Fusion of Threshold Rules for Target Detection in Sensor Networks

Mengxia Zhu, Song Ding, Richard R. Brooks, Qishi Wu, Nageswara S.V. Rao, and S. Sitharama Iyengar

---

We consider a network of sensors distributed in a target area providing environmental measurements that are subject to normally distributed, independent additive noise. Each sensor node applies a threshold rule to the measurements to decide the presence of a target; the distance to the target together with the threshold value determines its hit and false alarm probabilities or rates using a signal attenuation model. We propose a centralized threshold-OR fusion rule for combining the individual sensor node decisions. Under the statistical independence of sensor measurements, we derive fusion threshold bounds using Chebyshev's inequality based on individual hit and false alarm probabilities but without requiring *a priori* knowledge of the underlying probability distributions. We derive conditions to ensure that the fused method achieves a higher hit rate and lower false alarm rate compared to the weighted averages of individual sensor parameters. The simulations using Monte Carlo method illustrate significant detection performance improvements of the proposed fusion approach.

Categories and Subject Descriptors: I.2.9 [**Artificial Intelligence**]: Robotics–Sensors

General Terms: Wireless sensor network, binary decision fusion

Additional Key Words and Phrases: hit rate, false alarm rate, Bayes, Neyman-Pearson, Chebyshev inequality, Type I error, Type II error, ROC curve

---

1. INTRODUCTION

Wireless sensor networks are being increasingly deployed in a number of applications such as detection of missiles, identification of chemical, biological or nuclear plumes, monitoring of rain forests, and command and control operations in battlefield environments. In general, such sensor networks could be quite varied, ranging from thousands or millions of unattended tiny sensing devices to a small number of sensor nodes equipped with large instruments and workstations. Since sensor failures are common, redundant nodes are usually necessary to ensure uninterrupted and reliable operations. Measurements from individual sensors are sent to a fusion center over wireless channels to achieve a more accurate situation assessment.

We consider a network of sensors that are distributed in a three-dimensional environment to monitor a region of interest for possible intrusion by a target. The sensors collect

---

Zhu, Ding, Iyengar are with the Dept of Computer Science, Louisiana State University, Baton Rouge, LA 70803. Email: {mzhu, iyengar,sding}@bit.csc.lsu.edu. Brooks is with the Dept of Computer Engineering, Clemson University at Clemson, SC 61801, Email: rrb@acm.org. Wu, Rao are with Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, Email: {wuqn, raons}@ornl.gov.
This research is sponsored by the SensorNet Program funded by Office of Naval Research at Oak Ridge National Laboratory.

environmental measurements that are subject to independent additive noise. Each sensor node employs a threshold rule on the measurements to detect a target in the presence of random background noise. The threshold value employed by sensor nodes and their distances to the target determine both their hit and false alarm probabilities or rates using a signal attenuation model. When a target enters the monitoring region, it could be detected by multiple sensor nodes depending on their distances from the target. Typically, nearby sensors produce larger measurements while distant sensors receive less quantity of the measured signal. Our objective is to combine the decisions made by individual sensor nodes to achieve system detection performance beyond a weighted average of individual sensor nodes.

At the core, this problem is a specific instance of the classical distributed detection problem that has been studied extensively over past several decades [Varshney 1997]. The problem of fusing binary decisions taken at various sensor nodes has been solved using various schemes such as logical AND, OR, voting, Neyman-Pearson, and Bayes rule [Tenney and Sandell 1981; Sadjadi 1986; Thomopoulos et al. 1987; Reibman and Nolte 1987a; 1987b]. Logical AND, OR, and voting techniques are examples of simple fusion schemes that do not require sensor probability distribution functions. Voting schemes can be further classified into threshold schemes and plurality schemes. Unanimity voting, majority voting, and *m*-out-of-*n* voting are threshold schemes, which try to minimize the probability of wrong decisions. Plurality voting designates the hypothesis that receives the most votes to be the best possible decision. Although it involves only inexpensive computations and provides some degree of fault tolerance, a simple voting scheme generally lacks necessary performance guarantees in terms of error rates.

