Spoof Resilient Coordination for Distributed Multi-Robot Systems

Venkatraman Renganathan and Tyler Summers

***Abstract*— As cyber-physical networks become increasingly equipped with embedded sensing, communication, computation, and actuation capabilities, they are made vulnerable to malicious attacks by increasing the number of access points which are available for attackers. A particularly pernicious attack is spoofing, in which a malicious agent spawns multiple identities or impersonates legitimate agents to gain a disproportionate advantage. Spoofing attacks can easily compromise otherwise attack-resilient algorithms and network structures that assume an upper bound on the number of malicious agents in the network. We generalize a class of resilient consensus strategies, known as Weighted Mean-Subsequence-Reduced (W-MSR) consensus, to further provide spoof resilience by incorporating a physical fingerprint analysis of signals received from neighboring agents. By comparing the physical fingerprints of received signals, legitimate agents can identify and isolate malicious agents that attempt spoofing attacks. We quantify the effects of delays in detecting spoofing and inexact detection due to noise in the received signals. Numerical simulations illustrate the effectiveness of the proposed methods. Our framework is applicable to a variety of problems involving multi-robot systems coordinating via wireless communication, including coverage, distributed estimation, and formation control.**

I. Introduction

As cyber-physical networks become increasingly equipped with embedded sensing, communication, computation, and actuation capabilities, they are made vulnerable to malicious attacks by increasing the number of access points available for attackers. A large and growing literature has emerged on security, resilience, and robustness of cyber-physical systems in the presence of non-cooperative and adversarial agents [1], [2], [3], [4]. A particularly pernicious attack is spoofing[1], in which a malicious agent spawns multiple non-existent identities or impersonates existing legitimate agents to gain a disproportionate advantage in distributed algorithms that operate on the network. Spoofing is not just an abstract concern; successful attacks have been realized in several critical networks, such as civilian GPS [5], global navigation satellite systems [6], anti-lock braking systems [7], and others.

Autonomous multi-robot systems are a rapidly emerging type of cyber-physical network in which many tasks, including coverage, distributed estimation, cooperative manipulation, and formation control, utilize distributed consensus protocols to coordinate agreement on certain quantities of interest [15], [28]. Recent work has developed resilient consensus algorithms that prevent malicious agents from exerting undue influence on the network, effectively by designing sufficient redundancy in the algorithms and underlying network structures [12], [21], [13], [14], [16], [17], [18], [20]. These approaches have recently been applied to multi-robot systems [19], [10], [22]. However, spoofing attacks can easily compromise these otherwise attack-resilient algorithms and network structures, which assume an upper bound on the number of malicious agents in the network. Malicious information propagates through the network and can lead to severe performance deterioration or safety constraint violations.

It was argued in [8] that any defense against a spoofing attack requires either a trusted central authority to certify (perhaps cryptographically) the identities of all legitimate agents in the network, or a reliable method to distinguish physical fingerprints of signals received from neighboring agents. Reliance on a centralized authority is generally an undesirable feature in distributed multi-robot networks, so we focus here on the latter. Physical fingerprint analysis and discrimination has been used to detect spoofing in specific application contexts [5], [9], but not in the context of general distributed algorithms in cyber-physical networks or dynamic multi-robot systems.

Contributions: We propose an approach for spoof resilient coordination in cyber-physical and multi-robot networks. We generalize a class of resilient consensus strategies, known as Weighted Mean-Subsequence-Reduced (W-MSR) consensus, to provide spoof resilience by incorporating a physical fingerprint analysis of signals received from neighboring agents. By comparing the physical fingerprints of received signals, legitimate agents can detect and isolate malicious agents that attempt spoofing attacks. Our algorithm achieves resilient consensus despite an arbitrary number of spoofed agents in the network. We quantify the effects of delays in detecting spoofing and inexact detection due to noise in the received signals. We also demonstrate that it is possible to repair the effects of spoofing attacks after delayed detection by maintaining memory of received signals. Numerical simulations illustrate the effectiveness of the proposed methods. Our framework is applicable to a variety of problems involving multi-robot systems coordinating via wireless communication, including coverage, distributed estimation, and formation control.

The rest of the paper is organized as follows. Section II formulates a model for spoofing attacks in multi-robot networks and presents the attack detection technique using a physical fingerprint analysis. Section III proposes a new spoof resilient W-MSR algorithm by suitably modifying an existing standard W-MSR algorithm. Numerical simulations

V. Renganathan and T. Summers are with the Mechanical Engineering Department, University of Texas at Dallas, USA; email Addresses:{vrengana, tyler.summers}@utdallas.edu.

This work is supported by the National Science Foundation under Grant CNS-1566127.

[1] Also known as a "Sybil" attack [8].

results are then presented in Section IV. Finally, Section V summarizes the results and discusses future work directions.

