

# Social Learning Against Data Falsification in Sensor Networks

Fernando Rosas and Kwang-Cheng Chen, *Fellow, IEEE*

**Abstract**—Although surveillance and sensor networks play a key role in Internet of Things, sensor nodes are usually vulnerable to tampering due to their widespread locations. In this letter we consider data falsification attacks where an smart attacker takes control of critical nodes within the network, including nodes serving as fusion centers. In order to face this critical security thread, we propose a data aggregation scheme based on social learning, resembling the way in which agents make decisions in social networks. Our results suggest that social learning enables network resilience, even when a significant portion of the nodes have been compromised by the attacker. Finally, we show the suitability of our scheme to sensor networks by developing a low-complexity algorithm to facilitate the social learning data fusion rule in devices with restricted computational power.

**Index Terms**—Data fusion, sensor networks, surveillance networks, bizantine generals problem, social learning, network security, resilient networks.

## I. INTRODUCTION

Large distributed sensor networks typically provide surveillance services over extensive areas, such as activity monitoring in military or secure zones, protection of drinkable water tanks from chemical attacks, or intrusion detection [1], [2]. However, the reliability of these networks is in many cases limited due to the high vulnerability of the sensor nodes [3]. In reality, nodes are frequently located in unprotected locations and are susceptible to physical or cyber captures. Moreover, nodes are generally not tamper-proof due to cost concerns, and their limited computing power, memory, and energy capabilities do not allow sophisticated cryptographic techniques.

One serious threat to the reliability of distributed surveillance is the data falsification or “Byzantine” attack, where an adversary takes control over a number of authenticated nodes [4]. Following the classic *Byzantine Generals Problem* [5], Byzantine nodes can generate false sensing data, exhibit arbitrary behaviour or collude in order to create a networked malfunction. The effect of data falsification attacks over distributed detection has been intensely studied, characterizing the impact over the detection performance and also proposing various defense mechanisms (c.f. [6] for an overview, and also [7]–[9]). These works focus in networks with star or tree topology, where the data is gathered in a special node called “fusion center” (FC) that is responsible for the final decision.

A key assumption in the literature is that the adversary can compromise regular sensor nodes but not the FC itself.

However, in many scenarios the short range of the nodes’ transmissions force the FC to be installed in unsafe locations, being vulnerable to tampering as well. A tampered FC completely disables the detecting capabilities of the network, generating a single point of failure and hence becoming the weakest point of the system [10]. To address this serious security thread, this letter is novel in considering powerful topology-aware data falsification attacks, where the adversary knows the network topology and leverage this knowledge to take control of the most critical nodes of the network —either regular nodes or FCs. This represents a worst-case scenario, where the network structure has been disclosed e.g. from network tomography via traffic analysis [11].

The design of reliable distributed detection schemes is a challenging task. In effect, even though the distributed sensing literature is vast (see e.g. [1], [2] and references therein), the construction of optimal schemes is in general NP-hard [12]. Moreover, although in many cases the optimal schemes can be characterized as a set of thresholds for likelihood functions, the determination of these thresholds is usually an intractable problem [13]. For example, symmetric thresholds can be suboptimal even for networks with similar sensors arranged in star topology [14], being only asymptotically optimal when the network size increases [13], [15]. Moreover, symmetric strategies are not suitable for more elaborate network topologies, and hence heuristic methods are usually necessary.

To deal with this dilemma, in this letter we propose a low-complexity data aggregation scheme based on *social learning* principles, which resembles social decisions-making processes while avoiding fusion center functions [16]–[18]. The scheme is a threshold-based data fusion strategy related to the ones considered in [13]. However, its connection with social decision-making enables an intuitive understanding of its inner mechanisms, and also allows an efficient implementation that is suitable for the limited computational capabilities of a sensor node. For avoiding the security threads introduced by fusion centers, our scheme uses a tandem or serial topology [19]–[23]. Contrasting with the literature, our analysis does not focus on optimality issues of the data fusion but aims to illustrate how this scheme can enable network resilience against a powerful topology-aware data falsification attacker, even when a significant number of nodes have been compromised.

