Consensus-based Distributed Detection with Mitigating Outliers for Wireless Sensor Networks

Jinho Choi #1, Hyukjin Lee *2, Cheng-Chew Lim *3

College of Engineering, Swansea University
Singleton Park, Swansea, SA2 8PP, UK
[1] j.choi@swansea.ac.uk

* *School of Electrical and Electronic Engineering, The University of Adelaide*
Adelaide, Australia 5005
[2] hyukjin@eleceng.adelaide.edu.au
[3] cclim@eleceng.adelaide.edu.au

***Abstract*—In wireless sensor networks, intruders can manipulate some sensors' observations locally, which results in outliers in distributed detection. These outliers can be detected and removed by a fusion centre as all the sensors' observations are available. For wireless sensor networks without a fusion centre, however, the detection performance can be significantly degraded as distributed consensus algorithms are vulnerable to outliers. In this paper, we consider the outlier detection for wireless sensor networks without a fusion centre when a distributed consensus algorithm is employed for distributed detection.**[1]

I. Introduction

Wireless sensor networks (WSNs) are becoming more popular as they can be employed for various civil and military applications including environmental and industrial monitoring and surveillance. In general, a WSN consists of a number of sensor nodes that can perform sensing and wireless communications to send their measurements or data to a fusion centre (FC). Because of the limited capability of sensor nodes that are usually small and have limited power sources (e.g., batteries), the operations at sensors nodes have to be simple and efficient. For example, each sensor can decide the presence of a certain target using hypothesis testing and transmit its decision to a FC so that the FC can combine the decisions from sensor nodes [1].

In performing distributed detection for WSNs, there are various problems. As wireless channels from sensors to a FC are not ideal due to fading and noise, the FC can receive the decisions of sensor nodes with errors. For decision fusion, these channel impairments can be taken into account [2]. Another problem is that the required bandwidth for transmitting sensors' decisions increases with the number of sensor nodes if orthogonal channels are used. To avoid this problem, all the sensors can transmit signals to a FC in a common channel using multiple access schemes [3], [4]. In this case, the FC has a superposition of the transmitted signals that is a decision statistic for the FC to make a global decision.

Since WSNs are prone to munipulation by adversaries, outlier detection is required [5], [6], [7], [8]. For example, some sensor nodes may have biased observations and transmit manipulated decisions to a FC. In [5], a kernel-based technique is used to detect outliers provided that each sensor can have a sufficient number of observations in the absence of a parametric model for outliers. In [6], for an event boundary detection, an outlier detection problem is formulated. Using spatial correlation, it is shown that the event boundary can be detected. In [7], statistical hypothesis testing is employed for outlier detection with Markov models that can capture spatial structures of WSNs. In [8], an overview of outlier detection techniques for WSNs is provided. As discussed in [8], there are various problems where outlier detection is required in WSNs. Furthermore, the formulation of outlier detection relies on a given application and problem.

If sensor nodes can have sufficient computing power, the WSN can have a distributed architecture for the distributed detection without a FC. Using a certain iterative algorithm, sensor nodes can exchange their information so that each sensor node can have a common global decision. This kind of iterative algorithm is called distributed consensus algorithm [9], [10] and its application to distributed detection in WSNs can be found in [11]. The main advantage of distributed consensus algorithms for distributed detection is that *i)* no central processing is required; *ii)* the global decision is available at every sensor node. There are also disadvantages. For example, in order to converge to a final result (or to get a consensus), sensor nodes should repeatedly exchange and update their information, which might result in an excessive burden of communication between sensor nodes. Furthermore, if some sensors' observations are outliers (due to imperfect sensing or manipulation by intruders), these distributed consensus algorithms could provide biased results.

In this paper, we focus on the outlier detection for WSNs when a distributed consensus algorithm is employed for distributed detection. Outliers are the observations at some sensors where the observations are inconsistent with the observations of the other sensor nodes. In the context of distributed detection, the outliers are the observations that are generated from a different distribution. For example, the observations of some sensor nodes can be totally different from those of

[1]This work has been supported by EPSRC-DSTL, Grant No. EP/H011919/1.

the others due to sensors' malfunction or manipulation by an intruder. When a distributed consensus algorithm employed, these outliers can result in a wrong decision as averaging is prone to outliers (while the median is robust against outliers [6]). As there is no FC, central processing for the outlier detection is not available. Since most outlier detection techniques (c.f., [8]) are based on central processing, we need to consider a distributed outlier detection method that can be performed at each sensor node when a distributed consensus algorithm is employed. To the best of our knowledge, the outlier detection in the context of distributed detection by a distributed consensus algorithm has not been studied yet. In this paper, assuming that statistical models for outliers are available, we study the outlier detection in conjunction with a distributed consensus algorithm for distributed detection.

