Aditya Vempaty, Lang Tong, and Pramod K. Varshney

# Distributed Inference with Byzantine Data

## State-of-the-art review on data falsification attacks




In 1982, Lamport et al. presented the so-called *Byzantine generals problem* as follows [1]: "a group of generals of the Byzantine army camped with their troops around an enemy city. Communicating only by messenger, the generals must agree upon a common battle plan. However, one or more of them may be traitors who will try to confuse the others. The problem is to find an algorithm to ensure that the loyal generals will reach agreement." The authors gave a sharp characterization of the power of the Byzantine generals. It was shown that if the fraction of Byzantine generals is less than 1/3, there is a way for the loyal generals to reach a consensus agreement, regardless of what the Byzantine generals do. If the fraction is above 1/3, consensus can no longer be guaranteed.

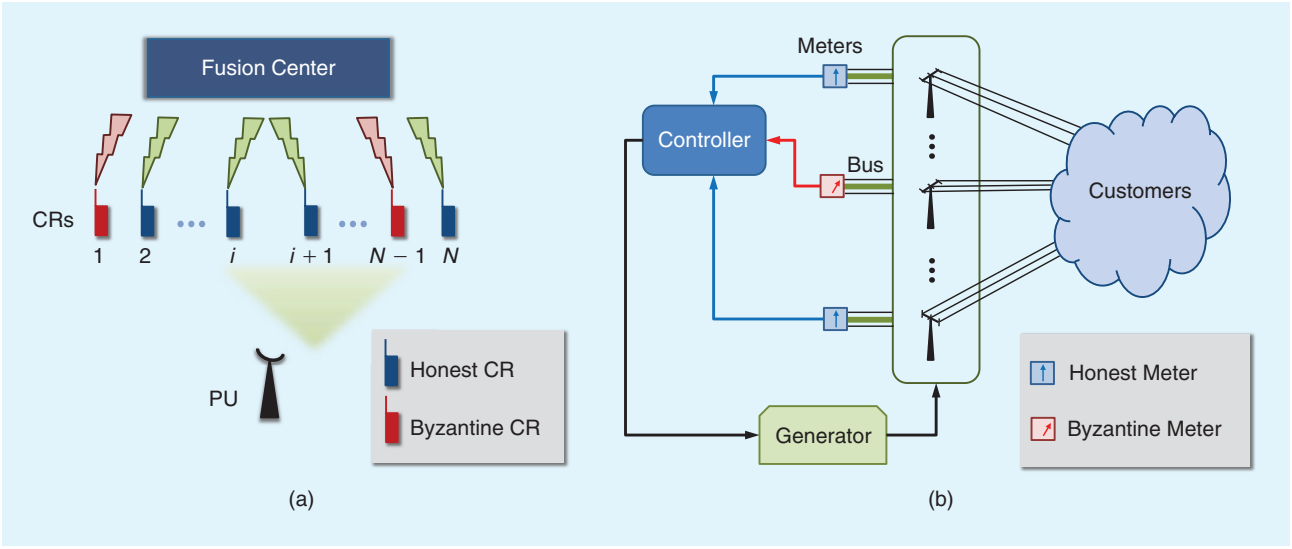
It is not difficult to relate the Byzantine generals problem to a variety of applications in cybersecurity, where Byzantine generals play the role of internal adversaries. There are many diverse behaviors that a Byzantine entity may engage in, such as a node (or sensor) may lie about connectivity, flood network with false traffic, attempt to subjugate control information, falsely describe opinions of another node (e.g., peer to peer), or capture a strategic subset of devices and collude. This article examines the Byzantine generals problem in the context of distributed inference [2], where data collected from remote locations are sent to a fusion center (FC) for processing and inference. The assumption is that the data are potentially tampered or falsified by some internal adversary who has the knowledge about the algorithm used at the FC. We refer to the problem considered in this tutorial article as distributed inference with Byzantine data.

The Byzantine data problem in statistical inference has to take into account the inherent randomness in the data. Even without the presence of an adversary, one cannot expect perfect inference; detections can at best be correct in probability as parameter estimates almost always are not equal to the true value. Therefore, there is a need for a probabilistic approach to the Byzantine data problem.

We shall focus on the two most basic forms of statistical inference: detection and estimation. Our objective is to introduce relevant problem formulations for each problem and present some related applications.

As an application to distributed detection with Byzantine data, we consider the problem of distributed spectrum sensing when some participants attack a cognitive radio network (CRN)

*Digital Object Identifier 10.1109/MSP.2013.2262116*
*Date of publication: 20 August 2013*




**[FIG1]** **Byzantine attacks on (a) distributed spectrum sensing in CRNs and (b) system state estimation in smart grids.**

by sending falsified data to the FC. In this context, Byzantine users [marked as red-colored CRs in Figure 1(a)] can affect decisions at the FC by reporting false data. This might result in a collision of secondary users with the primary user (PU) (if a busy PU is wrongly detected as idle) or in spectrum wastage (if an idle PU is detected as busy).

For distributed estimation, we consider the impact of Byzantine data on the state estimation of a power grid as illustrated in Figure 1(b). Here an adversary takes control of the "red" meters and launches a man-in-the-middle (MiM) attack by substituting actual measurements with falsified data. If undetected, state estimates at the FC will be altered and subsequent decisions using state estimates are affected.

## DISTRIBUTED DETECTION WITH BYZANTINE DATA

### *CHARACTERIZING EFFECTS OF BYZANTINE DATA*

We first consider the classical problem of distributed detection in a scenario where a fraction of the nodes may have been reprogrammed by an adversary. These adversary-controlled nodes may collaborate to mislead the FC by sending Byzantine data to the FC, causing increased detection error probabilities. The problem is similar to the classical Byzantine generals problem but has the following two important distinctions:

1) the information flow is from sensors to only the FC
2) the FC is the sole decision maker.

Thus, we need to assume that the FC is always honest.

One problem of interest is to characterize the degree to which the Byzantines can affect the detection performance of the FC. The parallel result in the classical Byzantine generals problem is that at most 1/3 of the generals can be Byzantine for possible consensus among loyal generals. Here we are interested in the minimum fraction $\alpha^*$ of Byzantine sensors that makes the detection at the FC no better than merely flipping a coin without using any data. We call $\alpha$ the attack power of the Byzantine sensors.

Should the attack power be more than 1/3 because we are only interested in the decision at the FC or should it be less because the presence of randomness in data makes it easier for Byzantine sensors to disguise their actions?

If the fraction of Byzantines is greater than 1/2, i.e., $\alpha \geq 1/2$, it is evident that sensor observations can be easily made useless. This is because the Byzantines may have 1/2 of the total sensors send false samples based on an incorrect hypothesis to the FC and have the rest of the Byzantines report the observed samples truthfully to the FC. Consequently, half of the samples are from one hypothesis and the other half from another. Therefore, the FC cannot use the sensors' data to make a final decision.

What happens when the attack power of the adversary is less than and equal to 1/2? Here we are interested in two related questions:

- What is the minimum power of the adversary to render sensor data useless to the FC?
- If the power of the adversary is less than this critical value, what should be the detection rule at the FC and to what degree is the performance affected?

While these questions are difficult to answer in general, some insights can be obtained by examining the classical binary distributed detection problem that we discuss next; see [3] for more details.

### HYPOTHESIS AND ATTACK MODELS

A classical distributed detection system consists of multiple remotely located sensors that observe a common phenomenon. Some data processing is carried out at the peripheral detectors and processed information is sent to a central unit that fuses this information to make a decision regarding the presence ($\mathcal{H}_1$) or absence ($\mathcal{H}_0$) of the phenomenon.