## II. SYSTEM MODEL AND PROBLEM STATEMENT

### A. System model

We consider a network of  $N$  sensor nodes that are deployed over an area where surveillance is needed. The output of the

sensor of the  $n$ -th node is denoted by  $S_n$ , taking values over a set  $\mathcal{S}$  that can be discrete or continuous. Based on these signals, the network needs to infer the value of the binary variable  $W$ , with events  $\{W = 1\}$  and  $\{W = 0\}$  corresponding to the presence or absence of an attack, respectively. No knowledge about of the prior distribution of  $W$  is assumed, as attacks are rare and might follow unpredictable patterns.

We consider nodes with equal sensing capabilities, and hence assume that the signals  $S_n$  are identically distributed. For the sake of tractability, it is assumed that the variables  $S_1, \dots, S_N$  are conditionally independent\* given the event  $\{W = w\}$ , following a probability distribution denoted by  $\mu_w$ . It is assumed that both  $\mu_0$  and  $\mu_1$  are absolutely continuous with respect to each other [25], i.e. no particular signal determines  $W$  unequivocally. The log-likelihood ratio of these two distributions is therefore given by the logarithm of the corresponding Radon-Nikodym derivative  $\Lambda_S(s) = \log \frac{d\mu_1}{d\mu_0}(s)^\dagger$ .

In addition to sensing hardware, each node is equipped with computing capability and a low-power transceiver to transmit and receive data. However, battery limitations impose severe restrictions over the communication bandwidth, and thus it is assumed that each node forward its data to others by broadcasting a binary variable  $X_n$ . Note that these signals could be appended to wireless control packages and viceversa.

The nodes transmit their signals sequentially according to their indices. Due to the nature of wireless broadcasting, which might be overlooked in some security literatures, nearby transmissions can be overheard. Therefore, it is assumed that the  $n$ -th node can generate  $X_n$  based on information provided by  $S_n$  and  $\mathbf{X}^{n-1} = (X_1, \dots, X_{n-1})$ . A *strategy* is a collection of functions  $\pi_n : \mathcal{S} \times \{0, 1\}^{n-1} \rightarrow \{0, 1\}$  such that  $X_n = \pi(S_n, \mathbf{X}^{n-1})$ . Although the burden of overhearing all the previously broadcasted signals can be reduced by designing smart network topologies and routing strategies, these networking functions are left for future studies.

The network operator collects the transmitted packages from a specific node labeled as  $n_c \in \{1, \dots, N\}$ , possibly employing unmanned ground or aerial vehicles that access a shared signal at a specific network location, or by using a shared communication channel. The network performance is quantified by the corresponding miss-detection and false alarm rates, given by  $\mathbb{P}\{\text{MD}\} = \mathbb{P}\{X_{n_c} = 0 | W = 1\}$  and  $\mathbb{P}\{\text{FA}\} = \mathbb{P}\{X_{n_c} = 1 | W = 0\}$ , respectively.

Finally, it is assumed that  $N^*$  Byzantine nodes are controlled by an adversary without being noticed by the network operator. The adversary can freely define the values of the binary signals transmitted by byzantine nodes in order to degrade the network performance. It is further assumed that the adversary is “topology-aware”, knowing the node sequence and the strategy that is in use. Therefore, the adversary could well control the  $N^*$  most critical nodes in terms of network

\*The conditional independency of sensor signals is satisfied when the sensor noise is due to local causes (e.g. thermal noise), but do not hold when there exist common noise sources (e.g. in the case of distributed acoustic sensors [24]).

$\dagger$ When  $S_n$  takes a finite number of values then  $\frac{d\mu_1}{d\mu_0}(s) = \frac{\mathbb{P}\{S_n=s|W=1\}}{\mathbb{P}\{S_n=s|W=0\}}$ , while if  $S_n$  is a continuous random variable with conditional p.d.f.  $p(S_n|w)$  then  $\frac{d\mu_1}{d\mu_0}(s) = \frac{p(s|w=1)}{p(s|w=0)}$ .

performance. However, the adversary has no knowledge about  $n_c$ , as it can be chosen at run-time and changed regularly.

### B. Problem statement

Our goal is to develop a network-resilient strategy to mitigate the effect from a powerful topology-aware adversary when the network operator (i.e. defender) has no knowledge of the number of Byzantine nodes or other attack’s statistics. Note that in most surveillance applications miss-detections are more important than false alarms, being difficult to estimate the cost of the worst-case scenario. Therefore, the system performance is evaluated following the Neyman-Pearson criteria by setting an allowable false alarm rate and focusing on the achievable miss-detection rate.

