Physical Layer Authentication Under Intelligent Spoofing in Wireless Sensor Networks[☆]

Ning Gao[a], Qiang Ni[b], Daquan Feng[c,*], Xiaojun Jing[d], Yue Cao[b]

[a] *National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China*
[b] *School of Computing and Communications, Lancaster University, Lancaster LA1 4WA, U.K.*
[c] *College of Information Engineering, Shenzhen University, Shenzhen 518060, China*
[d] *School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China*

Abstract

Location based access in wireless sensor networks (WSN) are vulnerable to location spoofing attacks. In this paper, we investigate the physical layer (PHY-layer) authentication in the threat of an intelligent location spoofing attack. The intelligent attack can emulate the legitimate channel information and maximize its long-term cumulative reward. First, we analyze the feasibility of this intelligent attack and investigate how it threats to the networks. Specifically, we derive the optimal transmit power allocation and find the worst case for the defenders, namely optimal intelligent attack, in which the attacker can learn the intelligent attack action based on the beamforming with optimal transmit power allocation. To defend against such an intelligent attack with high accuracy and low overhead, we develop a cooperative PHY-layer authentication scheme. Then, we provide an in-depth analysis on the belief and derive the belief bounds and the closed-form expression for the belief threshold. Furthermore, considering the whole computation complexity and the double counting problem in a loopy graph, we propose the cooperative neighbour selection algorithm to

[☆] Fully documented templates are available in the elsarticle package on CTAN.

[*] Corresponding author

Email addresses: gaoning788@gmail.com (Ning Gao), q.ni@lancaster.ac.uk (Qiang Ni), fdquan@szu.edu.cn (Daquan Feng), jxiaojun@bupt.edu.cn (Xiaojun Jing), e-mail: yue.cao@lancaster.ac.uk (Yue Cao)

accelerate belief convergence and reduce the overhead. Finally, the simulation results reveal that the proposed method can significantly improve the defense performance compared with the state-of-art methods.

Keywords: Physical layer authentication, intelligent location spoofing, WSN.

1. Introduction

Today, wireless sensor networks (WSN) have played an important role in Internet of mission critical things (IoMCT), i.e., battlefield, border patrol, search and rescue, etc. The location verification in WSN is key to location based security IoMCT services [1, 2]. For example, as the location based access obviates the need to establish shared secrets in advance, it can apace authenticate a transmitter via the received signal strength (RSS). However, the open air nature of wireless systems makes it vulnerable to physical layer (PHY-layer) security threats [3]. One serious threat is called location spoofing attack, which makes the attack impersonate the legitimate location to access networks. Such an attack can further cause denial-of-service (DoS), session hijacking, man-in-the-middle (MITM) attacks, which makes PHY-layer authentication extremely challenging.

Many location spoofing detection or robust localization algorithms have been developed to address the location spoofing threats. The key idea is to distinguish radio transmitters by exploiting uncorrelated PHY-layer spatial information between the legitimate users and the adversary, such as RSS [4, 5, 6, 7] and channel state information (CSI) [8, 9, 10, 11]. In [6], the optimal strategies to attack an RSS based wireless location verification system (LVS), have been analyzed for the spatially correlated shadowing channel. Similarly, the optimal attack strategy and the optimal LVS performance have been investigated in Rician fading channel [7]. To withstand the location spoofing attack, a robust localization algorithm has been developed in [12]. Compared with RSS, the CSI contains more location characteristic information [8], thus can improve the localization and spoofing detection performance. In [8], a user authentication approach

has been developed by exploiting power spectral densities, where the optimal test threshold for a specified false alarm probability is derived. In [11], a CSI based authentication scheme with optimal attack strategy has been proposed over multiple input multiple output correlated fading channel. In addition, machine learning techniques have emerged to integrate with RSS or CSI scheme to further optimize spoofing detection performance [13, 14, 15, 16]. In complex dynamic communication models, i.e., the hydraulic systems inspired communication models [17], the optimal solutions can be obtained by using metaheuristic algorithms [18, 19].

However, the existing work mainly focuses on optimizing the attack strategies and the detection performances with respect to a "blunt" location spoofing attack. The term "blunt" refers to attack action, i.e., whether launch attack, without changing with the communication environments. Nowadays, the machine learning is emerging not only to enhance WSN security [20, 21], but also to threaten WSN security. With the rapid development of artificial intelligence, the attackers can be smarter and more harmful than we have ever considered. For example, different from obtaining the conventional instantaneous reward, the attacker can use machine learning, i.e., Q-learning, to choose attack action based on the communication environments and to maximize the reward based on a series of time events. This reward is called long-term cumulative reward [22]. By using Q-learning, the maximum long-term cumulative reward can be obtained by an attacker over a period of time.

The intelligent location spoofing attack investigated in this paper is an attack that can emulate the legitimate channel information via beamforming and maximize its long-term cumulative reward. Specifically, the intelligent attacker can find the worst case for the defenders, namely optimal intelligent attack. That is, the intelligent attacker can falsify the legitimate CSI and RSS via beamforming with optimal transmit power allocation. Then, based on this optimal power allocation, the intelligent attacker further learns the intelligent attack action to maximize its long-term cumulative reward. Thereby, *the channel information is forged* and *attack action is shifty*, this intelligent location spoofing attack will

have a significant impact to the normal operation of WSN. In related work [23], one perfect location spoofing attack has been investigated, which can perfectly mimic the location of legitimate user. However, compared with the aforementioned intelligent location spoofing attack, this attack is not smart enough, i.e., the attack action cannot shift with the communication environments. Besides, the work of [23] focuses more on how to design one attack but inadequately tackles on how to defend against it. It is important to study the attack defense strategies. Inherently, once the performance and characteristics of a new attack are found, the emphasis is to propose the defense strategies with respect to this new attack. Thereby, motivated by the importance to study the attack defense strategy, we develop a PHY-layer authentication scheme under the threats of the investigated intelligent location spoofing attack and provide some detailed analysis.

In developing the aforementioned PHY-layer authentication scheme, some key factors should be concerned. First, since WSN is resource-limited, the PHY-layer authentication scheme should be with low overhead to prolong the life of the network. Then, the WSN are generally multi-hop networks with various topologies, which motivates us to consider a decentralized scheme to reduce maintenance cost [24]. Moreover, the PHY-layer authentication problem can be transformed into the signal detection problem, and cooperative detection can effectively improve the signal detection performance [25, 26, 27]. Whereas, there is lack of adequate attention to bring cooperation in PHY-layer authentication [4, 5, 6, 7, 8, 9, 10, 11].

Inspired by the above mentioned work [23, 24, 25, 26, 27], we propose a cooperative distributed PHY-layer authentication scheme to address intelligent location spoofing attack. To the best of our knowledge, the answers to the following questions are still missing:

- Is it possible to have an intelligent location spoofing attack to threaten WSN?
- How to address such intelligent location spoofing threats in PHY-layer

authentication?

The key contributions of this paper are summarized as follows:

- We study a new intelligent location spoofing attack, which can maximize the long-term cumulative reward. The feasibility of intelligent attack is analyzed and the optimal intelligent attack is exposed. Specifically,
  1. The beamforming is derived based on maximum likelihood estimator (MLE);
  2. The maximum long-term cumulative reward is obtained via Q-learning;
  3. The optimal transmit power allocation is derived by optimizing the Kullback-Leibler (KL) divergence.

- To address the intelligent attack, we propose a cooperative PHY-layer authentication scheme via belief forecasting propagation. The developed scheme only needs to communicate a short belief message with each other rather than a long message, which leads to little transmission overhead. Specifically:
  1. We design the local function and the compatibility function for Markov random field (MRF);
  2. We derive the belief bounds and obtain the closed-form expression for belief threshold;
  3. We propose the cooperative neighbour selection algorithm to accelerate the belief convergence and reduce the overhead.

The rest of the paper is organized as follows. In Section 2, we present the system model. In Section 3, the details of the investigated intelligent location spoofing attack are discussed. In Section 4, we propose the cooperative PHY-layer authentication scheme with respect to the intelligent attack discussed in Section 3. Simulations are presented in Section 5 and future work are discussed in Section 6. We summarize this paper in Section 7.

2. system model

In this section, we first introduce the channel model and then present the attack model. For ease of reference, important notations are summarized in Table 1.

