Switching Topology for Resilient Consensus using Wi-Fi Signals

Thomas Wheeler[1] and Ezhil Bharathi[1] and Stephanie Gil[2]

***Abstract*— Securing multi-robot teams against malicious activity is crucial as these systems accelerate towards widespread societal integration. This emerging class of "physical networks" requires new security methods that exploit their physical nature. This paper derives a theoretical framework for securing multi-agent consensus against the Sybil attack by using the physical properties of wireless transmissions. Our framework uses information extracted from the wireless channels to design a switching signal that stochastically excludes potentially untrustworthy transmissions from the consensus. Intuitively, this amounts to selectively ignoring incoming communications from untrustworthy agents, allowing for consensus to the true average to be recovered with high probability after a certain observation time T_0. This paper allows for arbitrary malicious node values and is insensitive to the initial topology of the network so long as a connected topology over legitimate nodes in the network is feasible. We show that our algorithm will recover consensus, and the true graph over the system of legitimate agents, with an error rate that vanishes exponentially with time.**

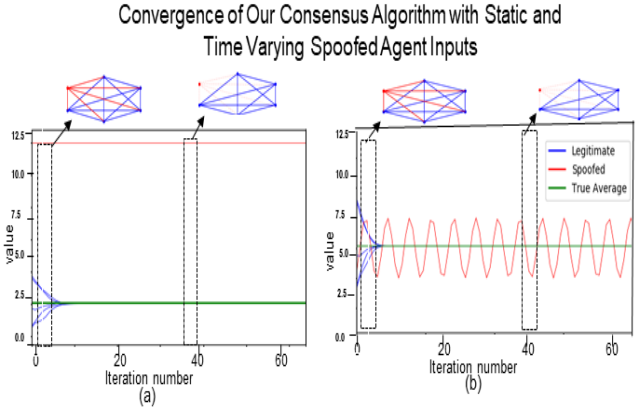

Fig. 1. Convergence of consensus using our protocol and switching function for static (left col) and time varying (right) spoofed node input.

I. Introduction

Multi-robot systems are at the horizon of wide-spread integration into our societies; as fleets of autonomous vehicles, delivery drones, and smart and mobile Internet of Things (IoT) devices. This promise has spurred a great amount of research both in academia and industry, accelerating the pace at which this vision can come to fruition. However, as these multi-robot systems become pervasive, security becomes paramount. At the heart of many multi-robot coordination tasks are algorithms such as coverage or consensus which are inherently vulnerable to adversarial action such as the Sybil attack. In the Sybil attack, a single adversarial node can "spoof" or generate a large number of spurious entities in the graph as a way of gaining a disproportionate influence in the network [1], [2]. In this way the Sybil attack can be detrimental to several critical multi-robot algorithms [3]. Additionally, the dynamic and distributed nature of multi-robot systems makes it particularly difficult to implement traditional methods of security such as authentication and key passing [4], [5], making it even more challenging to secure these systems against a Sybil attack. Interestingly, the physicality of these systems presents new opportunities for security. Recent developments in Wi-Fi characterization [6], [7], [8] demonstrate the ability to extract information from communicated wireless signals that can be used to detect malicious actors in the network. As a wireless signal propagates between communicating agents it is reflected, absorbed, and scattered by objects in the environment in a phenomenon called multipath [9]. This multipath signature can act as a "fingerprint" of the transmission [6], with clear implications for security. This suggests a new approach for the security of physical networks and multi-robot coordination that does not rely on additional data beyond communication signals present in the system, or additional overhead such as key passing for authentication. However, to fully exploit the use of wireless signals for thwarting attacks on multi-robot systems we need new theory that exploits this latent information in the network to improve achievable consensus results.

Towards this end, the focus of the current paper is to develop a theoretical framework for using observations of inter-agent wireless channels to secure consensus in the face of a Sybil attack and to characterize expected performance guarantees for this case. It is well known that in the presence of uncooperative or adversarial nodes, the ability to achieve consensus is severely impaired or lost [1]. Given a network of agents that includes an unknown subset of spoofed nodes, we wish to achieve convergence to a common consensus value over the legitimate nodes that is irrespective of spoofed node input to the system (see Figure 1). The key insight of our approach is to design a switching signal that selectively ignores or includes agents in a stochastic manner, according to an honesty metric [6], derived from multiple observations of their wireless channels (for each transmission in the network). Intuitively, this switching signal learns the legitimate network topology (one that excludes spoofed nodes) with an error rate that decreases exponentially fast in the number of

[1]T. Wheeler and E.Bharathi (tswheel1, epoongod)@asu.edu, are with the REACT Lab, Arizona State University, Tempe, Arizona 85281 USA

[2] S.Gil is the director of the REACT Lab and Assistant Professor of Computer Science, Arizona State University, Tempe, Arizona 85281 USA sgil4 at asu.edu

The authors gratefully acknowledge partial support from the Fulton Undergraduate Research Initiative (FURI).

observations of the wireless channel. By deriving both the switching signal and a modified consensus algorithm, we show that legitimate agents can achieve agreement and that this agreement value can be forced arbitrarily close to the true average with high probability that is also characterized in this paper.

This work improves upon existing work in adversarial consensus [10], [11], [7] in several ways: i) Resilience is achieved irrespective of the values transmitted by the spoofed nodes, ii) it does not assume knowledge of the number of spoofed nodes in the network, and iii) it is robust to different initial topology configurations.

Paper Contributions: The current paper provides the following contributions: 1) derivation of a consensus algorithm and switching signal for resilient consensus using wireless channel observations, 2) analysis demonstrating the convergence properties of this algorithm, and 3) an extensive simulation study demonstrating aggregate performance of our algorithm for different initial topologies and static vs. time varying spoofed node inputs.