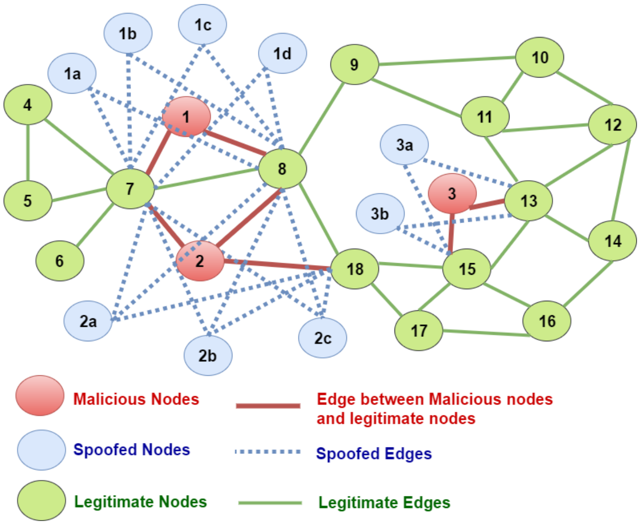

Fig. 1. **Spoofing Attack:** Malicious nodes in the network attempt to spoof multiple identities to gain disproportionate advantage over the network and disrupt the convergence of distributed consensus algorithms.

II. Spoofing Attacks In Multi-Robot Networks

A. *Spoofing Attack - Network Model*

Figure 1 illustrates an example of a spoofing attack we aim to address. We model the network with an undirected graph $\mathcal{G}$ comprising a node set $\mathcal{V}$ representing m agents and edge set $\mathcal{E}[t] \subset \mathcal{V} \times \mathcal{V}$ representing a set of (possibly time-varying) communication links amongst the agents. The node set is partitioned into two disjoint subsets $\mathcal{V} = S_l \cup S_a$. The set S_l represents the set of legitimate agents. A malicious agent attempts to disrupt the network by communicating subversive information to neighboring agents and *may* in addition attempt to perform a spoofing attack by creating multiple non-existent identities. Thus, the set of adversaries S_a is composed of both malicious and spoofed agents, so that $S_a = S_m \cup S_s$, where S_m denotes malicious agents and S_s denotes the agents spoofed by S_m. An upper bound of F number of malicious agents is assumed, whereas an arbitrary number of agents could be spoofed.

B. *Consensus Dynamics - Update Model*

We associate with each node $i \in \mathcal{V}$, a state $x_i[t] \in \mathbb{R}$ at time $t \in \mathbb{Z}_{\geq 0}$. The state may represent a position or some quantity to be estimated or optimized, depending on the application context. In order to achieve some objective, the nodes interact synchronously by exchanging their state value with neighbors in the network [24]. Each legitimate node updates its own state over time based on its current state and the state of neighboring agents according to a prescribed rule of the form

$$x_i[t+1] = f_i(x_j[t]),\ j \in \mathcal{J}_i[t] = \mathcal{N}_i[t] \cup \{i\},\ i \in S_l, \tag{1}$$

where $\mathcal{N}_i[t] = \{j \in \mathcal{V} : (j,i) \in \mathcal{E}[t]\}$ is the neighbor set of agent i at time t, whose states are available to agent i via communication links. The degree of i is denoted as $d_i[t] = |\mathcal{N}_i[t]|$, and every node is assumed to have access to its own state at time t. Resilient consensus algorithms specify a nonlinear function $f_i(.)$ that updates the states by suitably modifying which agents in the neighbor set (including i) $\mathcal{J}_i[t]$ are included in the update to provide resilience to malicious agents.

C. *Attack Detection Using Physical Fingerprint Analysis*

We imagine a scenario in which the agents in the network communicate amongst themselves using a wireless communication protocol. Physical properties of the received wireless signal profiles are leveraged to detect the spoofing attack. The physical fingerprint of an agent j received by agent i is modeled by a p-dimensional feature vector $\mathcal{F}_{ij} \in \mathbb{R}^p$ containing physical signal properties, such as angle-of-arrival, time-of-arrival and other features that can be used for discriminating between the signals received from two distinct agents [5], [9]. Since multiple spoofed agents may be generated by a single distinct malicious agent, the physical fingerprints associated with each of them will be similar. Of course, due to the random nature of wireless communication, there may be noise associated with the fingerprints of received signals. This situation can be modeled by associating a probability distribution with received signal fingerprints.

The neighbor set $\mathcal{N}_i$ of agent i includes the set of agents which can transmit signals to agent i. Based on the received signal fingerprints of pairs of neighboring agents, we define a *similarity metric*

$$\gamma_{ijk} = \frac{1}{1 + \|\mathcal{F}_{ij} - \mathcal{F}_{ik}\|}, \quad j,k \in \mathcal{N}_i, \tag{2}$$

which quantifies how similar the fingerprint of neighboring agent j is to that of neighboring agent k, as received by agent i. The authors in [9] deal with a similar setting for a coverage control problem and our development is inspired from their approach. Agent i computes these similarity metrics for each neighbor pair. From these similarity metrics, a *confidence weight* $\alpha_{ij} \in [0,1]$ can be associated with neighboring agent, which should be close 1 for legitimate neighbors and close to 0 for spoofed and spoofing neighbors. For example, in a deterministic setting,

