Scalable Self-Assembly for Ad Hoc Wireless Sensor Networks

Katayoun Sohrabi, William Merrill, Jeremy Elson, Lewis Girod, Fredric Newberg, and William Kaiser
Sensoria Corporation
200 Corporate Pointe, Suite 100
Culver City, CA 90230
310-641-1331x207
sohrabi@sensoria.com

Abstract— In distributed wireless sensing applications such as unattended ground sensor systems, remote planetary exploration, and condition based maintenance, where the deployment site is remote and/or the scale of the network is large, individual emplacement and configuration of the sensor nodes is difficult. Hence network self-assembly and continuous network self-organization during the lifetime of the network in a reliable, efficient, and scalable manner are crucial for successful deployment and operation of such networks. This paper provides an overview of the concept of network self-assembly for ad-hoc wireless sensor networks at the link level, with descriptions of results from implementation of a novel network formation mechanism for wireless un-attended ground sensor applications using a multi-cluster hierarchical topology and a novel dual-radio architecture.

1. Introduction

The goal of network assembly mechanism at the link level is to enable distributed formation of a connected wireless backbone and maintain this connectivity as the conditions in the network change, for example due to removal or addition of new nodes, over time. The network assembly is closely linked to the choice of the wireless channel access mechanism. We will briefly describe the fundamentals of this coupling. To ensure scalable operation of the network, a hierarchical network topology is preferred. This motivates a network self-assembly that generates a multi-clustered topology for a randomly deployed set of sensor nodes.

The multi-cluster mechanism enables the formation of a scalable network topology by allowing interconnected clusters in the network. Each cluster is formed independent of others, and is assigned a distinct channel. Certain nodes must be members of multiple clusters to allow network connectivity. This multi-cluster architecture enables the abstraction of the MAC dependent local channel operation so that each independent channel can be a fixed frequency, a TDMA schedule, or a CDMA spreading code, or even a local CSMA type channel on a fixed frequency. A node with a single radio must be switched between all the channels or clusters in which the node is a member of. This switching for most radios is not trivial, since it requires keeping accurate network synchronization on multiple channels in a serial fashion. For example, for frequency hopping radios, the transceiver must acquire the new code each time it switches to a different cluster. The switching time for commercially available radios may be as high as 2-10 seconds. Also, for most commercially available radios, the level of access to radio firmware that will allow the type of channel switching required is simply not available. To alleviate the need for switching between clusters a dual-radio node architecture has been implemented where each radio is able to participate in a different cluster.

2. Clustering

The network self-assembly mechanisms discussed here are part of a networking architecture that has been implemented on a hardware sensing platform [15]. The mechanisms discussed here have been demonstrated to work successfully using networks in excess of 10 nodes. Performance of variants of multi-cluster self-assembly in simulation and hardware implementation will be discussed is the rest of this paper.

We will describe two different distributed cluster formation mechanisms that take advantage of the dual radio architecture of the nodes. These mechanisms were developed for the DARPA/ATO's Self Healing Minefield application [15].

The first mechanism is called the "dual network clustering", (DNC). It uses only two distinct networks or channels in the entire randomly deployed field of nodes (RDF). The radios on a node are each tuned to a specific and well-known fixed channel. The first radio that wakes up and perceives a quiet channel becomes a cluster-head on that channel. The radios that arrive or tune to that channel at later times will become normal members of that cluster head. By simulation, it is shown that this mechanism is able to form a highly connected topology for general random topologies, for different sized networks. Simulation modeling of the DNC mechanism indicates that a dual-radio architecture is able to form large connected networks of size 400 and above in reasonable time intervals. The DNC has the advantage of being relatively simple to implement. However it lacks configurability to enable certain levels of control, in determining node membership in a cluster.

The second mechanism is called the "Rendezvous Clustering Algorithm" (RCA). This mechanism leverages one of the two radios on a node to tune it to a fixed signaling channel in the network. The node uses this channel to advertise it presence to those around it, gather advertisements from other nodes and clusters, and declare its choice(s) of nodes and clusters to attach. This mechanism takes advantage of a number of underlying utilities developed for the system, including a distributed time-synch mechanism, that allows precise time synchronization of processors on different nodes, and a link monitoring utility that allows reliable estimation of the state of radio links for the duration of lifetime of the nodes.

We will first provide a survey of some of the existing and relevant network formation mechanisms for sensor nodes as well as some existing related work. We will then describe the DNC and provide some simulation results on its performance. Next the RCA algorithms will be discussed along with its performance. A summary and closing will complete the paper.

3. Prior and Related Work

In this section a brief survey of prior or related work in the areas of multi-clustering and multiple radio or channel use for medium access will be provided. The section will also include references to some of the experimental and physically deployed ad-hoc networks reported in literature.

