Distributed Bayesian Detection in the Presence of Byzantine Data

Bhavya Kailkhura, *Student Member, IEEE*, Yunghsiang S. Han, *Fellow, IEEE*, Swastik Brahma, *Member, IEEE*, Pramod K. Varshney, *Fellow, IEEE*

***Abstract*—In this paper, we consider the problem of distributed Bayesian detection in the presence of Byzantines in the network. It is assumed that a fraction of the nodes in the network are compromised and reprogrammed by an adversary to transmit false information to the fusion center (FC) to degrade detection performance. The problem of distributed detection is formulated as a binary hypothesis test at the FC based on 1-bit data sent by the sensors. The expression for minimum attacking power required by the Byzantines to blind the FC is obtained. More specifically, we show that above a certain fraction of Byzantine attackers in the network, the detection scheme becomes completely incapable of utilizing the sensor data for detection. We analyze the problem under different attacking scenarios and derive results for different non-asymptotic cases. It is found that existing asymptotics-based results do not hold under several non-asymptotic scenarios. When the fraction of Byzantines is not sufficient to blind the FC, we also provide closed form expressions for the optimal attacking strategies for the Byzantines that most degrade the detection performance.**

***Index Terms*—Bayesian detection, Data falsification, Byzantine Data, Probability of error, Distributed detection**

I. Introduction

Distributed detection is a well studied topic in the detection theory literature [1]–[3]. In distributed detection systems, due to bandwidth and energy constraints, the nodes often make a 1-bit local decision regarding the presence of a phenomenon before sending it to the fusion center (FC). Based on the local decisions transmitted by the nodes, the FC makes a global decision about the presence of the phenomenon of interest. Distributed detection was originally motivated by its applications in military surveillance but is now being employed in a wide variety of applications such as distributed spectrum sensing (DSS) using cognitive radio networks (CRNs) and traffic and environment monitoring.

In many applications, a large number of inexpensive and less reliable nodes that can provide dense coverage are used to provide a balance between cost and functionality. The performance of such systems strongly depends on the reliability of the nodes in the network. The robustness of distributed detection systems against attacks is of utmost importance. The distributed nature of such systems makes them quite vulnerable to different types of attacks. In recent years, security issues of such distributed networks are increasingly being studied within the networking [4], signal processing [5] and information theory communities [6]. One typical attack on such networks is a Byzantine attack. While Byzantine attacks (originally proposed by [7]) may, in general, refer to many types of malicious behavior, our focus in this paper is on data-falsification attacks [8]–[15]. In this type of attack, an attacker may send false (erroneous) data to the FC to degrade detection performance. In this paper, we refer to such a data falsification attacker as a Byzantine and the data thus generated is referred to as Byzantine data.

We formulate the signal detection problem as a binary hypothesis testing problem with the two hypotheses H_0 (signal is absent) and H_1 (signal is present). We make the conditional i.i.d. assumption under which observations at the nodes are conditionally independent and identically distributed given the hypothesis. We assume that the FC is not compromised, and is able to collect data from all the nodes in the network via error free communication channels.[1] We also assume that the FC does not know which node is Byzantine, but it knows the fraction of Byzantines in the network.[2] We consider the problem of distributed Bayesian detection with prior probabilities of hypotheses known to both the FC and the attacker. The FC aims to minimize the probability of error by choosing the optimal fusion rule.

A. Related Work

Although distributed detection has been a very active field of research in the past, security problems in distributed detection networks gained attention only very recently. In [11], the authors considered the problem of distributed detection in the presence of Byzantines under the Neyman-Pearson (NP) setup and determined the optimal attacking strategy which minimizes the detection error exponent. This approach based on Kullback-Leibler divergence (KLD) is analytically tractable and yields approximate results in non-asymptotic cases. They also assumed that the Byzantines know the true hypothesis, which obviously is not satisfied in practice but does provide a bound. In [12], the authors analyzed the same problem in the context of collaborative spectrum sensing under Byzantine Attacks. They relaxed the assumption of perfect knowledge of

B. Kailkhura, S. Brahma and P. K. Varshney are with Department of EECS, Syracuse University, Syracuse, NY 13244. (email: bkailkhu@syr.edu; skbrahma@syr.edu; varshney@syr.edu)

Y. S. Han is with EE Department, National Taiwan University of Science and Technology, Taiwan, R. O. C. (email: yshan@mail.ntust.edu.tw)

[1] In this work, we do not consider how individual nodes deliver their data to the fusion center except that the Byzantines are not able to alter the transmissions of honest nodes.

[2] In practice, the fraction of Byzantines in the network can be learned by observing the data sent by the nodes at the FC over a time window; however, this study is beyond the scope of this work.

TABLE I
DIFFERENT SCENARIOS BASED ON THE KNOWLEDGE OF THE OPPONENT'S STRATEGIES

Cases	Attacker has the knowledge of the FC's strategies	FC has the knowledge of Attacker's strategies
Case 1	No	No
Case 2	Yes	No
Case 3	Yes	Yes
Case 4	No	Yes

the hypotheses by assuming that the Byzantines determine the knowledge about the true hypotheses from their own sensing observations. A variant of the above formulation was explored in [13], [16], where the authors addressed the problem of optimal Byzantine attacks (data falsification) on distributed detection for a tree-based topology and extended the results of [12] for tree topologies. By assuming that the cost of compromising nodes at different levels of the tree is different, they found the optimal Byzantine strategy that minimizes the cost of attacking a given tree. Schemes for Byzantine node identification have been proposed in [12], [15], [17], [18]. Our focus is considerably different from Byzantine node identification schemes in that we do not try to authenticate the data; we consider most effective attacking strategies and distributed detection schemes that are robust against attacks. Note that, the Byzantine attack model is similar to the scenario where local decisions are transmitted over a Binary Symmetric Channel (BSC) with a certain cross over probability. There are several papers that address the impact of transmission channels or faults on the distributed detection system and related problems [19]–[22]. However, Byzantine attacks are philosophically different from the BSC or the faulty sensor case. Byzantine attacks are intentional and, therefore, the attacker can optimize over the attack parameters. Thus, in contrast to channel aware detection, both the FC and the Byzantines can optimize their utility by choosing their actions based on the knowledge of their opponent's behavior. Study of these practically motivated scenarios in the presence of Byzantines is missing from the channel aware detection and fault tolerant detection literature because of the philosophical difference between these approaches.

B. Main Contributions

All the approaches discussed so far consider distributed detection under the Neyman-Pearson (NP) setup. In this paper, we consider the distributed Bayesian detection problems with known prior probabilities of hypotheses. We assume that the Byzantines do not have perfect knowledge about the true state of the phenomenon of interest. In addition, we also assume that the Byzantines neither have the knowledge nor control over the thresholds used to make local decisions at the nodes. Also, the probability of detection and the probability of false alarm of a node are assumed to be the same for every node irrespective of whether they are honest or Byzantines. In this paper, we focus on a *non-asymptotic* analysis for the Byzantine attacks on distributed Bayesian detection. First, we show that above a certain fraction of Byzantines in the network, the data fusion scheme becomes completely incapable (blind) and it is not possible to design a decision rule at the FC that can perform better than the decision rule based just on prior information. We find the minimum fraction of Byzantines that can blind the FC and refer to it as the *critical power*. Next, we explore the optimal attacking strategies for the Byzantines under different scenarios. In practice, the FC and the Byzantines will optimize their utility by choosing their actions based on the knowledge of their opponent's behavior. This motivates us to address the question: what are the optimal attacking/defense strategies given the knowledge of the opponent's strategies? Study of these practically motivated questions requires non asymptotic analysis, which is systematically studied in this work. By assuming the error probability to be our performance metric, we analyze the problem in the non asymptotic regime. Observe that, the probability of error is a function of the fusion rule, which is under the control of the FC. This gives us an additional degree of freedom to analyze the Byzantine attack under different practical scenarios where the FC and the Byzantines may or may not have knowledge of their opponent's strategies (For a description of different scenarios see Table I). It is found that results based on asymptotics do not hold under several non-asymptotic scenarios. More specifically, when the FC does not have knowledge of attacker's strategies, results for the non-asymptotic case are different from those for the asymptotic case. However, if the FC has complete knowledge of the attacker's strategies and uses the optimal fusion rule to make the global decision, results obtained for this case are the same as those for the asymptotic case. Knowledge of the behavior of the attacker in the non-asymptotic regime enables the analysis of many related questions, such as the design of the optimal detector (fusion rule) and effects of strategic interaction between the FC and the attacker. In the process of analyzing the scenario where the FC has complete knowledge of its opponent's strategies, we obtain a closed form expression of the optimal fusion rule. To summarize, our main contributions are threefold.

- In contrast to previous works, we carry out the non-asymptotic performance analysis of distributed Bayesian detection with Byzantines.
- We analyze the problem under different attacking scenarios and derive closed form expressions for optimal attacking strategies for different non-asymptotic cases.
- In the process of analyzing the scenario where the FC has complete knowledge of its opponent's strategies, we obtain a closed form expression for the optimal fusion rule.

The signal processing problem considered in this paper is closest to [12]. The approach in [12], based on Kullback-Leibler divergence (KLD), is analytically tractable and yields approximate results in non-asymptotic cases. Our results, however, are not a direct application of those of [12]. While as in [12] we are also interested in the optimal attack strategies, our objective function and, therefore, techniques of finding them are different. In contrast to [12], where only optimal strategies to blind the FC were obtained, we also provide closed form expressions for the optimal attacking strategies for

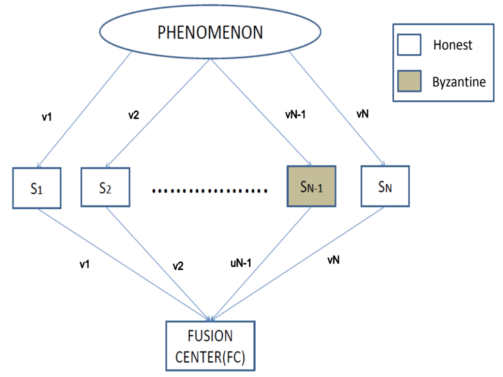


Fig. 1. System Model

the Byzantines that most degrade the detection performance when the fraction of Byzantines is not sufficient to blind the FC. In fact, finding the optimal Byzantine attacking strategies is only the first step toward designing a robust distributed detection system. Knowledge of these attacking strategies can be used to implement the optimal detector at the FC or to implement an efficient reputation based identification scheme [12], [23] (thresholds in these schemes are generally a function of attack strategies). Also, the optimal attacking distributions in certain cases have the minimax property and, therefore, the knowledge of these optimal attack strategies can be used to implement the robust detector.

The rest of the paper is organized as follows. Section II introduces our system model, including the Byzantine attack model. In Section III, we provide the closed form expression for the critical power above which the FC becomes blind. Next, we discuss our results based on non-asymptotic analysis of the distributed Bayesian detection system with Byzantine data for different scenarios. In Section IV, we analyze the problem when Byzantines do not have any knowledge about the fusion rule used at the FC. Section V discusses the scenario where Byzantines have the knowledge about the fusion rule used at the FC, but the FC does not know the attacker's strategies. Next in Section VI, we extend our analysis to the scenario where both the FC and the attacker have the knowledge of their opponent's strategies and act strategically to optimize their utilities. Finally, Section VII concludes the paper.

II. Distributed Detection in the Presence of Byzantines

Consider two hypotheses H_0 (signal is absent) and H_1 (signal is present). Also, consider a parallel network (see Figure 1), comprised of a central entity (known as the Fusion Center (FC)) and a set of N sensors (nodes), which faces the task of determining which of the two hypotheses is true. Prior probabilities of the two hypotheses H_0 and H_1 are denoted by P_0 and P_1, respectively. The sensors observe the phenomenon, carry out local computations to decide the presence or absence of the phenomenon, and then send their local decisions to the FC that yields a final decision after processing the local decisions. Observations at the nodes are assumed to be conditionally independent and identically distributed given the hypothesis. A Byzantine attack on such a system compromises some of the nodes which may then intentionally send falsified local decisions to the FC to make the final decision incorrect. We assume that a fraction α of the N nodes which observe the phenomenon have been compromised by an attacker. In this paper, we consider the Clairvoyant case and assume that, the fusion center knows the values of α and $(P_0,\ P_1)$. In appendix A we discuss some practical issues related to our assumptions regarding the knowledge of these parameters in the detection system. We consider the communication channels to be error-free. Next, we describe the modus operandi of the sensors and the FC in detail.