Most signal processing techniques for distributed detection rely on a FC(s) that gather data and generate estimators, and sensor nodes that provide informative signals to them [26]. Intuitively, if  $X_n$  is influenced by  $X_m$  with  $m < n$ , this would “double-count” the information provided by  $S_m$ . Therefore, in order to guarantee diversity, traditional distributed detection schemes choose to ignore previously broadcasted signals. However, as nodes don’t perform any data aggregation, each of their shared signals are not, by themselves, good estimations of the target variable. This generates a single point of failure in the network, as if the adversary compromises the FC(s) then the only accurate estimator that exist within the network is lost and hence the inference process fails.

## III. SOCIAL LEARNING AS A DATA AGGREGATION SCHEME

### A. Data fusion rule

Social learning models supply new directions to analyze the sequential decision processes where agents combine personal information and peers’ opinions [18]. Applied to a sensor network, each node can be considered as an agent that decides the presence of attacks based on measurements and overheard signals from other nodes. In this letter we consider rational agents that follow a *Bayesian strategy*, denoted as  $\pi_n^b(S_n, \mathbf{X}^{n-1})$ , which can be described by

$$\frac{\mathbb{P}\{W = 1 | S_n, \mathbf{X}^{n-1}\}}{\mathbb{P}\{W = 0 | S_n, \mathbf{X}^{n-1}\}} \stackrel{\pi_n^b=0}{\underset{\pi_n^b=1}{\leq}} \frac{u(0, 0) - u(1, 0)}{u(1, 1) - u(0, 1)}. \quad (1)$$

Above,  $u(x, w)$  is a cost assigned to the decision  $X_n = x$  when  $W = w$ , which can be engineered in order to match the relevance of miss-detections and false alarms [27]. Moreover, by noting that  $X^{n-1} = \pi_{n-1}^b(S_{n-1}, \mathbf{X}^{n-2})$  is influenced only by  $S_1, \dots, S_{n-1}$ , the conditional independency of the signals imply that  $S_n$  and  $\mathbf{X}^{n-1}$  are also conditionally independent given  $W = w$ . Therefore, using the Bayes rule, a direct calculation shows that (1) can be re-written as

$$\Lambda_S(S_n) + \Lambda_{\mathbf{X}^{n-1}}(\mathbf{X}^{n-1}) \stackrel{\pi_n^b=0}{\underset{\pi_n^b=1}{\leq}} \tau, \quad (2)$$

where  $\tau = \log \frac{\mathbb{P}\{W=0\}}{\mathbb{P}\{W=1\}} + \log \frac{u(0,0) - u(1,0)}{u(1,1) - u(0,1)}$  and  $\Lambda_{\mathbf{X}^{n-1}}(\mathbf{X}^{n-1})$  is the log-likelihood ratio of  $\mathbf{X}^{n-1}$ . As the prior distribution of  $W$  is usually unknown, the network operator needs to select

the lowest value of  $\tau$  that satisfies the required false alarm rate given by the Neyman-Pearson criteria (c.f. Section II-B).

As in a realistic scenario the statistical properties of the potential topology-aware data falsification attacks are not available to the defender, our approach is to make each node to follow a bayesian strategy ignoring the potential attack. Such an approach has three attractive features:

1. Provides a computation rule that does not need to adapt to different attacker's profiles.
2. Minimizes the average cost  $\mathbb{E}\{u(\pi_n(S_n, \mathbf{X}^{n-1}), W)\}$  when no attacks take place [27].
3. Enables network resilience (c.f. Section III-C and IV).

Clearly Byzantine nodes do not follow (2), as their interest is to degrade the network performance. Let us denote as  $\mathcal{B}$  the set of indices of the Byzantine nodes and  $N^*$  the cardinality of  $\mathcal{B}$ . As events  $\{W = 0\}$  are much more frequent than  $\{W = 1\}$ , any abnormal increase of the false alarm rate would be easily noted and hence provides no benefit to the adversary. Therefore, a rational strategy for the adversary is to increase the miss-detection rate by forcing  $X_n = 0$  for all  $n \in \mathcal{B}$ .

