Fusion of Threshold Rules for Target Detection in Wireless Sensor Networks

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

We propose a binary decision fusion rule that reaches a global decision on the presence of a target by integrating local decisions made by multiple sensors. Without requiring *a priori* probability of target presence, the fusion threshold bounds derived using Chebyshev's inequality ensure a higher hit rate and lower false alarm rate compared to the weighted averages of individual sensors. The Monte Carlo-based simulation results show that the proposed approach significantly improves target detection performance, and can also be used to guide the actual threshold selection in practical sensor network implementation under certain error rate constraints.

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, Chebyshev inequality, ROC curve

1. INTRODUCTION

Many existing non-model based or 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 [8; 7; 9; 5; 6; 11; 10; 4; 1; 3]. Data fusion is in general categorized as low-, intermediate-, or high- level fusion, depending on the stage where actual fusion processing takes place. In this technical note, we present a model based high-level hard fusion scheme, also known as decision fusion, where a final global decision is reached by integrating local binary decisions made by multiple sensor nodes that detect the same target from different distances. Without requiring *a priori* knowledge on the probability of target presence, this centralized fusion scheme uses Chebyshev's inequality to derive threshold bounds that ensure a better system performance compared with the weighted averages of all individual sensors. Simulation results based on Monte Carlo method show that the error probabilities in the fused system are significantly reduced to near zero. Furthermore, the simulation or experiment results are particularly useful in guiding practical implementation in which an upper bound on the false alarm rate and

Zhu is with Dept of Computer Science, Southern Illinois University, Carbondale, IL 62901. Email: mzhu@cs.siu.edu. Ding, Iyengar are with the Dept of Computer Science, Louisiana State University, Baton Rouge, LA 70803. Email: {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 is with the Dept of Computer Science, University of Memphis, Memphis, TN 38152. Email: qishiwu@memphis.edu. Rao is with Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, Email: raons@ornl.gov.
* Corresponding author. Tel: 618-453-6057

minimization of missing rate or vice versa are desired at the same time. The rest of the paper is organized as follows: Mathematical model is presented in Section 2. Section 3 discusses the technical approach to derive the proper system threshold bounds. Simulation results are given in Section 4. We conclude our work in Section 5.

2. PROBLEM FORMULATION

We consider N sensor nodes randomly deployed in a region of interest(ROI) with radius R. Noise at each local sensor follows the standard Gaussian distribution defined as: $n_i \sim \aleph(0,1)$. Each sensor makes a binary local decision on the target presence as:

$$H_1: \quad r_i = s_i + n_i \quad H_0: \quad r_i = n_i \tag{1}$$

where r_i is the total sensor reading of sensor i and n_i denotes the noise level observed by sensor i. The signal strength s_i decays as sensor moves away from the target and follows the isotropic attenuation power model as defined in Eq.2:

$$s_i = \frac{S_0}{\sqrt{1+\beta d_i^m}} \tag{2}$$

where S_0 is the original signal power from the target, β is a constant, and d_i represents the Euclidean distance between the target and sensor i. The signal attenuation exponent m ranges from 2 to 3. Here we assume the same threshold τ for every sensor node because of the fact that some simple sensor motes may not have the intelligence and necessary processing resources to adjust their thresholds dynamically. Thus, the hit rate p_{h_i} and false alarm rate p_{f_i} for sensor i can be defined as:

$$p_{h_i} = \int_\tau^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-(x-s_i)^2}{2}} dx, \quad p_{f_i} = \int_\tau^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx. \tag{3}$$

Please note that our fusion method can be applied to any signal attenuation model. A simplified model is used in Eq. 2 and Eq. 3 simply for discussion and simulation purposes. We assume that factory manufactured sensors are calibrated with specified ROC curves fitted for the characteristics of desired target under certain environment in real applications.

3. THRESHOLD FUSION METHOD

Sensor i makes an independent binary decision S_i as either 0 or 1. The fusion center uses a simple $0/1$ counting rule for convenience and collects local decisions and computes S as: $S = \sum_{i=1}^{N} S_i$, which is then compared with a system threshold T to make a final decision. [1]
For simplicity, we neglect covariance and assume that sensor measurements are conditionally independent under H_1. The mean and variance of S are given below when a target is present:

$$E(S|H_1) = \sum_{i=1}^{N} p_{h_i}, \quad Var(S|H_1) = \sum_{i=1}^{N} p_{h_i}(1-p_{h_i}). \tag{4}$$

Similarly, the mean and variance of S when a target is absent are defined as:

$$E(S|H_0) = \sum_{i=1}^{N} p_{f_i}, \quad Var(S|H_0) = \sum_{i=1}^{N} p_{f_i}(1-p_{f_i}). \tag{5}$$

[1]A simple suboptimal decision metric is used here without using weights to differentiate individual decisions S_i.

