Decision Fusion with Unbalanced Priors under Synchronized Byzantine Attacks: a Message-Passing Approach

Andrea Abrardo, Mauro Barni, Kassem Kallas, and Benedetta Tondi
University of Siena, Siena, Italy
abrardo@diism.unisi.it, barni@dii.unisi.it, k_kallas@hotmail.com, benedettatondi@gmail.com

***Abstract*—We consider a variant of the decision fusion problem in the presence of Byzantines where the two states of the system under observation are not equiprobable. In this setup, the Byzantines can not adopt a simple corruption strategy consisting in flipping the local decisions regardless of the estimated state of the system. Doing so, in fact, they would reveal their presence to the fusion center, since their reports would not follow the expected statistics. On its side, the fusion center can exploit the knowledge of the a-priori probabilities to improve its decision. In view of the above observations, we first introduce a new corruption strategy for the Byzantines, which permits them to make the statistics of their reports indistinguishable from those of the honest nodes. Then, we adopt the perspective of the fusion center and we propose a nearly-optimum, efficient, fusion strategy based on message passing, to face with the new attack. We do so in the most challenging scenario wherein the Byzantines are synchronised, i.e. they share a common source of randomness allowing them to submit wrong reports in a simultaneous way. We prove the validity of the proposed approach under several working conditions with regard to the percentage of byzantine nodes, the length of the observation window and the a priori-probabilities of the system states.**

I. Introduction

Decision fusion in the presence of Byzantines [1] has received an increasing attention in the last years due to its relevance in several scenarios including: wireless sensor networks [2], cognitive radio [3], distributed detection [4] and many others [5], [6], [7], [8]. In the most studied version of the problem, a Fusion Center (FC) has to make a binary decision about the status of an observed system, by collecting the decisions made locally by the nodes. In doing so, the FC must take into account the possible presence of Byzantines, that is nodes submitting a wrong decision in the attempt to induce a decision error. Most of the works published so far, assume that the corruption strategy adopted by the Byzantines consists in flipping the local decision made by the node with a certain probability P_{mal}, often assumed equal to 1.

In the non-adversarial version of the problem, the Bayesian optimal fusion rule (known as Chair-Varshney rule) has been derived in [9] . Extending the Chair-Varshney rule to consider the presence of the Byzantines requires that the FC knows the positions of the Byzantines as well as the flipping probability P_{mal}. This information is rarely available hence calling for the adoption of suboptimal fusion rules. In [4], for instance, by adopting a Neyman-Pearson setup and assuming that the Byzantines know the true system state, the asymptotic performance achievable by the FC when the size of the network (number of nodes) increases is analysed as a function of the percentage of Byzantines in the network.

In order to improve the estimation of the system states, the FC can make its decision by relying on a sequence of reports sent over an observation window of length m, referring to m subsequent states of the observed system. In this way, it is possible for the FC to isolate the Byzantines and consequently ignore their reports. In this vein, the analysis of [4] is extended in [10] to a situation in which the Byzantines are unaware of the true system state. Byzantines isolation is achieved by counting the mismatches between the reports received from each node and the global decision made by the FC. To overcome the lack of knowledge about the exact strategy adopted by the Byzantines, the authors adopt a game-theoretic setup in which each party makes its best choice without knowing the strategy of the other party. A soft isolation scheme is proposed in [11], where the reports from suspect nodes are given a lower reputation rather than being completely discarded. Even in [11], the lack of knowledge at the FC about the strategy adopted by the attacker (and viceversa) is tackled by adopting a game-theoretic formulation. A rather different approach is adopted in [12], where a tolerant scheme that mitigates the impact of Byzantines on the global decision is used rather that ignoring the reports submitted by suspect nodes. When the value of P_{mal} and the probability that a node is a Byzantine are known, the optimum fusion rule under multiple observations can be derived [13]. Since P_{mal} is usually unknown to the FC, in [13] the value of P_{mal} used within the optimum fusion rule and the value actually used by the Byzantines are strategically chosen in a game-theoretic setting. One of the main results in [13] is that the best option for the Byzantines is not to always flip the local decision (corresponding to $P_{mal} = 1$), since, once the malicious nodes are identified, a flipped report still brings useful information about the sate of the system. On the contrary, for certain combinations of the distribution of Byzantines within the network and the length of the observation window, it is better for the Byzantines to minimize the mutual information between the reports submitted to the FC and the system states ($P_{mal} = 0.5$). One of the main drawbacks of the optimum fusion rule proposed in [13] is that the computational cost grows exponentially with the size of the

observation window m. This problem is resolved in [14] using a nearly-optimum fusion scheme based on message passing (MP) that permits to reduce such exponential complexity to linear.

A common assumption in all the works discussed above is that the Byzantines do not talk to each other, which means that the Byzantine attack is not synchronized. However, in a recent work [15], it has been proven that when the Byzantines are synchronized, i.e. when they flip the local decisions simultaneously[1], the effect of the attack increases dramatically. For this reason, in the rest of this work we focus in the case of a synchronized attack, however the main conclusions of the paper still holds in the asynchronous case.

A. Contribution

In virtually all the scenarios considered so far, the a-priori probabilities of the two states of the observed system are supposed to be equal. Such an assumption, however, does not hold in many practical applications. In cognitive radio, for instance, the possibility of finding an unoccupied frequency band is due to the unbalanced prior probabilities of accessing the frequency spectrum by the stakeholder (Primary User). [16]. Distributed binary detection for monitoring and anomaly detection applications is another scenario wherein the assumption of balanced priors does not usually hold [17].