Suppose that, if a sensor is honest, its observation follows the conditional distribution $p$ under $\mathcal{H}_0$ and $q$ under $\mathcal{H}_1$. If a sensor is Byzantine, it generates false data with distribution $\tilde{p}$

under $\mathcal{H}_0$ and $\tilde{q}$ under $\mathcal{H}_1$. The FC, of course, does not know the attack distributions $(\tilde{p}, \tilde{q})$, nor does it know which sensor is a Byzantine.

If the FC knows that the maximum fraction of Byzantine sensors is $\alpha$, then the distributions of the data from a given sensor should come from a restricted set of possibilities, $\mathcal{F}(p;\alpha)$ under $\mathcal{H}_0$ and $\mathcal{F}(q;\alpha)$ under $\mathcal{H}_1$ where $\mathcal{F}(p;\alpha) \triangleq \{f: f = (1-\alpha)p + \alpha\tilde{p}\}$. Note that in the above model, $\tilde{p}$ is not fixed and, therefore, $\mathcal{F}(p;\alpha)$ does not just contain one distribution of $\tilde{p}$. In fact, it includes cases when Byzantines may use different attacking distributions, in which case $\tilde{p}$ and $\tilde{q}$ are composite distributions from all Byzantines.

## FC VERSUS BYZANTINE SENSORS

From the perspective of the FC, the problem can be viewed as one of robust detection; the data from a sensor is a mixture of good and attack distributions. In his seminal paper [4], Huber showed the striking result that the optimal detector, in the sense of minimizing the worst missed detection probability among all possible adversary attacking distributions, is a likelihood ratio test based on a pair of least favorable distributions. Intuitively, the least favorable distributions are a pair of distributions $f_0^* \in \mathcal{F}(p;\alpha)$ and $f_1^* \in \mathcal{F}(q;\alpha)$ such that they are most difficult to distinguish, resulting in the highest probabilities of error.

From the adversary's perspective, launching effective attacks requires a careful design of the attack distributions $\tilde{p}$ and $\tilde{q}$. In his paper [4], Huber identified the specific form of the least favorable distributions. The detection performance at the FC, however, cannot be evaluated easily.

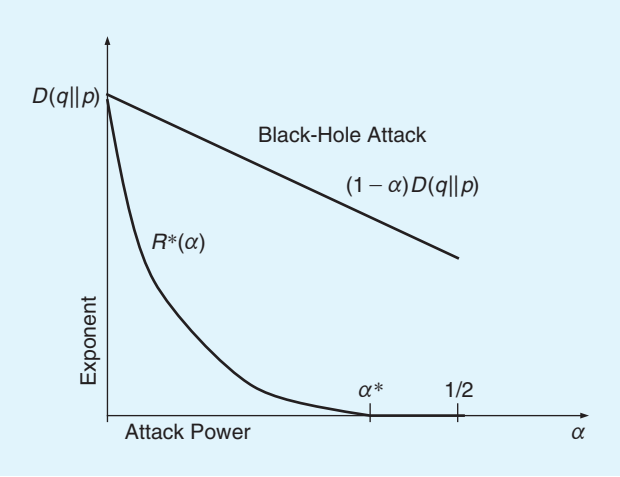
To gain insights into the degree to which an adversary can cause performance degradation, we can consider the asymptotic regime when the number of sensors $N$ is large. The large deviation analysis allows us to approximate the missed detection error probability by an exponential form. In particular, if the adversary uses attack distributions $\tilde{p}$ and $\tilde{q}$ (and assume that they are known to the FC), the optimal likelihood ratio detector will have the probability of missed detection that decays exponentially

$$P_{\text{miss}} = e^{(-R(\alpha, \tilde{p}, \tilde{q})N + O(\log N))},$$

where the rate of exponential decay of $P_{\text{miss}}$ is given by the Kullback-Leibler divergence (KLD) [5]. Thus, the adversary can use the rate $R(\alpha, \tilde{p}, \tilde{q})$ as a proxy to maximize its impact on the detection performance of the FC.

## CRITICAL POWER OF BYZANTINE SENSORS

We can now provide a characterization of the attack power in the asymptotic regime. Specifically, given the fraction $\alpha$ of Byzantine sensors, the adversary can choose the attack distributions $(\tilde{p}^*, \tilde{q}^*)$ from the optimization $R^*(\alpha) \triangleq R(\alpha, \tilde{p}^*, \tilde{q}^*) = \min_{\tilde{p}, \tilde{q}} R(\alpha, \tilde{p}, \tilde{q})$. Not surprisingly, perhaps, that the optimal attacking distributions $(\tilde{p}^*, \tilde{q}^*)$ are in fact Huber's least favorable distributions, and there is a "water-filling" interpretation for the construction of the optimal attack distributions [3].




**[FIG2] The optimized rate function $R^*(\alpha)$, which reaches zero at $\alpha = \alpha^*$. Note that the black-hole strategy corresponds to a straight line above $R^*(\alpha)$ and, therefore, is not effective.**

The most potent attack by the Byzantine sensors is to make the decay rate zero. In this case, it necessarily means that Byzantine sensors can make the two hypotheses indistinguishable, rendering the decision at the FC no better than flipping a biased coin. As in the classical Byzantine generals problem, the smallest $\alpha$ that makes $R^*(\alpha) = 0$ represents the critical power of the Byzantine sensors. Specifically, the critical power $\alpha^*$ is given by $\alpha^* = \min\{\alpha: R^*(\alpha) = 0\}$.

Figure 2 illustrates the general shape of the optimized rate function $R^*(\alpha)$. The function $R^*(\alpha)$ can be shown to be convex and monotonically decreasing, reaching zero at the critical power $\alpha^* \leq 1/2$. In addition, the black-hole strategy under which Byzantines simply do not send their measurements is not at all effective due to the convexity of $R^*(\alpha)$. It can be shown that when each sensor reports multiple independent observations, the gap between $\alpha^*$ and 1/2 shrinks to zero, which means that unless more than half of the sensors are made Byzantine, asymptotically, the FC can provide reliable detection.

## *COLLABORATIVE SPECTRUM SENSING*

An important recent application of distributed detection is the idea of dynamic spectrum access (DSA) using CRNs. In DSA, the spectrum is allocated to a licensed PU while the unlicensed secondary users, which are incorporated in the system as the cognitive radios (CRs), have the capability to sense the spectrum for availability. The CRs can transmit their data if PU is absent. To mitigate the effect of channels or hidden terminal problem on the process of spectrum sensing, collaborative spectrum sensing (CSS) has been proposed. CSS works on the parallel data fusion model of distributed detection where CRs transmit their decisions regarding the availability of the spectrum to an FC in parallel.