$$\begin{aligned} \gamma_{ijk} = 1 \quad &\Rightarrow \quad \alpha_{ij} = 0,\ \alpha_{ik} = 0, \\ \gamma_{ijk} < 1 \quad &\Rightarrow \quad \alpha_{ij} = 1,\ \alpha_{ik} = 1, \end{aligned} \tag{3}$$

i.e., the confidence weights for neighbors j and k are 0 if the neighbor j has the same fingerprint as neighbor k, and 1 otherwise. In general, we can write

$$\alpha_{ij} = \prod_{j \in \mathcal{N}_i,\ j \neq k} (1 - \gamma_{ijk}). \tag{4}$$

In a stochastic setting, we define a spoof detection threshold $\omega \in [0,1]$. If the likelihood that the physical fingerprints of two neighbors are different is below the threshold, the neighbors are classified as spoofed or spoofing agents, and otherwise they are classified as legitimate; for example,

$$\begin{aligned} \mathbf{E}[g(\alpha_{ij})] \leq \omega \quad &\Rightarrow \quad j \text{ is spoofed or spoofing}, \\ \mathbf{E}[g(\alpha_{ij})] > \omega \quad &\Rightarrow \quad j \text{ is legitimate.} \end{aligned} \tag{5}$$

where $g(\cdot)$ is some prescribed likelihood function. To recover the deterministic case, we can set $\omega = 0$, $g(x) = x$, and drop the expectation.

D. Example of a Physical Fingerprint Model

A specific example of a particular physical fingerprint model is discussed in [9]. The fingerprint is modeled by a directional signal strength profile that depends on wireless signal wavelengths, distances and relative angles between directional antennae, multiple possible signal paths, and random channel properties with additive Gaussian noise. Based on this stochastic channel model, similarity and confidence metrics can be explicitly defined, and quantitative bounds can be obtained on the expectation that received signals are coming from spoofed agents. Such models could be used to define spoof resilient algorithms tailored to the specific communication model and perform analyses of probabilistic algorithm properties. Here we focus mainly on deterministic detection settings and stochastic settings with simple generic thresholding. Incorporating bounds associated with such specific communication models is left for future work.

III. Design of a Spoof Resilient Coordination Algorithm

In this section, we describe a coordination algorithm that is resilient to malicious agents who share adversarial state values and *may* also attempt to spoof non-existent agent that also share adversarial state value. Since malicious agents do not *all* necessarily attempt to spoof, we build upon recent work on resilient consensus algorithms that do not handle spoofing. These algorithms achieve resiliency by effectively designing and exploiting redundancy in the underlying communication graph. We now review these resilient graph properties and an existing resilient consensus algorithm called Weighted Mean-Subsequence-Reduced (W-MSR) as described in [14]. We subsequently present our spoof resilient adaptation of W-MSR. We assume throughout that there are at most F malicious agents but that there may be an arbitrary number of spoofed agents.

Definition: A set $\mathcal{S}$ is *r-reachable*, if it contains a node that has at least r neighbors outside of $\mathcal{S}$. The parameter r quantifies the redundancy of information flow from nodes outside of $\mathcal{S}$ to some node inside $\mathcal{S}$. Intuitively, the r-reachability property captures the notion that some node inside the set is influenced by a sufficiently large number of nodes from outside the set.

Definition: A graph $\mathcal{D} = \{\mathcal{V}, \mathcal{E}\}$ on n nodes is said to be *r-robust*, with $r \in \mathbb{Z}_{\geq 0}$, if for every pair of disjoint nonempty subsets of $\mathcal{V}$, at least one of the subsets is r-reachable.

Definition: Given a graph $\mathcal{D}$ and a nonempty subset of nodes $\mathcal{S}$, we say that $\mathcal{S}$ is an *(r,s) - reachable set*, if there are at least s nodes in $\mathcal{S}$, each of which has at least r neighbors outside of $\mathcal{S}$, where $r, s \in \mathbb{Z}_{\geq 0}$. i.e., given $\mathcal{X}_{\mathcal{S}} = \{i \in \mathcal{S} : |\mathcal{V}_i \backslash \mathcal{S}| \geq r\}$, then $|\mathcal{X}_{\mathcal{S}}| \geq s$.