In this paper, we address the binary decision fusion problem in the presence of Byzantines, when the system states are not equiprobable. While this may seem a negligible assumption, its impact is a significant one. A first important consequence regards the strategy used by the Byzantines to induce a decision error. As we will see throughout the paper, they can no longer adopt a simple corruption strategy consisting in flipping the local decisions regardless of the estimated state of the system. In such a way, in fact, they would reveal their presence to the fusion center, since their reports would not follow the expected statistics. In turn, the fusion center can exploit the knowledge of the a-priori probabilities to improve its decision.

Following the above observations, this paper offers a two-fold contribution. First, we introduce a new corruption strategy for the Byzantines, which permits them to make the statistics of their reports indistinguishable from those of the honest nodes. According to the new strategy, the probability of flipping the local decision depends on the local estimate of the system state. As a second contribution, we derive the optimum decision fusion strategy by taking into account the a-priori unbalanced probabilities and, most of all, the new attacking strategy adopted by the Byzantines. We then introduce a nearly optimum decision fusion strategy based on the message passing approach (similarly to what has been done in [14]), which, at the price of a slight deterioration of the performance, simplifies greatly the fusion rule. Throughout the paper we assume that the fusion center has a perfect knowledge of all the parameters of the system, including the attacking strategy adopted by the Byzantines. We leave to a future work the study, in a game-theoretic setting, of a more realistic situation wherein the fusion center is not aware of the exact strategy adopted by the Byzantines. The soundness of the proposed solutions is proved through numerical simulations, aiming at showing the validity of the new attacking strategy and the effectiveness of the decision fusion rule based on message passing. We do so by considering several working conditions with regard to the percentage of Byzantine nodes, the length of the observation window and the a priori-probabilities of the system states.

The rest of this paper is organized as follows: in Section II, we formalise the problem at hand and we propose the new attack model, while in Section III we present the message passing algorithm with unbalanced a-priors. In Section IV we use simulations to analyze the performance of the message passing algorithm with unbalanced a-priors as well as the new attack model. Finally, we draw some conclusions and highlight directions for future work in Section V.

II. Problem Formulation

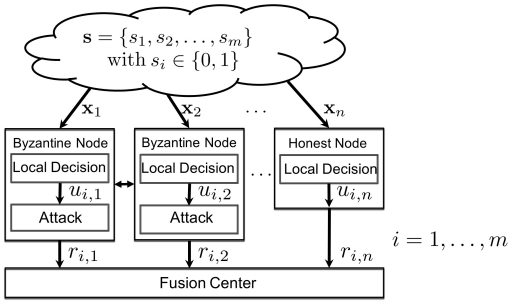

Fig. 1. Sketch of the adversarial decision fusion scenario considered in the paper.

A. Problem Setup

The adversarial decision fusion scenario considered in this paper is depicted in Figure 1. We let $\mathbf{s} = \{s_1, s_2, \ldots, s_m\}$ indicate a sequence of independent and identically distributed (i.i.d) system states, over an observation window of length m. We assume a binary system (i.e., $s_i \in \{0, 1\}$) with non-equiprobable states. We let P_0 and P_1 denote the prior probability of state 0 and 1 respectively (w.l.o.g., we assume $P_0 > P_1$). The n nodes collect information about the system through the vectors $\mathbf{x}_1, \mathbf{x}_2 \ldots \mathbf{x}_n$, with $\mathbf{x}_j$ indicating the observations available at node j. Based on its observation, node j makes a local decision $u_{i,j}$ about system state s_i. We assume that the local error probabilities are symmetric and independent of i and j, that is $p(u_{i,j} \neq s_i) = \varepsilon$. The state of the nodes in the network, indicating whether a node is a Byzantine or not, is given by the vector $\mathbf{h} = \{h_1, h_2, \ldots, h_n\}$ with $h_j = 1$ (res. 0) indicating that node j is honest (res. Byzantine). The reports received by the FC are collected into

[1]Of course the difference with respect to a non-synchronised attack makes sense only when $P_{mal} \neq 1$.

a matrix $\mathbf{R} = \{r_{i,j}\}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, where $r_{i,j}$ is the report sent by node j relative to s_i. For honest nodes, $u_{i,j} = r_{i,j}$ while, for Byzantines, $u_{i,j} \neq r_{i,j}$ with some positive probability. Specifically, by assuming an error-free transmission between the nodes and the FC, for honest nodes we have:

$$p\left(r_{i,j}|h_j = 1, s_i\right) = (1-\varepsilon)\delta(r_{i,j} - s_i) + \\ + \varepsilon(1 - \delta(r_{i,j} - s_i)), \quad (1)$$

where $\delta(a)$ is equal to 1 when its argument is 0 and 0 otherwise. Then, the probability of receiving a report $r_{i,j}$ at the FC is given by

$$p\left(r_{i,j} = 0|h_j = 1\right) = (1-\varepsilon)P_0 + \varepsilon P_1, \quad (2)$$

and, obviously, $p\left(r_{i,j} = 1|h_j = 1\right) = \varepsilon P_0 + (1-\varepsilon)P_1$. For malicious nodes, the probability that the FC receives a wrong report depends on the attack strategy adopted by the Byzantines and is discussed in the next section.

We assume that the states of the nodes are independent of each other and that the state of each node is a Bernoulli random variable with parameter α, that is, $p(h_j = 0) = \alpha \; \forall j$. Therefore, the number of Byzantine nodes in the network is a random variable following a binomial distribution, corresponding to the maximum entropy case [13] with $p\left(\mathbf{h}\right) = \prod_j p(h_j)$, where $p(h_j) = \alpha(1 - h_j) + (1-\alpha)h_j$.