II. Background

We discuss distributed detection using a distributed consensus algorithm in WSNs.

A. Distributed Detection

Suppose that a WSN consists of L sensors to detect a common target. We consider binary hypothesis testing, where H_0 and H_1 represent the absence and presence of the target, respectively. Let $\Pr(H_m)$ denote the a priori probability of H_m, $m = 0, 1$. Let x_l denote the log-likelihood ratio (LLR) from local observation(s) at sensor l. If each sensor has an observation that is conditionally independent, the sum of the LLRs at sensors is the global LLR that is obtained when all the observations are available. For convenience, the LLR at each sensor is called local LLR. For example, if sensor l has the following observation:

$$s_l = \begin{cases} \mu_0 + n_l, & \text{if } H_0 \text{ is true;} \\ \mu_1 + n_l, & \text{if } H_1 \text{ is true,} \end{cases} \tag{1}$$

where $\mu_1 > \mu_0$ and $n_l \sim \mathcal{N}(0, \sigma^2)$ is an independent noise, the local LLR at sensor l, denoted by x_l, is given by

$$\begin{aligned} \text{LLR}(s_l) = x_l &= \log \frac{f(s_l|H_1)}{f(s_l|H_0)} \\ &= \frac{\mu_1 - \mu_0}{\sigma^2} s_l - \frac{\mu_1^2 - \mu_0^2}{2\sigma^2}. \end{aligned} \tag{2}$$

The global LLR becomes

$$\begin{aligned} \text{LLR}(s_1, \ldots, s_L) &= \log \frac{f(s_1, \ldots, s_L|H_1)}{f(s_1, \ldots, s_L|H_0)} \\ &= \log \frac{\prod_{l=1}^L f(s_l|H_1)}{\prod_{l=1}^L f(s_l|H_0)} \\ &= \sum_{l=1}^L x_l. \end{aligned} \tag{3}$$

In the presence of a FC within the WSN, each sensor can transmit its local LLR to the FC so that it can compute the global LLR to make a final decision. This results in the optimal detection (e.g., the ML or MAP decision rule [12] can be carried out at the FC with the global LLR in (3)). Depending on communication limits imposed on the transmission from sensors to the FC, various suboptimal approaches can be employed [1]. For example, each sensor performs binary hypothesis testing with its observation only and sends a local binary decision to the FC for decision fusion.

B. Iterative Distributed Algorithm

In some WSNs, it may not be desirable to have a dedicated FC for central processing. Provided that each sensor has sufficient computing and communication power, it is possible that each sensor can make a final decision as a FC does through distributed consensus algorithms [9], [10]. Suppose that the network topology of a WSN is given by $G = (V, E)$, where $V = \{1, \ldots, L\}$ is the set of the sensor nodes and $E = \{(l, m)\}$ is the set of the edges. Here, an edge is the pair of two connected sensors that can communicate with each other directly. Throughout the paper we have the following assumptions.

T1) $G = (V, E)$ is a connected and undirected graph.

T2) the communication between two connected sensors is perfect.

Let N_l denote the set of the sensors that are connected with sensor l, i.e., $N_l = \{m | (l, m) \in E\}$.

Distributed consensus algorithms are iterative algorithms that are based on information exchange between neighbor sensor nodes at each iteration. Let $x_l(t)$ denote the updated LLR at sensor l at time l with $x_l(0) = x_l$, where t is the index for iteration. Then, each sensor can have the average of $\{x_l\}$ by using the following iterative distributed algorithm [9]:

$$x_l(t+1) = x_l(t) + \mu \sum_{m \in N_l} (x_m(t) - x_l(t)), \ l = 1, \ldots, L, \tag{4}$$

where μ is the gain for the disagreement, $x_m(t) - x_l(t)$. Although there are other approaches (e.g., generalized algorithms of (4) in [10]), we only consider the distributed consensus algorithm in (4) in this paper. The convergence properties of the iterative algorithm in (4) depends on the Laplacian matrix. For a given graph (V, E), the adjacency matrix is given by

$$[\mathbf{A}]_{l,m} = \begin{cases} 1, & \text{if } (l, m) \text{ or } (m, l) \in E; \\ 0, & \text{otherwise.} \end{cases}$$

