Social Learning Against Data Falsification in Sensor Networks

Fernando Rosas and Kwang-Cheng Chen

Abstract Sensor networks generate large amounts of geographically-distributed data. The conventional approach to exploit this data is to first gather it in a special node that then performs processing and inference. However, what happens if this node is destroyed, or even worst, if it is hijacked? To explore this problem, in this work we consider a smart attacker who can take control of critical nodes within the network and use them to inject false information. In order to face this critical security thread, we propose a novel scheme that enables data aggregation and decision-making over networks based on social learning, where the sensor nodes act resembling how agents make decisions in social networks. Our results suggest that social learning enables high network resilience, even when a significant portion of the nodes have been compromised by the attacker.

1 Introduction

Large networks of devices that monitor extensive geographical areas are widespread today, and will become pervasive in the near future. These networks enable critical services to society, including surveillance over military or secure zones, monitoring of drinkable water tanks and protection from chemical attacks, intrusion detection to private property, etc [27, 3]. However, the reliability of these networks is usually

Fernando Rosas
Centre of Complexity Science and Department of Mathematics
Department of Electrical and Electronic Engineering
Imperial College London, Kensington SW72AZ, London, UK
e-mail: f.rosas@imperial.ac.uk

Kwang-Cheng Chen
Department of Electrical Engineering
University of South Florida, Tampa, FL 33620, USA
e-mail: kwangcheng@usf.edu

limited due to the high vulnerability of the sensor nodes [23]. In reality nodes are frequently deployed in unprotected locations and can be damaged or destroyed, or can be subject of 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 of the most serious threats to the reliability of a large network of sensors is the data falsification or "Byzantine" attack, where an adversary takes control over a number of authenticated nodes [16]. Following the classic *Byzantine Generals Problem* [14], Byzantine nodes can generate false data, exhibit arbitrary behaviour or collude in order to create a networked malfunction. The effect of data falsification over distributed detection has been intensely studied, characterizing the impact over the detection performance and also proposing various defense mechanisms (c.f. [28] for an overview, and also [17, 32, 12, 11, 10] for some recent contributions). However, all these works focus in networks with star or tree topology, and rely on centralising the decision-making in a special node called "fusion center" (FC) which gathers all the sensed data. Note that these approaches rely on a strong division of labour: ordinary sensor nodes just sense and forward data while the processing is done exclusively at the FC, corresponding to a *distributed-sensing with centralized-processing* (DSCP) approach.

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 limited range of nodes' radios force the FC to be installed in unsafe locations, being vulnerable to tampering as well. A tampered FC completely disables the capabilities of the network, generating a single point of failure and hence becoming the weakest point of the system [19]. 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[6].

In order to address this issue one needs to consider *distributed-sensing with distributed-processing* (DSDP) schemes, which avoid FC functions while distributing processing tasks throughout the network. However, the design of reliable DSDP schemes is a challenging task. In effect, even though the distributed sensing literature is vast (see e.g. [27, 3] and references therein), the construction of optimal schemes is in general NP-hard [25]. 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 [26]. For example, symmetric thresholds can be suboptimal even for networks with similar sensors arranged in star topology [31], being only asymptotically optimal when the network size increases [26, 8]. 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 work we propose a DSDP scheme based on *social learning* principles, which resembles social decisions-making processes [5, 1, 13]. The scheme is a threshold-based data fusion strategy related to the ones

considered in [26]. 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 [29, 18, 24, 30, 2]. Contrasting with the literature, our analysis does not focus on optimality issues of the data fusion but aims to illustrate how the distribution of processing tasks can enable network resilience against a powerful topology-aware data falsification attacker. We show how the network resilience holds even when a significant number of nodes have been compromised.

The rest of this article is structured as follows. First Section 2 introduces the system model. Our social learning data fusion rule is then presented in Section 3, and it is then illustrated in a concrete scenario in Section 4. Finally, Section 5 presents our main conclusions.

2 System model and problem statement

2.1 System model

We consider a network of N nodes, where each node corresponds to an electronic device that has been deployed over a geographical area where sensing and surveillance is needed. The node are equipped with sensors that enables them to sense relevant variables from the environment. The output of the sensor of the n-th node is denoted by S_n, taking values over a set $\mathscr{S}$ 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 patters.

Nodes have equal sensing capabilities, and hence the signals S_n are assumed to be identically distributed. For the sake of tractability, it is assumed that the variables $S_1, \ldots, S_N$ are conditionally independent given the event $\{W = w\}$, following a probability distribution denoted by μ_w*. It is also assumed that both μ_0 and μ_1 are absolutely continuous with respect to each other [15], 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 $\Lambda_S(s) = \log \frac{d\mu_1}{d\mu_0}(s)$.†. It is also assumed that $\mu_0 \neq \mu_1$, so that $\Lambda(S_n)$ is not trivially equal to zero.