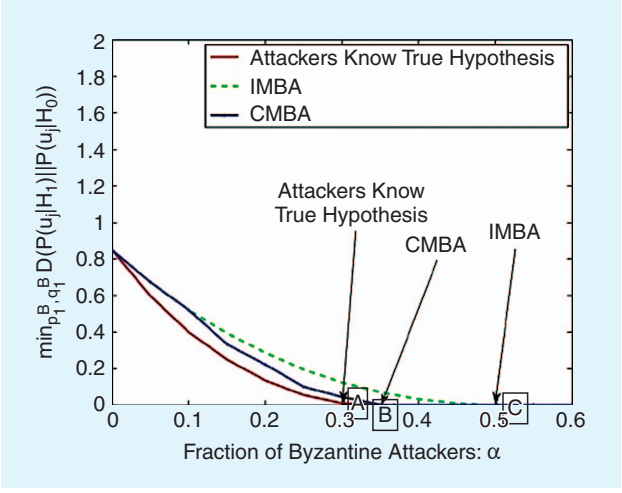
Rawat et al. in [6] consider the problem of Byzantines in CSS and analyze their effect using both KLD and probability of error ($P_e$) as the performance metrics. They have generalized the results presented by Marano et al. in [3] by relaxing the

assumption that the Byzantines have perfect knowledge about the true hypothesis. The critical power ($\alpha^*$) to "blind" the FC has been shown to be

$$\alpha^* = \frac{P_d^H - P_{fa}^H}{(P_d^B - P_{fa}^B) + (P_d^H - P_{fa}^H)}, \tag{1}$$

where $(P_{fa}^H, P_d^H)$ and $(P_{fa}^B, P_d^B)$ are the operating points on the receiver operating characteristics (ROC) of the honest sensors and Byzantines, respectively. Note that $P_d$ is the probability of detection and $P_{fa}$ is the probability of false alarm. The optimal strategy for the Byzantines is to flip their local decisions with probability "1" before transmitting to the FC. This results in the fewest number of Byzantines to blind the FC.

There are two types of Byzantine attacks [6] that involve two extremes of cooperation among Byzantine nodes: independent malicious Byzantine attacks (IMBA) and cooperative malicious Byzantine attacks (CMBA). In an independent attack, each Byzantine attacks the network independently relying on its own observation. Since the Byzantines do not know the identity of other Byzantines in the network, $P_d^B = P_d^H$ and $P_{fa}^B = P_{fa}^H$, which gives $\alpha^* = 0.5$ from (1). This implies that the number of Byzantines need to be at least 50% to blind the FC when the Byzantines attack the network independently. In a cooperative attack, Byzantines collaborate to make a decision regarding the true hypothesis and use this information to attack the network. Using collaboration, the Byzantines can reduce the minimum critical power $\alpha^*$ by increasing $(P_d^B - P_{fa}^B)$. The Byzantines are assumed to collude using the "$L$ out of $M$" rule to make their decision. In other words, if more than $L$ out of the $M$ Byzantines make a decision, say "1," then all the collaborating Byzantines in the network send a "0." The value of $L$ is often taken to be $M/2$, which corresponds to the majority rule. Figure 3 shows a plot of KLD against the fraction of Byzantines in the network. As the figure shows, $\alpha^*$ decreases with collaboration of the Byzantines. Note that one could also define




**[FIG3] The detection performance in terms of KLD as a function of the fraction of Byzantines for different attack strategies. (Figure used with permission from [6].)**

a partially cooperative malicious Byzantine attack (PCMBA), where some of the Byzantines collude while the others attack independently. Such an attack would be more severe than IMBA but less effective than CMBA.

Since Byzantines operate in an adversarial manner, one can formulate the problem as a game and find the best strategies for the Byzantines and the FC when the Byzantines cannot blindthe FC ($\alpha < \alpha^*$). The game can be represented as $\mathcal{G} = \{\mathcal{N}, \mathcal{S}, \mathcal{U}\}$, where $\mathcal{N}$ is a set of two players who are playing the game—Byzantines and the FC, $\mathcal{S}$ is the set of strategies of each player which in this case are the local decision thresholds ($\eta$) to be used, and $\mathcal{U}$ is the set of utilities of the players. The two possible utility functions are KLD and probability of error ($P_e$). The problem is a zero-sum game, and one can find the Nash-equilibrium as the saddle point of a minimax game, which is given by [6]

$$(\eta_h^*, \eta_b^*) = \min_{\eta_h} \max_{\eta_b} f(\eta_h, \eta_b), \tag{2}$$

where $f(\cdot, \cdot)$ is either $P_e$ or $-$KLD and $\eta_h, \eta_b$ are the local thresholds of honest sensors and Byzantines, respectively. Saddle-points can be determined numerically for both the performance metrics under both independent and cooperative attacks.

## *MITIGATION STRATEGIES*

The previous discussion addresses the issue of Byzantines from the attacker's perspective and the optimal attacking strategies are derived for distributed detection. However, one needs to look at the possible countermeasures from the network's perspective to protect the network from these Byzantines. Byzantines can be treated as outliers and, therefore, one can use signal processing techniques to mitigate their effects. Table 1 presents a comparison of the mitigation schemes proposed in the literature.

A simple and intuitive method to mitigate the effect of Byzantines is to identify them. For identification purposes, one needs to observe the sensors' behavior over time. We first discuss some of the schemes proposed in the literature that treat the FC as a watchdog to mitigate the effect of Byzantines.

### REPUTATION-BASED SCHEME

Rawat et al. in [6] have proposed a simple and effective scheme to identify the Byzantines. They define a reputation metric $\kappa_i$ for every sensor as the number of mismatches in a time interval $T$ between $i$th sensor's local decision and the global decision made at the FC using the majority rule. The local sensors sense the spectrum and transmit new data at every time instant $t$. The reputation metric after a time interval $T$ is given by $\kappa_i = \sum_{t=1}^{T} \mathcal{I}_{(u_i[t] \neq u_0[t])}$, where $u_i[t]$ is the $i$th sensor's local decision at time instant $t$, $u_0[t]$ is the global decision made at the FC at time instant $t$, and $\mathcal{I}_{(S)}$ is the indicator function over the set $S$. The sensors for which this reputation metric $\kappa_i$ is greater than a predetermined threshold $\kappa$ are tagged as Byzantines and removed from the fusion process. This scheme works only when the number of Byzantines in the network is less than 50% of the

**[TABLE 1] COMPARISON OF DIFFERENT BYZANTINE MITIGATION SCHEMES PROPOSED IN THE LITERATURE.**

| CONTRIBUTION | MITIGATION SCHEME | ADVANTAGES | DISADVANTAGES |
|---|---|---|---|
| **BYZANTINE IDENTIFICATION SCHEMES** | | | |
| RAWAT ET AL. [6] | REPUTATION-BASED SCHEME | SIMPLE AND EFFECTIVE | FAILS FOR $\alpha > 0.5$ |
| VEMPATY ET AL. [7] | ADAPTIVE LEARNING | WORKS FOR ANY $\alpha$ | REQUIRES HONEST SENSOR BEHAVIOR STATISTICS |
| HE ET AL. [8] | CONDITIONAL FREQUENCY CHECK | HIGH BYZANTINE SENSOR DETECTION ACCURACY | PRESENCE OF ONE TRUSTED SENSOR REQUIRED |
| **BYZANTINE TOLERANT SCHEMES** | | | |
| CHEN ET AL. [9] | WEIGHTED SEQUENTIAL PROBABILITY RATIO TEST | ROBUST DETECTOR DESIGN | MIGHT FAIL FOR $\alpha > 0.5$ |
| GAGRANI ET AL. [10] | STOCHASTIC RESONANCE | INCREASES $\alpha^*$ | BYZANTINE IDENTIFICATION NEEDED |

total number of sensors since the FC uses majority rule for fusion. If the Byzantines are in majority, the above reputation-based scheme identifies the honest sensors as outliers and removes them from the network and thereby worsens the inference performance of the network.

### ADAPTIVE LEARNING

An interesting method to improve the performance of the network is to use the information of the identified Byzantines to the network's benefit. Vempaty et al. in [7] use this idea to improve the network's performance. They propose a three-tier adaptive learning scheme that learns the parameters of the identified Byzantines and uses these parameters in the Chair-Varshney rule [11] to make the final decision. The three-tier scheme can be described as follows: 1) identification of Byzantines in the network, 2) estimation of parameters of the identified Byzantines at the FC, and 3) adaptation of fusion rule using the estimated parameters.