A. Modus Operandi of the Nodes

Based on the observations, each node i makes a one-bit local decision $v_i \in \{0, 1\}$ regarding the absence or presence of the phenomenon using the likelihood ratio test

$$\frac{p_{Y_i}^{(1)}(y_i)}{p_{Y_i}^{(0)}(y_i)} \overset{v_i=1}{\underset{v_i=0}{\gtrless}} \lambda, \tag{1}$$

where λ is the identical threshold[3] used at all the sensors and $p_{Y_i}^{(k)}(y_i)$ is the conditional probability density function (PDF) of observation y_i under the hypothesis H_k. Each node i, after making its one-bit local decision v_i, sends $u_i \in \{0, 1\}$ to the FC, where $u_i = v_i$ if i is an uncompromised (honest) node, but for a compromised (Byzantine) node i, u_i need not be equal to v_i. We denote the probabilities of detection and false alarm of each node i in the network by $P_d = P(v_i = 1|H_1)$ and $P_f = P(v_i = 1|H_0)$, respectively, which hold for both uncompromised nodes as well as compromised nodes. In this paper, we assume that each Byzantine decides to attack independently relying on its own observation and decision regarding the presence of the phenomenon. Specifically, we define the following strategies $P_{j,1}^H$, $P_{j,0}^H$ and $P_{j,1}^B$, $P_{j,0}^B$ ($j \in \{0, 1\}$) for the honest and Byzantine nodes, respectively:

Honest nodes:

$$P_{1,1}^H = 1 - P_{0,1}^H = P^H(x = 1|y = 1) = 1 \tag{2}$$

$$P_{1,0}^H = 1 - P_{0,0}^H = P^H(x = 1|y = 0) = 0 \tag{3}$$

Byzantine nodes:

$$P_{1,1}^B = 1 - P_{0,1}^B = P^B(x = 1|y = 1) \tag{4}$$

$$P_{1,0}^B = 1 - P_{0,0}^B = P^B(x = 1|y = 0) \tag{5}$$

$P^H(x = a|y = b)$ ($P^B(x = a|y = b)$) is the probability that an honest (Byzantine) node sends a to the FC when its actual local decision is b. From now onwards, we will refer to Byzantine flipping probabilities simply by $(P_{1,0}, P_{0,1})$. We also assume that the FC is not aware of the exact set of Byzantine nodes and considers each node i to be Byzantine with a certain probability α.

[3] It has been shown that the use of identical thresholds is asymptotically optimal [24].

B. Binary Hypothesis Testing at the Fusion Center

We consider a Bayesian detection problem where the performance criterion at the FC is the probability of error. The FC receives decision vector, $\mathbf{u} = [u_1, \cdots, u_N]$, from the nodes and makes the global decision about the phenomenon by considering the maximum a *posteriori* probability (MAP) rule which is given by

$$P(H_1|\mathbf{u}) \underset{H_0}{\overset{H_1}{\gtrless}} P(H_0|\mathbf{u})$$

or equivalently,

$$\frac{P(\mathbf{u}|H_1)}{P(\mathbf{u}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1}.$$

Since the u_is are independent of each other, the MAP rule simplifies to a K-out-of-N fusion rule [1]. The global false alarm probability Q_F and detection probability Q_D are then given by[4]

$$Q_F = \sum_{i=K}^{N} \left(\begin{array}{c} N \\ i \end{array} \right) (\pi_{1,0})^i (1-\pi_{1,0})^{N-i} \tag{6}$$

and

$$Q_D = \sum_{i=K}^{N} \left(\begin{array}{c} N \\ i \end{array} \right) (\pi_{1,1})^i (1-\pi_{1,1})^{N-i}, \tag{7}$$

where π_{j0} and π_{j1} are the conditional probabilities of $u_i = j$ given H_0 and H_1, respectively. Specifically, $\pi_{1,0}$ and $\pi_{1,1}$ can be calculated as

$$\pi_{1,0} = \alpha(P_{1,0}(1-P_f) + (1-P_{0,1})P_f) + (1-\alpha)P_f \tag{8}$$

and

$$\pi_{1,1} = \alpha(P_{1,0}(1-P_d) + (1-P_{0,1})P_d) + (1-\alpha)P_d, \tag{9}$$

where α is the fraction of Byzantine nodes.[5]

The local probability of error as seen by the FC is defined as

$$P_e = P_0\pi_{1,0} + P_1\left(1-\pi_{1,1}\right) \tag{10}$$

and the system wide probability of error at the FC is given by

$$P_E = P_0 Q_F + P_1\left(1-Q_D\right). \tag{11}$$

In our earlier work [25] on this problem, we analyzed the problem in the asymptotic regime. Adopting Chernoff information as our performance metric, we studied the performance of a distributed detection system with Byzantines in the asymptotic regime. We summarize our results in the following theorem.

Theorem 1 ([25]). *Optimal attacking strategies, $(P_{1,0}^*, P_{0,1}^*)$, which minimize the Chernoff information are*

$$(P_{1,0}^*, P_{0,1}^*) \begin{cases} (p_{1,0}, p_{0,1}) & \text{if } \alpha \geq 0.5 \\ (1,1) & \text{if } \alpha < 0.5 \end{cases},$$

where, $(p_{1,0}, p_{0,1})$ satisfy $\alpha(p_{1,0} + p_{0,1}) = 1$.

Notice that, the system wide probability of error P_E is a function of the parameter K, which is under the control of the FC, and the parameters $(\alpha, P_{j,0}, P_{j,1})$ are under the control of the attacker. The FC and the Byzantines may or may not have knowledge of their opponent's strategy. In this paper, we will analyze the problem of detection with Byzantine data under several different scenarios in the following sections. First, we will determine the minimum fraction of Byzantines needed to blind the decision fusion scheme.

III. Critical Power to Blind the fusion Center

In this section, we determine the minimum fraction of Byzantine nodes needed to make the FC "blind" and denote it by α_{blind}. We say that the FC is blind if an adversary can make the data that the FC receives from the sensors such that no information is conveyed. In other words, the optimal detector at the FC cannot perform better than simply making the decision based on priors.

Lemma 1. *In Bayesian distributed detection, the minimum fraction of Byzantines needed to make the FC blind is $\alpha_{blind} =$ 0.5.*

Proof: In the Bayesian framework, we say that the FC is "blind", if the received data $\mathbf{u}$ does not provide any information about the hypotheses to the FC. That is, the condition to make the FC blind can be stated as

$$P(H_i|\mathbf{u}) = P(H_i) \text{ for } i = 0, 1. \tag{12}$$

Applying Bayes' theorem, it can be seen that (12) is equivalent to $P(\mathbf{u}|H_i) = P(\mathbf{u})$. Thus, the FC becomes blind if the probability of receiving a given vector $\mathbf{u}$ is independent of the hypothesis present. In such a scenario, the best that the FC can do is to make decisions solely based on the priors, resulting in the most degraded performance at the FC. Now, using the conditional i.i.d. assumption, under which observations at the nodes are conditionally independent and identically distributed given the hypothesis, condition (12) to make the FC blind becomes $\pi_{1,1} = \pi_{1,0}$. This is true only when

$$\alpha[P_{1,0}(P_f - P_d) + (1-P_{0,1})(P_d - P_f)] + (1-\alpha)(P_d - P_f) = 0.$$

Hence, the FC becomes blind if

$$\alpha = \frac{1}{(P_{1,0} + P_{0,1})}. \tag{13}$$

α in (13) is minimized when $P_{1,0}$ and $P_{0,1}$ both take their largest values, i.e., $P_{1,0} = P_{0,1} = 1$. Hence, $\alpha_{blind} = 0.5$. ■

Next, we investigate how the Byzantines can launch an attack optimally considering that the parameter (K) is under the control of the FC. By assuming error probability to be our performance metric, we analyze the non-asymptotic regime. Observe that the probability of error is dependent on the fusion

[4] These expressions are valid under the assumption that $\alpha < 0.5$. Later in Section VI, we will generalize our result for any arbitrary α.

[5] The proposed analysis can be easily extended to the noisy channel case. For example, let us consider the Binary Symmetry Channel with crossover probabilities given by $(\hat{P_{1,0}}, \hat{P_{0,1}})$. Now, the conditional probability $\pi_{1,1}$ as given in (9) changes to:

$$\begin{aligned}\pi_{1,1} &= \alpha(1-\hat{P_{0,1}})[(1-P_{0,1})P_d + P_{1,0}(1-P_d)] + \alpha\hat{P_{1,0}}[P_{0,1}P_d + (1-P_{1,0})(1-P_d)] \\ &\quad +(1-\alpha)[(1-\hat{P_{0,1}})P_d + 1\hat{,}0(1-Pd)]\end{aligned}$$

and similarly the expression for $\pi_{1,0}$ can be obtained. Using theses expressions the proposed analysis can be extended to the noisy case (BSC).

TABLE II
SOULTION OF MAXIMIZING LOCAL ERROR P_e PROBLEM

$P_{1,0}$	$P_{0,1}$	Condition
0	0	$\frac{P_d}{P_f} < \frac{P_0}{P_1} < \frac{1-P_d}{1-P_f}$
0	1	$\frac{P_d}{P_f} > \frac{P_0}{P_1} < \frac{1-P_d}{1-P_f}$
1	0	$\frac{P_d}{P_f} < \frac{P_0}{P_1} > \frac{1-P_d}{1-P_f}$
1	1	$\frac{P_d}{P_f} > \frac{P_0}{P_1} > \frac{1-P_d}{1-P_f}$

rule. This gives us an additional degree of freedom to analyze the Byzantine attack under different scenarios where the FC and the Byzantines may or may not have knowledge of their opponent's strategies.

IV. OPTIMAL ATTACKING STRATEGIES WITHOUT THE KNOWLEDGE OF FUSION RULE

In practice, the Byzantine attacker may not have the knowledge about the fusion rule, i.e., the value of K, used by the FC. In such scenarios, we obtain the optimal attacking strategy for Byzantines by maximizing the local probability of error as seen by the FC, which is independent of the fusion rule K. We formally state the problem as

$$\begin{aligned} \underset{P_{1,0},P_{0,1}}{\text{maximize}} \quad & P_0\pi_{1,0} + P_1(1-\pi_{1,1}) \\ \text{subject to} \quad & 0 \le P_{1,0} \le 1 \\ & 0 \le P_{0,1} \le 1 \end{aligned} \tag{P1}$$

To solve the problem, we analyze the properties of the objective function, $P_e = P_0\pi_{1,0} + P_1(1-\pi_{1,1})$, with respect to $(P_{1,0}, P_{0,1})$. Notice that

$$\frac{dP_e}{P_{1,0}} = P_0\alpha(1-P_f) - P_1\alpha(1-P_d) \tag{14}$$

and

$$\frac{dP_e}{P_{0,1}} = -P_0\alpha P_f + P_1\alpha P_d. \tag{15}$$

By utilizing monotonicity properties of the objective function with respect to $P_{1,0}$ and $P_{0,1}$ ((14) and (15)), we present the solution of the Problem P1 in Table II. Notice that, when $\frac{P_d}{P_f} < \frac{P_0}{P_1} < \frac{1-P_d}{1-P_f}$, both (14) and (15) are less than zero. P_e then becomes a strictly decreasing function of $P_{1,0}$ as well as $P_{0,1}$. Hence, to maximize P_e, the attacker needs to choose $(P_{1,0}, P_{0,1}) = (0,0)$. However, the condition $\frac{P_d}{P_f} < \frac{P_0}{P_1} < \frac{1-P_d}{1-P_f}$ holds iff $P_d < P_f$ and, therefore, is not admissible. Similar arguments lead to the rest of results given in Table II. Note that, if there is an equality in the conditions mentioned in Table II, then the solution will not be unique. For example, $\left(\dfrac{dP_e}{P_{0,1}} = 0\right) \Leftrightarrow \left(\dfrac{P_0}{P_1} = \dfrac{1-P_d}{1-P_f}\right)$ implies that the P_e is constant as a function of $P_{0,1}$. In other words, the attacker will be indifferent in choosing the parameter $P_{0,1}$ because any value of $P_{0,1}$ will result in the same probability of error.