The concept of multiple linked clusters as an underlying topology structure for ad-hoc networks is not a new idea. Early examples include the classic Linked Clustering Algorithm of [1]. Later examples of the concept include the adaptive clustering of [2], and the generalized clustering algorithm of [3]. In the mechanisms described in this paper the clusters are formed directly as a consequence of the radio's MAC scheme, as will be described in later section under our first self assembly algorithm, or as is described under our second algorithm, the clusters are formed to satisfy a specific purpose, say to allow direct communication between closely located nodes. The cluster heads are chosen at random, and not based on their node ID.

There have been a number of empirical experiments that involve deployment of a number of wireless nodes in ad-hoc fashion, either as a MANET or a wireless sensor network. The list includes the system of [9] with 150-170 deployed motes, and those of the references therein, such as the system of [4], an ad-hoc network of 8 nodes using 802.11 radios, the system of [5], with five nodes and 802.11 radios, the network of [6] with 14 PC-104 nodes and RCP radios, the network of [7] with 11 Berkeley mote nodes, and that of [8] with 5 such nodes. The list of deployed sensor networks also includes an experimental network under the SensIT program for the Sitex02 exercises. This network was comprised of a 70 nodes deployed network of Sensoria WINS NG 2.0 nodes [10]. In these experiments, with the exception of [9] that is concerned with the emergent structure of the network under very simple rules of behavior, the purpose is usually to investigate an aspect of the network unrelated to initial self-assembly. For example performance of the routing mechanisms under mobility, or various MAC schemes were the focus. Therefore no automatic network self-assembly is performed. The nodes are either configured by hand ahead of time, for a specific topology, as in [20], or the issue of network assembly is not important because the number of nodes is small. So although ad-hoc wireless networks have been deployed, they did not use self-assembly in the sense discussed in this paper[1] and as is done for the SHM system.

The concept of using multiple channels simultaneously in the network has been investigated extensively in the literature. The general class of CDMA systems, and those that use spatial re-use fall in this category. However the notion of using multiple radios on the nodes, to have access to multiple channels simultaneously on the same node, in the context of ad-hoc multihop networks has not been examined very closely. An example of a MAC that proposes to use multiple channels is [11]. This system may be grouped together with the class of CDMA systems where the channel (or spreading code for the case of a CDMA system) to be used for each transmission is defined by the receiver. Each node in this scheme has its own virtual cluster, whose membership waxes and wanes as the traffic of the other members destined for the receiver node changes. In contrast, the clustering schemes described here generate a data bearing topology based on clusters that have fixed memberships and boundaries, also there is a local cluster head that controls the behavior of all the cluster members, regardless of their traffic load.

The sensor nodes described by [12] use two radios on board. However, one of the radios is a very simple and limited capability radio that uses very little power and is used solely to wake up the main radios when traffic arrives for it. This approach has the drawback that the limited channel sensing radio may easily be overwhelmed under realistic RF conditions, for example under intentional or un-intentional jamming.

---

[1] For example in the demonstration of http://webs.cs.berkeley.edu/800demo/ a large network of motes is reported to have been deployed. However since these motes use flooding, and a version of CSMA channel access scheme, they do not need to perform any network self-assembly so to speak, because a single packet will eventually reach some destination in the network via flooding or be dropped after multiple collisions. Due to lack of detailed description this is deployment is not included with the rest of the references.

Finally the class of MAC schemes that use a separate signaling channel for control are related to the second clustering and self-assembly mechanism described in this paper. A representative mechanism is the PAMAS scheme [13]. References to some other related schemes that use a separate signaling channel are included in the reference list of [13]. In PAMAS the signaling channel is used to coordinate packet transmission and reception, and also to turn off the radios when no traffic is forthcoming. By contrast, our signaling channel, which is called the rendezvous channel, is used to exchange connectivity and topology information between potential neighbors.

4. Dual Network Clustering (DNC)

In the DNC scheme, there are only two channels assigned to the entire network. Each cluster has a *cluster-head* radio that controls the cluster membership and parameters of channel access to the common medium of the cluster. The members of the cluster that are not a cluster-head are called *remotes*. Each node has two radios on board. The operation of the radios is coordinated at higher levels by the software that implements the clustering algorithm. However the two radio modems operate independent of each other, in terms of their over the air activities.

Each radio may be a *remote*, or a *cluster-head*. The two radios on the same node may not both be *cluster-heads* simultaneously. The two radios are tuned to two different well-known and fixed channels. We assume that transmissions on the two channels are orthogonal to each other, with the exception of occasional inter channel interference.

