Detection Performance in Self-Organized Wireless Sensor Networks

Robert E. Van Dyck
National Institute of Standards and Technology
Gaithersburg, Maryland, USA
vandyck@antd.nist.gov

Abstract— *One of the key design goals of a wireless sensor network is the ability to detect, in a distributed manner, events of interest. At the same time, it is desired that the network has the ability to self-organize, forming clusters of sensor nodes. In this work, we consider the interaction of the self-organization algorithm and the distributed detection processing. The effects of the connectivity of the wireless network and the sensitivity of the sensors on the detection performance are studied. Complete receiver operating curves are obtained for a number of scenarios.*

Keywords—**Wireless Sensor Networks, Self-Organization, Parley, Distributed Detection.**

I. Introduction

In the last few years, there has been much research on the best ways to design wireless sensor networks. Aspects of this work includes systems architecture and tradeoffs [1], [2], wireless network protocols, naming, and routing. While all of these components are essential for a practical system, we suggest [3] that one of the primary performance metrics should be the ability of the sensor network to detect events of interest. This concept is consistent with the approach taken by Iyengar and Prasad [4], who are primarily not concerned with the wireless component.

Since wireless sensor networks will often have many nodes and will be deployed in remote environments, it is important that they have the ability to self-organize. This process should be done in a way that optimizes the performance of the sensor network applications. The literature on wireless self-organization goes back at least twenty years to the Linked Cluster Algorithm (LCA) of Baker and Ephremides [5]. More recent work includes that of Gerla *et al.* [6], [7], Sharony [8], Amis *et al.* [9], Kozat *et al.* [10], and Mirkovic *et al.* [11]. In this paper, we study the interaction of the self-organization algorithm and the distributed detection processing. The primary goal is to determine the parameters that are responsible for the performance, and then to draw conclusions about how to optimize them. For concreteness, the self-organization is done using the LCA, while the distributed detection uses a robust version [12] of the parley algorithm [13].

[0]This work was supported by the Advanced Research and Development Activity (ARDA) under contract number 706400.

II. Self-Organization

The linked cluster algorithm [5] uses an underlying TDMA frame to allow the wireless sensor nodes to communicate among themselves. The total number of sensor nodes in the network is assumed to be fixed, and a slot in the frame is allocated for each node. This medium access control (MAC) structure is not efficient for large networks, say on the order of 1,000 or more nodes, and we discuss the implications of this fact in the conclusions section. The algorithm is distributed, with each node independently determining its neighbors.

The overall algorithm has two stages: (1) formation of the clusters, and (2) the linking of clusters using gateways. The aim of the first stage is to allow each node to construct part of the network's global connectivity matrix. Each local matrix is incomplete, but it is consistent with the global one. Two TDMA frames are used to exchange the required connectivity information. In the first frame, each node transmits a packet containing the nodes that it has heard from previously in the frame. During the second frame, each node broadcasts its full connectivity row.

Once the two frames have been transmitted and received, stage two begins. Each node uses its local connectivity matrix and its identification number to classify itself as either: (1) a cluster head, (2) a gateway, or (3) an ordinary node. At the completion of the algorithm, a backbone network is formed that links all the cluster heads, through gateways if necessary. One important heuristic of the algorithm is that if two nodes cover each other, the higher numbered node becomes the cluster head. This choice was made to reduce the transmission requirements during stage one, but it is not an optimal choice, both for forming the backbone network and for the distributed detection.

III. Distributed Detection

Each cluster runs a separate version of the soft decision parley algorithm [12]. This algorithm is a variant of the original algorithm proposed by Swaszek and Willett [13]. Here, we briefly review its operation.

A. Soft-decision Parley Algorithm

Assume each of N sensors measures a phenomenon and computes a likelihood function $\Lambda(r_i)$, for sensor $i = 1, \cdots, N$. The decentralized test performed at each sensor during each iteration of the original parley algorithm is given by

$$\Lambda(r_i) \underset{u_{i,m}=0}{\overset{u_{i,m}=1}{\gtrless}} \lambda_{i,m}. \tag{1}$$

In Eq. (1), $u_{i,m}$ is the hard decision and $\lambda_{i,m}$ is the local threshold, both for the ith sensor at iteration m. These thresholds need to be determined.