* 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 [4]).

† 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 $\frac{d\mu_1}{d\mu_0}(s) = \frac{p(s|w=1)}{p(s|w=0)}$.

In addition to sensing hardware, each node is equipped with computing capability and a wireless radio to transit and receive data. Two nodes in the network are said to be connected if they can exchange information wirelessly. Note that sensor nodes usually have very limited battery budget, which impose severe restrictions over the communication capabilities. Therefore, it is assumed that each node forward its data to others only by broadcasting a binary variable X_n[‡]. Without loss of generality, the nodes transmit their signals sequentially according to their indices (i.e. node 1 transmits first, then node 2, etc). Due to the nature of wireless broadcasting, which might be overlooked in some security literatures, nearby transmissions can be overheard. Therefore, a static fully-connected network topology is considered that allows the n-th node to generate X_n based on information provided by S_n and $X^{n-1} = (X_1, \ldots, X_{n-1})$. A *strategy* is a collection of functions $\pi_n : \mathscr{S} \times \{0,1\}^{n-1} \to \{0,1\}$ such that $X_n = \pi(S_n, X^{n-1})$. Although the burden of overhearing all the previously broadcasted signals can be reduced by designing smart network topologies, these networking functions are left for future studies.

The network operator collects the transmitted packages from a specific node labeled as $n_\mathrm{c} \in \{1, \ldots, N\}$, possibly employing unmanned ground or aerial vehicles that access a shared signal at a specific network location, or by using a shared communication channel. Therefore, X_{n_c} constitutes the output of the overall inference process. The network performance is quantified by the corresponding miss-detection and false alarm rates, given by $\mathbb{P}\{\mathrm{MD}\} = \mathbb{P}\{X_{n_\mathrm{c}} = 0 | W = 1\}$ and $\mathbb{P}\{\mathrm{FA}\} = \mathbb{P}\{X_{n_\mathrm{c}} = 1 | W = 0\}$, 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 specific strategy $\{\pi_n\}_{n=1}^N$ that is in use. Therefore, the adversary could well control the N^* most critical nodes in terms of network performance. However, the adversary has no knowledge about n_c, as it can be chosen at run-time and changed regularly.

2.2 Problem statement

Our goal is to develop a network-resilient strategy to mitigate the effect the attacks coming from a 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 [20]. Note than finding optimal solutions is a formidable challenge, even for the

[‡] These signals could be appended to wireless control packages and viceversa, or also could be shared by light, ultrasound or other media.

simple case of networks with start topology and no Byzantine attacks (see [9] and references therein).

Most signal processing techniques for distributed detection rely on a FC(s) that gather data and generate estimators using the data provided by passive sensor nodes [21]. Intuitively, if X_n is influenced by X_m with $m < n$, this would "double-count" the information provided by S_m. To avoid this, traditional distributed detection schemes choose to ignore previously broadcasted signals. This leads to good performance statistics, achieving exponentially decaying miss-detection rates with respect to the number of sensing nodes [7]. 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. The FC is the most critical node to the detection performance, and therefore would be the most endangered element of the network[§].

3 Social learning as a data aggregation scheme

3.1 Data fusion rule

Social learning models supply new directions to analyze sequential decision processes where agents combine personal information and peers' opinions [22]. Applied to a sensor network, each node can be considered as an agent that decides the presence of attacks based on their measurements and overheard signals from other nodes. In this work we consider rational agents that follow a *Bayesian strategy*, denoted as $\pi_n^{\mathrm{b}}(S_n, X^{n-1})$, which can be described by the following rule:

$$\frac{\mathbb{P}\left\{W=1|S_n,X^{n-1}\right\}}{\mathbb{P}\left\{W=0|S_n,X^{n-1}\right\}} \underset{\pi_n^{\mathrm{b}}=1}{\overset{\pi_n^{\mathrm{b}}=0}{\lessgtr}} \frac{u(0,0)-u(1,0)}{u(1,1)-u(0,1)} \quad . \tag{1}$$

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 [20]. Moreover, by noting that $X^{n-1} = \pi_{n-1}^{\mathrm{b}}(S_{n-1}, X^{n-2})$ is influenced only by $S_1, \ldots, S_{n-1}$, the conditional independency of the signals imply that S_n and 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

$$\Lambda_S(S_n) + \Lambda_{X^{n-1}}(X^{n-1}) \underset{\pi_n^{\mathrm{b}}=1}{\overset{\pi_n^{\mathrm{b}}=0}{\lessgtr}} \tau \ , \tag{2}$$

[§] For the case of wireless sensor networks the typical tranmission ranges are beyond 40 meters. It is therefore likely that the fusion center may also be deployed in a vulnerable location and hence be victim of tampering.