Definition: A graph $\mathcal{D} = \{\mathcal{V}, \mathcal{E}\}$ on n nodes $(n \geq 2)$ is *(r,s)-robust*, for nonnegative integers $r \in \mathbb{Z}_{\geq 0}$, $1 \leq s \leq n$, if for every pair of nonempty, disjoint subsets $\mathcal{S}_1$ and $\mathcal{S}_2$ of $\mathcal{V}$ such that $\mathcal{S}_1$ is $(r, s_{r,1})$-reachable and $\mathcal{S}_2$ is $(r, s_{r,2})$-reachable with $s_{r,1}$ and $s_{r,2}$ maximal (i.e., $s_{r,k} = |\mathcal{X}_{\mathcal{S}k}|$ where $\mathcal{X}_{\mathcal{S}k} = \{i \in \mathcal{S}_k : |\mathcal{V}_i \backslash \mathcal{S}_k| \geq r\}$ for k = $\{1, 2\}$), then at least one of the following hold:

- $s_{r,1} = |\mathcal{S}_1|$
- $s_{r,2} = |\mathcal{S}_2|$
- $s_{r,1} + s_{r,2} \geq s.$

The (r, s)-robustness property introduces information redundancy by specifying a minimum number of nodes that are sufficiently influenced from outside of their set. Note that (r, s)-robustness is a strict generalization of r-robustness.

A. Resilient Asymptotic Consensus

Let $x_M[t]$ and $x_m[t]$ denote the maximum and minimum values of the legitimate nodes at time t, respectively. The legitimate agents in the network are said to achieve *resilient asymptotic consensus* [18] in the presence of a particular threat model if for any initial conditions it holds

- $\exists$ L $\in \mathbb{R}$ such that $\lim_{t \to \infty} x_i[t] = L, \forall i \in S_l$
- the interval $[x_m[0], x_M[0]]$ is an invariant set (i.e., the legitimate values remain in the interval $\forall t$)

Resilient asymptotic consensus has three important properties [16]. First, the legitimate nodes must reach asymptotic consensus despite the presence of misbehaving nodes given a particular threat model and scope of threat (e.g., at most F malicious agents). This is a condition on *agreement*. Additionally, it is required that the interval containing the initial values of the legitimate nodes is an invariant set for the legitimate nodes; this is a *safety* condition, where the interval $[x_m[0], x_M[0]]$ is known to be safe. The agreement and safety conditions, when combined, imply a third condition on *validity*: the converged consensus value lies within the range of initial values of the legitimate nodes.

B. The Weighted Mean-Subsequence-Reduced (W-MSR) Algorithm

We now review a class of resilient consensus algorithms described in [14] that utilize a Weighted Mean-Subsequence-Reduced (W-MSR) update rule. At every time t, each legitimate node i obtains the values of other nodes in its neighborhood. Since there are at most F total malicious nodes in the network, some of node i's neighbors may misbehave; however, node i is unsure of which neighbors may be compromised. To ensure that node i updates its state in a safe manner, we consider a protocol where each node removes the extreme values with respect to its own value. Specifically, the W-MSR algorithm comprises the following steps:

1) At each time t, each legitimate node $i \in S_l$ obtains the state values of its neighbors, and forms a sorted list.
2) If there are less than F values strictly larger than its own value, $x_i[t]$, then legitimate node i removes all

values that are strictly larger than its own. Otherwise, it removes precisely the largest F values in the sorted list (breaking ties arbitrarily). Likewise, if there are less than F values strictly smaller than its own value, then node i removes all values that are strictly smaller than its own. Otherwise, it removes precisely the smallest F values.

3) Let $\mathcal{R}_i[t]$ denote the set of nodes whose values were removed by legitimate node i in step 2 at time t. Each legitimate node i applies the update

$$x_i[t+1] = \sum_{j \in \mathcal{J}_i[t] \setminus \mathcal{R}[t]} w_{ij}[t] x_j[t] \tag{6}$$

where $w_{ij}[t]$ is the weight[2] associated with node j's value by node i at time step t. The weights are chosen to satisfy the following conditions:

1) $w_{ij}[t]$ = 0 whenever j $\notin \mathcal{J}_i[t], i \in S_l, t \in \mathbb{Z}_{\geq 0}$
2) there exists a constant $\beta \in \mathbb{R}, 0 < \beta < 1$ such that $w_{ij}[t] \geq \beta, \forall j \in \mathcal{J}_i[t], i \in S_l, t \in \mathbb{Z}_{\geq 0}$
3) $\sum_{j=1}^{n} w_{ij}[t] = 1, \forall i \in S_l, t \in \mathbb{Z}_{\geq 0}$

Then if the underlying graph is $(F+1, F+1)$-robust, under the update protocol specified in equation (6), the legitimate agents in the network are guaranteed achieve resilient asymptotic consensus [14] despite the presence of at most F malicious agents, but assuming that there are no spoofed agents.

C. *Spoof Resilient W-MSR Algorithm*

A spoofing attack is capable of compromising the $(F+1, F+1)$-robust graph robustness property and W-MSR algorithm above, and hence the network resiliency. Our spoof resilient adaptation of the W-MSR algorithm here is summarized in Algorithm 1. Based on a pairwise comparison of physical fingerprints of signals received from neighboring agents and associated confidence weights, spoofed agents in the network are identified. Achieving resiliency then involves removing the identified spoofed and spoofing agents from the state update if the expectation of some likelihood of their confidence weight is at most equal to the spoofing threshold ω. In a deterministic setting, we have the following result.