Table 1: Summary of Notations

Symbols	Notations
$\mathbf{V}$	Matrix
$\mathbf{v}$	Vector
$\mathbf{V}^\top$	Transpose of matrix
$\mathbf{V}^\dagger$	Hermitian transpose of matrix
$\hat{\mathbf{v}}$	Estimation vector
$\mathbb{C}^{m\times n}$	Complex space
$\mathbb{R}^{m\times n}$	Real space
$\lvert\cdot\rvert$	Absolute value
$\mathbb{E}[\cdot]$	Expectation operator
$\lceil\cdot\rceil$	Upper bound
$\lfloor\cdot\rfloor$	Lower bound
$\triangleq$	Defined as
i.i.d.	Independent and identically distributed
N	Number of samples
M	Number of sensors
K	Number of channel states
γ	Threshold of likelihood ratio test
$\mathrm{Ei}(\cdot)$	Exponential integral function
$D_{KL}[\cdot]$	KL divergence
$\mathbf{S}$	State set
$\mathbf{a}$	Action set
$U(\mathbf{s}, a)$	Reward obtained when $a \in A$ is taken in state $\mathbf{s}$
ε	Learning rate of Q-learning
ϖ	Discount factor of Q-learning
ϵ	Probability that the attacker chooses the non-optimal action
O_i	Observation state of sensor i
S_i	Hidden state of sensor i
κ	Forgetting factor
b_T	Belief threshold

2.1. Channel Model

We consider a static WSN that consists of a sink node, multiple sensor nodes and an intelligent attacker. We assume that the sink node and the sensor nodes are with fixed locations and are resource-limited devices with a single antenna. The attacker is with fixed location once it is deployed and is equipped with multi-antennas. Let b be the sink node, $i \in \{1, \cdots, M\}$ be the i-th sensor node, s be the attacker. The channel coefficient from source a to destination d is denoted as $h_{ad} = \sqrt{d_{ad}^{-\eta}}\widetilde{h}_{ad}$ with $a, d \in \{b, i\}, a \neq d$, where d_{ad} represents the distance between source a and destination d, η is path loss exponent and $\widetilde{h}_{ad}$ represents small-scale block fading, which follows zero-mean complex Gaussian distribution with unit-variance. The channel h_{ad} eavesdropped by s via antenna c is denoted as $h_{ad\to s}(c)$. All of the channels are assumed to be reciprocal and spatially correlated.

The received signal at the i-th sensor is denoted as[1]

$$y_i = \sqrt{p_b}h_{bi}x_b + \Phi\sqrt{p_{si}}\mathbf{h}_{ib\to s}^{\dagger}\mathbf{w}_i x_{si} + n, \tag{1}$$

where p_b and p_{si} are the power budgets for the sink node and from the attacker to sensor i, respectively. x_b and x_{si} are the unit-energy genuine signal for the sink node and illegitimate signal from the attacker to sensor i, respectively. $\mathbf{w}_i \in \mathbb{C}^{\rho\times 1}$ is the beamforming using ρ antennas and $\mathbf{h}_{ib\to s} = [h_{ib\to s}(i\rho + 1 - \rho), \cdots, h_{ib\to s}(i\rho)] \in \mathbb{C}^{\rho\times 1}$ is the wiretap channel vector to sensor i, n is denoted as channel noise and is assumed to be independent of signal and follows complex Gaussian distribution with zero-mean and variance σ^2. $\Phi = 1$ and $\Phi = 0$ represent the present and absent of the attack, respectively.

[1]Note that this argument assumes perfect synchronization of sink node and attack's transmissions when $\Phi = 1$. Any imperfect synchronization is important in the detection process, but they are out of scope in this work.

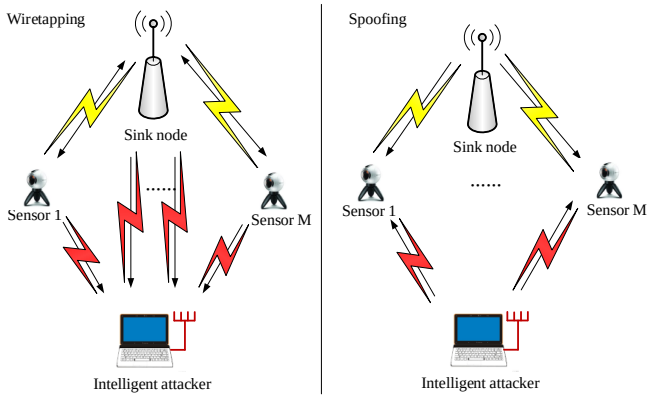

Figure 1: The intelligent location spoofing attack model.

2.2. Attack Model

The intelligent location spoofing attack is depicted in Fig. 1, which has two stages:

Wiretapping: When the sink node and sensors are communicating, the intelligent attacker wiretaps uplink and downlink channel which can be denoted as $\mathbf{H}_{i\to s} = [\mathbf{h}_{1b\to s}; \cdots; \mathbf{h}_{Mb\to s}]$ and $\mathbf{H}_{b\to s} = [\mathbf{h}_{b1\to s}; \cdots; \mathbf{h}_{bM\to s}]$, respectively[2].

Spoofing: The intelligent attacker emulates the legitimate channel information and learns the intelligent attack action to maximum long-term cumulative reward.

In WSN, the sink node and the sensors communicate frequently, the attacker do not wait to attack until the networks are idle. In other words, this intelligent attacker is an active attacker, which can launch attack when the sink node and sensors are communicating. To attack successfully, the attacker should guarantee the received deceiving signal power higher than the received legitimate signal power per sensor [2].

[2]We assume that the wiretap channel estimation are perfect and the channel statistics, i.e., channel correlation matrix, are known.

3. Intelligent Attack

In this section, we analyze the feasibility of intelligent attack which can emulate the legitimate channel information via beamforming and maximize long-term cumulative reward via Q-learning. Then, we investigate an optimal intelligent attack.

3.1. Feasibility of Intelligent Attack

3.1.1. Beamforming

From (1), sensor i can decode the deceiving signal only when the i-th mimic channel, $g_{bi} = \mathbf{h}_{ib\to s}^{\dagger}\mathbf{w}_i$, is very close to h_{bi}, that is $|g_{bi} - h_{bi}| \leq \beta$ for some $\beta > 0$. The beamforming to sensor i is to maximize the probability P that $|g_{bi} - h_{bi}| \leq \beta$, which can be denoted by

$$g_{bi}^{*} = \arg\max_{g_{bi}} P[|g_{bi} - h_{bi}| \leq \beta], \tag{2}$$

where g_{bi}^{*} is the optimal emulated channel.

By choosing the same column of uplink and downlink channel matrixes $\mathbf{H}_{i\to s}$ and $\mathbf{H}_{b\to s}$, respectively, i.e., the ρ-th column, we can get the vectors $\mathbf{h}_{s\uparrow} = [h_{1b\to s}(\rho), \cdots, h_{Mb\to s}(M\rho)]^{\dagger}$ and $\mathbf{h}_{s\downarrow} = [h_{b1\to s}(\rho), \cdots, h_{bM\to s}(M\rho)]^{\dagger}$. Denote vector $\mathbf{h} = [\mathbf{g}_b^{\top}, \mathbf{h}_{s\downarrow}^{\top}, \mathbf{h}_{s\uparrow}^{\top}]^{\top}$, where $\mathbf{h}$ follows zero-mean complex Gaussian distribution with correlation matrix

$$\mathbf{R} \quad = \quad \mathbb{E}[\mathbf{h}\mathbf{h}^{\dagger}]. \tag{3}$$

Then, for a given estimation vector $\hat{\mathbf{h}}(\hat{\mathbf{g}}_b) = [\hat{\mathbf{g}}_b^{\top}, \mathbf{h}_{s\downarrow}^{\top}, \mathbf{h}_{s\uparrow}^{\top}]^{\top}$, the optimal channel $\mathbf{g}_b^{*} = [g_{b1}^{*}, \cdots, g_{bM}^{*}]^{\dagger}$ can be achieved with MLE, which is given by [28, 11]

$$\mathbf{g}_b^{*} = \arg\min_{\hat{\mathbf{g}}_b} \hat{\mathbf{h}}^{\dagger}(\hat{\mathbf{g}}_b)\mathbf{R}^{-1}\hat{\mathbf{h}}(\hat{\mathbf{g}}_b), \tag{4}$$

where $\mathbf{R}^{-1}$ is the inverse of the block matrix $\mathbf{R}$. Denote matrix $\mathbf{V} = \mathbf{R}^{-1}$, and

we can write matrix $\mathbf{V}$ as

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} & \mathbf{V}_{13} \\ \mathbf{V}_{12}^{\dagger} & \mathbf{V}_{22} & \mathbf{V}_{23} \\ \mathbf{V}_{13}^{\dagger} & \mathbf{V}_{23}^{\dagger} & \mathbf{V}_{33} \end{bmatrix}. \tag{5}$$