A consensus occurs if all nodes choose the same hypothesis. The main concept of the original parley algorithm is given by the following proposition. "If the thresholds at each round of the parley satisfy $\prod_{i=1}^{N} \lambda_{i,m} = \lambda$, a consensus decision, if reached, matches the centralized decision exactly" [13]. From this, one can show that the optimal local thresholds are

$$\begin{aligned} \lambda_{i,m+1} &= (\lambda \prod_{k=1}^{N} \frac{Pr(s_{k,m} \leq \Lambda(r_k) \leq t_{k,m}|H_0)}{Pr(s_{k,m} \leq \Lambda(r_k) \leq t_{k,m}|H_1)})^{1/N} \\ &\quad \times \frac{Pr(s_{i,m} \leq \Lambda(r_i) \leq t_{i,m}|H_1)}{Pr(s_{i,m} \leq \Lambda(r_i) \leq t_{i,m}|H_0)}. \end{aligned} \tag{2}$$

Here, $s_{k,m}$ and $t_{k,m}$ are the minimum and maximum values of $\Lambda(r_k)$, given the past hard decisions $u_{k,n}, n = 1, \cdots, m-1$. λ is the threshold of the centralized detection problem, and its choice selects the point on the receiver operating curve (ROC).

The soft decision parley algorithm [12] makes two main changes to the original algorithm: Firstly, it uses a two bit quantizer instead of making hard decisions. To ensure convergence, the quantizer is designed so that the first bit determines which hypothesis is selected, as in the original algorithm. The second bit then gives a confidence measure of this decision. Secondly, it relaxes the constraint that all sensor nodes agree on the hypothesis. If some fraction of then, say 80%, agree, then the parley process stops.

The algorithm works as follows. The minimum and maximum values are chosen for each node's likelihood function, and the parley process begins. At each iteration, the likelihood function at each node is computed and compared to the local threshold. This threshold is calculated using the *a posteriori* probabilities determined by numerically integrating the probability density functions of the other nodes' likelihood functions between the minimum and maximum values, as shown in Eq. (2). A soft decision is made, and this value is then broadcast to all other nodes. Since each node knows the previous minimum and maximum values, it can reduce these values for the next iteration by using the two information bits received from each of the other sensor nodes and its knowledge of the quantizer.

B. Robustness: Sensor Performance versus Distance

One of the fundamental assumptions of both versions of the parley algorithm is that the observations received at the sensor nodes (in a cluster) are all independent and identically distributed. In a wireless sensor network, this assumption will almost always be violated. For example, the sound pressure received at an acoustic sensor will be a function of the distance between the acoustic emitter and the sensor. In this paper, we study the Gaussian shift in mean problem, where the two hypotheses are given by

$$p_R(r) = \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{-(r \pm s)^2}{2\sigma^2}, \tag{3}$$

where the mean is s for hypothesis one and $-s$ for hypothesis zero. Alternatively, one can shift these so that for hypothesis zero the mean is zero and for hypothesis one the mean is $m = 2s$. To account for the distance from the emitter to the sensor, Eq. (3) is modified so that the mean distance between the two hypotheses is given by

$$m' = \frac{m}{\alpha\, d^{\beta}}, \tag{4}$$

where d is the physical distance from the emitter to the sensor, and α and β are attenuation factors that determine the propagation properties of the sensed medium. For example, acoustic, seismic, and magnetic sensors will generally all have different values. The sensitivity of the results to α and β is considered in the section on simulation results.

Consequently, there is a fundamental mismatch between the probability density functions assumed at each sensor node and the actual density functions. For nodes close to the emitter, the signal will be stronger, and the two hypotheses will have a larger mean separation, m'. Conversely, for nodes far from the emitter, the hypotheses will be statistically much closer. The use of soft decisions helps make the algorithm robust to the mean distance between the density functions.

IV. Simulation Results

In this section, we look at two sample networks. First, let us define a few terms. The neighborhood, Γ_v, of node v is the set of all nodes that are connected to v; the number of these nodes is denoted by k_v. We use two statistics to characterize the networks in a graph-theoretic sense. The

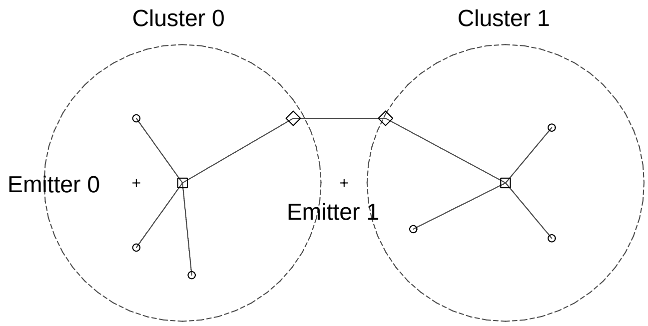

Fig. 1. Self-organized sensor network 1. The cluster heads are squares, the gateway nodes are diamonds, and the ordinary sensor nodes are circles. The transmission areas of the two cluster heads are indicated by the two large circles.

