=1}^{T-K}(j+K-1)(j+K-2)c_{j+K-1}+T(T-1)(1-F_{T-1}), \quad (63)$$

$$\bar{\eta}=\sum_{j=K}^{T-1}j(1-F_j), \quad (64)$$

And,

$$y(1)=\frac{\lambda}{1-\lambda\tau}. \quad (65)$$

The distribution of decoding time becomes as,

$$F_j=\begin{cases}\sum_{i=K}^{j}c_i, & j\geq K,\\ 0 & Otherwise.\end{cases} \quad (66)$$

Moreover,

$$1-F_j=\prod_{i=K}^{j}(1-\mu_{i-(K-1)}), \qquad j\geq K. \quad (67)$$

III. ANAYTICAL RESULTS

Fig. 2 shows the effect of operating different upper bound decoding limits, packet arriving probabilities, as well as channel conditions on the buffer size when working on a value of two for lower bound decoding limit (notice the *semilogy* shape adapted). However, the increase in the average number of packets, when packet arriving probability increases, is noticed. This can be explained by more arriving packets than the buffer can serve. The upper bound decoding limit has a noticed impact on the average number of packets in the buffer. Once this limit becomes larger, the number of packets waiting for decoding gets also larger. For example, the average buffer size, when $\lambda=0.0489$, is 528.3, 5.43, and 1.324 for $T=90, 60,$ and 30 respectively (given $\beta=0.6$). When comparing plots (a) with (b), the effect of changing the channel condition (β) is shown. The decrease in β means the channel gets worse. Thus, higher decoding time is needed. Consequently, all packets may reach the upper bound decoding limits and this leads to increase in the average number of packets waiting in the buffer. As values selected from Fig. 2 to verify that, the average buffer size, when $\lambda=0.2084$, is 333 and 5.622 for $\beta=0.2$ and 0.6 respectively (given $T=5$).

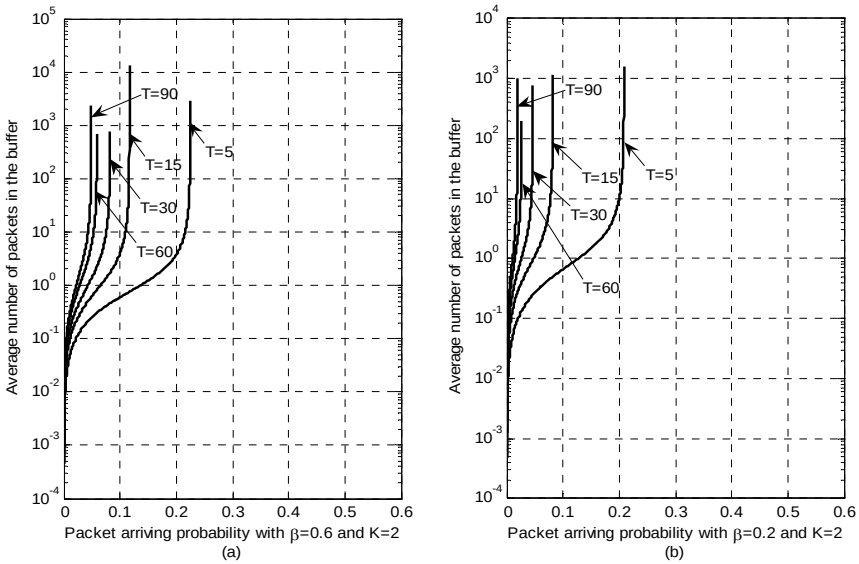

Figure 2: Average number of packets in the buffer versus packet arriving probability (λ) for different T and β. (a) $\beta = 0.6$. (b) $\beta = 0.2$.

Figs. 3 and 4 show the average number of packets in the decoder's buffer for different lower bound decoding limit (K), upper bound decoding limit (T), channel condition (β), and packet arriving probability (λ). Our major goal is to see the effect of changing K on the average buffer size for different values of T, λ, and β. The increase in the average buffer size is clearly seen as long as K increases for fixed values of T, λ, and β. This is expected since increasing K leads to having more packets in the buffer waiting for service. The decrease in β as seen in Fig. 3 causes a noticed increase in the average buffer size for all values of K. Table I provides values for the average buffer size when choosing different values of λ to be compared in both subplots of Fig. 3 under different K. As a summary, the average buffer size increases for any value of K when β decreases. In Fig. 4, the effect of decreasing T which leads to a decrease in the average buffer size is noticed when compared to Fig. 3 for the same values of K, λ, and β. Table I also presents values for the average buffer size for different values of T (60 and 15). For fixed values of λ and β, the average buffer size decreases as T decreases for all values of K.

There is one more observation is shown in Table I (also can be extracted from Figs. 3 and 4). The value of average buffer size is 365.3 when $(T = 60, \beta = 1.7)$ and 6.304 when $(T = 15, \beta = 1.7)$ while the value of average buffer size is 1.744 when $(T = 60, \beta = 2.7)$ and 1.591 when $(T = 15, \beta = 2.7)$. Does that mean, for $\lambda = 0.19330$ and $K = 2$, we should employ a buffer size near 7 for $\beta = 1.7$ and 2 for $\beta = 2.7$ regardless of the value of T? The answer is yes for the case of $\beta = 2.7$ but is no for $\beta = 1.7$. It is yes because there is no much difference obtained in the average buffer size and this is a good indication that the channel is not so noisy because of relatively high value of β. For the other part of the answer, which is no, and that is because if we select 7 when T=60, then many packets will get partially decoded and accordingly will be considered to be lost and need to be retransmitted since the buffer size is too small. This is an indication that the channel becomes so noisy since β is low. Thus, we can summarize that the value of T becomes less sensitive to the change of average buffer size when β gets larger.

It is shown from Figs. 2, 3, and 4 that the average number of packets goes to infinity after reaching certain limit of packet arriving probability (λ_{max}). The increment chosen for λ in Figs. 2, 3, and 4 is 0.00001 just to be able to find exactly the value of λ_{max}. Table II provides these limits for Figs. 3 and 4. As shown in this table, $\lambda_{\max}$ is decreased when K increases for fixed values of T and β. On the other hand, λ_{max} increases as T decreases for fixed values of K and β. Lastly, λ_{max} increases when channel condition gets better for fixed values of T and K.

One important comment about the values of λ_{max}, when $T = 15$ and $K = 12$, which are 0.06829 and 0.06913 for β =1.7 and β =2.7 respectively. These values are the closest pair found in Table II. Does that mean that the values of the average buffer size around these values of λ_{max} do not differ much? We would like to state that although these values seems to be close to each other, it is necessary to know that when λ_{max} gets low, the average buffer size will differ with just a small increase of packet arriving probability (λ). For example, when $\lambda = 0.06826$ (a bit lower than λ_{max} for both cases), the average buffer size is about 897.7 and 37.1 for β =1.7 and β =2.7 respectively.

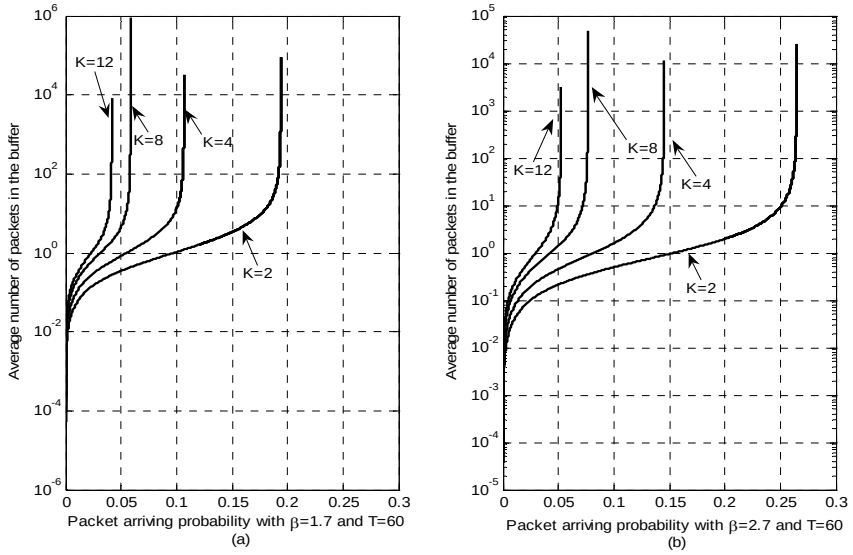

Figure 3: Average number of packets in the buffer versus packet arriving probability including fixed $T = 60$ and different working K and β.

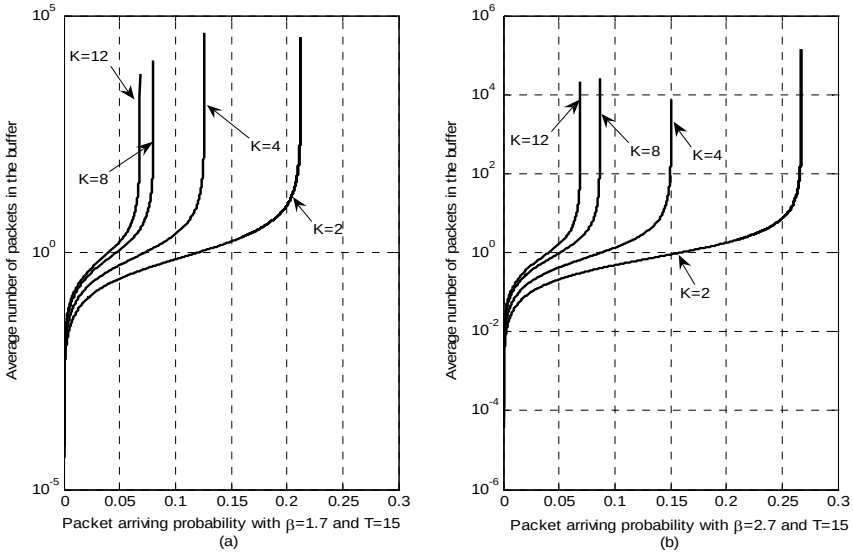

Figure 4: Average number of packets in the buffer versus packet arriving probability including fixed $T = 15$ and different working K and β.

TABLE I. AVERAGE NUMBER OF PACKETS IN THE BUFFER FOR VARIOUS SYSTEM PARAMETERS.

		$\beta = 1.7, T = 60$	$\beta = 2.7, T = 60$	$\beta = 1.7, T = 15$	$\beta = 2.7, T = 15$
Packet arriving probability (λ)	Lower bound decoding limit (K)	Average buffer size	Average buffer size	Average buffer size	Average buffer size
0.19330	2	365.3	1.744	6.304	1.591
0.10640	4	799.2	1.899	3.242	1.547
0.05813	8	196.3	2.186	1.616	1.318
0.04151	12	611.2	2.551	1.048	1.022

TABLE II. UPPER LIMIT OF PACKET ARRIVING PROBABILITY (λ_{max}) FOR VARIOUS SYSTEM PARAMETERS.

	$\beta = 1.7, T = 60$	$\beta = 2.7, T = 60$	$\beta = 1.7, T = 15$	$\beta = 2.7, T = 15$
Lower bound decoding limit (K)	λ_{max}	λ_{max}	λ_{max}	λ_{max}
2	0.19380	0.26470	0.21220	0.26730
4	0.10650	0.14530	0.12610	0.15050
8	0.05834	0.07646	0.08057	0.08678
12	0.04155	0.05232	0.06829	0.06913

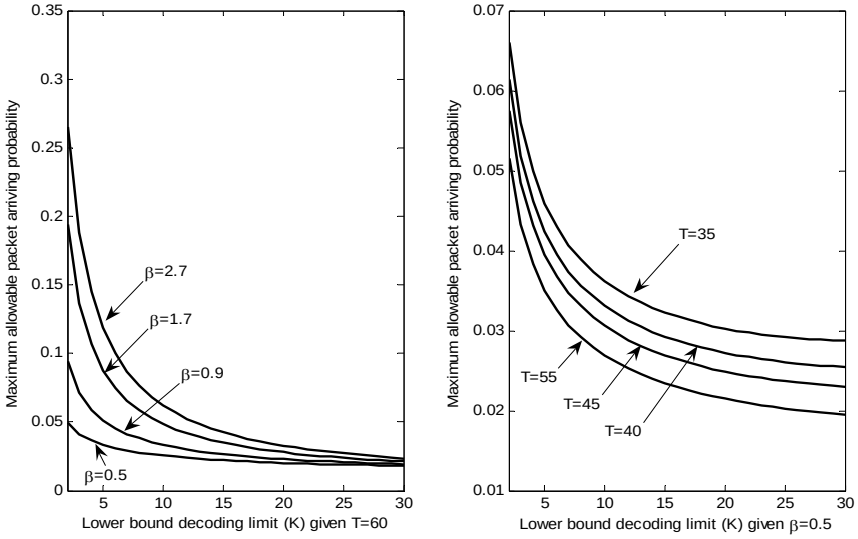

Figure 5: Maximum allowable packet arriving probability versus lower bound decoding limit (K) for different values of T and β.

To illustrate the case of existing λ_{max}, we can refer to load or system utilization which is defined in general [25] as the result of multiplying packet arrival rate with packet service rate for any system (server with its buffer). Also it is very important that the result of this product should be less than or equal to one in order to keep the buffer stable [14], [25]. In our proposed model, the system utilization is the packet arriving probability times the average decoding time, which is $\lambda\tau$. Hence, we conclude that there is λ_{max} which equals to $(1/\tau)$ for system stability. Consequently, when $\lambda > \lambda_{max}$, then the average queue size will go to infinity and this is what happens in Figs. 2, 3, and 4. The expression for average decoding time (τ) is found at (61).

Fig. 5 shows the maximum allowable packet arriving probability in order to preserve the stability of the system versus lower bound decoding limits. The decrease in λ_{max} is noticed when K increases. This indicates that the buffer gets larger. For a fixed value of K and T, the decrease in λ_{max} is seen as long as β decreases. Also, it is shown that λ_{max} is decreased, for a fixed value of K and β, when T increases. Hence, we can see that there is an effect of changing β and T on the buffer size by getting, as a result, various values of λ_{max}. Moreover, the results just explained of Fig. 5 (a) and (b) totally agree with our previous results and explanations about Figs. 2, 3, and 4. Table III is done to prove that through providing real numbers. It can be clearly seen that the results of this table, which are taken by applying the expression $1/\tau$, for λ_{max} are completely the same as the values found in Table II, which are obtained when applying our general form for the average buffer size, for λ_{max} with the same selected values of β, K, and T.

TABLE III.
MAXIMUM ALLOWABLE PACKET ARRIVING PROBABILITY FOR FIXED T AND DIFFERENT K AND β.

	$\beta = 1.7, T = 60$	$\beta = 2.7, T = 60$
Lower decoding limit (K)	λ_{max}	λ_{max}
2	0.19380	0.26470
4	0.10650	0.14530
8	0.05834	0.07646
12	0.04155	0.05232

IV. SIMULATION RESULTS

The simulation of the previously modeled system is done through Matlab. A software-based approach is considered in order to validate our analytical observations. Throughout the duration of the simulation, Bernoulli RNG (random number generator) is invoked to simulate packets' arriving process as well as Pareto RNG which is programmed through utilizing the approach of inverse cumulative distribution function [27] to simulate packets' decoding time. As soon as a packet is recorded, the decoder starts decoding it at the beginning of the next time slot if it is not busy. The number of packets in the system buffer is managed until the duration of the simulation by defining a vector which is updated in every time slot where there is a possible packet arrival, a packet gets decoding, or both together at the same time slot. Consequently, the average number of packets can be found. However, Fig. 6 describes the average system buffer size obtained through simulation when employing different channel conditions, packet arriving probabilities, and upper bound decoding limits. The chosen value of lower bound decoding limit is the same as used in Fig. 2. The simulation time chosen is $4x10^5$ time slots. It is important to know that smooth results are obtained when choosing larger simulation time. This simulation has been run for four days with an increment of 0.01 for λ.

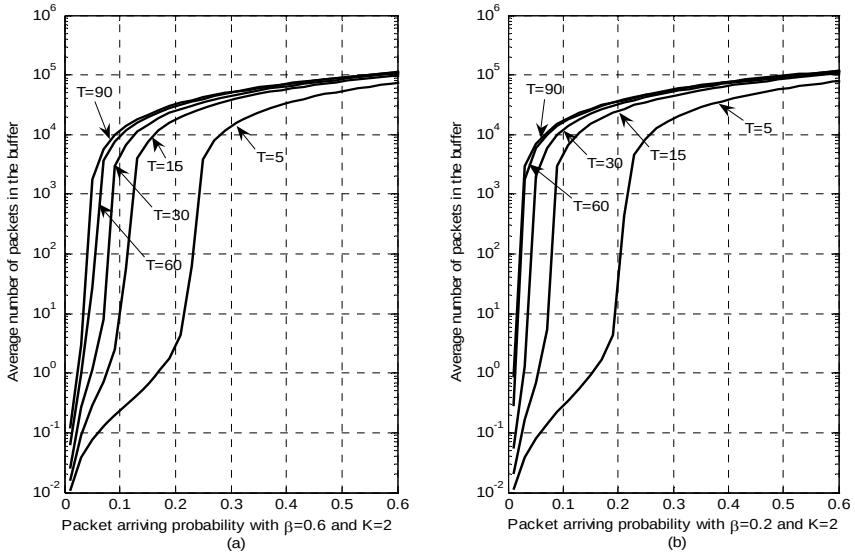

Figure 6: Average number of packets in the buffer versus packet arriving probability (λ) for different upper bound decoding limits and channel condition. Simulation time = $4x10^5$. (a) $\beta = 0.6$. (b) $\beta = 0.2$.