Theorem 1: Given an undirected network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ represents the set of agents and $\mathcal{E}$ represents the set of communication links between them. Suppose the network is $(F+1, F+1)$-robust, assuming an upper bound of F total malicious agents in the network, *some* of which may spoof. Then the network achieves resilient asymptotic consensus under Algorithm 1 in the presence of any spoofing attack.

Proof: In a deterministic setting, the physical fingerprints of signals of spoofed agents are identical to that of the spoofing agent. So any spoofed and spoofing agents are exactly detected and removed from the network state updates in lines 17-20 of the proposed Algorithm 1. Moreover, since the initial underlying graph is $(F+1, F+1)$-robust, the use of the W-MSR protocol for the state updates guarantees resilient asymptotic consensus in the presence of up to F malicious nodes who do not spoof but may behave in other adversarial ways. Thus, the overall protocol is resilient to *both* an arbitrary number of spoofed agents and up to F non-spoofing malicious agents in the network. ■

Algorithm 1 Spoof Resilient W-MSR (SR-W-MSR)

```
 1: procedure SR-W-MSR($\omega$)
 2:     // Input: spoofing threshold $\omega$, convergence threshold $\epsilon$, initial states x[0], received signal fingerprints $\mathcal{F}_{ij}$ for each agent at each time
 3:     $t \leftarrow 0$
 4:     while $\|x[t+1] - x[t]\| \geq \epsilon$ do
 5:         $i \leftarrow 1$
 6:         // Iterate through all legitimate nodes
 7:         while $i \leq |S_l|$ do
 8:             for each $j \in \mathcal{N}_i$ do
 9:                 $\alpha_{ij} \leftarrow 1$
10:                 for each $k \in \mathcal{N}_i$ do
11:                     // Pairwise comparison of neighbors
12:                     if $j \neq k$ then
13:                         $\gamma_{ijk} = \frac{1}{1+\|\mathcal{F}_{ij}-\mathcal{F}_{ik}\|}$
14:                     end if
15:                     $\alpha_{ij} \leftarrow \alpha_{ij}(1-\gamma_{ijk})$
16:                 end for
17:                 if $\mathbf{E}[g(\alpha_{ij})] \leq \omega$ then
18:                     // Spoof attack is detected
19:                     $\mathcal{Z}[t] \leftarrow \mathcal{R}[t] \cup \{j\}$
20:                     $x_i[t+1] \leftarrow \sum_{j \in \mathcal{J}_i[t] \setminus \mathcal{Z}[t]} w_{ij}[t] x_j[t]$
21:                 else
22:                     // No spoof attack is detected
23:                     $x_i[t+1] \leftarrow \sum_{j \in \mathcal{J}_i[t] \setminus \mathcal{R}[t]} w_{ij}[t] x_j[t]$
24:                 end if
25:             end for
26:             $i \leftarrow i+1$
27:         end while
28:         $t \leftarrow t+1$
29:     end while
30: end procedure
```

IV. Numerical Simulation Results

We now illustrate our spoof resilient W-MSR algorithm using the 7-node (2,2)-robust network with 6 legitimate agents, 1 malicious agent who spoofs 1 agent as shown in Figure 2. The six legitimate agents were given random initial states, and with states of the malicious and spoofed agents set to 300. Malicious node 4 performs spoofing attack by emulating another non-existent spoofed identity called 4a and sends the message to all his neighbors. The similarity metric (γ_{ijk}) between neighbors j and k recorded by the agent i was drawn from an uniform distribution on the interval [0,1]. A spoofing attack is evaluated both in deterministic and stochastic settings. In the deterministic setting, all the physical fingerprints are obtained without any noise. As a

[2] In this case, a simple choice for the weights [13] is to let $w_{ij}[t] = 1/(1 + d_i[t] - |\mathcal{R}_i[t]|)$, for $j \in \mathcal{J}_i[t] \setminus \mathcal{R}[t]$

result, the spoofed nodes are exactly identified and removed from the network using the spoof resilient W-MSR algorithm. In the stochastic setting, a spoofing threshold is employed and spoofed agents are only identified and removed probabilistically in the spoof resilient W-MSR algorithm.

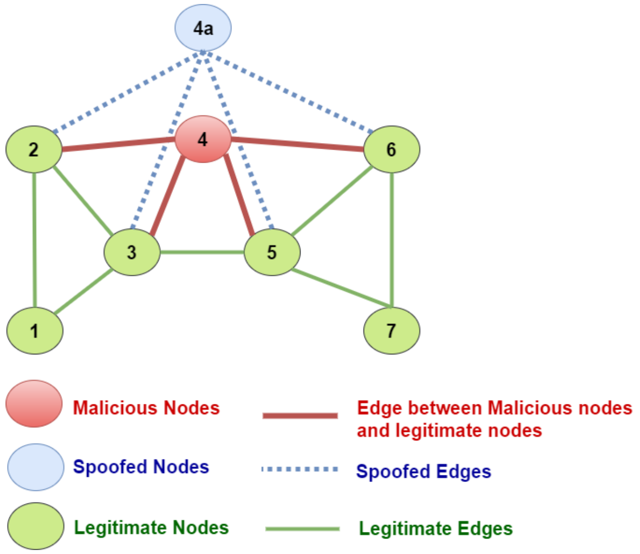

Fig. 2. A (2,2)-robust graph with node 4 being malicious and it spoofs 1 more agent (4a) to gain disproportionate advantage over the network.

