Distributed Detection in Wireless Sensor Networks under Byzantine Attacks

Junhai Luo and Zan Cao

School of Electronic Engineering, University of Electronic Science and Technology of China, Xiyuan Avenue, Chengdu 611731, China

Correspondence should be addressed to Junhai Luo; junhai_luo@uestc.edu.cn

Received 21 May 2015; Revised 21 August 2015; Accepted 28 September 2015

1. Introduction

Wireless sensor networks (WSNs) consist of a large number of tiny power-limited sensors that are densely and spatially deployed to monitor physical phenomena. When detecting a target in the region of interest (ROI), all the sensors in network report their findings to the fusion center (FC) where a global-final decision is made. For the advantage of easy deployment and fast self-organization, WSNs have been widely used [1]. Due to the increasing importance of being used in both military and civilian applications, it is imperative to incorporate secure localization and detection into WSNs. However, limited by both the processing capability and power supply of sensor nodes, secure detection in WSNs has been a challenging task. WSNs are also vulnerable to tampering. A serious threat to WSNs is Byzantine attacks where some authenticated sensor nodes have been fully controlled by an intelligent adversary. These compromised sensors are dispatched to disrupt or confuse the FC. While Byzantine attacks may, in general, refer to many types of Byzantine behaviors [2, 3], our focus in this paper is on Byzantine attacks in terms of data-falsification. In this type of attack, compromised nodes are reprogrammed and then forced to send falsified data to the FC in order to undermine the inference performance of network. The main goal of Byzantine attackers is to havoc performance of the FC as much as possible so that decoder at the FC is unable to utilize sensors’ information to determine the presence of target correctly.

An important task that WSNs need to perform is target detection, which is imperative for an accurate tracking of target. In the context of Byzantine attackers attempting to disrupt the network, a reliable algorithm of detection needs to be introduced. In [4, 5], several algorithms have been developed for secure detection and localization in WSNs. The techniques based on direction of arrival (DOA) and time of arrival (TOA) (or time-difference-of-arrival (TDOA)) have been investigated in [4] and [5], respectively. However, the TDOA is not suitable for detecting target because sensor nodes are narrow band and lack accurate synchronization. Several researchers have focused on developing techniques that do not suffer from imperfect time synchronization. For example, scheme of measuring intensity of signal energy is
used to detect target. Therefore, convenience of the energy-based method gives feasibility to detect and localize a target through detecting energy. In [6, 7], each cognitive radio adopts an effective detection scheme based on harvesting energy to make its decision and determine whether there is a licensed user requesting to occupy the spectrum band. Each cognitive radio sensor detects the intensity of signal and measures the energy received. If the measured energy is greater than a properly predefined threshold, the current spectrum is busy. Otherwise, the current spectrum is idle. We adopt this scheme in this paper to detect whether there is a target in the ROI.

Marano et al. have investigated distributed detection in the presence of Byzantine attacks in [8, 9], where Byzantine attackers are assumed to have a complete knowledge about the true hypotheses. In their system model, the authors assumed that the FC did not know which sensor node was Byzantine attacker. Though the FC did not know which sensor node was Byzantine attacker, it knew average percentage of compromised sensor nodes or upper bound of the average. Vempaty et al. have also analyzed the problem of distributed detection under Byzantine attacks in [10]. However, they did not assume that the Byzantine sensors had a complete knowledge about true hypotheses. Instead, they assumed that the Byzantine sensors made decisions about the presence of target through their own observations. In other words, each Byzantine sensor potentially flips the local decision made at the node. The performance of network has also been analyzed in the context of presence of independent and collaborative Byzantine attacks, respectively, in [10]. In addition to the analysis of distributed detection in the presence of Byzantine attacks, an adaptive learning scheme that mitigated the effect brought by Byzantine attackers has been proposed by Vempaty et al. in [10]. The scheme identifies Byzantine sensors through observing the ratio value of deviations between the estimated behavior of ith sensor and the expected behavior of an Honest sensor to the estimated behavior of ith sensor and the expected behavior of a Byzantine sensor. If the value is greater than 1, the FC declares the sensor to be Byzantine. Otherwise, the FC labels the sensor as Honest. In order to maximize the performance of decision fusion, the FC then removes those decisions that are from sensors tagged as Byzantine when the FC makes a global decision at the next time.

In literature [10, 11], analysis of performance of the network under independent and collaborative malicious Byzantine attacks has been performed, respectively. When a target enters into the monitoring region, the ith compromised sensor node has a detection and makes a local/original decision. In the case of independent malicious Byzantine attacks (IMBA), the local decision of ith compromised sensor completely depends on its own observation. In the case of collaborative Byzantine attacks (CMB), the local decision of ith compromised sensor depends on not only its own observation but also decisions from the remaining compromised sensors. In the scenario of IMBA, though Byzantine sensor can make multiple observations in a proper time window and average these observations in order to reduce the effect of additive noise, it is not effective significantly when Byzantine sensor nodes stay far away from a target in the ROI. In addition, a sensor that adopts IMBA has no way to conquer a certain degree of blind flipping of its own decision. In the situation of CMB, each Byzantine sensor communicates with the left compromised sensors and refers to their decisions before determining its own local decision. Although CMB produces remarkable improvement in attacking effect, much energy has been consumed in communication among Byzantine sensors, especially when Byzantine sensor nodes are deployed sparsely. In addition to analysis of performance of distributed detection in the presence of Byzantine attacks, the blinding attacking power has been obtained. The blinding attacking power (\(\alpha_{blind}\)) is equal to 0.5 and 0.35 in the case of IMBA and CMB, respectively [11]. Motivated by this, we propose a new Byzantine attacks model named after neighborhood malicious Byzantine attacks (NMBA). NMBA is such a kind of attacks model where each compromised sensor node in the network determines whether there is a target or not in the ROI depending on not only its own local decision but also a certain amount of decisions coming from Honest sensors which are nearest around the compromised sensor. Then each Byzantine sensor node employs a majority strategy among decisions to make a final local decision. At last the Byzantine sensor node flips the final local decision confidently and sends false decision to the FC.

Attacking power \(\alpha\) which is also termed indicator of the vulnerability of the sensor networks is a crucial performance metric [12, 13]. We assume that there are a large number of sensor nodes deployed in the ROI. According to the law of large numbers, \(\alpha\) is equal to the ratio of number of compromised sensors to the total number of sensors in the network (\(\alpha \in [0, 1]\)). In practice, although the FC knows the presence of Byzantine sensors in the network, decider at the FC can hardly determine the exact attacking power [14]. If the decider knows the attacking power, it is convenient for the FC to adopt a robust strategy to mitigate the negative effect caused by Byzantine attackers [15, 16]. In this paper, we propose a simple and effective scheme to determine the attacking power in the perspective of decider at the FC. After the attacking power is estimated, two kinds of discrepancy distance which are used to help in identifying Byzantine sensors are constructed in this paper.

An effective scheme of decision fusion plays an important role in the FC [17–19]. Many literatures focused on the mitigation of Byzantine attacks and developed algorithms to design a static and identical threshold for decision making at the FC [11, 14]. Several authors have proposed online learning of normal trajectory patterns for detection in trajectory [20]. In this paper, we propose an effective scheme based on both dynamic threshold and identifying Byzantine attackers for decision fusion at the FC.

The paper is organized as follows. In Section 2, we describe our system model including detection model and NMBA attacking model. In order to formulate the problem of distributed detection in WSNs clearly, we divide the process of decision fusion into three hierarchies or stages in this section. The attacks model of NMBA is proposed at the first stage which is different from independent Byzantine attacks and collaborative Byzantine attacks. The performance
metric is also presented. In Section 3, we determine the optimal attacking strategy in the perspective of Byzantine attackers and closed-form expression for blinding region is derived. Comparison among IMBA, CMBA, and NMBA is also performed and numerical results are provided at the same time. From the perspective of network designer, we propose a fusion scheme based on dynamic threshold to make a reliable global decision and analyze how the FC identifies Byzantine attackers to enhance the fusion performance in Section 4. The attacking power is also estimated. Finally, we present our conclusion in Section 5.