TX_{rad}	1.5	4.5
γ	0.667	0.933
L	2.715	1.000

TABLE I
STATISTICS FOR NETWORK 1.

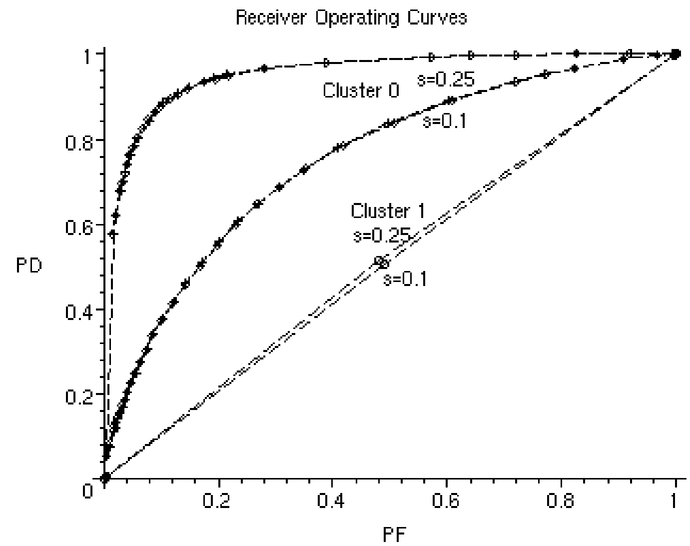

Fig. 2. Receiver operating curves for emitter 0. Crosses indicate cluster 0, circles indicate cluster 1, and diamonds indicate the single cluster of 10 sensor nodes. $\alpha = 1.0$ and $\beta = 2.0$. 100,000 trials per data point.

primary one is the clustering coefficient, which is defined for each node by

$$\gamma_v = \frac{|E(\Gamma_v)|}{\begin{pmatrix} k_v \\ 2 \end{pmatrix}}, \tag{5}$$

where $|E(\Gamma_v)|$ is the number of edges in the neighborhood of node v, and $\begin{pmatrix} k_v \\ 2 \end{pmatrix}$ is the total number of possible edges in Γ_v [15, p. 33]. The average of γ_v over all nodes gives the overall clustering coefficient γ, which satisfies $0 \leq \gamma \leq 1$. A value of one indicates that the network consists of a number of fully connected components, while a value of zero says that no neighbors of any node are adjacent to each other.

The second statistic is the characteristic path length, L. It is defined to be the median of the means of the shortest path lengths connecting each node to all other nodes. One calculates the distance from a node v to all other nodes and finds the average distance, $\overline{d}_v$. L is the median over the set $\{\overline{d}_v\}$ [15, p. 29].

A. Network 1

Figure 1 shows a ten node network after it has been clustered using the LCA. The radio transmission radius, TX_{rad}, is set to a value of 1.5 (normalized) units, which results in two clusters. The figure also shows the location of two emitters. Only the links from the ordinary nodes to the cluster heads and the backbone links are shown. These are the links actually used to route messages *after* a detection decision is made. During an iteration of the parley algorithm, each node in a cluster sends a message that is received by all the other nodes in the cluster (relayed through the cluster head if necessary). Table I shows the statistics of this network for two different values of the radio transmission radius [1].

The complete receiver operating curves for emitter 0 are shown in Figure 2 for both radio transmission radii and for two different values of the nominal mean, m. When $TX_{rad} = 1.5$, clusters 0 and 1 make independent decisions, in each case using five sensors, and when $TX_{rad} = 4.5$, only a single cluster using all ten nodes is formed. For the first set of experiments, the sensor nodes assume that the underlying probability density functions are separated by a mean of $m = 0.5$ ($s = 0.25$); this value is correct if the sensor is one unit away from the emitter, but the actual observations are changed according to Eq. (4). Since cluster 1 is far from the sensor, it is ineffective in detecting the target. However, cluster 0 does quite a good job, seen by examining the top curve indicated by crosses. The performance of the ten node cluster is also shown in the figure; its curve overlaps that of cluster 0.