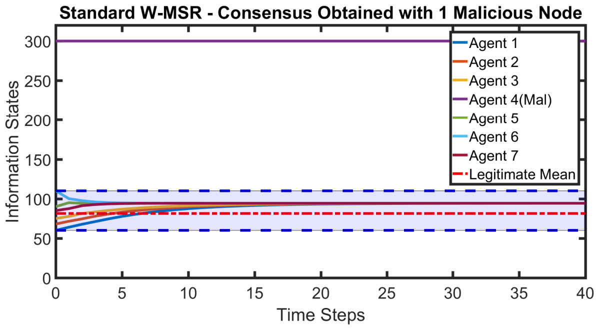

Fig. 3. Under the W-MSR algorithm *without* spoofing, the legitimate agents achieve resilient asymptotic consensus despite the malicious agent 4 and converge to a value within the range of their initial values shown as the shaded region

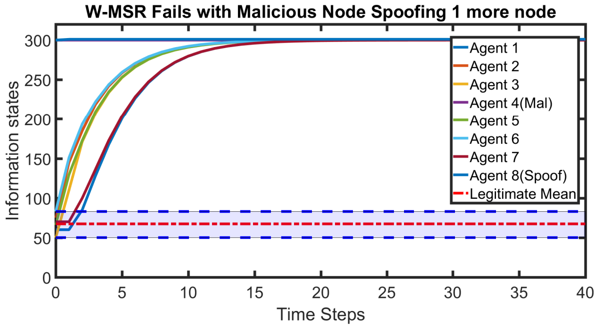

Fig. 4. When the malicious agent 4 spoofs a single additional identity, the legitimate agents fail to achieve resilient asymptotic consensus, and the state values are pulled to the value of agent #4, showing failure of W-MSR algorithm under spoofing. Also, node 4 malicious and node 8 (spoofed) take the same value 300 in the figure.

A. *The W-MSR Algorithm Fails Under a Spoofing Attack*

Consider a spoofing attack on the network shown in Figure 2, where agent 4 is malicious. When a standard linear consensus protocol is used, the malicious agent is able to pull the consensus values of other agents in the network to his desired value even without spoofing, just by maintaining its state to a constant value. When the W-MSR resilient consensus algorithm [14] is used, the legitimate agents achieve resilient asymptotic consensus, as shown in Figure 3, despite the presence of the malicious agent and despite them being unaware about the identity of the malicious agent. However, when malicious agent spoofs a single additional agent identity, the legitimate agents fail to achieve resilient asymptotic consensus, as shown in Figure 4 where all states converge to the malicious agent's state value. This shows that spoofing attacks are capable of compromising graph robustness properties and thereby the network resiliency. Under the proposed Spoof Resilient W-MSR algorithm, the legitimate agents achieve resilient asymptotic consensus, as shown in Figure 5.

B. *Variations: Delayed and Probabilistic Spoof Detection*

In practical settings, it may take a non-trivial amount of time to detect a spoofing attack, and spoofing may not be detected perfectly due to noise and uncertainty in the received signal properties. Here we provide some preliminary quantifications of these effects. First, Figure 6 shows the effect of detection delays on the final consensus value. As expected, with longer delays in detection of spoofing, the malicious agents can have a larger effect on the final value. Further, it is shown that a malicious agent's location in the networks affects how influential it can be in perturbing the state of the network; a spoof attack by agent 7 with lower degree than agent 4 has a lower impact. However, it is also possible to repair the impact of a spoofing attack on the network after delayed detection by maintaining memory of neighboring state transmission histories and subtracting out modifications made by spoofed and spoofing agents, as shown in Figure 7. Of course, such a bookkeeping effort would be limited by memory constraints, and may be cumbersome to implement in complicated networks. We are pursuing how this might be achieved in general for future work.

Finally, Figure 8 illustrates the effects of inexact detection of a spoofing attack, where a Monte Carlo estimate of the expected value of the final consensus value is plotted against the probability of spoof detection. As expected, it can be seen that when spoofing is detected with lower probability, the malicious agents have a larger effect on the state of the network. Sharper probabilistic analyses and algorithmic modifications in the stochastic detection setting are also being pursued in future work.

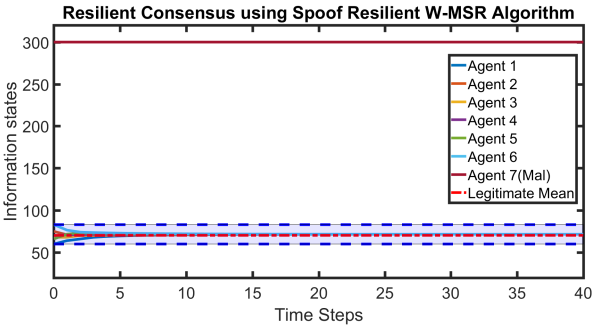

Fig. 5. The proposed Spoof Resilient W-MSR Algorithm achieves resilient asymptotic consensus.