The Laplacian matrix is given by

$$\mathbf{L} = \text{diag}(d_1, \ldots, d_L) - \mathbf{A}.$$

where $d_l = |N_l|$ is the degree of node l. The iterative algorithm in (4) is now rewritten as

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{x}(t) - \mu \mathbf{L} \mathbf{x}(t) \\ &= (\mathbf{I} - \mu \mathbf{L}) \mathbf{x}(t), \end{aligned} \tag{5}$$

where $\mathbf{x}(t) = [x_1(t) \ \ldots \ x_L(t)]^{\mathrm{T}}$. While the largest eigenvalues of $\mathbf{I} - \mu\mathbf{L}$ is 1, the magnitudes of the other eigenvalues are less than 1 if $0 < \mu < \frac{1}{\max_l d_l}$. Since the eigenvector corresponding to the largest eigenvalue of $\mathbf{I} - \mu\mathbf{L}$ is $\frac{1}{\sqrt{L}}[1 \ \ldots \ 1]^{\mathrm{T}}$

[13], it can be shown that

$$\lim_{t\to\infty} x_l(t) = \frac{1}{L}\sum_{l=1}^{L} x_l = \bar{x},\ l = 1,\dots,L. \quad (6)$$

Each sensor can make a global decision through an optimal decision rule as the global LLR, $\bar{x}$, is available.

III. System Model with Outliers

In a WSN, the observations of some nodes can be manipulated by intruders. In this case, the local LLRs of these nodes become outliers and can lead to a biased result when a consensus-based algorithm is used for distributed detection. If no intruders exist, each sensor has the observation as in (1). On the other hand, if intruders are present, they can manipulate the observations of some sensors. As a result, a sensor can have the following observation:

$$s_l = \begin{cases} \mu_m + n_l; & \text{if there is no interference under } H_m; \\ A + n_l; & \text{if there is interference (outlier),} \end{cases} \quad (7)$$

where A is the level of interference. Throughout the paper, for the sake of simplicity, we assume that $\mu_1 = -\mu_0 = \mu > 0$. Depending on the value of A, there will be different impact on the global decision. In general, it is expected that the magnitude of A is large and the sensors' local LLRs with interfering signals become outliers. Consensus-based distributed detection algorithms are vulnerable to these outliers. To model local observations when intruders exist, we consider the following assumptions:

A1) s_l is conditionally independent as in (1) and (7).

A2) Let P_U denote the probability that intruders exist. In the presence of intruders, each sensor can have interference independently with P_U. That is, $s_l = \mu_m + n_l$ with probability $1 - P_U$ and $s_l = A + n_l$ with probability P_U if intruders exist.

The nodes that have strong interference, i.e., whose observation is $A + n_l$, are called outlier nodes in this paper. Under **A1**), we can see that the local LLR is a scaled observation. Thus, for convenience, we assume that

$$x_l = s_l,\ l = 1,\dots,L.$$

Based on the model in (7) and the associated assumptions (i.e., **A1**) and **A2**)), there can be three different hypotheses: H_0, H_1, and U, where U denotes the hypothesis that there exist manipulated sensors by intruders. The ML decision rule is given by

$$\hat{\theta} = \arg \max_{\theta\in\{H_0,H_1,U\}} f(\mathbf{s}|\theta), \quad (8)$$

where $f(\mathbf{s}|\theta)$ is the likelihood function of $\theta \in \{H_0, H_1, U\}$ for given $\mathbf{s} = \{s_1,\dots,s_L\}$. Provided that prior probabilities are available, the MAP decision rule is applicable. A Bayesian approach can also be obtained by assigning costs. These approaches may not be practical due to the following reasons: *i*) $\mathbf{s}$ may not be available for each sensor node (this usually requires extensive inter-communications between sensors); *ii*) $f(\mathbf{s}|U)$ does not have a simple expression.