When a node wakes up its two radios are powered on, each tuned to one of the two networks (for example assigned the hopping pattern for a network if a frequency-hopping radio is used). Each radio will listen on its channel for a random period of time. If during this time, it hears a *cluster-head* on that channel, and gets attached to its cluster, the radio's search is over. However, if the listening period is expired, and no *cluster-head* is heard, or successful attachment to a cluster has not occurred and if the other radio on the node has not become a *cluster-head*, the radio will become a *cluster-head* on its own channel. If the radio cannot become a *cluster-head*, it will choose another random listening period and will do another search for this new listening period.

When the radio becomes a *cluster-head*, it will listen for a random period of time, to acquire new *remotes*. If within this period it does not acquire any, it will then go back to being a *remote* on the same channel at the end of listening time. The details of the scheme are given in Figure 2.

The specific radio used to implement DCA uses a TDMA access along with slow frequency hopping. The network topology is a star network, with the TDMA controller at the center of the star. The TDMA controller in this case corresponds to the cluster head described in previous paragraphs. For this specific radio, the cluster head will be called the Base. The Base radio sets the epoch and the length of each TDMA frame based on the number of remotes it has attached. The carrier frequency hops at the beginning of each TDMA frame. The Base uses a small period at the beginning of each TDMA frame to send out a synchronization signal. The *remotes* that are listening on this channel will be able to acquire the hopping pattern based on this synchronization signal. The Base will then transmit its own data during an assigned slot immediately following the synchronization period. After that each remote will transmit any packets it may have during its own fixed assigned slot. A TDMA frame structure is shown in Figure 1.

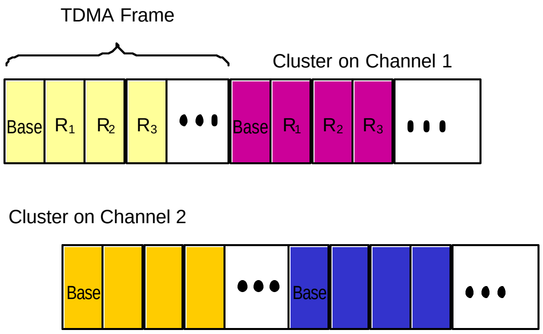

Figure 1 TDMA Frame Structure.

4.1 DNC Performance

The DNC was simulated using the PARSEC simulation environment [14]. The PARSEC models include the entire operation of the node, including the radio access scheme. A simple radio propagation model with a fixed maximum range for the radios was used[2]. Omni-directional propagation was assumed. Also a packet collision model was used, such that any packet overlap in time results in packet being dropped. This simulation model does not model the complexities of the hardware platform on which the code will run on, neither the overhead of the processing environment and the operating system (OS). The strength of this model is in its ability to show a proof of concept for the algorithms in early design stages.

The parameters of the DNC algorithm, such as the length of listening times for each mode as well as the maximum number of remotes per Base (or the cluster size) may be

---

[2] Note that the purpose of this simulation activity was to test and investigate the protocol internals, not to assess the performance of the self-assembly mechanism under various environmental conditions. Therefore the details of the radio channel propagation were not modeled here.

varied. The simulation model takes a text file that describes the node locations in a two-dimensional field, and runs the DNC algorithm for each node that is generated according to the topology file. A number of statistics are then generated for each run, including the length of time it take for 95% of the nodes to get connected together, time until the entire network is connected, average number of neighbors per node, total number of bases generated in the network, and the average number of hops between neighboring Bases. In the results given in this paper, the radio ranges are assumed to be 30 m.