II. Related Work

Consensus is a critical algorithm for cooperative multi-agent decision making and control. Traditional assumptions of cooperation and trust within the network leave consensus algorithms inherently vulnerable to adversarial attacks. The Sybil attack is one such classic example. Cryptographic and key passing schemes are one common approach to securing multi-agent consensus [4], [5], [12]. These schemes provide excellent security for typical networks but require computational and communication overhead and trust in a central authority. Alternatively, data-based methods for resilient consensus can be characterized by analyzing transmitted values to deduce the malicious nature of the agents, leaving them vulnerable to sophisticated attackers that may intentionally manipulate their values to avoid detection. Research in Mean Sub-sequence Reduced (MSR) [13], [14], [15] focus on transmitted values in order to remove adversarial agent influence from the consensus.

Methods relying on the physics of the communication signals themselves offer an interesting alternative to data-based methods. Mobile robot systems often communicate over wireless radio signals and it has been shown that useful information can be extracted from these signals [16], [17]. Traditionally, these types of approaches would require expensive hardware, like multi-antenna arrays, that cannot be mounted on small robotic platforms. However, recent techniques based on Synthetic Aperture Radar (SAR) [18] have been adapted to cheap commodity Wi-Fi radios suitable for robotic platforms [8], [19]. This opens up a theoretically rich set of physical sensing capabilities, without needing to rely on additional hardware. Moreover, prior work has already demonstrated some practical applications in localization [19], coverage [6], [20], and consensus [7].

This paper builds upon advances in adversary detection using Wi-Fi signals [6], [21], and rich theory in consensus over switching or time-varying graphs [22], [23], [24], [25], [26], [27], [28], [29], [30], [31] to derive a mathematical framework for recovering average consensus in the case of a Sybil attack. This work is different from data-based methods [10], [11] because it makes observations over the wireless network, allowing our algorithm to be largely independent of transmitted values. This is generalizable to a larger set of attacker behaviors including dynamically changing input values as shown in our simulation studies in the Simulation Section.

III. Problem

We consider the problem of distributed consensus for multi-agent systems where a subset of nodes in the graph are adversarial and attempt to disturb the consensus ability of the network by inputting false values. Specifically, the multi-robot system can be described by a weighted, undirected, state-dependent graph $\mathbb{G}(t) = (\mathbb{V}, \mathcal{E}(t))$, where $\mathbb{V} = \{1, \dots, N\}$ denotes the set of node indices for N robots and $\mathbb{W} : \mathbb{V} \times \mathbb{V} \times \mathbb{R}_+ \to \mathbb{R}_+$ denotes the set of edge weights at time t such that $w_{ij}(t) = \mathbb{W}(i, j, t)$ for $i, j \in \mathbb{V}$. The set $\mathcal{E}(t) = \{(i, j) | w_{ij}(t) > 0\}$ is called the set of undirected edges of $\mathbb{G}$. The set of *neighbors* of node i is denoted by $\mathcal{N}_i(t) = \{j \in \mathbb{V} : (i, j) \in \mathcal{E}(t)\}$ where agent $i \in \mathbb{V}$ has value denoted by $x_i(t) \in \mathbb{R}$. Note that we assume undirected edges and symmetric neighborhoods such that if $j \in \mathcal{N}_i(t)$ then $i \in \mathcal{N}_j(t)$. Furthermore, the *algebraic connectivity* of a graph $\mathbb{G}$ is the second smallest eigenvalue of the Laplacian matrix of the graph and is denoted $\lambda_2(\mathbb{G})$. It is well known that a positive algebraic connectivity, $\lambda_2(\mathbb{G}) > 0$, is possible if and only if the graph $\mathbb{G}$ is connected [32]. We consider the case where a subset of nodes with indices denoted by $\mathcal{S}$, $\mathcal{S} \subset \mathbb{V}$, are spoofed, such that $n_\mathcal{S} = |\mathcal{S}|$ is assumed to be unknown. Nodes in the set $\mathcal{L} \equiv \mathbb{V} \backslash \mathcal{S}$ are not spoofed, which we denote as *legitimate*. As a result, $N = n_\mathcal{L} + n_\mathcal{S}$, where $n_\mathcal{L} = |\mathcal{L}|$. We denote by $\mathbb{G}_\mathcal{S} = (\mathbb{V}_\mathcal{S}, \mathcal{E}_\mathcal{S})$ the subgraph induced by $\mathcal{S}$, where $\mathbb{V}_\mathcal{S} = \mathcal{S}$ and $\mathcal{E}_\mathcal{S} = \{(i, j) | i \in \mathcal{S} \text{ or } j \in \mathcal{S}\}$. Similarly, we denote by $G_\mathcal{L} = (\mathbb{V}_\mathcal{L}, \mathcal{E}_\mathcal{L})$ the subgraph induced by $\mathcal{L}$, where $\mathbb{V}_\mathcal{L} = \mathcal{L}$ and $\mathcal{E}_\mathcal{L} = \{(i, j) | i, j \in \mathcal{L}\}$. Thus $\mathbb{G} = \mathbb{G}_\mathcal{S} \cup \mathbb{G}_\mathcal{L}$.

A. Threat Model

Our threat model considers one or more adversarial agents with one omnidirectional Wi-Fi antenna each. Adversarial agents perform the "Sybil Attack" to inject packets emulating $n_\mathcal{S}$ non-existent clients according to the following definition:

Definition III.1 (Sybil Attack)**.** An adversary in the network can control the values of one or more "spoofed" nodes in the network by simultaneously sending various messages over the network with unique IDs $\{j_1, j_2, \dots\} \in \mathcal{S}$ in order to gain a disproportionate influence in the network. We assume that graph $\mathbb{G}$ is known, but the set of clients that are spoofed (i.e., in $\mathcal{S}$) is unknown. If an agent j is a spoofed node then its value at time t, denoted by $x_j(t)$, can be arbitrarily controlled by an adversarial agent in the network. This value is assumed to be finite for all time so that $|x_j(t)| \leq \eta$ for some $\eta > 0$ for all t and for all j. We also assume that spoofed nodes will remain spoofed at all times and legitimate nodes will remain legitimate at all times $t > 0$.