We have

$$f(\mathbf{s}|U) = f(\mathbf{s}|U,H_0)\Pr(H_0) + f(\mathbf{s}|U,H_1)\Pr(H_1), \quad (9)$$

where $f(\mathbf{s}|U,H_m)$ is the likelihood function when outliers exist among $\{s_l\}$ under H_m, which is given by

$$f(\mathbf{s}|U,H_m) = \sum_{k=1}^{L} f(\mathbf{s}|U_k,H_m)P_k, \quad (10)$$

where $f(\mathbf{s}|U_k,H_m)$ denotes the likelihood function when k sensors have interference under H_m and P_k denotes the probability that k sensors have interference. Since

$$f(\mathbf{s}|U_k,H_m) = \sum_{\mathcal{I}_k}\prod_{l\in\mathcal{I}_k} f_I(s_l)\prod_{l\in\mathcal{I}_k^c} f_I(s_l|H_m);$$
$$P_k = P_U^k(1-P_U)^{L-k}, \quad (11)$$

where $\mathcal{I}_k$ is the index set of all the possible combinations of k sensors with interference, we can see that $f(\mathbf{s}|U)$ in (9) is a Gaussian mixture. Note that each sensor node needs to know A, P_U, and $\{\Pr(H_0),\Pr(H_1)\}$. While the conventional approaches to detect the presence of outliers are applicable when there exists a FC for central processing and it knows $f(\mathbf{s}|U,H_m)$, they may not be applicable when each sensor node has to perform the outlier detection using distributed consensus algorithms due to these difficulties.

IV. Self-aware Outlier Detection

We present outlier detection approaches that can be carried out at each sensor in WSNs using the test statistics that are available using distributed consensus algorithms.

A. Approximation for Outlier Detection

As pointed out earlier, the standard approaches for the outlier detection at each sensor are not feasible. Thus, we can consider an approximation. We assume that the average of local LLRs is a Gaussian random variable (rather than a Gaussian mixture) under U and H_m. From **A1**), the conditional sample mean of local LLRs has the mean and variance as follows:

$$\mathbb{E}[\bar{x}|U,H_m] = \mu_{U,m} = \mu_m(1-P_U) + AP_U;$$
$$\text{Var}(\bar{x}|U,H_m) = \sigma^2_{U,m} = \frac{1}{L}\left(\sigma^2 + (A-\mu_m)^2 P_U(1-P_U)\right). \quad (12)$$

Then, we have

$$f(\bar{x}|H_m) = \mathcal{N}\left(\mu_m, \frac{\sigma^2}{N}\right);$$
$$f(\bar{x}|U,H_m) = \mathcal{N}\left(\mu_{U,m}, \sigma^2_{U,m}\right). \quad (13)$$

As shown above, in the presence of outliers, the average of local LLRs depends on H_m. From this, including U, we can consider 4-ary hypothesis testing. Let $\theta = \{H_0, H_1, (U,H_0), (U,H_1)\}$. Then, the MAP decision rule is given by

$$\hat{\theta} = \arg \max_{\theta=\{H_0,H_1,(U,H_0),(U,H_1)\}} f(\bar{x}|\theta)\Pr(\theta), \quad (14)$$

where

$$\Pr(\theta) = \begin{cases} (1-P_U)\Pr(H_0), & \text{if } \theta = H_0; \\ (1-P_U)\Pr(H_1), & \text{if } \theta = H_1; \\ P_U \Pr(H_0), & \text{if } \theta = (U, H_0); \\ P_U \Pr(H_1), & \text{if } \theta = (U, H_1). \end{cases} \tag{15}$$

If A is not known, the ML estimate of A can be used as in the generalized ML detection [14], which is given by

$$\begin{aligned} \hat{A}_m &= \arg\max_A f(\bar{x}|U, H_m, A) \\ &= \arg\min_A \frac{(\bar{x}-\mu_{U,m})^2}{\sigma_{U,m}^2} \\ &= \frac{\bar{x}-\mu_m(1-P_U)}{P_U}. \end{aligned} \tag{16}$$

The third equality in (16) results from the fact that $\frac{(\bar{x}-\mu_{U,m})^2}{\sigma_{U,m}^2} \geq 0$ for any A. In this case, since $(\bar{x}-\mu_{U,m})^2 = 0$ when $\hat{A}_m$ replaces A, the resulting likelihood functions, $f(\bar{x}|U, H_m, \hat{A}_m)$, become

$$\begin{aligned} f(\bar{x}|U, H_m, \hat{A}_m) &= \sqrt{\frac{L}{2\pi\left(\sigma^2 + \left(\hat{A}_m - \mu_m\right)^2 P_U(1-P_U)\right)}} \\ &= \sqrt{\frac{L}{2\pi\left(\sigma^2 + \frac{1-P_U}{P_U}(\bar{x}-\mu_m)^2\right)}}. \end{aligned} \tag{17}$$