Chair and Varshney [Chair and Varshney. 1986] proposed an optimal fusion rule based on Bayes rule, requiring *a priori* probabilities. However, in practice, these *a priori* probabilities are often unavailable. Niu *et. al.* [Niu et al. 2004] presented a fusion method that chooses an optimal sensor threshold to achieve the maximal system hit rate given a certain system false alarm rate. This method demands a very large number of sensor nodes to apply the Central Limit Theorem. Unfortunately, this criterion cannot be always satisfied in real scenarios. Neyman-Pearson decision rule provides an optimal solution minimizing Type II error when subject to a chosen upper bound on Type I error probability or vice versa based on the likelihood ratio test. Although no *a priori* probability is required, Neyman-Pearson fusion rule requires sensor readings and continuous sensor probability distribution functions to apply the notion of confidence level, and these information may incur high computational cost or even not be available.

Other researchers focused on correlated sensor data fusion. Chen and Ansari [Chen and Ansari 1998] developed an adaptive fusion algorithm to estimate *a priori* and conditional probability through reinforcement learning. Although their method does not rely on *a priori* probability, it requires certain number of iterations for system convergence. In addition, the weight rate is updated using an approximation method that ignores certain variables dependence in partial derivative calculations. Rao [Rao 1996] developed fusion methods that do not require the knowledge of sensor probability distribution functions but need to be trained using measurements. Such training measurements may not be available in some cases, and in others too many measurements may be needed to ensure reasonable levels of performance guarantees.

We propose a centralized threshold-OR fusion rule for combining the individual sensor

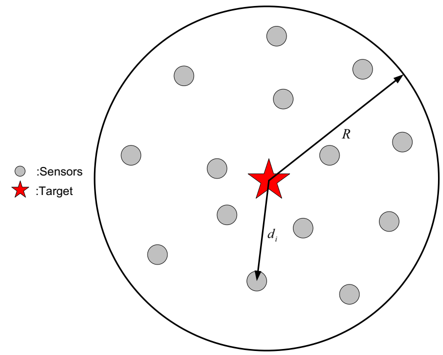

Fig. 1. Sensor deployment for target detection.

node decisions in the sensor network described above. We derive threshold bounds for accumulated decisions using Chebyshev's inequality based on individual hit and false alarm probabilities but without requiring an *a priori* knowledge of the underlying probability densities. We analytically show that the fused method achieves a higher hit rate and lower false alarm rate compared to the weighted averages of individual sensor nodes. Restraining conditions are also derived. Simulation results using Monte Carlo method show that the error probabilities in the fused system are significantly reduced to near zero. Our method lies between the voting method and Neyman-Pearson based fusion method in terms of the amount of information required for computation and the detection performance: it requires more information than the voting method, but less information than Neyman-Pearson; meanwhile, it achieves better fusion performance than weighted averages of individual sensors. Note that the voting method provides no performance guarantee and Neyman-Pearson ensures optimality in terms of Type I and Type II errors.

The rest of the paper is organized as follows: an analytical model of the proposed sensor network is presented in Section 2. In Section 3, we present a technical solution that derives the proper system threshold bounds for hard fusion systems. Simulation results are given in Section 4. We finally conclude our work in Section 5.

2. PROBLEM FORMULATION

We consider N sensor nodes deployed in a three-dimensional region of interest (ROI), centering around a target within radius R, as shown in Fig. 1. At sensor i, the noise n_i in a sensor measurement is independently and identically distributed (iid) according to the normal distribution.

$$n_i \sim \aleph(0,1). \tag{1}$$

The sensor measurements are subjective to an additive term due to such noise. Each sensor i makes a binary local decision as:

$$\begin{array}{ll} H_1: & r_i = w_i + n_i \\ H_0: & r_i = n_i \end{array} \tag{2}$$

where r_i is the actual sensor reading, w_i is the ideal sensor measurement, and n_i is the noise term.