2. System Model

2.1. Detection Model. A network with N sensor nodes which are spatially deployed in the ROI is considered. All sensor nodes in this network are independent on functionality. Each sensor makes a decision independently after detection. As illustrated in Figure 1, the sensor nodes which are denoted as symbol of plus are shown to be deployed on a regular grid and intensity of energy attenuated as the distance from the target that is represented as blue star increases. It is worth mentioning that the detection scheme based on harvesting energy is capable of handling any kind of deployment as long as the location information of each sensor node is available at the FC. The uniform sensor deployment shown in Figure 1 is only a special case. In any one kind of deployment, N sensor nodes can correctly detect a target when the target intrudes at the position \( \theta = (x_i, y_i) \), where \( x_i \) and \( y_i \) denote the coordinate of this target location in 2D Cartesian. We introduce an isotropic intensity of signal attenuation model as follows:

\[
a_i^2 = P_0 \left( \frac{d_0}{d_{Ei}} \right)^n, \tag{1}
\]

where \( a_i \) is the signal amplitude received at \( i \)th sensor and \( P_0 \) is the emitted power measured at a reference distance \( d_0 \). \( n \) is the power decay exponent, and \( d_{Ei} \) is the Euclidean distance between the target and \( i \)th sensor:

\[
d_{Ei} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \quad i = 1, 2, \ldots, N, \quad j = 1, 2, \ldots, N, \tag{2}
\]

in which \((x_j, y_j)\) are the coordinate of \( j \)th sensor. For simplicity but without loss of generality in this paper, we let \( n = 2, d_0 = 1 \) [10]. As a result, (1) can be expressed as

\[
a_i^2 = \frac{P_0}{(d_{Ei})^2}, \quad i = 1, 2, \ldots, N. \tag{3}
\]

Equation (3) is a quite general model for signal attenuation of electromagnetic wave that propagates isotropically in free space. However, when the signal of energy arrives at \( i \)th sensor, it has been contaminated by additive white Gaussian noise in practice. Therefore, the signal amplitude received at \( i \)th sensor is expressed as \( r_i = a_i + n_i \), in which \( n_i \) is Gaussian noise which follows standard normal distribution. Here, we assume that all sensors in the network have the identical additive white Gaussian noise, that is, \( n_i \sim N(\mu, \sigma^2) \), \( i = 1, 2, \ldots, N \).

Each sensor node needs to quantize the received signal of energy because of its limitations of bandwidth and energy and sends quantized binary measurements to the FC. Threshold of quantizers is adopted in this work for its simplicity of both easy implementation and analysis as follows:

\[
\bar{d}_i = \begin{cases} 
1, & r_i > \zeta_i, \\
0, & r_i < \zeta_i, 
\end{cases}
\]

\[
\zeta_i = \zeta, \quad i = 1, 2, \ldots, N, \tag{4}
\]

where \( \bar{d}_i \) and \( \zeta_i \) are local decisions made by \( i \)th sensor after quantizing the received signal and a predefined threshold adopted by \( i \)th sensor, respectively. In this paper, we assume that all of the sensors share the identical threshold; that is, \( \zeta = \zeta, i = 1, 2, \ldots, N \).

In this work, the classical distribution detection model is taken into account where two hypotheses are considered. Each sensor solves hypothesis testing problem and makes a local decision on either hypothesis \( H_0 \) (target is absent) or \( H_1 \) (target is present). We consider the scenario that the adversary knows the complete information about the location of sensors and is capable of attacking all the sensors simultaneously. Due to the constraint of budget, the Byzantine attackers conquer only a part of nodes in the network to deteriorate capability of inference performance of network. These Byzantine sensors transmit false decision to the FC in order to deteriorate inference performance of the network. We assume that the channel between the FC and local sensors is error-free. The original or local one-bit decision generated at \( i \)th sensor node is denoted as \( \tilde{d}_i \in \{0, 1\} \), \( i = 1, 2, \ldots, N \). Then the \( i \)th sensor reports one-bit decision \( d_i \) to the FC where \( \tilde{d}_i = \tilde{d}_i \) if \( i \)th sensor is Honest. For a Byzantine sensor, the local original decision \( d_i \) need not be equal to \( \tilde{d}_i \) in our attacks model.

Let \( N_H \) and \( N_B \) be the number of Honest and Byzantine sensors, respectively. The total number of sensors can be expressed as \( N = N_H + N_B \) and the number of Byzantine sensor nodes \( N_B \) is equal to \( \alpha \cdot N \). In the perspective of
Byzantine attackers, conquering $N$ sensors is not a wise strategy for the adversary itself at the risk of exposed activity. The main goal of adversary is to compromise a fraction of the sensors to degrade the performance of the FC instead of capturing the network with a huge cost. Therefore, we have $N_B < N$. We use $P_d^H(i) = Pr(d_i = 1 \mid H_1, H)$ and $P_{fa}^H(i) = Pr(d_i = 1 \mid H_0, H)$ to denote the probability of detection and false-alarm of $i$th sensor, respectively. We use $H$ to present a sensor node to be Honest and $i \in \{1, 2, \ldots, N_H\}$. The decision probability of $i$th sensor can be expressed as

$$P_d^H(i) = Pr(\tilde{d}_i = 1 \mid H_1, H) = Pr(a_i + n_i > \varsigma_i) = Q\left(\frac{\varsigma_i - a_i - \mu}{\sigma}\right),$$

Similarly, the false-alarm probability of $i$th sensor can be expressed as

$$P_{fa}^H(i) = Pr(n_i > \varsigma_i) = Q\left(\frac{\varsigma_i - \mu}{\sigma}\right),$$

where $Q(\cdot)$ is the complementary distribution function of the standard Gaussian

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt.$$

When a target intrudes into the ROI, each sensor node starts to sense and record the energy propagated from the target using detection scheme based on harvesting energy [21]. We let each sensor perform $K$ observations in a small time window $T$ where target is assumed to be static. This is a reasonable assumption. For example, if the sampling rate of each sensor is $6000$ Hz, a target with a speed of $100$ km/h only moves $0.25$ m during $T = 54$ sampling intervals [22]. The $j$th observation at $i$th sensor node can be expressed as $\tilde{d}_{ij}$, $i \in \{1, 2, \ldots, N\}$ and $j \in \{1, 2, \ldots, K\}$. A local/original decision matrix $\tilde{D} = [\tilde{d}_1, \tilde{d}_2, \ldots, \tilde{d}_N]^T$ is generated where $\tilde{d}_i = (\tilde{d}_{i1}, \tilde{d}_{i2}, \ldots, \tilde{d}_{iK})^T$ is the vector of local/original decision at the $i$th sensor node. And, $\tilde{d}_{ij} \in \{0, 1\}$, $i \in \{1, 2, \ldots, N\}$ and $j \in \{1, 2, \ldots, K\}$. The FC receives $N$ vectors of decisions from local sensors. Then, a decision matrix $D = [d_1, d_2, \ldots, d_N]^T$ is formulated at the FC; that is, $D = (d_{ij})_{N \times K}$, where $d_{ij} \in \{0, 1\}$, $i \in \{1, 2, \ldots, N\}$ and $j \in \{1, 2, \ldots, K\}$. The local/original decision matrix $\tilde{D}$ is equal to $D$ if there is no presence of Byzantine attackers.

In order to formulate the problem in the process of decision fusion, we divide the process into three hierarchies/stages. As illustrated in Figure 2, $i$th sensor makes a local/original vector of $\tilde{d}_i$ and sends the vector $d_i$ into the FC after $\tilde{d}_i$ is probably “attacked” at the first stage. A decision matrix $D$ is formulated from which the vector of global decision $z = (z_1, z_2, \ldots, z_K)$ is mapped at the second stage. At the last stage, a global-final decision $z$ is mapped from vector $z$ at the FC.