B. Detecting Sybil Attacks using Wireless Signals

Our previous work developed a method for measuring *directional signal profiles* using channel state information (CSI) from the wireless messages over each link (i, j) in the network [8], [19]. These profiles measure signal strength arriving from every direction in the 3D plane. This paper builds upon our previous work that derived a measurable scalar value α_{ij} that captures the likelihood that a transmission between two communicating agents i and j is legitimate or spoofed in the sense of a Sybil attack:

Definition III.2. (Confidence Weights α_{ij}) Our previous work [6] theoretically derived and rigorously tested (in hardware experiments) the existence of a scalar value $\alpha_{ij} \in (-1/2, 1/2)$ for the wireless channel between any two communicating agents i and j, capturing the likelihood that the transmission is spoofed in the Sybil sense. Importantly, it was shown that in expectation, these α scalars will be bounded less than zero for spoofed transmissions and above zero for legitimate transmissions within an epsilon bound such that $\mathbb{E}[\alpha_{ij}] \leq -1/2+\epsilon_{\mathcal{S}}$ if $j \in \mathcal{S}$ and $\mathbb{E}[\alpha_{ij}] \geq 1/2-\epsilon_{\mathcal{L}}$ if $j \in \mathcal{L}$. Here the bounds, $\epsilon_{\mathcal{S}}$, and $\epsilon_{\mathcal{L}}$ are error terms characterized in [6] that indicate the quality of the α-measurements and can be determined in closed form as a function of signal to noise ratio (SNR) of the channel, number of agents, and channel constants. Intuitively, the higher the signal quality, the smaller the $\epsilon_{\mathcal{S}}$ and $\epsilon_{\mathcal{L}}$, and the easier it is to detect spoofed nodes transmitting messages in the system. In this paper, we only utilize the fact that SNR of the wireless messages is high enough (i.e. the link is of good quality) such that $\epsilon_{\mathcal{S}}, \epsilon_{\mathcal{L}} \in (0, 1/2)$.

The objective of the current work is to develop a theoretical framework for using derived wireless channel information as captured by the α-measurements for robust consensus of multi-robot systems in the face of a Sybil Attack with switching network topologies. More information on the nature and derivation of the α values is included in the appendix.

C. Threat Resilient Consensus

We consider the distributed linear consensus protocol

$$x_i(t+1) = \mathbb{W}(i,i,t)x_i(t) + \sum_{j\in\mathcal{N}_i} \mathbb{W}(i,j,t)x_j(t). \quad (1)$$

Our recent work [7] derived a weight matrix $\mathbb{W}(t)$ using all observations on the wireless channels,$\{\alpha(0), \ldots, \alpha(t)\}$, such that the disturbance of the spoofed nodes on reaching average consensus could be bounded by a computable bound $\Delta_{\max}(\mathcal{P}, \delta)$ with user-specified probability $\delta \in (0, 1)$ and problem parameters $\mathcal{P} = (n_{\mathcal{S}}, n_{\mathcal{L}}, \eta, \epsilon_{\mathcal{S}})$ so that

$$\mathbb{P}\left(\lim_{t\to\infty} \left| x_{\mathcal{L}}(t) - \frac{vv^T x(0)}{n_{\mathcal{L}}} \right| \leq \Delta(\mathcal{P}, \delta)\right) \geq 1-\delta,$$

for some finite $\Delta(\mathcal{P}, \delta) \leq \Delta_{\max}(\mathcal{P}, \delta)$, where

$$v_i = \begin{cases} 1, & \text{if node } i \text{ is legitimate} \\ 0 & \text{otherwise} \end{cases}.$$

(See Theorem 1 in [7]). The current paper relaxes two critical assumptions from [7] with significant effects on the resiliency of the resulting consensus. In particular, the current paper does not assume a static network topology or knowledge of the number of spoofed nodes in the network. The objective of this paper is to determine which edges in the network to switch on or off over the evolution of the consensus in order to eliminate spoofed node influence in the network in an asymptotic sense. By making observations over the wireless channels (i.e. a sequence of $\alpha_{ij}(t)$ values), our problem is to learn the topology of the graph over which to perform consensus such that the influence of the spoofed node values on the consensus vanishes at an exponential rate.

Given an allowable edge set $\bar{\mathcal{E}} = \{(i,j)|i,j \in \bar{\mathbb{V}}\}$, where $\bar{\mathbb{V}} \subseteq \{1, \ldots, N\}$, let $s(\beta_{ij}(t)) : \mathbb{R} \to \{0, 1\}$ denote a switching signal for a particular edge in $\bar{\mathcal{E}}$. We define $\beta_{ij}(t) = \sum_{l=0}^{t} \alpha_{ij}(l)$. Note that the allowable edge set $\bar{\mathcal{E}}$ contains all the possible edges that can be connected in the graph. This need not be the complete graph but should include a connected graph over legitimate nodes.