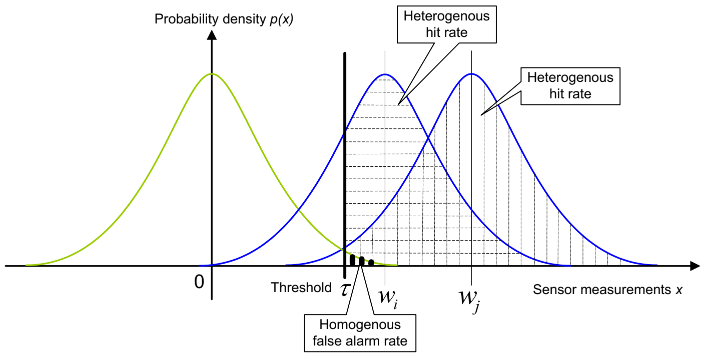

Fig. 2. Normal distribution based hit rate and false alarm rate calculation for heterogenous system

We consider an isotropic signal attenuation power model defined by:

$$w_i = \frac{w_0}{\sqrt{1+\beta d_i^n}}, \tag{3}$$

where w_0 is the original signal power emitted from the target located at point (x_0, y_0, z_0), β is a system constant, and d_i represents the Cartesian distance between the target and the sensor node, which is defined in Eq. 4. Parameter n is the signal attenuation exponent typically ranging from 2 to 3, and the distance is defined as:

$$d_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2}. \tag{4}$$

This model describes a three-dimensional unobstructed region monitored by a set of sensors that detect the signal emitted from a target within the monitoring area.

Suppose that every sensor node employs the same threshold τ for decision making regardless of its distance to the target. The expected signal strength w_i for sensor i can be computed from Eq. 3 according to its distance to the target. Thus, the hit rate p_{h_i} and false alarm rate p_{f_i} for sensor i can be derived as follows [Niu et al. 2004]:

$$p_{h_i} = \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-(x-w_i)^2}{2}} dx, \tag{5}$$

$$p_{f_i} = \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx. \tag{6}$$

Fig. 2 illustrates the use of a normal distribution function for computing individual sensor hit rate and false alarm rate in a heterogenous system. A simplified homogeneous system ignores the impact of various distances to the target on sensor detection capabilities so that every sensor has the same hit rate and false alarm rate. The heterogenous sensor network system considers different hit rates and the same false alarm rates as computed by Eq. 5 and Eq. 6, respectively.

3. THRESHOLD-OR FUSION METHOD

Each sensor i makes an independent binary decision S_i as either 0 or 1. The fusion center collects local decisions and computes S as:

$$S = \sum_{i=1}^{N} S_i, \tag{7}$$

which is then compared with a threshold T to make a final decision. Under the assumption that sensor measurements are statistically independent under H_1, the mean and variance of S are given as follows when a target is present:

$$\begin{aligned} E(S|H_1) &= \sum_{i=1}^{N} p_{h_i}, \\ Var(S|H_1) &= \sum_{i=1}^{N} p_{h_i}(1-p_{h_i}). \end{aligned} \tag{8}$$

Similarly, under the assumption that sensor measurements are statistically independent under H_0, the mean and variance of S when a target is absent are defined as:

$$\begin{aligned} E(S|H_0) &= \sum_{i=1}^{N} p_{f_i}, \\ Var(S|H_0) &= \sum_{i=1}^{N} p_{f_i}(1-p_{f_i}). \end{aligned} \tag{9}$$

The threshold value T is critical to the system performance. Let P_h and P_f denote the hit rate and false alarm rate of the fused system. If we let $P_h > 0.5$ and $P_f < 0.5$, T should be bounded by $\sum_{i=1}^{N} p_{f_i} < T < \sum_{i=1}^{N} p_{h_i}$.