2.2. Byzantine Attacks Model. In the attacks model of NMBA, the $i$th Byzantine sensor has exactly $M_i - 1$ ($i = 1, 2, \ldots, N_B$) neighbors to consult and $M_i \leq N_H$. In order to facilitate analysis, we assume that the scenario of many Byzantine sensors flocking together does not happen. Namely, the whole Byzantine sensor nodes are deployed sparsely by intelligent adversary in the ROI. In the case of $N$ sensors deployed on a regular grid, NMBA has several neighborhood types including diamond type and square type. For each Byzantine sensor, its neighborhood nodes are those sensors that are the nearest and Honest around it in specific neighborhood type. We assume that each Byzantine sensor knows the identifications of the remaining compromised sensors. Each Byzantine sensor consults all of its neighborhood nodes to make a wise and tricky decision. In Figure 3, the type of square neighborhood is presented and the case of $M_i = M = 9$ is considered. Clearly, each Byzantine sensor node consults its eight neighbors and makes a decision based on decisions from its neighbors.

We make the conditional i.i.d. assumption under which observations from sensors are conditionally independent and identically distributed. The $j$th observation at $i$th sensor then has the distributions

$$H_0 \; : \; v_{ij}(k) = Pr(\hat{d}_{ij} = k \mid H_0) = (1 - \alpha) \cdot Pr(\hat{d}_{ij} = k \mid H_0, H) + \alpha \cdot Pr(\hat{d}_{ij} = k \mid H_0, B)$$

$$H_1 \; : \; u_{ij}(k) = Pr(\hat{d}_{ij} = k \mid H_1) = (1 - \alpha) \cdot Pr(\hat{d}_{ij} = k \mid H_1, H) + \alpha \cdot Pr(\hat{d}_{ij} = k \mid H_1, B).$$

Figure 2: Model of three hierarchies. $\tilde{d}_i$ is the vector of local decision made by $i$th sensor $S_i$, $d_i$ is the vector of decision sent to the FC, $i = 1, 2, \ldots, N$. $D_{N \times K}$ is the decision matrix formulated at the FC, $z$ is global decision vector, and $z$ is global-final decision.
If the sensor is Honest, its observation $k$ follows distributions $p$ and $q$ under hypotheses $H_0$ and $H_1$, respectively. Therefore, we have

$$H_0 : \Pr(\bar{d}_{ij} = k \mid H_0, H) = q_{ij}(k),$$
$$H_1 : \Pr(\bar{d}_{ij} = k \mid H_1, H) = p_{ij}(k),$$

$$i = 1, 2, \ldots, N_H.$$

According to (5), (6), and (7), we get

$$p_{ij}(1) = Q\left(\frac{\zeta - \alpha_i - \mu}{\sigma}\right),$$
$$p_{ij}(0) = 1 - p_{ij}(1),$$
$$q_{ij}(1) = Q\left(\frac{\zeta - \mu}{\sigma}\right),$$
$$q_{ij}(0) = 1 - q_{ij}(1).$$

Similarly, we have distributions $x$ and $y$ under the same hypotheses for Byzantine sensor as follows:

$$H_0 : \Pr(\bar{d}_{ij} = k \mid H_0, B) = y_{ij}(k),$$
$$H_1 : \Pr(\bar{d}_{ij} = k \mid H_1, B) = x_{ij}(k),$$

$$i = 1, 2, \ldots, N_B.$$

In the attacks model of NMBA, the $i$th Byzantine sensor makes an initial decision $c_{ij}$ independently and get the $M_i - 1$ decisions from its neighborhood sensors. As a result, a set of decisions $\{c_{ij} : l \in \{0, 1, \ldots, M_i - 1\}, c_l \in \{0, 1\}\}$ is obtained where the $c_l$ represents the decision from the $l$th neighbor of the $i$th Byzantine sensor. Then, the $i$th Byzantine sensor makes its local or original decision using a majority strategy, that is, the original local decision $\bar{d}_{ij} = \text{IF}(\sum_{l=0}^{M_i-1} c_l > \eta_i)$, where $\text{IF}(\cdot)$ and $\eta_i$ are indicator function and threshold adopted by the $i$th Byzantine sensor, respectively. Therefore, we have the following equations:

$$y_{ij}(k) = \sum_{m=\eta_i}^{M_i-1} \prod_{l=1}^{m} q_{\pi(l)j}(k) \prod_{l=m+1}^{M_i} (1 - q_{\pi(l)j}(k)),$$
$$x_{ij}(k) = \sum_{m=\eta_i}^{M_i-1} \prod_{l=1}^{m} p_{\pi(l)j}(k) \prod_{l=m+1}^{M_i} (1 - p_{\pi(l)j}(k)),$$

where $M_i \in \{1, 2, \ldots, N_H\}$, $1 \leq \eta \leq M_i$, and $\Gamma$ denotes the set of all permutations of the $M_i$ sensors. After using majority strategy to make a local decision, the $i$th Byzantine sensor flips confidently its decision with probability of $P_{flip} = 1$. Specifically, we have

$$\Pr(d_{ij} = k \mid \bar{d}_{ij} = l, B) = \begin{cases} 1, & \text{when } l \neq k \\ 0, & \text{when } l = k \end{cases}$$

$$k, l \in \{0, 1\}.$$

Thus, we get

$$\Pr(d_{ij} = k \mid H_0, B) = \sum_{l \neq k} \Pr(d_{ij} = k \mid \bar{d}_{ij} = l, H_0, B)$$
$$+ \sum_{l = k} \Pr(d_{ij} = k \mid \bar{d}_{ij} = l, H_0, B),$$
$$\Pr(d_{ij} = k \mid H_1, B) = \sum_{l \neq k} \Pr(d_{ij} = k \mid \bar{d}_{ij} = l, H_1, B)$$
$$+ \sum_{l = k} \Pr(d_{ij} = k \mid \bar{d}_{ij} = l, H_1, B).$$

Therefore, we have

$$\Pr(d_{ij} = 0 \mid H_0, B) = \Pr(\bar{d}_{ij} = 1 \mid H_0, B) = y_{ij}(1),$$
$$\Pr(d_{ij} = 1 \mid H_0, B) = \Pr(\bar{d}_{ij} = 0 \mid H_0, B) = y_{ij}(0),$$
$$\Pr(d_{ij} = 0 \mid H_1, B) = \Pr(\bar{d}_{ij} = 1 \mid H_1, B) = x_{ij}(1),$$
$$\Pr(d_{ij} = 1 \mid H_1, B) = \Pr(\bar{d}_{ij} = 0 \mid H_1, B) = x_{ij}(0).$$
Substituting (9) and (15) in (8) and after simplification, we obtain
\[ v_{ij}(k) = \Pr(d_{ij} = k \mid H_0) \]
\[ = (1 - \alpha) [kp_{ij}(k) + (1 - k) (1 - p_{ij}(k))] \]
\[ + \alpha y_{ij}(1 - k), \]
\[ u_{ij}(k) = \Pr(d_{ij} = k \mid H_1) \]
\[ = (1 - \alpha) [kp_{ij}(k) + (1 - k) (1 - p_{ij}(k))] \]
\[ + \alpha x_{ij}(1 - k). \] (16)

2.3. Performance Metric. In the perspective of Byzantine attackers, the primary objective is to deteriorate the inference performance of FC as much as possible. On the contrary, the FC wants to make inference performance as much highly excellent as possible in order to guarantee valid detection. In this paper, we adopt Kullback-Leibler divergence (KLD) as the network performance that characterizes inference performance of FC as much as possible. Byzantine attackers have the optimal strategy of Byzantine attacks that have been compromised have small KL divergence so that the FC can hardly make a right decision between normal and compromised sensors. We refer to this case as the FC being blinded completely when
\[ u_{ij}(k) = \Pr(d_{ij} = k \mid H_1) = \Pr(d_{ij} = k \mid H_0) \]
\[ = v_{ij}(k). \] (18)