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 $\Lambda_{X^{n-1}}(X^{n-1})$ is the log-likelihood ratio of X^{n-1}¶. In simple words, (2) states how the the n-th node should fuse the knowledge coming from S_n and X^{n-1}: it should only infere the presence of an attack when the sum of the log-likelihood terms is larger than τ.

As in a realistic scenario the statistical properties of attacks are usually not available to the defender, our approach is for each node to follow a bayesian strategy that ignores the potential attack. Such an approach has three attractive features:

1. Provides a scheme that does not need to adapt to different attacker's profiles.
2. Minimizes the average cost when no attacks take place.
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 $\mathscr{B}$ the set of indices of the Byzantine nodes and N^* the cardinality of $\mathscr{B}$. 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 $n \in \mathscr{B}$.

3.2 An algorithm for computing the social log-likelihood

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

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

with the understanding that $X^0 = x^0$ is null. Then, following the discussion presented in Section 3.1, we compute $\mathbb{P}\{X_k = x_k \,|\, X^{k-1} = x^{k-1}, W = w\}$ ignoring potential attacks. Assuming that the k-th node is not a Byzantine node, one obtains

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

¶ As the prior distribution of W is usually unknown, the network operator needs to select the lowest value of τ that satisfies the required false alarm rate given by the Neyman-Pearson criteria (c.f. Section 2.2).

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

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

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 $\Lambda_{X^n}(x^n)$. Note that the algorithm's complexity scales gracefully, as it grows linearly with the length of 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^{Λ}.

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

```
function LOGLIKELIHOOD(x^n, τ)
    L_1 = λ(x_1, τ).
    for k = 2, ..., n do
        L_k = L_{k-1} + λ(x_{k+1}, τ - L_{k-1}).
    end for
    return L_n
end function
```

3.3 Information cascades as strength or weakness

The term "social learning" refers to the fact that X_n becomes a better predictor of W as n grows, and hence n_{c} is usually chosen as one of the last nodes in the decision sequence. However, as the number of shared signals increases the growing "social pressure" can make nodes to ignore their individual measurements and blindly follow the dominant choice, generating a herd behaviour [5]. This phenomenon, known as *information cascade*, introduces severe limitations in the asymptotic performance of social learning [1].

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 aggregated data. This avoids the existence of a single point of failure and allows to robustly face topology-aware attacks. In fact 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[‖]. 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" depends on the signals distribution, and is explored in the next section.

4 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 geographical 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.

For finding the posterior distributions of S_n, first note that

$$\mathbb{P}\{S_n = 1|W = 0\} = q,$$

as a sensor false-alarm can only be due to noise. The probability of detecting an event is given by

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

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

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

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

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

‖ 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.

The inference problem is hence to distinguish between two Bernoulli variables with parameters q and $r+q-2rq$, respectively. Note that the only non-trivial strategy based on a single measurement is to choose $X_n = S_n$. However, if $r = 5\%$ and $q = 10^{-3}$ this strategy give a miss-detection rate of 0.949, indicating that without collaboration each node is extremely unreliable.

We studied a network of $N = 200$ sensor nodes, generating X^n sequentially following (3) and using Algorithm 1 to compute $\Lambda_{X^n}(X^n)$. Following Section 3.3, we considered a topology-aware attacker who tampered the first N^* nodes of the decision sequence and uses them to increase the miss-detection rate by setting $X_n = 0$ for $n = 1, \ldots, N^*$. 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 [14]. Moreover, Figure 1 also suggest that our scheme can still provide network resilience within the 10% most unfavorable cases. These results confirms that our data aggregation scheme effectively avoids having single points of failure.

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 to the network may not introduce significant improvements to the asymptotic performance, as the asymptotic estimator is practically attained already by the 150-th node, being 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 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.

** Simulations showed that if $\tau < 0$ then $X_n = 1$ for all $n \in \mathbb{N}$ independently of the value of W, triggering a premature information cascade.