In a second set of experiments, the (nominal) mean is

[1] Since there are ten nodes, the value of L for $TX_{rad} = 1.5$ is the average of the middle two $\overline{d}_v$ values.

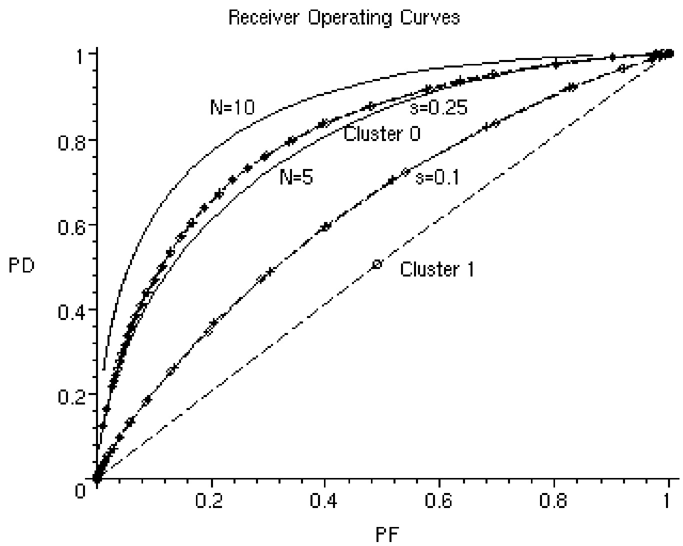

Fig. 3. Receiver operating curves for emitter 0. Crosses indicate cluster 0, circles indicate cluster 1, and diamonds indicate the single cluster of 10 sensors. $\alpha = 2.0$ and $\beta = 2.0$. 100,000 trials per data point. The solid curves indicate the best performance that can be obtained with the given number of sensors using centralized detection.

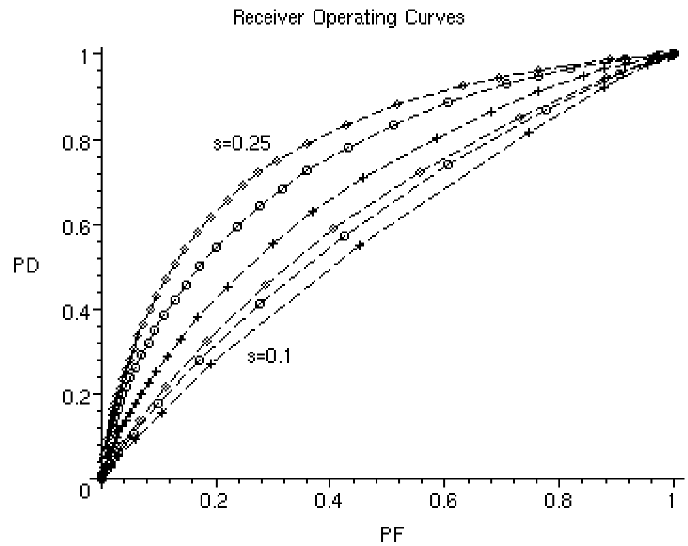

Fig. 4. Receiver operating curves for emitter 1. Crosses indicate cluster 0, circles indicate cluster 1, and diamonds indicate the single cluster of 10 sensors. $\alpha = 1.0$ and $\beta = 2.0$. 100,000 trials per data point.

$m = 0.2$. Again, the actual value varies at each sensor with the distance to the emitter. Since the two hypotheses are now statistically much closer to each other, the performance is degraded. Cluster 0 and the ten node cluster give equally good results, while cluster 1 is still of no use. We suggest that part of the reason that these two configurations give the same results is because of the use of soft decisions in the parley algorithm. The sensors that are closer to the emitter have better data, and their decisions are given more weight in the distributed detection process.