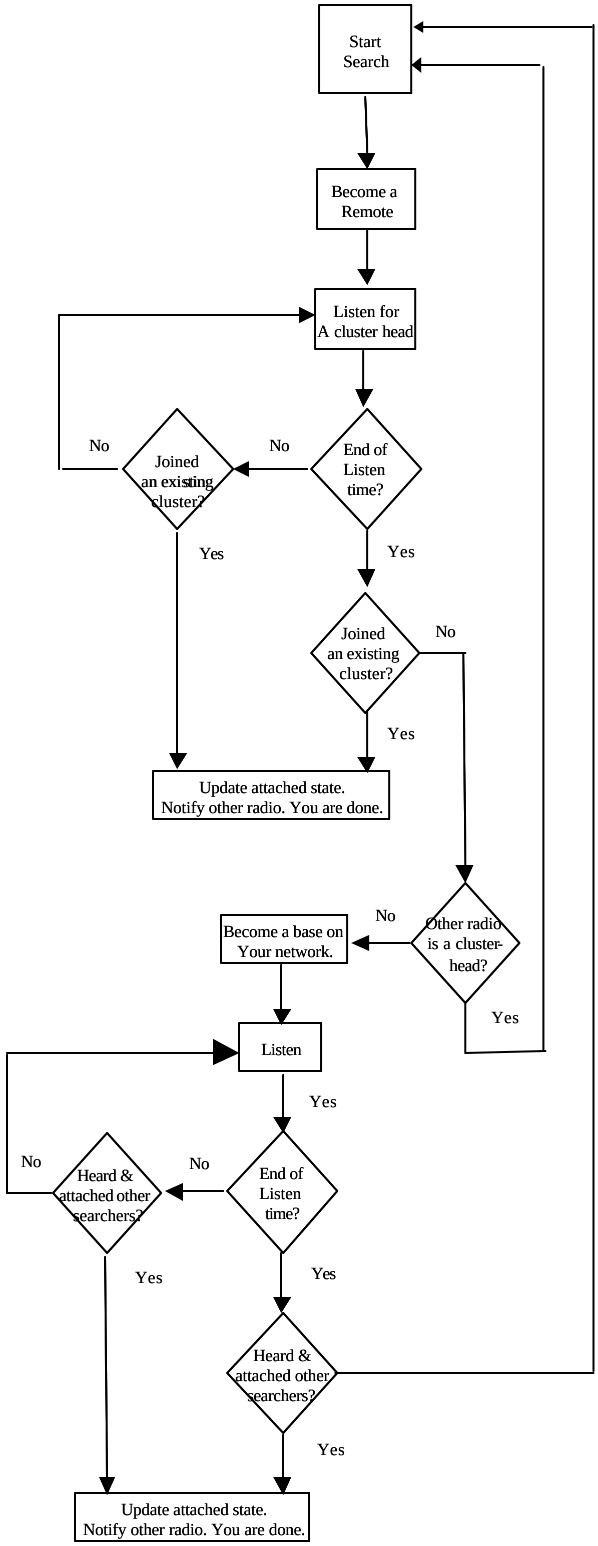

Figure 2. The flow chart of the DNC algorithm.

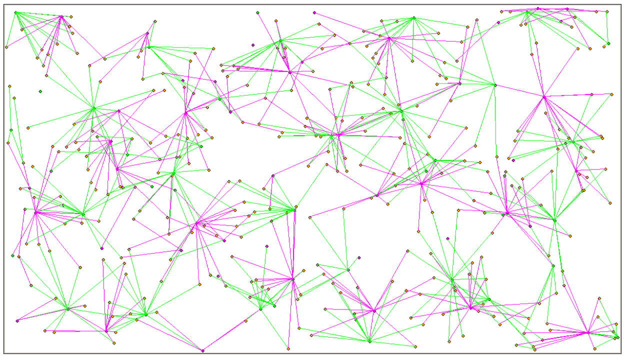

Figure 3 Graphical representation of a self-assemble network of 400 nodes, using the DNC scheme.

Figure 3 depicts a graphical visualization of a network of 400 nodes that was simulated with the DNC algorithm. In this graph, each link is colored to indicate the network to which it belongs. A cluster head is the node that is the center of convergence of multiple links. The first question that needed to be answered was whether the scheme is able to successfully connect random network topologies. Diagram of Figure 4 shows the average time needed to connect a network of size N with the DNC, where N=10,50, 200.

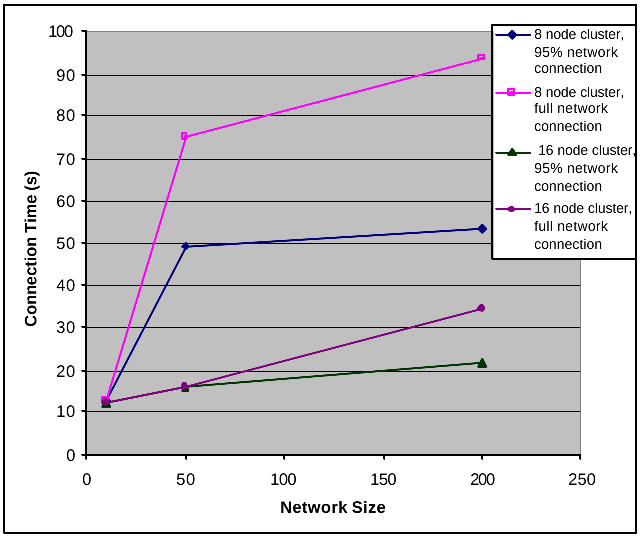

Figure 4. Average connection time for various network sizes under DNC.

These runs were conducted using a single topology file, and on the order of hundred runs to generate for different node start up times. The percentages of success of the schemes in connecting the network are given in Figure 5.