Actually, the simulation period is dependent on the increment selected for λ. The simulation has been performed on Intel Core 2 Duo 3.0 GHZ with a 2MB cache and 2 GB RAM. However, it is noticed that the average number of packets in the buffer increases as the packet arriving probability and the upper bound decoding limit increase. It is also seen an increase in the average number of packets as channel condition gets worse (i.e., β decreases).

One interesting observation can be made, which is the average number of waiting packets in the buffer reaches gradually to the working simulation time (i.e., maximum number of simulation slots) after being around a certain value of packet arriving probability which is mentioned in our analytical results as λ_{max}. In other words, the average buffer size moves gradually towards the duration of the simulation around $\lambda = 0.2$ for $(T = 5, \beta = 0.2)$ and $\lambda = 0.23$ for $(T = 5, \beta = 0.6)$. Also, it does so before $\lambda = 0.1$ for (T = 30, 60, and 90) when $\beta = 0.2$ and 0.6 with considering the notice that a faster move is the case when β decreases. While the case of T=15, the average buffer size moves around a limit after $\lambda = 0.1$ for $\beta = 0.6$ and before $\lambda = 0.1$ for $\beta = 0.2$. All of these simulation facts agree fairly with those shown in Fig. 2 as a result of theory.

The trend of moving gradually towards the duration of the simulation can be explained due to the finite number of simulation time slots $(4x10^5)$ where there is a maximum one packet that can arrive during any certain time slot. Hence, in the best case, a maximum of $4x10^5$ packets is the result. It is necessary to mention that if we increase the duration of the simulation (no matter what is the increase), the average buffer size will move gradually to that duration also around the mentioned limits of λ. This means reaching to infinity which verifies the same trend prescribed in theory as shown in Fig. 2.

V. CONCLUSION

In this paper, a hybrid type 1 ARQ with Fano decoding at the MAC layer of wireless access and end points is considered as an end-to-end improvement over wireless networks. A queuing analysis and simulation study are proposed for the system (queue and decoder) of wireless access and end points to obtain results about the expected number of packets in the system's buffer when Fano decoding or any other variable complexity decoding algorithms is implemented. This performance metric (average buffer size) has a severe impact on the wireless system performance and overall wireless network performance when it is chosen randomly. In the analytical study we derive a general form expression for the average size of the buffer that belongs to the Fano decoder, which is bounded by maximum and minimum variable decoding limits (T and K respectively), due to the unpredictable and noisy nature of wireless networks. This formulated expression is a function of not only (T and K), but also (β and λ). Analytical results show that the average buffer size increases dramatically when channel condition decreases (i.e., gets worse) by reaching λ_{max} which then goes to infinity. The effect of an increasing in λ on the average buffer size becomes less when β increases. Both variables T and K become highly sensitive to the increase of average buffer size when β decreases. Also, we provide results for a new derived form of λ_{max} that are totally the same as the λ_{max} results obtained through our general form expression of the average buffer size. On the other hand, we provide results obtained through simulation for the average buffer size. We show that our results and explanations for both

analytical and simulation studies agree fairly with each other.

REFERENCES

[1] K.A. Darabkh and R. S. Aygün, "TCP Traffic Control Evaluation and Reduction over Wireless Networks Using Parallel Sequential Decoding Mechanism," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, Article ID 52492 (SCI Expanded), 16 pages, 2007.

[2] Ye Tian, Kia Xu, and N. Ansari, "TCP in wireless environments: Problems and Solutions," *Communications Magazine*, IEEE, Volume 43, Issue 3, pp. S27 - S32, Mar. 2005.

[3] M. Gast, *802.11 Wireless Networks: The Definitive Guide*, O'Reilly, 2005.

[4] C. Rinaldi, "Link-Layer Error Recovery Techniques to Improve TCP Performance over Wireless Links," *Master's Thesis*, Royal Institute of Technology, Stockholm, Sweden, 2005.

[5] H. Balakrishnan and R. H. Katz, "Explicit loss notification and wireless web performance," *Proc. of IEEE GLOBECOM '98,* Internet Mini-Conference, Sydney, Australia, Nov. 1998.

[6] Y.Wang, L.Pan, and J. Li, "The Necessity of Combining ELN and SACK to Improve TCP Performance over Heterogeneous Networks," *Proc. of third International IEEE Conference on Signal-Image Technologies and Internet-Based System*, Shanghai, China, Dec. 2007, pp. 137-142.

[7] Y. S. Han, P.-N. Chen, and H.-B. Wu, "A maximum-likelihood soft-decision sequential decoding algorithm for binary convolutional codes," *IEEE Trans. on Commun.,* vol.50, no.2, pp.173-178, Feb. 2002.

[8] S. Lin, and D.J. Costello, *Error Control Coding: Fundamentals and Applications*, Pearson Education, 2004.

[9] J. B. Anderson, and S. Mohan, "Sequential Coding Algorithms: A Survey and Cost Analysis," *IEEE Trans. on Commun.,* vol. COM-32, no.2, pp.169-176, Feb. 1984.

[10] J. C. Moreira, and P. G. Farrell, *Essentials of Error-Control Coding*, John Wiley and Sons, 2006.

[11] R.O. Ozdag and P.A. Beerel, "A Channel Based Asynchronous Low Power High Performance Standard-Cell Based Sequential Decoder Implemented with QDI Templates," *Proc. of 10th International Symposium on Asynchronous Circuits and Systems*, Crete, Greece, Apr. 2004, pp.187-197.

[12] K. A. Darabkh, and W. D. Pan, "Queuing Simulation for *Fano* Decoders with Finite Buffer Capacity," *Proc. of the 9th Communications and Networking Simulation Symposium (CNSS'06)*, Huntsville, Alabama, Apr. 2006.

[13] W. D. Pan, and A. Ortega, "Adaptive Computation Control of Variable Complexity Fano Decoders," *IEEE Trans. on Commun.*, vol. 57, no. 6, pp. 1556 – 1559, June 2009.

[14] S. P. Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability*, Cambridge University Press, 2005.

[15] T. Hashimoto, "Bounds on a probability for the heavy tailed distribution and the probability of deficient decoding in sequential decoding," *IEEE Trans. on Inform. Theory*, vol. 51, issue 3, pp. 990 – 1002, Mar. 2005.

[16] R. Sundaresan and S. Verdu, "Sequential Decoding for the Exponential Server Timing Channel," *IEEE Trans. on Commun.*, vol. 46, no. 2, pp. 705-709, March 2000.

[17] F. Jelinek, "An Upper Bound on Moments of Sequential Decoding Effort," *IEEE Trans. on Inform. Theory*, vol. 15, no. 1, pp. 140-149, Jan. 1969.

[18] W.D. Pan and S.M. Yoo, "Queuing Analysis of sequential decoders with buffers," *Proc. of Huntsville Simulation Conference,* Huntsville, Alabama. Nov. 2004.

[19] W. D. Pan, and S.M. Yoo, "Fano Decoding with Timeout: Queuing Analysis," *ETRI Journal*, vol. 28, no. 3, pp. 301-310, June 2006.

[20] N. Shacham, "ARQ with Sequential Decoding of Packetized Data," *IEEE Trans. on Commun.,* vol. COM-32, pp. 1118-1127, Oct. 1984.

[21] K. A. Darabkh, and W.D. Pan, "Queuing Simulation for Sequential Decoders with Timeout," *Proc. of the 2005 Huntsville Simulation Conference*, Huntsville, Alabama, Nov. 2005.

[22] S. Kallel, and D. Haccoun, "Sequential decoding with an efficient partial retransmission ARQ strategy," *IEEE Trans. on Commun.,* vol. 39, Issue 2, pp. 208 – 213, Feb. 1991.

[23] K. A. Darabkh, and R. S. Aygun, "Performance Evaluation of Sequential Decoding System for UDP-Based Systems for Wireless Multimedia Networks," *Proc. of 2006 International Conference on Wireless Networks,* Las Vegas, Nevada, June 2006, pp. 365-371.

[24] K. Darabkh and B. Abu-Jaradeh, "Bounded Fano Decoders over Intermediate Hops Excluding Packet Retransmission," *Proc. of IEEE 24th International Conference on Advanced Information Networking and Applications,* Perth, Australia, Apr. 2010.

[25] D. Cross, and C. M. Harris, *Fundamentals of Queuing Theory*, John Wiley and Sons, 1998.

[26] S. Halfin, "The shortest queue problem", *Journal of Applied Probability*, vol. 22, pp.865-878, 1985.

[27] J. Gentle, *Random Number Generation and Monte Carlo Methods (Statistics and Computing),* Springer, 2004.

Khalid A. Darabkh received his Ph.D. degree in Computer Engineering from University of Alabama in Huntsville, Alabama, USA in 2007. He is currently an assistant professor and the assistant dean for computer affairs in the Faculty of Engineering and Technology at the University of Jordan, Amman, Jordan. His main research interests include wireless and mobile communications, queuing theory, traffic management, multimedia systems and networking, congestion control architectures and resource allocation, parallel computing, and pattern recognition.

Cooperative Networks: Bit-Interleaved Coded Modulation with Iterative Decoding

Shujaat Ali Khan Tanoli, Imran Khan and Nandana Rajatheva
School of Engineering and Technology, Asian Institute of Technology, Thailand
Email: {shujaat.ali.khan,imran.khan,rajath} @ait.ac.th

***Abstract*— In this paper the performance of bit-interleaved coded modulation-iterative decoding (BICM-ID) based cooperative network is analyzed over Rayleigh, Nakagami-m and Rician fading channels. In this system coding diversity is obtained through BICM and spatial diversity through cooperative relaying network. The analysis is performed for the relays operating in both amplify-and-forward (AF) and decode-and-forward (DF) modes. The bit error rate (BER) bound of multiple-relay cooperative system with the serial concatenated BICM-ID using the set partitioning (SP) and Gray mapped (GM) labeling for M-ary Phase Shift Keying (M-PSK) modulations over Nakagami-m fading channels is obtained. For the numerical results of union bound on BER, moment generating function (MGF) based approach is used. The expression of MGF derived for Nakagami distribution is then extended to Rayleigh and Rician fading channels. The comparisons of BER theoretical and simulation results are also shown. The performance of bit-interleaved low-density parity-check coded modulation with iterative decoding (BILDPCM-ID) based cooperative system, using low-density parity-check (LDPC) code instead of convolutional code, is also analyzed in term of bit-error rate (BER). The performance comparisons of BICM-ID and BILDPCM-ID based systems are shown using the set partitioning (SP) labeling for 8-PSK modulation scheme by performing Monte Carlo simulation.**

***Index Terms*— Bit-interleaved coded modulation, Bit-interleaved low-density Parity-Check coded modulation with iterative decoding (BILDPCM-ID), Cooperative Networks, Nakagami-m distribution, Moment generating function.**

I. INTRODUCTION

Wireless communication faces the main challenges of spectral efficiency, link reliability, power on the terminal and complexity. In order to address all such problems, the cooperative communication is most potential candidate in the next generation wireless systems [1] [2].

In cooperative communication, the spatial diversity as in multiple-input multiple-output (MIMO) is achieved by involving a number of relays between source and destination. This involvement helps to achieve the high data rate and make the overall system more reliable in term of bit error rate (BER) and throughput. The relay node operates over two strategies i.e. Amplify and forward (AF) and decode and forward (DF). Both modes of transmission are well-matched for high order modulation (i.e. 8-PSK, 16-QAM etc) and coding rates at source and relay. To obtain the maximum throughput, the transmitter and relay both raise its modulation order depending on the reliability of the channel condition between source-relay and relay-destination [3]. Hence, the communication system efficiently exploits the higher order modulation to further enhance the system throughput, which makes the system more efficient.

Bit-interleaved coded modulation (BICM) is a coding scheme that obtains the code diversity by effectively utilizing the hamming distance structure of binary codes when used in the combination of higher order modulation over fading channels. Hence, the reliability of the coded modulation in fading scenario could be further improved. The performance evaluation and principles of the design for bit-interleaved coded modulation is explained in [4].

In [5], the author identified that pitfall of BICM is reduced free Euclidean distance because of random modulation inbuilt in bit-interleaved scheme. To address this problem, a simpler approach of iterative decoding (ID) is used with a serial concatenation of encoding, bit-wise interleaving and high order modulation. In [6], bit-interleaved space time coded modulation with iterative decoding (BI-STCM-ID) is introduced for MIMO system. Hence, the performance of the BICM can be further improved by iterative decoding (BICM-ID) using hard decision feedback [7] and achieves significantly reduced receiver complexity as compared to turbo codes, as BICM-ID needs only one set of encoder/decoder.

The BICM-ID switches a 2^M-ary signaling channel to M parallel binary channels. With appropriate bit labeling, hence, a large binary Hamming distance between coded bits can be indirectly interpreted in to a large Euclidean distance. Consequently, this comprehends high diversity order, large free Euclidean distance and efficiently combines powerful binary codes with bandwidth efficient modulation. BICM-ID improves BICM by more than 1-dB and provides excellent performance over both Gaussian and fading channels, with iterative decoding [8] [9].

In this regard, BICM-ID will be most promising candidate when cooperative communication employs the higher order modulations over fading channels. The BICM-ID based cooperative system further allows different modulation i.e. 8-PSK, QAM, 16-QAM and also support different types of coding for cooperative transmission instead of conventional convolutional codes like space-time codes (STC) [6], Turbo codes [10] [11] and low density parity

Manuscript received February 11, 2010; revised April 15, 2010.

doi:10.4304/jcm.5.7.562-570

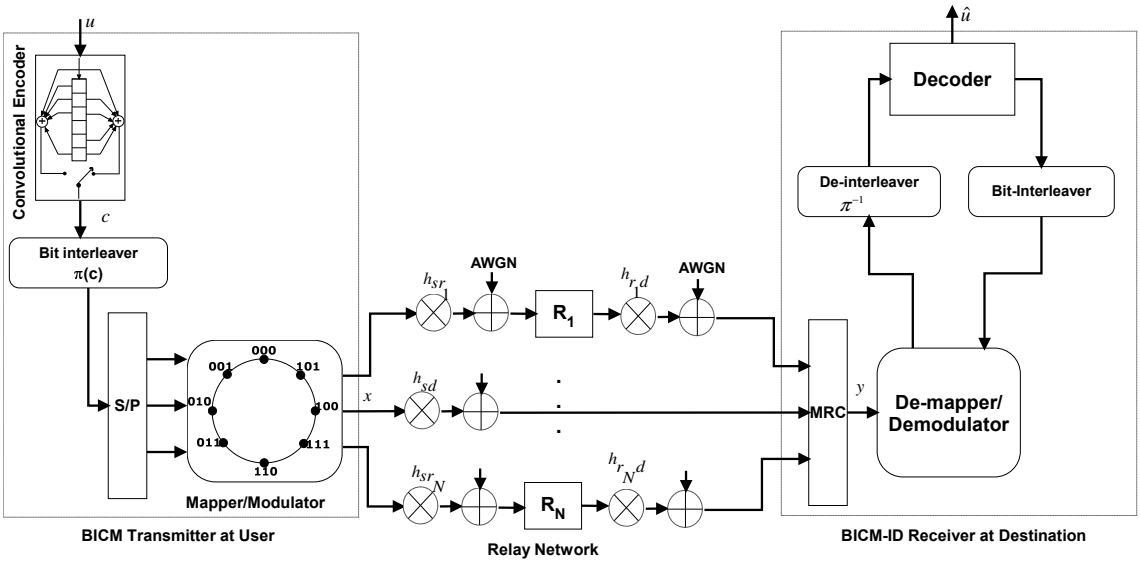

Figure 1. BICM-ID Based Cooperative Network.

check codes (LDPC) [12], [13]. In coded cooperation [14]- [15] the users are considered as the relay devices, that receive and decode the information bits of its neighboring user, re-encode and retransmit to the destination with additional parity bit for its neighbor users. If it is not possible for user to decode the information of its partner then it reverts to the non-cooperative mode.

The BICM based relay network using bilayer LDPC codes is introduced in [12]. Where, the fading scenario is ignored. In [16], we briefly analyze the performance of BICM based cooperative networks over AWGN and Rayleigh fading channels. The results are further extended to bit-interleaved space time-coded modulation for user cooperation diversity (CO-BISTCM) and two STC based transmission protocols are proposed for the system [17]. The idea of BICM and LDPC for DF mode is also presented in [18]. In [13], a bit-interleaved low density parity-check (LDPC) coded modulation with iterative demapping and decoding is introduced, using LDPC codes.