By calculating the gradient of $\hat{\mathbf{h}}^{\dagger}(\hat{\mathbf{g}}_b)\mathbf{V}\hat{\mathbf{h}}(\hat{\mathbf{g}}_b)$ with respect to $\hat{\mathbf{g}}_b$ and setting it to zero, i.e., $\frac{\partial \hat{\mathbf{h}}^{\dagger}(\hat{\mathbf{g}}_b)\mathbf{V}\hat{\mathbf{h}}(\hat{\mathbf{g}}_b)}{\partial \hat{\mathbf{g}}_b} = 0$, we have

$$\mathbf{g}_b^* = -\mathbf{V}_{11}^{-1}\left(\mathbf{V}_{12}\mathbf{h}_{s\downarrow} + \mathbf{V}_{13}\mathbf{h}_{s\uparrow}\right). \tag{6}$$

The optimal beamformer vector of sensor i is obtained by

$$\mathbf{w}_i^* = (\mathbf{h}_{ib\to s}^{\dagger})^{-1} g_{bi}^*, \quad i = 1, 2, \cdots, M \tag{7}$$

where $(\mathbf{h}_{ib\to s}^{\dagger})^{-1}$ is the Moore-Penrose inverse of $\mathbf{h}_{ib\to s}^{\dagger}$.

3.1.2. Q-learning

The objective of attack is to learn the intelligent attack action to maximize the long-term cumulative reward. We quantize the channel h_{ad} with $a, d \in \{b, i\}, a \neq d$ into K levels and model it as an i.i.d. K state Markov chain with $h_{ad} \in \{h_1, \cdots, h_K\}$ [29]. In one time slot, the attacker chooses an action to decide whether launch an attack or not. The action is determined by the feedback reward of the communication environments, i.e., channel coefficient. Thus, we model action choice process as a finite Markov decision process (MDP), which can be denoted as a 4-tuple $\langle \mathbf{S}, \mathbf{a}, U, P(\cdot|\mathbf{s}, a)\rangle$

- $\mathbf{S}$ is the state set with $\mathbf{s} = (\mathbf{h}_b, \mathbf{g}_b^*, \mathbf{p}_s) \in \mathbf{S}$, where the channel vector $\mathbf{h}_b = [h_{b1}, \cdots, h_{bM}]^{\dagger}$, $\mathbf{p}_s = \{p_{si} | i = 1, \cdots, M\}$ and the number of states is K^{2M}. The size of states can be reduced to K^M with $\mathbf{s} = (\mathbf{h}_b, \mathbf{p}_s)$[3].

[3]This simplification is reasonable since the beamforming makes $\mathbf{g}_b^* \approx \mathbf{h}_b$. If the system states are large, the deep Q-network can be used [24].

- $\mathbf{a} = \{0, 1\}$ is the action set, which represents whether to launch an attack or not.

- $U(\mathbf{s}, a)$ is the immediate reward obtained when $a \in \mathbf{a}$ is taken in state $\mathbf{s}$. As in [30], we define the immediate reward as[4]

$$U(\mathbf{s}, a) = a \sum_{i=1}^{M} \underbrace{\log_2 \left(1 + \frac{p_{si}|g_{bi}^*|^2}{p_b|h_{bi}|^2 + \sigma^2}\right)}_{\text{Throughput}} - \underbrace{p_{si}C_s}_{\text{Transmit cost}}, \tag{8}$$

  where C_s is the unit transmit cost of the attack.

- $P(\cdot|\mathbf{s}, a)$ is the transition probability of the next state, conditioned on action a being chosen in state $\mathbf{s}$.

Q-learning can be used to generate a near-optimal solution to MDP [22]. Furthermore, Q-learning is a model-free reinforcement learning algorithm, and we can obtain the optimal solution without knowing the state transition probability of MDP. We define $Q(\mathbf{s}, a)$ as the Q-function of state $\mathbf{s}$ with action a and define $V(\mathbf{s}) = \max_{a \in \mathbf{a}} Q(\mathbf{s}, a)$ as the maximum long-term cumulative reward of state $\mathbf{s}$ with action a, respectively. Q-function is then iteratively via the iterative Bellman equation,

$$\begin{aligned} Q(\mathbf{s}^{t-1}, a) &= (1 - \varepsilon)Q(\mathbf{s}^{t-1}, a) + \varepsilon \left[U(\mathbf{s}^{t-1}, a) + \varpi V(\mathbf{s}^t)\right] \\ V(\mathbf{s}^t) &= \max_{a \in \mathbf{a}} Q(\mathbf{s}^t, a), \end{aligned} \tag{9}$$

where $\varepsilon \in (0, 1]$ is the learning rate, $\varpi \in (0, 1]$ is the discount factor, $\mathbf{s}^t$ is the next state. To adequately explore the state set, the attacker utilizes ϵ-greedy policy [15]. On this occasion, the probability we choose action a^* can be

[4]Since the ratio between the throughput and the signal-to-interference-plus-noise ratio (SINR) is close to a constant throughout long range of bit rates [31], the SINR can be directly used to represent the throughput in the simulations.

expressed as

$$P(a = a^*) = \begin{cases} 1 - \epsilon, & a^* = \arg\max_{a \in \mathbf{a}} Q(\mathbf{s}, a) \\ \epsilon, & \text{otherwise} \end{cases}, \tag{10}$$

where $\epsilon \in (0, 1)$ is a small positive value, i.e., $\epsilon = 0.1$.

3.2. Optimal Intelligent Attack

From (1), when the attack is present ($\Phi = 1$), the attacker and the sink node are co-existence and the receive power is larger than normal ($\Phi = 0$). Particularly, the larger power the deceiving signal is, the higher risk the attack being detected. Hence, there is a tradeoff between transmit power and detection probability of local observation. We find the worst case for the defenders, namely optimal intelligent attack, in which the attacker can learn the intelligent attack action based on the beamforming with optimal transmit power allocation. In the following, we derive the optimal transmit power allocation.

We define the local observation of sensor i as

$$\psi(\mathbf{y}_i) = \mathbb{E}[\mathbf{y}_i^\dagger \mathbf{y}_i], \tag{11}$$

where $\mathbf{y}_i \in \mathbb{C}^{N \times 1}$ is the receive signal in (1) with sampling number N. Let $f(\psi(\mathbf{y}_i)|\mathcal{H}_\Phi)$ be the probability density function (PDF) of the observation $\psi(\mathbf{y}_i)$ under hypothesis $\mathcal{H}_\Phi$ for $\Phi = 0, 1$. The generalized likelihood ratio test is given by

$$\Lambda(\psi(\mathbf{y}_i)) \triangleq \frac{f(\psi(\mathbf{y}_i)|\mathcal{H}_1)}{f(\psi(\mathbf{y}_i)|\mathcal{H}_0)} \underset{O_i=1}{\overset{O_i=0}{\lessgtr}} \gamma, \tag{12}$$

where $\Lambda(\psi(\mathbf{y}_i))$ is the test statistic, γ is the threshold, $O_i = 0$ and $O_i = 1$ are the observation states of the attack. If the test statistic is less than the threshold, the sensor accepts the hypothesis $\mathcal{H}_0$, otherwise, the sensor accepts the hypothesis $\mathcal{H}_1$. If each term y_i of $\mathbf{y}_i$ is an i.i.d. complex Gaussian random variable with zero-mean and variance σ_y^2, for large N, $\psi(\mathbf{y}_i)$ can converge with

probability one to a Gaussian distribution with mean $\mu_\infty = \sigma_y^2$ and variance $\sigma_\infty^2 = \frac{\sigma_y^4}{N}$. Under $\mathcal{H}_0$, $\psi(\mathbf{y}_i)$ approximately follows a Gaussian distribution, given by

$$f(\psi(\mathbf{y}_i)|\mathcal{H}_0) \sim \mathcal{N}\left(p_b d_{bi}^{-\eta} + \sigma^2, \frac{(p_b d_{bi}^{-\eta} + \sigma^2)^2}{N}\right). \tag{13}$$

Similarly, under $\mathcal{H}_1$, $\psi(\mathbf{y}_i)$ is given by

$$f(\psi(\mathbf{y}_i)|\mathcal{H}_1) \sim \mathcal{N}\left(p_{si} d_{bi}^{-\eta} + 2d_{bi}^{-\eta}\sqrt{p_b p_{si}} + \mathcal{K}, \frac{(p_{si} d_{bi}^{-\eta} + 2d_{bi}^{-\eta}\sqrt{p_b p_{si}} + \mathcal{K})^2}{N}\right), \tag{14}$$

where $\mathcal{K} = p_b d_{bi}^{-\eta} + \sigma^2$. The optimal transmit power allocation can be found by minimizing the detection probability of local observation, subject to the ergodic transmit rate $\bar{\mathcal{R}}_{si} \geq \tau$, which can be expressed as

$$\begin{aligned} \arg\min_{p_{si}} \quad & P_{d,i}^{loc} \\ \text{s.t.} \quad & \bar{\mathcal{R}}_{si} \geq \tau, \end{aligned} \tag{15}$$

where $P_{d,i}^{loc} = \int_\gamma^{+\infty} f(\psi(\mathbf{y}_i)|\mathcal{H}_1) d\psi(\mathbf{y}_i)$.