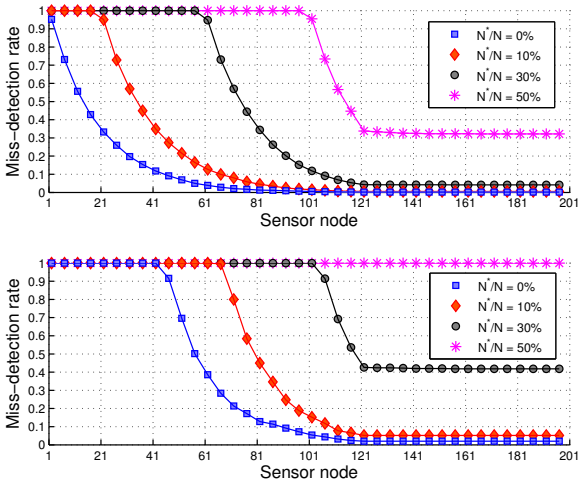

Fig. 1 *Above:* Performance of a surveillance network based on social learning, with binary signals of range $r = 5\%$ and error rate $q = 10^{-4}$, when N^* out of N nodes are compromised by an attacker. *Bellow:* Performance considering the 10% most unfavorable cases.

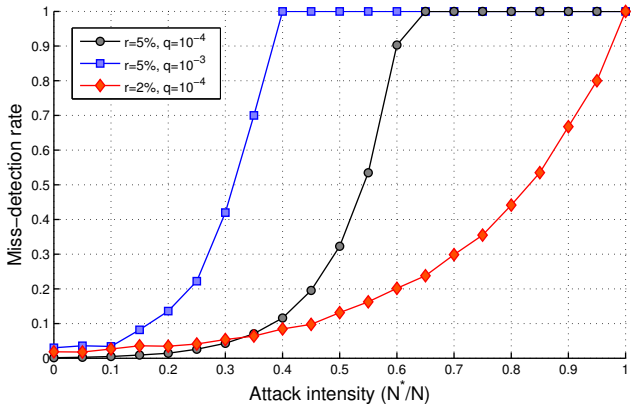

Fig. 2 Asymptotic average performance of a surveillance system. A smaller sensor error rate (q) or large sensing range (r) improves the performance under a low N^*, but the latter also makes the performance degradation less graceful when N^* grows.

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.

5 Conclusions

Traditional data aggregation schemes over sensor networks posses a single point of failure, which is caused by the fact that the actual processing is performed by few particular nodes. This weakness can be overcome by aggregating the data following social learning principles, which distributes the processing tasks throughout the network. This approach avoids single points of failure by generating a large number of nodes from where aggregated data can be accessed. A social learning algorithm was presented, which is simple and susceptible of being implemented in devices with limited computational capabilities.

Our social learning data processing scheme enables resilience against topology-aware data falsification attacks, which totally disable the detection capabilities of traditional sensor networks. Furthermore, results suggest that the network resilience persists even when the attacker has compromised an important number of nodes.

We hope that these results can motivate further explorations on the interface between distributed decision making, inference and signal processing over technological and social networks.

6 Acknowledges

This project was supported by the European Union's H2020 research and innovation programme, under the Marie Skłodowska-Curie grant agreement No. 702981.

References

1. Acemoglu, D., Dahleh, M.A., Lobel, I., Ozdaglar, A.: Bayesian learning in social networks. The Review of Economic Studies **78**(4), 1201–1236 (2011)
2. Bahceci, I., Al-Regib, G., Altunbasak, Y.: Serial distributed detection for wireless sensor networks. In: Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on, pp. 830–834. IEEE (2005)
3. Barbarossa, S., Sardellitti, S., Lorenzo, P.D.: Distributed Detection and Estimation in Wireless Sensor Networks, vol. 2, pp. 329–408. Academic Press Library in Signal Processing, Vol. 2, Communications and Radar Signal Processing (2013)
4. Bertrand, A.: 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), pp. 1–6 (2011). DOI 10.1109/SCVT.2011.6101302
5. Bikhchandani, S., Hirshleifer, D., Welch, I.: A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of political Economy pp. 992–1026 (1992)
6. Castro, R., Coates, M., Liang, G., Nowak, R., Yu, B.: Network tomography: recent developments. Statistical science pp. 499–517 (2004)
7. Chamberland, J.F., Veeravalli, V.V.: Decentralized detection in sensor networks. IEEE Transactions on Signal Processing **51**(2), 407–416 (2003)
8. Chamberland, J.F., Veeravalli, V.V.: Asymptotic results for decentralized detection in power constrained wireless sensor networks. IEEE Journal on selected areas in communications **22**(6), 1007–1015 (2004)