In this article, our main contribution is to analyze a single user serial concatenation of BICM-ID with multiple-relay cooperative system. The performance analysis of the proposed system is carried out by obtaining the theoretical bounds and verified with simulation results. The analytical results consist of BER curves of BICM-ID based cooperative system over Rayleigh, Nakagami-m and Rician fading channels for M-PSK modulation schemes. The capacity of the system is analyzed by performing Monte Carlo simulations for various fading channels. The analysis is performed for both AF and DF relaying schemes.

Currently, an explosion of attention has been seen in the field of relay coding to further enhance the capacity performance. So, in this regards we also present a novel approach of bit-interleaved low-density parity-check coded modulation with iterative decoding (BILDPCM-ID) based cooperative system in which we explore the possibility of using LDPC codes (instead of convention convolutional codes) with combination of BICM-ID and AF half duplex relay protocol. LDPC codes are very famous for their capacity-approaching performance for conventional single user communication channels. The main inspiration of LDPC codes is to practically apply the random coding theorem of Shannon by implementing a set of random parity check constraints on information bits [19].

This paper is further organized as follows: the next section presents the system model; section III shows the theoretical results for the MGF of branch metric. The simulation and analytical discussion is made in section V. Finally, section VI includes the conclusion of our work.

II. System Model

The system is composed of three main components: user (U), relay network (R) and destination (D). A single user BICM-ID based multiple-relay cooperative system model is given in Fig. 1 and each block is explained below.

A. BICM Transmitter at User and Relay:

The BICM Transmitter has further three modules discussed as follow:

Encoder: Two encoders are employed for U and R represented by Enc-U and Enc-R respectively. The Enc-U first takes k_c bit information block u from the user and generates n_c-bit codeword $\bar{c}_1 \in C^1$ and similarly $\bar{c}_2 \in C^2$ form Enc-R, where both C^1 and C^2 are binary codes. Note that, each encoder has two types, The convolutional encoder is used for BICM-ID and LDPC encoder for BILDPCM-ID scenarios.

Interleaver: Both code-words $\bar{c}_1$ and $\bar{c}_2$ are then bit interleaved by π-U and π-R respectively, to obtain the interleaved codeword. A one-to-one correspondence is

setup by interleaver in BICM technique i.e. $\pi : t \rightarrow (\acute{t}, i)$ here i identifies the bit position in the symbol label, $\acute{t}$ is the time ordering of the modulated symbol and t represents the time ordering of bit sequence before interleaving. Here, bit wise interleaving has the main objective, of maximizing the diversity order of the system by breaking the correlation of sequential fading coefficient [12].

Modulator: After interleaving, the bits are mapped onto symbols by M-PSK binary labeling maper at U and R $\mu_1 : (0,1)^{\ell_1} \rightarrow \chi_1$ and $\mu_2 : (0,1)^{\ell_2} \rightarrow \chi_2$ and modulated over the signal sets $\chi_1 \subset V^{q_1}$ and $\chi_2 \subset V^{q_2}$, where ℓ is the number of bits per symbol, q is the dimension of complex Euclidean spaces V^{q_1} and V^{q_2}. Both modulators at the U and R are memory less and denoted by Mod-U$= (\mu_1, \chi_1)$ and Mod-R$= (\mu_2, \chi_2)$ respectively. Where the size of the signal set is given by $|\chi| = M = 2^{\ell}$.

B. The Tanner graph of BILDPCM-ID

The Tanner graph of BILDPCM-ID system shown in Fig.2 represents both BICM-ID and LDPC codes. The corresponding bits and symbols have some restrictions. Here, the message passing occurs back and forth between symbol nodes and bit nodes that can be considered as the procedure of iterative decoding. The interesting thing which can be noticed form BILDPCM-ID Tanner graph is that different bit nodes are attach to each symbol node but there are no cycles between the symbol node and bit nodes except some minimum length of the cycles that termed as girth as shown in Tanner graph. These girths affects the BILDPCM-ID error performance at low coding rates.

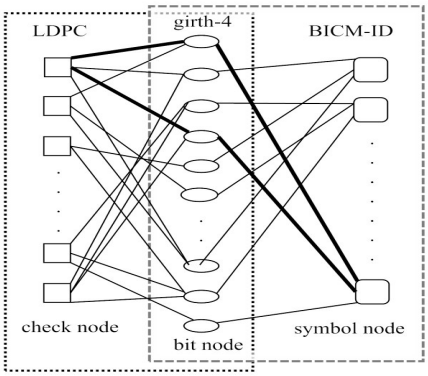

Figure 2. Tanner graph of BILDPCM-ID

C. Relay Network

The modulated signal is transmitted over the relay network. Consider the multiple-relay cooperative system as shown in the Fig. 3. There are Total N number of relays. The relays are operating in amplify-and-forward (AF) mode (transparent relaying) with fixed gain or decode-and-forward (DF) mode (regenerative relaying). In AF mode each relay receives a signal from user and retransmits after amplifying the signal to the destination. In the DF mode, the signal is decoded, re-encoded and finally retransmitted to the destination.

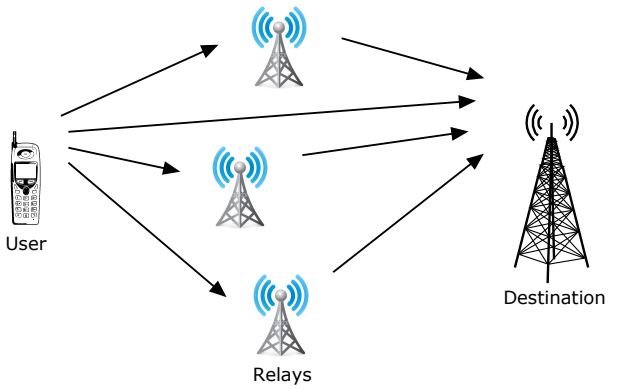

Figure 3. Relay Network.

1) Relay Links: We consider that all components operate on TDMA based two time frame protocol for transmission and reception as discussed in Section 2.2.2. The proposed system consist of three types of links $U \rightarrow R_n$, $U \rightarrow D$, and $R_n \rightarrow D$ where n is n-th relay among the multiple relays. For simplicity we assume that the links are itemize with j$= (0,1,2)$ respectively.

Each terminal U, R and D are equipped with single antenna for transmission and reception. The mechanism of various signal transmission is designed on the bases of positive integers q_1, q_2, and q_D dimensional complex Euclidean spaces V^{q_1},V^{q_2} and V^{q_D} respectively and transmission-reception of the single antenna can be modeled by allocating all these constraints to one. Similar to the MIMO channel model, the relay link with high order complex modulated signal (i.e. MPSK, and MQAM) can also be designed by setting all dimension $q_1 = 1$,$q_2 = 1$ and $q_D = 1$ for a single antenna on U, R and D.

Now the transmission/reception is carried out in two time frames. In first time frame, U transmits signal sequence $\bar{x}_1 = (x_{1,1}, \cdots, x_{1,f_1})$ to R and D, while in the second time frame R forwards the signal sequence $\bar{x}_1 = (x_{2,1}, \cdots, x_{2,f_2})$ to the destination. Where f_1 and f_2 are the transmissions during first and second time frames.

y_{sr_n} and y_{sd} are the signals sequences received at R and D respectively in the first time frame while $y_{r_n d}$ is the received signals sequences at D in the second time frames (discussed detail in later section) from corresponding links $U \rightarrow R_n$, $U \rightarrow D$, and $R_n \rightarrow D$. Here the number of relays $n = 1 \cdots N$ where n is n-th relay and N denotes the total number of relays. These links considered to be follow the frequency-flat fading channel distribution. The perfect channel state information (CSI) is assumed at the receiver of relay and destination and transmitter have no channel knowledge.

2) Channel Model: As shown by [5], [4], It is assuming in BICM system model that, ideal interleaving permits the transmission of code sequence through parallel binary input channel, where each bit position in the labeling map corresponds to that of parallel binary input channel.

Keeping in view, we can also consider the BICM with ideal interleaving assumption in our BICM-ID based cooperative communication system. So, in this regards our system is composed of $N+1$ sets of orthogonal parallel channels corresponding to $U \rightarrow D$ and $R_n \rightarrow D$ links with characteristics of independent and memory less binary input, where each bit location in the χ_j corresponds to each channel in a set of parallel channels. The conditional probability density function (PDF) of binary input channel when selected i bit position in label mapping is given by [4]:

$$p_{\theta^j}(y^j|b,i) = \frac{1}{2^{\ell_j - 1}} \sum_{z \in \chi_j(i;b)} p_{\theta^j}(y^j|z) \tag{1}$$

where θ^j denotes channel state, $i = (1, 2, \cdots, \ell_j)$, $b \in (0,1)$, $j = 1, 2$ and y^j is received signal sequence at D from j-th link.

It is assumed that the fading channels remain the same over the duration of two time slots i.e. slow-flat fading coefficients. Channel coefficients follow Rayleigh, Nakagami-m and Rician distributions. For Nakagami fading factor, $m = 1$ and $m = (K+1)^2/(2K+1)$ [20] (where K is the Rice factor), the channel coefficients follows Rayleigh and Ricean distribution, respectively. Nakagami-m distribution is a versatile statistical representation, that can model a variety of fading channels, like Rayleigh, Rician and one-sided Gaussian distribution. It is also assumed that perfect synchronization and channel state information (CSI) are available at the receiver. h_{xy} is the path gain through $X \rightarrow Y$ link and for $|h_{xy}|$ follows Nakagami distribution with probability density function (pdf) given by [21]:

$$f(|h_{xy}|) = \left(\frac{m}{E(|h_{xy}|^2)}\right)^m \frac{2|h_{xy}|^{2m-1}}{\Gamma(m)} \times \exp\left(\frac{-m|h_{xy}|^2}{E\{|h_{xy}|^2\}}\right) \tag{2}$$

where $E\{|h_{xy}|^2\} = 1$, Let $\alpha_{sd} = |h_{sd}|^2$, $\alpha_{rd} = |h_{rd}|^2$ and $\alpha_{sr} = |h_{sr}|^2$ are gamma distributed random variables and its pdf is given by:

$$f(\alpha_{xy}) = \frac{2m^m \alpha_{xy}^{m-1}}{\Gamma(m)} \exp(-m\alpha_{xy}) \tag{3}$$

where $\Gamma(.)$ is a gamma function. It is assumed that all the terminals transmit through orthogonal channels [22] using time division, frequency division or code division. The received signals at the destination are $(N+1)$ independent copies of transmitted signal.

3) Transmission Protocol: The transmission protocol used here is proposed by Nabar *et al.* [23]. In first time frame U broadcasts its information to R and D, and in second time frame only R operating in AF or DF mode, forwards the signal to D and U kept silence during second time frame. By doing so, we save the resources in second time frame.

The received signal at the destination in time slot 1 is given as:

$$y_{sd} = h_{sd}\sqrt{E_{sd}}x + w_{sd} \tag{4}$$

where , E_{sd} is the transmitted bit energy and w_{sd} is additive white Gaussian noise at terminal D, i.e. $w_{sd} \sim C\mathcal{N}(0, N_0)$. The received signal at the n-th relay in time slot 1 is given by:

$$y_{sr_n} = h_{sr_n}\sqrt{E_{sr_n}}x + w_{r_n} \tag{5}$$

where, $w_{r_n} \sim C\mathcal{N}(0, N_0)$.

Amplify-and-Forward (AF) scheme: By using the AF relay scheme in second time slot, the relay normalizes the received signal from source and forwards to the destination. The received signal through n-th relay (R_n) in time slot 2 is given as [24]:

$$y_{r_n d} = \frac{1}{\omega_n}\sqrt{\frac{E_{rd}E_{sr}}{E_{sr}+N_0}}h_{r_n d}h_{sr_n}x + w_{r_n d} \tag{6}$$

where $\omega_n = \sqrt{\frac{E_{rd}|h_{r_n d}|^2}{E_{sr}+N_0} + 1}$, is a noise normalization factor for n-th relay path, $w_{r_n d} \sim C\mathcal{N}(0, N_0)$ for $n = 1, 2, \cdots, N$ and $E\{|h_{sr_n}|^2\} = 1$. Source (S) transmits same signal to all the relays (R) with same power (E_{sr}). It is assumed that all the relays are transmitting the signal to the destination with the same power i.e. E_{rd}. Hence the amplification factor is the same for all the relays (N). N is also the number of independent signals received at D in the second time slot.

Decode-and-Forward (DF) scheme: In DF scheme, the relay forwards the decoded signal, when it successfully decodes the signal in time slot 2. Here we assume ideal case the relay knows whether the transmitted symbol is decoded correctly or not. The received signal at the destination from n-th relay is given by [24]:

$$y_{r_n d} = \sqrt{\hat{E}_{r_n d}}h_{r_n d}x + w_{r_n d} \tag{7}$$

where $\hat{E}_{r_n d} = E_{rd}$ on correctly decoding the signal, otherwise $\hat{E}_{r_n d} = 0$.

General: In general the input/output equations can also be written as:

$$\mathbf{Y} = \mathbf{H}\mathbf{x} + \mathbf{W} \tag{8}$$

where

$$\mathbf{Y}^T = \begin{pmatrix} y_{sd} & y_{r_1 d} & \cdots & y_{r_N d} \end{pmatrix}_{1\times(N+1)}$$

The channel gain is represented as:

$$\mathbf{H}^T = \begin{pmatrix} A & B_1 & \cdots & B_N \end{pmatrix}_{1\times(N+1)}$$

$$\mathbf{W}^T = \begin{pmatrix} w_{sd} & w_{r_1 d} & \cdots & w_{r_N d} \end{pmatrix}_{1\times(N+1)}$$

where $A = \sqrt{E_{sd}}h_{sd}$ and

$$B_n = \begin{cases} \frac{1}{\omega_n}\sqrt{\frac{E_{rd}E_{sr}}{E_{sr}+N_0}}h_{r_n d}h_{sr_n}; & \text{for AF} \\ \sqrt{\hat{E}_{r_n d}}h_{r_n d}; & \text{for DF} \end{cases}$$

where $n = 1, 2, \cdots, N$ and ^{T} denotes the transpose of the matrix.

D. BICM-ID Receiver at Relay and Destination

Similar to BICM Transmitter, The reversed components are installed at the R and D such as:

Demodulator: Bit metric calculation is carried out by employing two demodulators Dem-R and Dem-D at R and D respectively.

Interleaver: The third interleaver π-D is installed at the D to reduce the error propagation in iterative decoding as shown in Fig.1 by removing the correlation among sequentially coded bits and the bit associated with the same channel symbol [7].

Deinterleaver: Two deinterleavers π^{-1}-R and π^{-1}-D are at R and D respectively.

Decoder: The decoder Dec-R is at R and Dec-D is at D. Note that the iterative decoding implements on D, hence only Dec-D involves in iterative decoding but both decoders Dec-R and Dec-D separately decodes convolutional and LDPC codes for BICM-ID and BILDPCM-ID scenarios repectively.

At the destination maximal ratio combiner (MRC) is used to combine the received signal coherently and then noise normalization is performed. The BICM decoder with iterative decoding is installed at the destination. The block diagram of the BICM-ID decoder is shown in the Fig. 1. The hard decision of the information bits is made from the soft information. The de-mapper generates the extrinsic logarithmic likelihood ratios (LLRs). The LLRs are then de-interleaved and fed to the SISO a-posteriori decoder. The iterative decoding is carried by feeding back the posterior probability (MAP) through the interleaver to de-mapper for the next iteration. The output of demapper will be given as:

$$L_k = \log \frac{\sum_{x_k \in S_k^{(1)}} f(Y|x_k) p(x_k)}{\sum_{X_k \in S_k^{(0)}} f(Y|x_k) p(x_k)}$$

where $p(x_k)$ is the probability of signal $x_k \in S$ and $f(Y|x_k) = Dp(Y|x_k)$, D is a constant and $p(Y|x_k)$ is the probability that Y is received given that x_k is transmitted. $S_k^{(1)}$ and $S_k^{(0)}$ represent the set of symbols having k-th bit equal to 1 and 0, respectively. Fig. 8 shows the gray and set partitioning (SP) labeling maps for 8-PSK modulation. At the second pass decoding, given the feedback of bits 2 and 3, the constellation of bit 2 is confined to a pair of points shown in Fig. 8. Therefore as far as bit 1 is concerned, 8-PSK channel is translated into a binary channel with a BPSK constellation selected by the two feedback bits from the four possible signal pairs. Similarly we can proceed for bits 2 and 3.