The basic idea of the proposed identification scheme is to compare every sensor's observed behavior over time with the expected behavior of an honest sensor. The sensors whose observed behavior is too far from the expected behavior are tagged as Byzantines. This scheme works even when the Byzantines are in the majority since it does not use the global decision for identification purposes. The behavior of every sensor is characterized by the probability of sending a "1" to the FC. This value is a function of the operating point of the sensor ($P_{fa}$, $P_d$) and the prior probabilities of the hypotheses ($\pi_0$, $\pi_1$), which are assumed to be known at the FC for honest sensors.

At the FC, the expected behavior is estimated for every sensor over time by averaging the number of times a particular decision (0 or 1) is made over a time interval of $T$ instants. These probabilities can be updated after every time instant. The test statistic $\Lambda_i^T$ for the $i$th sensor after time $T$ is the deviation between the expected and observed behavior for every sensor. The FC declares a sensor as a Byzantine if $\Lambda_i^T$ is greater than a particular threshold $\lambda$. This threshold $\lambda$ is determined as the minimum value when the Byzantine's operating point is in the region below the $P_d = P_{fa}$ line on the ROC.

After identifying the Byzantines, their parameters can be estimated by assuming that all the Byzantines have the same operating point. This assumption is typically made in the literature since it is assumed that a single adversary has attacked some of the sensors in the network and has reprogrammed them to behave as Byzantines. Therefore, it can be assumed that all these Byzantine sensors have the same operating point on the ROC. These estimated parameters are used in the Chair-Varshney optimal fusion rule [11] in an adaptive manner to find the global decision. It is important to note that this scheme works for any fraction of Byzantines in the network but assumes the knowledge of honest sensor's behavior and the PU statistics.

### CONDITIONAL FREQUENCY CHECK

In distributed spectrum sensing (DSS), there is usually memory in the spectrum state evolution and researchers fail to exploit this state evolution. He et al. in [8] model spectrum state transitions using a Markov model and propose a global decision independent method, conditional frequency check (CFC), to counter Byzantines in DSS in CRNs.

Let the true spectrum, which has two states: 0 (idle) and 1 (occupied), be modeled as a homogeneous Markov model with state transition matrix $A = [a_{ij}]$, where $a_{ij} = P[s_{t+1} = j \,|\, s_t = i]$ and $s_t$ denotes the spectrum state at time $t$. Also assume that the FC knows about the presence of one known trusted honest sensor and let $\phi = [\phi_{01}, \phi_{10}]$ denote the flipping probabilities of the local sensors where $\phi_{lm}$ is the probability of flipping the local decision from $l$ to $m$, $\{l, m\} \in \{0, 1\}$. A sensor is Byzantine if and only if $\{\phi_{01}, \phi_{10}\} \neq \{0, 0\}$. Using this idea, the authors propose a two-phase Byzantine sensor detection method: conditional frequency check and an auxiliary Hamming distance check.

In the CFC phase, the FC evaluates the two conditional frequency statistics $\psi_1 = P[r_t = 1 \,|\, r_{t-1} = 1]$ and $\psi_0 = P[r_t = 0 \,|\, r_{t-1} = 0]$ for every sensor individually where $r_t$ denotes the sensor's report at time $t$. These statistics are related to the system parameters: the Markov state transition matrix $A$ and the local sensor operating point ($P_{fa}$, $P_d$), which are assumed to be identical for all the sensors. Since these parameters are not known in

practice, they are found as an average over a window of time interval $T$ as follows:

$$\hat{\psi}_1 = \frac{\left(\sum_{t=1}^{T} \delta_{r_{t+1},1}\delta_{r_t,1}\right)}{\left(\sum_{t=1}^{T-1} \delta_{r_t,1}\right)}, \tag{3}$$

$$\hat{\psi}_0 = \frac{\left(\sum_{t=1}^{T} \delta_{r_{t+1},0}\delta_{r_t,0}\right)}{\left(\sum_{t=1}^{T-1} \delta_{r_t,0}\right)}, \tag{4}$$

where $\delta_{i,j} = 1$ if and only if $i = j$. These histogram estimators converge to the true values as $T \to \infty$.

The Byzantine users can be detected after time $T$ using the error function $e(\phi) = \| \psi^{(tr)} - \psi^{(M)} \|_2$ where $\| \cdot \|_2$ is the two-norm (Euclidean distance), $\psi^{(tr)} = [\psi_1^{(tr)}, \psi_0^{(tr)}]$ and $\psi^{(M)} = [\psi_1^{(M)}, \psi_0^{(M)}]$ are the CFC statistics of the trusted sensor and the Byzantine sensor, respectively. If the error function value is greater than a predetermined threshold $\beta_{\text{CFC}}$, the sensor is declared as Byzantine. However, the CFC approach fails when the flipping probabilities of the sensor are $\phi = [1, 1]$. To address this concern, the sensors identified as honest go through a second phase: the Hamming distance check, where a normalized Hamming distance between the sensor $k$ and the trusted sensor, defined as $d_h(k, tr) = (1/T)\sum_{t=1}^{T} \delta_{r_t^{(k)}, r_t^{(tr)}}$ is compared against a prespecified threshold $\beta_{\text{HDC}}$. In this manner, a high Byzantine user detection accuracy can be achieved using a trusted sensor without relying on the global decision.

### WEIGHTED SEQUENTIAL RATIO PROBABILITY TEST

The schemes discussed above focus on identifying the malicious sensors/Byzantines in the network. One may also consider approaches that make the system more robust to adversarial actions. Chen et al. [9] consider detector design at the FC, which is robust to the Byzantine attacks for DSS in CRNs. They propose a weighted sequential probability ratio test (WSPRT) for data fusion at the FC. In DSS, one wants to control both the false alarm and missed detection probabilities, so the sequential probability ratio test (SPRT) can be used for data fusion at the FC. In a nonadversarial environment, the SPRT for data fusion is defined as

$$S_n = \prod_{i=1}^{n} \frac{P[u_i \,|\, H_1]}{P[u_i \,|\, H_0]}, \tag{5}$$

where $S_n$ is the test statistic after $n$ measurements and $u_i \in \{0, 1\}$ is the local decision of the secondary user. Also, note that $n$, which is the variable number of samples, is different from the total number of sensors in the network. The final fusion decision is based on the following criterion:

$$\begin{cases} S_n \geq \eta_1 \Rightarrow \text{accept } H_1, \\ S_n \leq \eta_0 \Rightarrow \text{accept } H_0, \\ \eta_0 < S_n < \eta_1 \Rightarrow \text{take another local measurement.} \end{cases} \tag{6}$$

The values of the thresholds are a function of the tolerated false alarm and missed detection probabilities.

Based on the SPRT, the authors propose a two-step WSPRT. In the first step, called the reputation maintenance step, a reputation value $r_i$ is allocated to the local sensors based on their consistency with the final global decision made by the FC. The initial reputation value is zero for all sensors and is either increased or decreased based on its consistency with the final global decision. In the second step, the hypothesis test is performed at the FC based on modified SPRT [cf. (5)] called WSPRT

$$W_n = \prod_{i=1}^{n} \left(\frac{P[u_i \,|\, H_1]}{P[u_i \,|\, H_0]}\right)^{w_i}, \tag{7}$$

where $w_i$ is the weight given to the $i$th measurement and is given as $w_i = f(r_i)$. The authors in [9] analyze the properties required by the function $f(\cdot)$ and propose the following function:

$$w_i = f(r_i) = \begin{cases} 0 & r_i \leq -g, \\ \dfrac{r_i + g}{\max r_i + g} & r_i > -g, \end{cases} \tag{8}$$

where the positive variable $g$ is chosen to ensure that a "good" sensing terminal is not severely penalized when it sends incorrect data only due to randomness.