### B. An algorithm for computing the social log-likelihood

The only challenge for implementing (2) in a sensor node as a data fusion rule is to have an efficient algorithm for computing  $\Lambda_{\mathbf{X}^{n-1}}(\mathbf{x}^{n-1})$ . For finding such an algorithm, a direct application of the chain rule of probabilities shows that

$$\Lambda_{\mathbf{X}^n}(\mathbf{x}^n) = \log \prod_{k=1}^n \frac{\mathbb{P}\{X_k = x_k | \mathbf{X}^{k-1} = \mathbf{x}^{k-1}, W = 1\}}{\mathbb{P}\{X_k = x_k | \mathbf{X}^{k-1} = \mathbf{x}^{k-1}, W = 0\}},$$

with the understanding that  $\mathbf{X}^0 = \mathbf{x}^0$  is null. Then, following the discussion presented in Section III-A, we compute  $\mathbb{P}\{X_k = x_k | \mathbf{X}^{k-1} = \mathbf{x}^{k-1}, W = w\}$  ignoring potential attacks. Assuming that the  $k$ -th node is not a Byzantine node, one obtains

$$\begin{aligned} & \mathbb{P}\{X_k = 0 | \mathbf{X}^{k-1} = \mathbf{x}^{k-1}, W = w\} \\ &= \int_{\mathcal{S}} \mathbb{P}\{X_k = 0 | \mathbf{X}^{k-1} = \mathbf{x}^{k-1}, W = w, S_k = s\} d\mu_w(s) \\ &= \int_{\mathcal{S}} \mathbb{1}\{\pi_k^b(s, \mathbf{x}^{k-1}) = 0\} d\mu_w(s) \\ &= \mathbb{P}_w\{\Lambda_S(S_k) + \Lambda_{\mathbf{X}^{k-1}}(\mathbf{x}^{k-1}) < \tau\} \\ &= F_w^\Lambda(\tau - \Lambda_{\mathbf{X}^{k-1}}(\mathbf{x}^{k-1})), \end{aligned} \quad (3)$$

where  $F_w^\Lambda(\cdot)$  is the c.d.f. of the variable  $\Lambda_S(S_n)$  conditioned to  $W = w$ . Using the above results, it can be shown that

$$\Lambda_{\mathbf{X}^{n+1}}(\mathbf{x}^{n+1}) - \Lambda_{\mathbf{X}^n}(\mathbf{x}^n) = \lambda(x_n, \tau - \Lambda_{\mathbf{X}^n}(\mathbf{x}^n)),$$

where  $\lambda(\cdot, \cdot)$  is defined as

$$\lambda(x, a) = x \log \frac{F_1^\Lambda(a)}{F_0^\Lambda(a)} + (1-x) \log \frac{1 - F_1^\Lambda(a)}{1 - F_0^\Lambda(a)}.$$

Leveraging above derivations, we develop Algorithm 1 as a simple iterative procedure for computing  $\Lambda_{\mathbf{X}^n}(\mathbf{x}^n)$ . Note that the algorithm's complexity scales gracefully, as it grows linearly with the length of  $\mathbf{x}^n$ . Moreover, the algorithm does not need any information about potential attack, only requiring knowledge of the signals statistics as given by  $F_w^\Lambda$ .

---

### Algorithm 1 Computation of $\Lambda_{\mathbf{X}^n}(\mathbf{x}^n)$

---

```

1: function LOGLIKELIHOOD( $\mathbf{x}^n, \tau$ )
2:    $L_1 = \lambda(x_1, \tau)$ .
3:   for  $k = 2, \dots, n$  do
4:      $L_k = L_{k-1} + \lambda(x_{k+1}, \tau - L_{k-1})$ .
5:   end for
6:   return  $L_n$ 
7: end function

```

---

### C. Information cascades as strength or weakness

The term ‘‘social learning’’ refers to the fact that the accuracy of  $X_n$  as a predictor of  $W$  grows with  $n$ , and hence  $n_c$  is usually chosen as one of the last nodes in the decision sequence. However, as the number of shared signals grows the increasing ‘‘social pressure’’ can make the nodes to ignore their individual measurements and blindly follow the dominant choice [16]. This phenomenon, known as *information cascade*, introduces severe limitations in the achievable asymptotic performance of social learning [17].