Next, to gain insight into the solution, we present illustrative examples that corroborate our results.

A. Illustrative Examples

In Figure 2(a), we plot the local probability of error P_e as a function of $(P_{1,0}, P_{0,1})$ when $(P_0 = P_1 = 0.5)$. We assume that the local probability of detection is $P_d = 0.8$ and the local probability of false alarm is $P_f = 0.1$ such that $\frac{P_d}{P_f} = 8$, $\frac{1-P_d}{1-P_f} = .2222$, and $\frac{P_0}{P_1} = 1$. Clearly, $\frac{P_d}{P_f} > \frac{P_0}{P_1} > \frac{1-P_d}{1-P_f}$ and it implies that the optimal attacking strategy is $(P_{1,0}, P_{0,1}) = (1,1)$, which can be verified from Figure 2(a).

In Figure 2(b), we study the local probability of error P_e as a function of the attacking strategy $(P_{1,0}, P_{0,1})$ when $(P_0 = 0.1, P_1 = 0.9)$. We assume that the local probability of detection is $P_d = 0.8$ and the local probability of false alarm is $P_f = 0.1$ such that $\frac{P_d}{P_f} = 8$, $\frac{1-P_d}{1-P_f} = .2222$, and $\frac{P_0}{P_1} = .1111$. Clearly, $\frac{P_d}{P_f} > \frac{P_0}{P_1} < \frac{1-P_d}{1-P_f}$ implies that the optimal attacking strategy is $(P_{1,0}, P_{0,1}) = (0,1)$, which can be verified from the Figure 2(b). These results corroborate our theoretical results presented in Table II.

In the next section, we investigate the scenario where Byzantines are aware of the fusion rule K used at the FC and can use this knowledge to provide false information in an optimal manner to blind the FC. However, the FC does not have knowledge of Byzantine's attacking strategies $(\alpha, P_{j,0}, P_{j,1})$ and does not optimize against Byzantine's behavior. Since majority rule is a widely used fusion rule [14], [26], [27], we assume that the FC uses the majority rule to make the global decision.

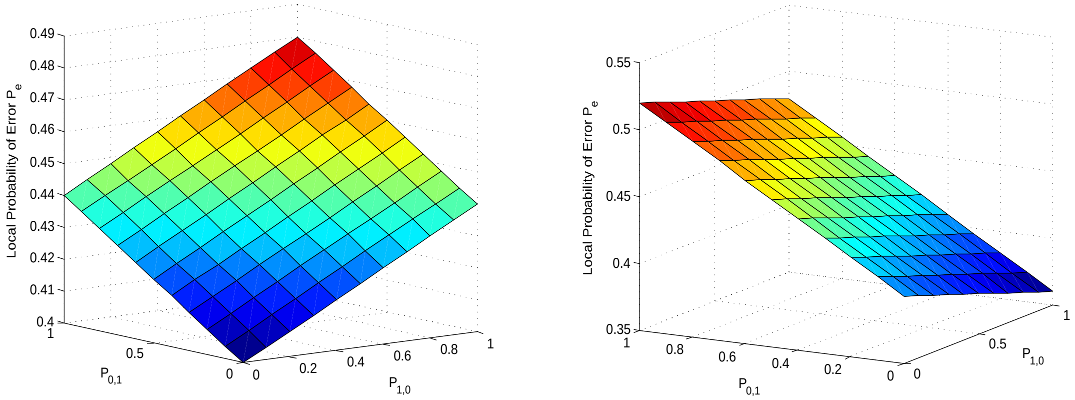
V. OPTIMAL BYZANTINE ATTACKING STRATEGIES WITH KNOWLEDGE OF MAJORITY FUSION RULE

In this section, we investigate optimal Byzantine attacking strategies in a distributed detection system, with the attacker having knowledge about the fusion rule used at the FC. However, we assume that the FC is not strategic in nature, and uses a majority rule, without trying to optimize against the Byzantine's behavior. We consider both the FC and the Byzantine to be strategic in Section VI. The performance criterion at the FC is assumed to be the probability of error P_E.

For a fixed fusion rule (K^*), which, as mentioned before, is assumed to be the majority rule $K^* = \lceil \frac{N+1}{2} \rceil$, P_E varies with the parameters $(\alpha, P_{j,0}, P_{j,1})$ which are under the control of the attacker. The Byzantine attack problem can be formally stated as follows:

$$\begin{aligned} \underset{P_{j,0},P_{j,1}}{\text{maximize}} \quad & P_E(\alpha, P_{j,0}, P_{j,1}) \\ \text{subject to} \quad & 0 \le P_{j,0} \le 1 \\ & 0 \le P_{j,1} \le 1. \end{aligned} \tag{P2}$$

For a fixed fraction of Byzantines α, the attacker wants to maximize the probability of error P_E by choosing its attacking strategy $(P_{j,0}, P_{j,1})$ optimally. We assume that the attacker is aware of the fact that the FC is using the majority rule for making the global decision. Before presenting our main results



(a)

(b)

Fig. 2. (a) P_e as a function of $(P_{1,0}, P_{0,1})$ when $P_0 = P_1 = 0.5$. (b) P_e as a function of $(P_{1,0}, P_{0,1})$ when $P_0 = 0.1, P_1 = 0.9$.

for Problem P2, we make an assumption that will be used in the theorem.

Assumption 1. *We assume that* $\alpha < \min\{(0.5 - P_f), (1 - (m/P_d))\}$,[6] *where* $m = \frac{N}{2N-2}$.

A consequence of this assumption is $\pi_{1,1} > m$, which can be shown as follows. By (9), we have

$$\begin{aligned}\pi_{1,1} &= \alpha(P_{1,0}(1-P_d) + (1-P_{0,1})P_d) + (1-\alpha)P_d \\ &= \alpha P_{1,0}(1-P_d) - \alpha P_d P_{0,1} + P_d \\ &\geq -\alpha P_d P_{0,1} + P_d \geq P_d(1-\alpha) > m. \qquad (16)\end{aligned}$$

Eq. (16) is true because $\alpha < \min\{(0.5-P_f), (1-(m/P_d))\} \leq (1-(m/P_d))$. Another consequence of this assumption is $\pi_{1,0} < 0.5$, which can be shown as follows. From (8), we have

$$\begin{aligned}\pi_{1,0} &= \alpha(P_{1,0}(1-P_f) + (1-P_{0,1})P_f) + (1-\alpha)P_f \\ &= \alpha P_{1,0} - \alpha P_f(P_{1,0} + P_{0,1}) + P_f \\ &\leq \alpha + P_f < 0.5. \qquad (17)\end{aligned}$$

Eq. (17) is true because $\alpha < \min\{(0.5-P_f), (1-(m/P_d))\} \leq (0.5 - P_f)$.

Next, we analyze the properties of P_E with respect to $(P_{1,0}, P_{0,1})$ under our assumption that enable us to find the optimal attacking strategies.

Lemma 2. *Assume that the FC employs the majority fusion rule* K^* *and* $\alpha < \min\{(0.5-P_f), (1-(m/P_d))\}$, *where* $m = \frac{N}{2N-2}$. *Then, for any fixed value of* $P_{0,1}$, *the error probability* P_E *at the FC is a quasi-convex function of* $P_{1,0}$.

Proof: A function $f(P_{1,0})$ is quasi-convex if, for some $P^*_{1,0}$, $f(P_{1,0})$ is non-increasing for $P_{1,0} \leq P^*_{1,0}$ and $f(P_{1,0})$ is non-decreasing for $P_{1,0} \geq P^*_{1,0}$. In other words, the lemma is proved if $\frac{dP_E}{dP_{1,0}} \leq 0$ (or $\frac{dP_E}{dP_{1,0}} \geq 0$) for all $P_{1,0}$, or if for some $P^*_{1,0}$, $\frac{dP_E}{dP_{1,0}} \leq 0$ when $P_{1,0} \leq P^*_{1,0}$ and $\frac{dP_E}{dP_{1,0}} \geq 0$ when $P_{1,0} \geq P^*_{1,0}$. First, we calculate the partial derivative of P_E with respect to $P_{1,0}$ for an arbitrary K as follows:

$$\frac{dP_E}{dP_{1,0}} = P_0 \frac{dQ_F}{dP_{1,0}} - P_1 \frac{dQ_D}{dP_{1,0}}. \qquad (18)$$

The detailed derivation of $\frac{dP_E}{dP_{1,0}}$ is given in Appendix C and we present a summary of the main results below.

$$\frac{dQ_F}{dP_{1,0}} = \alpha(1-P_f)N \begin{pmatrix} N-1 \\ K-1 \end{pmatrix} (\pi_{1,0})^{K-1}(1-\pi_{1,0})^{N-K}, \qquad (19)$$

$$\frac{dQ_D}{dP_{1,0}} = \alpha(1-P_d)N \begin{pmatrix} N-1 \\ K-1 \end{pmatrix} (\pi_{1,1})^{K-1}(1-\pi_{1,1})^{N-K}, \qquad (20)$$

and

$$\begin{aligned}\frac{dP_E}{dP_{1,0}} &= -P_1\alpha(1-P_d)N \begin{pmatrix} N-1 \\ K-1 \end{pmatrix} (\pi_{1,1})^{K-1}(1-\pi_{1,1})^{N-K} \\ &+ P_0\alpha(1-P_f)N \begin{pmatrix} N-1 \\ K-1 \end{pmatrix} (\pi_{1,0})^{K-1}(1-\pi_{1,0})^{N-K}. \qquad (21)\end{aligned}$$

$\frac{dP_E}{dP_{1,0}}$ given in (21) can be reformulated as follows:

$$\frac{dP_E}{dP_{1,0}} = g(P_{1,0}, K, \alpha)\left(e^{r(P_{1,0},K,\alpha)} - 1\right), \qquad (22)$$

where

$$g(P_{1,0}, K, \alpha) = N \begin{pmatrix} N-1 \\ K-1 \end{pmatrix} P_1\alpha(1-P_d)(\pi_{1,1})^{K-1}(1-\pi_{1,1})^{N-K} \qquad (23)$$

and

$$\begin{aligned}r(P_{1,0}, K, \alpha) &= \ln\left(\frac{P_0}{P_1}\frac{1-P_f}{1-P_d}\left(\frac{\pi_{1,0}}{\pi_{1,1}}\right)^{(K-1)}\left(\frac{1-\pi_{1,0}}{1-\pi_{1,1}}\right)^{(N-K)}\right) \\ &= \ln\left(\frac{P_0}{P_1}\frac{1-P_f}{1-P_d}\right) + (K-1)\ln\frac{\pi_{1,0}}{\pi_{1,1}} + (N-K)\ln\frac{1-\pi_{1,0}}{1-\pi_{1,1}}. \qquad (24)\end{aligned}$$

[6] Condition $\alpha < \min\{(0.5-P_f), (1-(m/P_d))\}$, where $m = \frac{N}{2N-2} > 0.5$, suggests that as N tends to infinity, $m = \dfrac{N}{2N-2}$ tends to 0.5. When P_d tends to 1 and P_f tends to 0, the above condition becomes $\alpha < 0.5$.

It can be seen that $g\left(P_{1,0}, K, \alpha\right) \geq 0$ so that the sign of $\frac{dP_E}{dP_{1,0}}$ depends only on the value of $r\left(P_{1,0}, K, \alpha\right)$. To prove that P_E is a quasi-convex function of $P_{1,0}$ when the majority rule K^* is used at the FC, it is sufficient to show that $r\left(P_{1,0}, K^*, \alpha\right)$ is a non-decreasing function. Differentiating $r\left(P_{1,0}, K^*, \alpha\right)$ with respect to $P_{1,0}$, we get

$$\begin{aligned}
&\frac{dr\left(P_{1,0}, K^*, \alpha\right)}{dP_{1,0}} = \\
&\quad (K^*-1)\left(\frac{\alpha(1-P_f)}{\pi_{1,0}} - \frac{\alpha(1-P_d)}{\pi_{1,1}}\right) \\
&\qquad\qquad +(N-K^*)\left(\frac{\alpha(1-P_d)}{1-\pi_{1,1}} - \frac{\alpha(1-P_f)}{1-\pi_{1,0}}\right) \\
&\quad = (K^*-1)\alpha\left(\frac{1-P_f}{\pi_{1,0}} - \frac{1-P_d}{\pi_{1,1}}\right) \\
&\qquad\qquad -(N-K^*)\alpha\left(\frac{1-P_f}{1-\pi_{1,0}} - \frac{1-P_d}{1-\pi_{1,1}}\right).
\end{aligned}$$

It can be shown that $\dfrac{dr\left(P_{1,0}, K^*, \alpha\right)}{dP_{1,0}} > 0$ (see Appendix B) and this completes the proof. ■

Quasi-convexity of P_E over $P_{1,0}$ implies that the maximum of the function occurs on the corners, i.e., $P_{1,0} =$ 0 or 1 (may not be unique). Next, we analyze the properties of P_E with respect to $P_{0,1}$.