Based on simulations, it has been shown that the WSPRT design [cf. (7)] improves the robustness of data fusion against Byzantines as compared to the traditional SPRT since it includes weights given to the data. These weights are a function of the reputation and, therefore, make the detector robust to Byzantines. However, the authors have also pointed out that the proposed WSPRT technique requires an increased number of local sensing reports for improved robustness. Also, since WSPRT depends on $r_i$, which is based on the final global decision, the detector might fail if the Byzantines are in majority.

### NOISE-ENHANCED SIGNAL PROCESSING

In [10], Gagrani et al. propose the use of stochastic resonance (SR) to make the network robust against Byzantines. SR [12] is a counterintuitive physical phenomenon, where the performance of some suboptimal, nonlinear systems can be improved by adding suitable noise to the input. Gagrani et al. look at the possible improvement in system performance in terms of two metrics: security metric ($\alpha^*$) and detection performance metric (KLD). When local sensors use threshold quantizers such as energy detectors that are nonlinear and suboptimal, the use of SR noise at the local sensors can improve the local sensors' detection performance. This is achieved by randomization induced by the use of optimal additive SR noise as discussed in more detail in [10].

## *OTHER DISTRIBUTED DETECTION FRAMEWORKS*

In this section, we briefly discuss some other detection frameworks in the presence of Byzantines. The goal here is to expose the reader to possibilities of performance metrics beyond the probability of error related metrics and to network topologies other than the parallel topology.

## FALSE DISCOVERY RATE-BASED DISTRIBUTED DETECTION

Recently, Ray and Varshney in [13] have proposed the use of false discovery rate (FDR) to design distributed detection systems. Traditionally, multiple comparison procedures and classical comparison procedures control the family wide error rate (FWER). From Table 2, FWER $\beta_F$ for $k$ tests is

$$\beta_F = P(F \geq 1) = 1 - (1-\beta)^k, \tag{9}$$

where $F$ is the total number of false alarms.

A different and more liberal detection approach controls the FDR, defined as the fraction of false rejections among those hypotheses rejected. Formally, FDR is defined as the expected ratio of the number of false alarms (declared $H_1$ when $H_0$ is true) to the total number of detections ($H_1$ declarations consisting of both true and false detections). From Table 2, the ratio of false alarms to the total number of detections can be viewed as the random variable,

$$Q = \begin{cases} \dfrac{F}{F+S} & \text{if } F+S \neq 0, \\ 0 & \text{if } F+S = 0. \end{cases} \tag{10}$$

FDR ($Q_e$) is defined to be the expectation of $Q$,

$$Q_e = E(Q). \tag{11}$$

Ray and Varshney [13] have proposed a distributed algorithm to control the FDR value at $\gamma$, which uses linearly increasing thresholds on the ordered p-values of the local sensor observations. The p-value is defined as $p = \int_s^\infty f_0(t)\,dt$, where, $f_0(t)$ is the pdf of the observation under $H_0$ and $s$ is the observation at the local sensor.

Vempaty et al. [14] have analyzed the system in the presence of Byzantines. The Byzantines are modeled so that they employ a flipping strategy based on false p-values as $q = h(p) = 1-p$ and send binary quantized data based on the transformed p-values. This transformation ensures that the Byzantines attack the network in a covert manner. They have shown with analytical and numerical results that this attack strategy reduces the detection performance of the network while controlling the FDR value at the predetermined threshold $\gamma$. Hence, the attack strategy $h(p) = 1-p$ allows the Byzantines to reduce the network performance while not changing their behavior in a manner that they can be detected by observers such as the FC. To improve system performance, they also proposed a learning based adaptive system design approach which estimates the fraction of Byzantines ($\alpha$) in the network and adaptively changes the system design parameter to improve the system detection performance. Figure 4 shows the improvement in detection performance when an adaptive scheme is used against a nonadaptive scheme.

## DISTRIBUTED DETECTION IN TREE-BASED TOPOLOGIES

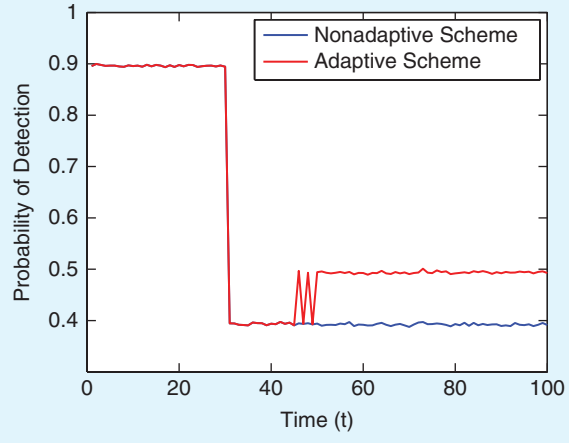
In a scenario when the FC may be outside the communication range of local sensors, a multihop network, where sensors are organized hierarchically into multiple levels (tree networks) are needed. Kailkhura et al. in [15] examine tree-topology-based distributed detection in the presence of Byzantines and analyze the susceptibility of such tree networks to Byzantines. The sensors in the network act as relays and forward the decisions of their children sensors to the FC. The Byzantines flip all their children's decisions and due to the tree topology, there are more parameters for analysis than the parallel fusion network. The authors come up with the condition on these parameters that blind the FC and make it incapable of making a decision. They have shown that 50% or more sensors in the network need to be covered by Byzantines to blind the FC where the covered sensors for a sensor $i$ are defined as the sensors in the subtree formed with sensor $i$ as the root.

**[TABLE 2] NOTATIONS TO DEFINE FDR.**

| | DECLARED $H_0$ | DECLARED $H_1$ | TOTAL |
|---|---|---|---|
| $H_0$ TRUE | W | F | $N_0$ |
| $H_1$ TRUE | T | S | $N-N_0$ |
| **TOTAL** | $N-R$ | R | N |

Due to the tree nature of the topology, there are many ways to have the same impact in terms of inference performance degradation. If the cost of each sensor being a Byzantine is different, there are tradeoffs that one can study. The authors formulate a minimum Byzantine attack problem as a bounded Knapsack problem. Let there be $K$ levels in the tree with $N_k$ sensors at each level $k = 1, \ldots, K$. The minimum Byzantine attack is to determine $B_k$ for $k = 1, \ldots, K$ such that at least 50% of the sensors in the network are covered and the total cost incurred is minimized. Hence, the bounded knapsack optimization problem is

$$\begin{aligned} \underset{\{B_k\}_{k=1}^K}{\text{minimize}} \quad & \sum_{k=1}^{K} c_k B_k \\ \text{subject to} \quad & \sum_{k=1}^{K} P_k B_k \geq \sum_{k=1}^{K} \frac{N_k}{2} \\ & B_k \text{ is an integer with } 0 \leq B_k \leq N_k, \quad \forall k, \end{aligned} \tag{12}$$




**[FIG4] The probability of detection versus time when $\alpha$ changes from 0 to 0.7 at $t = 30$. (Figure used with permission from [14].)**

where $c_k$ is the cost associated with attacking each sensor of level $k$ and $P_k$ is the profit associated with attacking each sensor of level $k$, which is the number of children it covers. This problem is NP-hard in general, but by exploiting the specific cost and profit structure, the authors propose a polynomial time algorithm for finding the optimal Byzantine attack.