Note that the ML estimate of A in (16) is unbiased under the correct hypothesis. That is, under H_m, using (12), we can show that

$$\mathbb{E}[\hat{A}_m|H_m] = \mathbb{E}\left[\frac{\bar{x}-\mu_m(1-P_U)}{P_U}|H_m\right] = A.$$

Under the incorrect hypothesis, the mean of the estimate of A is given by

$$\begin{aligned} \mathbb{E}[\hat{A}_m|H_{m'}] &= \mathbb{E}\left[\frac{\bar{x}-\mu_m(1-P_U)}{P_U}\middle|H_{m'}\right] \\ &= A + \frac{(\mu_m-\mu_{m'})(1-P_U)}{P_U}. \end{aligned} \tag{18}$$

In this case, we can observe that the estimate of A is less biased if P_U approaches 1. This implies that the performance of the outlier detection in (14) can be better as P_U increases.

Note that if local LLRs or s_l are assumed to be Gaussian as in (13), $\bar{x}$ becomes a sufficient statistic. Thus, each sensor can perform the outlier detection with $\bar{x}$ individually. Furthermore, when A is not known, the estimate of A can be used. Since this estimate in (16) is a function of $\bar{x}$, each sensor node can have the same estimate of A and perform the same outlier detection.

B. Mitigation of the Influence of Outliers

Through the outlier detection in Subsection IV-A, each sensor can detect the presence of outliers within the WSN. Although the MAP decision rule in (14) can provide a decision on the target (if H_m or (U, H_m) is accepted from the MAP decision rule, we assume that H_m is true), the decision performance can be improved if outliers are removed. In order to remove outliers, each sensor has to know the presence of interference in its observation or local LLR. This requires another hypothesis test.

Suppose that the outlier detection in Subsection IV-A accepts hypothesis (U, H_0) or (U, H_1). This outlier detection is referred to as the overall outlier detection (OOD). Then, each sensor can perform binary hypothesis testing to see the presence of interference in its observation. This detection is referred to as the sensor-level outlier detection (SOD) with the two hypotheses in (7). Thus, the proposed approach in this paper is a two-step approach: i) in the first step, the OOD is performed once $\bar{x}$ is available; ii) in the second step, the SOD is performed if the OOD chooses (U, H_0) or (U, H_1). For convenience, let G and $\bar{G}$ denote the hypotheses of no interference under either H_0 or H_1 and of interference, respectively. The MAP decision rule for SOD can be given by

$$\hat{\theta} = \arg\max_{\theta\in\{G,\bar{G}\}} f(s_l|\theta)\Pr(\theta), \tag{19}$$

where, from (7),

$$f(s_l|\theta) = \begin{cases} \mathcal{N}(\mu_0,\sigma^2)\Pr(H_0) + \mathcal{N}(\mu_1,\sigma^2)\Pr(H_1), & \text{if } \theta = G; \\ \mathcal{N}(A,\sigma^2), & \text{if } \theta = \bar{G} \end{cases} \tag{20}$$

and

$$\Pr(\theta) = \begin{cases} (1-P_U), & \text{if } \theta = G; \\ P_U, & \text{if } \theta = \bar{G}. \end{cases} \tag{21}$$

Note that hypothesis G is a composite hypothesis as $G = H_0 \cup H_1$. If the interference level, A, is not available, the ML estimate in (16) can be used. In this case, the conditional distribution, $f(s_l|\bar{G})$ is replaced by

$$f(s_l|\bar{G}) = \mathcal{N}(\hat{A}_0,\sigma^2)\Pr(H_0,U) + \mathcal{N}(\hat{A}_1,\sigma^2)\Pr(H_1,U). \tag{22}$$

Through the SOD, each sensor can decide that whether or not s_l is an outlier. With this decision, the sensor nodes can carry out the *second* run of the iterative distributed algorithm to have a better average of local LLRs after removing outliers. If sensor l accepts G (i.e., its observation does not have interference), this sensor will set the original initial value as $x_l(0) = s_l$ in the second run. On the other hand, if sensor l accepts $\bar{G}$, this sensor should not participate in the second run. Note that if this sensor sets the initial value to zero, i.e., $x_l(0) = 0$ for the purpose of removing its corrupted observation, this will result in a biased decision. Since the iterative distributed algorithm is to find the average of local LLRs, the average will be biased to 0 as the local LLRs of the corrupted observations are set to zero. To avoid this problem, an adjacent sensor accepting $\bar{G}$ should be merged to a sensor accepting G. However, this merging process reduces the number of effective sensor nodes.