B. A New (Synchronized) Attack Strategy

The attack strategy for the byzantine nodes generally consists in flipping the local decision with a certain probability P_{mal}, independently of each other and regardless of the observed value of the local decision, that is $p(r_{i,j} \neq u_{i,j}) = P_{mal}$ [10]-[14]. In the setup with unbalanced priors considered in this paper, a similar attack strategy would make the statistics of the byzantine reports different from that of the honest nodes, thus easing the task of the FC, which can improve the decision by exploiting the knowledge he has on the system (see Section IV-A). To avoid this problem, we introduce a new attack strategy according to which the probability of flipping the local decision depends on the value of the decision itself.

As we said in the introduction, we assume that the Byzantines are synchronized, since in this way the effectiveness of the attack increases significantly at the price of a minor complication [15]. In a fully symmetric setup, like the one considered in [15] (the system states are equiprobable and the local decision errors are symmetric), the Byzantines can coordinate the attacks by generating, locally, a binary sequence and then deciding to flip the reports based on the value assumed by such a sequence. The generation of the same sequence for all the nodes can be achieved, for instance, by means of a pseudo random number generator fed with a common seed.

Here we propose to generalize the synchronized attack strategy proposed in [15] by considering two random sequences, $\hat{\mathbf{s}}_0$ and $\hat{\mathbf{s}}_1$, of length m, based on which the Byzantines flip the local decision when such a decision is 0 and 1 respectively. Specifically, at any time instant i, the local decision is flipped when $\hat{s}_{0,i} = 1$ and the local decision is 0, and when $\hat{s}_{1,i} = 1$ and the local decision is 1. We also assume that subsequent observations are flipped independently, that is $\hat{\mathbf{s}}_0$ and $\hat{\mathbf{s}}_1$ are i.i.d. sequences. Let P^0_{mal} and P^1_{mal} denote the probability that $\hat{s}_{0,i} = 1$ and $\hat{s}_{1,i} = 1$ respectively, that is the probability that the local decision is flipped when $u_{i,j} = 0$ (res. $u_{i,j} = 1$). The above attack strategy corresponds to applying a binary asymmetric channel (BAC) with crossover probabilities (P^0_{mal}, P^1_{mal}) to the local decisions of the byzantine nodes. In the above setting, for the reports received from the byzantine nodes (under error-free transmission), we have

$$p\left(r_{i,j}|h_j = 0, s_i, \hat{s}_{1,i}, \hat{s}_{0,i}\right) = \\ \begin{cases} (1-\varepsilon)\delta(r_{i,j} - s_i) + \varepsilon(1 - \delta(r_{i,j} - s_i)), & \hat{s}_{1,i} = \hat{s}_{0,i} = 0 \\ (1-\varepsilon)(1 - \delta(r_{i,j} - s_i)) + \varepsilon\delta(r_{i,j} - s_i), & \hat{s}_{1,i} = \hat{s}_{0,i} = 1 \\ \delta(r_{i,j}), & \hat{s}_{1,i} = 1, \hat{s}_{0,i} = 0 \\ \delta(r_{i,j} - 1), & \hat{s}_{1,i} = 0, \hat{s}_{0,i} = 1, \end{cases} \quad (3)$$

where ε is the error probability of the local decisions at the nodes. The probability of receiving a report $r_{i,j}$ from a byzantine node j, i.e., $p\left(r_{i,j}|h_j = 0\right)$, can be obtained from (3), by applying the law of total probability. Alternatively, it can be directly derived by considering the cascade of BSC and BAC channels accounting, respectively, for the local decision errors and the action of the Byzantines (see Figure 2). For $r_{i,j} = 0$, we have:

$$p\left(r_{i,j} = 0|h_j = 0\right) = P_0(1-\varepsilon)(1 - P^0_{mal}) + P_0\varepsilon P^1_{mal} + \\ + P_1(1-\varepsilon)P^1_{mal} + P_1\varepsilon P_1(1 - P^0_{mal}), \quad (4)$$

and, similarly, for $r_{i,j} = 1$.

We now determine the relationship between P^0_{mal} and P^1_{mal} which ensures that the statistics of the reports submitted by the byzantine nodes are equal to that of the honest nodes.

1) Relationship between P^0_{mal} and P^1_{mal}: for a Byzantine node, the report received by the FC can be regarded as the output of the cascade of a BSC with crossover probability ε and a BAC with crossover probabilities P^0_{mal} and P^1_{mal}, see Figure 2. To lighten the notation, let us define

$$\rho = p(u_{i,j} = 0) = (1-\varepsilon)P_0 + \varepsilon P_1. \quad (5)$$

Then, $1 - \rho = p(u_{i,j} = 1)$. We observe that since $\varepsilon < 0.5$ is small and $P_0 > P_1$, we have $\rho > 1 - \rho$.

As said before, the choice of the pair (P^0_{mal}, P^1_{mal}) must be made in such a way that $r_{i,j}$ is statistically indistinguishable from $u_{i,j}$. This corresponds to impose that $p(r_{i,j} = 0) = \rho$. Therefore, the Byzantines choose the pair of flipping probabilities in such a way that $\rho(1 - P^0_{mal}) + (1-\rho)P^1_{mal} = \rho$, yielding:

$$P^0_{mal} = \frac{1-\rho}{\rho} P^1_{mal}. \quad (6)$$

With the above choice, the only degree of freedom for the Byzantines is the choice of P^1_{mal}, the two synchronization sequences being generated according to $(P^0_{mal}(P^1_{mal}), P^1_{mal})$. Notice that in our setting $P^0_{mal} < P^1_{mal}$, and hence when

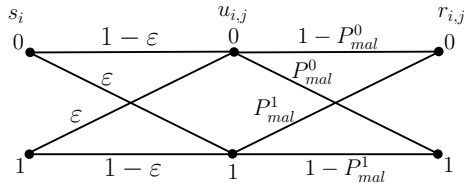

Fig. 2. Cascade of BSC with error probability ε (local decision), and a BAC with error probabilities P^0_{mal} and P^1_{mal} (attack) characterizing the report r_{ij} from a Byzantine node j.