Figure 3 shows the receiver operating curves when the propagation constant $\alpha = 2.0$. This scenario represents a sensor network where each sensor is less effective for a given distance from the emitter. Another way to think of this network is that for a given propagation constant, the network is less dense; that is, the nodes are relatively further apart. Again, cluster 0 and the ten node cluster have effectively the same performance for both $m = 0.5$ and $m = 0.2$. The location of the emitter is, to a large extent, responsible for this result. Also shown in the figure are analytical curves for centralized detection using the observations from five and ten sensors, respectively [2]. The general trends are similar to the case above where $\alpha = 1.0$ and the mean is smaller, *i.e.* $m = 0.2$. This result shows that the performance degrades similarly with distance and with the sensitivity of the sensors. To some extent, one can trade off these factors. Increasing β will lead to a faster decrease in performance vs. increasing α, which may be a problem given certain terrain.

Figure 4 shows the curves for emitter 1, which is located between clusters 0 and 1. Now, the use of all ten sensors provides superior performance, since it is able to use sensors that surround the emitter. Cluster 1 does better than cluster 0, primarily because it has more sensors closer to the emitter. The advantage of the ten node cluster over cluster 1 is greater when $m = 0.5$ ($s = 0.25$) then when $m = 0.2$, mainly because the two bit quantization in the parley algorithm is more accurate when the means are further apart. Increasing the number of bits in the quantizer may further improve performance

B. Network 2

Now, we consider a 100 node network to see how well the above results generalize. The nodes are randomly placed in a square of size 10×10, according to a uniform distribution in each dimension. Table II shows the clustering coefficient for four values of the transmission radius.

[2] The reason that cluster 0 does somewhat better than the "optimal" performance for five sensors is due to the difference in probability distributions. The centralized detection curves assume that all the observations are independent and identically distributed with a value of $m' = m = 0.25$. The optimal detector using these observations is derived in [14, pp. 36-38]. In the actual system, the value of m' depends on the distance from the sensor node to the emitter, as given by Eq.(4). For $\alpha = 2.0$ and $\beta = 2$, the value of m' is sometimes greater than 0.25, and the performance is improved.

TX_{rad}	1.5	2.5	3.5	4.5
γ	0.597	0.671	0.707	0.748

TABLE II
STATISTICS FOR NETWORK 2.

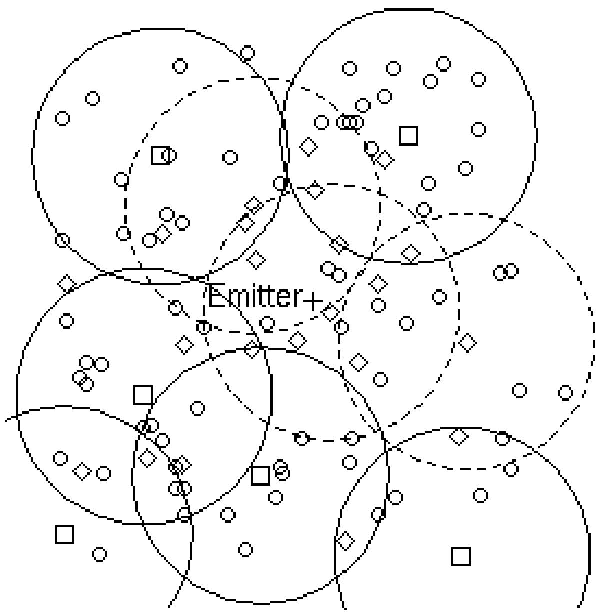

Fig. 5. Self-organized sensor network 2. The cluster heads are squares, the gateway nodes are diamonds, and the ordinary sensor nodes are circles. For $TX_{rad} = 2.5$, there are fifteen clusters, but only six nodes that classify themselves as cluster heads; the transmission areas of these nodes are indicated by the solid large circles. Nine gateway nodes also act as cluster heads. The transmission areas for three of these nodes are indicated by the dashed large circles.

A radius of 2.5 gives a value of 0.671, which is about the same as the ten node network with $TX_{rad} = 1.5$. Figure 5 shows the resulting clustered network for $TX_{rad} = 2.5$. Because of the larger number of nodes, the links are not plotted. The emitter is placed in the center of the square.