The merging process of outlier nodes plays a key role in improving the performance of distributed detection by removing potential outliers in the second run of the iterative distributed algorithm. Suppose that sensor l is an outlier node and there is a sensor node in N_l which is not an outlier node. For convenience, let m denote the index of this non-outlier node. In the merging process, sensor l becomes a relay node of sensor node m and does not involve in the iterative distributed algorithm. The resulting merged node can have the index m, while the index l (for the outlier node) is removed. In this case, N_m is updated as

$$N_m \Leftarrow N_m \cup N_l.$$

All the outlier node can be merged with non-outlier nodes. As a result, if the number of outlier nodes is L_{out}, the number of effective nodes for the distributed detection becomes $L - L_{\text{out}}$. Clearly, this merging process changes the network topology in terms of the graph and the associated adjacency matrix.

V. Simulation Results

We present simulation results with the system model under **A1**) and **A2**) when $\mu_0 = -\mu_1 = 1$. We consider the following environments.

- L sensor nodes are uniformly located within a square area and they are connected. In simulations, we assume that $L = 100$.
- Outlier sensor nodes are uniformly located. A sensor node can be an outlier node with probability P_U independently. Thus, the spatial correlation cannot be exploited to identify outlier nodes.
- $\Pr(H_0) = \Pr(H_1) = \frac{1}{2}$.

The probability of error of the target detection is given by

$$\begin{aligned} P_{\text{et}} &= \Pr(\text{accept } H_1|H_0)\Pr(H_0) + \Pr(\text{accept } H_0|H_1)\Pr(H_1) \\ &= \frac{1}{2}\left(\Pr(\text{accept } H_1|H_0) + \Pr(\text{accept } H_0|H_1)\right). \end{aligned}$$

In simulations, we will assume that $A \gg 0$. Thus, $\Pr(\text{accept } H_1|H_0)$ would be negligible. It follows that

$$P_{\text{et}} \approx \frac{\Pr(\text{accept } H_0|H_1)}{2}.$$

For convenience, denote by $P_{\text{et},1}$ and $P_{\text{et},2}$ the probabilities of error of the target detection after the first and second runs of the iterative distributed algorithm, respectively. Note that if the OOD fails to detect outliers, the second run is not carried out and the result of the first run becomes the final one. Thus, under $P_U > 0$, we have

$$P_{\text{et},2} = P_{\text{et},1}P_{\text{e}-\text{ood}} + (1 - P_{\text{e}-\text{ood}})P_2, \tag{23}$$

where $P_{\text{e}-\text{ood}}$ denotes the probability of error of OOD and P_2 denotes the probability of error of the distributed detection in the second run of the iterative distributed algorithm after each sensor node performs SOD. If $P_2 \ll P_{\text{et},1}$, we have $P_{\text{et},2} \approx P_{\text{et},1}P_{\text{e}-\text{ood}}$. Thus, $P_{\text{et},2}$ is strongly related to the probability of error of OOD.