Lemma 3. *Assume that the FC employs the majority fusion rule K^* and $\alpha < \min\{(0.5-P_f),(1-(m/P_d))\}$, where $m = \frac{N}{2N-2}$. Then, the probability of error P_E at the FC is a quasi-convex function of $P_{0,1}$ for a fixed $P_{1,0}$.*

Proof: For a fixed $P_{1,0}$, we have

$$(\pi_{1,0})' = d\pi_{1,0}/dP_{0,1} = \alpha(-P_f). \tag{25}$$

By a similar argument as given in Appendix C, for an arbitrary K we have

$$\begin{aligned}
\frac{dP_E}{dP_{0,1}} &= P_1\alpha P_d N\begin{pmatrix} N-1 \\ K-1 \end{pmatrix}\left(\pi_{1,1}\right)^{K-1}\left(1-\pi_{1,1}\right)^{N-K} \\
&-P_0\alpha P_f N\begin{pmatrix} N-1 \\ K-1 \end{pmatrix}\left(\pi_{1,0}\right)^{K-1}\left(1-\pi_{1,0}\right)^{N-K}.
\end{aligned} \tag{26}$$

$\dfrac{dP_E}{dP_{0,1}}$ given in (26) can be reformulated as follows:

$$\frac{dP_E}{dP_{0,1}} = g\left(P_{0,1}, K, \alpha\right)\left(e^{r\left(P_{0,1}, K, \alpha\right)} - 1\right), \tag{27}$$

where

$$g\left(P_{0,1}, K, \alpha\right) = N\begin{pmatrix} N-1 \\ K-1 \end{pmatrix} P_0\alpha P_f(\pi_{1,0})^{K-1}(1-\pi_{1,0})^{N-K} \tag{28}$$

and

$$\begin{aligned}
r\left(P_{0,1}, K, \alpha\right) &= \ln\left(\frac{P_1}{P_0}\frac{P_d}{P_f}\left(\frac{\pi_{1,1}}{\pi_{1,0}}\right)^{(K-1)}\left(\frac{1-\pi_{1,1}}{1-\pi_{1,0}}\right)^{(N-K)}\right) \\
&= \ln\frac{P_1}{P_0}\frac{P_d}{P_f} + (K-1)\ln\frac{\pi_{1,1}}{\pi_{1,0}} + (N-K)\ln\frac{1-\pi_{1,1}}{1-\pi_{1,0}}.
\end{aligned} \tag{29}$$

It can be seen that $g\left(P_{0,1}, K, \alpha\right) \geq 0$ such that the sign of $\dfrac{dP_E}{dP_{0,1}}$ depends on the value of $r\left(P_{0,1}, K, \alpha\right)$. To prove that P_E is a quasi-convex function of $P_{1,0}$ when the majority rule K^* is used at the FC, it is sufficient to show that $r\left(P_{0,1}, K^*, \alpha\right)$ is a non-decreasing function. Differentiating $r\left(P_{0,1}, K^*, \alpha\right)$ with respect to $P_{0,1}$, we get

$$\begin{aligned}
&\frac{dr\left(P_{0,1}, K^*, \alpha\right)}{dP_{0,1}} = \\
&(K^*-1)\left(\frac{\alpha P_f}{\pi_{1,0}} - \frac{\alpha P_d}{\pi_{1,1}}\right) + (N-K^*)\left(\frac{\alpha P_d}{1-\pi_{1,1}} - \frac{\alpha P_f}{1-\pi_{1,0}}\right)
\end{aligned} \tag{30}$$

$$= (N-K^*)\alpha\left(\frac{P_d}{1-\pi_{1,1}} - \frac{P_f}{1-\pi_{1,0}}\right) - (K^*-1)\alpha\left(\frac{P_d}{\pi_{1,1}} - \frac{P_f}{\pi_{1,0}}\right). \tag{31}$$

In the following, we show that

$$\frac{dr\left(P_{0,1}, K^*, \alpha\right)}{dP_{0,1}} > 0, \tag{32}$$

i.e., $r\left(P_{0,1}, K^*, \alpha\right)$ is non-decreasing. It is sufficient to show that

$$(N-K^*)\left(\frac{P_d}{1-\pi_{1,1}} - \frac{P_f}{1-\pi_{1,0}}\right) > (K^*-1)\left(\frac{P_d}{\pi_{1,1}} - \frac{P_f}{\pi_{1,0}}\right). \tag{33}$$

First, we consider the case when there are an even number of nodes in the network and majority fusion rule is given by $K^* = \dfrac{N}{2}+1$. Since $0 \leq \pi_{1,0} < \pi_{1,1} \leq 1$ and $N \geq 2$, we have

$$\begin{aligned}
&\left(1-\frac{2}{N}\right)\frac{\pi_{1,1}\pi_{1,0}}{(1-\pi_{1,1})(1-\pi_{1,0})} > -1 \\
\Leftrightarrow\ &\left(1-\frac{2}{N}\right)\left[\frac{1}{1-\pi_{1,1}} - \frac{1}{1-\pi_{1,0}}\right] > \left[\frac{1}{\pi_{1,1}} - \frac{1}{\pi_{1,0}}\right] \\
\Leftrightarrow\ &\left(1-\frac{2}{N}\right)\frac{1}{1-\pi_{1,1}} - \frac{1}{\pi_{1,1}} > \left(1-\frac{2}{N}\right)\frac{1}{1-\pi_{1,0}} - \frac{1}{\pi_{1,0}}
\end{aligned} \tag{34}$$

Using the fact that $\dfrac{P_d}{P_f} > 1$, $\pi_{1,1} > \frac{N}{2N-2}$, and $K^* = \dfrac{N}{2}+1$, (34) becomes

$$\begin{aligned}
&\frac{P_d}{P_f}\left[\left(1-\frac{2}{N}\right)\frac{1}{1-\pi_{1,1}} - \frac{1}{\pi_{1,1}}\right] > \\
&\qquad\qquad \left[\left(1-\frac{2}{N}\right)\frac{1}{1-\pi_{1,0}} - \frac{1}{\pi_{1,0}}\right]
\end{aligned} \tag{35}$$

$$\begin{aligned}
\Leftrightarrow\ &\left(1-\frac{2}{N}\right)\frac{P_d}{1-\pi_{1,1}} - \frac{P_d}{\pi_{1,1}} > \left(1-\frac{2}{N}\right)\frac{P_f}{1-\pi_{1,0}} - \frac{P_f}{\pi_{1,0}} \\
\Leftrightarrow\ &(N-K^*)\left(\frac{P_d}{1-\pi_{1,1}} - \frac{P_f}{1-\pi_{1,0}}\right) > (K^*-1)\left(\frac{P_d}{\pi_{1,1}} - \frac{P_f}{\pi_{1,0}}\right)
\end{aligned} \tag{36}$$

Next, we consider the case when there are odd number of nodes in the network and majority fusion rule is given by $K^* = \dfrac{N+1}{2}$. By using the fact that $\frac{\pi_{1,0}}{\pi_{1,1}} > \frac{P_f}{P_d}$, it can be seen that the right-hand side of (36) is nonnegative. Hence, from (36), we have

$$\begin{aligned}
&\left(\frac{N}{2}-1\right)\left(\frac{P_d}{1-\pi_{1,1}} - \frac{P_f}{1-\pi_{1,0}}\right) > \frac{N}{2}\left(\frac{P_d}{\pi_{1,1}} - \frac{P_f}{\pi_{1,0}}\right) \\
\Leftrightarrow\ &\left(\frac{N-1}{2}\right)\left(\frac{P_d}{1-\pi_{1,1}} - \frac{P_f}{1-\pi_{1,0}}\right) > \\
&\qquad\qquad \left(\frac{N-1}{2}\right)\left(\frac{P_d}{1-\pi_{1,1}} - \frac{P_f}{1-\pi_{1,0}}\right) \\
\Leftrightarrow\ &(N-K^*)\left(\frac{P_d}{1-\pi_{1,1}} - \frac{P_f}{1-\pi_{1,0}}\right) > (K^*-1)\left(\frac{P_d}{\pi_{1,1}} - \frac{P_f}{\pi_{1,0}}\right).
\end{aligned} \tag{37}$$

This completes our proof. ■

Theorem 2. *$(1,0)$, $(0,1)$, or $(1,1)$ are the optimal attacking strategies $(P_{1,0}, P_{0,1})$ that maximize the probability of error P_E, when the majority fusion rule is employed at the FC and $\alpha < \min\{(0.5-P_f), (1-(m/P_d))\}$, where $m = \frac{N}{2N-2}$.*

Proof: Lemma 2 and Lemma 3 suggest that one of the corners is the maximum of P_E because of quasi-convexity. Note that $(0,0)$ cannot be the solution of the maximization problem since the attacker does not flip any results. Hence, we end up with three possibilities: $(1,0)$, $(0,1)$, or $(1,1)$. ■

Next, to gain insights into Theorem 2, we present illustrative examples that corroborate our results.

A. Illustrative Examples

In Figure 3(a), we plot the probability of error P_E as a function of the attacking strategy $(P_{1,0}, P_{0,1})$ for even number of nodes, $N = 10$, in the network. We assume that the probability of detection is $P_d = 0.8$, the probability of false alarm is $P_f = 0.1$, prior probabilities are $(P_0 = 0.4, P_1 = 0.6)$, and $\alpha = 0.37$. Since $\alpha < \min\{(0.5-P_f), (1-(m/P_d))\}$, where $m = \frac{N}{2N-2}$, quasi-convexity can be observed in Figure 3(a). Figure 3(b) shows the probability of error P_E as a function of attacking strategy $(P_{1,0}, P_{0,1})$ for odd number of nodes, $N = 11$, in the network. Similarly, quasi-convexity can be observed in Figure 3(b).

It is evident from Figures 3(a) and 3(b) that the optimal attacking strategy $(P_{1,0}, P_{0,1})$ is either of the following three possibilities: $(1,0)$, $(0,1)$, or $(1,1)$. These results corroborate our theoretical results presented in Theorem 2.

Observe that the results obtained for this case are not the same as the results obtained for the asymptotic case (Please see Theorem 1). This is because the asymptotic performance measure (i.e., Chernoff information) is the exponential decay rate of the error probability of the "optimal detector". In other words, while optimizing over Chernoff information, one implicitly assumed that the optimal fusion rule is used at the FC.

Next, we investigate the case where the FC has the knowledge of attacker's strategies and uses the optimal fusion rule K^* to make the global decision. Here, the attacker tries to maximize its worst case probability of error $\underset{K}{min} P_E$ by choosing $(P_{1,0}, P_{0,1})$ optimally.

VI. Optimal Byzantine Attacking Strategies with Strategy-aware FC

In this section, we analyze the scenario where the FC has the knowledge of attacker's strategies and uses the optimal fusion rule K^* to make the global decision. The Byzantine attack problem can be formally stated as follows:

$$\begin{aligned} \underset{P_{j,0},P_{j,1}}{\text{maximize}} \quad & P_E(K^*, \alpha, P_{j,0}, P_{j,1}) \\ \text{subject to} \quad & 0 \le P_{j,0} \le 1 \\ & 0 \le P_{j,1} \le 1, \end{aligned} \tag{P3}$$

where K^* is the optimal fusion rule. In other words, K^* is the best response of the FC to the Byzantine attacking strategies. Next, we find the expression for the optimal fusion rule K^* used at the FC.

A. Optimal Fusion Rule

First, we design the optimal fusion rule assuming that the local sensor threshold λ and the Byzantine attacking strategy $(\alpha, P_{1,0}, P_{0,1})$ are fixed and known to the FC.