However, the problem (15) is non-convex, since the ergodic transmit rate $\bar{\mathcal{R}}_{si}$ is not a convex function. To make (15) tractable, we use the Jensen's inequality [32] to obtain the upper bound on $\bar{\mathcal{R}}_{si}$. The upper bound $\lceil \bar{\mathcal{R}}_{si} \rceil$ is given in (16),

where $\mathrm{Ei}(z) = -\int_{-z}^{+\infty} \frac{\exp(-\xi)}{\xi} d\xi$.

$$
\begin{aligned}
\bar{\mathcal{R}}_{si} &= \mathbb{E}\left[\log_2\left(1 + \frac{p_{si}\mathbf{w}_i^{\dagger}\mathbf{h}_{ib\to s}\mathbf{h}_{ib\to s}^{\dagger}\mathbf{w}_i}{p_b|h_{bi}|^2 + \sigma^2}\right)\right] \\
&= \frac{1}{\ln 2}\left\{\mathbb{E}\left[\ln\left(1 + \frac{(p_b + p_{si})d_{bi}^{-\eta}}{\sigma^2}|\tilde{h}_{bi}|^2\right)\right] - \mathbb{E}\left[\ln\left(1 + \frac{p_b d_{bi}^{-\eta}}{\sigma^2}|\tilde{h}_{bi}|^2\right)\right]\right\} \\
&\leq \frac{1}{\ln 2}\left\{\ln\left(1 + \mathbb{E}\left[\frac{(p_b + p_{si})d_{bi}^{-\eta}}{\sigma^2}|\tilde{h}_{bi}|^2\right]\right) - \int_0^{+\infty} \frac{1}{z_1 + 1}\exp\left(-\frac{z_1\sigma^2}{p_b d_{bi}^{-\eta}}\right) dz_1\right\} \\
&= \frac{1}{\ln 2}\left\{\ln\left(1 + \frac{(p_b + p_{si})d_{bi}^{-\eta}}{\sigma^2}\right) + \exp\left(\frac{\sigma^2}{p_b d_{bi}^{-\eta}}\right)\mathrm{Ei}\left(-\frac{\sigma^2}{p_b d_{bi}^{-\eta}}\right)\right\} \qquad (16)
\end{aligned}
$$

Then, the problem (15) can be transformed to minimize the KL divergence from $f(\psi(\mathbf{y}_i)|p_{si}, \mathcal{H}_1)$ to $f(\psi(\mathbf{y}_i)|\mathcal{H}_0)$, subject to $\lceil\bar{\mathcal{R}}_{si}\rceil \geq \tau$. As a result, we have

$$
\begin{aligned}
\arg\min_{p_{si}} \quad & D_{KL}[f(\psi(\mathbf{y}_i)|p_{si}, \mathcal{H}_1)||f(\psi(\mathbf{y}_i)|\mathcal{H}_0)] \\
\text{s.t.} \quad & \lceil\bar{\mathcal{R}}_{si}\rceil \geq \tau, \qquad (17)
\end{aligned}
$$

where

$$
\begin{aligned}
&D_{KL}[f(\psi(\mathbf{y}_i)|p_{si}, \mathcal{H}_1)||f(\psi(\mathbf{y}_i)|\mathcal{H}_0)] \\
&\quad = \int_{-\infty}^{+\infty} \ln\frac{f(\psi(\mathbf{y}_i)|p_{si}, \mathcal{H}_1)}{f(\psi(\mathbf{y}_i)|\mathcal{H}_0)} f(\psi(\mathbf{y}_i)|p_{si}, \mathcal{H}_1) d\psi(\mathbf{y}_i) \\
&\quad = \frac{N\mathcal{A}^2}{2\mathcal{K}^2} + \frac{1}{2}\left[\left(\frac{\mathcal{K} + \mathcal{A}}{\mathcal{K}}\right)^2 - \ln\left(\frac{\mathcal{K} + \mathcal{A}}{\mathcal{K}}\right)^2 - 1\right], \qquad (18)
\end{aligned}
$$

with $\mathcal{A} = p_{si}d_{bi}^{-\eta} + 2d_{bi}^{-\eta}\sqrt{p_b p_{si}}$. Based on the closed-form expression for function D_{KL}, the first derivative of D_{KL} with respect to p_{si} is derived as

$$
\begin{aligned}
&\frac{dD_{KL}[f(\psi(\mathbf{y}_i)|p_{si}, \mathcal{H}_1)||f(\psi(\mathbf{y}_i)|\mathcal{H}_0)]}{dp_{si}} \\
&\quad = \frac{N\mathcal{A}\mathcal{A}'}{\mathcal{K}^2} + \frac{\mathcal{A}\mathcal{A}'(2\mathcal{K} + \mathcal{A})}{\mathcal{K}^2(\mathcal{K} + \mathcal{A})}, \qquad (19)
\end{aligned}
$$

where $\mathcal{A}' = (d_{bi}^{-\eta} + d_{bi}^{-\eta}\sqrt{p_b}/\sqrt{p_{si}})$. Following (19), the second derivative of D_{KL}

with respect to p_{si} is denoted as

$$\begin{aligned}&\frac{d^2 D_{KL}[f(\psi(\mathbf{y}_i)|p_{si},\mathcal{H}_1)||f(\psi(\mathbf{y}_i)|\mathcal{H}_0)]}{d^2 p_{si}}\\&=\frac{N(\mathcal{A}'^2+\mathcal{A}\mathcal{A}'')}{\mathcal{K}^2}+\frac{2(\mathcal{A}'^2+\mathcal{A}\mathcal{A}'')}{(\mathcal{K}+\mathcal{A})^2}\\&+\frac{\mathcal{K}^3\mathcal{A}(2\mathcal{A}'^2+3\mathcal{A}\mathcal{A}'')+\mathcal{K}^2\mathcal{A}^2(\mathcal{A}'^2+\mathcal{A}\mathcal{A}'')}{\mathcal{K}^4(\mathcal{K}+\mathcal{A})^2},\end{aligned}\tag{20}$$

where $\mathcal{A}''=-d_{bi}^{-\eta}\sqrt{p_b}/2p_{si}^{\frac{3}{2}}$. Obviously, we can obtain $\frac{d^2 D_{KL}[f(\psi(\mathbf{y}_i)|p_{si},\mathcal{H}_1)||f(\psi(\mathbf{y}_i)|\mathcal{H}_0)]}{d^2 p_{si}}>0$, which indicates that D_{KL} is a convex function in p_{si}. Following a similar procedure, we have $\frac{d^2\lceil\bar{\mathcal{R}}_{si}\rceil}{d^2 p_{si}}<0$, thus $\lceil\bar{\mathcal{R}}_{si}\rceil$ is a concave function. Then, problem (17) is a convex optimization problem and the Lagrangian $\mathcal{L}_{p_{si},\alpha}$ is defined as

$$\mathcal{L}_{p_{si},\alpha}=D_{KL}[f(\psi(\mathbf{y}_i)|p_{si},\mathcal{H}_1)||f(\psi(\mathbf{y}_i)|\mathcal{H}_0)]-\alpha(\bar{\mathcal{R}}_{si}-\tau),$$

where α is a Lagrange multiplier. Setting the derivative of $\mathcal{L}(p_{si},\alpha)$ with respect to p_{si} to be zero, the necessary and sufficient Karush-Kuhn-Tucker (KKT) conditions (21), (22), (23) are given as follows

$$\text{Dual feasibility:}\alpha\geqslant 0\tag{21}$$

Complementary slackness:

$$\alpha\left(\frac{1}{\ln 2}\left\{\ln\left(1+\frac{(p_b+p_{si})d_{bi}^{-\eta}}{\sigma^2}\right)+\exp\left(\frac{\sigma^2}{p_b d_{bi}^{-\eta}}\right)\mathrm{Ei}\left(-\frac{\sigma^2}{p_b d_{bi}^{-\eta}}\right)\right\}-\tau\right)=0\tag{22}$$

$$\text{Stationarity:}\frac{N\mathcal{A}\mathcal{A}'}{\mathcal{K}^2}+\frac{\mathcal{A}\mathcal{A}'(2\mathcal{K}+\mathcal{A})}{\mathcal{K}^2(\mathcal{K}+\mathcal{A})}-\frac{\alpha}{\ln 2}\left\{\frac{d_{bi}^{-\eta}}{\sigma^2+(p_b+p_{si})d_{bi}^{-\eta}}\right\}=0.\tag{23}$$

For the problem of (17), the optimal transmit power allocation is given in Theorem 1, proved in Appendix A.