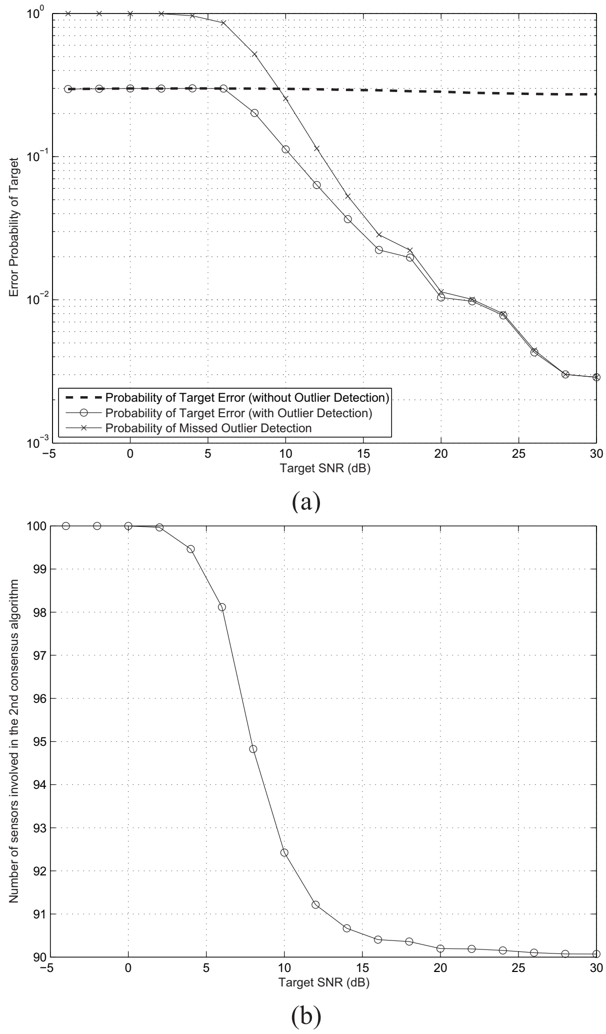
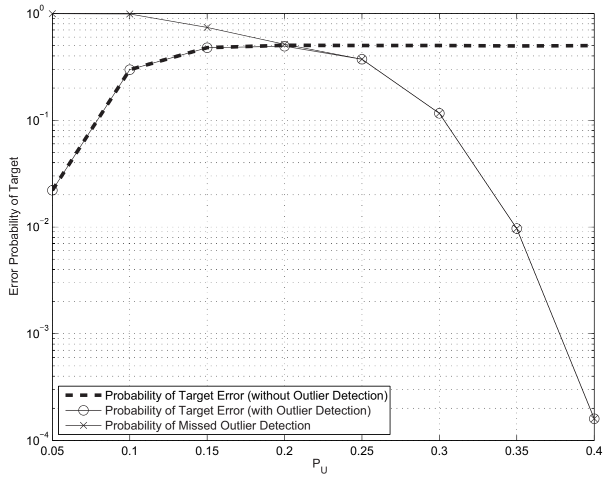


Fig. 1. Performance of the distributed detection with/without outlier detection for different values of the target SNR ($P_U = 0.1$ and $A = 10$): (a) error probabilities; (b) the number of sensors in the second run of the iterative distributed algorithm.

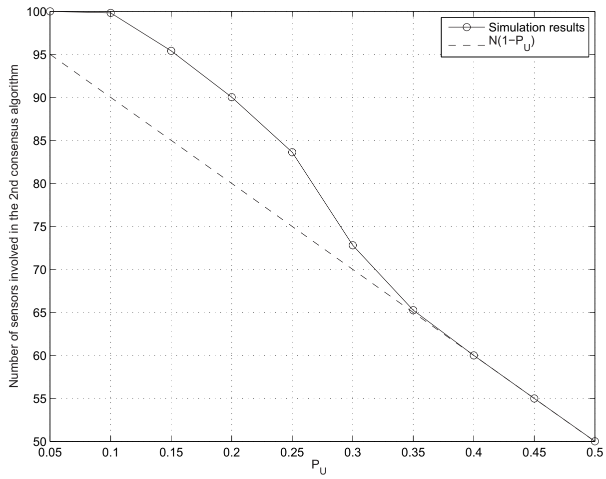
Fig. 1 shows the simulation results for different values of the target SNR, $\frac{|\mu_0|^2}{\sigma^2}$, when $P_U = 0.1$ and $A = 10$. The dashed line represents the probability of error of the target detection without the outlier detection, $P_{\text{et},1}$, (i.e., the results after the first run of the iterative distributed algorithm), while the solid line with circle marks represents that with the outlier detection, $P_{\text{et},2}$ (i.e., the results after the second run of the iterative distributed algorithm). The solid line with cross marks represents the probability of a miss in OOD. Without the outlier detection, there is almost no performance improvement by increasing the target SNR. However, when the outlier detection is employed, the performance is significantly improved and this performance improvement is closely related to the performance of OOD as shown in Fig. 1 (a). This shows that the outlier detection is essential in the distributed detection when intruders exist to manipulate sensors' observations. Note that as shown in Fig. 1 (b), the number of the sensors in the second run of the iterative distributive algorithm approaches $N(1 - P_U) = 90$, which is the average number of sensors without outliers, as the target SNR increases. Clearly, this means that each sensor can perform SOD well as the target SNR increases and this results in an improved performance in the target detection as shown in Fig. 1 (a).