Denoting by $s(t)$ the collection of all switching signals for each edge at time t so that $s(t) = \{s(\beta_{ij}(t))$ for all $(i,j) \in \bar{\mathcal{E}}\}$, we define $\mathcal{E}(s(t)) = \{(i,j) : i,j \in \bar{\mathbb{V}}$ and $s_{ij}(t) = 1\}$ to be the selected edge set and $\mathbb{G}(t) = (\mathbb{V}(s(t)), \mathcal{E}(s(t)))$ to be the graph defined by this edge set. The problem that this paper addresses is that of finding $s(t)$, and subsequently $\mathbb{G}(t)$, such that executing the consensus algorithm over the graph $\mathbb{G}(t)$ will converge for all legitimate nodes in the graph. Specifically,

Problem 1 (Switching Signal for Sybil-Resilient Consensus)**.** Given an allowable edge set $\bar{\mathcal{E}} = \{(i,j)|i,j \in \bar{\mathbb{V}}\}$, where $\bar{\mathbb{V}} \subseteq \{1, \ldots, N\}$, find a switching signal $s(\beta_{ij}(t)) : \mathbb{R} \to \{0, 1\}$ for each edge $(i,j) \in \bar{\mathcal{E}}$ such that the dynamics in Equation (1) executed over $\mathbb{G}(t)$ (chosen by the switching signal) will converge for all legitimate nodes to a scalar value $x^* \in \mathbb{R}$, i.e. $x_i \to x^*$, with high probability. Additionally, $x^* \neq x_{\mathcal{S}}$ with $x_{\mathcal{S}}$ being the spoofed node values.

We are also interested in the characterization of a time T_0 such that if observations over the wireless channels begin at time $t = 0$, and network topology is controlled from time $t = 0$ using the switching signal found in Problem 1, but consensus is allowed to start at a later time $t = T_0$, then with high probability we can recover Sybil-free consensus.

Problem 2 (Characterization of T_0)**.** We wish to characterize a time $T_0(\delta)$ such that for any $\delta \in (0, 1)$, the network topologies selected by the switching signal from Problem 1 result in $\mathbb{G}(t) \subset \{(\mathbb{V}, \mathcal{E}(t)) : \mathcal{E}(t) = \mathcal{E}_{\mathcal{L}}\}$ for all $t \geq T_0$ with probability $1-\delta$. In other words, a consensus algorithm starting at any time $t \geq T_0$ over the graphs selected by the switching function, $\mathbb{G} = \{\mathbb{V}, \mathcal{E}(s(t))\}$, can achieve Sybil-free consensus (where spoofed node influence is eliminated) with probability of at least $1-\delta$.

In this paper we will use a slight modification of the weights defined in [7] where we do not assume knowledge of the cardinality of the spoofed node set $n_{\mathcal{S}}$ or legitimate node set $n_{\mathcal{L}}$. We choose each element in our weighting matrix

$\mathbb{W}(t)$ in the following way, where $\beta_{ij}(t) = \sum_{l=0}^{t} \alpha_{ij}(l)$:

$$\mathbb{W}(i,j,t) = \begin{cases} \frac{1}{n}(1 - e^{-\beta_{ij}(t)/2}), & \text{if } \beta_{ij}(t) > 0, s_{ij}(t) = 1 \\ \frac{1}{2n} e^{\beta_{ij}(t)}, & \text{if } \beta_{ij}(t) \leq 0, s_{ij}(t) = 1 \\ 1 - \sum_l \mathbb{W}(i,l,t), & \text{if } i = j, s_{ij}(t) = 1 \\ 0, & \text{else} \end{cases} \tag{2}$$

Where n is the number of nodes included in the consensus graph, i.e. having an edge $(i,j) \in \mathcal{E}(s(t))$, as determined by the switching function from Problem 1 being positive $s(\beta_{ij}(t)) > 0$ and we assume symmetry such that $\mathbb{W}(i,j,t) = \mathbb{W}(j,i,t)$.

We summarize the assumptions used throughout the paper here for ease of reference:

Assumption 1. *1) **Nature of agents:** We assume that legitimate agents remain legitimate for all time t and spoofed agents remain spoofed for all time t.*

*2) **Undirected edges with symmetric weights:** We assume that edges are undirected and that observations over the wireless channel for a given edge are symmetric so that $w_{ij}(t) = w_{ji}(t)$ for all time t and all edges $(i,j) \in \mathcal{E}(t)$.*

*3) **Graph topology:** The allowable edge set $\bar{\mathcal{E}}$ over which the derived switching signal $s(\beta_{ij}(t))$ is defined in Problem 1 includes a connected graph over legitimate nodes.*

IV. Analysis

In this section we derive a switching function $s(\beta_{ij}(t))$, that determines which edges of the graph to include in the consensus and which to delete, and that keeps legitimate nodes and excludes the spoofed nodes with high probability. In particular we define our switching function as:

$$s(\beta_{ij}(t)) = \begin{cases} 1, & \text{if } \beta_{ij}(t) > 0 \\ 0, & \text{else} \end{cases} \tag{3}$$

where $\beta_{ij}(t) = \sum_{l=0}^{t} \alpha_{ij}(l)$ is the cumulative sum of repeated observations $\alpha_{ij}(t)$ of the wireless channel between two agents i and j. In the sequel, we will use a shorthand notation of $s_{ij}(t)$ for $s(\beta_{ij}(t))$. By applying this switching function to the graph we have that our edge set is composed of only edges (i,j) for which $s_{ij}(t) = 1$ so that (by Assumption 1.ii we also have the symmetric condition that $s_{ji}(t) = 1$) $\mathcal{E}(t) = \{(i,j) : (i,j) \in \bar{\mathcal{E}}$ and $s_{ij}(t) = 1\}$. Notice that this graph may not be connected over all legitimate nodes at some times t and that spoofed nodes may or may not be included in the edge set $\mathcal{E}(t)$ for some times t. We will show that this choice of switching function however, allows for asymptotic convergence of the legitimate nodes to x^* where $x^* \neq x_{\mathcal{S}}$ and $x_{\mathcal{S}}$ is the spoofed node value. This is our first result summarized below as Proposition 1 and is the solution to Problem 1 from Section III.