Theorem 1. *For $\bar{\mathcal{R}}_{si}\geq\tau$, the optimal transmit power allocation to minimize*

the KL divergence is given by

$$p_{si}^{*}=\left\{\left(\frac{2^{\tau}}{\exp\left(\exp\left(\frac{\sigma^{2}}{p_{b}d_{bi}^{-\eta}}\right)Ei\left(-\frac{\sigma^{2}}{p_{b}d_{bi}^{-\eta}}\right)\right)}\right)-1\right\}d_{bi}^{\eta}\sigma^{2}-p_{b}. \tag{24}$$

Remark 1. We note that (24) explicitly captures the effects of the transmitter location on p_{si}^{*}. For example, when the distance between the sink node and the i-th sensor is smaller, the optimal transmit power allocation will be higher than the case when the distance is longer. It is interesting to show that the optimal transmit power allocation is irrelevant to the location of the attacker (location-free). The reason is that the attacker utilizes the beamforming to emulate the legitimate channel. This result emphasizes that this attacker can perfectly hide its physical location, so it cannot be well detected via PHY-layer spatial decorrelation information, i.e., RSS, CSI.

4. Cooperative PHY-Layer Authentication

In this section, we propose the cooperative PHY-layer authentication scheme with respect to the intelligent location spoofing attack discussed in section 4, which is shown in Fig. 2. In the following, our analysis is based on the optimal intelligent attack, which is the worst case for defenders. Specifically, we formulate the cooperative detection model as MRF and provide the location function and the compatibility function. Then, we develop the complete scheme and analyze the performance. Finally, we propose the cooperative neighbour selection algorithm to accelerate belief convergence and reduce the overhead.

4.1. Cooperative Detection Model

When a sink node requests to access the sensors via location based protocol, the sensors cooperatively authenticate the sink node. We model sensors as random nodes and multi-hop communication links as edges in set $E=\{e_{d}|d=1,\cdots,\mathcal{U}\}$. For example, if two random nodes i and j can communicate mutually, it is called neighbour nodes and there is an edge e_{d} connects them, otherwise

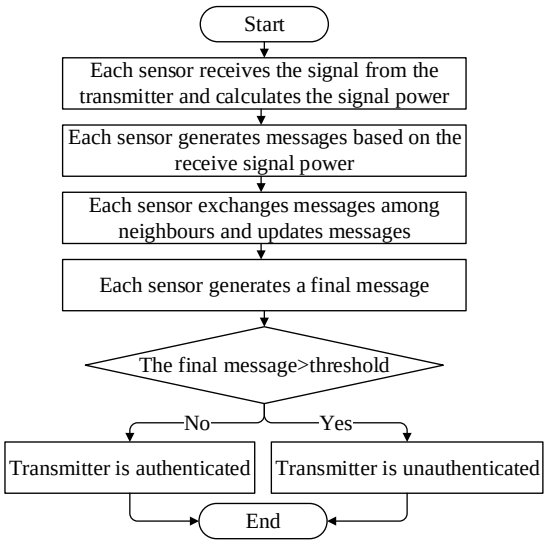

Figure 2: The process of cooperative PHY-layer authentication scheme.

there is no connections between them. We assume each random node i is independent of non-neighbour nodes. In this case, the sensor networks are regarded as an undirected graph $G = (\mathcal{V}, E)$, which can be further represented as a MRF [33].

In MRF, it defines two types of potential function, which are the local function $\phi_i(S_i|\Lambda(\psi(\mathbf{y}_i)))$ and the compatibility function $\varphi_{ij}(S_i, S_j|\Lambda(\psi(\mathbf{y}_i)), \Lambda(\psi(\mathbf{y}_j)))$, respectively. The first one defines how confidence the sensor has to infer the hidden state (real-life state) from the observation state. The latter one represents the correlation between hidden state S_i and hidden state S_j. For simplify, let $\phi_i(S_i|\Lambda(\psi(\mathbf{y}_i))) \triangleq \phi_i(S_i|\Lambda_i)$ and $\varphi_{ij}(S_i, S_j|\Lambda(\psi(\mathbf{y}_i)), \Lambda(\psi(\mathbf{y}_j))) \triangleq \varphi_{ij}(S_i, S_j|\Lambda_i, \Lambda_j)$. We also define the potential function index set $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$, where $\mathcal{F}_1 = \mathcal{V}$ for local function and $\mathcal{F}_2 \subseteq \{\{i, j\} : i, j \in \mathcal{V}, i \neq j\}$ for compatibility function. Next, we analyze the two functions in details.

4.2. Local Function and Compatibility Function

The local function $\phi_i(S_i|\Lambda_i)$ denotes the inference relation between test statistic Λ_i (observation states O_i) and hidden state S_i. In the following, we

analyze the design of local function. Denote $P(a^{t*}|a^{t-1})$ and $P(\tilde{a}^{t*}|a^{t-1})$ as the probabilities that the attack action will be change from an action $a^{t-1} \in \{a^{t-1*}, \tilde{a}^{t-1*}\}$ to the optimal action a^{t*} and non-optimal action $\tilde{a}^{t*}$, respectively. Then

$$\begin{aligned} P(a^{t*}|a^{t-1}) &= P(\mathbf{s}^t, a^{t*}|\mathbf{s}^{t-1}, a^{t-1}) \\ &= P(\mathbf{s}^t|\mathbf{s}^{t-1}, a^{t-1})P(a^{t*}) \\ &= P(\mathbf{h}_b^{t-1}, \mathbf{p}_s^{t-1*}|\mathbf{h}_b^t, \mathbf{p}_s^{t*})(1-\epsilon), \qquad (25) \\ P(\tilde{a}^{t*}|a^{t-1}) &= P(\mathbf{s}^t, \tilde{a}^{t*}|\mathbf{s}^{t-1}, a^{t-1}) \\ &= P(\mathbf{s}^t|\mathbf{s}^{t-1}, a)P(\tilde{a}^{t*}) \\ &= P(\mathbf{h}_b^{t-1}, \mathbf{p}_s^{t-1*}|\mathbf{h}_b^t, \mathbf{p}_s^{t*})\epsilon, \qquad (26) \end{aligned}$$

where $\mathbf{s} = (\mathbf{h}_b, \mathbf{p}_s^*)$ is the state with respect to the optimal intelligent attack. Since the intelligent attack action is time correlated and the current state is correlated with the state in the previous time slot. Thus, it can be exploited for intelligent attack detection.

To simplify analysis, we set $\epsilon = 0$, that is $P(a = a^*) = 1$. Note that this simplification is reasonable since ϵ is a small positive value, i.e., $\epsilon = 0.01$, $P(\tilde{a}^{t*}|a^{t-1}) \approx 0$, and our analyses can be easily extended to the case of $\epsilon \neq 0$. When $\epsilon = 0$, we can regard the intelligent attack action as a two-state Markov chain. In such Markov chain, $P(a^{t*}|a^{t-1*})$ represents an intelligent attack action transition probability from time slot $t-1$ to t.

In MRF, the belief can be used to estimate the marginal probability which decides the real-life state of the attack. For example, in time slot t, if the belief $b_{i,t}$ is less than a threshold, we use $S_{i,t} = 0$ to represent the absence of the attack, otherwise, we use $S_{i,t} = 1$ to represent the presence of the attack. For this intelligent attack, the transition probability of the hidden state between time slot $t-1$ and t is given by $P(S_{i,t}|S_{i,t-1})$, where $S_{i,t}, S_{i,t-1} \in \{0, 1\}$ with $P(1|0) + P(0|0) = 1$ and $P(0|1) + P(1|1) = 1$. It is seen that the state transition probability of the hidden state is equal to the transition probability of the

intelligent attack action.