Substituting (16) in (18) and after simplification, the condition to make KLD(u_{ij}(k) \parallel u_{ij}(k)) = 0 is equivalent to
\[ \alpha = \frac{p_{ij}(1) - q_{ij}(1)}{(p_{ij}(1) - q_{ij}(1)) + (1 - q_{ij}(1))} \]
\[ = \frac{p_{ij}(1) - q_{ij}(1)}{(p_{ij}(1) - q_{ij}(1)) + f(M_i, \eta, p_{ij}(1), q_{ij}(1))}, \] (19)
where the close-form expression of function \( f(\cdot) \) is denoted as the following equation:
\[ f(M_i, \eta, p_{ij}(1), q_{ij}(1)) = \sum_{m=\eta_i, \pi \in \Gamma} \sum_{i=1}^{m} p_{\pi(i)j}(1) \prod_{i=m+1}^{M_i} (1 - p_{\pi(i)j}(1)) \]
\[ - \sum_{m=\eta_i, \pi \in \Gamma} \sum_{i=1}^{m} q_{\pi(i)j}(1) \prod_{i=m+1}^{M_i} (1 - q_{\pi(i)j}(1)), \] (20)
where \( M_i \) and \( \eta_i \) are the number of neighborhood nodes and threshold adopted by \( i \)th Byzantine sensor, respectively.

As mentioned above, the KL distance between \( v_{ij}(k) \) and \( u_{ij}(k) \) is equal to zero; that is, KLD(v_{ij} \parallel u_{ij}) = 0, if and only if
\[ v_{ij}(k) = u_{ij}(k). \] The FC is incapable of distinguishing the two distributions under \( H_0 \) and \( H_1 \) when KLD is equal to zero. The attackers then project interests in the minimum attacking power that can just make the ability of inference of the FC destroyed. Thus, the minimum attacking power in the context of NMBA is denoted as
\[ \alpha_{\text{blind}} = \min \{ \alpha; \alpha \text{ that make } KLD(u_{ij} \parallel v_{ij}) = 0 \}. \] (21)

For the sake of minimizing \( \alpha \) to reach \( \alpha_{\text{blind}} \), we have the following equation depending on operating point \((p_{ij}(1), q_{ij}(1))\):
\[ \alpha_{\text{blind}} = \frac{p_{ij}(1) - q_{ij}(1)}{(p_{ij}(1) - q_{ij}(1)) + \max \{ f(M_i, \eta, p_{ij}(1), q_{ij}(1)) \}}. \] (22)
Because of \( 0 < p_{ij}(1) - q_{ij}(1) \leq 1 \), we have the following inequality:
\[ \alpha_{\text{blind}} \leq \frac{1}{1 + \max \{ f(M_i, \eta, p_{ij}(1), q_{ij}(1)) \}}. \] (23)
To prove inequality (23), we apply the monotonic property of the function of \( x/(x + 1) \). Due to the function possessing differentiability, we have the following inequality:
\[ \frac{d}{dx} \left( \frac{x}{x + 1} \right) = \frac{1}{(x + 1)^2} > 0. \] (24)
Therefore, \( x/(x + 1) \) is a monotonically increasing function when \( 0 \leq x \leq 1 \). As a result, inequality (23) is certified. After certifying (23), we have the following equation:
\[ \eta_{\text{opt}} = \text{ceil} \left( \frac{M_i}{2} \right), \] (25)
in (19) reaches the maximum value point only when \( \eta = \eta_{\text{opt}} \). Therefore, for a pair of fixed operating points \((p_{ij}(1), q_{ij}(1))\), we have

\[
\max f = f\left(M_i, \text{ceil}\left(\frac{M_i}{2}\right), p_{ij}(1), q_{ij}(1)\right)
\]

and (22) can be represented as

\[
\alpha_{\text{blind}} = \frac{p_{ij}(1) - q_{ij}(1)}{(p_{ij}(1) - q_{ij}(1)) + \max f}.
\]

When the intelligent adversary poses attacking power \( \alpha \) which is greater than \( \alpha_{\text{blind}} \), the FC is trapped into a dilemma where the decision cannot utilize any information from sensors to make a correct decision. The situation refers to the fusion center’s entering into the blinding region where the FC is blinded completely.

3.2 Numerical Results. In this subsection, numerical results are presented to support our analysis of optimal attacking strategy for Byzantine attackers. We consider the network where \( N = 100 \) sensors are deployed on regular grid in the ROI under the Byzantine attacks model of NMBA. We set the power at the reference point \((d_0 = 1)\) as 200 and the signal amplitude arriving at the local sensor is contaminated by AWGN with mean value \( \mu = 0 \) and standard deviation \( \sigma = 3 \). A target intrudes at position \( \theta = (90, 90) \). We consider the square type of NMBA where all the Byzantine sensors adopt identical number of neighborhood nodes and optimal threshold; that is, \( M = M_i \) and \( \eta = \eta_{\text{opt}}, i = 1, 2, \ldots, N_B \). Attacking power \( \alpha \) is observed in the case of square type of NMBA over \( M \in \{1, 2, \ldots, N_B\} \) and \( \eta \in \{1, 2, \ldots, M\} \). In the blinding region, the FC is incapable of utilizing any information to make a decision to determine whether there is a target in the ROI. Byzantine users try to pay as much low cost as possible when the FC has been completely blinded by them. Attacking power \( \alpha \) in blinding region is a convex function of threshold \( \eta \)

under the NMBA attacking model. Therefore, it is possible for Byzantine attackers to adopt the least attacking power to blind the FC. From the numerical results presented in Figure 4, we see that the attacking power \( \alpha \) reaches minimum value only when \( \eta = 5 \) in the case of NMBA when \( M = 9 \). The minimum value of \( \alpha \) is equal to 0.3844. The result is intuitive and also indicates that the optimal threshold associates with the number of neighborhood nodes, \( M \). The other cases of NMBA are presented in Table 1. We observe that the minimum value of \( \alpha_{\text{blind}} \) is just obtained at the point where the optimal threshold \( \eta \) is got. In Figure 5, we present that optimal threshold increases with the increase in the value of \( M \). We observe that \( \eta_{\text{opt}} \) is always equivalent to \( M/2 \) when \( M \) is even and \( \eta_{\text{opt}} \) is equal to the maximum integer that is not larger than \( M/2 \) when \( M \) is odd. It is clear that optimal threshold Byzantine attacker associates with the number of its neighborhood. In Figures 6 and 7, it is shown that \( \alpha_{\text{blind}} \) is the monotonically decreasing function of \( M \) in the condition that \( \eta \) is always equal to \( \eta_{\text{opt}} \). As illustrated in Figures 5 and 6, IMBA is a particular case when \( M = 1 \). Under condition of \( M = 1 \), each Byzantine sensor makes a judgement only depending on its own observation and flips its decision confidently prior to sending to the FC. When \( M = 2 \), each Byzantine sensor consults with two neighborhood nodes around it before making a decision and the minimum attacking power in the blinding region is equal to 0.5. When \( M \) is equal to 2, the attacking effect brought by NMBA matches with IMBA. \( \alpha_{\text{blind}} \) of NMBA decreases monotonically with the increase of \( M \) and it reaches 0.3844 when \( M \) increases to 9. \( \alpha_{\text{blind}} \) decreases to 0.3752 when \( M \) increases to 25. The exact value of \( \alpha_{\text{blind}} \) corresponding to \( M \) and \( \eta_{\text{opt}} \) is presented in Table 2. In addition, as we can see from Figures 6 and 7, the number of neighborhood nodes that each Byzantine sensor node consults including itself should be odd for better attacking efficiency. To further analyze the attacks of NMBA, we compare it with IMBA and CMBA in terms of KL distance about attacking power. As we discussed above, IMBA is a special case of NMBA when \( M = 1 \).
Table 1: $\alpha$ at $(M, \eta)$, $M \in \{1, 2, \ldots, 9\}$ and $\eta \leq M$.