P^1_{mal} spans the [0,1] interval, P^0_{mal} ranges from 0 to $(1-\rho)/\rho$. It is also worth observing that the mutual information conveyed by the Byzantines towards the FC depends on the choice of P^1_{mal}. With reference to the scheme in Figure 2, it is easy to argue that the condition of zero mutual information between the system states and the malicious reports, that is, the condition $p(r_{ij}|s_{ij}=0) = p(r_{ij}|s_{ij}=1)$, is achieved when $P^1_{mal} = (1-P^0_{mal})$. In fact, for $r_{ij}=0$ (similarly, for $r_{ij}=1$), we have:

$$\begin{aligned} p(r_{ij}=0|s_{ij}=0) &= (1-\varepsilon)(1-P^0_{mal}) + \varepsilon P^1_{mal}, \\ p(r_{ij}=0|s_{ij}=1) &= (1-\varepsilon)P^1_{mal} + \varepsilon(1-P^0_{mal}). \end{aligned} \quad (7)$$

Given the relationship between P^1_{mal} and P^0_{mal} established in (6), the condition $P^1_{mal} = (1-P^0_{mal})$ is satisfied when $P^1_{mal} = \rho$, and $P^0_{mal} = (1-\rho)$.

III. MP-Based, Decision Fusion with Unbalanced Priors

Given the sequence of reports $\mathbf{R}$, the optimum decision at the FC can be taken by looking at the *bitwise* Maximum A Posteriori Probability (MAP) estimation of the system states s_i, as follows:

$$\begin{aligned} s_i^* &= \arg\max_{s_i\in\{0,1\}} p(s_i|\mathbf{R}) \\ &= \arg\max_{s_i\in\{0,1\}} \sum_{\{\mathbf{s},\hat{\mathbf{s}}_0,\hat{\mathbf{s}}_1,\mathbf{h}\}\setminus s_i} p(\mathbf{s},\hat{\mathbf{s}}_0,\hat{\mathbf{s}}_1,\mathbf{h}|\mathbf{R}) \\ &= \arg\max_{s_i\in\{0,1\}} \sum_{\{\mathbf{s},\hat{\mathbf{s}}_0,\hat{\mathbf{s}}_1,\mathbf{h}\}\setminus s_i} p(\mathbf{R}|\mathbf{s},\hat{\mathbf{s}}_0,\hat{\mathbf{s}}_1,\mathbf{h})\, p(\mathbf{s})p(\hat{\mathbf{s}}_0)p(\hat{\mathbf{s}}_1)p(\mathbf{h}) \\ &= \arg\max_{s_i\in\{0,1\}} \sum_{\{\mathbf{s},\hat{\mathbf{s}}_0,\hat{\mathbf{s}}_1,\mathbf{h}\}\setminus s_i} \prod_{i,j} p(r_{i,j}|s_i,\hat{s}_{0,i},\hat{s}_{1,i},h_j) \prod_i p(s_i) \\ &\prod_i p(\hat{s}_{0,i}) \prod_i p(\hat{s}_{1,i}) \prod_j p(h_j), \end{aligned} \quad (8)$$

where the notation $\sum_{\setminus}$ denotes a summation over all the variables contained in the expression except the one listed after the operator.

The objective function of the optimal fusion rule expressed in (8) can be seen as a marginalization of a sum product of functions of integer variables, and, as such, it can be computed by resorting to Message Passing (MP) [14]. Specifically, in our problem, the variables are the system states s_i, the local random sequences $\hat{s}_{0,i}$ and $\hat{s}_{1,i}$, and the status of the nodes h_j, while the functions are the probabilities of the reports (3) and (1), and the a-priori probabilities $p(s_i)$, $p(\hat{s}_{0,i})$, $p(\hat{s}_{1,i})$ and $p(h_j)$. Hence, similarly to the approach proposed [14], it is possible to associate the bipartite graph shown in Fig. 3 to problem (8), from which we can easily derive the corresponding MP algorithm that proceeds iteratively according to the general message passing rules, until all variable nodes are able to compute the respective marginals. However, by comparing the graph representing the problem at hand with that described in [14], the presence of shorter cycles (cycles of order 3) readily comes out, due to the fact that variables s_i, $\hat{s}_{0,i}$ and $\hat{s}_{1,i}$ are connected to the same function nodes. This may be a problem, since many previous works in the field of channel coding, e.g., see [18], has proven that, in order to get good performance, the factor graph should not contain short cycles. To circumvent this problem, we propose a variable grouping approach that allows to reduce the cycles' length at the expenses of an slight increase of the computational complexity.

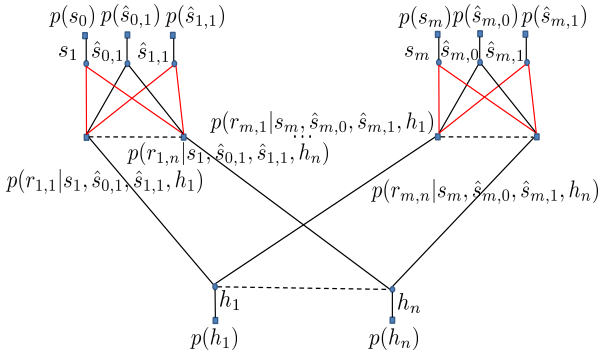

Fig. 3. Factor graph for problem (8): the presence of a cycle of order 3, wherein a message crosses three different nodes before returning to the sender, is highlighted.