Remark 2. Here, we obtain an important insight that the state transition probability is proportional to the channel time-vary speed which is related to the variation of communication environments. Thus, the intelligent location spoofing attack is more serious in time-vary complex communication environments, such as crowed urban areas.

As analyzed, the current hidden state is correlated with the hidden state in the previous time slot, we develop a temporal dimension on the local function of MRF. Fig. 3 shows the relations among the current hidden state $S_{i,t}$, the current test statistic $\Lambda_{i,t}$, the previous belief $b_{i,t-1}$, and compatibility function. As in [34], if the test statistic $\Lambda_{i,t}$ belongs to $O_{i,t} = 0$, no attack is detected

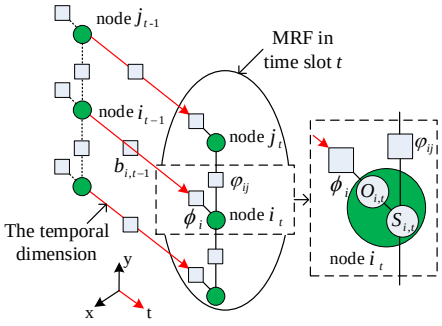

Figure 3: The MRF with the temporal dimension. The green node represents random node which is composed by the observation state $O_{i,t}$ and the hidden state $S_{i,t}$.

at sensor i. In this case, it is given a small value to represent the hidden state $S_{i,t} = 0$. When $\Lambda_{i,t}$ belongs to $O_{i,t} = 1$, the attack is detected at sensor i. In this case, it is given a large value to show the confidence of the hidden state $S_{i,t} = 1$.

Based on the above analysis, the local function can be written as

$$\phi_i(S_{i,t}|\Lambda_{i,t}, b_{i,t-1}) = \begin{cases} \varsigma(\frac{1}{\Omega} + \kappa\hat{b}_{i,t}), & O_{i,t} = 0 \\ \varsigma(\delta(O_{i,t} = 1) + \kappa\hat{b}_{i,t}), & O_{i,t} = 1 \end{cases}, \tag{27}$$

where

$$\hat{b}_{i,t} = \sum_{S_{i,t-1}} P(S_{i,t}|S_{i,t-1}) b_{i,t-1} \tag{28}$$

represents the one-step ahead belief prediction of the hidden state $S_{i,t}$ based on Bayesian forecasting [35], $\frac{1}{\Omega}$ with $\Omega \geq 2$ is a uniform distribution, $\delta(O_{i,t} = 1)$ is an impulse distribution, $\varsigma = \frac{1}{1+\kappa}$ is a normalization factor, and $\kappa \in (0, 1)$ is a forgetting factor represents how sensor views the importance of the previous belief[5].

Remark 3. Note that the local function gives a large value to the node when infers the hidden state $S_i = 1$ and gives a small value to the node, i.e., $\varsigma(\frac{1}{\Omega} + \kappa\hat{b}_{i,t})$, when infers the hidden state $S_i = 0$.

In the following, if no confusions occur, we omit the symbol $b_{i,t-1}$ in $\phi_i(\cdot)$ and the time slot index t in the subscript of $S_{i,t}$, $O_{i,t}$ and $\Lambda_{i,t}$. Note that the local function has to update in each time slot, we propose the update process in Fig. 4. We find that the algorithm is recursive, that is, the local function ϕ_i in time slot t is calculated by the belief $b_{i,t-1}$ in time slot $t-1$. For the initial belief $b_{i,0}$ in time slot $t = 0$, we consider an arbitrary value, i.e., $b_{i,0} = 0.5$.

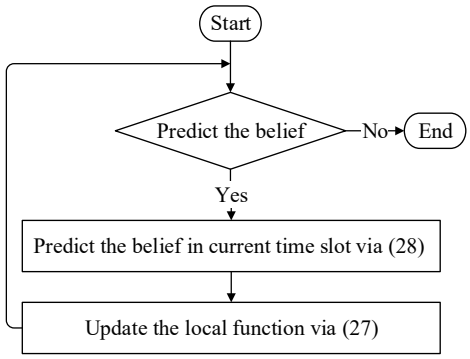

Figure 4: The process of local function update.

[5]Note that our proposed authentication scheme can be used to authenticate traditional "blunt" location spoofing attack by setting the forgetting factor κ to be zero.

The compatibility function $\varphi_{ij}(S_i, S_j|\Lambda_i, \Lambda_j)$ represents the correlation between sate S_i and state S_j. Sensor i and sensor j that are close to each other are more likely to have correlated receive signal power, which means the observations from these users are more related than that from further away. For illustrative purposes, we use the receive signal power correlation between sensor i and sensor j to represent $\varphi_{ij}(S_i, S_j|\Lambda_i, \Lambda_j)$. Considering a less computationally complex solution, we define the following simple compatibility function between two neighbours via the Potts model [34]

$$\varphi_{ij}(S_i, S_j|\Lambda_i, \Lambda_j) = \begin{cases} \theta_{ij}, & S_i = S_j \\ \frac{1-\theta_{ij}}{\Omega-1}, & S_i \neq S_j \end{cases}, \tag{29}$$

where $\theta_{ij} = \exp\left(-\psi(\mathbf{y}_i) \log \frac{\psi(\mathbf{y}_i)}{\psi(\mathbf{y}_j)}\right)$ with $0 < \theta_{ij} < 1$, $\theta_{ij} \gg (1-\theta_{ij})/(\Omega-1)$.

Remark 4. The larger θ_{ij} is, the higher correlation the neighbour sensors will be. We see that the compatibility function encourages neighbour sensor to have the same state.

4.3. Complete Scheme and Performance Metrics

Here, we propose the complete scheme and analyze the related performance metrics. Belief propagation is an information passing algorithm, which operates in a pairwise MRF to compute marginal probability associated with the joint probability [36]. Since the cooperation of the sensors, all neighbouring sensors can predict and exchange the information with each other using Bayesian forecasting and belief propagation, namely, *belief forecasting propagation*.

The joint probability $\mathbb{S}$ conditional on $\mathbf{\Lambda}$ can be written as

$$P(\mathbb{S}|\mathbf{\Lambda}) = \prod_{i=1}^{M} \phi_i(S_i|\Lambda_i) \prod_{i \neq j} \varphi_{ij}\{S_i, S_j|\Lambda_i, \Lambda_j\}. \tag{30}$$

where $\mathbb{S} = \{S_i|\ i = 1, \cdots, M\}$ is the hidden state set, and $\mathbf{\Lambda} = \{\Lambda_i|i = 1, \cdots, M\}$ is the test statistic set. The goal is to independently compute the

marginal probability $P(S_i|\mathbf{\Lambda})$ refers to belief for each sensor, and make decisions. We define an information passing from sensor j to sensor i as $m_{ij}(S_i)$. Since belief forecasting propagation can be used to iteratively calculate the marginal probability for each hidden state, conditional on any observation states. In the l-th iteration, the information $m_{ij}^l(S_i)$ that sensor j transmits to sensor i can be updated by

$$m_{ij}^l(S_i) \quad = \quad C_j \sum_{S_j} \phi_j(S_j|\Lambda_j)\varphi_{ij}(S_i, S_j|\Lambda_i, \Lambda_j) \prod_{k\neq i,j} m_{kj}^{l-1}(S_j), \tag{31}$$

where C_j is a normalization factor so that $m_{ij}^l(S_i = 1) + m_{ij}^l(S_i = 0) = 1$, m_{ij}^l means the belief about the state of sensor i, which is estimated by sensor j. When the algorithm convergence, i.e., n iterations, each sensor obtains its final belief which is represented as

$$b_i(S_i) = c_i\phi_i(S_i|\Lambda_i) \prod_{i\neq j} m_{ij}^n(S_i), \tag{32}$$

where c_i is a normalization factor. Then the marginal probability $P(S_i|\mathbf{\Lambda})$ can be approximately estimated via belief $b_i(S_i)$. Theorem 2 proves the convergence of the algorithm, and the proof is in Appendix B.

Theorem 2. *For binary variables with pairwise interactions, if*

$$|J_{ij}| < \arctan\left(\frac{1}{M-1}\right), \tag{33}$$

where $|J_{ij}| = -\ln\theta_{ij}$ is the "couplings" of sensor i and sensor j, then belief forecasting propagation is an ℓ_1-contraction and converges with probability one to an unique fixed point irrespective of the initial information.