The size of a cluster has some effect on the behavior of the self-assembly algorithm. From the curves of Figure 4, we see that, as the cluster size is reduced, the time to connect the network is generally increased. It is also clear that the ability of the algorithm to get the network connected successfully is reduced, as is indicated in Figure 5. However in certain instances having a smaller sized cluster is preferred. For example, a large cluster size will increase the packet propagation delay in the cluster, since the traffic from many nodes must go through the cluster-head to reach its destination.

5. Rendezvous Clustering Algorithm (RCA)

Although the DNC is capable of generating a connected network topology that is at least 95% connected, it does not provide a lot of control over the choice of nodes that become members of a cluster. It was also seen that having a larger cluster size, allows a faster network connection time. However in many situations, it is desired to have a topology that has relatively smaller cluster sizes (this allows better throughput locally). The smaller size, potentially will allow lower transmit power levels, it will also allow a better level of distribution. At any rate we would like to have the means for realizing this capability.

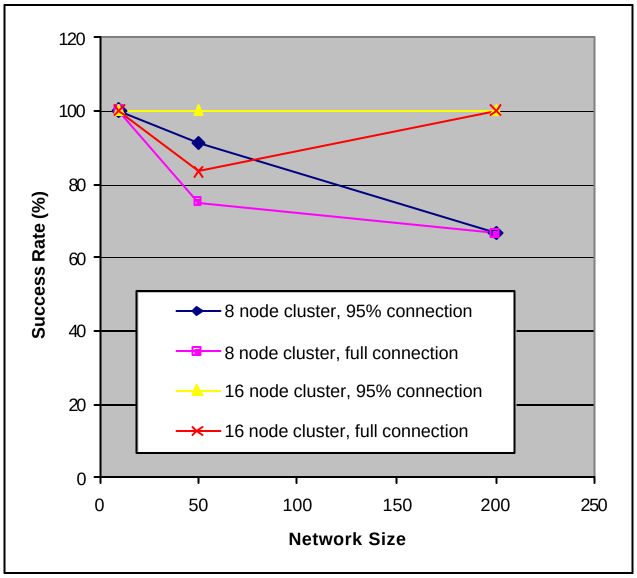

Figure 5. Success rates of connection of DNC for various network sizes for a single node lay-down topology.

Also, due to the peculiarity of the radios and the fact that the links are not always reliable, we need a means for monitoring the link quality. Generally we would also like to increase the connectivity of the networks to 100% under all conditions, when the topology and radio links warrant a connected network. These all motivated the design and evolution of the DNC to the Rendezvous Clustering algorithms. The main feature of the RCA is the use of the control channel to support a separate signaling mechanism for exchange of connectivity information separate from the channel used for exchange of data. The premise is that nodes will use a specific channel to exchange information about themselves. Once enough information is gleaned, the node will make a decision about its role (whether to become a cluster head or become ordinary cluster member on one or two different clusters), and also to what network its radios must be tuned. The cluster heads use one of their radios to surf the rendezvous channel (R-channel), while the other radio is used to talk to the cluster head's cluster members, after the cluster is formed. After getting connected, a remote will use both its radios to talk to two different clusters.

A newcomer node or a node that wishes to change its connectivity pattern will join the R-channel. Once on this channel[3][4], if it hears traffic from others, i.e. hear from a cluster head on the R-channel, it will get all the connectivity information it requires fromthat base. If no traffic is heard, the remote will itself become a temporary cluster-head on the R-channel, and start advertising its own existence and those of other remotes it has heard over the R-channel. This behavior will continue for a certain length of time, and then the radio will drop out of the R-channel.

At this point, the node will see if it has heard any interesting advertisements from other nodes. The advertisements are first pruned to keep only the eligible candidates. For example the advertisements from nodes with low RSSI may be dropped. The eligible advertisements are ranked according to a metric. For example, if in each advertisement the nodes also give their total number of attached neighbors, then both the RSSI and the other node's neighbor count may be combined to represent a metric. For example we may give priority to nodes that are near and not heavily connected.

At the end of the listening period and armed with the ordered list of received advertisements the new-comer node will decide as follows:

---

[3] The details of joining the channel are specific to the radios. For example in our implementation of star topology and TDMA base and remotes, a radio will become a remote and try to be acquired by a base on a specific network number that is assigned as the R-channel first.

[4] The rendezvous channel is a special and amorphous cluster. The membership in this cluster is constantly in flux, as new nodes join it to find eligible cluster for attachments for the first time, or to advertise their availability to accept new members as cluster-heads.

If the node has received information from a large enough number of other nodes that are also searching and with whom it can directly connect, it will become a cluster head on a specific new network. It will advise the other searching nodes to join it as members on the specific network. A new cluster if formed. If not enough searching nodes are heard, but it is possible to steal members from other clusters, the radio will become a cluster head, and advise the cluster heads in charge of the members to allow them to join the new cluster head on the newly forming network. Finally, the radio may become a cluster head, if its other radio is not attached, and it has heard somewhat smaller but still a non-zero number of searching members.