The weighted averages of p_{h_i} and p_{f_i}, $i = 1, 2, \ldots, N$ are defined as follows:

$$\sum_{i=1}^{N} \frac{p_{h_i}}{\sum_{j=1}^{N} p_{h_j}} p_{h_i} = \frac{\sum_{i=1}^{N} p_{h_i}^2}{\sum_{i=1}^{N} p_{h_i}}, \tag{10}$$

$$\sum_{i=1}^{N} \frac{1-p_{f_i}}{\sum_{j=1}^{N} (1-p_{f_j})} p_{f_i} = \frac{\sum_{i=1}^{N} (1-p_{f_i}) p_{f_i}}{\sum_{i=1}^{N} (1-p_{f_i})}. \tag{11}$$

We desire better detection performance of the fused system than the corresponding weighted averages in terms of higher hit rate and lower false alarm rate such that:

$$P_h > \frac{\sum_{i=1}^{N} p_{h_i}^2}{\sum_{i=1}^{N} p_{h_i}}, \tag{12}$$

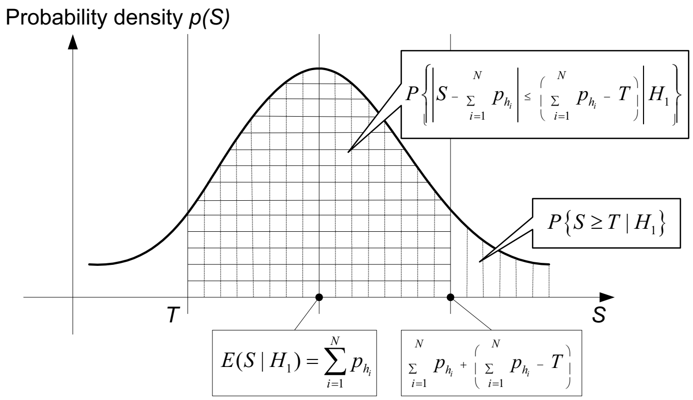

Fig. 3. Application of Chebyshev's inequality in calculating the lower bound of the system hit rate.

$$P_f < \frac{\sum_{i=1}^{N}(1-p_{f_i})p_{f_i}}{\sum_{i=1}^{N}(1-p_{f_i})}. \tag{13}$$

We first consider a lower bound of the hit rate of the fused system :

$$\begin{aligned} P_h &= P\{S \geq \mathrm{T}|\mathrm{H}_1\} \\ &\geq P\{|S - \sum_{\mathrm{i=1}}^{\mathrm{N}} p_{h_i}| \leq (\sum_{1}^{\mathrm{N}} p_{h_i} - \mathrm{T})|\mathrm{H}_1\} \\ &\geq 1 - \frac{\sigma^2}{\mathrm{k}^2} \\ &= 1 - \frac{\sum_{i=1}^{N} p_{h_i}(1-p_{h_i})}{(\sum_{i=1}^{N} p_{h_i} - T)^2} \end{aligned} \tag{14}$$

where we applied Chebyshev's inequality in the third step as shown in Fig. 3 and k equals $(\sum_{1}^{\mathrm{N}} p_{h_i} - \mathrm{T})$. Now the condition Eq. 12 can be ensured by the following sufficient condition.

$$1 - \frac{\sum_{i=1}^{N} p_{h_i}(1-p_{h_i})}{(\sum_{i=1}^{N} p_{h_i} - T)^2} \geq \frac{\sum_{i=1}^{N} p_{h_i}^2}{\sum_{i=1}^{N} p_{h_i}}. \tag{15}$$

Following that, an upper bound of T can be derived from Eq. 15 as follows:

$$T \leq \sum_{i=1}^{N} p_{h_i} - \sqrt{\sum_{i=1}^{N} p_{h_i}}. \tag{16}$$

Similarly, for the false alarm rate, we carry out a similar procedure to compute the lower bound from Eq. 17 to Eq. 21. Again, Chebyshev inequality is applied in the second step in Eq. 19.