A positive effect of information cascades, which has been overlooked before, is to make a large number of agents/nodes to hold equally qualified estimator(s), generating many locations where the network operator can collect and aggregate the data. This property avoids the existence of a single point of failure to robustly face topology-aware attacks. An attempt to blindly guess  $n_c$  in order to tamper the  $n_c$ -node would be inefficient due to the large number of potential candidates.

However, an attacker can also leverage the information cascade phenomenon. A rational attacking strategy is to tamper the first  $N^*$  nodes of the decision sequence, setting their signals in order to push the networked decisions towards a misleading cascade<sup>‡</sup>. If  $N^*$  is large enough an information cascade can be triggered almost surely, making the learning process to fail. However, if  $N^*$  is not large enough then the network may undo the initial pool of wrong opinions and end up triggering a correct cascade anyway. This capability of ‘‘resilience’’ is explored in the next section.

## IV. PROOF OF CONCEPT

To illustrate the application of social learning against topology-aware data falsification attacks, we consider a network of randomly distributed sensors over a sensitive area following a Poisson Point process (PPP). The ratio of the area that is within the range of each sensor is denoted by  $r$ . If attacks occur uniformly over the surveilled area, then  $r$  is also the probability of an attack taking place under the coverage area of a particular sensor is. It is further assumed that each node is equipped with a binary sensor (i.e.  $S_n \in \{0, 1\}$ ), whose probability of generating a wrong measurement due to electronic and other imperfections is denoted by  $q$ .

<sup>‡</sup>Intuitively, it is more likely for a node to follow a misleading cascade if all the previous  $N^*$  nodes have been tampered and act homogeneously, than for a node of higher index if the previous decisions are non-homogeneous.

For finding the posterior distributions of  $S_n$ , first note that  $\mathbb{P}_0\{S_n = 1\} = q$ , as a sensor false-alarm can only be due to noise. The probability of detecting an event is given by

$$\begin{aligned} \mathbb{P}\{S_n = 1|W = 1\} &= \mathbb{P}\{\text{attack in range, good measurement}|W = 1\} \\ &\quad + \mathbb{P}\{\text{attack out of range, bad measurement}|W = 1\} \\ &= r + q - 2rq. \end{aligned}$$

Therefore, the sensor miss-detection rate is  $\mathbb{P}_1\{S_n = 0\} = 1 - r - q + 2rq$ . The signal log-likelihood is hence given by

$$\Lambda_S(S_n) = S_n \log \frac{r + q - 2rq}{q} + (1 - S_n) \log \frac{1 - r - q + 2rq}{1 - q}.$$

Note that  $\Lambda_S(1) > \Lambda_S(0)$ , which is consequence of  $r + q - 2rq > q$  and  $q < 1/2$ . Correspondingly, the c.d.f. of  $\Lambda_S$  is

$$F_w^\Lambda(l) = \begin{cases} 0 & \text{if } l < \Lambda(0), \\ \mathbb{P}\{S_n = 0|W = w\} & \text{if } \Lambda(0) \leq l < \Lambda(1), \\ 1 & \text{if } \Lambda(1) \geq l. \end{cases}$$

We studied a network composed by  $N = 200$  sensor nodes, generating  $\mathbf{X}^n$  sequentially following (3) and using Algorithm 1 to compute  $\Lambda_{\mathbf{X}^n}(\mathbf{X}^n)$ . Following Section III-C, it is assumed that a topology-aware attacker tampered the first  $N^*$  nodes of the decision sequence and uses them to increase the miss-detection rate by setting  $X_n = 0$  for  $n = 1, \dots, N^*$ . Finally, in order to favour the reduction of miss-detections over false alarms,  $\tau = 0$  is chosen as is the lowest value that still allows a non-trivial inference process. For each set of parameter values,  $10^4$  simulation runs are performed.

Simulations demonstrate that the proposed scheme enables strong network resilience in this scenario, allowing the sensor network to maintain a low miss-detection rate even in the presence of an important number of Byzantine nodes (see Figure 1). In contrast, if a traditional distributed detection scheme is used, a topology-aware attacker can cause a miss-detection rate of 100% by just compromising the few nodes that perform data aggregation, i.e. the FC(s). Figure 1 shows that nodes aggregating data by social learning can achieve an average asymptotic miss-detection rate of less than 5% even when 30% of the most critical nodes are under the control of the attacker, having some resemblance with the well-known 1/3 threshold of the Byzantine generals problem [5]. Moreover, Figure 1 also suggest that our scheme can still provide network resilience within the 10% most unfavorable cases.