If none of the above applies the radio will try to become a non-cluster-head member of one of the networks it has heard from. If the radio is not successful, it will either start a new search, or turn itself off, and let the other radio on the node to start a search.

This process continues until all of the radios on the node become attached either as a cluster head in charge of a new cluster or as members of an existing cluster. Once a radio becomes a cluster-head, it will force the other radio on its node to switch to the R-channel, and advertise the state of the node on that channel periodically. The diagram of Figure 7 depicts the flow diagram of the RCA scheme. This diagram gives the general logical flow of actions for RCA.

5.1 RCA Performance

To investigate the behavior of the RCA algorithm for large network sizes an embedded code simulation/emulation software tool has been developed in house. This simulation tool allows emulation of the algorithms of interest by running the same embedded code that implements the algorithm for the embedded platform, on a different host environment such as a desktop machine. The details of the implementation of this simulation tool are outside the scope of this paper.

The results of network simulations using the simulation tool are given next. Figure 8 depicts the network for a simulation run of 100 nodes while the network is being formed. The R-channel here is depicted with the red network. At the point of time when the snapshot was taken, 19 different networks were assigned to different clusters in this system. The radio attached to node 80 in the center of figure is currently a base on the R-channel and is waiting to hear from some of the other radio on the R-channel (the squares with hollow red outline).

Figure 9 shows the time it takes for the networks to get connected under RCA for 10, 50, and 100, 120 node topologies. The important factor to be observed here is that the connection time tapers off with larger network sizes and appears to reach a plateau. This indicates that the network formation is indeed scalable with larger network sizes. Note that for all the simulation runs the algorithm was able to connect the network, and the percentage of success of the algorithm is 100% for all network sizes.

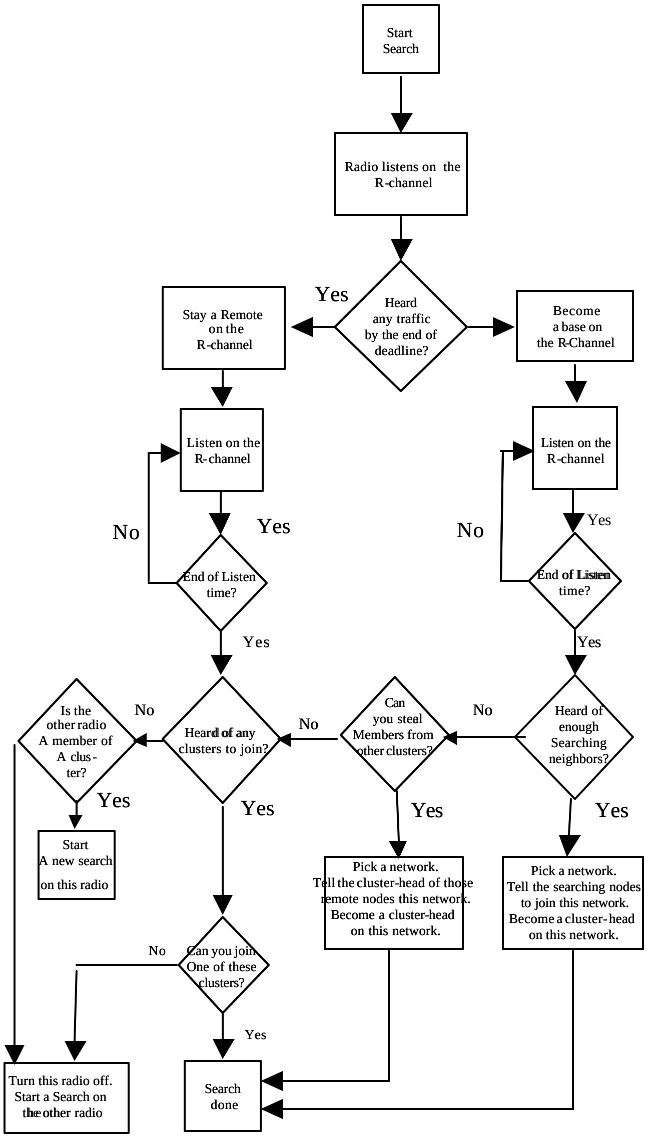

Figure 7.The Flow chart of the RCA mechanism for network self-assembly.

The simulation results indicate that the algorithm is capable of operating on relatively large network sizes. This data was gathered for a single node topology, over one hundred different simulation runs. This topology is the same used for the DNC algorithm.