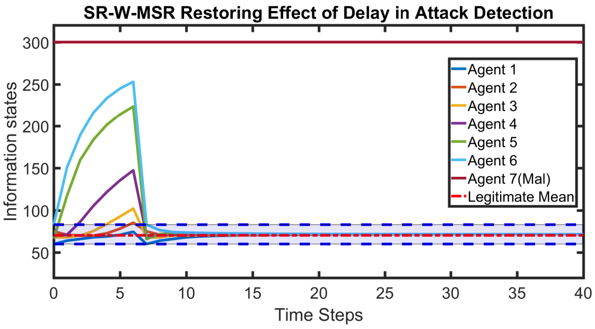

Fig. 7. Spoof Resilient W-MSR restoring the effect of delay of 7 time steps of attack detection.

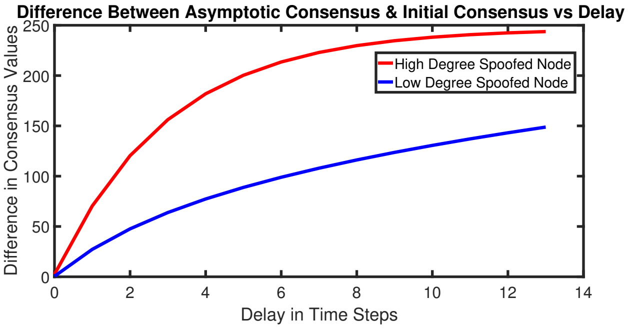

Fig. 6. Difference between asymptotic consensus value and ideal consensus value when a spoofing attack is launched by agent 4 with high degree or agent 7 with lower degree in the network shown in figure 2.

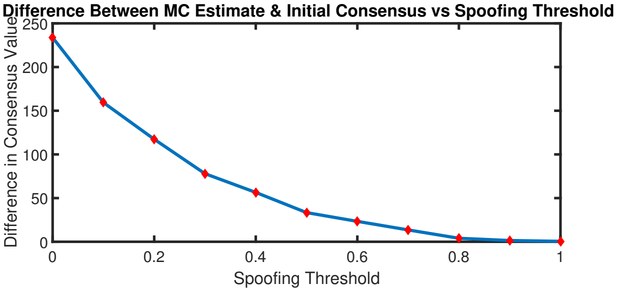

Fig. 8. Difference between Monte Carlo estimate of the asymptotic consensus value and ideal consensus value for different spoof classification threshold values with zero delay in attack detection.

V. Conclusions

We proposed a spoof resilient consensus algorithm that extends a class of resilient consensus strategies, known as Weighted Mean-Subsequence-Reduced (W-MSR) consensus, to provide resilience to malicious agents that may both adversely update state values and spoof non-existent agent identities. Physical fingerprint comparisons of received signals are used by legitimate agents to identify and isolate malicious agents that attempt spoofing attacks. The proposed algorithm using physical fingerprint approach guarantees resiliency despite the presence of a certain number of malicious agents and an arbitrary number of spoofed agents in the network. The proposed framework is applicable to a variety of problems involving multi-robot systems coordinating via wireless communication, including coverage, distributed estimation, and formation control. Future research involves investigating the resiliency against spoofing attack with different fault models and using specific probabilistic physical fingerprint models.

References

[1] Q. Zhu and T. Başar, "Game-Theoretic Methods for Robustness, Security, and Resilience of Cyberphysical Control Systems," *IEEE Control Systems Magazine*, February 2015.

[2] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. Sastry, "Challenges for securing cyber physical systems," in *Workshop on future directions in cyber-physical systems security*, 2009.

[3] A. Banerjee, K. Venkatasubramanian, T. Mukherjee, and S. Gupta, "Ensuring safety, security, and sustainability of mission-critical cyberphysical systems," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 283-299, 2012.

[4] R. Albert, H. Jeong and A. L. Barabasi, "Error and attack tolerance of complex networks," *Letters to Nature*, Vol. 406, July 2000.

[5] T. E. Humphreys, B. M. Ledvina, M. L. Psiaki, B. W. O'Hanlon and P. M. Kintner Jr., "Assessing the Spoofing Threat: Development of a Portable GPS Civilian Spoofer," *Proceedings of the 21st International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2008)*, Savannah, GA, Sept 2008, pp. 2314-2325.

[6] M. L. Psiaki and T. E. Humphreys, "GNSS Spoofing and Detection," *in Proceedings of the IEEE*, vol. 104, no. 6, pp. 1258-1270, June 2016.