III. Moment Generating Function (MGF) of Branch Metric

We review the bit metric for the viterbi decoding in BICM-ID based cooperative system over Nakagami-m fading channels and evaluate the conditional MGF of the metric difference. The BICM-ID based cooperative system uses viterbi decoding with branch metric, given

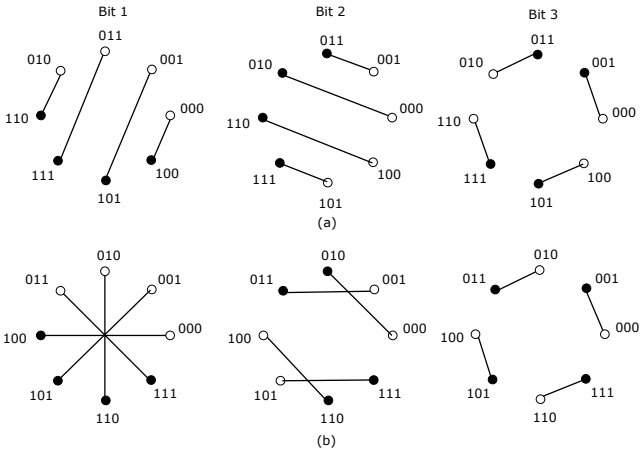

Figure 4. Labeling methods for 8-PSK a) Gray mapped (GM) b) Set partitioning (SP) [8].

as [4]:

$$\lambda_b^i = \min_{x \in \chi_b^i} \|\mathbf{Y} - \mathbf{Hx}\|^2 \tag{9}$$

where $\|.\|$ denotes the Euclidean norm of matrix. The metric difference relative to component x and z is given as:

$$\Delta(x,z) = \|\mathbf{Y} - \mathbf{Hx}\|^2 - \|\mathbf{Y} - \mathbf{Hz}\|^2 \tag{10}$$

For BER calculation, the MGF of metric difference can be given as [25]:

$$\Phi_{\Delta(x,z)}(s) = E\{\exp(-s\Delta(x,z))\} \tag{11}$$

The total MGF can be written as:

$$\Phi_{\Delta(x,z)} = \Phi_{\Delta(x,z)_{direct}} \times \Phi_{\Delta(x,z)_{2-hop}} \tag{12}$$

where $\Phi_{\Delta(x,z)_{direct}}$ is the MGF for direct path and $\Phi_{\Delta(x,z)_{2-hop}}$ is the MGF for 2-hop path. The MGF for direct path is given as:

$$\Phi_{\Delta(x,z)_{direct}} = E_{\alpha_{sd}} \left\{ e^{-sE_{sd} d_E^2 \alpha_{sd}(1-2sN_0)} \right\} \tag{13}$$

where, $d_E^2 = \|x - z\|^2$ is the squared Euclidean distance.

AF Scheme: For AF scheme, the MGF for 2-hop path is given as:

$$\Phi_{\Delta(x,z)_{2-hop}} = E_{\alpha_{sr_n}} \left\{ e^{-\sum_{n=1}^{N} \frac{se_1 d_E^2}{e_2+1} \alpha_{sr_n} \alpha_{r_n d}(1-2sN_0)} \right\}$$

where $e_1 = \frac{E_{rd}E_{sr}}{E_{sr}+N_0}$ and $e_2 = \frac{E_{rd}}{E_{sr}+N_0}$. Equation (12) should be averaged over all channel realizations to get the unconditional MGF, by assuming that α_{sd}, α_{sr} and α_{rd} to be independent random variables. When α_{rd} is given, (11) can be written as:

$$\Phi_{\Delta(x,z)|\alpha_{r_n d}}(s) = E_{\alpha_{sd}} \left\{ e^{-sE_{sd} d_E^2 \alpha_{sd}(1-2sN_0)} \right\} \times E_{\alpha_{sr_n}} \left\{ e^{-\sum_{n=1}^{N} \frac{se_1 d_E^2}{e_2 \alpha_{r_n d}+1} \alpha_{sr_n} \alpha_{r_n d}(1-2sN_0)} \right\} \tag{14}$$

$$\Phi_{\Delta(x,z)|\alpha_{r_nd}}(s) = \frac{1}{\left(1-\frac{sE_{sd}d_E^2}{m}(2sN_0-1)\right)^m} \times \frac{1}{\left(1-\frac{se_1d_E^2}{e_2m}(2sN_0-1)\right)^{mN}} \prod_{n=1}^{N}\left(1+\sum_{v=1}^{m}\frac{C_v}{(\alpha_{r_nd}+\lambda(s))^v}\right) \quad (17)$$

$$\Phi_{\Delta(x,z)}(s) = \frac{1}{\left(1-\frac{sE_{sd}d_E^2}{m}(2sN_0-1)\right)^m\left(1-\frac{se_1d_E^2}{e_2m}(2sN_0-1)\right)^{mN}} \times \left(1+\sum_{v=1}^{m} m^m C_v \lambda^{m-v}\psi(m,m-v+1,m\lambda)\right)^N \quad (18)$$

$$\Phi_{\Delta(x,z)} = \left(\frac{1}{1-sE_{sd}d_E^2(2sN_0-1)}\right) \times \left(-\frac{(1+e_1)}{k}.e^{\frac{1}{k}}.E_i(-\frac{1}{k})\right)^N \quad (19)$$

After averaging (11) with respect to α_{sd} and α_{sr}, (14) can be written as:

$$\Phi_{\Delta(x,z)|\alpha_{r_nd}}(s) = \frac{1}{\left(1-\frac{sE_{sd}d_E^2}{m}(2sN_0-1)\right)^m} \times \prod_{n=1}^{N} \frac{1}{\left(1\frac{se_1d_E^2}{(e_2\alpha_{r_nd}{}^{+1})m}\alpha_{r_nd}(2sN_0-1)\right)^m} \quad (15)$$

After some simple mathematical steps, (15) can be written as:

$$\Phi_{\Delta(x,z)|\alpha_{r_nd}}(s) = \frac{1}{\left(1-\frac{sE_{sd}d_E^2}{m}(2sN_0-1)\right)^m} \times \frac{1}{\left(1-\frac{se_1d_E^2}{e_2m}(2sN_0-1)\right)^{mN}} \prod_{n=1}^{N}(1+G(\alpha_{r_nd})) \quad (16)$$

where $G(\alpha_{r_nd}) = \frac{g_{m-1}\alpha_{r_nd}^{m-1}+\cdots+g_1\alpha_{r_nd}+g_0}{(\alpha_{r_nd}+\lambda(s))^m}$ with $\lambda(s) = \frac{1}{e_2-\frac{se_1d_E^2}{m}(2sN_0-1)}$ and $g_{m-1},\cdots,g_1,g_0$ are real constants. After decomposing into partial fraction, (16) can be written as (17). where $C_v = \frac{1}{(m-v)!}\frac{d^{m-v}}{d\alpha_{r_nd}^{m-v}}\left(\left(\alpha_{r_nd}+\lambda(s)\right)^m G(\alpha_{r_nd})\right)|_{\alpha_{r_nd}=-\lambda(s)}$.
By averaging (17) over α_{r_nd}, we get the unconditional MGF as given in (18) for Nakagami-m fading channels. where $k = e_1 + se_2d_E^2(2sN_0-1)$, $\lambda(s) = \frac{1}{e_2-\frac{se_1d_E^2}{m}(2sN_0-1)}$, $E_i(.)$ is an exponential integral function and $\psi(.,.,.)$ is the confluent hypergeometric function of the second kind [26].
For $m=1$, i.e. Rayleigh distribution, (18) takes the form as (19). $E_i(.)$ is an exponential integral function [26]. For Rician distribution, $m=(K+1)^2/(2K+1)$ [20] in (18).

DF Scheme: For DF scheme, the MGF for 2-hop path is given as:

$$\Phi_{\Delta(x,z)_{2-hop}} = E_{\alpha_{r_nd}}\left\{e^{-\sum_{n=1}^{N}\hat{E}_{r_nd}\alpha_{r_nd}(1-2sN_0)}\right\}$$

After averaging (12) with respect to α_{sd} and α_{r_nd}, the unconditional MGF can be written as:

$$\Phi_{\Delta(x,z)}(s) = \frac{1}{\left(1-\frac{sE_{sd}d_E^2}{m}(2sN_0-1)\right)^m} \times \prod_{n=1}^{N}\frac{1}{\left(1-\frac{s\hat{E}_{r_nd}d_E^2}{m}(2sN_0-1)\right)^m} \quad (20)$$

In most of the cases the relay decodes correctly at high SNR values, i.e. $\hat{E}_{r_nd} = E_{rd}$. Hence at higher SNR (20) can be written as:

$$\Phi_{\Delta(x,z)}(s) = \frac{1}{\left(1-\frac{sE_{sd}d_E^2}{m}(2sN_0-1)\right)^m} \times \frac{1}{\left(1-\frac{sE_{rd}d_E^2}{m}(2sN_0-1)\right)^{mN}} \quad (21)$$

In (21), $m=1$ and $m=(K+1)^2/(2K+1)$ [20] for Rayleigh and Rician distributions, respectively.

Here we obtain the expressions for multiple relay network and for both AF and DF operating modes. It is clear from the MGF expressions that a diversity order of $m(N+1)$ is obtained for N-relay network in Nakagami-m fading.

IV. Bit Error Rate (BER) Bound

The performance analysis in our work is based on the union bound analysis assuming Error free feedback [30] and moment generating function (MGF) approach used by [9], [27], [28] and [29]. The probability to decode a received sequence as a codeword x with an error weight d (hamming distance) given that a transmitted codeword is z is known as pairwise error probability (PEP). The PEP union bound for BICM can also be expressed in the form of moment generating function (MGF) approach, given as [4]:

$$f(d,\mu,\chi) \leq \frac{1}{2\pi j} \times \int_{\alpha-j\infty}^{\alpha+j\infty}[\psi_{ub}(s)]^d\frac{ds}{s} \quad (22)$$

where d is hamming distance of code and

$$\psi_{ub}(s) = \frac{1}{\ell 2^{\ell}} \sum_{i=1}^{\ell} \sum_{b=0}^{1} \sum_{x \in \chi_b^i} \sum_{z \in \chi_{\bar{b}}^i} \Phi_{\Delta_{(x,z)}}(s) \qquad (23)$$

$\Phi_{\Delta_(x,z)}(s)$ is the Laplace transform (MGF) of the metric difference $\Delta(x,z)$ between x and z.

As the iterative decoding give us a significant gain, so it is very much interesting for us to evaluate the analytical bound for the error free feedback performance which we term as error floor (*ef*)to which the BICM-ID performance converges at low BER.

Known ideal feedback for each $x \in \chi_b^i$, as there is only one term in $x \in \chi_{\bar{b}}^i$ whose label has the same binary bit values as those of x except at the i-th bit location that term is $\tilde{z} = \tilde{z}(x)$, where $\bar{b}$ is the compliment of b and $\tilde{z} = \tilde{z}(x)$denotes the nearest neighbor of x. Therefore, the PEP of the error floor of BICM-ID can be obtained by removing the innermost summation in (11), and can be written as [9]:

$$f(d,\mu,\chi) \leq \frac{1}{2\pi j} \times \int_{\alpha - j\infty}^{\alpha + j\infty} [\psi_{ef}(s)]^d \frac{ds}{s} \qquad (24)$$

where

$$\psi_{ef}(s) = \frac{1}{\ell 2^{\ell}} \sum_{i=1}^{\ell} \sum_{b=0}^{1} \sum_{x \in \chi_b^i} \Phi_{\Delta_{(x,\tilde{z})}}(s) \qquad (25)$$

The union bound of probability of bit error as shown in the Fig. 5 and 6 code of rate $R = k_c/n_c$ is given as [4]:

$$P_b \leq \frac{1}{k_c} \sum_{d=d_H}^{\infty} W_1(d) f(d,\mu,\chi) \qquad (26)$$

where the minimum Hamming distance d_H and $W_1(d)$ is the total input weight of error events at d . As the harmonic mean of the minimum squared Euclidean distance can also be increased by increasing in the euclidean distance between signals through iterative decoding, therefore, the error floor of BICM-ID is the horizontally shifted version of the performance curve of BICM without feedback.

V. Results and Discussion

Now in this section, we present the analytical and simulation results. The analysis is performed in terms of BER and achievable rates for the system with $1/2$ convolutional encoder of generator sequences $g = [133 \quad 171]$, QPSK and 8-PSK modulation schemes with set partitioning (SP) and Gray labeled mapping. The simulation results are obtained for uncorrelated Rayleigh, Nakagami and Rician fading channels by simulating 10^7 information bits using MATLAB and the numerical results of error floors are calculated from (18) and (21). The simulation results are taken for SISO decoder using the log-MAP algorithm with iteration for decoding. The analysis are based on Monte Carlo simulations.

Fig. 5 shows simulation results with analytical error free feedback (EF) bounds for Nakagami-2 fading channels. It presents the performance of the system with 2 relays operating in AF mode. These curves are simulated upto 5 iterations using set-partitioning (SP) mapped 8-PSK modulation. The results show that iterative decoding helps converging the BER curves to the bound obtained, at 3-rd iteration the BER converges to the theoretical EF bound given in (18). The error floor effect occurs at BER less than 10^{-4}. It is clear from the results that the BER curves are very tight to the asymptotic performances at medium and high SNR. In these regions the theoretical expression can be used.

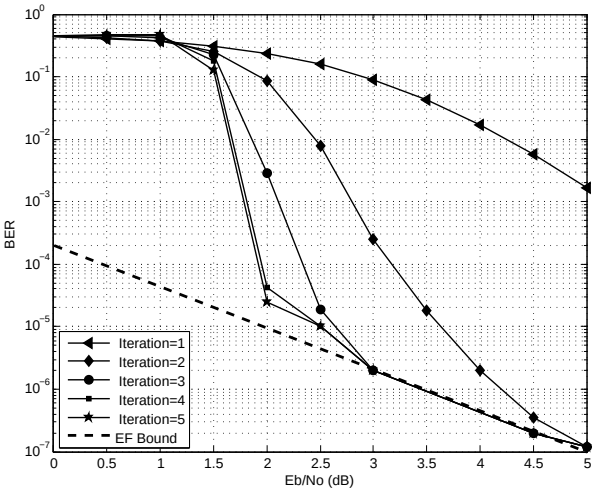

Figure 5. BER curves of BICM based 2-relay cooperative network over Nakagami-2 fading channels, 8-PSK and SP labeling.

Fig. 6 shows the EF bound and the simulation results for the same system over Rician fading channel for different number of iterations. The BER curves converges at SNR 2dB in 3-rd iteration. In this case the error floor occurs at SNR less than 10^{-4}, similar to the previous case.

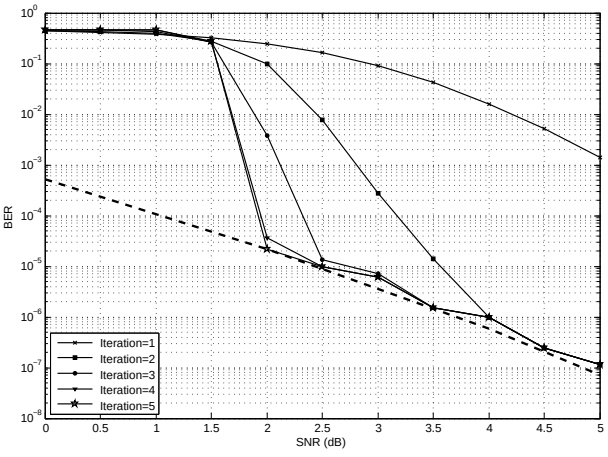

Figure 6. BER curves of BICM based 2-relay cooperative network for over Rician fading channels, $K = 10$, 8-PSK and SP labeling.

Fig. 7 shows the comparison of BER curves for partially and perfectly estimated decode-and-forward (DF) cooperative network with $m = 1$ and $m = 3$. In perfect DF, we assume the ideal case the relay knows whether

the transmitted symbol is decoded correctly or not. The variations of BER is shown for iterations 1 and 2.

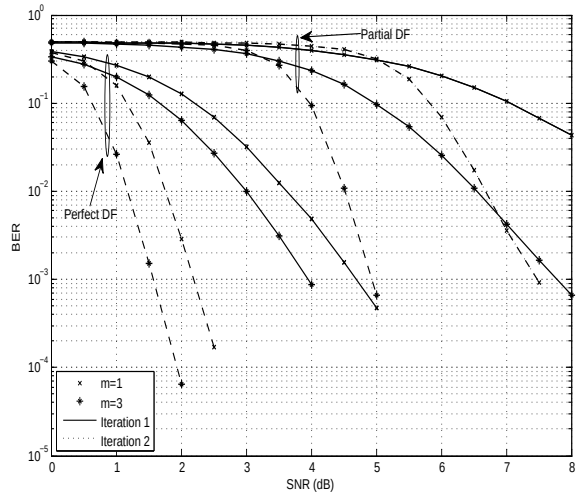

Figure 7. Comparison of Perfect and Partial DF over Nakagami-m fading channels ($m = 1, 3$), with SP labeling 8PSK and $N = 1$

For the performance comparison, the Monte Carlo simulation results are provided for BICM-ID and BILDPCM-ID based multiple-relay cooperative system over Nakagami-m fading channels. LDPC codes of rate $1/2$ length 2304 bits (LDPC short codes (SC)) from WiMax standard and 64800 bits (LDPC long codes (LC)) from DVB-S2 are used for BILDPCM-ID system. The LDPC internal iterations are set to be 20 for BILDPCM-ID system. The performance comparison of BILDPCM-ID and BICM-ID based single-relay cooperative network is shown in Fig. **??**. The simulation results are obtained for Nakagamim $(m = 1, 2)$ fading channels. From the curves, it is clear that in the BILDPCM-ID scenario the increase the m factor will increase the gap of about 1dB which is much larger than that of BICM-ID. The BILDPCM-ID gives significant improvement in the performance as compare to BICM-ID at low SNR. The results are shown for a single iteration of decoding at the end-receiver.