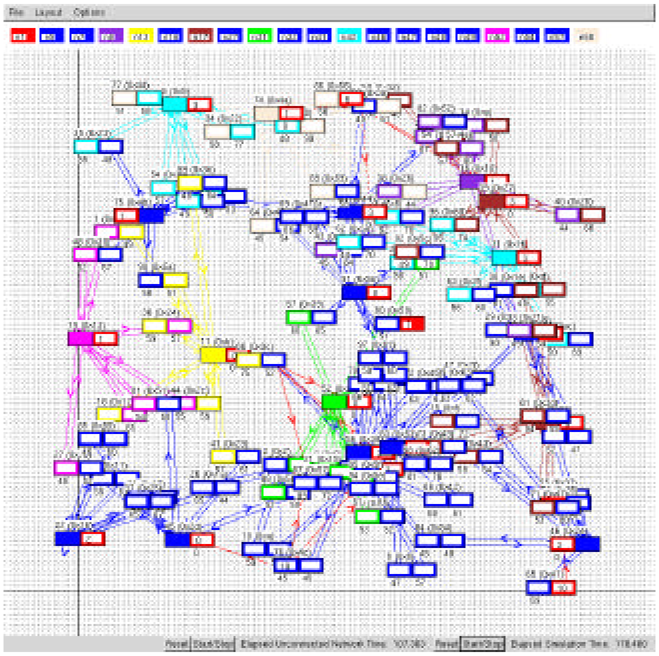

Figure 8. A simulated network of 100 nodes during network assembly.

The algorithm was also implemented on hardware nodes. The RCA was tested on networks of up to 20 nodes indoors and outdoors. Figure 10 shows the cumulative distribution Function (CDF) of the network connection time for a set of measurements of performance of RCA conducted on real hardware, for a network of 20 nodes. This network was setup inside the Sensoria offices building. For this set of runs the full connection time and time for 95% network connection were identical.

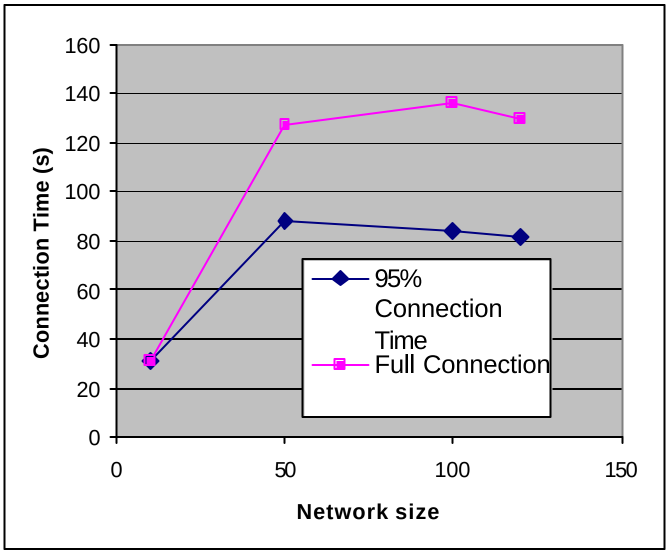

Figure 9. Average simulated connection time for 10, 50, 100, 120 node networks with RCA.

Finally the entire network was tested outdoors, with the nodes placed on the ground. The results for a network formation experiment with 20 nodes are given in Figure 12. This network was 95% connected after 79 sec, and full connectivity was achieved after 122 seconds. Note that the nodes were all booted at about the same time, so no significant time span was elapsed between the time the first and last nodes of the network were powered up and attempted to self-assemble.

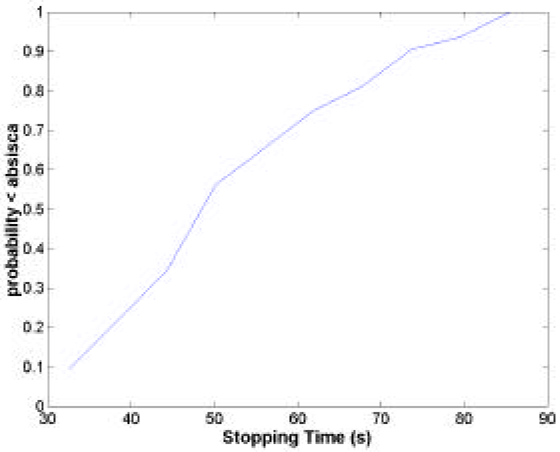

Figure 10. Cumulative distribution function of the full connection time of the RCA on real hardware.

6. Conclusion

The problem of network self-assembly for sensor networks was discussed. A short summary of related and prior work was provided. The concept of multi-clustered topologies for ad-hoc multihop wireless networks network was discussed, and the dual radio implementation were discussed. Two variants of multi-clustered network assembly for ad-hoc sensor networks were described. It was shown that both schemes are capable of forming connected networks for a wide variety of network sizes and topologies. The schemes scale well with network size.