| $M/\alpha$ | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9     |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1          | 0.5000| 0.5000|       |       |       |       |       |       |       |
| 2          | 0.5000| 0.5000| 0.5434| 0.5434|       |       |       |       |       |
| 3          | 0.5952| 0.4310| 0.5434| 0.4310| 0.5952|       |       |       |       |
| 4          | 0.6469| 0.4509| 0.4042| 0.4509| 0.6469|       |       |       |       |
| 5          | 0.7100| 0.5100| 0.4145| 0.3913| 0.4145| 0.5100| 0.7100|       |       |
| 6          | 0.7814| 0.5438| 0.4298| 0.3913| 0.3913| 0.4298| 0.5438| 0.7814|       |
| 7          | 0.8171| 0.5790| 0.4484| 0.3970| 0.3844| 0.3970| 0.4484| 0.5790| 0.8171|

$L$ out of $N_B$ fusion rule has been used for CMBA among the Byzantine sensors. And Byzantine sensors collaborate together to make decisions about the presence of target. If the attacking power is greater than 0.5, the FC can be blinded completely by any kinds of attacking strategies adopted by Byzantine attackers. Therefore, our analysis about detection performance is under the condition of $\alpha \leq 0.5$. In Figures 8 and 9, we plot the KLD of three attacking models against $\alpha$ in the case $M = 9$ and $M = 25$, respectively. From the numerical results presented in Figures 8 and 9, we can see that the KLD of NMBA is very close to CMBA and outperforms IMBA significantly. Though the performance of NMBA is very close to CMBA, the difference of KL distance between them varies with attacking power. In Figures 8 and 9, KL distance of CMBA is greater than that of NMBA in the region of $\alpha \in [0, 0.09]$ and $\alpha \in [0, 0.25]$, respectively. However, CMBA performs better when $\alpha$ meets the condition of $\alpha \geq 0.09$ and $\alpha \geq 0.025$ in Figures 8 and 9, respectively. For CMBA, the blinding point C which is marked by black-square in Figures 8 and 9 is determined by the total number of Byzantine sensor nodes instead of $M$. For NMBA, the blinding point B which is marked by blue-square in Figures 8 and 9 is determined by the number of neighborhood nodes. For NMBA, $M$ is larger and $\alpha_{\text{blind}}$ is smaller. Furthermore, $\alpha_{\text{blind}} = 0.3844$ and $\alpha_{\text{blind}} = 0.3751$ when $M$ is equal to 9 and 25, respectively.

4. Fusion Center Decision Strategy

In this section, let us solve the problem of identifying Byzantine attackers to enhance inference performance of the FC. In order to explain the problem well, we have defined three types of decision which include local/original decision, global decision, and global-final decision. As we discussed in Section 2, a decision matrix $D'_{N \times K} = (d_j(t))_{N \times K}$ is formulated at the FC after $K$ observations in a time window $T$ at $t$th global-final decision making, $t \in N^+$. We let the FC make a corresponding global decision $z_j(t)$ over vector of decision $d_j$, at $j$th observation. And a vector of global decision $z(t) = (z_1(t), z_2(t), \ldots, z_K(t))$ is formulated over $\{d_1, d_2, \ldots, d_K\}$, $z_j(t) \in \{0, 1\}$ for any $t \in N^+$, $j = 1, 2, \ldots, K$. Majority vote which is simple and effective scheme is adopted to make a global-final decision $z(t) \in \{0, 1\}$ over the vector $z(t)$, where “1” represents the presence of target and “0” represents the absence of target.
Table 2: $\alpha_{\text{blind}}$ at $(M, \eta_{\text{opt}})$, $M \in \{1, 2, \ldots, 9\}$ and $\eta_{\text{opt}} = \text{ceil}(M/2)$.

| $M/\eta_{\text{opt}}$ | 1/1 | 2/1 | 3/2 | 4/2 | 5/3 | 6/3 | 7/4 | 8/4 | 9/5 |
|---------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| $\alpha_{\text{blind}}$ | 0.5000 | 0.5000 | 0.4310 | 0.4310 | 0.4042 | 0.4042 | 0.3913 | 0.3913 | 0.3844 |

represents the absence of target, respectively. However, we project our focus on designing a rule to construct vector of global decision $z(t)$. We propose a scheme of making global decision that the FC can adaptively adjust its threshold $\eta_{fc}(t,j)$ for $d_{ij}^j$ to make a corresponding global decision $z_j(t)$. The information of elements in decision matrix $D_N^{N \times K}$ is also utilized to estimate probability of miss detection and false alarm, respectively. In order to evaluate the fusion scheme, performance metric introduced in this work is the accuracy of fusionscheme interms of identifying Byzantine sensors.

We define an intuitive distance between the global-final decision and local/original decisions as

$$|d_i^j| = \sum_{i=1}^{N} |d_{ij}(t) - z(t)|.$$  \hspace{1cm} (28)

Similarly, another intuitive distance is also defined as the following equation:

$$|d_i^j| = \sum_{j=1}^{K} |d_{ij}(t) - z(t)|.$$  \hspace{1cm} (29)

In (29), $|d_i^j|$ measures the degree of discrepancy between the global-final decision and $K$ local/original decisions that are from $i$th sensor. Similarly, $|d_i^j|$ in (28) measures the degree of discrepancy between the global-final decision and $N$ local/original decisions generated at $j$th observation. The distance $|d_i^j|$ is larger and the $i$th sensor is closer to behavior of Byzantine. On the contrary, the distance of $|d_i^j|$ is smaller and the probability of $i$th sensor being tagged as Byzantine attacker by the FC is lower. Similarly, when the distance of $|d_i^j|$ is larger, the decision generated at $j$th observation is worse. The FC regards the $j$th observation as excellent and we believe that the network performs well when $|d_i^j|$ is small.

For simplicity, we let probability of miss detection equal probability of false alarm; that is, $P_m = P_{fa}$ in the context of the attacks model of NMBA.

4.1. Scheme of Identifying Byzantine Attackers. If one sensor’s local/original decision is different from global-final decision, it is labeled as Byzantine. It is worth mentioning that Honest sensors are classified probably into Byzantine group when they behave like Byzantine. Similarly, the Byzantine attackers can also be probably labeled as Honest. We use $N_0(t,j)$ and $N_1(t,j)$ to denote the total effective number of elements “0” and “1” in decision matrix, respectively. $N_0(t,j)$ and $N_1(t,j)$ vary with the order of $j$ and take part in making global decision. Here, we represent (28) and (29) again as follows:

$$|d_i^j| = \sum_{i \in I(t,j)} |d_{ij}(t) - z(t)|,$$

$$|d_i^j| = \sum_{k=1}^{K} |d_{jk}(t) - z(t)|, \quad i \in I(t,j),$$

where $I(t,j)$ is defined at (37).
**Proposition 1.** Suppose the network with $N$ sensors including Byzantine and Honest, where the probability of miss detection and false-alarm is zero, that is, $P_m = P_{fa} = 0$, and Byzantine attackers always know the true natural state. Attacking power brought by the attackers is estimated through the following expression:

$$\hat{\alpha}(t) = \frac{\min\{N_0(t), N_1(t)\} - K \cdot \max_2(|d^m_i|) - \min_2(|d^m_j|)}{n(t,j) \cdot K}.$$  \hspace{1cm} (31)

where $N_0(t)$ and $N_1(t)$ are denoted as the total number of "0" and "1," respectively.