[7] Y. Shoukry, P. Martin, P. Tabuada, and M. Srivastava, "Non-invasive Spoofing Attacks For Anti-Lock Braking Systems," *Proceedings of the 15th International Workshop on Cryptographic Hardware and Embedded Systems 2013 (CHES 2013)*, Santa Barbara, CA, USA, Aug. 20-23, 2013, pp. 55-72.

[8] J. R. Douceur, "The Sybil Attack," *In Revised Papers from the First International Workshop on Peer-to-Peer Systems (IPTPS '01)*, Springer-Verlag 2002, London, UK, pp. 251-260.

[9] S. Gill, S. Kumar, M. Mazumdar, D. Katabi and D. Rus, "Guaranteeing Spoof-Resilient Multi-Robot Networks," *Robotics: Science and Systems Conference*, 2015.

[10] L. Guerrero-Bonilla, A. Prorok and V. Kumar, "Formations for Resilient Robot Teams," *IEEE International Conference on Robotics and Automation (ICRA)*, 2017.

[11] R. Olfati-Saber, J. A. Fax and R. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," *In Proceedings of the IEEE*, Vol. 95, No. 1, pp. 215-233, Jan. 2007.

[12] H. Zhang and S. Sundaram, "Network Robustness: Diffusing Information Despite Adversaries," *Technical report for WICI*, 2013.

[13] H. Zhang, "Network Robustness: Diffusing Information Despite Adversaries," *MS Thesis*, University of Waterloo, Canada, 2012.

[14] H. Zhang, E. Fata and S. Sundaram, "A Notion of Robustness in Complex Networks," *IEEE Transactions on Control of Network Systems*, Vol.2, No. 3, pp. 310-320, 2015.

[15] W. Ren, R. W. Beard and E. M. Atkins, “Information Consensus in Multi-vehicle Cooperative Control,” *IEEE Control Systems Magazine*, April 2007.

[16] H. J. LeBlanc, H. Zhang, X. Koutsoukos and S. Sundaram, “Resilient asymptotic consensus in robust networks,” *IEEE Journal on Selected Areas in Communications*, Vol. 31, No. 4, pp. 766-781, 2013.

[17] H. J. LeBlanc, H. Zhang, S. Sundaram and X. Koutsoukos, “Consensus of multi-agent networks in the presence of adversaries using only local information,” *Proceedings of the* 1^{st} *ACM International Conference on High Confidence Networked Systems*, 2012, pp. 1-10.

[18] H. Zhang and S. Sundaram, “A simple median-based resilient consensus algorithm,” *Annual Allerton Conference on Communication, Control, and Computing*, pp. 1734-1741, 2012.

[19] K. Saulnier, D. Salda$\tilde{n}$a, A. Prorok, G. J. Pappas and V. Kumar, “Resilient Flocking for Mobile Robot Teams,” *IEEE Robotics and Automation Letters*, Vol. 2, No. 2, pp. 1039-1046, April 2017.

[20] F. Pasqualetti, A. Bicchi and F. Bullo, “Consensus Computation in Unreliable Networks: A System Theoretic Approach,” *IEEE Transactions on Automatic Control*, Vol. 57, No. 1, January 2012.

[21] H. Zhang and S. Sundaram, “Robustness of Information Diffusion Algorithms to Locally Bounded Adversaries,” *American Control Conference (ACC)*, 2012.

[22] D. Salda$\tilde{n}$a, A. Prorok, M. F. M. Campos and V. Kumar, “Triangular Networks for Resilient Formations,” 13^{th} *International Symposium on Distributed Autonomous Robotic Systems (DARS)*, 2016.

[23] S. Sundaram and C. N. Hadjicostis, “Distributed Function Calculation via Linear Iterative Strategies in the Presence of Malicious Agents,” *IEEE Transactions on Automatic Control*, Vol.56, No. 11, July 2011.

[24] J. N. Tsitsiklis, D. P. Bertsekas and M. Athans, “Distributed Asynchronous Deterministic and Stochastic Gradient Optimization Algorithms,” *IEEE Transactions on Automatic Control*, Vol. AC-31, No. 9, pp. 803-812, September 1986.

[25] A. Jadbabaie, J. Lin and S. Morse, “Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules,” *IEEE Transactions on Automatic Control*, Vol. 48, No. 6, June 2003, pp. 988-1001.

[26] Y. Chen, W. Trappe and R. P. Martin, “Detecting and Localizing Wireless Spoofing Attacks,” *2007 4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, San Diego, CA, 2007, pp. 193-202.

[27] D. Moser, P. Leu, V. Lenders, A. Ranganathan, F. Ricciato, and S. Capkun. 2016, “Investigation of multi-device location spoofing attacks on air traffic control and possible countermeasures,” *In Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking (MobiCom ’16) ACM*, New York, NY, USA, pp. 375-386.

[28] J. Cortes, S. Martinez, T. Karatas, F. Bullo, “Coverage control for mobile sensing networks,” *IEEE Transactions on Robotics and Automation*, Vol. 20, No. 2, pp. 243-255, IEEE, 2004.