**A KEY INSIGHT IS TO CONNECT THE PROBLEM OF DATA FALSIFICATION ATTACK WITH THE CLASSICAL NOTION OF OBSERVABILITY.**

## DISTRIBUTED ESTIMATION WITH BYZANTINE DATA

We now consider the problem of distributed estimation with Byzantine data. As before, we assume that a fraction of sensors are controlled by an adversary who substitutes measurements at these sensors by data designed to perturb the estimate at the FC. Such an attack is also referred to as the MiM attack.

Similar to the corresponding detection problem, we would like to quantify the critical power of Byzantine sensors. The classical Byzantine generals problem and distributed detection are decision problems on a finite action space. For the distributed estimation problem, the decision space is uncountable, which requires a different measure of performance. To this end, the critical power of a set of Byzantine sensors is characterized by the condition under which the presence of Byzantine sensor makes the estimation error arbitrarily large.

In this section, we discuss the effect of Byzantine data on state estimation, using power system state estimation as an example followed by a brief discussion on the problem of estimating the location of a target.

### *POWER SYSTEM STATE ESTIMATION WITH BYZANTINE DATA*

The power grid is an interconnections of large number of generators, loads, transmission lines, and transformers. It can be abstracted as a graph $G$ with vertices representing buses (substations) of the grid and edges representing the transmission lines that connect the buses.

The state of the power grid is defined by the vector $x$ of voltage phasors at the network buses. Efficient operation of the power grid depends critically on monitoring of the system state, which is accomplished by using measurements collected from sensors deployed throughout the network. Typically, sensor measurements include the real and reactive power injections and branch power flows. When phasor measurement units (PMUs) are used, the power system states are measured directly.

A fundamental property of a power grid is the instantaneous balance of power flow governed by the Kirchoff current and voltage laws. The steady state representation of the power flow is given by a nonlinear equation

$$z = h(x; G) + w, \tag{13}$$

where $z$ is a vector of power flows at branches of the power grid, $x$ the state vector, $G$ the graph representing the power grid, and $w$ the measurement noise.

A fundamental problem of power system operation is state estimation [16] in which the supervisory control and data acquisition (SCADA) system collects measurement data $z$ from a large geographical area and form a state estimate $\hat{x}$. A standard technique is the weighted least squares based on (13)

$$\hat{x} = \arg\min_x (z - h(x; G))^T W (z - h(x, G)),$$

where $W$ is the weighting matrix.

### BYZANTINE DATA MODEL

We model Byzantine data by adding an attack vector $a$

$$z = h(x, G) + a + w,$$

where $a \in \mathcal{A}$ is a sparse vector representing data falsely injected by the adversary. For the nonlinear model, the design of attack vector $a$ and analysis of the impact of attack is difficult. In [17], Liu et al. first considered the problem of data falsification attack based on the linearized power flow equation [the so-called *direct current* (DC) model] given by

$$z = Hx + a + w, a \in \mathcal{A}.$$

The key observation is that, if $a$ can be chosen in the column space of $H$, i.e., $a = Hc \in \mathcal{A}$, then $z = H(x + c) + w$. Thus, the state estimator will not be able to distinguish between the actual state $x$ and the adversary intended value $x + c$. Indeed, the magnitude of $c$ can be scaled arbitrarily, and the resulting estimation error at the FC can be arbitrarily large.

A key insight is to connect the problem of data falsification attack with the classical notion of observability [18]. Intuitively, the adversary can simply remove the attacked meters. In the absence of these meter data, the model becomes one with rows of $H$ removed. When enough rows are removed, the state vector becomes unidentifiable. Of course, if the adversary does take such an approach, the control center would have detected and remedial actions would have been taken. It is shown in [19] that choosing $a = Hc$ is simply a smart way of evading detection by the state estimator. The Byzantine data attack is equivalent to removing rows corresponding to the attacked meters and make matrix $H$ column rank deficient.

A number of questions can be formulated from here.

- Given the network topology, what is the minimum number of meters (and which meters) that the adversary has to attack?
- If the adversary has only access to a fixed set of meters, to what degree can the adversary affect the performance of the state estimator?
- Since the network topology is also estimated from data collected from the field, what happens when the adversary attacks the topology?
- How effective are such attacks when the linearity assumptions are removed?

These questions have been addressed to different degrees of satisfaction recently. We highlight a few such results.

## CRITICAL POWER OF BYZANTINE SENSORS

We can define the critical power $\chi$ of Byzantine sensors as the minimum number of meters that have to be compromised so that the network becomes unobservable. Finding $\chi$ for a given network is nontrivial and is made possible through the connection of a Byzantine attack and network observability. Checking the rank of $H$ by exhaustively searching rows to remove has exponential complexity and is also sensitive to numerical errors.

From the classical results of Krumpholz et al. [20], Kosut et al. developed a graph theoretic approach [19] to find $\chi$ with polynomial complexity. Interestingly, this approach uses only the network topology and does not depend on the specific system parameters. In [19], Kosut et al. find $\chi$ and the set of most vulnerable meters by maximizing a submodular function, which has a polynomial time solution.

If the attacker cannot alter enough meter data, the attacker is operating in the weak attack regime, and the power grid remains observable to the control center but the accuracy of the estimator is affected. Furthermore, the power system state estimator is equipped with a bad data detector that detects the presence of data anomalies. Thus, the adversary risks the possibility of being detected. The adversary then needs to make the tradeoff on the curve of attack operating characteristic that relate the mean squared error (MSE) of the state estimator and the probability of being detected, while the bad data detector operates on the ROC curve [19].

**THEREFORE, A BYZANTINE ATTACK CAN AFFECT THE REAL-TIME ELECTRICITY PRICE, PROVING AN ECONOMIC INCENTIVE TO THE ADVERSARY BY RAISING ELECTRICITY PRICES AT CERTAIN LOCATIONS.**

## TOPOLOGY ATTACKS

In addition to meter data, the control center also needs circuit breaker status for state estimation. In particular, the circuit breaker status data are used to construct network topology based on which state estimation follows. In fact, circuit breaker status changes from time to time, and the state estimator has to update the network topology accordingly. The breaker status is measured based on the level of physical contacts and is subject to error, which causes topology errors at the state estimator. The impact of topology error on state estimation has long been recognized. The state of the art is to detect topology error as part of the bad data detection and, in the event of topology error, jointly estimate topology and the state.

A more sophisticated data falsification attack is to attack both the meter and the circuit breaker status data. Such an attack can be quite effective in both evading detection by the control center and affecting system operation. For instance, the adversary may disguise a connected transmission line as disconnected, causing the control center to shed load to ensure stability. Similarly, the adversary may falsely mask a disconnected line thereby delay the necessary response by the control center to prevent cascading blackout. The effect of topology attacks on real-time pricing was also shown to be significant [21].

## ATTACK ON REAL-TIME ELECTRICITY MARKET

While state estimation serves as an important monitoring function, its use in real-time control is limited. However, the state estimate plays a critical role in computing the real-time price of electricity in the wholesale market. Therefore, a Byzantine attack can affect the real-time electricity price, proving an economic incentive to the adversary by raising electricity prices at certain locations. It is shown [22] that the state space of the power system is partitioned into price regions, and the price in each region depends on the congestion pattern of transmission lines and network topology. The adversary can affect the price by moving the estimated system state from one price region to another or creating a different congestion pattern thereby changing the price in a price region. Mechanisms of attacking real-time electricity market and counter measures are considered in [22]–[24].