**Proof.** See Appendix A. □

Based on Proposition 1 and constrained by its conditions, we have the following equation for estimating the attacking power at the global decision making:

$$\hat{\alpha}(t,j) = \frac{\min\{N_0(t,j), N_1(t,j)\}}{N_0(t,j) + N_1(t,j)}.$$  \hspace{1cm} (32)

Taking the scenario of Byzantine attackers without knowledge about true natural state into consideration in practical application, we have the following.

**Proposition 2.** Suppose the network of size $N$ containing both Honest sensors and Byzantine sensor nodes, where decision matrix $D^{\hat{t}}_{N \times K}$ is made at $t$th global-final decision after $N$ sensors taking $K$ observations in the time window $T$. The network is not completely blinded by adversary. The estimation of least attacking power has a lower bound which can be expressed as

$$\hat{\alpha}(t,j) \geq \frac{\min\{N_0(t,j), N_1(t,j)\} - K \cdot \left[\max_k(|d^m_i|) - \min_k(|d^m_j|)\right]}{n(t,j) \cdot K},$$  \hspace{1cm} (33)

$$k \in \{1,2,\ldots,K\}.$$  

It is clear that $\hat{\alpha}(t,0) = 0$ and $n(t,0) = N$, $t \in N^+$. In (33), $[\max_k(|d^m_i|) - \min_k(|d^m_j|)]$ is used to compute the maximum distance of decisions between two certain vectors of observation which are from two sensors tagged as Honest at the $(j-1)$th global decision making.

**Proof.** See Appendix B. □

Here, we let

$$\hat{\alpha}(t,j) = \frac{\min\{N_0(t,j), N_1(t,j)\} - K \cdot \left[\max_j(|d^m_i|) - \min_j(|d^m_j|)\right]}{n(t,j) \cdot K}$$  \hspace{1cm} (34)

during the process of identifying Byzantine sensors. After estimating the attacking power, $\hat{\alpha}(t,j)$, the FC approximately knows how many sensors have been compromised by an adversary among $n(t,j)$ sensors that participate in the $j$th global decision making. The FC then takes an action to identify $\hat{\alpha}(t,j) \cdot n(t,j)$ Byzantine sensors to help the next global decision making. $L(t,j)$ is defined as the set of identities of sensors that are tagged as Honest at the $(j-1)$th. These sensors will participate in the $j$th global decision making. It is clear that $L(t,0) = \{i \mid i = 1,2,\ldots,N\}$, $t \in N^+$. We define a sequence $I_N(t,j)$ over $|d^m_i| \mid i \in L(t,j-1)$, in which $|d^m_i|$ is used for presenting the element with order $m$, $m \in \{1,2,\ldots,n(t,j)\}$. If $|d^m_i| > |d^m_j|$, then $m > n$. The identity $i$ is greater than $l$ when $m > n$ if $|d^m_i| = |d^m_j|$ or $i < l$ when $m < n$ otherwise, $i,l \in L(t,j-1)$, $m,n \in \{1,2,\ldots,n(t,j)\}$. The sequence can be expressed as

$$I_N(t,j) = \left\{|d^m_i| \mid i \in L(t,j-1)\right\},$$  \hspace{1cm} (35)

where $\forall m \in \{1,2,\ldots,n(t,j)\}, |d^m_i|$ matches only a specific identity. There are $n(t,j)$ identities and $|I_N(t,j)| = n(t,j)$. In order to find the sequence of $d^m_i$, we define a function $ID(\cdot)$ over $I_N(t,j)$:

$$n = ID(d^m_i) = \sum_{t \in L(t,j-1)} \text{IF}(|d^m_i| > |d^m_j|)$$  \hspace{1cm} (36)

$$+ \sum_{t \in L(t,j-1)} \text{IF}(|d^m_i| = |d^m_j| \mid l > j),$$

$$n \in \{1,2,\ldots,n(t,j)\},$$

where $\text{IF}(\cdot)$ is an indicator function, and $\text{IF}(true) = 1$ and $\text{IF}(false) = 0$.

The FC has obtained the approximate knowledge about the attacking power posed by the adversary at $j$th global decision making. The FC has also known each sensor’s sequence in $I_N(t,j)$. Therefore, sensors whose sequences are dropped into the region from $n(t,j)$ to $[n(t,j) \cdot [1 - \hat{\alpha}(t,j)] + 1]$ are judged to be Byzantine attackers. The decisions from Byzantine attackers are removed at the $(j+1)$th global decision making. The identities of remaining sensors which are trusted by the FC at $j$th global decision making can be expressed as

$$I(t,j) = \{i \mid \text{ID}(d^m_i) < \hat{\alpha}(t,j) \cdot n(t,j), \ i \in L(t,j-1)\}$$  \hspace{1cm} (37)

and $I(t,j) \subseteq L(t,j-1)$. The attacking power at $t$th global-final decision making is estimated by the following equation:

$$\hat{\alpha}(t) = \frac{N - n(t,K)}{N}.$$  \hspace{1cm} (38)

### 4.2. Rule of Decision Fusion Based on Dynamic Threshold

As we described in Section 2, the process of making a global-final decision is divided into three stages. A local/original decision matrix $D^{\hat{t}}_{N \times K}$ is generated at the first stage. A decision matrix $D^{\hat{t}}_{N \times K}$ is formulated after $D^{\hat{t}}_{N \times K}$ being probably attacked. A vector of global decision $z(t) = (z_1(t), z_2(t), \ldots, z_K(t))$ is computed and obtained over vectors of decision $[d^m_i, d^m_j, \ldots, d^m_K]$ through applying a policy of fusion
at the middle stage. At last, the decider at the FC has a global-final decision \( z(t) \). In this paper, we put our focus on proposing an effective method of decision fusion to get an excellent vector of global decision \( z(t) \), which can easily help to make a global-final decision \( z(t) \) at the last stage. The rule of decision fusion based on dynamic threshold is that threshold \( \eta_{hc}(t, j) \) for \( j \)th global decision making varies with \( j \). The number of sensors which participate in \( j \)th global decision making is denoted as \( n(t, j) \). In addition, \( n(t, j) \) always associates with \( n(t, j - 1) \) tightly, which is expressed as

\[
n(t, j) = n(t, j - 1) \cdot [1 - \tilde{\alpha}(t, j - 1)].
\]  

Here, \( q \) out of \( m \) fusion rule in [13] has been used for designing threshold \( \eta_{hc}(t, j) \), which is given by

\[
\eta_{hc}(t, j) = \text{floor} \left( a \cdot n(t, j) + \sqrt{b \cdot n(t, j)}Q^{-1}(1 - \beta) \right),
\]

where floor(\( \cdot \)) is a floor function and \( Q^{-1}(\cdot) \) is the inverse function of \( Q(\cdot) \) which has been introduced in Section 2. \( \beta \) is a predefined constraint of miss detection. \( a \) and \( b \) are given by

\[
a = [1 - \tilde{\alpha}(t, j)] \cdot p_{ij}(1) + \tilde{\alpha}(t, j) \cdot x_{ij}(1),
\]

\[
b = [1 - \tilde{\alpha}(t, j)] \cdot p_{ij}(1) \cdot (1 - p_{ij}(1)) + \tilde{\alpha}(t, j) \cdot y_{ij}(1) \cdot (1 - y_{ij}(1)),
\]

respectively. Therefore, we get the \( j \)th global decision \( z_j(t) \):

\[
z_j(t) = \text{IF} \left( \sum_i d_{ij}(t) > \eta_{hc}(t, j) \mid i \in L(t, j - 1) \right).
\]

4.3. Numerical Results. In this subsection, we analyze the performance of scheme of identifying Byzantine attackers through numerical results. There are also 100 sensor nodes that contain Byzantine and Honest nodes which are considered in our simulation. Meanwhile, global-final decision making is performed 100 times in our simulation. The power at the reference point \( (d_0 = 1) \) is set as 200 and the signal amplitude arriving at the local sensor has been contaminated by AWGN with mean value \( \mu = 0 \) and standard deviation \( \sigma = 3 \). A target intrudes at position \( \theta = (90, 90) \). The true attacking power is fixed to 0.3 and \( \beta = 0.2 \).