To elaborate, let us consider the composite variable $\mathbf{g}_i = \{s_i, \hat{s}_{0,i}, \hat{s}_{1,i}\}$, with an alphabet of cardinality $2^3 = 8$. These variables can be further super-grouped into variables $\mathbf{g}_{i,w} = \{\mathbf{g}_i, \mathbf{g}_{i+1}, \ldots, \mathbf{g}_{i+w-1}\}$, with cardinality 2^{3w}. Leveraging on the definition of $\mathbf{g}_{i,w}$, assuming for simplicity $m = w \times q$, i.e., the number of states in the super-group is a multiple of w, and introducing the set of q indexes $\mathcal{I} = \{1, w+1, 2w+1, \ldots, m-w+1\}$, we can rewrite problem (8) as:

$$\begin{aligned} s_i^* = \arg\max_{s_i\in\{0,1\}} \sum_{\{\mathbf{g},\mathbf{h}\}\setminus s_i} \prod_{k\in\mathcal{I}} \prod_{j=1}^{n} p([\mathbf{R}]_{k:k+w,j}|\mathbf{g}_{k,w},h_j) \\ \prod_{k\in\mathcal{I}} p(\mathbf{g}_{k,w}) \prod_{j=1}^{n} p(h_j) \end{aligned} \quad (9)$$

where $[\mathbf{R}]_{k:k+w,j} = \{r_{k,j}, r_{k+1,j}, \ldots, r_{k+w,j}\}$, $\mathbf{g} =$

$\cup_{i\in\mathcal{I}}\mathbf{g}_{i,w}$, and

$$p\left([\mathbf{R}]_{k:k+w,j}|\mathbf{g}_{k,w},h_j\right)=\prod_{c=1}^{w}p\left(r_{k+c,j}|s_c,\hat{s}_{0,c},\hat{s}_{1,c},h_j\right)$$
$$p\left(\mathbf{g}_{k,w}\right)=\prod_{c=1}^{w}p(s_c)p(\hat{s}_{0,c})p(\hat{s}_{1,c}). \quad (10)$$

It is now possible to associate the bipartite graph shown in Fig. 4 to problem (10), and derive the corresponding MP algorithm that allows all variable nodes to compute the respective marginals. The details of the MP algorithm are omitted for the sake of brevity, but they can be easily derived considering the general rules of MP, e.g., see [19] and [20]. It is worth noting that grouping the variables nodes allows to get - in this specific case - a minimum cycle length equal to 7. Moreover, it is straightforward to observe that when $w=m$ the graph contains a single super-grouped variable $\mathbf{g}_{1,w}$ and, accordingly, it reduces to a tree-graph, i.e., no cycles are present. In this case, the MP algorithm allows to get the optimal solution of (10), with a complexity that grows exponentially with m. Moreover, it is easy to verify that for $w<m$, i.e., for $q>1$, the graph contains cycles of minimum order 7, but the number of minimum length cycles increases exponentially with q. Hence, we expect that the performance of the MP algorithm improves with the increase of w. Eventually, the value of w must be established by considering the trade-off between complexity (that depends exponential on $3w$) and performance. In the next section, we will show that letting $w=3$ achieves such a trade-off.

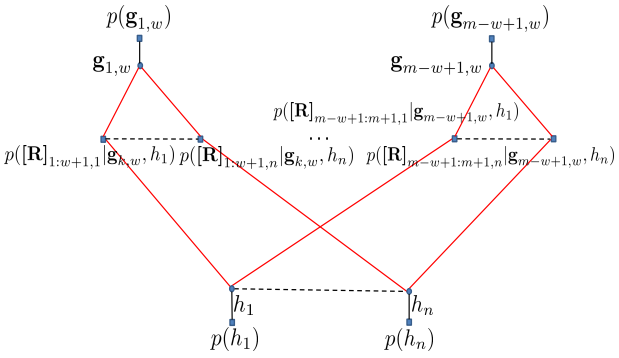

Fig. 4. Factor graph for problem (10): Cycle of order 7, a message before returning to the sender crosses seven different nodes.

IV. Simulation Results and Discussion

In this section, we evaluate the performance of the new attack and the refined message passing-based detector for the case of unbalanced priors. First, we show that the new (synchronized) attack strategy is more harmful than the symmetric flipping attack strategy. Then, we evaluate and discuss the performance of the MP-based decision fusion under various settings. Finally, we explain an unexpected effect according to which - in the setting considered in this paper - increasing the fraction of Byzantines α sometimes could be less harmful and results in a lower error probability at the FC.

Throughout the rest of the paper, we consider the following setting: a network with $n=20$ nodes, prior probabilities of the state $(P_0,P_1)=(0.7,0.3)$ and $(0.9,0.1)$, and a local error probability $\varepsilon=0.1$. The fraction of Byzantines in the network is $\alpha\in[0,\ldots,0.45]$ with a quantization step of 0.05, and $P^1_{mal}\in[0,\ldots,1]$ with a step of 0.1. Then, from (6), $P^0_{mal}=((1-\rho)/\rho)P^1_{mal}$, where $\rho=0.66$ when $(P_0,P_1)=(0.7,0.3)$ and $\rho=0.82$ when $(P_0,P_1)=(0.9,0.1)$. For the decision fusion, we consider two different observations window corresponding to $m=12$ and $m=21$. With regard to the MP algorithm, we consider the cardinalities $w\in\{1,2,3,4\}$ for the super-grouped variables $\mathbf{g}_{i,w}$.