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 respectively in Eq. 6. We also made reasonable value bounds for T as $\sum_{\mathrm{i=1}}^{N} p_{f_i} < T < \sum_{\mathrm{i=1}}^{N} p_{h_i}$.

$$P_h = P\{S \geq \mathrm{T}|\mathrm{H}_1\}, \quad P_f = P\{S \geq \mathrm{T}|\mathrm{H}_0\} = 1 - P\{S < \mathrm{T}|\mathrm{H}_0\}, \tag{6}$$

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

$$\sum_{i=1}^{N} \frac{p_{h_i}}{\sum_{j=1}^{N} p_{h_j}} p_{h_i} = \frac{\sum_{i=1}^{N} p_{h_i}^2}{\sum_{i=1}^{N} p_{h_i}}, \tag{7}$$

$$\sum_{i=1}^{N} \frac{1 - p_{f_i}}{\sum_{j=1}^{N} (1 - p_{f_j})} p_{f_i} = \frac{\sum_{i=1}^{N} (1 - p_{f_i}) p_{f_i}}{\sum_{i=1}^{N} (1 - p_{f_i})}. \tag{8}$$

We wish to achieve better system detection performance than the corresponding weighted averages in terms of higher hit rate and lower false alarm rate. Prior work on distance-based weighted average has been proposed in [2]. Thus, the following inequalities should hold:

$$P_h > \frac{\sum_{i=1}^{N} p_{h_i}^2}{\sum_{i=1}^{N} p_{h_i}}, \quad P_f < \frac{\sum_{i=1}^{N} (1-p_{f_i}) p_{f_i}}{\sum_{i=1}^{N} (1-p_{f_i})}. \tag{9}$$

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

$$P_h \geq P\{|S - \sum_{\mathrm{i=1}}^{N} p_{h_i}| \leq (\sum_{\mathrm{i=1}}^{N} p_{h_i} - \mathrm{T})|\mathrm{H}_1\} \geq 1 - \frac{\sigma^2}{\mathrm{k}^2} = 1 - \frac{\sum_{i=1}^{N} p_{h_i}(1-p_{h_i})}{(\sum_{i=1}^{N} p_{h_i} - T)^2} \tag{10}$$

where we apply Chebyshev's inequality in the second step and denote $(\sum_{1}^{\mathrm{N}} p_{h_i} - \mathrm{T})$ by k. Now the inequality of P_h in Eq. 9 can be ensured by the following sufficient condition:

$$1 - \frac{\sum_{i=1}^{N} p_{h_i}(1 - p_{h_i})}{(\sum_{i=1}^{N} p_{h_i} - T)^2} \geq \frac{\sum_{i=1}^{N} p_{h_i}^2}{\sum_{i=1}^{N} p_{h_i}}. \tag{11}$$

Following that, an upper bound on T can be derived from Eq. 11 as follows:

$$T \leq \sum_{i=1}^{N} p_{h_i} - \sqrt{\sum_{i=1}^{N} p_{h_i}}. \tag{12}$$

For the false alarm rate, we follow a similar procedure from Eq. 13 to Eq. 16 to compute the lower bound. Note that Chebyshev's inequality is applied in the second step in Eq. 14.

$$P\{S < T|\mathrm{H}_0\} \geq P\{|S - \sum_{i=1}^{N} p_{f_i}| \leq (T - \sum_{i=1}^{N} p_{f_i})|\mathrm{H}_0\}, \tag{13}$$

Table I. Fusion system performance with different deployment radius.

$N=25$	$W_{p_{h/f}}$	$P^{8}_{h/f}$	$P^{9}_{h/f}$	$P^{10}_{h/f}$	$P^{11}_{h/f}$	$P^{12}_{h/f}$	$P^{13}_{h/f}$	$P^{14}_{h/f}$	$P^{15}_{h/f}$
R=1	.64/.2	.99/.11	.99/.05	**.99/.02**	**.98/.01**	.96/.0	.9/.0	.81/.0	.68/.0
R=3	.51/.2	.96/.11	**.90/.05**	.81/.02	.68/.01	.52/.0	.35/.0	.21/.0	.11/.0
R=5	.49/.2	**.89/.11**	**.79/.05**	.65/.02	.47/.01	.31/.0	.18/.0	.09/.0	.04/.0

$$P_f \leq 1 - P\{|S - \sum_{i=1}^{N} p_{f_i}| \leq (T - \sum_{i=1}^{N} p_{f_i})|\mathrm{H}_0\} \leq \frac{\sum_{i=1}^{N} p_{f_i}(1-p_{f_i})}{(T - \sum_{i=1}^{N} p_{f_i})^2}. \quad (14)$$

Now we consider the sufficient condition that ensures the system false alarm rate to be smaller than that of weighted average:

$$\frac{\sum_{i=1}^{N} p_{f_i}(1-p_{f_i})}{(T - \sum_{i=1}^{N} p_{f_i})^2} \leq \frac{\sum_{i=1}^{N} (1-p_{f_i})p_{f_i}}{\sum_{i=1}^{N} (1-p_{f_i})}, \quad (15)$$

$$T \geq \sum_{\mathrm{i}=1}^{N} p_{f_i} + \sqrt{\sum_{\mathrm{i}=1}^{N} (1-p_{f_i})}. \quad (16)$$