$$P_f = P\{S \geq T|\mathrm{H}_0\} = 1 - P\{S < T|\mathrm{H}_0\}, \tag{17}$$

$$P\{S < T|\mathrm{H}_0\} \geq P\{|S - \sum_{i=1}^{N} p_{f_i}| \leq (T - \sum_{i=1}^{N} p_{f_i})|\mathrm{H}_0\}, \tag{18}$$

$$P_{\mathrm{f}} \leq 1 - P\{|S - \sum_{i=1}^{N} p_{f_i}| \leq (T - \sum_{i=1}^{N} p_{f_i})|\mathrm{H}_0\} \leq \frac{\sum_{i=1}^{N} p_{f_i}(1-p_{f_i})}{(T - \sum_{i=1}^{N} p_{f_i})^2}, \tag{19}$$

Now we consider the condition that ensure false probability of fuser is smaller than that of weighted average given by:

$$\frac{\sum_{i=1}^{N} p_{f_i}(1-p_{f_i})}{(T - \sum_{i=1}^{N} p_{f_i})^2} \leq \frac{\sum_{i=1}^{N} (1-p_{f_i})p_{f_i}}{\sum_{i=1}^{N} (1-p_{f_i})}, \tag{20}$$

$$T \geq \sum_{i=1}^{N} p_{f_i} + \sqrt{\sum_{i=1}^{N} (1-p_{f_i})}. \tag{21}$$

Therefore, we define the range of T using the upper bound in Eq. 16 and lower bound in Eq. 21 as follows:

$$[\sum_{i=1}^{N} p_{f_i} + \sqrt{\sum_{i=1}^{N} (1-p_{f_i})}, \sum_{i=1}^{N} p_{h_i} - \sqrt{\sum_{i=1}^{N} p_{h_i}}]. \tag{22}$$

For the case of a homogeneous system, the bounds of threshold T are simplified as follows:

$$\left[Np_f + \sqrt{N(1-p_f)},\ Np_h - \sqrt{Np_h}\right]. \tag{23}$$

To ensure that the upper bound is larger than the lower bound, we have the following restrictions on individual hit rates, individual false alarm rates, and the number of sensor nodes, for the heterogeneous and homogeneous systems, respectively :

$$\sum_{i=1}^{N} p_{f_i} + \sqrt{\sum_{i=1}^{N} (1-p_{f_i})} - \sum_{i=1}^{N} p_{h_i} + \sqrt{\sum_{i=1}^{N} p_{h_i}} \quad \leq 0, \tag{24}$$

$$p_h - p_f \geq \frac{\sqrt{p_h} + \sqrt{1-p_f}}{\sqrt{N}}. \tag{25}$$

Note that for the homogeneous system, the number of sensor nodes N can be set large enough to satisfy Eq. 25.

4. SIMULATION RESULTS

Our simulation based on 100,000 runs[1] produced the receiver operative characteristic (ROC) curve, a plot of the system hit rate against the false alarm rate for different possible thresholds. There is a tradeoff between sensitivity and specificity, namely any increase in sensitivity of hit rate will be accompanied by an increase in non-specificity of false alarm rate.

---

[1]To simplify calculation, we restrict the sensor deployment in a 2-D plane by setting coordinate *z* to be zero. The simulation results for a 3-D region is qualitatively similar to the 2-D case.