VI. Conclusion and Future Direction

We analyze the performance of BICM-ID based cooperative network over versatile Nakagami-m fading channels in terms of BER and achievable rates. The same theoretical results are extended to Rayleigh and Rician fading channels. The derivation of expression for the theoretical bounds is based on the MGF approach. MGFs of the metric difference of BICM-ID based multiple relay network with orthogonal channels are derived. Maximal ratio combining (MRC) is used at the destination to get the advantage of spatial diversity. The analysis is obtained for M-PSK modulation schemes. It is clear from simulation results that a significant gain is achieved in the performance of the system due to the cooperative and code diversities by introducing the number of cooperating relays and bit-interleaved coded modulation with iterative decoding on user and destination side, respectively.

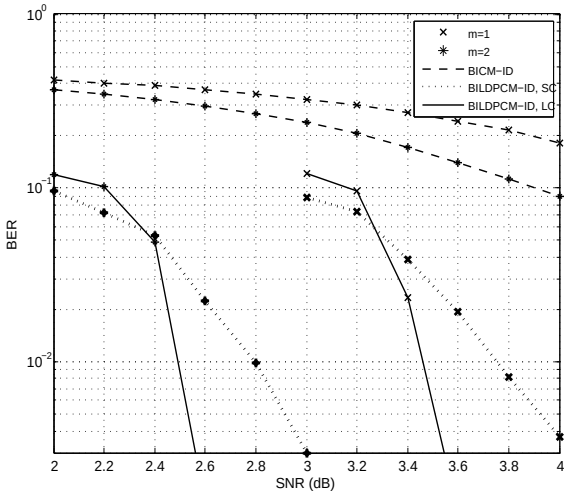

Figure 8. Performance comparison of BILDPCM-ID (short and long LDPC codes) and BICM-ID based single-relay cooperative network, for $m = 1, 2$ and 8-PSK.

Similarly, BICM-ID and cooperation brings a significant improvement to the performance of the system. The impact on the BER curves due to various number of iterations are also shown. The results show that iterative decoding converges the performance to the theoretical bound obtained.

The performance comparisons of BICM-ID and BILDPCM-ID based cooperative systems are also apresented, using the set partitioning (SP) labeling for 8-PSK modulation scheme by performing Monte Carlo simulation. we explored a novel approach of capacity-approaching performance for cooperative single user communication channels by exploiting iteratively decoded BI-LDPC codes and got significant improvement in the performance. Our work can be extended for higher order modulation like 16-QAM, 64-QAM. We can extend this analysis under the effects of other constraints like signal labeling and block length.

Acknowledgment

The first and second authors would like to thank the Higher Education Commission, Pakistan and UET Peshawar, Pakistan respectively, for providing the funds to carry out their higher studies.

References

[1] A. Nosratinia, T.E. Hunter and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, pp. 74–80, 2004.

[2] A. Sendonaris, E. Erkip and B. Aazhang, "User cooperation diversity, part I: System description," *IEEE Trans. on Commun.*, vol. 51, no.11, pp. 1927–1938, Nov. 2003.

[3] R. Hoshyar and R. Tafazolli, "BER performance analysis of a cooperative BICM system based on post-BSC model," Proc. IEEE 19th International Symposium PIMRC 2008. pp. 1–5, Sept. 2008.

[4] G. Caire, G. Taricco and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theo.*, vol. 44, no. 3, May 1998.

[5] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, pp. 873–884, May 1992.
[6] I. Shahid and N. Rajatheva, "Asymptotic BER bounds of iteratively decoded bit-interleaved space-time coded modulation under correlated Rayleigh and Rician fading channels," *IEEE Commun. Lett.*, vol. 12, no. 10, pp. 770–772, Oct. 2008.
[7] X. Li and J.A. Ritcey, "Bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 1, pp. 169–171, Nov. 1997 .
[8] X. Li and J.A. Ritcey, "Trellis-coded modulation with bit-interleaving and iterative decoding," *IEEE J. on Select. Areas in Commun.*, vol. 17, no. 4, April 1999.
[9] X. Li, A. Chindapol and J.A. Ritcey, "Bit-interleaved coded modulation with iterative decoding and 8PSK signaling," *IEEE Trans. Commun.*, vol. 50, no. 8, pp. 1250–1257, Aug. 2002.
[10] S. Lin and D.J. Costello, *Error control coding*, Prentice Hall, 2004.
[11] H.T. Nguyen, H.H. Nguyen and T. Le-Ngoc, "A bandwidth-efficient coded cooperative communications system," Proc. of the 64th IEEE VTC '06, pp. 1–5, Montreal, Canada, Sep. 2006.
[12] P. Razaghi, M. Aleksic and W. Yu, "Bit-interleaved coded modulation for the Relay channel using Bilayer LDPC codes," 10th Canadian Workshop on Info. Theory, 2007.
[13] Q. Xie, K. Peng, J. Song and Z. Yang, "Bit-Interleaved LDPC-coded modulation with iterative demapping and decoding," IEEE 69th Vehicular Technology Conference, 2009 (VTC Spring 2009), Barcelona, April 2009.
[14] M. Janani, A. Hedayat, T.E. Hunter and A. Nosratinia, "Coded cooperation in wireless communications: Space-time transmission and iterative decoding," *IEEE Trans on Sig. Proc.*, vol. 52, pp. 362–372, 2004.
[15] A. Stenfanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Trans. on Commun.*, vol. 52, pp. 1470–1476, 2004.
[16] S. A. K. Tanoli, I. Khan and N. Rajatheva, "Asymptotic BER bounds for BICM-based multiple relay cooperative network under AWGN and Rayleigh fading Channels," Proc. 4-th Int. Conf. on Commun. and Net. in China (ChinaCom'09), August 2009.
[17] I. Khan, S. A. K. Tanoli and N. Rajatheva, "Performance bounds for bit-interleaved space-time coded modulation with iterative decoding based cooperative network," *Journal of Communications*, Vol 5, No 4 (2010), 297-306, Apr 2010.
[18] G. Kramer, "Distributed and layered codes for relaying," The 39th Asilomar Conf. on Sig., Sys. and Comp., pp. 1752–1756, Oct. 2005.
[19] P. Razaghi and W. Yu, " Bilayer LDPC codes for the Relay Channel," IEEE International Conference on Communications, ICC '06. vol. 4, pp. 1574–1579, June 2006.
[20] R. Hasanizadeh and S. Zokaei, "Optimal power allocation to diversity branches of cooperative MISO sensor networks," *Intl. J. of Sig.*, IJSP, vol. 3, no. 4, pp. 281–288, 2006.
[21] M. Nakagami, "The m-distribution -A general formula for intensity distribution of rapid fading," Statistical Methods in Radio Wave Propagation, W. G. Hoffman, Ed. Oxford, U. K.: Pergamon, 1960.
[22] S. Ikki and M.H. Ahmed, "Performance analysis of cooperative diversity wireless networks over Nakagami-m fading channels," *IEEE Commun. Lett.*, vol. 11, no. 4, pp. 334–336, April 2007.
[23] R.U. Nabar, H. Bolckei and F.W. Kneubuhler, "Fading relay channels performance limits and space-time signal design," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.
[24] K.J.R. Liu and A.K. Sadek, *Cooperative communications and networking*, Cambridge University Press, 2009.
[25] A. Nasri, R. Schober and L. Lampe, "Performance evaluation of BICM-OFDM systems impaired by UWB interference," Proc. IEEE Intl. Conf. on Commun. (ICC 2008), Bejing, May 2008.
[26] I.S. Gradshteyn and I.M. Ryzhik, *Table of integrals, series and products*, 6th Ed., New York: Academic Press, 2000.
[27] C. Polprasert and J.A. Ritcey, "Effect of imperfect CSI on iteratively decoded BICM over Nakagami fading channels," Proc. IEEE Commun. Comp. and signal Processing, (PACRIM 2005), pp. 89–92, Aug. 2005.
[28] N.H. Tran and H.H. Nguyen, "Signal mapping of 8-ary constellations for BICM-ID systems over a Rayleigh fading channel," *IEICE Trans. Commun.*, vol. E88-B, no. 10, October 2005.
[29] N.H. Tran and H.H. Nguyen, "A novel multi-dimensional mapping of 8-PSK for BICM-ID," *IEEE Trans. on Commun.*, vol. 6, no. 3, March 2007.
[30] X. Li and J.A. Ritcey, "Bit-interleaved coded modulation with iterative decoding–Approaching Turbo–TCM performance without code concatenation," Proc. CISS'98, March 1998.

Shujaat Ali Khan Tanoli received his B.Sc. degree in electrical engineering from COMSATS Institute of Information Technology, Pakistan in 2006 and M.Sc. in Telecommunications Engineering from the Asian Institute of Technology, Thailand, in 2009. He is currently working towards the PhD degree at the School of Engineering and Technology, Asian Institute of Technology, Thailand. He is student member of IEEE.

Mr. Tanoli's research interests include performance analysis of wireless communications systems, OFDM, OFDMA, MIMO, BICM-ID based systems and cooperative networks.

Imran Khan received his B.Sc. degree (Honors) in electrical engineering from NWFP University of Engineering and Technology, Peshawar, Pakistan in 2003 and M.Sc. in telecommunications engineering from the Asian Institute of Technology, Thailand, in 2007. He is currently working towards the PhD. degree at School of Engineering and Technology, Asian Institute of Technology, Thailand. Earlier he has been working as Lecturer at NWFP University of Engineering and Technology, Peshawar, Pakistan since 2004. He is student member of IEICE and IEEE.

Mr. Khan's research interests include performance analysis of wireless communications systems, OFDM, OFDMA, MIMO, BICM-ID based systems and cooperative networks.

Nandana Rajatheva received the B.Sc. degree in electronic and telecommunication engineering (with first class honors) from the University of Moratuwa, Moratuwa, Sri Lanka, and the M.Sc. and Ph.D. degrees from the University of Manitoba,Winnipeg, MB, Canada, in 1987, 1991, and 1995, respectively. Currently, he is an Associate Professor of telecommunications in the School of Engineering and Technology, Asian Institute of Technology, Pathumthani, Thailand. Earlier, he was with the University of Moratuwa, Sri Lanka, where he became a Professor of Electronic and Telecommunication Engineering in June 2003. From May 1996 to December 2001, he was with TC-SAT as an associate professor. He is an editor of International Journal of Vehicular Technology (Hindawi) and senior member of IEEE (since 2001), Comsoc and VTS.

Dr. Rajatheva's research interests include digital and mobile communications, cooperative diversity, relay systems, OFDMA resource allocation, cognitive radio: detection/estimation techniques, space time processing MIMO systems and distributed video coding (DVC).

An Adaptive IEEE 802.15.4a TH-TDMA UWB Industrial Field Level Network

Farah Haroon, Kazi. M. Ahmed
Telecommunications Field of Study, Asian Institute of Technology, Thailand
Email: {Farah.Haroon, kahmed}@ait.ac.th

Abstract— **We introduce the deployment of IEEE 802.15.4a in industrial field level communication, which is unexplored so far. Impulse radio-time hopping ultra wide band (IR-TH UWB) physical layer is robust against dense multipath fading and interferences. However, we propose an adaption in medium access control (MAC) to meet the major constraints of reliability and timeliness. The self configuring MAC overcomes the difficulties and limitations of specified carrier sense multiple access/collision avoidance (CSMA/CA), slotted Aloha and guaranteed time slots (GTS) mechanisms. Addressable TH codes in fixed time slots of modified superframe structure are used to handle asynchronous alarm requests during ongoing synchronous transmissions. Traffic modeling is presented and we derive the expressions for accessing delays of both the synchronous and asynchronous data traffics. The realistic industrial non line of sight (NLOS) channel environment of IEEE 802.15.4a is employed and the respective power delay profile (PDP) and cumulative density function (CDF) of instantaneous signal to noise ratio (SNR) are found. Moreover, for large number of resolvable multipath components (MPCs), we also propose a reduced complexity adaptive SRake receiver, which efficiently recovers the weak IR-TH UWB signals in dense multipath propagation and strong noise. In addition, we evaluate the error performance of the proposed receiver architecture in comparison with conventional SRake and additive white Gaussian noise (AWGN) correlation receivers. Our simulation results show a significant performance improvement with very less number of SRake fingers.**

Index Terms— **Adaptive SRake receiver, field level industrial network, IEEE 802.15.4a, industrial NLOS environment, synchronous and asynchronous data.**

I. Introduction

An industrial network is characterized as a three level network. The lower most field level contains hardware, software and protocols for an interconnection of sensors and actuators with a master controller. It handles two types of data traffic: synchronous and asynchronous with both reliability and timeliness. The synchronous data is updated periodically and deals with the transfer of sensor and actuator states. Comparatively on the occurrence of alarms and emergency events, the low latency asynchronous data appears aperiodically and is served on high priority basis.

Among the most popular license free wireless technologies such as Bluetooth, ZigBee and IEEE 802.11, IEEE 802.15.4a impulse radio time hopping ultra wide band (IR-TH UWB) physical layer (PHY) has a strong potential to bear harsh and rugged industrial environment. The peculiar characteristics reduce intersymbol interference (ISI) and the inherent discontinuity mitigates the effects of large electromagnetic interventions. With small duty cycle and very fine pulse durations, it is recommended in dense multipath propagated industrial field level communication [1]. Independent time hopping (TH) codes of different data streams allow multiple access communication without catastrophic collisions. The resultant drop off in data rate is not so critical for small size data packets of field level, where the major constraint is meeting deadlines with correctness.

The wireless activities in industrial scenario were mostly directed towards the involvement of IEEE 802.11 and its alternatives. Prospects and significance of UWB in industrial applications were indicated in [2]- [4]. Deployment of IR-TH UWB in industrial adhoc networks was focused in [5] and [6], but the efforts were merely confined to a position based routing strategy. A packet aggregating medium access control (MAC) protocol for direct sequence spread spectrum (DSSS) UWB in field level network was suggested in [7]. Carrier sense multiple access/collision avoidance (CSMA/CA) was used as a channel access mechanism to reduce the large acquisition time (time required by the receiver to achieve bit synchronization with the transmitter) of the PHY.

However, the selected high data rate DSSS PHY is more suitable for visualization and monitoring services in upper levels of an industrial network [1]. CSMA/CA requires clear channel assessment (CCA) by energy detection (ED), which is challenging in UWB PHY having low power spectral density (PSD). Decorrelation and preamble detection carrier sensing are also much more complex. It was therefore declared in [8], that CSMA/CA is not appropriate for such ultra wide bandwidth signals and the use of other channel access mechanisms should be investigated.

IEEE 802.15.4a medium access control (MAC) relies on a superframe structure of 16 time slots. The first slot is always reserved for synchronization and transmission of control information through beacons. The remaining slots are occupied either by a contention access period (CAP) or a contention free period (CFP). Due to inherent restraints of CSMA/CA in UWB technology, use of slotted Aloha is also allowed in CAP [9]. But, both the channel access mechanisms for synchronous data transfer between single master and abundant of slave nodes suffer from

Manuscript received January 31, 2010; revised April 18, 2010; accepted May 11, 2010.

doi:10.4304/jcm.5.7.571-580

severe imperfections and render implementation complexities and very high acquisition times. For low latency applications, superframe structure provides a maximum of seven time division multiple access (TDMA) guaranteed time slots (GTS) in CFP. On request of slave nodes, they are reserved and allocated by the main master controller. But, due to the involvement of two way handshaking signals and insufficient availability, their provision is inapt for asynchronous data transfer.

Scattering and reflections with heavy metallic objects in industrial (non line of sight) NLOS surroundings generate plenty of resolvable multipath components (MPCs) and a dense power delay profile (PDP). In contrast to office and residential environments, it imposes serious complexity on optimum Rake receiver architectures. The first ray also does not carry the maximum amount of energy. Selective-Rake (SRake) utilizing the strongest MPCs is therefore preferable over all-Rake (ARake) and partial-Rake (PRake) receivers. SRake reception in IEEE 802.15.4a industrial NLOS channel model (CM8) was simulated in [10]. It indicated an obligation of hundreds of SRake fingers, but no solution to solve this problem is yet proposed.

Motivated by above considerations and constrictions, we bring in IEEE 802.15.4a in low data rate industrial field level networks. However, some relevant adaption is recommended to meet the deadlines of asynchronous requests. A modified superframe structure with addressable TH codes in TDMA based time slots is suggested. It overcomes the problems of traditional CSMA/CA in UWB technology and is capable of handling both the synchronous and asynchronous data. We propose a self configuring MAC protocol to transfer data without any CCA and ED. Expressions of the respective access delays are derived along with traffic modeling. Another substantial contribution of our work is the evaluation of error performance of IR-TH UWB in realistic IEEE 802.15.4a industrial non line of sight (NLOS) environment of channel model 8 (CM8). The respective channel cumulative distribution function (CDF) is found, which indicates the distribution of instantaneous signal to noise ratio (SNR) among very large number of resolvable MPCs. We include an adaptive noise cancellation principle of [11] to conventional SRake reception of IR-TH UWB. The proposed adaptive SRake architecture cancels the noise from weak UWB signals and improves SNR through an adaptive recursive least square (RLS) algorithm. It overcomes the complexity of [10] and shows a substantial performance improvement as compared to conventional SRake and AWGN correlation receivers.