## THE USE OF NONLINEAR ESTIMATORS

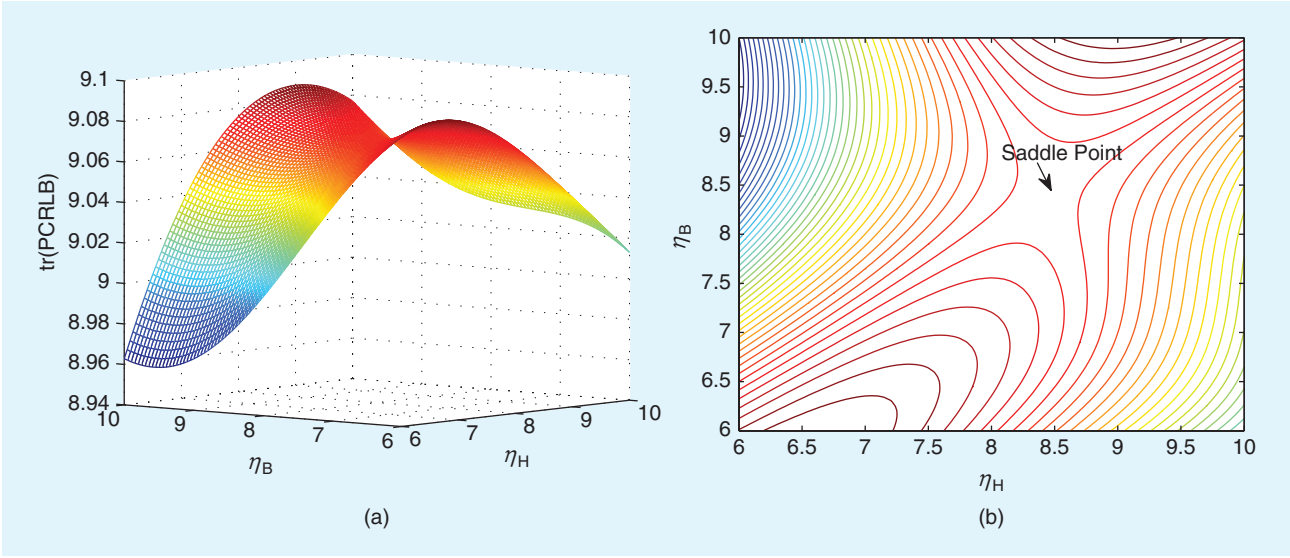
A less understood problem is the efficacy of attack in the presence of network nonlinearity and the use of nonlinear state estimator and the bad data detector. To this end, there is a lack of theoretical understanding of many issues mentioned above. There are, however, practical approaches that have been proposed, and simulation studies seem to suggest that some of the conclusions based on linear (DC) models may not hold. Indeed, most sensors used today (with the exception of the PMUs) give nonlinear measurements, and the state estimators used in the control center are variants of nonlinear least squares types. It is shown in some simulations that the effect of data falsification attacks on state estimates is limited when the nonlinear least squares method is used. The effect of topology attack remains substantial [25], [21].

### *LOCALIZATION USING QUANTIZED DATA*

Here, we provide another illustrative distributed estimation problem with Byzantine data. In particular, we discuss target localization in a wireless sensor network, where sensors send quantized data to the FC [26]. Vempaty et al. in [27] analyze the effect of Byzantines on target localization. Under an isotropic attenuation model of signal power from the target, a Monte Carlo-based minimum mean square error (MMSE) estimator is employed at the FC using the binary quantized data ($\mathbf{u}$) from the local sensors. Since the local sensors send binary quantized data, the strategy of the Byzantines is to flip their local decisions. Let the Byzantines flip their local decisions with probability "p." For a Bayesian estimation problem, the performance is evaluated by the estimation error, which is lower bounded by the posterior Cramér Rao lower bound (PCRLB),

$$E\{[\hat{\theta}(\mathbf{u})-\theta][\hat{\theta}(\mathbf{u})-\theta]^T\} \geq \mathbf{F}^{-1}, \tag{14}$$

where $\hat{\theta}(\mathbf{u})$ is the estimate of the target location given by $\theta$ and $\mathbf{F}$ is the Fisher information matrix (FIM). Using the PCRLB and




**[FIG5] Surface and contour plots of trace of PCRLB versus honest and Byzantine sensor's threshold showing the saddle point. (Figure used with permission from [28].)**

FIM as the performance metrics, the FC is defined to be "blind" when the data's contribution to the Fisher information (FI) becomes zero. The FIM consists of data's contribution $\mathbf{F}_D$ and prior's contribution $\mathbf{F}_P$ as follows:

$$\begin{aligned}\mathbf{F} &= -\mathrm{E}_{\theta,\mathbf{u}}[\nabla_\theta \nabla_\theta^{\mathrm{T}} \ln \mathrm{P}(\mathbf{u},\theta)],\\ &= -\mathrm{E}_{\theta,\mathbf{u}}[\nabla_\theta \nabla_\theta^{\mathrm{T}} \ln \mathrm{P}(\mathbf{u}\mid\theta)] - \mathrm{E}_{\theta}[\nabla_\theta \nabla_\theta^{\mathrm{T}} \ln \mathrm{p}_0(\theta)],\\ &= \mathbf{F_D} + \mathbf{F_P},\end{aligned}$$

where $P(\mathbf{u}\mid\theta)$ is the conditional probability density of the quantized data and $p_0(\theta)$ is the prior probability density function of the target location. Therefore, the FC becomes blind when the posterior FI becomes the same as the prior FI. For an independent attack model of the Byzantines, the optimal attacking strategy of the Byzantines has been found by modeling the effect of Byzantines as a binary symmetric channel (BSC). The optimal strategy is to flip with probability "1" and the minimum fraction of sensors that need to be Byzantines (attack power) to blind the FC is $\alpha^* = 0.5$. When the fraction of Byzantines $\alpha < \alpha^*$, there exists a zero-sum game between the honest sensors and the Byzantines and the optimal thresholds to be used by the local sensors are found as the saddle-point of the minimax game. Figure 5 shows the saddle point of the zero-sum game representing the optimal strategies of both honest and Byzantine sensors. Vempaty et al. [27] have also found a lower bound on $\alpha^*$ for the collaborative attack case by assuming that the Byzantines can learn the true location of the target perfectly.

## MITIGATION TECHNIQUES

In [27], Vempaty et al. propose two mitigation techniques to nullify the effect of Byzantines. In the first approach, an identification scheme similar to [7] has been proposed for the target localization framework. In such a scheme, each sensor's behavior is observed over time and the sensors that do not behave as expected are tagged as the Byzantine sensors. The rationale behind such a formulation is that in traditional classification/pattern recognition problems, the decision regarding the type (of a sensor) is made by observing the behavior (of the sensor). A sensor is declared as Type A, if it behaves closer to the expected behavior of Type A. They show with simulations that the proposed scheme detects most of the Byzantines. However, the proposed scheme has some shortcomings and declares a few sensors to be ambiguous when a clear decision regarding a sensor being Byzantine cannot be made. To mitigate the effect of Byzantines completely, they formulate the design of dynamic nonidentical thresholds using calculus of variation that maximizes the posterior FI. The threshold design corresponds to $\eta_i^{T+1} = \hat{a}_i^T$, where $\hat{a}_i^T$ is the estimated amplitude at this sensor at the time instant $T$. This means that the threshold of the $i$th sensor at time $(T+1)$ is the estimated amplitude at this sensor at the previous time instant $T$. This amplitude is estimated by using the previous time instant's location estimate, $\hat{\theta}^T$, which is broadcast by the FC to the local sensors. They show that this design of nonidentical thresholds can be implemented in a dynamic manner and the design not only reduces the location estimation error but also makes the Byzantines "ineffective" in their attack strategy.