Proposition 1 (Convergence of Legitimate Nodes to Consensus)**.** *Using the switching signal $s_{ij}(t)$ from Equation (3) the consensus dynamics from the proposed protocol with weights as described in Equation (2) yield node values $x_i \to x^*$ as $t \to \infty$ for all $i \in \mathcal{L}$ where $x^* \neq x_{\mathcal{S}}$. In other words, the system composed of legitimate nodes converges to a single consensus value irrespective of the input of spoofed nodes in the system.*

Proof. Theorem 2 from the paper [33] provides for convergence of discrete consensus dynamics over switching graph topologies with mild requirements on connectivity where consensus is attainable on a graph $G = (\mathbb{V}, \mathcal{E})$ so long as $\mathbb{W}(i,j,t) > 0$, i.e. $(i,j) \in \mathcal{E}$ infinitely often. We will show that for our choice of switching signal from Equation (3) this condition is true for the edge set over legitimate nodes $\mathcal{E}_{\mathcal{L}}$ and *not* for $\mathcal{E}_{\mathcal{S}}$. Specifically, to satisfy the condition of Theorem 2 from [33] we must show that for all i,j,t:

i Our node weights $\mathbb{W}(i,j,t)$ are always positive so that $\exists a > 0$ s.t. if $\mathbb{W}(i,j,t) > 0$ then $\mathbb{W}(i,j,t) > a$.

ii The weight matrix has a positive diagonal such that $\mathbb{W}(i,i,t) > a$ for all i and t.

iii The choice of weights satisfy a cut-balance assumption defined in [33] as: for any nonempty proper subset $\mathbb{S}$ of $\{1,\ldots,N\}$ there exists $i \in \mathbb{S}$ and $j \notin \mathbb{S}$ with $\mathbb{W}(i,j,t) > 0$ if and only if there exists $i' \in \mathbb{S}$ and $j' \notin \mathbb{S}$ with $\mathbb{W}(j',i',t) > 0$.

iv For the graph $G(t) = (\mathbb{V}, \mathcal{E}(t))$ in which $(i,j) \in \mathcal{E}(t)$ if $s_{ij}(t) = 1$ we have that $s_{ij}(t) = 1$ occurs infinitely often for $(i,j) \in \mathcal{E}_{\mathcal{L}}$ and moreover, $s_{ij}(t) = 1$ *does not occur* infinitely often for $(i,j) \in \mathcal{E}_{\mathcal{S}}$.

We have that conditions (i)-(ii) are satisfied by the definition of $\mathbb{W}(i,j,t)$ in Equation (2) for all i,j,t. The cut balance condition (iii) holds true by Proposition 1 in [33] since our definition of $\mathbb{W}$ is average-preserving at each time t and we assume symmetric weights and undirected edges (see Assumption 1). Lastly, condition (iv) holds by a straightforward application of Corollary 1 from our previous paper [7] in the following way. For each $t \in \mathbb{N}$ let $A_{ij}(t)$ denote the undesirable event that $\beta_{ij}(t) \geq 0$ if $j \in \mathcal{S}$ and alternatively the (undesirable) event $\beta_{ij} < 0$ if $j \in \mathcal{L}$. By Corollary 1 from [7] we have that $\sum_{t=0}^{\infty} \mathbb{P}(A_{ij}(t)) \leq \sum_{t=0}^{\infty} \exp(-tc) < \infty$ since it is a geometric series with $c = (1/2 - \epsilon_{\mathcal{S}})^2$ where $\epsilon_{\mathcal{S}} < 1/2$ by the definition of $\alpha_{ij}(t)$ (Definition III.2). Therefore we have that the probability that the event $A_{ij}(t)$ (i.e. $s_{ij}(t) = 1$ for $(i,j) \in \mathcal{E}_{\mathcal{S}}$ or $s_{ij}(t) = 0$ for $(i,j) \in \mathcal{E}_{\mathcal{L}}$) occurs for infinitely many values of t is 0. Thus the event that $s_{ij}(t) = 1$ for $(i,j) \in \mathcal{E}_{\mathcal{L}}$ or $s_{ij}(t) = 0$ for $(i,j) \in \mathcal{E}_{\mathcal{S}}$ occurs for infinitely many values of t with probability 1. In other words, 1) the number of times that legitimate agents share an edge $(i,j) \in \mathcal{E}(t)$ for all $(i,j) \in \mathcal{E}_{\mathcal{L}}$ is unbounded, and 2) the number of times that legitimate nodes share an edge with spoofed nodes is bounded, and condition (iv) is satisfied. This establishes the claim. ■

Remark (Nature of x^*)**.** *Proposition 1 states that legitimate nodes will converge to the same value. Because this paper uses a similar definition for $w_{ij}(t)$ as [7] where convergence to a value within bounded distance to the true average is*

guaranteed, we expect that x^* will also be bounded close to the true average though we do not prove this here.

Proposition 1 implies that in the limit as $t \to \infty$ we expect that the topology chosen by the switching function $s_{ij}(t)$ converges to the true topology over the legitimate nodes. More specifically, the probability of our switching signal choosing a spoofed edge as part of the graph topology $G(t)$ decays exponentially with t. Therefore we can expect that there exists a time T_0 such that for $t > T_0$ the topology $\bar{\mathcal{E}}(s(t))$ will be equivalent to the subgraph over legitimate nodes $\mathcal{E}_{\mathcal{L}}$ and thus that a consensus algorithm run over the graph defined by the edges $\bar{\mathcal{E}}(s(t))$ will be free of influence from the spoofed nodes with high probability. We formalize this in the following proposition which is the solution to Problem 2 from Section III:

Proposition 2 (Characterization of consensus start time T_0)**.** *For uncorrelated α_{ij} values, $G(t)$ selected by the switching signal $s(t)$ from Equation (3), and some user defined allowable probability of error $\delta \in (0, 1)$, there is some finite T_0 such that consensus over $G(t)$ started at time $t \geq T_0$ converges to the true average with high probability. Specifically, running Algorithm 1 over the graph $G(t)$ that is started at time $t \geq T_0$ converges to the true average with probability at least $1 - \delta$. Furthermore, this consensus start time is found to be $T_0 = -\frac{1}{c}\ln(\delta c)$ where $c = (\epsilon_{\mathcal{S}} - 1/2)^2$.*

Proof. By Lemma 1 from our previous work [7] we have that for any link (i, j) with $j \in \mathcal{S}$, $t \in \mathbb{N}$, and a sequence of weights $(\alpha_{ij}(0), \ldots, \alpha_{ij}(t))$, the probability that $\beta_{ij}(t)$ is non-negative decays exponentially fast with t such that $\mathbb{P}(\beta_{ij}(t) \geq 0) \leq \exp(-t(\epsilon_{\mathcal{S}} - 1/2)^2) = \exp(-ct)$ where $c = (\epsilon_{\mathcal{S}} - 1/2)^2$ and reflects the quality of the wireless channels. Lemma 1 from [7] also states that for a legitimate node where $j \in \mathcal{L}$ we have that $\mathbb{P}(\beta_{ij}(t) < 0) \leq \exp(-tc)$. Thus by defining A_t as the undesirable event that $\beta_{ij}(t) \geq 0$ for any link (i, j) with $i \in \mathcal{S}$ or $j \in \mathcal{S}$, or $\beta_{ij}(t) < 0$ for any link $(i, j), i, j \in \mathcal{L}$, we have that the probability of A_t occurring for any time in $t \geq T_0$ is $\mathbb{P}(\bigcup_{t=T_0}^{\infty} A_t)$. By the union bound we have that

$$\mathbb{P}\Big(\bigcup_{t=T_0}^{\infty} A_t\Big) \leq \sum_{t=T_0}^{\infty} \mathbb{P}(A_t) \leq \sum_{t=T_0}^{\infty} \exp(-tc) \tag{4}$$

Setting the above probability to be at most δ and upper bounding the infinite sum (a geometric series) gives $\frac{1}{c}\exp(-cT_0) < \delta$. Solving for T_0 gives $T_0 > -\frac{1}{c}\ln(\delta c)$ which establishes the claim. ■

V. SIMULATION

Through a comprehensive simulation study, we compare the consensus protocol described in this paper with the resilient consensus protocol from [7], as well as a general consensus protocol without resilience from [23]. We carry out each simulation as follows. Legitimate agents are initialized with values $x(t)$ sampled uniformly in the interval $(0, 10)$, and at each time step each node i updates its state $x_i(t + 1)$ according to the local information available to it.

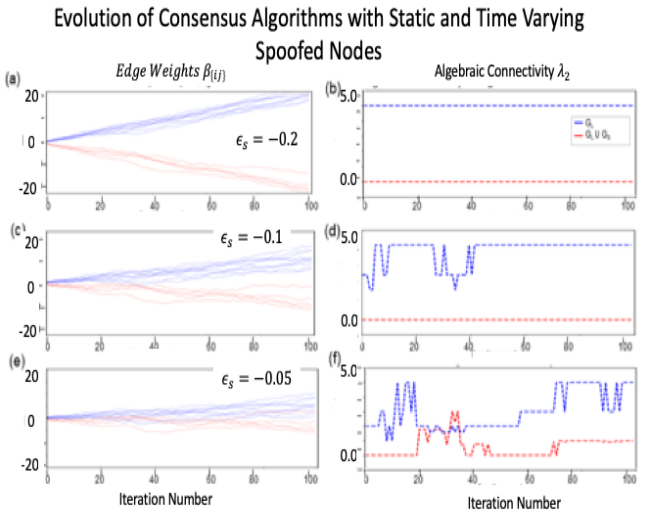

Fig. 2. Spoofed and legitimate $\beta_{ij}(t)$ values of legitimate links (blue) and spoofed links (red) and the algebraic connectivity λ_2 of the legitimate sub-graph (blue) and malicious subgraph (red).

The spoofed nodes adhere to one of two strategies, **Strategy 1:** broadcast a constant value or **Strategy 2:** use a time varying sinusoidal function to generate broadcasted values. Spoofed nodes do not cooperate in the consensus update Algorithm 1.

We simulate observations $\alpha_{ij}(t)$ by sampling from a Gaussian distribution with $\mathbb{E}[\alpha_{ij}(t)] = \epsilon_L \ \forall j \in L$ and $\mathbb{E}[\alpha_{ij}(t)] = \epsilon_S \ \forall j \in S$. The mean ϵ_S is a simulation parameter but the standard deviation of the distributions are held at $\sigma = 0.33$. Intuitively, for ϵ_S close to zero, it is harder to distinguish between legitimate and spoofed nodes and will take longer to converge to the legitimate sub-graph G_L. Figure 2 plots the β values and the algebraic connectivity of the spoofed graph G_S and the legitimate graph G_L over time for simulations with $\epsilon_S = -0.3$, $\epsilon_S = -0.1$, $\epsilon_S = -0.05$. Figure 3 shows the average and standard deviation of disagreement over time for 1000 simulations, obtained from sweeping through different values of $\epsilon_S \in [-0.4, -0.1]$ for each of the algorithms mentioned on Network 1 and Network 2. Figure 4 shows the values over time for different runs for each of the 3 different algorithms, with spoofing Strategy 1 & 2.

The simulations are consistent with Proposition 1, as we see that as time $t \to \infty$ the node values approach consensus $x_i \to x^* \ \forall i \in \mathcal{L}$. Additionally we see that the results in Figure 2, consistent with Proposition 2, as time increases, the probability of a spoofed node being classified as legitimate approaches zero, and the switching signal in Equation (3) converges to the legitimate sub-graph $G_L(t)$. Additionally, although we do not prove this, the empirical evidence from simulation bolsters our remark that the bounds on this algorithm may be tighter than that of [7].