7. Acknowledgements

This work was partially supported by DARPA/ATO contract DAAE30-00-C-1055. We would like to acknowledge the substantial guidance and advice provided by Thomas Altshuler with the DARPA/ATO office. We would also like to thank Sri Kumar with the DARPA/IXO office and the SensIT community.

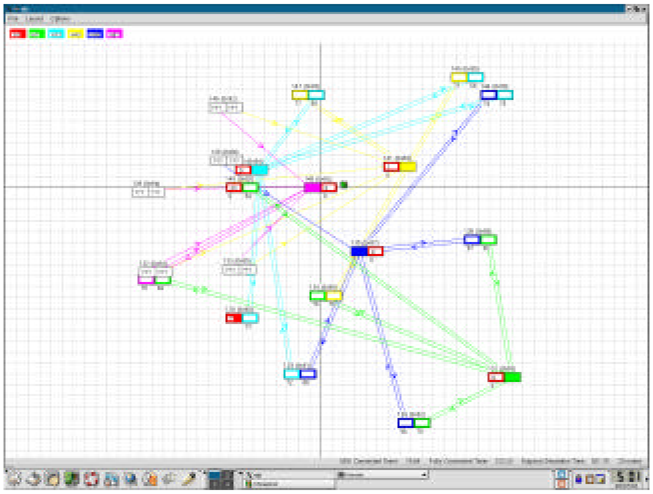

Figure 12 A snapshot of the network that shows a 20-node network out in the field being formed.

8. References

[1] D.J. Baker and A. Ephremides, "The Architectural Organization of a Mobile Radio Network vie a Distributed Algorithm, " *IEEE Trans. Comm.*, no11, Nov, 1981, pp. 1694-1701

[2] M. Gerla and J.T. Tsai, " Multicluster, Mobile, Multimedia Radio Networks", *Wireless Networks*, 1995, pp.255-65.

[3] Basagni, I. Chlamtac, and A. Farago, "A Generalized Clustering Algorithm for Peer-to-Peer Networks", *Workshop on Algorithmic Aspects of Communication* (satellite workshop of ICALP), Bologna, Italy, July 1997.

[4] D.A. Maltz, J. Broch, D.B. Johnson, "Experiences Designing and Building a Multi-Hop Wireless Ad Hoc Network Testbed", *CMU School of Computer Science Technical Report* CMU-CS-99-116, March 1999.

[5] S. Desilva and S.R. Das, "Experimental Evaluation of a Wireless Ad Hoc Network", *Proceedings of the 9th Int. Conf. On Computer Communications and Networks (IC3N)*, Las Vegas, October 2000.

[6] J. Heidemann et al., "Building Efficient Wireless Sensor Networks with Low-Level Naming", in *Proceedings of the Symposium on Operating Systems Principles*, pp. 146-159. October 2001.

[7] A. Woo, and D. Culler, "A Transmission Control Scheme for Media Access in Sensor Networks", *Mobicom 2001*.

[8] W. Ye, J. Heidemann, and D. Estrin, "An Energy Efficient MAC Protocol for Wireless Sensor Networks", IEEE INFOCOM, June 2002.

[9] D. Ganesan, B. Krishnamachari, A. Woo, D. Culler, D. Estrin, S. Wicker, "An Empirical Study of Epidemic Algorithms in Large Scale Multihop Wireless Networks", Intel Research, IRB-TR-02-003, Mar. 14, 2002.

[10] W. Merrill, K. Sohrabi, L. Girod, J. Elson, F. Newberg, and W. Kaiser, "Open Standard Development Platforms for Distributed Sensor Networks", *Proceeding of SPIE Aero Sense Conference 2002*.

[11] A. Muir, J. J. Garica-Luna-Aceves, "A Channel Access Protocol for Multihop Wireless Networks with Multiple Channels", *Proceedings of IEEE ICC '98*, Atlanta, Georgia, June 7--11, 1998.

[12] L. C. Zhong, R. Shah, C. Guo, and J. Rabaey, "An Ultra-Low Power And Distributed Access Protocol For Broadband Wireless Sensor Networks", IEEE Broadband Wireless Summit, Las Vegas, May 2001.

[13] S. Singh, M. Woo, and C. S. Raghavendra, "Power-Aware Routing in Mobile Ad Hoc Networks", The Fourth Annual ACM/IEEE International Conf. on Mobile Computing and Networking, pp. 181-190, 1998.

[14] http://pcl.ucla.edu/projects/parsec/

[15] W. M. Merrill, "Summary of the communication unit features developed in Phase I", Phase I summary for Sensoria Corporation under contract DAAE 30-00-O-1055, DARPA/ATO Self Healing Minefield, April 15, 2002.