In Figures 10, 11, 12, and 13, we show that \( n(t, j) \) decreases following the increase of \( j \) and has stable tendency to a fixed level. The sum of \( N_B(t, j) \) and \( N_H(t, j) \) is always equal to \( K \cdot N(t, j) \) and \( N_H(t, j) \) is greater than \( N_B(t, j) \). \( N_B(t, j) \) and \( N_H(t, j) \) verge to \( K \cdot N_H \) and \( K \cdot N_B \), respectively. And \( \tilde{\alpha}(t, j) \rightarrow 0 \) when \( j \) exceeds a certain number and \( \tilde{\alpha}(t, j) = 0 \) when \( j \geq 7 \) in the scheme of identifying Byzantine attackers. Threshold for global decision making varies with \( j \) and \( \eta_{hc}(t, j) \) verges to constant value when \( \tilde{\alpha}(t, j) \) is equal to zero. The attacking power is estimated after each global decision making. Part of sensors is judged to be Byzantine from \( n(t, j) \) sensors. The left sensors then participate in next global decision making. With the conducting of global decision making, there are no sensors tagged as Byzantine until \( j \)th reaches a certain number. At the FC, the total number of sensors tagged as Byzantine is computed as \( N - n(t, K) \). We plot \( y_{HH}, y_{HB}, y_{BH}, \) and \( y_{BB} \) versus \( t \) in Figures 14 and 15 under different cases. As illustrated in Figures 14 and 15, \( y_{HH} \) is close to 0.3, which is black line with cross. And \( y_{HH} \) is close to 0.7, which is denoted as blue line with cross. \( y_{HH} \) and \( y_{BB} \) are presented as black line with star and blue line with star in the Figures 14 and 15, respectively. It is clear that few Byzantine sensors are judged to be Honest by this identifying scheme. In other words, the Byzantine attackers are almost found out. We also can find that the scheme performs well at identifying Byzantine attackers.

5. Conclusion

We have divided the process of data fusion into three hierarchies/stages in this paper. In addition, two problems have been put forward at the first and second stage, respectively. At the first stage, we proposed neighborhood malicious Byzantine attacks model. Distributed detection under neighborhood malicious Byzantine attacks has been taken into consideration. We have shown that NMBA is an intelligent strategy adopted by an adversary. The attacking performance of NMBA has also been analyzed. We have also proved that attacking effect of NMBA matches with CMBA's when the attacking power is lower 0.3844 and NMBA always outperforms IMBA in any case. It has been analyzed that data fusion is incapable when the attacking power enters the blinding region and the closed-form expression for blinding region has been derived. At the second stage, a data fusion rule based on dynamic threshold has been put forward and we have proposed an effective way of identifying Byzantine attackers. Consequently, we have shown that most Byzantine attackers are identified through the scheme and significant improvement of performance of this scheme in terms of
(a) \( n(t, j) \) decreases with the increment of \( j \) and converges to a stable level when \( j \geq 6 \)

(b) \( N_1(t, j) \) decreases with \( j \) and converges to a stable level when \( j \geq 6 \)

(c) \( N_0(t, j) \) decreases with \( j \) and converges to a stable level when \( j \geq 6 \)

(d) \( \eta_{fc}(t, j) \) decreases with \( j \) and converges to stable level when \( j \geq 6 \)

(e) \( \alpha(t, j) \) versus \( j \). The attacking power \( \alpha \) estimated at \( j \)th global decision making decreases with \( j \). \( \alpha(t, j) \) converges to 0 when \( j \geq 6 \). It is clear that more and more Byzantine sensors are identified with increase of \( j \)

Figure 10: \( n(t, j) \), \( N_0(t, j) \), \( N_1(t, j) \), \( \alpha(t, j) \), and \( \eta_{fc}(t, j) \) are plotted versus \( j \) when \( N = 100, K = 10, M = 9 \).
The attacking power $\alpha(t, j)$ estimated at $j$th global decision making decreases with $j$. $\alpha(t, j)$ converges to 0 when $j \geq 7$. It is clear that more and more Byzantine sensors are identified with increase of $j$. (e) $\alpha(t, j)$ versus $j$. The attacking power $\alpha$ estimated at $j$th global decision making decreases with $j$. $\alpha(t, j)$ converges to 0 when $j \geq 7$. It is clear that more and more Byzantine sensors are identified with increase of $j$.
(a) $n(t, j)$ decreases with the increment of $j$ and converges to a stable level when $j \geq 6$

(b) $N_i(t, j)$ decreases with $j$ and converges to a stable level when $j \geq 6$

(c) $N_0(t, j)$ decreases with $j$ and converges to a stable level when $j \geq 6$

(d) $\eta_{fc}(t, j)$ decreases with $j$ and converges to stable level when $j \geq 6$

(e) $\alpha(t, j)$ versus $j$. The attacking power $\alpha$ estimated at $j$th global decision making decreases with $j$. $\alpha(t, j)$ converges to 0 when $j \geq 6$. It is clear that more and more Byzantine sensors are identified with increase of $j$.

Figure 12: $n(t, j)$, $N_0(t, j)$, $N_1(t, j)$, $\alpha(t, j)$, and $\eta_{fc}(t, j)$ are plotted versus $j$ when $N = 100$, $K = 15$, $M = 9$. 
(a) $n(t, j)$ decreases with the increment of $j$ and converges to a stable level when $j \geq 7$

(b) $N_1(t, j)$ decreases with $j$ and converges to a stable level when $j \geq 7$

(c) $N_0(t, j)$ decreases with $j$ and converges to a stable level when $j \geq 7$

(d) $\eta(t, j)$ decreases with $j$ and converges to stable level when $j \geq 7$

(e) $\alpha(t, j)$ versus $j$. The attacking power $\alpha$ estimated at $j$th global decision making decreases with $j$. $\alpha(t, j)$ converges to 0 when $j \geq 7$. It is clear that more and more Byzantine sensors are identified with increase of $j$

Figure 13: $n(t, j), N_0(t, j), N_1(t, j), \alpha(t, j)$, and $\eta(t, j)$ are plotted versus $j$ when $N = 100, K = 15, M = 25$. 
accuracy is obtained. Based on the scheme of identifying Byzantine attackers, a data fusion scheme with dynamic threshold has been explored at this stage.

Appendices

A. Proof for Proposition 1

If Byzantine attackers always know the true hypothesis, each decision is flipped by Byzantine attacker with probability of 1 prior to sending it to the FC [8, 9]. Because of $P_m = P_{ih} = 0$, each Honest sensor can detect the natural state correctly. It is clear that the attacking power is $N_0(t)/N$ when global-final decision is “1.” Similarly, the attacking power is $N_1(t)/N$ when global-final decision is “0.” In other words, $\min\{N_0(t), N_1(t)\}$ is always the number of decisions from Byzantine sensors. Therefore, the attacking power can be estimated through expression

$$\bar{\alpha}(t) = \frac{\min\{N_0(t), N_1(t)\}}{N}. \quad (A.1)$$

B. Proof for Proposition 2

The way that attacking power is estimated through the ratio between the number of decisions that are different from global-final decision to the total number of decisions is adopted. Under the condition that the network has not been blinded completely, several Byzantine sensor nodes may have made mistakes and then behaved like Honest sensors. Similarly, some Honest sensors may have behaved
like Byzantine, for example, situation of $P_m = P_{fa} \neq 0$. Therefore, we have
\[
\min \{N_0(t, j), N_1(t, j)\} \quad \text{n (t, j)} \cdot K
\]
\[
= \alpha(t, j) + \frac{N^{B}_{01}(t, j)}{n(t, j) \cdot K} + \frac{N^{H}_{01}(t, j)}{n(t, j) \cdot K}.
\] (B.1)