Lemma 4. *For a fixed local sensor threshold λ and $\alpha < \dfrac{1}{P_{0,1}+P_{1,0}}$, the optimal fusion rule is given by*

$$K^* \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\ln\left[(P_0/P_1)\{(1-\pi_{1,0})/(1-\pi_{1,1})\}^N\right]}{\ln\left[\{\pi_{1,1}(1-\pi_{1,0})\}/\{\pi_{1,0}(1-\pi_{1,1})\}\right]}. \tag{38}$$

Proof: Consider the maximum a *posteriori* probability (MAP) rule

$$\frac{P(\mathbf{u}|H_1)}{P(\mathbf{u}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1}.$$

Since the u_is are independent of each other, the MAP rule simplifies to

$$\prod_{i=1}^{N} \frac{P(u_i|H_1)}{P(u_i|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1}.$$

Let us assume that K^* out of N nodes send $u_i = 1$. Now, the above equation can be written as

$$\frac{\pi_{1,1}^{K^*}(1-\pi_{1,1})^{N-K^*}}{\pi_{1,0}^{K^*}(1-\pi_{1,0})^{N-K^*}} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0}{P_1}.$$

Taking logarithms on both sides of the above equation, we have

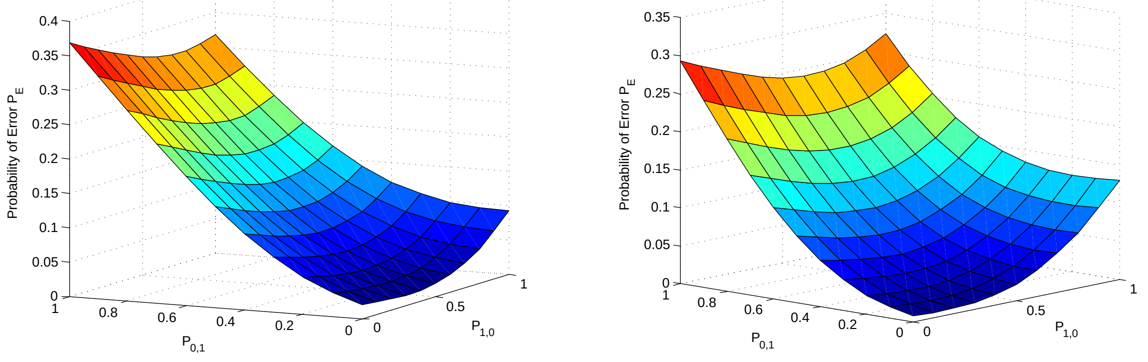
$$\begin{aligned} & K^* \ln \pi_{1,1} + (N-K^*)\ln(1-\pi_{1,1}) - K^* \ln \pi_{1,0} \\ & \qquad -(N-K^*)\ln(1-\pi_{1,0}) \underset{H_0}{\overset{H_1}{\gtrless}} \ln\frac{P_0}{P_1} \\ \Leftrightarrow \quad & K^*[\ln(\pi_{1,1}/\pi_{1,0}) \\ & +\ln((1-\pi_{1,0})/(1-\pi_{1,1}))] \underset{H_0}{\overset{H_1}{\gtrless}} \ln\frac{P_0}{P_1} + N\ln\frac{1-\pi_{1,0}}{1-\pi_{1,1}} \\ \Leftrightarrow \quad & K^* \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\ln\frac{P_0}{P_1} + N\ln((1-\pi_{1,0})/(1-\pi_{1,1}))}{[\ln(\pi_{1,1}/\pi_{1,0}) + \ln((1-\pi_{1,0})/(1-\pi_{1,1}))]} \qquad (39) \\ \Leftrightarrow \quad & K^* \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\ln\left[(P_0/P_1)\{(1-\pi_{1,0})/(1-\pi_{1,1})\}^N\right]}{\ln\left[\{\pi_{1,1}(1-\pi_{1,0})\}/\{\pi_{1,0}(1-\pi_{1,1})\}\right]}, \end{aligned}$$

where (39) follows from the fact that, for $\pi_{1,1} > \pi_{1,0}$ or equivalently, $\alpha < \dfrac{1}{P_{0,1}+P_{1,0}}$, $[\ln(\pi_{1,1}/\pi_{1,0}) + \ln((1-\pi_{1,0})/(1-\pi_{1,1}))] > 0$. ■

The probability of false alarm Q_F and the probability of detection Q_D for this case are as given in (6) and (7) with $K = \lceil K^* \rceil$. Next, we present our results for the case when the fraction of Byzantines $\alpha > \dfrac{1}{P_{0,1}+P_{1,0}}$.

Lemma 5. *For a fixed local sensor threshold λ and $\alpha > \dfrac{1}{P_{0,1}+P_{1,0}}$, the optimal fusion rule is given by*

$$K^* \underset{H_1}{\overset{H_0}{\gtrless}} \frac{\ln\left[(P_1/P_0)\{(1-\pi_{1,1})/(1-\pi_{1,0})\}^N\right]}{[\ln(\pi_{1,0}/\pi_{1,1}) + \ln((1-\pi_{1,1})/(1-\pi_{1,0}))]}. \tag{40}$$



(a)

(b)

Fig. 3. (a) P_E as a function of $(P_{1,0}, P_{0,1})$ for $N = 10$. (b) P_E as a function of $(P_{1,0}, P_{0,1})$ for $N = 11$.

Proof: This can be proved similarly as Lemma 4 and using the fact that, for $\pi_{1,1} < \pi_{1,0}$ or equivalently, $\alpha > \dfrac{1}{P_{0,1} + P_{1,0}}$, $[\ln(\pi_{1,0}/\pi_{1,1}) + \ln((1-\pi_{1,1})/(1-\pi_{1,0}))] > 0$. ■

The probability of false alarm Q_F and the probability of detection Q_D for this case can be calculated to be

$$Q_F = \sum_{i=0}^{\lfloor K^* \rfloor} \binom{N}{i} (\pi_{1,0})^i (1-\pi_{1,0})^{N-i} \tag{41}$$

and

$$Q_D = \sum_{i=0}^{\lfloor K^* \rfloor} \binom{N}{i} (\pi_{1,1})^i (1-\pi_{1,1})^{N-i}. \tag{42}$$

Next, we analyze the property of P_E with respect to Byzantine attacking strategy $(P_{1,0}, P_{0,1})$ that enables us to find the optimal attacking strategies.

Lemma 6. *For a fixed local sensor threshold λ, assume that the FC employs the optimal fusion rule $\lceil K^* \rceil$,* [7] *as given in* (38). *Then, for $\alpha \leq 0.5$, the error probability P_E at the FC is a monotonically increasing function of $P_{1,0}$ while $P_{0,1}$ remains fixed. Conversely, the error probability P_E at the FC is a monotonically increasing function of $P_{0,1}$ while $P_{1,0}$ remains fixed.*

Proof: Observe that, for a fixed λ, $P_E(\lceil K^* \rceil)$ is a continuous but not a differentiable function. However, the function is non differentiable only at a finite number (or infinitely countable number) of points because of the nature of $\lceil K^* \rceil$. Now observe that, for a fixed fusion rule K, $P_E(K)$ is differentiable. Utilizing this fact, to show that the lemma is true, we first find the condition that a fusion rule K should satisfy so that P_E is a monotonically increasing function of $P_{1,0}$ while keeping $P_{0,1}$ fixed (and vice versa) and later show that $\lceil K^* \rceil$ satisfies this condition. From (22), finding those K that satisfy $\dfrac{dP_E}{dP_{1,0}} > 0$[8] is equivalent to finding those value of K that make

$$\begin{aligned} & r\left(P_{1,0}, K, \alpha\right) > 0 \\ \Leftrightarrow\quad & \ln \frac{P_0}{P_1} \frac{1-P_f}{1-P_d} + (K-1) \ln \frac{\pi_{1,0}}{\pi_{1,1}} + (N-K) \ln \frac{1-\pi_{1,0}}{1-\pi_{1,1}} > 0 \\ \Leftrightarrow\quad & K < \frac{\ln \dfrac{P_0}{P_1} + N \ln \dfrac{(1-\pi_{1,0})}{(1-\pi_{1,1})} + \ln \dfrac{1-P_f}{1-P_d} - \ln \dfrac{\pi_{1,0}}{\pi_{1,1}}}{\ln\left[\{\pi_{1,1}(1-\pi_{1,0})\}/\{\pi_{1,0}(1-\pi_{1,1})\}\right]}. \end{aligned} \tag{43}$$

Similarly, we can find the condition that a fusion rule K should satisfy so that P_E is a monotonically increasing function of $P_{0,1}$ while keeping $P_{1,0}$ fixed. From (27), finding those K that satisfy $\dfrac{dP_E}{dP_{0,1}} > 0$ is equivalent to finding those K that make

$$\begin{aligned} & r\left(P_{0,1}, K, \alpha\right) > 0 \\ \Leftrightarrow\quad & \ln \frac{P_1}{P_0} \frac{P_d}{P_f} + (K-1) \ln \frac{\pi_{1,1}}{\pi_{1,0}} + (N-K) \ln \frac{1-\pi_{1,1}}{1-\pi_{1,0}} > 0 \\ \Leftrightarrow\quad & K > \frac{\ln \dfrac{P_0}{P_1} + N \ln \dfrac{(1-\pi_{1,0})}{(1-\pi_{1,1})} + \ln \dfrac{P_f}{P_d} - \ln \dfrac{\pi_{1,0}}{\pi_{1,1}}}{\ln\left[\{\pi_{1,1}(1-\pi_{1,0})\}/\{\pi_{1,0}(1-\pi_{1,1})\}\right]}. \end{aligned} \tag{44}$$

From (43) and (44), we have

$$A > K > B \tag{45}$$

where

$$A = \frac{\ln \dfrac{P_0}{P_1} + N \ln \dfrac{(1-\pi_{1,0})}{(1-\pi_{1,1})} + \ln \dfrac{1-P_f}{1-P_d} - \ln \dfrac{\pi_{1,0}}{\pi_{1,1}}}{\ln\left[\{\pi_{1,1}(1-\pi_{1,0})\}/\{\pi_{1,0}(1-\pi_{1,1})\}\right]}$$

and

$$B = \frac{\ln \dfrac{P_0}{P_1} + N \ln \dfrac{(1-\pi_{1,0})}{(1-\pi_{1,1})} + \ln \dfrac{P_f}{P_d} - \ln \dfrac{\pi_{1,0}}{\pi_{1,1}}}{\ln\left[\{\pi_{1,1}(1-\pi_{1,0})\}/\{\pi_{1,0}(1-\pi_{1,1})\}\right]}.$$

Next, we show that the optimal fusion rule $\lceil K^* \rceil$ given in (38)

[7] Notice that, K^* might not be an integer.