The rest of the paper is organized as follows. Section II gives an overview of the system model with complete details of signal construction and industrial NLOS channel model 8. Section III covers the proposed adaption and modifications at the MAC layer. Traffic modeling and the derivation of access delays are also included. Proposed receiver architecture is specifically discussed in Section IV. Numerical and simulation results are presented in Section V. Section VI concludes the final remarks.

II. System Model

We deploy IEEE 802.15.4a in field level communication between master controller acting like a personal area network (PAN) coordinator and number of slave nodes. They are all connected in star topology in IEEE 802.15.4a industrial NLOS environment of CM8 as shown in Figure 1. All the nodes are fixed and consist of a transceiver which either acts as a transmitter or a receiver. An antenna array with a main lobe and a side lobe is utilized to obtain the primary and reference inputs at the proposed adaptive SRake receiver. Following [11], at the primary input, the main lobe is used to supply the distant transmitted signal corrupted by noise. Where as, the required correlated version of primary noise at the the reference input is acquired by its respective side lobe. Uncoded IR-TH

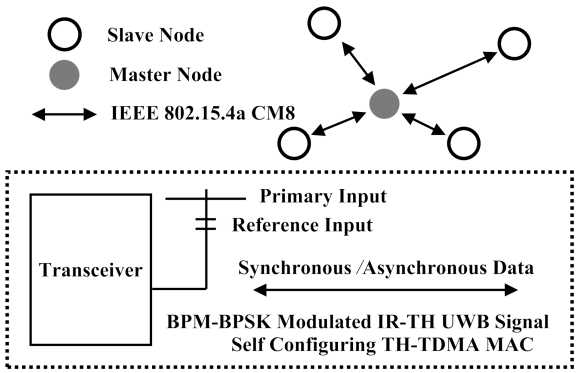

Figure 1. System Model

UWB, modulated using burst position modulation-binary phase shift keying (BPM-BPSK) is employed for both the synchronous and asynchronous data transfer at a center frequency of 7.987 GHz of channel 11 of high band plan [9]. In addition, a self configuring TH-TDMA MAC is proposed to reduce the accessing delay of asynchronous data classified into $q = 1, ..., Q$ priorities.

The detail of signal construction using addressable TH codes and IEEE 802.15.4a CM8 is as follows

A. Signal Construction using Addressable TH Codes

According to [9], data for coherent detection is modulated as a two bit symbol using BPM-BPSK. It generates a burst of duration T_{burst} using N_{cpb} consecutive pulses with each pulse occupying a small chip interval of T_c=2 ns. The burst can be present in any of the possible N_{hop} positions identified by independent TH codes, assigned to synchronous and each priority of asynchronous data. The code represents a sequence $[c_0^q, c_1^q, ..., c_{N_{hop-1}}^q]$ of integers defining respective burst locations in a data stream.

Although, specified pseudo-random TH codes with an unlimited number of independent sequences are very popular but due to the absence of any structure are also very difficult to address. Therefore, we assigned code to

each q^{th} priority class and synchronous data based on code construction 1 as mentioned in [12].

$$c_j^q = (q + j - 1)\ mod\ N_q \tag{1}$$

where N_q should be a prime number indicating the code length or period with a value taken between 0 and $N_{hop} - 1$.

The first data bit d_{jBPM}^{qc} encodes the position of the burst in either half of the symbol duration T_{BPM}, while the second data bit d_{jBPSK}^{qc} indicates its polarity. It generates a set of M=4 symbols using $m_b = 1, ..., M$ biorthogonal transmitted signals. For symbol energy E_s and $p_{tx}(t)$ as the first derivative Gaussian pulse, $s_{txm_b}^{qc}(t)$ is the m_b^{th} biorthogonal signal of q^{th} priority using c^{th} TH code given as below

$$s_{txm_b}^{qc}(t) = \sqrt{\frac{E_s}{N_{cpb}}} \sum_{j=-\infty}^{\infty} (1 - 2d_{jBPSK}^{qc})$$
$$\sum_{x=1}^{N_{cpb}} p_{tx}(t - c_j^q T_{burst} - xT_c - d_{jBPM}^{qc} T_{BPM}) \tag{2}$$

These M=4 signals can be completely constructed by $m = \frac{M}{2} = 2$ orthogonal signals and their inverses. If $d_{jBPSK}^{qc} = 0$ is kept in the above Eq.(2), $s_{txm}^{qc}(t)$ represents the m^{th} equivalent orthogonal signal expressed as

$$s_{txm}^{qc}(t) = \sqrt{\frac{E_s}{N_{cpb}}} \sum_{j=-\infty}^{\infty} \sum_{x=1}^{N_{cpb}} p_{tx}$$
$$(t - c_j^q T_{burst} - xT_c - d_{jBPM}^{qc} T_{BPM}) \tag{3}$$

In vector notation, these m signals can be shown by nonzero value in their m^{th} dimension as

$$\mathbf{s_{tx1}^{qc}} = (\sqrt{E_s}, 0) \tag{4}$$

$$\mathbf{s_{tx2}^{qc}} = (0, \sqrt{E_s}) \tag{5}$$

B. Industrial NLOS Channel Model 8

IEEE 802.15.4a industrial NLOS environment is available as CM8. The analysis is based on modified Saleh and Valenzuela (SV) model for indoor multipath propagation, which indicates that the grouping of objects in the surroundings leads to clustering of MPCs. Using a_{kl} and τ_{kl} as the gain and delay of k^{th} component of the l^{th} cluster and T_l as the arrival time of l^{th} cluster, the channel impulse response is given by

$$h(t) = \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} a_{kl}\ \delta(t - T_l - \tau_{kl}) \tag{6}$$

It is evident from Eq.(6) that for a carrier less baseband IR-UWB, using pulse based transmitter and receiver, the channel response is represented as a real pass band system without considering the phase angles. It is in contrast to complex baseband technique which is used to express channel impulse responses of carrier modulated signals [13]. Their reception involve the generation of in phase and quadrature components. For amplitude variations, complex Gaussian best accounts over a small area, with an equivalent complex baseband Rayleigh distribution.

In the presence of large excess delays prevailing in NLOS industrial, the conventional distribution is no more valid for UWB bandwidth having fine time resolution of delay bins. Alternatively, Nakagami distribution with probability distribution function (PDF) shown in Eq.(7) has been suggested for small scale fading with a Lognormally distributed Nakagami m_l factor [14].

$$pdf(z) = \frac{2}{\Gamma(m_l)} (\frac{m_l}{\Omega})^{m_l}\ z^{2m_l - 1}\ exp\ (-\frac{m_l}{\Omega} z^2) \tag{7}$$

Where $\Gamma(m_l)$ is the gamma function and Ω is the mean square value of the amplitude.

In other environments, the PDP appears with Poisson distributed cluster and ray arrival times. In contrast, with a single cluster in industrial NLOS channel, it is expressed as

$$E\{|\ a_{kl}\ |^2\} = (1 - \chi\ exp\ (-\frac{\tau_{kl}}{\gamma_{rise}}))$$
$$exp\ (-\frac{\tau_{kl}}{\gamma_1})\ \frac{\gamma_1 + \gamma_{rise}}{\gamma_1} \frac{\Omega_1}{\gamma_1 + \gamma_{rise}(1 - \chi)} \tag{8}$$

Where χ represent the attenuation of first path. γ_{rise} and γ_1 respectively pertaining the rising and decaying time constants with Ω_1 as the integrated energy of first single cluster.

For the statistics of small scale fading and complete description of modified path loss functions, [15] can be consulted.

III. THE PROPOSED MAC FRAMEWORK

The data and management services under IEEE 802.15.4a are respectively provided through physical layer data-service access point (PD-SAP) and physical layer management entity-service access point (PLME-SAP).

Each node either master or slave acts as a transceiver present within the communication range of all the other nodes of the network. They remain in receiving state

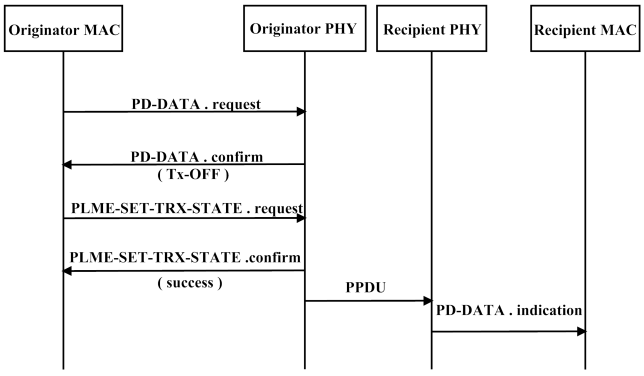

Figure 2. IEEE 802.15.4a message sequence for data transfer

with their transmitters disabled. However, if one wants to transmit, a PD-DATA.request primitive is assigned by a local MAC to its PHY as shown in Figure 2. In response, the local PHY delivers the PD-DATA.confirm primitive

to its MAC sublayer entity with a TX-OFF status. It leads MAC to issue a PLME-SET-TRX-STATE.request primitive to its PLME entity to disable the receiver (RX-OFF) and enable the transmitter (TX-ON). On successful change of state, physical layer service data unit (PSDU) is generated and transferred to the peer PHY. Finally the PD-DATA.indication primitive is generated by the receiving PHY to its MAC sublayer entity.

A. Data Traffic

The asynchronous and synchronous data in our system are modeled as

1) Asynchronous Data: It is classified into $q = 1, ..., Q$ classes with Q indicating the lowest asynchronous priority. The traffic arrival is modeled as a random Poisson (Markovian) process with exponentially independent and identically distributed inter arrival times. Each q^{th} class defines a separate queue with non preemptive Head On Line (HOL) discipline. They are served by one server, the master controller within fixed deterministic service time of T_s and modeled as Markovian/Deterministic/single server represented as M/D/1 traffic.

2) Synchronous Data: It has the lowest $Q+1$ priority and can be interrupted by any asynchronous request. Due to cyclic nature, it is updated periodically and is preferred to be buffered rather then being queued [16]. The old data present in the buffers is continuously replaced by the new ones in every cycle. To achieve data efficiency, periodic data does not need to be acknowledged as the receiver in case of occurrence of any error can wait for the correct data until the next cycle. The service time allowed is still equivalent to duration T_s. Although it occurs periodically but when executed with random asynchronous requests, is also treated as M/D/1 traffic.

B. Modified Superframe Structure

The MAC is based on a modified superframe structure and provides channel sensing functionality without incurring any extra complexity and overheads of ED and CCA. We assume the entire duration of the beacon enabled modified superframe structure of Figure 3 as an active period. All the slots are of fixed duration Ts with the first slot reserved for beacons from the master controller. The last 15 slots instead of being divided into CAP and CFP are all occupied by TDMA based contention free period. In normal operation, fixed scheduled synchronous

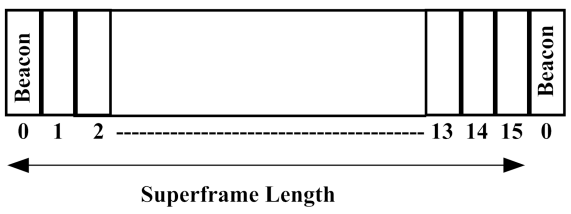

Figure 3. Modified contention free superframe architecture: Using TH codes asynchronous access is ensured in the next time slot without CCA and ED

data is transmitted on common TH code. On the occurrence of any alarm, the self configuring MAC delays the synchronous transmission and assigns the very next slot to the highest priority asynchronous request using a distinct TH code sequence. It avoids the use of standard GTS mechanism of [9] for low latency operations and hence overcomes the delay d_{GTS} expressed as a sum of

- d_{req}: time consumed in requesting a GTS from a slave node.
- d_{alloc}: on the availability and requirement, time spend in allocation of GTS by the master controller.
- d_{ack}: time wasted in number of acknowledgment transmissions.

Thus, the total delay d_{GTS} incurred in GTS management is averted, which is given as below

$$d_{GTS} = d_{req} + d_{alloc} + d_{ack} \tag{9}$$

C. Self Configuring TH-TDMA MAC

All the nodes are assumed within the communication range of each other having a complete knowledge of assigned TH codes of synchronous and asynchronous data. Although, the selected PHY efficiently supports multiple access communication based on distinct TH code sequences. But, for one to many or many to one directed master slave field level communication, single detection technique is more suitable. The operation starts from synchronous transmissions on common code sequence and continues as follows.

1) Identification: All the transmissions are completed within interval of time slot T_s. The header of the data packet holds the information about the used TH code. In addition, respective TH sequence is always interspersed in the end of data portion of every ongoing synchronous or asynchronous transmission. It provides channel sensing functionality and priority identification by all the receiving nodes without any ED and CCA.

2) Arrival: On the occurrence of alarm, the corresponding node acquires the transmitting state and signals the arrival of asynchronous request during the interspersed coded portion of the data packet. It first uses the TH sequence of the ongoing transmission which corrupts the embedded common code and indicates an arrival. Later, the alarm generating node employs respective self TH code which is registered in non preemptive HOL queue at all the receiving nodes. The transmission of cyclic data is postponed from the next scheduled node for one time slot and the alarm generating node transmits the data mixed with its code sequence in the forth coming slot.

3) Self Organization: If the synchronous processing has been resumed, the appearance of second asynchronous request is dealt as previously. Otherwise, on the basis of registered TH sequences, prioritized self configuring transmitting decisions at each node are taken and the respective alarms are sequentially processed. However, every new asynchronous request delays the synchronous transmission by one time slot.

D. Formulation of Delay

Requests of same priority generated from all nodes, form a separate queue with its respective arrival rate λ_q and service rate μ_q. Due to TDMA, mean service time of class q, $\overline{T}_q$ is constant and equals to slot duration T_s. It relates with μ_q and offered Poison traffic A_q as

$$\mu_q = \frac{1}{\overline{T}_q}, \ for \ \forall \ q = 1, ..., Q \tag{10}$$

$$Aq = \lambda_q \overline{T}_q \tag{11}$$

Thus the total asynchronous traffic offered to cyclic processing is given by

$$A_{asy} = \sum_{q=1}^{Q} A_q \tag{12}$$

The corresponding accessing delays of asynchronous and synchronous data on the basis of the proposed MAC are analyzed and formulated as below

1) Delay in Accessing of Asynchronous Data: On behalf of [17], the delay faced by an arbitrary asynchronous request of priority 'n' in being served is the sum of three factors

w_c = delay due to current ongoing transmission.
w_b = delay due to equal and higher priority requests that arrive before 'n'.
w_a = delay due to higher priority requests that arrive after 'n'.

a) Delay due to current transmission: This type of delay is always equal to mean residual service time of continuing request and can be determined using mean value analysis [18].
In our case, the slot assignment for an asynchronous alarm is either delayed due to ongoing synchronous or asynchronous transmission. The sum of the probabilities of being in either of two states would be 1 and the residual time by using mean value analysis can be written as

$$w_c = \overline{T}_{res} = p_{syn}\overline{T}_{syn} + p_{asy}\overline{T}_{asy} \tag{13}$$

The probability of occurrence of asynchronous data p_{asy} actually represents the total offered asynchronous traffic A_{asy}. The average time values of both the synchronous and asynchronous data respectively represented as $\overline{T}_{syn}$ and $\overline{T}_{asy}$ are equivalent to slot duration T_s. Thus, w_c is given by

$$w_c = (1 - \sum_{q=1}^{Q} A_q)T_s + \sum_{q=1}^{Q} A_q T_s = T_s \tag{14}$$

b) Delay due to equal and higher priority requests that arrive before 'n': Considering the arrival of an arbitrary asynchronous request with priority 'n', before which the requests of equal and higher priorities are already present. Let $\overline{N}_{qb}$ be the expected number of requests from individual priority q when $q \leq n$, then each of such request will arrive with an arrival rate λ_q and require to be serviced. Its value using Little equation for a period w_q (time up to the occurrence of q) is given as

$$\overline{N}_{qb} = \lambda_q w_q \tag{15}$$

Thus, the waiting time due to requests from all such classes equals to

$$w_b = \sum_{q=1}^{n} \overline{N}_{qb} \ \overline{T}_q = \sum_{q=1}^{n} A_q w_q \tag{16}$$

c) Delay due to higher priority requests that arrive after 'n': Equal priority requests are not considered due to HOL discipline with in one priority class. Let $\overline{N}_{qa}$ be the expected number of requests from individual priority q, when $q < n$ with a duration that can extend up to w_{asy}. The respective waiting time is expressed as

$$w_a = \sum_{q=1}^{n-1} \overline{N}_{qa} \ \overline{T}_q = \sum_{q=1}^{n-1} A_q w_{asy} \tag{17}$$

Now the overall delay in processing of asynchronous data of n^{th} priority will be given by

$$w_{asy} = T_s + \sum_{q=1}^{n} A_q w_q + \sum_{q=1}^{n-1} A_q w_{asy} \tag{18}$$