## SUMMARY AND OPEN PROBLEMS

In this article, we have presented a tutorial discussion on distributed inference with Byzantine data and surveyed some recent results. While the results describe the importance for security in distributed networks, there are still a number of interesting questions that need to be addressed. Some examples are as follows:

- the issue of uncertainties such as imperfect channels, i.e., consideration of the reliability/imperfectness of the messengers in the Byzantine generals problem

- ensuring reliability in the presence of Byzantine sensors, which requires more sophisticated design across multiple layers of the networking protocol stack: advanced distributed inference at the physical layer, sophisticated network coding schemes for large networks, and a variety of cryptographic techniques for different applications
- development of a coherent theory and methodology that guides practical design for networks robust to Byzantines, and fundamental characterizations of communication rate in the presence of Byzantines in the general case
- development of network-wide capability to detect and mitigate Byzantines by deriving new networking protocols that facilitate distributed inference at all internal sensors under the assumptions that some internal sensors may themselves be Byzantines
- development of complex Byzantine misbehavior models and methods to detect and mitigate such Byzantines.

## AUTHORS

*Aditya Vempaty* (avempaty@syr.edu) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, in 2011. He is currently pursuing the Ph.D. degree at Syracuse University. He is with the Sensor Fusion Group. He was a visiting research intern at Syracuse University from May to July 2010. His research interests include wireless sensor networks, network security, statistical signal processing, crowdsourcing, and data fusion.

*Lang Tong* (ltong@ece.cornell.edu) received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 1985, and the Ph.D. degree in electrical engineering from the University of Notre Dame, Indiana in 1991. He joined Cornell University in 1998, where he is now the Irwin and Joan Jacobs Professor in Engineering and the Cornell site director of the Power Systems Engineering Research Center. Prior to joining Cornell University, he was on the faculty at West Virginia University and the University of Connecticut. He has received numerous awards including the 1996 Young Investigator Award from the Office of Naval Research and several Best Paper Awards. He was named a 2009 Distinguished Lecturer by the IEEE Signal Processing Society. His research is in the general area of statistical signal processing, communications, and complex networks. He is a Fellow of the IEEE.

*Pramod K. Varshney* (varshney@syr.edu) received the B.S. degree in electrical engineering and computer science 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. Since 1976, he has been with Syracuse University, where he is currently a distinguished professor of electrical engineering and computer science and the director of the Center for Advanced Systems and Engineering. His current research interests are in distributed sensor networks and data fusion, detection and estimation theory, and wireless communications. He was the 2001 president of International Society of Information Fusion. He is a Fellow of the IEEE and has received numerous awards including the IEEE Judith A. Resnik Award in 2012.

## REFERENCES

[1] L. Lamport, R. Shostak, and M. Pease, "The Byzantine generals problem," *ACM Trans. Program. Lang. Syst.*, vol. 4, no. 3, pp. 382–401, 1982.

[2] V. V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Philos. Trans. Roy. Soc. London A, Math. Phys. Eng. Sci.*, vol. 370, no. 1958, Jan. 2012, pp. 100–117.

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

[4] P. J. Huber, "A robust version of the probability ratio test," *Ann. Math. Stat.*, vol. 36, no. 6, pp. 1753–1758, 1965.

[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[6] A. Rawat, P. Anand, H. Chen, and P. K. Varshney, "Collaborative spectrum sensing in the presence of Byzantine attacks in cognitive radio networks," *IEEE Trans. Signal Processing*, vol. 59, pp. 774–786, Feb. 2011.

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

[8] X. He, H. Dai, and P. Ning, "A Byzantine attack defender: The conditional frequency check," in *Proc. IEEE Int. Symp. Information Theory*, July 2012, pp. 975–979.

[9] R. Chen, J.-M. Park, and K. Bain, "Robust distributed spectrum sensing in cognitive radio networks," in *Proc. 27th Conf. Computer Communication, IEEE INFOCOM*, Apr. 2008, pp. 1876–1884.

[10] M. Gagrani, P. Sharma, S. Iyengar, V. Nadendla, A. Vempaty, H. Chen, and P. Varshney, "On noise-enhanced distributed inference in the presence of Byzantines," in *Proc. 49th Annu. Allerton Conf. Communications Control Computing*, Sept. 2011, pp. 1222–1229.

[11] Z. Chair and P. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, pp. 98–101, Jan. 1986.

[12] H. Chen, "Noise enhanced signal detection and estimation," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Syracuse Univ., New York, 2007.

[13] P. Ray and P. Varshney, "False discovery rate based sensor decision rules for the network-wide distributed detection problem," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 3, pp. 1785–1799, July 2011.

[14] A. Vempaty, P. Ray, and P. K. Varshney. (2012, Dec.). False discovery rate based distributed detection in the presence of Byzantines. *IEEE Trans. Aerosp. Electron. Syst.* [Online]. Available: http://arxiv.org/pdf/1212.5654.pdf

[15] B. Kailkhura, S. Brahma, and P. K. Varshney, "Optimal Byzantine attacks on distributed detection in tree-based topologies," in *Proc. Int. Conf. Computing, Networking and Communications (ICNC)*, San Diego, CA, Jan. 2013, pp. 227–231.

[16] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. Boca Raton, FL: CRC Press, 2004.

[17] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. ACM Conf. Computer and Communications Security*, 2009, pp. 21–32.

[18] A. Monticelli and F. Wu, "Network observability: Theory," *IEEE Trans. Power Apparatus Syst.*, vol. PAS-104, no. 5, pp. 1042–1048, May 1985.

[19] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.

[20] G. Krumpholz, K. A. Clements, and P. W. Davis, "Power system observability: A practical algorithm using network topology," *IEEE Trans. Power Apparatus Syst.*, vol. PAS-99, no. 4, pp. 1534–1542, July 1980.

[21] J. Kim and L. Tong, "On topology attack of a smart grid: Undetectable attacks and countermeasures," *IEEE J. Select. Areas Commun.*, to be published.

[22] L. Jia, J. Kim, R. Thomas, and L. Tong, "Impacts of data quality on real-time locational marginal price: A worst case analysis," submitted for publication. [Online]. Available: http://arxiv.org/pdf/1212.6662.pdf

[23] L. Xie, Y. Mo, and B. Sinopoli, "False data injection attacks in electricity markets," in *Proc. 2010 1st IEEE Int. Conf. Smart Grid Communication (SmartGridComm)*, Gaithersburg, MD, pp. 226–231.

[24] L. Jia, R. J. Thomas, and L. Tong, "Malicious data attack on real-time electricity market," in *Proc. 2011 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, pp. 5952–5955.

[25] L. Jia, R. J. Thomas, and L. Tong, "On the nonlinearity effects on malicious data attack on power system," in *Proc. 2012 Power and Energy Society General Meeting*, pp. 1–8.

[26] R. Niu and P. K. Varshney, "Target location estimation in sensor networks with quantized data," *IEEE Trans. Signal Processing*, vol. 54, no. 12, pp. 4519–4528, Dec. 2006.

[27] A. Vempaty, O. Ozdemir, K. Agrawal, H. Chen, and P. Varshney, "Localization in wireless sensor networks: Byzantines and mitigation techniques," *IEEE Trans. Signal Processing*, vol. 61, no. 6, pp. 1495–1508, Mar. 2013.

[SP]