Interestingly, the data aggregation is performed node by node independently of the network size. Hence, in a very large network the first 200 nodes would exhibit the same performance as the one shown in Figure 1. Adding more nodes may not introduce significant improvements to the asymptotic performance, as the asymptotic estimator is copied by later nodes following an information cascade. Nevertheless, in a large network information cascades provide the fundamental benefit of creating a large number of nodes from where the network operator can access aggregated data.

The network resilience provided by our scheme is influenced by the sensor statistics, which are determined by  $q$  and  $r$  (see Figure 2). Intuitively, the achievable miss-detection rate under a low number of Byzantine nodes is reduced by a smaller  $q$

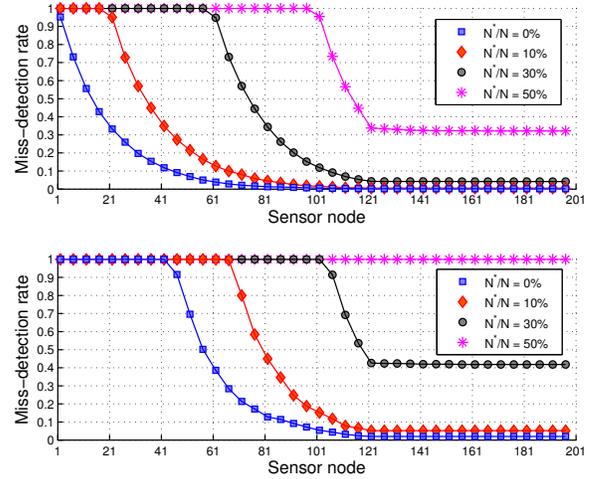


Fig. 1: *Above*: Performance of a surveillance network based on social learning, with binary signals of range  $r = 5\%$  and error rate  $q = 10^{-4}$ , when  $N^*$  out of  $N$  nodes are compromised by an attacker. *Below*: Performance considering the 10% most unfavorable cases.

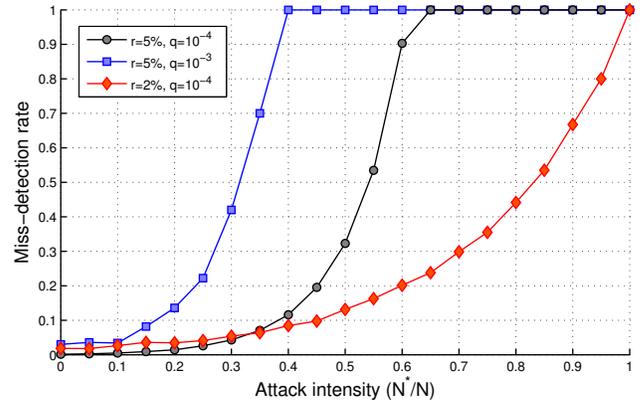


Fig. 2: Asymptotic average performance of a surveillance system. A smaller sensor error rate ( $q$ ) or large sensing range ( $r$ ) improves the performance under a low  $N^*$ , but the latter also makes the performance degradation less graceful when  $N^*$  grows.

or larger  $r$ . Furthermore, our numerical results suggest that the number of Byzantine nodes affects the miss-detection rate exponentially with a rate of growth inversely proportional to  $r$ , as nodes with smaller  $r$  trust each others decisions less and hence are less affected by “social pressure”. Consequently, it is desirable to deploy sensors with smaller probability of malfunction ( $q$ ) than larger coverage ( $r$ ), as a larger coverage makes the network more vulnerable to Byzantine nodes and subsequent misleading information cascades.

Our scheme does not require knowledge about attack statistics, being well-suited for practical scenarios as operation in large scale or mobile scenarios suggest dynamically changing topology. Moreover, simulations show that if the adversary tamper not the initial nodes but a different set of the same cardinality, then the attack has less impact over the system performance. This suggests that our scheme can provide further resilience against attackers who are not topology-aware.