To evaluate the performance of the MP algorithm, we estimate the error probability P_e at the FC over 5000 simulations.

A. Effectiveness of the new attack strategy

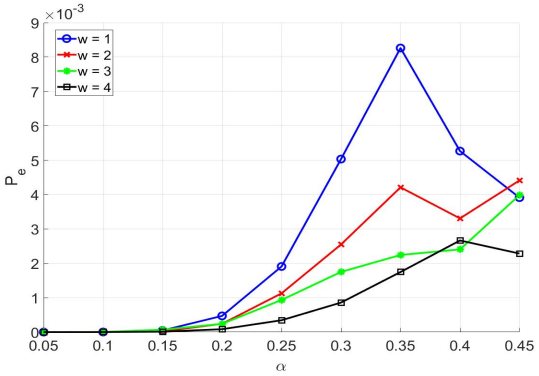

Fig. 5. P_e vs α for symmetric attack with $P_{mal}=0.5$ and $(P_0,P_1)=(0.9,0.1)$ for $w=\{1,2,3,4\}$.

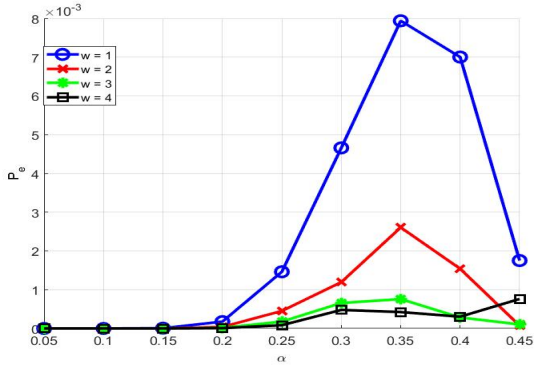

Fig. 6. P_e vs α for symmetric attack with $P_{mal}=1.0$ and $(P_0,P_1)=(0.9,0.1)$ for $w=\{1,2,3,4\}$.

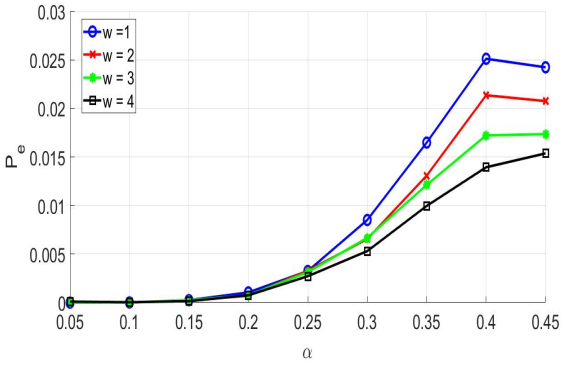

Fig. 7. P_e vs α for new asymmetric attack with $P^1_{mal} = 0.5$ and $(P_0, P_1) = (0.9, 0.1)$ for $w = \{1, 2, 3, 4\}$.

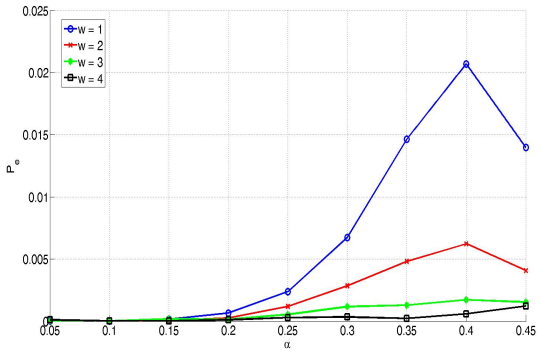

Fig. 8. P_e vs α for new asymmetric attack with $P^1_{mal} = 1.0$ and $(P_0, P_1) = (0.9, 0.1)$ for $w = \{1, 2, 3, 4\}$.

In the first simulation, we show that, when the system states are not equiprobable, the task of the FC is easier if the Byzantines adopt a symmetric attack channel, instead of the asymmetric attack described in Section II-B. In both cases we assume that the FC knows the attack strategy adopted by the Byzantines. Figures 5 through 8 show the performance of the decision fusion under the two different attack strategies as a function of the percentage of Byzantines nodes present in the network. For these experiments we use the following settings: $(P_0, P_1) = (0.9, 0.1)$ and $w = \{1, 2, 3, 4\}$.

To compare the two cases, we considered $P_{mal} = 0.5$ and $P_{mal} = 0.1$, for the case of symmetric attack, whereas, under the new attack, we report the results obtained with $P^1_{mal} \in \{0.5, 1.0\}$. This is fair comparison since P^1_{mal} corresponds to the largest between the two error probabilities in the BAC, so for instance, the cases $P_{mal} = 1$ and $P^1_{mal} = 1$, both correspond to the highest flipping rate that can be applied by the Byzantines. In the latter case, $P^0_{mal} = ((1-\rho)/\rho)P^1_{mal} \cong 0.15$ (which preserves the statistics).

As it can be seen by comparing Figure 5 to Figure 7, and Figure 6 to Figure 8, the asymmetric attack is more harmful than the symmetric attack. For instance, with $\alpha = 0.45$ and $w = 1$, for the case with $P_{mal}, P^1_{mal} = 1$, the error probability at the FC is $Pe = 0.0018$ under the symmetric attack and it increases up to $Pe = 0.014$ under the new attack as reported in Figure 6 and Figure 8, respectively. These results confirm the intuition that the symmetric attack does not work well in the case of unbalanced priors; hence, justifying the introduction of a new attack which keeps the statistics of the reports submitted by the Byzantines equal to those of the reports produced by the of the honest nodes.