9. Chamberland, J.F., Veeravalli, V.V.: Wireless sensors in distributed detection applications. IEEE signal processing magazine **24**(3), 16–25 (2007)
10. Kailkhura, B., Brahma, S., Dulek, B., Han, Y.S., Varshney, P.K.: Distributed detection in tree networks: Byzantines and mitigation techniques. IEEE Transactions on Information Forensics and Security **10**(7), 1499–1512 (2015). DOI 10.1109/TIFS.2015.2415757
11. Kailkhura, B., Brahma, S., Han, Y.S., Varshney, P.K.: Distributed detection in tree topologies with byzantines. IEEE Transactions on Signal Processing **62**(12), 3208–3219 (2014)
12. Kailkhura, B., Han, Y.S., Brahma, S., Varshney, P.K.: Distributed bayesian detection in the presence of byzantine data. IEEE Transactions on Signal Processing **63**(19), 5250–5263 (2015). DOI 10.1109/TSP.2015.2450191
13. Krishnamurthy, V., Poor, H.V.: Social learning and bayesian games in multiagent signal processing: How do local and global decision makers interact? IEEE Signal Processing Magazine **30**(3), 43–57 (2013)
14. Lamport, L., Shostak, R., Pease, M.: The byzantine generals problem. ACM Transactions on Programming Languages and Systems (TOPLAS) **4**(3), 382–401 (1982)
15. Loeve, M.: Probability Theory I. Springer (1978)
16. Marano, S., Matta, V., Tong, L.: Distributed detection in the presence of byzantine attacks. IEEE Transactions on Signal Processing **57**(1), 16–29 (2009)
17. Nadendla, V.S.S., Han, Y.S., Varshney, P.K.: Distributed inference with m-ary quantized data in the presence of byzantine attacks. IEEE Transactions on Signal Processing **62**(10), 2681–2695 (2014). DOI 10.1109/TSP.2014.2314072
18. Papastavrou, J.D., Athans, M.: Distributed detection by a large team of sensors in tandem. IEEE Transactions on Aerospace and Electronic Systems **28**(3), 639–653 (1992)
19. Parno, B., Perrig, A., Gligor, V.: Distributed detection of node replication attacks in sensor networks. In: 2005 IEEE Symposium on Security and Privacy (S&P'05), pp. 49–63. IEEE (2005)
20. Poor, H.V.: An introduction to signal detection and estimation. Springer Science & Business Media (2013)
21. Rajagopalan, R., Varshney, P.K.: Data-aggregation techniques in sensor networks: A survey. IEEE Communications Surveys Tutorials **8**(4), 48–63 (2006). DOI 10.1109/COMST.2006.283821
22. Rosas, F., Hsiao, J.H., Chen, K.C.: A technological perspective on information cascades via social learning. IEEE Access **PP**(99), 1–1 (2017). DOI 10.1109/ACCESS.2017.2687422
23. Shi, E., Perrig, A.: Designing secure sensor networks. IEEE Wireless Communications **11**(6), 38–43 (2004)
24. Swaszek, P.F.: On the performance of serial networks in distributed detection. IEEE transactions on aerospace and electronic systems **29**(1), 254–260 (1993)
25. Tsitsiklis, J., Athans, M.: On the complexity of decentralized decision making and detection problems. IEEE Transactions on Automatic Control **30**(5), 440–446 (1985)
26. Tsitsiklis, J.N., et al.: Decentralized detection. Advances in Statistical Signal Processing **2**(2), 297–344 (1993)
27. Veeravalli, V.V., Varshney, P.K.: Distributed inference in wireless sensor networks. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **370**(1958), 100–117 (2012)
28. Vempaty, A., Tong, L., Varshney, P.K.: Distributed inference with byzantine data: State-of-the-art review on data falsification attacks. IEEE Signal Processing Magazine **30**(5), 65–75 (2013)
29. Viswanathan, R., Thomopoulos, S.C., Tumuluri, R.: Optimal serial distributed decision fusion. IEEE Transactions on Aerospace and Electronic Systems **24**(4), 366–376 (1988)
30. Viswanathan, R., Varshney, P.K.: Distributed detection with multiple sensors i. fundamentals. Proceedings of the IEEE **85**(1), 54–63 (1997)
31. Warren, D., Willett, P.: Optimum quantization for detector fusion: some proofs, examples, and pathology. Journal of the Franklin Institute **336**(2), 323–359 (1999)
32. Zhang, J., Blum, R.S., Lu, X., Conus, D.: Asymptotically optimum distributed estimation in the presence of attacks. IEEE Transactions on Signal Processing **63**(5), 1086–1101 (2015). DOI 10.1109/TSP.2014.2386281