Remark 5. Here, we obtain an insight that when the number of sensors is large-scale, a smaller θ_{ij} is required to satisfy the belief convergence condition. We find that the convergence constraint is relaxed with the increasing of the number of sensors. In other words, the more sensors within the WSN, the better

convergence performance the algorithm will be.

When all the sensors obtain the final beliefs, they compare the belief with a belief threshold b_T and the decision rule is given by

$$b_i \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} b_T. \tag{34}$$

If the belief of sensor i is higher than the threshold b_T, sensor i considers the attack is present, i.e., $S_i = 1$, otherwise not. The complete cooperative PHY-layer authentication scheme is presented in Algorithm 1.

Algorithm 1 Cooperative PHY-layer authentication

1: Initialize M, $b_{i,0}$, $S_{i,0}$, κ, b_T, n;
2: Each sensor calculates the local observation via (12);
3: Each sensor infers the observation state O_i via (11);
4: Each sensor calculates the local function value and the compatibility function value via (27) and (29);
5: **for** iteration $\leftarrow 1$ to n **do**
6: Each sensor exchanges messages among neighbours and updates the information $m_{ij}^l(S_i)$ via (31);
7: **end for**
8: Calculates the final belief via (32);
9: Each sensor finally decides the state (hidden state) of the attack via (34);
10: **if** $S_i = 1$ **then**
11: The transmitter is unauthenticated;
12: **else**
13: The transmitter is authenticated.
14: **end if**

The detection probability and the false alarm probability are presented as follows

$$\begin{aligned} P_{f,i} &= Pr(b_i > b_T|\mathcal{H}_0) \\ &= \frac{b_i(0) \cap b_i(1)}{b_i(0)}, \end{aligned} \tag{35}$$

$$\begin{aligned} P_{d,i} &= Pr(b_i > b_T|\mathcal{H}_1) \\ &= 1 - \frac{b_i(0) \cap b_i(1)}{b_i(1)}, \end{aligned} \tag{36}$$

where $Pr(\cdot|\cdot)$ is the conditional probability, b_T is the belief threshold, $b_i(S_i)$

is the belief of the state with $S_i \in \{0, 1\}$. After belief iterating, the belief of sensor i under the attack and without the attack are given by (C.2) and (C.4) in Appendix C, respectively. After some algebraic computations, the bounds on belief can be expressed as

$$\lfloor b_i(1) \rfloor = B_1 \left[\left(\frac{1}{\Omega} + \kappa \hat{b}_{j,t} \right) \frac{1}{\Omega - 1} \right]^m, \tag{37}$$

$$\lceil b_i(1) \rceil = B_1 \left[\delta + \kappa \hat{b}_{j,t} \right]^m, \tag{38}$$

$$\lfloor b_i(0) \rfloor = \begin{cases} B_0 \left[\frac{1}{\Omega} + \kappa \hat{b}_{j,t} \right]^m, & \text{if } \hat{b}_{j,t} < \frac{1}{\kappa(\Omega^2 - 2\Omega)}, \\ B_0 \left[\left(\delta + \kappa \hat{b}_{j,t} \right) \frac{1}{\Omega - 1} \right]^m, & \text{if } \hat{b}_{j,t} \geq \frac{1}{\kappa(\Omega^2 - 2\Omega)}, \end{cases} \tag{39}$$

$$\lceil b_i(0) \rceil = \begin{cases} B_0 \left[\frac{1}{\Omega} + \kappa \hat{b}_{j,t} \right]^m, & \text{if } \hat{b}_{j,t} \geq \frac{1}{\kappa(\Omega^2 - 2\Omega)}, \\ B_0 \left[\left(\delta + \kappa \hat{b}_{j,t} \right) \frac{1}{\Omega - 1} \right]^m, & \text{if } \hat{b}_{j,t} < \frac{1}{\kappa(\Omega^2 - 2\Omega)}, \end{cases} \tag{40}$$

where $B_1 = c_i \varsigma^{m+1} \left[\delta + \kappa \hat{b}_{i,t} \right]$, $B_0 = c_i \varsigma^{m+1} \left[\frac{1}{\Omega} + \kappa \hat{b}_{i,t} \right]$, and we omit $O_i = 1$ in $\delta(\cdot)$ for brevity. Please see the derivation in Appendix C.

Remark 6. From (37)-(40), We can deeply understand the algebraic relation among current belief, the previous belief, state transition probability and the number of neighbours. For example, the bounds are exponential functions with respect to the number of neighbours m.

The closed-form expression for belief threshold b_T and the specific value of it can be given in Theorem 3 which is proved in Appendix D.

Theorem 3. *For $\forall$ $P_{f,i}$, the closed-form expression for belief threshold b_T is obtained by*

$$b_T = P_{f,i} \lfloor b_i(0) \rfloor + (1 - P_{f,i}) \lceil b_i(0) \rceil, \tag{41}$$

and the specific value of belief threshold b_T can be obtained by predetermining a false alarm probability, i.e., $P_{f,i} = 0.1$.

4.4. Cooperative Neighbour Selection

Since the whole computation complexity takes a time proportional to the number of links in the graph, and the belief propagation are not exact due to the double counting problem in a loopy graph [36], we propose the cooperative neighbour selection algorithm to accelerate belief convergence and reduce the overhead. In WSN, the quality of service (QoS) of each multi-hop link is different [37]. Our optimization goal is to select the neighbour sensor who has the optimal multi-hop link to maximize global QoS, subject to no loop in the network.

Definition 1. *Let the weight of edge is a mapping $\omega : E \to W$ with the weight set $W = \{w_d | d = 1, \cdots, \mathcal{U}\}$. Define the weight complement of edge e_d as $w_d^c = \sum_{d=1}^{\mathcal{U}} w_d - w_d$, and define the new graph G^c as the complement graph of G.*

Based on graph theory [38], we formulate the optimization problem as finding the maximum spanning tree $\mathbb{T}$, which can be written as

$$\begin{aligned} &\max_{w_d} \quad \mathbb{T} \\ &\text{s.t.} \quad \forall e_d \in E, e_d \cup \mathbb{T} \text{ contains a loop.} \end{aligned}$$

This optimization problem can be solved by exploiting the following theorem, which is proved in Appendix E.

Theorem 4. *Solving the minimum spanning tree $\mathbb{T}^c$ of complement graph G^c is equal to search the maximum spanning tree of graph G. In other words, the greedy algorithm can be utilized to maximize spanning tree $\mathbb{T}$.*

The weight of the edge e_d is designed as

$$w_d = \mathcal{C}_\omega |h_{id}|^2, \qquad d = 1, \cdots, \mathcal{U} \tag{42}$$

where $\mathcal{C}_\omega$ is the unit weight, and $|h_{id}|^2$ is the instantaneous channel gain between the i-th sensor and its d-th neighbour. The complete algorithm is described in Algorithm 2.

Remark 7. Complexity analysis: The standard implementation of the information passing on the loopy graph takes $\mathcal{O}(ML^2n)$ run time, where n is the number of iterations, and L is the number of the possible state for sensors, i.e., here $L = 2$. While, the run time reduces to $\mathcal{O}(ML^2)$ by using the proposed cooperative neighbour selection algorithm in that the iteration passes through each sensor on the tree only once. We find that the run time is linear increasing with the number of sensors, which suggests that the overhead is low even in large-scale networks. Furthermore, in information passing, only a short belief message is needed to communicate with neighbours, which takes a low overhead.

Remark 8. Efficiency analysis: To calculate the receive signal power, we assume that each sensor takes $\mathcal{O}(1)$ run time. Then, with Remark 7, the total run time of a single sensor is $\mathcal{O}(1)+\mathcal{O}(L^2)$. Similarly, we can calculate the total run time of M cooperative sensors as $\underbrace{\mathcal{O}(1)+\cdots+\mathcal{O}(1)}_{M}+\underbrace{\mathcal{O}(L^2)+\cdots+\mathcal{O}(L^2)}_{M} = M\mathcal{O}(1)+M\mathcal{O}(L^2)$. For M cooperative sensors, the average run time for each sensor is $\frac{M\mathcal{O}(1)+M\mathcal{O}(L^2)}{M} = \mathcal{O}(1)+\mathcal{O}(L^2)$, which means the run time is linear increasing with the number of sensors and the run time for each sensor remains unchanged. In addition, with the number of sensors increases, the detection performance has been significantly improved, which proves the efficiency.