## REFERENCES

- [1] V. V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1958, pp. 100–117, 2012.
- [2] S. Barbarossa, S. Sardellitti, and P. D. Lorenzo, *Distributed Detection and Estimation in Wireless Sensor Networks*. Academic Press Library in Signal Processing, Vol. 2, Communications and Radar Signal Processing, Oct. 2013, vol. 2, pp. 329–408.
- [3] E. Shi and A. Perrig, "Designing secure sensor networks," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 38–43, 2004.
- [4] S. Marano, V. Matta, and L. Tong, "Distributed detection in the presence of byzantine attacks," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 16–29, 2009.
- [5] L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," *ACM Transactions on Programming Languages and Systems (TOPLAS)*, vol. 4, no. 3, pp. 382–401, 1982.
- [6] A. Vempaty, L. Tong, and P. K. Varshney, "Distributed inference with byzantine data: State-of-the-art review on data falsification attacks," *IEEE Signal Processing Magazine*, vol. 30, no. 5, pp. 65–75, 2013.
- [7] V. S. S. Nadendla, Y. S. Han, and P. K. Varshney, "Distributed inference with m-ary quantized data in the presence of byzantine attacks," *IEEE Transactions on Signal Processing*, vol. 62, no. 10, pp. 2681–2695, May 2014.
- [8] J. Zhang, R. S. Blum, X. Lu, and D. Conus, "Asymptotically optimum distributed estimation in the presence of attacks," *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1086–1101, March 2015.
- [9] B. Kailkhura, Y. S. Han, S. Brahma, and P. K. Varshney, "Distributed bayesian detection in the presence of byzantine data," *IEEE Transactions on Signal Processing*, vol. 63, no. 19, pp. 5250–5263, Oct 2015.
- [10] B. Parno, A. Perrig, and V. Gligor, "Distributed detection of node replication attacks in sensor networks," in *2005 IEEE Symposium on Security and Privacy (S&P'05)*. IEEE, 2005, pp. 49–63.
- [11] R. Castro, M. Coates, G. Liang, R. Nowak, and B. Yu, "Network tomography: recent developments," *Statistical science*, pp. 499–517, 2004.
- [12] J. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Transactions on Automatic Control*, vol. 30, no. 5, pp. 440–446, 1985.
- [13] J. N. Tsitsiklis *et al.*, "Decentralized detection," *Advances in Statistical Signal Processing*, vol. 2, no. 2, pp. 297–344, 1993.
- [14] D. Warren and P. Willett, "Optimum quantization for detector fusion: some proofs, examples, and pathology," *Journal of the Franklin Institute*, vol. 336, no. 2, pp. 323–359, 1999.
- [15] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE Journal on selected areas in communications*, vol. 22, no. 6, pp. 1007–1015, 2004.
- [16] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, pp. 992–1026, 1992.
- [17] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *The Review of Economic Studies*, vol. 78, no. 4, pp. 1201–1236, 2011.
- [18] V. Krishnamurthy and H. V. Poor, "Social learning and bayesian games in multiagent signal processing: How do local and global decision makers interact?" *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 43–57, 2013.
- [19] R. Viswanathan, S. C. Thomopoulos, and R. Tumuluri, "Optimal serial distributed decision fusion," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 4, pp. 366–376, 1988.
- [20] J. D. Papastavrou and M. Athans, "Distributed detection by a large team of sensors in tandem," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 639–653, 1992.
- [21] P. F. Swaszek, "On the performance of serial networks in distributed detection," *IEEE transactions on aerospace and electronic systems*, vol. 29, no. 1, pp. 254–260, 1993.
- [22] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors i. fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, 1997.
- [23] I. Bahceci, G. Al-Regib, and Y. Altunbasak, "Serial distributed detection for wireless sensor networks," in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*. IEEE, 2005, pp. 830–834.
- [24] A. Bertrand, "Applications and trends in wireless acoustic sensor networks: A signal processing perspective," in *2011 18th IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT)*, Nov 2011, pp. 1–6.
- [25] M. Loeve, *Probability Theory I*. Springer, 1978.
- [26] R. Rajagopalan and P. K. Varshney, "Data-aggregation techniques in sensor networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 8, no. 4, pp. 48–63, Fourth 2006.
- [27] H. V. Poor, *An introduction to signal detection and estimation*. Springer Science & Business Media, 2013.