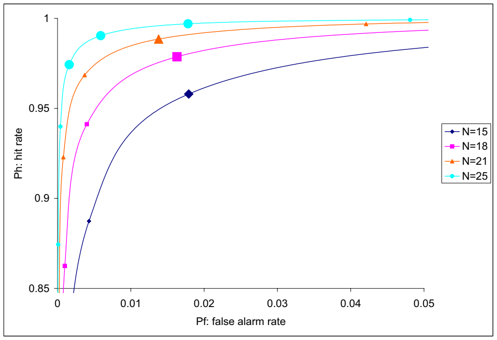

Fig. 4. Homogeneous system ROC curve with different sensor node number

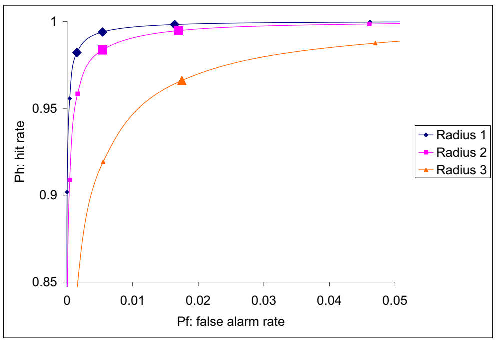

Fig. 5. Heterogenous system ROC curve with different deployment radius

The curvature of the ROC curve determines the detection accuracy of the system, namely, the closer does the curve follow the left-hand border and then the top border of the ROC space, the more accurate is the system. The closer does the curve approach the 45-degree diagonal of the ROC space, the less accurate is the test.

For the homogeneous case, we set sensor hit rate to be 0.65, and sensor false alarm rate to be 0.2. In Fig. 4, four system ROC curves with sensor node number going from 15 to 25 are plotted. The desirable segment on the ROC curve is located by restricting thresholds selection. Due to the discrete nature of our binary decision system, we identify the selected ROC segment as individual enlarged markers. From Fig. 4, we observe that all selected points fall at the top left corner of the ROC space, and bear hit rate and false alarm rate that is by far superior to that of a single sensor node. When the number of sensor nodes increases, our system performances are further enhanced due to ample available resources.

Table I. Numeric results with different deployment radius for heterogenous system.

$N=25$, $W_{p_f}=0.2$	T_l	T_u	W_{p_h}	P_{h_l}	P_{h_u}	P_{f_l}	P_{f_u}
R=1	10	12	.6788	.9821	.9982	.0015	.0164
R=2	10	11	.628	.9836	.9947	.0054	.017
R=3	10	10	.5612	.966	.966	.0175	.0175

Table II. Comparison of different decision rules based fusion.

Rules	Fusion type	Hypothesis	pdf	*a priori* probability	Performance
Voting	Hard	Multiple	No	No	No guarantee
Bayes Criterion	Soft/Hard	Multiple	Yes	Yes	Minimize cost function
MAP	Soft/Hard	Multiple	Yes	Yes	Minimize error probability
Neyman-Pearson	Soft	Binary	Yes	No	Fix one error prob. minimize the other
Proposed method	Hard	Binary	Yes	No	Significantly reduce error prob.

For the heterogeneous case, from Eq. 5 and Eq. 6, we conclude that sensors have the same false alarm rate and different hit rates due to various distances to the target. In our simulation program, the given hit rate and false alarm rate for sensors in close proximity to the target are represented as p_h and p_f, respectively. Sensor threshold τ can be calculated from Eq. 6 given a known false alarm rate p_f. Consequently, original signal power w_0 can be computed from Eq. 5. w_i can be derived from Eq. 3 and therefore, p_{h_i} can be computed according to Eq. 5. Our approach is also capable of handling systems with heterogenous false alarm rates if necessary.

We set the total number of sensor nodes to be 25. The p_h is set to be 0.75 and p_f is set to be 0.2 in simulation tests. Sensor deployment radius ranges from 1 to 3 to compare different deployment strategies. Results are tabulated in Table 1 as system threshold bounds(T_l,T_u), weighted average hit rate and false alarm rate as W_{p_h}, and W_{p_f} respectively, system hit rate bounds(P_{h_l},P_{h_u}), and system false alarm rate bounds(P_{f_l},P_{f_u}) under different radius R. Dispersed sensor deployment negatively affects the system performance as a result of depressed detection performance and truncated scope of threshold bounds. It further validates the rule that sensors should be deployed as close to potential targets as possible. From Fig. 5 with radius 3, our fusion system achieves excellent hit rate close to 0.97 and low false alarm rate below 0.02 as compared to weighted average of 0.5612 and 0.2, respectively, before fusion.