As w_{asy} is the delay in accessing of n^{th} asynchronous request represented by w_n, then the above equation is simplified as

$$w_{asy} = \frac{T_s + \sum_{q=1}^{n-1} A_q w_q}{1 - \sum_{q=1}^{n} A_q}, \quad n = 1, ..., Q \tag{19}$$

E. Delay in Processing of Synchronous Data

The delay in synchronous data due to the arrival of n^{th} asynchronous request is the sum of the accessing delay w_{asy} and the mean service time of n^{th} arrival.

$$w_{syn} = w_{asyn} + T_s \tag{20}$$

where T_s is the time spent during the transmission of n^{th} asynchronous request. The resultant delay in synchronous processing is therefore given by

$$w_{syn} = \frac{T_s(2 - \sum_{q=1}^{n} A_q) + \sum_{q=1}^{n-1} A_q w_q}{1 - \sum_{q=1}^{n} A_q}, \ n = 1, ..., Q \tag{21}$$

IV. An Adaptive SRake Receiver

In frequency selective IEEE 802.15.4a UWB channel, under the slow fading assumption the gain a_k and the delay τ_k of the k^{th} MPC are taken constant over a symbol duration. If all such K paths exhibit negligible correlations, then $\{a_k\}_{k=0}^{K-1}$ are treated statistically independent random variables with their PDF presenting Nakagami distribution. In the presence of AWGN, which is assumed as independent of fading amplitudes a_k having N_0 (W/Hz) as single sided power spectral density, the instantaneous SNR of k^{th} path per symbol having E_s energy is given by

$$\gamma_k = (a_k^2 E_s)/N_0 \tag{22}$$

and the respective total received instantaneous SNR would be

$$\gamma = \sum_{k=0}^{K-1} \gamma_k \tag{23}$$

In our system, a set of M=4 biorthogonal signals of uncoded IEEE 802.15.4a IR-TH UWB is transmitted through an industrial NLOS environment of CM8 in the presence of AWGN as shown in Figure 4. Addition

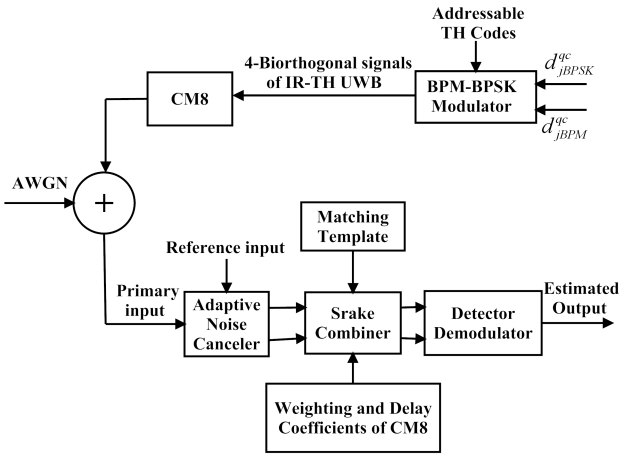

Figure 4. A transceiver system with proposed receiver architecture

of noise engulf the dense multipath propagated weak UWB signals which results in an increased complexity of SRake reception. A receiver architecture is proposed that includes an adaptive noise canceler (ANC) with primary and reference inputs, SRake combiner and a detector followed by a demodulator. Instead of utilizing M=4 biorthogonal signals, the SRake combiner is comprised of two independent Rake units for the respective $\frac{M}{2}$ orthogonal signals which reduces the required Rake units by half [19]. In addition to weighting and delay coefficients, the matching template also rely on the impulse response of CM8. The detector estimates one of $\frac{M}{2}$ signals based on highest magnitude of correlation sum. The sign of correlation sum finally decides one of M biorthogonal signals.

The sections of the proposed receiver architecture of Figure 5 utilizing K_s strongest MPCs include

A. RLS Adaptive Noise Canceler

For m^{th} transmitted signal, the respective mutually and statistically independent and identically distributed (iid) AWGN random variable has zero mean and variance given by

$$\sigma_m^2 = \frac{N_o}{2}, \quad m = 1, ..., \frac{M}{2} \tag{24}$$

While neglecting the antenna effects and assuming $h(t)$ as the impulse response of CM8, the received signal at the primary input of ANC is represented as $s_{rxm}^{qc}(t)$. If $*$ indicates the convolution operation and $n_{pm}(t)$ as the

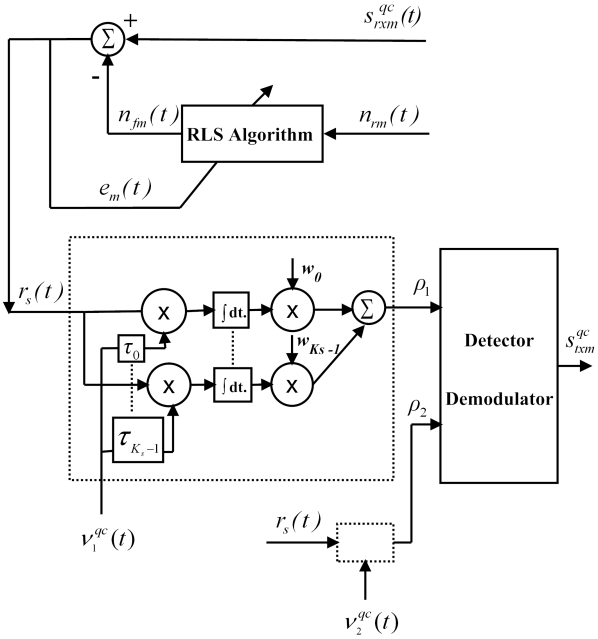

Figure 5. Adaptive SRake receiver operating in IEEE 802.15.4a industrial NLOS CM8, using an equivalent $\frac{M}{2}$ orthogonal signals instead of M biorthogonal signals

primary noise added into the m^{th} signal, $s_{rxm}^{qc}(t)$ is given by

$$s_{rxm}^{qc}(t) = s_{txm}^{qc}(t) * h(t) + n_{pm}(t) \tag{25}$$

Using i as time index the above equation is written as

$$s_{rxm}^{qc}(i) = s_{txm}^{qc}(i) * h(i) + n_{pm}(i) \tag{26}$$

where $s_{txm}^{qc}(i) * h(i)$ designates our desired response $d_m(i)$ free from noise $n_{pm}(i)$. The noise $n_{rm}(i)$, a correlated version of primary noise appears at the reference input. It was applied at the first tap input of the M_w order tap weight vector. After being adaptively filtered as $n_{fm}(i)$, it was subtracted from the primary input to produce an error signal $e_m(i)$ at the output of an ANC.

$$e_m(i) = s_{rxm}^{qc}(i) - n_{fm}(i) \tag{27}$$

It minimizes the squared error via RLS algorithm of [20]. The filtered noise $n_{fm}(i)$ was adjusted such that the noise power present in the primary input was reduced keeping the signal power unchanged. Thus, for the length of the observed data y, our objective function $\varepsilon_m(y)$ is given by

$$\varepsilon_m(y) = \sum_{i=1}^{y} \beta_m(y,i) \mid e_m(i) \mid^2 \tag{28}$$

To minimize the objective function, the algorithm on the basis of initial known conditions and feed back samples recursively updates the old estimates of tap weight vector $\mathbf{w_m}(y)$.

During the observation interval $1 \leq i \leq y$, the forgetting factor $0 \leq \beta_m(y,i) \leq 1$ of Eq.(29) is introduced to exponentially ignore the effect of past data.

$$\beta_m(y,i) = \alpha_m^{y-i} \tag{29}$$

where α_m is a positive constant close to, or equal to 1. The tap weight vector modifies the error signal according to

$$\mathbf{e_m}(i) = \mathbf{s^{qc}_{rxm}}(i) - \mathbf{w_m}^H(y)\mathbf{n_{rm}}(i) \quad (30)$$

when $\mathbf{s^{qc}_{rxm}}(i)$ indicates the time indexed sampled values of the received signal and $\mathbf{w_m}^H(y)$ at $(t = y)$ represents the conjugate transpose of $\mathbf{w_m}(y)$.
Both the tap weight vector $\mathbf{w_m}(y)$ and the time indexed reference noise vector $\mathbf{n_{rm}}(i)$ are expressed as

$$\mathbf{w_m}(y) = [w_{m0}(y), w_{m1}(y), ..., w_{m(M_w-1)}(y)]^T \quad (31)$$

$$\mathbf{n_{rm}}(i) = [n_{rm}(i), n_{rm}(i-1), ..., n_{rm}(i-M_w+1)]^T \quad (32)$$

The optimum value of the tap weight vector, at which our objective function acquires the minimum, is represented as $\hat{\mathbf{w}}_\mathbf{m}(y)$. According to [20], it is the product of the inverse of $\mathbf{\Phi_m}(y)$ and $\theta_\mathbf{m}(y)$ written as

$$\hat{\mathbf{w}}_\mathbf{m}(y) = \mathbf{\Phi^{-1}_m}(y)\theta_\mathbf{m}(y) \quad (33)$$

where $\mathbf{\Phi^{-1}_m}(y)$ is the M_w by M_w time averaged correlation matrix of tap inputs given by

$$\mathbf{\Phi^{-1}_m}(y) = \sum_{i=1}^{y} \alpha_m^{y-i}\mathbf{n_{rm}}(i)\mathbf{n_{rm}}^H(i) \quad (34)$$

and $\theta_\mathbf{m}(y)$ of m^{th} signal is the M_w by 1 time averaged cross correlation vector between the tap inputs and the complex conjugation of desired response $\mathbf{d_m}(i)$ indicated as

$$\theta_\mathbf{m}(y) = \sum_{i=1}^{y} \alpha_m^{y-i}\mathbf{n_{rm}}(i)\mathbf{d_m}^*(i) \quad (35)$$

The received signal after being adaptively filtered at the output of ANC is represented as $r_s(t)$. It was applied to both the Rake units, where the resultant increased SNR was effectively utilized to reduce the required number of fingers.

B. SRake Combiner

The branch statistic of SRake Combiner was found by generating an impulse response of CM8 according to [21]. 100 realizations of channel impulse response $h(t)$ were produced at each sampling interval T_s whose average gave the channel vector $\mathbf{h}$ with the corresponding time index vector $\mathbf{t}$.
Representing the strength of all the received K MPCs, $\mathbf{h}$ was then sorted for K_s strongest MPCs in $\mathbf{h_s}$ with time index $\tau_\mathbf{s}$ as

$$\mathbf{h_s} = [h_0(\tau_0), h_1(\tau_1), ..., h_{(K_s-1)}(\tau_{K_s-1})]^T \quad (36)$$

$$\tau_\mathbf{s} = [\tau_0, \tau_1, ..., \tau_{(K_s-1)}]^T \quad (37)$$

The above two vectors respectively represent the branch weighting coefficients $\mathbf{w_s}$ and delay elements of SRake combiner. A set of K_s fingers (branches) in each m^{th} Rake unit is indicated by a dashed box in Figure 5.

The locally generated matching template $\nu_m^{qc}(t)$ after neglecting the antenna effects was selected for one complete symbol duration. For m^{th} signal using c^{th} TH code assigned to q^{th} priority class, we selected it as

$$\nu_m^{qc}(t) = s_{txm}^{qc}(t) * h(t) \quad (38)$$

The received signal in each branch was correlated with the matching template delayed by τ_s and weighted by w_s. The sum of all the branch outputs corresponds to a maximal ratio combiner (MRC), which was treated as a decision metric ρ_m given by

$$\rho_m = \sum_{s=0}^{K_s-1} w_s.C_m(r_s(t), \nu_m^{qc}(t-\tau_s)) \quad (39)$$

where $C_m(a, b)$ represents the correlation between a and b entities.
Both of these ρ_m values for the two orthogonal signals fed into a detector/demodulator for final estimation of the transmitted symbol.

C. Detector

Assuming perfectly synchronized transmitter and receiver, the signal detection was analyzed and simulated as a single q^{th} data stream using c^{th} TH code. The estimated signal $\tilde{s}_{txm}^{qc}(t)$ was based on maximum likelihood (ML) detection expressed as

$$\tilde{s}_{txm}^{qc}(t) = \arg\max_m \mid \rho_m \mid \quad (40)$$

Later, either of $s_{txm}^{qc}(t)$ or -$s_{txm}^{qc}(t)$ was selected on behalf of the sign of the weighted largest correlation sum. Thus employing an equivalent orthogonal set, one of M signal was selected using $\frac{M}{2}$ cross correlators. The corresponding estimated symbol was finally provided by a demodulator.

V. Numerical and Simulation Results

A. Delay Evaluation

The accessing delays of $n = 1$ to Q asynchronous priorities were found numerically using Eq.(19). These delays are respectively shown in Figure 6 for slot duration $T_s = 0.01$ s, $Q = 5$, $w_q = 100$ hrs and the arrival vector $\lambda_\mathbf{q}$ of Eq.(41) in requests/hr.

$$\lambda_\mathbf{q} = [0.001, 0.005, 0.001, 0.002, 0.015]^T \quad (41)$$

It can be observed that an arbitrary n^{th} request belonging to q = 1 highest priority faces a delay equivalent to slot duration T_s. Where as, the delays faced by other priorities also depend on their arrival rates and time of occurrences.

Delay due to equal and higher priority requests that appear before n^{th} customer directly affects w_{asy} and depends on the period w_q. Using the same arrival vector, the asynchronous access delays for $n = 1$ to Q priority requests for the period w_q ranging from 100 to 600 hrs were found numerically and plotted in Figure 7. It can be observed that delay faced by the highest priority test customer is independent of the variations of w_q and is

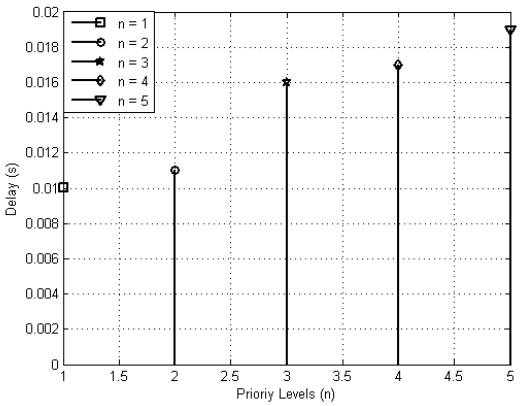

Figure 6. Accessing delays of $n = 1, ..., Q$ priority asynchronous data at w_q = 100 hrs

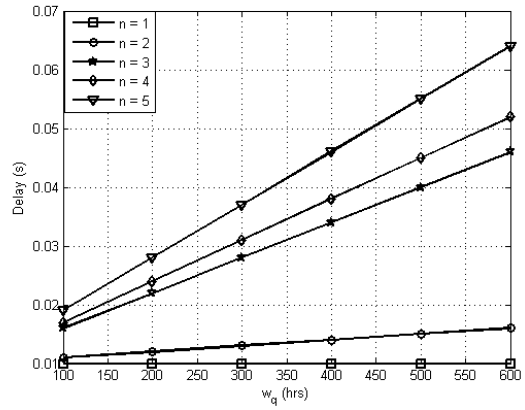

Figure 7. Accessing delays of $n = 1, ..., Q$ priority asynchronous requests as a function of w_q

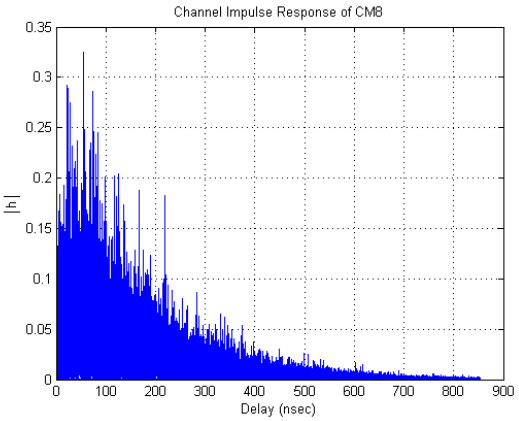

Figure 8. Channel impulse response: Average over 100 CM8 channels

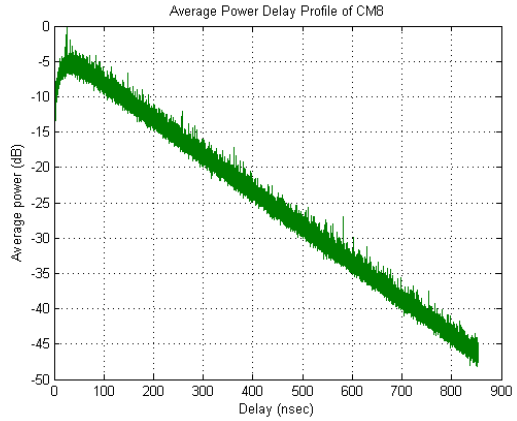

Figure 9. Power delay profile of CM8

always equivalent to slot duration T_s. However, at small arrival rates, effect of w_q dominates especially for $n = 3$ to 5 lower priorities .