[8] Observe that, for $\alpha < 0.5$, the function $g\left(P_{1,0}, K^*, \alpha\right) = 0$ (as given in (23)) only under extreme conditions (i.e., $P_1 = 0$ or $P_d = 0$ or $P_d = 1$). Ignoring these extreme conditions, we have $g\left(P_{1,0}, K^*, \alpha\right) > 0$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2015.2450191, IEEE Transactions on Signal Processing

is within the region (A, B). First we prove that $\lceil K^* \rceil > B$ by showing $K^* > B$. Comparing K^* given in (38) with B, $K^* > B$ iff

$$0 > \ln \frac{P_f}{P_d} - \ln \frac{\pi_{1,0}}{\pi_{1,1}}. \tag{46}$$

Since $P_d > P_f$, to prove (46) we start from the inequality

$$\begin{aligned}
& \frac{(1-P_d)}{P_d} < \frac{(1-P_f)}{P_f} \\
\Leftrightarrow \quad & \frac{\alpha P_{1,0}(1-P_d) + P_d(1-P_{0,1}\alpha)}{P_d} < \\
& \qquad \frac{\alpha P_{1,0}(1-P_f) + P_f(1-P_{0,1}\alpha)}{P_f} \\
\Leftrightarrow \quad & \frac{\pi_{1,1}}{P_d} < \frac{\pi_{1,0}}{P_f} \\
\Leftrightarrow \quad & 0 > \ln \frac{P_f}{P_d} - \ln \frac{\pi_{1,0}}{\pi_{1,1}}.
\end{aligned}$$

Now, we show that $A > \lceil K^* \rceil$. Observe that,

$$\begin{aligned}
& A > \lceil K^* \rceil \\
\Leftrightarrow \quad & \frac{\ln \dfrac{1-P_f}{1-P_d} - \ln \dfrac{\pi_{1,0}}{\pi_{1,1}}}{\ln\left[\{\pi_{1,1}(1-\pi_{1,0})\}/\{\pi_{1,0}(1-\pi_{1,1})\}\right]} > \lceil K^* \rceil - K^*.
\end{aligned}$$

Hence, it is sufficient to show that

$$\frac{\ln \dfrac{1-P_f}{1-P_d} - \ln \dfrac{\pi_{1,0}}{\pi_{1,1}}}{\ln\left[\{\pi_{1,1}(1-\pi_{1,0})\}/\{\pi_{1,0}(1-\pi_{1,1})\}\right]} > 1 > \lceil K^* \rceil - K^*.$$

$1 > \lceil K^* \rceil - K^*$ is true from the property of the ceiling function. By (56), we have

$$\begin{aligned}
& \frac{1-P_f}{1-P_d} > \frac{1-\pi_{1,0}}{1-\pi_{1,1}} \\
\Leftrightarrow \quad & \ln \frac{1-P_f}{1-P_d} > \ln \frac{1-\pi_{1,0}}{1-\pi_{1,1}} \\
\Leftrightarrow \quad & \ln \frac{1-P_f}{1-P_d} - \ln \frac{\pi_{1,0}}{\pi_{1,1}} > \ln\left[\{\pi_{1,1}(1-\pi_{1,0})\}/\{\pi_{1,0}(1-\pi_{1,1})\}\right] \\
\Leftrightarrow \quad & \frac{\ln \dfrac{1-P_f}{1-P_d} - \ln \dfrac{\pi_{1,0}}{\pi_{1,1}}}{\ln\left[\{\pi_{1,1}(1-\pi_{1,0})\}/\{\pi_{1,0}(1-\pi_{1,1})\}\right]} > 1
\end{aligned}$$

which completes the proof. ■

Based on Lemma 6, we present the optimal attacking strategies for the case when the FC has the knowledge regarding the strategies used by the Byzantines.

Theorem 3. *The optimal attacking strategies, $(P^*_{1,0}, P^*_{0,1})$, which maximize the probability of error, $P_E(\lceil K^* \rceil)$, are given by*

$$(P^*_{1,0}, P^*_{0,1}) \begin{cases} (p_{1,0}, p_{0,1}) & \text{if } \alpha > 0.5 \\ (1,1) & \text{if } \alpha \leq 0.5 \end{cases}$$

where $(p_{1,0}, p_{0,1})$ satisfies $\alpha(p_{1,0} + p_{0,1}) = 1$.

Proof: Note that, the maximum probability of error occurs when the posterior probabilities are equal to the prior probabilities of the hypotheses. That is,

$$P(H_i|\mathbf{u}) = P(H_i) \text{ for } i = 0, 1. \tag{47}$$

Now using the result from (13), the condition can be simplified to

$$\alpha(P_{1,0} + P_{0,1}) = 1. \tag{48}$$

Eq. (48) suggests that when $\alpha \geq 0.5$, the attacker can find flipping probabilities that make $P_E = \min\{P_0, P_1\}$. When $\alpha = 0.5$, $P_{1,0} = P_{0,1} = 1$ is the optimal attacking strategy and when $\alpha > 0.5$, any pair which satisfies $P_{1,0} + P_{0,1} = \dfrac{1}{\alpha}$ is optimal. However, when $\alpha < 0.5$, (48) cannot be satisfied. In this case, by Lemma 6, for $\alpha < 0.5$, $(1, 1)$ is an optimal attacking strategy, $(P_{1,0}, P_{0,1})$, which maximizes probability of error, $P_E(\lceil K^* \rceil)$. ■

Next, to gain insight into Theorem 3, we present illustrative examples that corroborate our results.

B. Illustrative Examples

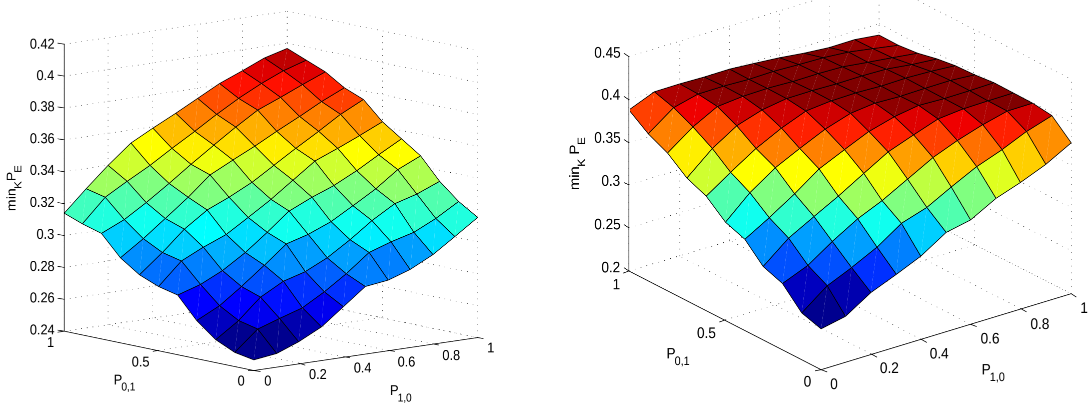
In Figure 4, we plot the minimum probability of error as a function of attacker's strategy $(P_{1,0}, P_{0,1})$, where P_E is minimized over all possible fusion rules K. We consider a $N = 11$ node network, with the nodes' detection and false alarm probabilities being 0.6 and 0.4, respectively. Prior probabilities are assumed to be $P_0 = 0.4$ and $P_1 = 0.6$. Observe that, the optimal fusion rule as given in (38) changes with attacker's strategy $(P_{1,0}, P_{0,1})$. Thus, the minimum probability of error $\min_K P_E$ is a non-differentiable function. It is evident from Figure 4(a) that $(P_{1,0}, P_{0,1}) = (1, 1)$ maximizes the probability of error, $P_E(\lceil K^* \rceil)$. This corroborates our theoretical results presented in Theorem 3, that for $\alpha < 0.5$, the optimal attacking strategy, $(P_{1,0}, P_{0,1})$, that maximizes the probability of error, $P_E(\lceil K^* \rceil)$, is $(1, 1)$.

In Figure 4(b) we consider the scenario where $\alpha = 0.8$ (i.e., $\alpha > 0.5$). It can be seen that the attacking strategy $(P_{1,0}, P_{0,1})$, that maximizes $\min_K P_E$ is not unique in this case. It can be verified that any attacking strategy which satisfies $P_{1,0} + P_{0,1} = \frac{1}{0.8}$ will make $\min_K P_E = \min\{P_0, P_1\} = 0.4$. This corroborates our theoretical results presented in Theorem 3. Observe that the results obtained for this case are consistent with the results obtained for the asymptotic case. This is because the optimal fusion rule is used at the FC and the asymptotic performance measure (i.e., Chernoff information) is the exponential decay rate of error probability of the "optimal detector", and thus, implicitly assumes that the optimal fusion rule is used at the FC.

When the attacker does not have the knowledge of the fusion rule K used at the FC, from an attacker's perspective, maximizing its local probability of error P_e is the optimal attacking strategy. The optimal attacking strategy in this case is either of the three possibilities: $(P_{1,0}, P_{0,1}) = (0, 1)$ or $(1, 0)$ or $(1, 1)$ (see Table II). However, the FC has knowledge of the attacking strategy $(\alpha, P_{1,0}, P_{0,1})$ and thus, uses the optimal fusion rule as given in (38) and (40).

VII. Conclusion and Future Work

We considered the problem of distributed Bayesian detection with Byzantine data, and characterized the power of attack analytically. For distributed detection for a binary hypothesis testing problem, the expression for the minimum attacking power above which the ability to detect is completely



(a)

(b)

Fig. 4. Minimum probability of error ($\min_K P_E$) analysis. (a) $\min_K P_E$ as a function of $(P_{1,0}, P_{0,1})$ for $\alpha = 0.4$. (b) $\min_K P_E$ as a function of $(P_{1,0}, P_{0,1})$ for $\alpha = 0.8$.

destroyed was obtained. We showed that when there are more than 50% of Byzantines in the network, the data fusion scheme becomes blind and no detector can achieve any performance gain over the one based just on priors. The optimal attacking strategies for Byzantines that degrade the performance at the FC were obtained. It was shown that the results obtained for the non-asymptotic case are consistent with the results obtained for the asymptotic case only when the FC has the knowledge of the attacker's strategies, and thus, uses the optimal fusion rule. However, results obtained for the non-asymptotic case, when the FC does not have knowledge of attacker's strategies, are not the same as the results obtained for the asymptotic case. There are still many interesting questions that remain to be explored in the future work such as an analysis of the scenario where Byzantines can also control sensor thresholds used for making local decisions. Other questions such as the case where Byzantines collude in several groups (collaborate) to degrade the detection performance can also be investigated.

Acknowledgment

This work was supported in part by ARO under Grant W911NF-14-1-0339, AFOSR under Grant FA9550-10-1-0458 and National Science Council of Taiwan, under grants NSC 99-2221-E-011-158 -MY3, NSC 101-2221-E-011-069 -MY3. Han's work was completed during his visit to Syracuse University from 2012 to 2013.

Appendix A
Sensitivity to Imperfect Knowledge

In this section, we discuss the sensitivity of system performance to imperfect knowledge regarding the fraction of Byzantines α in the network and the prior probability of hypotheses, i.e., $(P_0,\ P_1)$. We limit the analysis to a couple of illustrative examples.

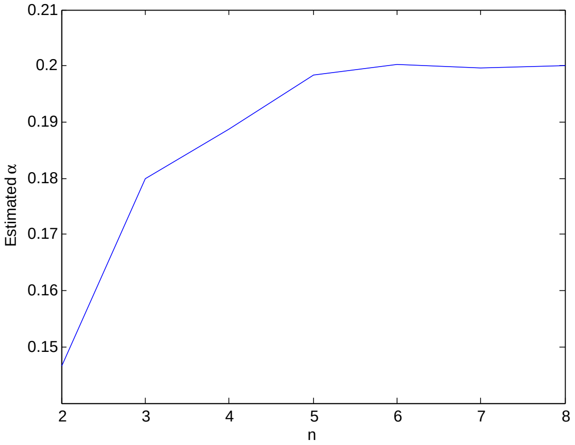


Fig. 5. Estimation of the fraction of Byzantines as a function of $N = 10^n$ when the true value of $\alpha = 0.2$.

In many practical scenarios, the value of the fraction of Byzantines α in the network might not be known a-priori. In such scenarios, α may be estimated (learned) by observing decisions at the FC over a fixed duration. Next, we present a rather simple estimation procedure and some numerical results to corroborate our claim.

We assume that $P_d = 0.8$, $P_f = 0.2$ and the fraction of Byzantines is $\alpha = 0.2$ with $(P_{1,0}, P_{0,1}) = (1, 1)$. Based on the received decisions under hypothesis H_1, the FC can estimate $\hat{\alpha}$ as follows:

$$\hat{\alpha} = \frac{P_d - \pi_{1,1}}{2P_d - 1},$$

where $\pi_{1,1}$ is the fraction of 1's received at the FC. In Figure 5, we plot the value of estimated α at the FC as a function of the number of decisions at the FC, i.e., $N = 10^n$.

It can be seen from Figure 5 that the estimated α approaches the true value of α as the number of decisions N at the FC increases.

Next, we look at the sensitivity of the performance of the detection scheme to the uncertainty regarding the fraction of Byzantines α and the prior probability of hypotheses, i.e.,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2015.2450191, IEEE Transactions on Signal Processing

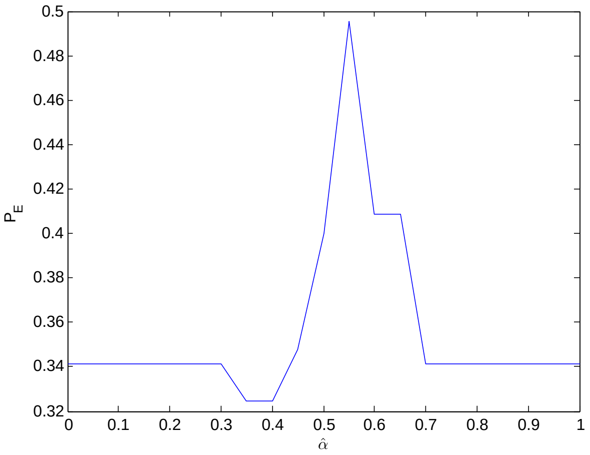


Fig. 6. Error probability in the presence of imperfect knowledge of α.