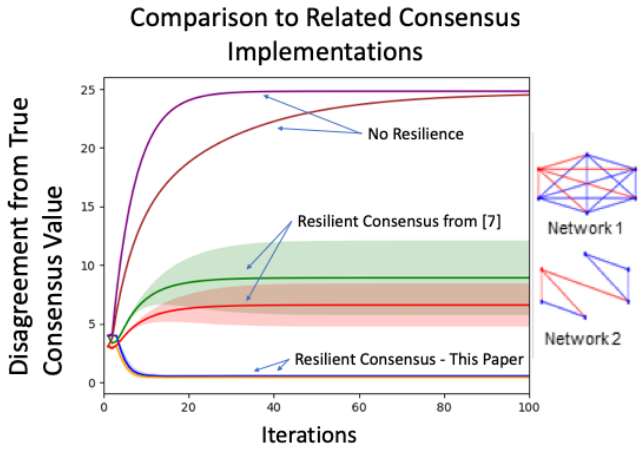

Fig. 3. Plot of the average discrepancy over time for each algorithm on both topologies. Discrepancy here is defined as the root mean squared error $\sqrt{\frac{1}{N_L}\sum_{i\in\mathcal{L}}[x_i(t)-\frac{1\mathbb{1}^T}{N_L}x_L(0)]^2}$.

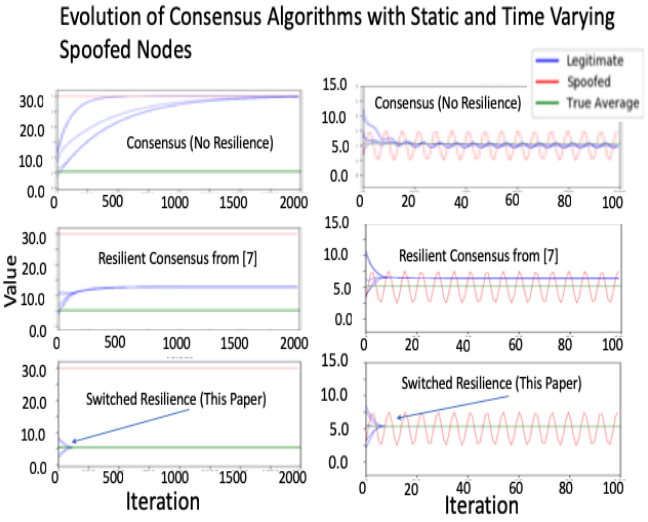

Fig. 4. Several example runs demonstrating how each protocol responds to spoofing Strategy 1 (left) and Strategy 2 (right).

Algorithm 1: Wi-Fi Resilient Consensus

input : $G(t-1)$,$\beta(t-1)$,$S(t-1)$,$x(t-1)$, $\alpha_{ij}(t-1)$,t
output: $x(t)$,$\beta(t)$

for *all* $(i,j) \in |V| \times |V|$ *where* $i \neq j$ **do**
$\beta_{ij}(t)$=$\beta_{ij}(t-1)$ + $\alpha_{ij}(t-1)$
if $t \geq T_0$ **then**
$\mathbb{W}_{ij}(t) =$

$$\begin{cases} \frac{1}{n}(1-e^{-\beta_{ij}(t)/2}), & \text{if } \beta_{ij}(t)>0 \text{ and } s_{ij}(t)=1 \\ \frac{1}{2n}e^{\beta_{ij}(t)}, & \text{if } \beta_{ij}(t)\leq 0 \text{ and } s_{ij}(t)=1 \\ 1-\sum_l \mathbb{W}_{il}(t), & \text{if } i=j \\ 0 & \text{otherwise} \end{cases}$$

$W_{ji}(t) = W_{ij}(t)$
$x_i(t) = W_{ii}(t)x_i(t-1) + \sum_{j\in\mathcal{N}_i} W_{ij}(t)x_j(t-1)$
end
return $x(t)$,$\beta(t)$

VI. Conclusion

This paper provides a novel algorithm for resilient consensus in Wi-Fi networks by controlling the network topology. Results are based on wireless channel observations (signal profiles), created by complex multi-path fading. Analytical results of this paper provide conditions for achieving consensus over legitimate node values within a user-defined probability. Simulation results validate this analysis wherein the network converges to the average of the legitimate node values while asymptotically disconnecting the spoofed nodes.

Appendix

Our previous work developed a method for measuring *directional signal profiles* using channel state information (CSI) from the wireless messages over each link (i,j) in the network [8], [19]. These profiles measure signal strength arriving from every direction in the 3D plane. Directional signal profiles display two important properties: 1) transmissions originating from the same physical agent have very similar profiles and 2) energy can be measured coming from the direct-line path between physical agents. For simultaneous transmissions from a spoofed node these transmission profiles are similar from the same physical node. This is captured quantitatively by α_{ij}. The paper [6] quantifies these properties, providing an analysis that shows both analytically and experimentally, that a single scalar value $\alpha_{ij} \in (-0.5, 0.5)$ (shifted by -0.5 from [6]) can be computed for each signal profile that quantifies the likelihood that the transmission is coming from the same physical (spoofed) node, or a unique (legitimate) node; a property critical for thwarting Sybil Attacks. Intuitively, the α_{ij} was shown experimentally and theoretically to be close to -0.5 if one of the agents j is a spoofed node and close to 0.5 if both agents are legitimate nodes in the network [6]. This is captured quantitatively by the bounds on the expectation of the α_{ij}. Note that the quality of the bounds, as captured by the epsilons, will depend on the number of spoofed nodes (becoming more lose as the number of spoofed agents increases). For the purposes of the consensus algorithm presented in this paper we do not need to know the number n_L of spoofed agents, however this number is used in [6] and can be estimated from the number of similar transmission profiles in the system as characterized by the similarity metric in [6].