Finally, we define the range of T using the upper bound in Eq. 12 and lower bound in Eq. 16 as follows:

$$[\sum_{\mathrm{i}=1}^{N} p_{f_i} + \sqrt{\sum_{\mathrm{i}=1}^{N} (1-p_{f_i})}, \sum_{i=1}^{N} p_{h_i} - \sqrt{\sum_{i=1}^{N} p_{h_i}}]. \quad (17)$$

In order to ensure that the upper bound is larger than the lower bound, we have the following restriction on individual hit rates, individual false alarm rates, and the number of sensor nodes:

$$\sum_{\mathrm{i}=1}^{N} p_{f_i} + \sqrt{\sum_{\mathrm{i}=1}^{N} (1-p_{f_i})} - \sum_{i=1}^{N} p_{h_i} + \sqrt{\sum_{i=1}^{N} p_{h_i}} \quad \leq 0. \quad (18)$$

4. SIMULATION RESULTS

Our simulation based on one million runs produces the system's receiver operative characteristic (ROC) curve, a plot of the system hit rate against the system false alarm rate under all possible thresholds. The curvature of the ROC curve indicates the detection accuracy of the system: the closer does the curve follow the left-top border of the ROC space, the higher detection accuracy does the system achieve. In the simulation, we randomly deployed 25 sensor nodes in ROI. The original individual sensor hit rate and false alarm rate (sensor in the immediate vicinity of the target) are found to be 0.75 and 0.2, respectively. [2] The deployment radius ranges from 1 to 5 for comparing different deployment strategies.

[2]manufactured sensor motes have multiple ROC curves under different signal power levels. With a specified threshold and detected power level, false alarm and hit rates can be found out from a look-up table.

The individual sensor's hit rate and false alarm rate will be determined by their distance to the target. Performance comparisons in terms of fusion system hit rate and false alarm rate under different thresholds and deployment radius are tabulated in Table I. The weighted average hit rate and false alarm rate are denoted as $W_{p_{h/f}}$ for each deployment radius. The system hit rate and false alarm rate with threshold k are represented by $P^k_{h/f}$. From the table, we observe that the legitimate threshold bounds are from 8 to 15 for radius 1, from 8 to 12 for radius 3, and from 8 to 10 for radius 5, since all thresholds from those ranges produce a better system performance than the weighted averages $W_{p_{h/f}}$. However, the calculated threshold bounds are indicated in bold as a subset of the legitimate bounds because our fusion method imposes more stringent requirements (due to sufficient and not necessary conditions posed by our inequalities) with a higher expectation for detection performance. On the other hand, if a specific system false alarm rate, e.g. 0.05, needs to be satisfied under a deployment radius of 5, the best threshold must be set to 9 in order to achieve the best possible system hit rate of 0.79, according to Table I.

5. CONCLUSIONS

We proposed a threshold fusion method for sensor networks wherein each node decides the presence of a target and sends its binary decision to the fusion center for final decision making. Our method is a centralized hard fusion scheme accepting discrete sensor decisions without requiring *a priori* probability of target presence. In addition to achieving better system performance than corresponding weighted averages, the determined threshold bounds allow users certain freedom in fine tuning between sensitivity and specificity. The ROC curve acquired from off-line simulation or experiments can be used to maximize the system hit rate under the constraint of a given system false alarm rate or vice versa, and hence to guide the practical implementation of sensor networks. Our approach also has a low computational cost thus which makes practical deployment feasible with limited computing resource.

REFERENCES

D. Chen and P.K. Varshney. A survey of void handling techniques for geographic routing in wireless networks. *IEEE Communications Surveys and Tutorials*, 9:50–67, 2007.

M. Duarte and Y. H. Hu. Distance-based decision fusion in a distributed wirelesssensor networks. *Telecommunication Systems*, 26(2-4):339–350, 2004.

N. Katenka, E. Levina, and G. Michailidis. Local vote decision fusion for target detection in wireless sensor networks. *IEEE Transactions on Signla Processing*, 56(1):329–338, Jan 2008.

R. Niu, P. K. Varshney, and Q. Cheng. Distributed detection in a wireless sensor network with a large number of sensors. *Information Fusion*, 7:380–394, Dec 2006.

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

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

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

R. R. Tenney and N. R. Sandell. Detection with distributed sensors. *IEEE Trans. Aerosp. Electron. Syst.*, AES(17):501–510, July 1981.

F. A. Thomopoulos, R. Viswanathan, and D. C. Bougoulias. Optimal decision fsion in multiple sensor systems. *IEEE Trans. Aerosp. Electron. Syst.*, AES(23):644–653, Sept 1987.

J. N. Tsitsiklis. Decentralized detection. *Advances in Statistical Signal Processing*, 2:297–343, 1993.

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