In order to see the impact of P_U, simulations are carried out with different values of P_U and the results are shown in Fig. 2. It is assumed that $A = 10$ and the target SNR is 3 dB. Without the outlier detection, the performance of the distributed detection becomes worse as P_U increases (i.e., as more sensors have outlier observations). With the outlier detection, the performance of the distributed detection is improved as P_U increases as shown in Fig. 2 (a). This is an interesting result as it could imply that the performance is improved if more outlier nodes exists. This interesting behavior results from the performance of OOD (the solid lines with cross marks). Since the performance of OOD becomes better as P_U increases, outliers nodes are excluded in the second run of the iterative distributed algorithm and it results in a better performance. Fig. 2 (b) shows that the number of the sensors in the second run of the iterative distributive algorithm approaches $N(1 - P_U)$ as P_U increases.

In Fig. 2 (a), the probability of error of the target detection is shown to be a concave function of P_U. Since the outlier detection is not successful when P_U is low, $P_{\mathrm{et},2}$ could be close to $P_{\mathrm{et},1}$, which is increasing with P_U. As P_U increases, the probability of error of the outlier detection decreases and $P_{\mathrm{et},2}$ can also decrease. As a result, $P_{\mathrm{et},2}$ is a concave function of P_U. This means that there exists a certain value of P_U that maximizes $P_{\mathrm{et},2}$ or makes the worst performance.



(a)



(b)

Fig. 2. Performance of the distributed detection with/without outlier detection for different values of P_U (Target SNR = 3 dB and $A = 10$): (a) error probabilities; (b) the number of sensors in the second run of the iterative distributed algorithm.

VI. Concluding Remarks

We studied outlier detection for WSNs when consensus-based distributed detection algorithms are used. Since central processing is not available, each sensor node has to perform outlier detection with an iterative distributed detection algorithm. We showed that the proposed two-step approach for the outlier detection can effectively detect the presence of outliers within a WSN and mitigate them at each sensor level without central processing. Through simulation results, we have observed that the performance of the outlier detection decides the performance of target detection in the second run of the iterative distributed detection algorithm. Furthermore, it was shown that the performance of target detection after removing outliers is improved as P_U, the probability of intruders, increases.

References

[1] P. K. Varshney, *Distributed Detection and Data Fusion*, New York: Springer-Verlag, 1997.

[2] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Channel aware decision fusion in wireless sensor networks," *IEEE Trans. Signal Proc.*, vol. 52, no. 12, pp. 3454-3458, December 2004.

[3] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access over multiaccess fading channels," *IEEE Trans. Signal Proc.*, vol. 55(3), pp. 1081-1092, March 2007.

[4] H. Lee, C.C. Lim, and J. Choi, "Joint backoff control in time and frequency for multichannel wireless networks and its Markov model for analysis," to appear, Discrete and Continuous Dynamical Systems - Series B, May 2011.

[5] T. Palpanas, D. Papadopoulos, V. Kalogeraki, and D. Gunopulos, "Distributed deviation detection in sensor networks," in *Proc. Sigmod*, 2003, pp. 77-82.

[6] W. Wu, X. Cheng, M. Ding, K. Xing, F. Liu, and P. Deng, "Localized outlying and boundary data detection in sensor networks," *IEEE Trans. Knowledge and Data Engineering*, vol. 19(8), pp. 1145-1157, August 2007.

[7] I. C. Paschalidis and Y. Chen, "Statistical anomaly detection with sensor networks," *ACM Trans. Sensor Networks*, pp. 17:1-17:22, August 2010.

[8] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: A survey," *IEEE Commun. Surveys and Tutorials*, vol. 12, no. 2, pp. 159-170, 2nd Quarter 2010.

[9] R. O. Saber and R. M. Murray, "Consensus protocols for networks of dynamic agents," in *Proc. IEEE American Control Conf.*, 2003, pp. 951-956.

[10] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems and Control Letters*, vol. 53, pp. 65-78, 2004.

[11] S. A. Aldosari and J. M. F. Moura, "Distributed detection in sensor networks: connectivity graph and small world networks," in *Proc. IEEE 39'th Asilomar Conf. on Signals, Systems, and Computers*, 2005, pp. 230-234.

[12] H. V. Poor, *An Introduction to Signal Detection and Estimation: 2/e*, New York, NY: Springer-Verlage, 1994.

[13] F. Chung, *Spectral Graphs Theory*, American Mathematical Society, 1997.

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing II: Detection Theory*, Englewood Cliffs, NJ: Prentice Hall, 2002.