References

[1] I. Sargeant and A. Tomlinson, "Modelling malicious entities in a robotic swarm," in *Digital avionics systems conference (DASC), 2013 IEEE/AIAA 32nd*. IEEE, 2013, pp. 7B1–1.
[2] S. Ghildiyal, H. Goel, and S. K. Jangra, "Sybil attack in wireless sensor networks," 2018.
[3] Y. Sheng, K. Tan, G. Chen, D. Kotz, and A. Campbell, "Detecting 802.11 mac layer spoofing using received signal strength," in *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications*, April 2008, pp. 1768–1776.
[4] L. Lamport *et al.*, "Paxos made simple," *ACM Sigact News*, vol. 32, no. 4, pp. 18–25, 2001.
[5] L. Lamport, "Fast paxos," *Distrib Comput*, vol. 19, no. 2, pp. 79–103, 2006.

[6] S. Gil, S. Kumar, M. Mazumder, D. Katabi, and D. Rus, "Guaranteeing spoof-resilient multi-robot networks," *Autonomous Robots*, vol. 41, no. 6, pp. 1383–1400, 2017.

[7] S. Gil, C. Baykal, and D. Rus, "Resilient multi-agent consensus using wi-fi signals," 2018, preprint.

[8] S. Gil, S. Kumar, D. Katabi, and D. Rus, "Adaptive communication in multi-robot systems using directionality of signal strength," *The International Journal of Robotics Research*, vol. 34, no. 7, pp. 946–968, 2015.

[9] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005.

[10] D. Saldana, A. Prorok, S. Sundaram, M. F. M. Campos, and V. Kumar, "Resilient consensus for time-varying networks of dynamic agents," in *2017 American Control Conference (ACC)*, May 2017, pp. 252–258.

[11] H. J. LeBlanc and X. D. Koutsoukos, "Consensus in networked multi-agent systems with adversaries," in *Proceedings of the 14th International Conference on Hybrid Systems: Computation and Control*, ser. HSCC '11. New York, NY, USA: ACM, 2011, pp. 281–290.

[12] Y. Wang, G. Attebury, and B. Ramamurthy, "A survey of security issues in wireless sensor networks," *IEEE Communications Surveys Tutorials*, vol. 8, no. 2, pp. 2–23, Second 2006.

[13] R. M. Kieckhafer and M. H. Azadmanesh, "Reaching approximate agreement with mixed-mode faults," *IEEE Transactions on Parallel and Distributed Systems*, vol. 5, no. 1, pp. 53–63, Jan 1994.

[14] 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, April 2013.

[15] S. M. Dibaji and H. Ishii, "Resilient consensus of second-order agent networks: Asynchronous update rules with delays," *Automatica*, vol. 81, pp. 123–132, 2017.

[16] J. Xiong and K. Jamieson, "Securearray: improving wifi security with fine-grained physical-layer information," in *MobiCom*, 2013.

[17] J. Yang, Y. Chen, W. Trappe, and J. Cheng, "Detection and localization of multiple spoofing attackers in wireless networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 1, pp. 44–58, Jan 2013.

[18] J. P. Fitch, *Synthetic Aperture Radar*. Berlin, Heidelberg: Springer-Verlag, 1988.

[19] S. Kumar, S. Gil, D. Katabi, and D. Rus, "Accurate indoor localization with zero start-up cost," in *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '14. New York, NY, USA: ACM, 2014, pp. 483–494.

[20] S. Gil, S. Kumar, D. Katabi, and D. Rus, "Adaptive communication in multi-robot systems using directionality of signal strength," *The International Journal of Robotics Research*, vol. 34, no. 7, pp. 946–968, 2015.

[21] S. Gil, C. Baykal, and D. Rus, "Resilient multi-agent consensus using wi-fi signals," *IEEE Control Systems Letters*, vol. 3, no. 1, pp. 126–131, 2019.

[22] N. R. Chowdhury, S. Sukumar, and D. Chatterjee, "A new condition for asymptotic consensus over switching graphs," *Automatica*, vol. 97, pp. 18–26, 2018.

[23] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, Sept 2004.

[24] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on automatic control*, vol. 48, no. 6, pp. 988–1001, 2003.

[25] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," in *Proceedings of the 44th IEEE Conference on Decision and Control*, Dec 2005, pp. 2996–3000.

[26] S. Martin and J. M. Hendrickx, "Continuous-time consensus under non-instantaneous reciprocity," *IEEE Transactions on Automatic Control*, vol. 61, no. 9, pp. 2484–2495, 2016.

[27] B. Touri and A. Nedic, "Distributed consensus over network with noisy links," in *2009 12th International Conference on Information Fusion*, 2009, pp. 146–154.

[28] J. Liu, A. S. Morse, A. Nedić, and T. Başar, "Internal stability of linear consensus processes," in *53rd IEEE Conference on Decision and Control*, Dec 2014, pp. 922–927.

[29] A. Nedich, "Convergence rate of distributed averaging dynamics and optimization in networks," *Foundations and Trends in Systems and Control*, vol. 2, no. 1, pp. 1–100, 2015.

[30] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation*. Prentice Hall, 1989.

[31] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[32] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak Mathematical Journal*, vol. 23, no. 2, pp. 298–305, 1973. [Online]. Available: http://eudml.org/doc/12723

[33] J. M. Hendrickx and J. N. Tsitsiklis, "Convergence of type symmetric and cut-balanced consensus seeking systems," *CoRR*, vol. abs/1102.2361, 2011. [Online]. Available: http://arxiv.org/abs/1102.2361