The linked cluster algorithm creates 30 clusters when $TX_{rad} = 1.5$. Figure 6 shows the receiver operating curves for the four best clusters. Cluster 17 does by far the best. This is a cluster of seven nodes with the cluster head located at (5.515, 6.140); cluster 19 consists of three nodes, with the cluster head at (5.380, 4.817). Figure 7 shows the ROCs when $TX_{rad} = 2.5$; the cluster numbers are new for this figure. Now, there are only 15 clusters. Cluster 9, containing six nodes and with its cluster head also at (5.515, 6.140), gives almost the same performance as cluster 17 does. It is interesting to note that even

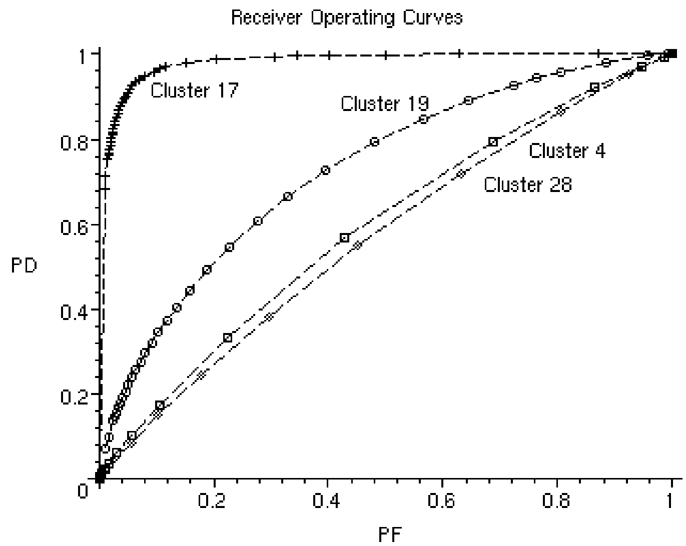

Fig. 6. Receiver operating curves for network 2. $TX_{rad} = 1.5$. $\alpha = 1.0$ and $\beta = 2.0$. 100,000 trials per data point.

though the transmission radius is now larger, the cluster actually has fewer nodes. This phenomenon is caused by the heuristic in the LCA that higher numbered nodes are more likely to become cluster heads. This cluster head has an identification number of 86. Cluster heads 91, 98, and 99 have clusters of 16, 26, and 19 nodes, respectively.

The main conclusion that can be drawn from these experiments is that a properly placed cluster with fewer nodes can provide performance as good as that of a larger number of nodes spread over a bigger geographic area. Additionally, fewer iterations are needed for convergence of the parley algorithm for the smaller cluster [3], and a lower transmitter power is used. This not only saves battery life, but it also reduces the probability that the network itself is detected by an adversary. However, the larger cluster is generally able to provide better performance when the emitter is further away.

V. Conclusions and Present Work

The linked cluster algorithm was designed for a moderate sized network, perhaps up to about 100 nodes. The requirement of reserving a TDMA slot for each node in the network, at least during the self-organization phase, is a restriction that we would like to remove. Moreover, the heuristic that the higher numbered nodes are more likely

[3]Reference [12] shows plots of the number of iterations necessary to achieve convergence as a function of the detection threshold. It also shows similar plots when the radio transmissions are corrupted by channel errors. In all cases studied, fewer iterations are needed for a five node cluster vs. a ten node one. Also, fewer total bits are transmitted per iteration.

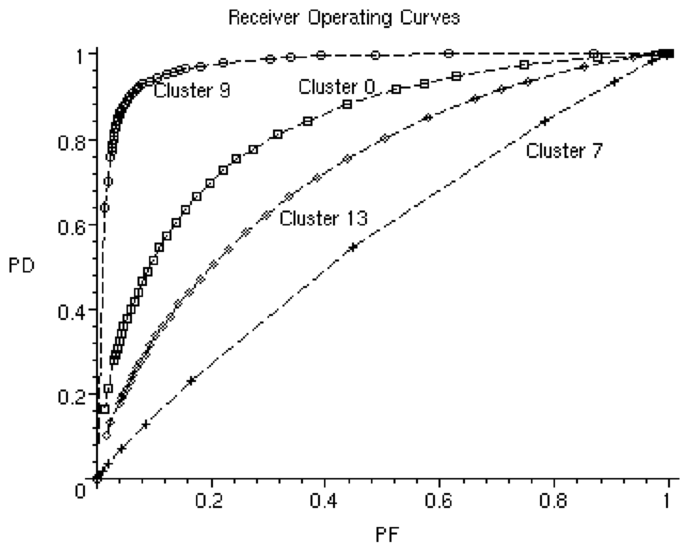

Fig. 7. Receiver operating curves for network 2. $TX_{rad} = 2.5$. $\alpha = 1.0$ and $\beta = 2.0$. 100,000 trials per data point.

to become cluster heads can lead to cluster formations that are not particularly good for distributed detection. Present work is taking the architecture proposed by Subramanian and Katz [2] and turning it into a complete self-organization implementation. This algorithm will then be integrated with the distributed detection processing.