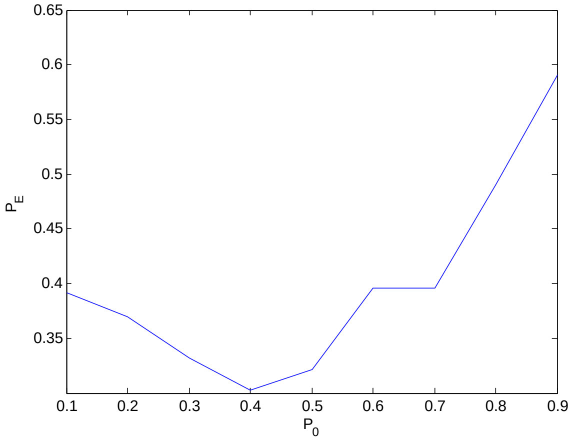


Fig. 7. Error probability in the presence of imperfect knowledge of P_0.

$(P_0,\ P_1)$. Sensitivity of the performance of the scheme to the uncertainty of parameter values is a model mismatch problem. In general, finding the analytical expressions for performance degradation due to model mismatch is a difficult problem, thus, we limit our analysis to numerical results. However, one can expect that the performance of the scheme will improve as the estimated parameter values approach their true value.

In Figure 6, we plot the probability of error as the estimated α, $\hat{\alpha}$, at the FC is varied from 0 to 1 when the actual value of α is 0.4, $N = 10$ and $(P_d, P_f) = (0.8, 0.1)$ with $(P_{1,0}, P_{0,1}) = (1, 1)$. Note that, the error probability is minimum when $\hat{\alpha}$ is equal to the actual α.

In Figure 7, we plot the probability of error as the value of P_0 at the FC is varied from 0.1 to 0.9 when the actual value of P_0 is 0.4, $N = 10$ and $(P_d, P_f) = (0.8, 0.1)$ with $(P_{1,0}, P_{0,1}) = (1, 1)$. Note that, the error probability is minimum when the estimated P_0 is equal to the actual P_0.

Appendix B
Proof of $\dfrac{dr\left(P_{1,0}, K^*, \alpha\right)}{dP_{1,0}} > 0$

Differentiating both sides of $r\left(P_{1,0}, K^*, \alpha\right)$ with respect to $P_{1,0}$, we get

$$\begin{aligned}\frac{dr\left(P_{1,0}, K^*, \alpha\right)}{dP_{1,0}} &= (K^*-1)\alpha\left(\frac{1-P_f}{\pi_{1,0}} - \frac{1-P_d}{\pi_{1,1}}\right)\\ &\quad -(N-K^*)\alpha\left(\frac{1-P_f}{1-\pi_{1,0}} - \frac{1-P_d}{1-\pi_{1,1}}\right).\end{aligned}$$

In the following we show that

$$\frac{dr\left(P_{1,0}, K^*, \alpha\right)}{dP_{1,0}} > 0 \tag{49}$$

i.e., $r\left(P_{1,0}, K^*, \alpha\right)$ is non-decreasing. Observe that in the above equation,

$$\frac{(1-P_f)}{\pi_{1,0}} > \frac{(1-P_d)}{\pi_{1,1}}. \tag{50}$$

To show that the above condition is true, we start from the inequality

$$P_d > P_f \tag{51}$$

$$\Leftrightarrow \quad \frac{P_d}{1-P_d} > \frac{P_f}{1-P_f} \tag{52}$$

$$\Leftrightarrow \quad \alpha P_{1,0} + (1-P_{0,1}\alpha)\frac{P_d}{1-P_d} > \alpha P_{1,0} + (1-P_{0,1}\alpha)\frac{P_f}{1-P_f} \tag{53}$$

$$\Leftrightarrow \quad \frac{\alpha P_{1,0}(1-P_d) + P_d(1-P_{0,1}\alpha)}{(1-P_d)} > \frac{\alpha P_{1,0}(1-P_f) + P_f(1-P_{0,1}\alpha)}{(1-P_f)} \tag{54}$$

$$\Leftrightarrow \quad \frac{\pi_{1,1}}{(1-P_d)} > \frac{\pi_{1,0}}{(1-P_f)} \tag{55}$$

$$\Leftrightarrow \quad \frac{(1-P_f)}{\pi_{1,0}} > \frac{(1-P_d)}{\pi_{1,1}} \tag{56}$$

Similarly, it can be shown that

$$\frac{1-\pi_{1,1}}{1-P_d} > \frac{1-\pi_{1,0}}{1-P_f} \tag{57}$$

Now from (50) and (57), to show that $\dfrac{dr\left(P_{1,0}, K^*, \alpha\right)}{dP_{1,0}} > 0$ is equivalent to show that

$$(K^*-1)\left(\frac{1-P_f}{\pi_{1,0}} - \frac{1-P_d}{\pi_{1,1}}\right) > (N-K^*)\left(\frac{1-P_f}{1-\pi_{1,0}} - \frac{1-P_d}{1-\pi_{1,1}}\right) \tag{58}$$

Next, we consider two different cases, first when there are odd number of nodes in the network and second when there are even number of nodes in the network.

Odd Number of Nodes: When there are odd number of nodes in the network, the majority fusion rule is $K^* = (N+1)/2$. In this case (58) is equivalent to show that

$$\frac{N-1}{2}\left(\frac{1-P_f}{\pi_{1,0}} - \frac{1-P_d}{\pi_{1,1}}\right) > \frac{N-1}{2}\left(\frac{1-P_f}{1-\pi_{1,0}} - \frac{1-P_d}{1-\pi_{1,1}}\right). \tag{59}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2015.2450191, IEEE Transactions on Signal Processing

To show that the above condition is true, we start from the following inequality

$$
\begin{aligned}
& \frac{(1-\pi_{1,0})(1-\pi_{1,1})}{\pi_{1,0}\pi_{1,1}} > -1 \\
\Leftrightarrow\quad & \left[\frac{1}{\pi_{1,0}} - \frac{1}{\pi_{1,1}}\right] > \left[\frac{1}{1-\pi_{1,0}} - \frac{1}{1-\pi_{1,1}}\right] \\
\Leftrightarrow\quad & \left[\frac{1}{\pi_{1,0}} - \frac{1}{1-\pi_{1,0}}\right] > \left[\frac{1}{\pi_{1,1}} - \frac{1}{1-\pi_{1,1}}\right]
\end{aligned}
$$

Since $\frac{1-P_f}{1-P_d} > 1$, $\pi_{1,0} < 0.5$ (consequence of our assumption) and $N \geq 2$, the above condition is equivalent to

$$
\begin{aligned}
& \frac{1-P_f}{1-P_d}\left[\frac{1}{\pi_{1,0}} - \frac{1}{1-\pi_{1,0}}\right] > \left[\frac{1}{\pi_{1,1}} - \frac{1}{1-\pi_{1,1}}\right] \\
\Leftrightarrow\quad & \left(\frac{1-P_f}{\pi_{1,0}} - \frac{1-P_d}{\pi_{1,1}}\right) > \left(\frac{1-P_f}{1-\pi_{1,0}} - \frac{1-P_d}{1-\pi_{1,1}}\right) \\
\Leftrightarrow\quad & \frac{N-1}{2}\left(\frac{1-P_f}{\pi_{1,0}} - \frac{1-P_d}{\pi_{1,1}}\right) > \\
& \qquad \frac{N-1}{2}\left(\frac{1-P_f}{1-\pi_{1,0}} - \frac{1-P_d}{1-\pi_{1,1}}\right) \qquad (60)
\end{aligned}
$$

which implies that $\frac{dr\left(P_{1,0}, K^*, \alpha\right)}{dP_{1,0}} > 0$ for odd number of nodes case. Next, we consider the even number of nodes case.

Even Number of Nodes: Now, we consider the case when there are even number of nodes in the network and majority fusion rule is given by $K^* = \frac{N}{2} + 1$. Condition (58) is equivalent to show that

$$
\left(\frac{N}{2}\right)\left(\frac{1-P_f}{\pi_{1,0}} - \frac{1-P_d}{\pi_{1,1}}\right) > \left(\frac{N}{2} - 1\right)\left(\frac{1-P_f}{1-\pi_{1,0}} - \frac{1-P_d}{1-\pi_{1,1}}\right).
$$

Which follows from the fact that

$$
\left(\frac{N}{2}\right)\left(\frac{1-P_f}{\pi_{1,0}} - \frac{1-P_d}{\pi_{1,1}}\right) > \left(\frac{N}{2} - 1\right)\left(\frac{1-P_f}{\pi_{1,0}} - \frac{1-P_d}{\pi_{1,1}}\right)
$$

and the result given in (59). This completes our proof.

Appendix C
Calculating partial derivative of P_E w.r.t. $P_{1,0}$

First, we calculate the partial derivative of Q_F with respect to $P_{1,0}$. Notice that,

$$
Q_F = \sum_{i=K^*}^{N} \binom{N}{i} (\pi_{1,0})^i (1-\pi_{1,0})^{N-i} \qquad (61)
$$

where

$$
\begin{aligned}
\pi_{1,0} &= \alpha(P_{1,0}(1-P_f) + (1-P_{0,1})P_f) + (1-\alpha)P_f \qquad (62) \\
(\pi_{1,0})' &= d\pi_{1,0}/dP_{1,0} = \alpha(1-P_f). \qquad (63)
\end{aligned}
$$

Differentiating both sides of (61) with respect to $P_{1,0}$, we get

$$
\begin{aligned}
& \frac{dQ_F}{dP_{1,0}} = \\
& \binom{N}{K^*}\left(K^*(\pi_{1,0})^{K^*-1}(\pi_{1,0})'(1-\pi_{1,0})^{N-K^*}\right. \\
& \qquad \left. -(\pi_{1,0})^{K^*}(N-K^*)(1-\pi_{1,0})^{N-K^*-1}(\pi_{1,0})'\right) \\
& + \binom{N}{K^*+1}\left((K^*+1)(\pi_{1,0})^{K^*}(\pi_{1,0})'(1-\pi_{1,0})^{N-K^*-1}\right. \\
& \qquad \left. -(\pi_{1,0})^{K^*+1}(N-K^*-1)(1-\pi_{1,0})^{N-K^*-2}(\pi_{1,0})'\right) \\
& \quad + \cdots + \binom{N}{N}\left(N(\pi_{1,0})^{N-1}(\pi_{1,0})' - 0\right) \\
= \; & (\pi_{1,0})'(\pi_{1,0})^{K^*-1}(1-\pi_{1,0})^{N-K^*} \\
& \qquad \left[\binom{N}{K^*}\left(K^* - \frac{\pi_{1,0}}{1-\pi_{1,0}}(N-K^*)\right)\right. \\
& + \left.\binom{N}{K^*+1}\left((K^*+1)\frac{\pi_{1,0}}{1-\pi_{1,0}} - (N-K^*-1)\frac{\pi_{1,0}^2}{(1-\pi_{1,0})^2}\right) + \cdots\right] \\
= \; & (\pi_{1,0})'(\pi_{1,0})^{K^*-1}(1-\pi_{1,0})^{N-K^*} \\
& \left[\binom{N}{K^*}(K^* - \frac{\pi_{1,0}}{1-\pi_{1,0}}(N-K^*)) + \frac{\pi_{1,0}}{1-\pi_{1,0}}\binom{N}{K^*+1}\right. \\
& \qquad \left.\left((K^*+1) - (N-K^*-1)\frac{\pi_{1,0}}{1-\pi_{1,0}}\right) + \cdots\right] \\
= \; & (\pi_{1,0})'(\pi_{1,0})^{K^*-1}(1-\pi_{1,0})^{N-K^*}\left[\binom{N}{K^*}K^* + \left[-\frac{\pi_{1,0}}{1-\pi_{1,0}}\right.\right. \\
& \left.\left.\binom{N}{K^*}(N-K^*) + \frac{\pi_{1,0}}{1-\pi_{1,0}}\binom{N}{K^*+1}(K^*+1)\right] + \cdots\right]
\end{aligned}
$$