5. CONCLUSIONS AND FUTURE WORK

We proposed a threshold-OR fusion method for a sensor network wherein each node employs a threshold to the measurement to decide the presence of a target in the field of view. Current non-model and model-based fusion methodologies are derived from some variants of decision rules such as Voting, Bayes Criterion, Maximum a Posterior Criterion(MAP), and Neyman-Pearson. A comparison of fusion types in terms of input sensor data, hypothesis, sensor probability density functions(pdfs), *a priori* probability, and performance is provided in Table 2. Despite the simplicity, non-model based voting rule gives no system performance guarantee in terms of probability of errors. Bayes criterion minimizes the cost function based on the knowledge on *a priori* probability which is not always available. MAP, a similar decision rule to Bayes Criterion, simply minimizes the overall probability of errors since it seeks to maximize the posterior probability. However, *a priori* probability is still desired. Neyman-Pearson(NP) is attractive since it does not require knowledge of priors and a cost function as opposed to Bayes criterion [Varshney 1997]. Nevertheless, NP based fusion rule is essentially a type of soft fusion, which demands continuous local sensor readings and pdfs to control one error probability under certain confidence level.

Our method falls into the category of hard fusion accepting discrete sensor decisions, and does not require the knowledge of the *a priori* probability. Since our conditions are sufficient and not necessary in deriving the threshold bounds, simulation results under such restrictive conditions yield significantly reduced probability of errors compared to that of single sensor or weighted average with both Type I and Type II errors close to zero.

The given threshold bounds allow users certain freedom in shifting between sensitivity and specificity instead of a single cutting point. For example, if Type I error is not tolerable, user can choose a threshold close to the upper bound. If Type II error needs to be minimized, user may tend to pick a threshold close to the lower bound. In addition, our approach has a low computational cost making practical deployment feasible with limited resource. For future work, we plan to apply weighted factors to local decisions based on various distances and signal noise ratios (SNR) when calculating binomial distribution in Eq. 7. Deployment of the proposed fusion rule in practical sensor network application is also of our future interest.

REFERENCES

CHAIR, Z. AND VARSHNEY., P. 1986. Optimal data fusion in multiple sensor detection systems. *IEEE Trans. Aerosp. Electron. Syst AES,* 22 (Jan), 98–101.

CHEN, J. G. AND ANSARI, N. 1998. Adaptive fusion of correlated local decisions. *IEEE Trans. Systems. Man. Cybernetics 28,* 2 (May), 276–281.

NIU, R., VARSHNEY, P., MOORE, M., AND KLAMER, D. 2004. Decision fusion in a wireless sensor network with a large number of sensors. In The 7th International Conference on Information Fusion. *Proc. The 7th International Conference on Information Fusion*, 21–27.

RAO, N. S. V. 1996. Distributed decision fusion using empirical estimation. *IEEE Transactions on Aerospace and Electronic Systems 33,* 4, 1106–1114.

REIBMAN, A. R. AND NOLTE, L. W. 1987a. Design and performance comparison of distributed detection networks. *IEEE Trans. Aerosp. Electron. Syst. AES,* 23 (Nov), 789–797.

REIBMAN, A. R. AND NOLTE, L. W. 1987b. Optimal detection and performance of distributed sensor systems. *IEEE Trans. Aerosp. Electron. Syst AES,* 23 (Jan), 24–30.

SADJADI, F. A. 1986. Hypothesis testing in a distributed environment. *IEEE Trans. Aerosp. Electron. Syst. AES,* 22 (March), 134–137.

TENNEY, R. R. AND SANDELL, N. R. 1981. Detection with distributed sensors. *IEEE Trans. Aerosp. Electron. Syst. AES,* 17 (July), 501–510.

THOMOPOULOS, F. A., VISWANATHAN, R., AND BOUGOOULIAS., D. C. 1987. Optimal decision fsion in multiple sensor systems. *IEEE Trans. Aerosp. Electron. Syst. AES,* 23 (Sept), 644–653.

VARSHNEY, P. 1997. *Distributed detection and data fusion*. Springer-Verlag.