B. Characteristics of Industrial NLOS Channel

Simulations were carried out in IEEE 802.15.4a industrial NLOS environment of CM8 at a center frequency $f_c = 7.987$ GHz of channel 11 of high band plan. Using a sampling frequency of $f_s = 50$ GHz, the channel impulse response and average PDP were generated as shown in Figure 8 and Figure 9. Both of them confirm the previous discussion and clearly indicate that the first arriving component in industrial NLOS does not exhibit the highest magnitude.

The CDFs of instantaneous received SNR from 20 and 100 strongest MPCs in CM8 were acquired for 1000 channel realizations in Figure 10. For it, discrete probability distribution functions (PDFs) were obtained from respective histograms for 0 dBs transmitted symbol energy. As a result, a gain of only 5 dBs is achieved when strongest MPCs are increased from 20 to 100 in channel environment following Nakagami probability distribution. It suggests an obligation of large number of Rake fingers required to attain an appreciable increase in energy collection in dense multipath propagated industrial NLOS environment.

C. Performance of Adaptive Noise Canceler

For simulation purpose instead of obtaining the primary and reference signals from an antenna array, 1000 random samples of primary noise $n_{pm}(i)$ were generated at ($i = 1$ to 100) time instants. At each i^{th} time instant they were passed through a 32 order low pass filter (LPF) to obtain the respective correlated reference noise samples of $n_{rm}(i)$. In response, a correlation strength $\varphi_m(i)$ ranging from -0.08 to +0.07 was obtained between the two types of noise samples as depicted in Figure 11.

Both of these noise samples were fed into an ANC. For the simulation of RLS algorithm in $M_w = 32$ order tap weight vector, Identity matrix $\mathbf{I}$ was selected as $\mathbf{\Phi_m}(y)$ of Eq.(34). The tap weight vector $\mathbf{w_m}(y)$ was initialized using all zeros vector with the exponential weighting factor $\alpha_m = 1$. To keep the recursions in progress and for the initial setting of the correlation matrix, $\mathbf{\Phi_m}(y)$ was multiplied with a small positive value $\delta_m = 0.1$.

The impulse response of ANC is adapted via RLS algorithm to remove the noise traces from the primary

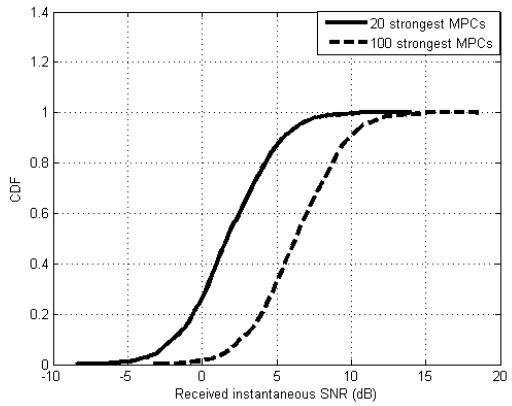

Figure 10. The CDF of received instantaneous SNR obtained from 20 and 100 strongest MPCs using 1000 channel realizations of CM8. At E_s = 0 dB and N_0= 1 W/Hz

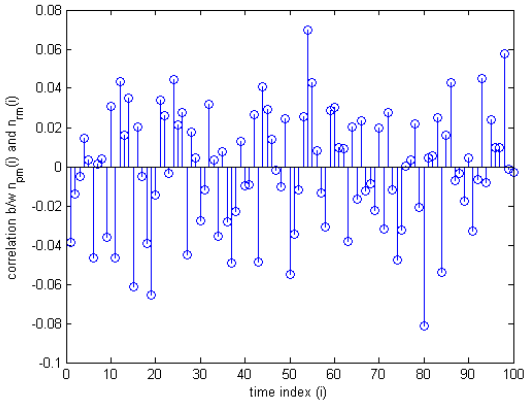

Figure 11. Correlation strength $\varphi_m(i)$ between primary and reference noise applied to an ANC at (t=i) time instants

signal. The obtained correlation strength between the desired response $d_m(i)$ and the output $r_s(i)$ varies between 0.65 to 0.9 as evidenced in Figure 12. It can be seen that even for $\phi_m(i)$ of less than 0.01, a correlation strength of greater then 0.6 between the desired response and the output of ANC is received.

D. BER Performance

According to [9], the values for channel 11 for f_c the center frequency, R_b the bit rate, T_c the chip duration, N_{cpb} the number of chips per burst, N_c the number of chips per symbol and N_{hop} the number of hops available per symbol are shown in Table I.

TABLE I.
PARAMETERS FOR CHANNEL 11

f_c(GHz)	R_b(Mb/s)	T_c(ns)	N_{cpb}	N_c	N_{hop}
7.987	0.85	2	16	256	8

Assuming synchronized transmitter and receiver and using the above mentioned parameters, the error performance of uncoded BPM-BPSK modulated IR-TH UWB

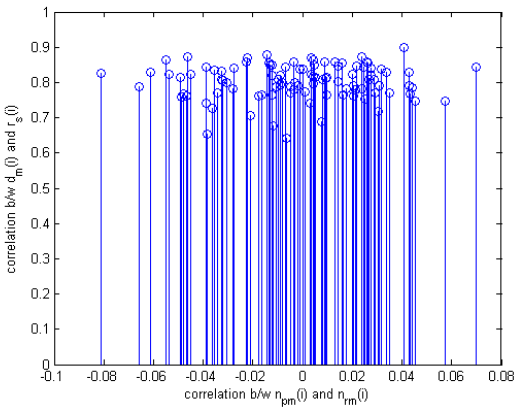

Figure 12. Correlation strength between desired response $d_m(i)$ and the output of ANC $r_s(i)$ due to applied $\phi_m(i)$

M biorthogonal signals was found in IEEE 802.15.4a industrial NLOS environment of CM8 in the presence of AWGN. Both the conventional SRake and adaptive SRake receivers employ 20 fingers in each Rake unit. The results are then compared in Figure 13 with those acquired via an AWGN correlation receiver. A significant performance improvement is achieved with an adaptive SRake which provides a gain of approximately 33 dBs and 5 dBs respectively with the conventional SRake and AWGN correlation receivers. It is therefore suggested, that instead of using hundreds of Rake fingers, an addition of an ANC drastically improves the error performance of conventional SRake in dense multipath propagated environment.

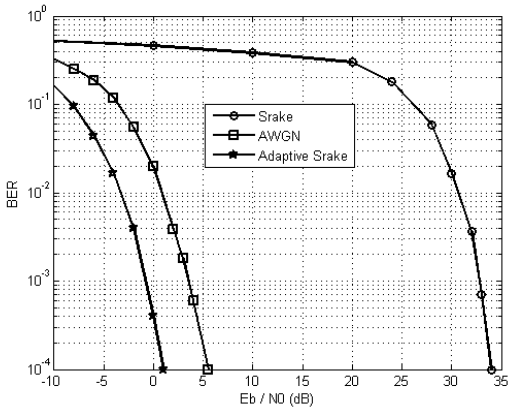

Figure 13. BER vs SNR of BPM-BPSK modulated uncoded IEEE 802.15.4a IR-TH UWB M-biorthogonal signals. SRake and adaptive SRake reception with 20 fingers: Average over 100 CM8 channel in the presence of AWGN in comparison over AWGN

VI. CONCLUSION

With an increasing demand of wireless technologies, our work serves as a first step toward deployment of IEEE 802.15.4a IR-TH UWB in dense multipath propagated noisy industrial environment. However, some adaption

was proposed to ensure required reliability and timeliness constraints of field level networks. The delay analysis indicates, that the proposed TH-TDMA MAC provides an urgent accessibility of asynchronous data during on going synchronous transmissions. It not only overcomes the delay of GTS management, but also avoids the use of CSMA/CA and slotted Aloha in UWB PHY. The channel characteristics were simulated and discussed. The proposed adaptive SRake architecture received the weak BPM-BPSK modulated IR-TH UWB biorthogonal signals with much decreased intricacy. The error performance depicts an appreciable improvement as compared to conventional SRake and AWGN receivers. As a direction for further research, the presented work should focus toward the ranging capability of the selected PHY and include mobility. Increase in the communication range based on multihop UWB networks and packet level simulations are another likely avenues for future research.

ACKNOWLEDGMENT

The first author expresses her gratitude to Higher Education Commission (HEC) of Pakistan, for providing scholarship opportunity and constant support for research funding.

REFERENCES

[1] X. Carcelle, T. Dang and C. Devic, "Industrial Wireless Technologies: Applications for the electrical utilities," *in Proc. IEEE International Conference on Industrial Informatics*, Aug. 2006, pp. 108–113.

[2] G. P. Hancke and B. Allen, "Ultrawideband as an Industrial Wireless Solution," *Pervasive Computing, IEEE Com Soc*, 2006, no. 4, pp. 78-85.

[3] D. Pellegrini, D. Miorandi, S. Vitturi and A. Zanella, "On the Use of Wireless Networks at Low Level of Factory Automation Systems," *IEEE Transactions on Industrial Informatics*, May 2006, vol. 2, no. 2, pp. 129– 143.

[4] G. Scheible, D. Dzung, J. Endresen and J. E. Frey, "Unplugged but connected - Design and Implementation of a Truly Wireless Real-Time Sensor/Actuator Interface," *IEEE Industrial Electronics Magazine*, 2007, vol. 1, no. 2, pp. 25–34.

[5] W. Zeng, H. Wang H. Yu and A. Xu,"The Research and Application of UWB Based Industrial Network," *in Proc. Ultrawideband and Ultrashort Impulse Signals, The Third International Conference*, Sept. 2006, pp. 153–155.

[6] W. Zeng, H. Wang H. Yu and A. Xu,"The Research of UWB Ad-Hoc Industrial Network," *in Proc. Communications and Networking in China, (ChinaCom)*, 25-27 Oct. 2006, pp. 1–4.

[7] J. Wang, W. Chen, H. Wang, and A. Xu, "Optimization on Delay and Slot Utilization for Fieldbus Control System Based on UWB," *in Proc. IEEE Conference on Industrial Electronics and Applications, ICIEA 2007*, May. 2007, pp. 1329-1333.

[8] A. Gupta and P. Mohapatra, "A survey on ultra wide band medium access control schemes," *Computer Networks 51, 2967-2993; available online at www.sciencedirect.com; Elesvier*, 2007.

[9] IEEE 802.15.4a-2007,"Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs); Amendment 1: Add Alternate PHYs," Mar. 2007.

[10] H. Gong, H. Nie and Z. Chen, "Performance Comparisons of UWB Selective Rake and Transmitted Reference Receivers under IEEE802.15.4a Industrial Environments," *in Proc. IEEE Wireless and Microwave Technology Conference (WAMICON)*, pp. 1–5, 2006.

[11] B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. R. Williams, E. Dong and R. C. Goodlin, "Adaptive Noise Cancelling: Principles and Applications," *Proceedings of the IEEE*, vol. 63, no. 12, Dec. 1975.

[12] I. M. S. Iacobucci and D. M. Benedetto, "Multiple access design for impulse radio communication systems," *in Proc. IEEE Communications Conference, ICC '02*, vol. 2, pp. 817–820, 2002.

[13] S. G. Glisic and P. A. Leppanen, *Wireless Communications:TDMA vs CDMA*, Kluwer Academic Publishers, 1997.

[14] J. Karedal, S. Wyne, P. Almers, F. Tufvesson and A. F. Molisch, "A measurement Based Statistical Model for Industrial Ultra Wideband Channels," *IEEE Transactions on Wireless Communication*, vol. 6, no. 8, 2007.

[15] A. F. Molisch, K. Balakrishnan, D. Cassioli, C. C. Chong, S. Emami, J. Karedal, A. Fort, B. Kanan, J. Kunisch, H. G. Schantz, K. Siwiak and M. Z. Win, "A Comprehensive Standardized Model for Ultrawideband Propagation Channels," *IEEE Transactions on Antenna and Propagation*, vol. 54, no. 11 Part 1, pp. 3151–3166, 2006.

[16] J. P. Thomesse, "Fieldbus Technology in Industrial Automation," *Proceedings of the IEEE*, vol. 93, no. 6, pp. 1073-1101, 2005.

[17] W. C. Chan, *Performance Analysis of Telecommunications and Local Area Networks*, Kluwer International Series in Engineering and Computer Science, 2002.

[18] H. Yang and B. Sikdar,"Performance Analysis of Polling based TDMA MAC Protocols with Sleep and Wakeup Cycles," *in Proc. IEEE International Communications Conference, ICC'07*, pp. 241–246, 2007.

[19] H. Zhang, T. A. Gulliver, "Biorthogonal pulse position modulation for time-hopping multiple access UWB communications," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1154–1162, May 2005.

[20] S. Haykin, *Adaptive Filter Theory; 3rd Edition*, Prentice Hall Information and System Sciences Series, 1996.

[21] A. F. Molisch, K. Balakrishnan, D. Cassioli, C. C. Chong, S. Emami and J. Karedal, "IEEE 802.15.4a Channel Model-final report," *Taskgroup 4a (TG4a),Tech. Report*, 2004.

Farah Haroon received her Masters of Electrical Engineering degree from NED University of Engineering and Technology, Pakistan in 2003. She is currently a PhD student in Telecommunications Field of Study, Asian Institute of Technology, Thailand. Her research interests are ultra wide band communications, propagation and reception in multipath channel and industrial wireless networks.

Kazi Mohiuddin Ahmed received his M.Sc. Engg degree in Electrical Engineering from the Institute of Communications, Leningrad, USSR, and the PhD degree from the University of Newcastle, NSW, Australia, in 1978 and 1983, respectively. Currently, he is a Professor of telecommunications in Asian Institute of Technology, Thailand. His current research interests are in wireless communications and communication networks. Mr. Ahmed is a member of IEEE, IEICE and life member of Bangladesh Electronic Society (BES).

Call for Papers and Special Issues

Aims and Scope.

Journal of Communications (JCM) is a scholarly peer-reviewed international scientific journal published monthly, focusing on theories, systems, methods, algorithms and applications in communications. 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 communications.

JCM invites original, previously unpublished, research, survey and tutorial papers, plus case studies and short research notes, on both applied and theoretical aspects of communications. These areas include, but are not limited to, the following topics:

- Signal Processing for Communications
- Multimedia Processing and Communications
- Communication QoS and Performance Modeling
- Cross-layer Design and Optimization
- Communication and Information Theory
- Communication Software and Services
- Protocol and Algorithms for Communications
- Wireless Communications and Networking
- Wireless Ad-hoc and Sensor Networking
- Broadband Wireless Access
- Cooperative Communications and Networking
- Optical Communications and Networking
- Broadband Networking and Protocols
- Internet Services, Systems and Applications
- P2P Communications and Networking
- Pervasive Computing and Grid Networking
- Communication Network Security
- Cognitive Radio Communications and Networking
- Hardware Architecture for Communications and Networking
- Parallel and Distributed Computing
- Satellite and Space Communications
- Emerging Communication Technology and Standards

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 8 to 12 papers, with each paper 8 to 12 pages of length, and the papers include:

- A Guest Editorial;
- 2-3 Invited Survey papers from world well-known scientists in the specific area;
- 6-10 Research papers reflecting the latest advances in the specific area.

The following information should be included as part of the proposal:

- Proposed title for the Special Issue
- An initial version of Call for Papers with specific topics covered in the Special Issue
- Name, contact, position, affiliation, and biography of the Guest Editor(s)
- Tentative time-table for the call for papers and reviews
- Potential authors and topics for the Invited Survey papers
- List of potential reviewers
- Plans for advertising the Call for Paper and attracting high-quality paper submissions

If a proposal is accepted, the guest editor will be responsible for:

- Submitting a final “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.
- Leading a fair and strict review process for the paper submissions, collecting 2-3 reviews for each paper before the final decision making. Authors should be informed the Author Instructions.
- Providing JCM 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.

In the Guest Editor Team building process, it is highly recommended that a world well-known scientist (e.g., IEEE or ACM Fellow) in the area is involved in this effort to promote the visibility of the Special Issue in the society. On the other hand, it is suggested to consider the geographic coverage of the team. Due to conflict-of-interest, the Guest Editors are not encouraged to submit their own papers to the Special Issue.

Recommended Papers from an International Conference

JCM accept recommendations from well-known International conferences. The conference organizer can recommend the best papers (top 5% of the accepted papers) to be considered in either JCM regular issue or a JCM special issue. A fast-track review process would be conducted by a JCM Editor for these recommended papers and the final decisions are made based on the review feedback.

The following information should be included as part of the proposal:

- The name of the conference/workshop, and the URL of the event
- A brief description of the event, including: number of submitted and accepted papers, and number of attendees. If these numbers are not yet available, please refer to previous events. First time conference would NOT be considered.
- Tentative time-table for the paper submission.

If a proposal is accepted, the conference organizer needs to submit the following items at a later stage:

- The list of the best papers (top 5% of the accepted papers) for recommendation
- The submitted conference paper draft and review feedbacks

If a conference contributes more than 5 papers that are finally accepted by JCM, the organizer would be invited to serve as a Guest Co-Editor of a JCM annual Special Issue, “SI on the Latest Advances in Communications and Networking”, and these accepted papers would be included in the same Special Issue. Otherwise, the accepted papers would be published in JCM regular issues.

More information is available on the web site at http://www.academypublisher.com/jcm/.