Since, $\binom{N}{K^*}\frac{K^*}{N} = \binom{N-1}{K^*-1}$, the above equation can be written as

$$
\begin{aligned}
& \frac{dQ_F}{dP_{1,0}} = (\pi_{1,0})'(\pi_{1,0})^{K^*-1}(1-\pi_{1,0})^{N-K^*}\left[\binom{N-1}{K^*-1}N\right. \\
& \left. + \frac{\pi_{1,0}}{1-\pi_{1,0}}\left\{\binom{N}{K^*+1}(K^*+1) - \binom{N}{K^*}(N-K^*)\right\} + \cdots\right].
\end{aligned} \qquad (64)
$$

Notice that, for any positive integer t

$$
\left(\frac{\pi_{1,0}}{1-\pi_{1,0}}\right)^t\left[\binom{N}{K^*+t}(K^*+t) - \binom{N}{K^*+t-1}(N-K^*-t+1)\right] = 0. \qquad (65)
$$

Using the result from (65), (64) can be written as

$$
\begin{aligned}
& \frac{dQ_F}{dP_{1,0}} = (\pi_{1,0})'(\pi_{1,0})^{K^*-1}(1-\pi_{1,0})^{N-K^*}\left[\binom{N-1}{K^*-1}N + \frac{\pi_{1,0}}{1-\pi_{1,0}}[0]\right] \\
\Leftrightarrow\quad & \frac{dQ_F}{dP_{1,0}} = \alpha(1-P_f)N\binom{N-1}{K^*-1}(\pi_{1,0})^{K^*-1}(1-\pi_{1,0})^{N-K^*}.
\end{aligned}
$$

Similarly, the partial derivative of Q_D w.r.t. $P_{1,0}$ can calculated to be

$$
\frac{dQ_D}{dP_{1,0}} = \alpha(1-P_d)N\binom{N-1}{K^*-1}(\pi_{1,1})^{K^*-1}(1-\pi_{1,1})^{N-K^*}.
$$

References

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

[2] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I - Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54 –63, Jan 1997.

[3] V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, pp. 100–117, 2012.

[4] B. Wu, J. Chen, J. Wu, and M. Cardei, "A survey of attacks and countermeasures in mobile ad hoc networks," *Wireless/Mobile Network Security, Springer*, vol. 17, pp. 103–135, 2007.

[5] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, no. 3, pp. 433–481, 1985.

[6] S. Jaggi, M. Langberg, S. Katti, T. Ho, D. Katabi, and M. Medard, "Resilient network coding in the presence of byzantine adversaries," in *Proc. 26th IEEE Int. Conf. on Computer Commun., INFOCOM, (Anchorage, AK)*, 2007, pp. 616–624.

[7] L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," *ACM Trans. Program. Lang. Syst.*, vol. 4, no. 3, pp. 382–401, Jul. 1982. [Online]. Available: http://doi.acm.org/10.1145/357172.357176

[8] A. Vempaty, L. Tong, and P. Varshney, "Distributed Inference with Byzantine Data: State-of-the-Art Review on Data Falsification Attacks," *Signal Processing Magazine, IEEE*, vol. 30, no. 5, pp. 65–75, 2013.

[9] A. Fragkiadakis, E. Tragos, and I. Askoxylakis, "A survey on security threats and detection techniques in cognitive radio networks," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 428–445, 2013.

[10] H. Rifà-Pous, M. J. Blasco, and C. Garrigues, "Review of robust cooperative spectrum sensing techniques for cognitive radio networks," *Wirel. Pers. Commun.*, vol. 67, no. 2, pp. 175–198, Nov. 2012. [Online]. Available: http://dx.doi.org/10.1007/s11277-011-0372-x

[11] S. Marano, V. Matta, and L. Tong, "Distributed detection in the presence of byzantine attacks," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 16 –29, Jan. 2009.

[12] A. Rawat, P. Anand, H. Chen, and P. Varshney, "Collaborative spectrum sensing in the presence of byzantine attacks in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 774 –786, Feb 2011.

[13] B. Kailkhura, S. Brahma, and P. K. Varshney, "Optimal byzantine attack on distributed detection in tree based topologies," in *Proc. International Conference on Computing, Networking and Communications Workshops (ICNC-2013)*, San Diego, CA, January 2013, pp. 227–231.

[14] B. Kailkhura, S. Brahma, Y. S. Han, and P. K. Varshney, "Optimal distributed detection in the presence of byzantines," in *Proc. The 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2013)*, Vancouver, Canada, May 2013.

[15] A. Vempaty, K. Agrawal, H. Chen, and P. K. Varshney, "Adaptive learning of byzantines' behavior in cooperative spectrum sensing," in *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC)*, march 2011, pp. 1310 –1315.

[16] B. Kailkhura, S. Brahma, Y. S. Han, and P. K. Varshney, "Distributed Detection in Tree Topologies With Byzantines," *IEEE Trans. Signal Process.*, vol. 62, pp. 3208–3219, June 2014.

[17] R. Chen, J.-M. Park, and K. Bian, "Robust distributed spectrum sensing in cognitive radio networks," in *Proc. 27th Conf. Comput. Commun., Phoenix, AZ*, 2008, pp. 1876–1884.

[18] E. Soltanmohammadi, M. Orooji, and M. Naraghi-Pour, "Decentralized hypothesis testing in wireless sensor networks in the presence of misbehaving nodes," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 1, pp. 205–215, 2013.

[19] B. Chen, L. Tong, and P. Varshney, "Channel-aware distributed detection in wireless sensor networks," *Signal Processing Magazine, IEEE*, vol. 23, no. 4, pp. 16–26, July 2006.

[20] X. Luo, M. Dong, and Y. Huang, "On distributed fault-tolerant detection in wireless sensor networks," *Computers, IEEE Transactions on*, vol. 55, no. 1, pp. 58–70, Jan 2006.

[21] T.-Y. Wang, L.-Y. Chang, D.-R. Duh, and J.-Y. Wu, "Distributed fault-tolerant detection via sensor fault detection in sensor networks," in *Information Fusion, 2007 10th International Conference on*, July 2007, pp. 1–6.

[22] T. Clouqueur, K. Saluja, and P. Ramanathan, "Fault tolerance in collaborative sensor networks for target detection," *Computers, IEEE Transactions on*, vol. 53, no. 3, pp. 320–333, Mar 2004.

[23] B. Kailkhura, Y. Han, S. Brahma, and P. Varshney, "On Covert Data Falsification Attacks on Distributed Detection Systems," in *Communications and Information Technologies (ISCIT), 2013 13th International Symposium on*, Sept 2013, pp. 412–417.

[24] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors*," *Math. control, Signals, and Systems*, vol. 1, pp. 167–182, 1988.

[25] B. Kailkhura, Y. S. Han, S. Brahma, and P. K. Varshney, "Asymptotic Analysis of Distributed Bayesian Detection with Byzantine Data," *CoRR*, vol. abs/1408.3434, 2014. [Online]. Available: http://arxiv.org/abs/1408.3434

[26] W. Shi, T. W. Sun, and R. D. Wesel, "Optimal binary distributed detection," in *Proc. The 33rd Asilomar Conference on Signals, Systems, and Computers*, 1999, pp. 24–27.

[27] Q. Zhang, P. Varshney, and R. Wesel, "Optimal bi-level quantization of i.i.d. sensor observations for binary hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 2105 –2111, jul 2002.

Bhavya Kailkhura (S'12) received the M.S. degree in electrical engineering from Syracuse University, Syracuse, NY. Since 2012, he has been pursuing the Ph.D. degree with the Department of Electrical Engineering and Computer Science, Syracuse University. His research interests include high-dimensional data analysis, signal processing, machine learning and their applications to solve inference problems with security and privacy constraints.

Yunghsiang S. Han (S'90-M'93-SM'08-F'11) was born in Taipei, Taiwan, 1962. He received B.Sc. and M.Sc. degrees in electrical engineering from the National Tsing Hua University, Hsinchu, Taiwan, in 1984 and 1986, respectively, and a Ph.D. degree from the School of Computer and Information Science, Syracuse University, Syracuse, NY, in 1993. He was from 1986 to 1988 a lecturer at Ming-Hsin Engineering College, Hsinchu, Taiwan. He was a teaching assistant from 1989 to 1992, and a research associate in the School of Computer and Information Science, Syracuse University from 1992 to 1993. He was, from 1993 to 1997, an Associate Professor in the Department of Electronic Engineering at Hua Fan College of Humanities and Technology, Taipei Hsien, Taiwan. He was with the Department of Computer Science and Information Engineering at National Chi Nan University, Nantou, Taiwan from 1997 to 2004. He was promoted to Professor in 1998. He was a visiting scholar in the Department of Electrical Engineering at University of Hawaii at Manoa, HI from June to October 2001, the SUPRIA visiting research scholar in the Department of Electrical Engineering and Computer Science and CASE center at Syracuse University, NY from September 2002 to January 2004 and July 2012 to June 2013, and the visiting scholar in the Department of Electrical and Computer Engineering at University of Texas at Austin, TX from August 2008 to June 2009. He was with the Graduate Institute of Communication Engineering at National Taipei University, Taipei, Taiwan from August 2004 to July 2010. From August 2010, he is with the Department of Electrical Engineering at National Taiwan University of Science and Technology as Chair Professor. He is also a Chair Professor at National Taipei University from February 2015. His research interests are in error-control coding, wireless networks, and security.

Dr. Han was a winner of the 1994 Syracuse University Doctoral Prize and a Fellow of IEEE. One of his papers won the prestigious 2013 ACM CCS Test-of-Time Award in cybersecurity.

Swastik Brahma (S'09-M'14) received the B.Tech degree in Computer Science and Engineering from West Bengal University of Technology, Kolkata, India in 2005. He received the M.S. and Ph.D. degrees in Computer Science from the University of Central Florida, Orlando, FL, in 2008 and 2011 respectively. In 2011, he joined the Department of Electrical Engineering and Computer Science (EECS) at Syracuse University, Syracuse, NY, as a Research Associate, where he is currently a Research Scientist. His research interests include communication systems, detection and estimation theory, cyber-security, and game theory. He serves as a technical program committee member for several conferences. Dr. Brahma won the best paper award in the IEEE Globecom conference in 2008, and was the recipient of the best Ph.D. forum award in the IEEE WoWMoM conference in 2009.

Pramod K. Varshney (S'72-M'77-SM'82-F'97) was born in Allahabad, India, on July 1, 1952. He received the B.S. degree in electrical engineering and computer science (with highest honors), and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign in 1972, 1974, and 1976 respectively.

From 1972 to 1976, he held teaching and research assistantships with the University of Illinois. Since 1976, he has been with Syracuse University, Syracuse, NY, where he is currently a Distinguished Professor of Electrical Engineering and Computer Science and the Director of CASE: Center for Advanced Systems and Engineering. He served as the Associate Chair of the department from 1993 to 1996. He is also an Adjunct Professor of Radiology at Upstate Medical University, Syracuse. His current research interests are in distributed sensor networks and data fusion, detection and estimation theory, wireless communications, image processing, radar signal processing, and remote sensing. He has published extensively. He is the author of Distributed Detection and Data Fusion (New York: Springer-Verlag, 1997). He has served as a consultant to several major companies.

Dr. Varshney was a James Scholar, a Bronze Tablet Senior, and a Fellow while at the University of Illinois. He is a member of Tau Beta Pi and is the recipient of the 1981 ASEE Dow Outstanding Young Faculty Award. He was elected to the grade of Fellow of the IEEE in 1997 for his contributions in the area of distributed detection and data fusion. He was the Guest Editor of the Special Issue on Data Fusion of the IEEE PROCEEDINGS January 1997. In 2000, he received the Third Millennium Medal from the IEEE and Chancellor's Citation for exceptional academic achievement at Syracuse University. He is the recipient of the IEEE 2012 Judith A. Resnik Award and Doctor of Engineering degree honoris causa from Drexel University in 2014. He is on the Editorial Boards of the Journal on Advances in Information Fusion and IEEE Signal Processing Magazine. He was the President of International Society of Information Fusion during 2001.