In (B.1), $N^{B}_{01}(t, j)$ is denoted as the total number of decisions that come from Byzantine attackers and are identical with the FC’s global-final decision. $N^{H}_{01}(t, j)$ in (B.1) is denoted as the total number of decisions that come from Honest sensors and are different from the FC’s global-final decision. $\alpha(t, j)$ is the ratio of the number of decisions that are from Byzantine attackers which have attacked successfully to the total number of decisions. Equation (B.1) can be represented as
\[
\min \{N_0(t, j), N_1(t, j)\} \quad \text{n (t, j)} \cdot K
\]
\[
= \alpha(t, j) + \frac{N^{B}_{01}(t, j)}{n(t, j) \cdot K} + \alpha_i(1)
\] (B.2)
in the absence of target or
\[
\min \{N_0(t, j), N_1(t, j)\} \quad \text{n (t, j)} \cdot K
\]
\[
= \alpha(t, j) + \frac{N^{B}_{01}(t, j)}{n(t, j) \cdot K} + 1 - \alpha_i(1)
\] (B.3)
in the presence of target, where $\alpha_i(1)$ and $\alpha_i(1)$ are the estimation of probability of detection and false alarm, respectively. Therefore, they are given by
\[
\alpha_i(1) = \frac{N^{H}_{01}(t, j)}{n(t, j) \cdot K}
\] (B.4)
or
\[
\alpha_i(1) = 1 - \frac{N^{H}_{01}(t, j)}{n(t, j) \cdot K}
\] (B.5)
in the perspective of decider at the FC, it is unknown which sensor is Byzantine. However, all the Byzantine attackers possess high efficiency of attacking according to (14) in the context of adopting strategy of N MBA. Therefore, we have
\[
\alpha(t, j) = \min \{N_0(t, j), N_1(t, j)\} \quad \text{n (t, j)} \cdot K
\]
\[
= \min \{N_0(t, j), N_1(t, j)\} \quad \text{n (t, j)} \cdot K
\]
\[
= \alpha_i(1) - \alpha_i(1)
\] (B.6)
Because of
\[
\alpha_i(1) \geq 1 - \frac{\max_k (\{d_k^i\}) - \min_k (\{d_k^i\})}{n(t, j)}
\] (B.7)

\[
\alpha_i(1) \geq 1 - \frac{\max_k (\{d_k^i\}) - \min_k (\{d_k^i\})}{n(t, j)}
\] (B.8)

\[
\alpha_i(1) \geq 1 - \frac{\max_k (\{d_k^i\}) - \min_k (\{d_k^i\})}{n(t, j)}
\] (B.9)

\section*{Conflict of Interests}
The authors declare that there is no conflict of interests regarding the publication of this paper.

\section*{Acknowledgments}
The authors specifically acknowledge the financial support through the National Science Foundation of China (Grant no. 61001086), the Fundamental Research Funds for the Central University (Grant no. ZYGX2011X004), and the project Sponsored by Academic Training Funds, University of Electronic Science and Technology of China (Grant no. 201506075013).

\section*{References}
[1] J.-F. Chamberland and V. V. Veeravalli, “Wireless sensors in distributed detection applications,” IEEE Signal Processing Magazine, vol. 24, no. 3, pp. 16–25, 2007.
[2] L. Lamport, R. Shostak, and M. Pease, “The byzantine generals problem,” ACM Transactions on Programming Languages and Systems, vol. 4, no. 3, pp. 382–401, 1982.
[3] H. Chen, V. P. Jilkov, and X. R. Li, “Optimizing decision fusion in the presence of byzantine data,” in Proceedings of the 17th International Conference on Information Fusion, pp. 1–8, July 2014.
[4] P. Stoica and A. Nehorai, “Performance study of conditional and unconditional direction-of-arrival estimation,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 38, no. 10, pp. 1783–1795, 1990.
[5] D. Liu, K. Liu, Y. Ma, and J. Yu, “Joint TOA and DOA localization in indoor environment using virtual stations,” IEEE Communications Letters, vol. 18, no. 8, pp. 1291–1294, 2014.
[6] S. Althunibat and F. Granelli, “An objection-based collaborative spectrum sensing for cognitive radio networks,” IEEE Communications Letters, vol. 18, no. 8, pp. 1291–1294, 2014.
[7] X. Lu, P. Wang, D. Niyato, and E. Hossain, “Dynamic spectrum access in cognitive radio networks with RF energy harvesting,” IEEE Wireless Communications, vol. 21, no. 3, pp. 102–110, 2014.
[8] S. Marano, V. Matta, and L. Tong, “Distributed detection in the presence of Byzantine attacks,” IEEE Transactions on Signal Processing, vol. 57, no. 1, pp. 16–29, 2009.
[9] S. Marano, V. Matta, and L. Tong, “Distributed inference in the presence of byzantine sensors,” in Proceedings of the 46th Asilomar Conference on Signals, Systems, and Computers (ACSSC’06), pp. 281–284, Pacific Grove, Calif, USA, November 2006.
[10] A. Vempaty, O. Ozdemir, K. Agrawal, H. Chen, and P. K. Varshney, “Localization in wireless sensor networks: byzantines and mitigation techniques,” *IEEE Transactions on Signal Processing*, vol. 61, no. 6, pp. 1495–1508, 2013.

[11] A. S. Rawat, P. Anand, H. Chen, and P. K. Varshney, “Collaborative spectrum sensing in the presence of byzantine attacks in cognitive radio networks,” *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 774–786, 2011.

[12] X. F. He, H. Y. Dai, and P. Ning, “A Byzantine attack defender: the conditional frequency check,” in *Proceedings of the IEEE International Symposium on Information Theory (ISIT ’12)*, pp. 975–979, IEEE, Cambridge, Mass, USA, July 2012.

[13] M. Abdelhakim, L. E. Lightfoot, J. Ren, and T. Li, “Distributed detection in mobile access wireless sensor networks under byzantine attacks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 4, pp. 950–959, 2013.

[14] E. Soltanmohammadi, M. Orooji, and M. Naraghi-Pour, “Decentralized hypothesis testing in wireless sensor networks in the presence of misbehaving nodes,” *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 1, pp. 205–215, 2012.

[15] B. Kailkhura, S. Brahma, Y. S. Han, and P. K. Varshney, “Distributed detection in tree topologies with byzantines,” *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3208–3219, 2014.

[16] S. Talarico, N. A. Schmid, M. Alkhweldi, and M. C. Valenti, “Distributed estimation of a parametric field: algorithms and performance analysis,” *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1041–1053, 2014.

[17] P. Zhang, J. Y. Koh, S. Lin, and I. Nevat, “Distributed event detection under byzantine attack in wireless sensor networks,” in *Proceedings of the 9th IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP ’14)*, pp. 1–6, IEEE, Singapore, April 2014.

[18] Q. Liu, X. Wang, and N. S. V. Rao, “Fusion of state estimates over long-haul sensor networks with random loss and delay,” *IEEE/ACM Transactions on Networking*, vol. 23, no. 2, pp. 644–656, 2015.

[19] M. H. El-Ayadi, “Nonstochastic adaptive decision fusion in distributed-detection systems,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 4, pp. 1158–1171, 2002.

[20] R. Laxhammar and G. Falkman, “Online learning and sequential anomaly detection in trajectories,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 6, pp. 1158–1173, 2013.

[21] Y. Chen, J. Yang, W. Trappe, and R. P. Martin, “Detecting and localizing identity-based attacks in wireless and sensor networks,” *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2418–2434, 2010.

[22] R. Niu and P. K. Varshney, “Target location estimation in sensor networks with quantized data,” *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4519–4528, 2006.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.

[24] S. Kullback, *Information Theory and Statistics*, Dover, New York, NY, USA, 1997.