We also notice that, in all the figures, the decision accuracy at the FC increases as the number of grouped variables increases. This is expected since the behavior of the message passing algorithm gets closer to the optimum fusion rule as the number of grouped variables increases. This improvement is achieved at the cost of an exponential growth of the execution time. For instance, in Figure 8, the execution times in seconds are: $5.6035 \times 10^{+3}s$ for $w = 1$, $1.3677 \times 10^{+4}s$ for $w = 2$, $5.7810 \times 10^{+4}s$ for $w = 3$, and $3.8611 \times 10^{+4}s$ for $w = 4$. Based on these results, we decided to let $w = 3$, which provides a good tradeoff between decision accuracy and computational complexity.

B. Performance of the decision fusion and optimum attacking strategy

In this section, we analyze the performance of the decision fusion for various settings. Specifically, we show the performance of P_e vs P^1_{mal} for the cases $\{P_0, P_1\} = \{0.7, 0.3\}$ and $\{0.9, 0.1\}$, $m = \{12, 21\}$ and $\alpha = [0.3, 0.4, 0.45]$. Then, based on such analysis, we discuss the choice of the optimal P^1_{mal}.

The error probabilities of the MP-based detector in the various cases are reported in Figures 9 through 12. As a first observation, we notice that when $(P_0, P_1) = (0.7, 0.3)$ the error probability is higher than in the case $(P_0, P_1) = (0.9, 0.1)$ for both observation windows $m = \{12, 21\}$. This is an expected behaviour since in general when $(P_0, P_1) = (0.9, 0.1)$ the FC has more a-priori information about the system state hence the decision is easier. Moreover, in all cases, using a larger observation results in a lower P_e.

From the figures, we observe that, thanks to the synchronization among the Byzantines, the attack is very powerful also for low values of P^1_{mal}. At a first sight, it may seem strange that the Byzantines tend to do not use all their available power and prefer to attack with a low P^1_{mal}, the optimum P^1_{mal} (the one resulting in the largest error probability) being well below 0.5 in all the cases. Such values are also smaller than the values corresponding to a zero mutual information between the reports submitted to the FC and the system state, which are $P^1_{mal} = 0.66$ and 0.82 for $(P_0, P_1) = (0.7, 0.3)$ and $(0.9, 0.1)$ respectively. Such a counter-intuitive behaviour can be explained as follows. The zero mutual information is clearly the best strategy for the Byzantines if we assume that the FC can correctly identify them. In such cases, in fact, the FC can still get some useful information about the state of the observed system, unless the mutual information between the reports and the system state is zero. However, by keeping

the flipping probability low, the Byzantines can avoid being identified and still induce an error due to the synchronisation of the attack. This would not be true with an asynchronous attack, since in such a case the low flipping probability would have no effect on the decision at the FC most of the times.

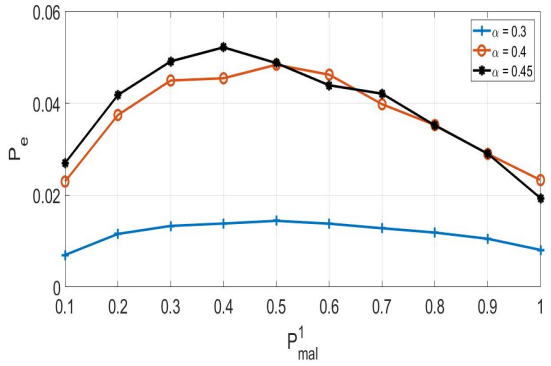

Fig. 9. P_e vs P^1_{mal} for $(P_0, P_1) = (0.7, 0.3)$, $m = 12$ and $\alpha = [0.3, 0.4, 0.45]$.

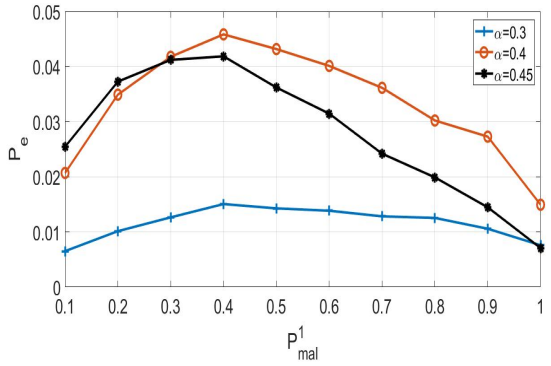

Fig. 10. P_e vs P^1_{mal} for $(P_0, P_1) = (0.7, 0.3)$, $m = 21$ and $\alpha = [0.3, 0.4, 0.45]$.

In some cases, most noticeably in Figure 10, we observe an unexpected behaviour of the error probability as a function α, i.e, the percentage of Byzantine nodes in the network: increasing the number of Byzantines in the network results in a lower error probability.

A possible explanation of this behaviour may be rooted in the fact that the FC has full knowledge of the attack strategy including P^1_{mal} and α, and it exploits such a knowledge to implement the decision fusion rule by means of the MP algorithm (see Section III). In a certain sense, a larger α forces the FC to be more cautious in the interpretation of the reports provided by the nodes (the MP algorithm is applied with a larger α) and this, in turn, allows to handle better the cases where the actual number of Byzantines is larger than αn, a value which represents only the average number of Byzantines

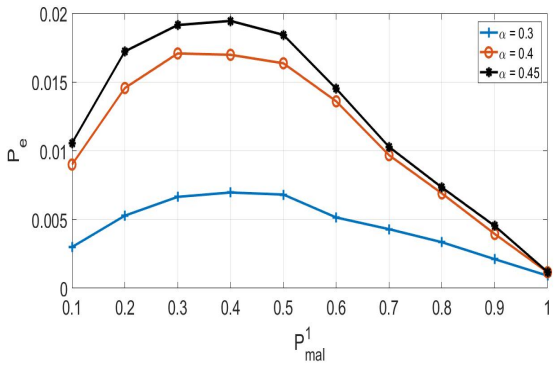

Fig. 11. P_e vs P^1_{mal} for $(P_0, P_1) = (0.9, 0.1)$, $m = 12$ and $\alpha = [0.3, 0.4, 0.45]$.