The choice of the soft decision parley algorithm for the distributed detection is motivated by three factors: (1) When the observations are independent and identically distributed and the underlying probability density functions are known, the detection performance of the algorithm matches that of a centralized detector having access to all the observations. (2) The algorithm is robust to uncertainty in the physical distance between the sensor nodes and the emitter, which leads to a change in the mean distance between the hypotheses. It is also robust to bit errors in transmission of the quantized information. Ongoing work is studying robustness to the choice of probability density functions. (3) By sending only very short packets (presently two bits) among the sensor nodes, the sensor network itself is harder to detect. These bits can be transmitted using low probability of intercept communication techniques. Only after a target is detected will a larger packet be sent through the backbone network, and even this packet does not need to be particularly large.

Two other issues are also of interest. While the self-organization algorithm forms the clusters and creates a backbone network, it does not specify how to route the decision through the network. The choice of a good routing protocol is therefore important. Perhaps even more critical is the design of a suitable MAC layer for use with the sensor network. While there is a considerable literature on this topic for wireless networks, we suggest that it is important to study how the self-organization, the distributed detection algorithm, and the MAC layer interact. The use of clusters should simplify this problem to some extent. Since the packet transmissions associated with the parley iterations are only intended for the sensor nodes of the given cluster, they can be sent at a low enough power to reduce interference with other clusters.

REFERENCES

[1] G. J. Pottie, "Wireless sensor networks," *IEEE Information Theory Workshop*, pp. 139-140, June. 1998.

[2] L. Subramanian and R. H. Katz, "An architecture for building self-configurable systems," *Proc. IEEE MobiHoc*, 2000.

[3] R. E. Van Dyck and L. E. Miller, "Distributed sensor processing over an ad-hoc wireless network: simulation framework and performance criteria," *IEEE Milcom*, Mclean, VA, Oct. 2001.

[4] S. S. Iyengar and L. Prasad, "A general computational framework for distributed sensing and fault-tolerant sensor integration," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 25, pp. 643-650, Apr. 1995.

[5] D. J. Baker and A. Ephremides, "The architectural organization of a mobile radio network via a distributed algorithm," *IEEE Trans. on Comm.*, vol. COM-29, pp. 1694-1701, Nov. 1981.

[6] M. Gerla and J. T.-C. Tsai, "Multicluster, mobile, multimedia radio network," *Wireless Networks*, pp. 255-265, 1995.

[7] C. R. Lin and M. Gerla, "Adaptive clustering for mobile wireless networks," *IEEE J. Sel. Areas in Comm.*, Vol. 15, No. 7, pp. 1265-1275, Sept. 1997.

[8] J. Sharony, "An architecture for mobile radio networks with dynamically changing topology using virtual subnets," *Mobile Networks and Applications*, pp. 75-86, 1996.

[9] A. D. Amis, R. Prakash, T. H. P. Vuong, and D. T. Huynh, "Max-min D-cluster formation in wireless ad hoc networks," *Proc. IEEE Infocom*, pp. 32-41, 2000.

[10] U. C. Kozat, G. Kondylis, B. Ryu, and M. K. Marina, "Virtual dynamic backbone for mobile ad hoc networks," *Proc. IEEE Int. Conf. on Comm.*, Helsinki, Finland, 2001.

[11] J. Mirkovic, G. P. Venkataramani, S. Lu, and L. Zhang, "A self-organizing approach to data forwarding in large-scale sensor networks," *IEEE Int. Conf. on Comm.*, Helsinki, Finland, 2001.

[12] R. E. Van Dyck, "Distributed detection using parley with soft decisions," *Proc. Conf. on Information Sciences and Systems (CISS)*, Baltimore, MD, pp. 537-542, March 2001.

[13] P F. Swaszek and P. Willett, "Parley as an approach to distributed detection," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 31. no. 1, pp. 447-457, Jan. 1995.

[14] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, vol. 1*, John Wiley & Sons, New York, 1968.

[15] D. J. Watts, *Small Worlds the Dynamics of Networks between Order and Randomness*, Princeton Univ. Press, Princeton, NJ, 1999.