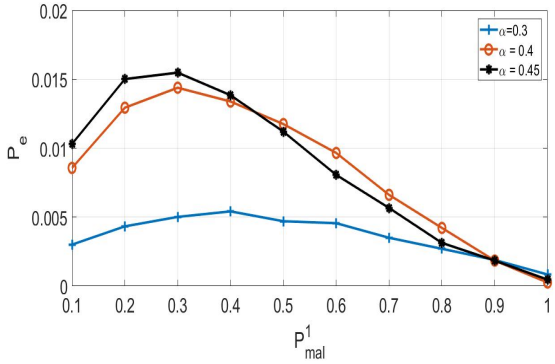

Fig. 12. P_e vs P^1_{mal} for $(P_0, P_1) = (0.9, 0.1)$, $m = 21$ and $\alpha = [0.3, 0.4, 0.45]$.

in the network. The impossibility of running the optimum fusion algorithm for the values of m when such an effect occurs ($m = 21$ in Figure 10), did not allow us to investigate whether this effect is due to the non-optimality of the MP algorithm or it is an intrinsic characteristic of the problem at hand.

V. Conclusions and Future Works

We have addressed a new variant of the decision fusion problem in the presence of Byzantines wherein the two states of the system under observation are not equiprobable. We have introduced a new unbalanced attack strategy for which the statistics of the reports submitted by the Byzantines are equal to those produced by the honest nodes, thus making it more difficult for the FC to isolate the corrupted nodes. Then, we have introduced a new fusion strategy, based on message passing, specifically thought to cope with the new unbalanced attack.

The analysis of the simulation results highlights some specific peculiarities of the unbalanced-priors setting (with synchronized attacks), the most relevant of which being the

possibility for the Byzantines to attack the system with a relatively small strength (low P^1_{mal}), yet causing a significant deterioration of the detection accuracy. The possibility that a smaller number of Byzantines results is a larger error probability is also worth attention and calls for further investigation.

Throughout the paper we have assumed that the FC is aware of the percentage of Byzantines present in the network and the flipping probability used to corrupt the reports. In future works, we plan to use game theory to relax such hypotheses.

REFERENCES

[1] A. Vempaty, T. Lang, and P. 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, Sept 2013.

[2] M. Abdelhakim, L. Lightfoot, J. Ren, and T. Li, "Distributed detection in mobile access wireless sensor networks under byzantine attacks," *IEEE Transactions on Parallel and Distributed Systems,*, vol. 25, no. 4, pp. 950–959, April 2014.

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

[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] W. Wang, H. Li, Y. Sun, and Z. Han, "Securing collaborative spectrum sensing against untrustworthy secondary users in cognitive radio networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, p. 4, 2010.

[6] B. Kailkhura, S. Brahma, and P. Varshney, "Optimal byzantine attacks on distributed detection in tree-based topologies," in *IEEE International Conference on Computing, Networking and Communications (ICNC)*, Jan 2013, pp. 227–231.

[7] M. Barni and B. Tondi, "Multiple-observation hypothesis testing under adversarial conditions," in *Proc. of WIFS'13, IEEE International Workshop on Information Forensics and Security*, Guangzhou, China, Nov 2013, pp. 91–96.

[8] M. Barni and F. Pérez-González, "Coping with the enemy: Advances in adversary-aware signal processing," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 8682–8686.

[9] Z. Chair and P. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-22, no. 1, pp. 98–101, Jan 1986.

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

[11] A. Abrardo, M. Barni, K. Kallas, and B. Tondi, "Decision fusion with corrupted reports in multi-sensor networks: A game-theoretic approach," in *53rd IEEE Conference on Decision and Control*, Dec 2014, pp. 505–510.

[12] R. Chen, J. M. Park, and K. Bian, "Robust distributed spectrum sensing in cognitive radio networks," in *Proc. of INFOCOM 2008, 27th IEEE Conference on Computer Communications*, April 2008, pp. –.

[13] A. Abrardo, M. Barni, K. Kallas, and B. Tondi, "A game-theoretic framework for optimum decision fusion in the presence of byzantines," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 6, pp. 1333–1345, June 2016.

[14] A. Abrardo, M. Barni, K. Kallas, and B. Tondi, "A Message Passing Approach for Decision Fusion in Adversarial Multi-Sensor Networks," *submitted to the Information Fusion Journal. ArXiv e-prints 1702.08357*, Feb 2017.

[15] ——, "A message passing approach for decision fusion of hidden-markov observations in the presence of synchronized attacks," in *Int. Conf. Advances in Multimedia (MMEDIA), Special Track on Models and Algorithms for Spatially and Temporally Correlated Data(STCD), Venice, Italy*, 2017.

[16] , "Spectrum policy task force report et docket no. 02-135," *US Federal Communications Commission*, 2002.

[17] V. V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Phil. Trans. R. Soc. A*, vol. 370, no. 1958, pp. 100–117, 2012.

[18] Y. Mao and A. H. Banihashemi, "A heuristic search for good low-density parity-check codes at short block lengths," in *Communications, 2001. ICC 2001. IEEE International Conference on*, vol. 1. IEEE, 2001, pp. 41–44.

[19] D. J. MacKay, *Information theory, inference and learning algorithms*. Cambridge university press, 2003.

[20] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.