Algorithm 2 Cooperative neighbour selection

1: Initialize probe messages;
2: **for** sensor $\leftarrow 1$ to M **do**
3: Broadcast probe message to the neighbours;
4: Feed back probe messages from neighbours to sensor;
5: Calculate the weight of each edge via (42).
6: **end for**
7: Obtain the maximum spanning tree via greedy algorithm.

5. Simulations and Performance Analysis

We verify the theoretical analysis and show the performance of the proposed cooperative authentication scheme by simulations. In the simulations, we set

the transmit power $p_b = 30$ dBm, the number of samples $N = 200$, the noise variance $\sigma = -10$ dBm, the path loss exponent $\eta = 2$ [39]. We set the WSN coverage radius to be 2 km, the location of the sink node to be $(0, 0)$, the location of the attack to be $(0.5, -0.5)$. We set state transition probability to be $P(1|0) = P(0|1) = 0.4$, $P(1|1) = P(0|0) = 0.6$, initial belief $b_{i,0} = 0.5$. The optimal spanning tree topologies are obtained by using Algorithm 2. For example, as shown in Fig. 5 with $M = 9, 11, 13, 15$, the topologies are optimal spanning trees with no loop in networks. These optimal topologies are obtained via random generated sensors, and the following analysis are based on these topologies.

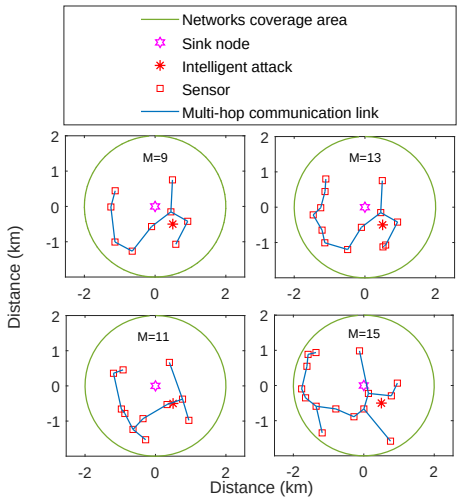

Figure 5: The network topologies via cooperative neighbour selection algorithm, where $M = 9, 11, 13, 15$.

Fig. 6 is obtained via Q-learning on ϵ-greedy policy with learning rate $\varepsilon = 0.8$, discount factor $\varpi = 0.9$ and $\epsilon = 0.15, 0.1, 0.05$, $C_s = 1$ dB $= 1.3$ mW, $K = 3$. As shown in the figure, the long-term cumulative reward converges to 105.2 dB, 117.4 dB and 119.1 dB after 900 time slots, respectively. We conclude that the intelligent attack can achieve a stable long-term cumulative reward via Q-learning. Specifically, a high exploration, i.e., $\epsilon = 0.15$, can converge to a higher cumulative reward than a low exploration. This is because the high

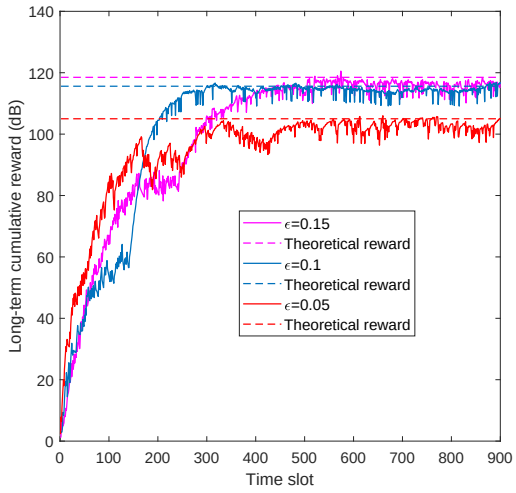

Figure 6: The long-term cumulative reward via Q-learning with different ϵ, instantaneous reward vs. theoretical reward.

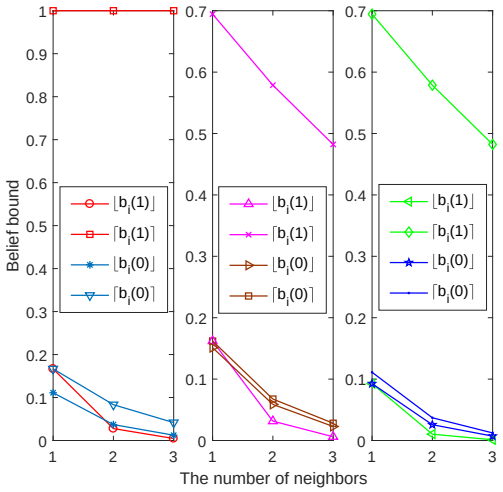

Figure 7: The belief bounds vs. the number of neighbours.

exploration can avoid converging to the local optimum at the training process. However, the high exploration reduces the Q-learning convergence speed. For example, the algorithm converges to theoretical reward at 500 time slot with $\epsilon = 0.15$, while it converges to theoretical reward at 300 time slot with $\epsilon = 0.05$.

Fig. 7 demonstrates the lower bound and the upper bound on belief of sensor i with $\Omega = 3, 3, 4$ and $\kappa = 0, 0.5, 0.5$, respectively. With regard to these three subfigures, we find that the belief bounds are inversely proportional to the number of its neighbours, except the upper bound on $b_i(1)$ with $\Omega = 3, \kappa = 0$. From the figure, we can conclude that the upper bound on $b_i(1)$ depends on κ rather than Ω. We can also observe that the gap of the belief bounds become smaller, with the increasing of Ω and forgetting factor κ. It implies that by considering the belief of the previous time slot, the uncertainty of the current belief is decline. Furthermore, we can see that the bound $b_i(0)$ belongs to the bound $b_i(1)$, which suggests that the false alarm probability $P_{f,i}$ and the miss detection probability $1 - P_{d,i}$ always exist. These phenomenons are also confirmed by the Theorem 3.

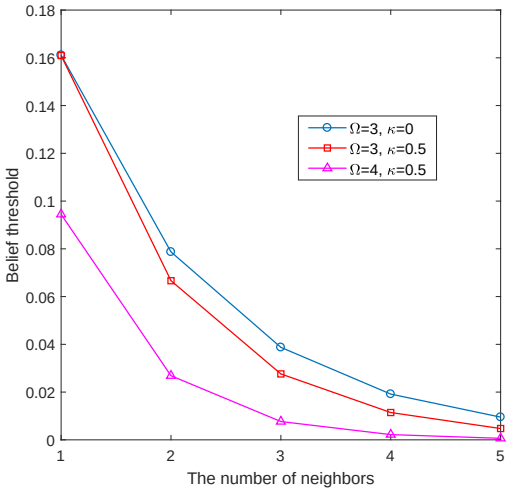

Figure 8: The belief threshold b_T vs. the number of neighbours.

Fig. 8 analyzes the variation trend of the belief threshold. As the number of neighbours increase, we see that the belief threshold is decreased exponentially, which can also be confirmed by the insights in Remark 6. We also see that a higher Ω and forgetting factor κ can lead to a lower belief threshold.

Fig. 9 illustrates the effect of different distance from the sink node to the

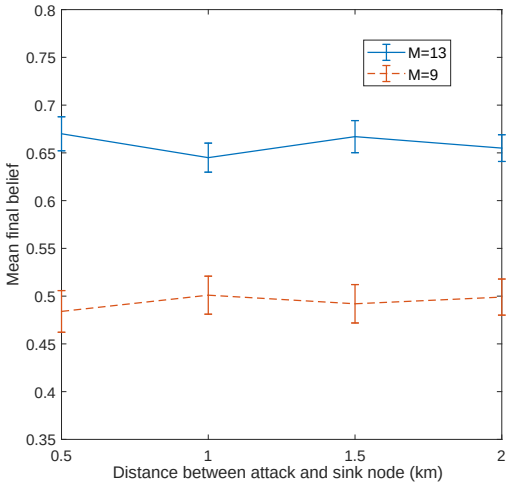

Figure 9: The final belief mean, variance vs. distance between attack and sink node.

attacker on the final belief with $M = 9, 13$. The simulation result is obtained from 1000 times Monte-Carlo simulations. It is shown that when the distance increase from 0.5 km to 2 km, the mean final belief and variance are almost stable, which means the attack is location-free as analyzed in Remark 1. We also find that the more sensors cooperation, the higher final belief the algorithm has. For example, the mean final belief is around 0.65 when the sensor number is $M = 13$, while the mean final belief is around 0.5 when the sensor number is $M = 9$. The result shows that the more sensors cooperation, the higher robust the algorithm is. Note that although the intelligent attack is location-free, the long distance between sink node and attack will increase the difficulty of beamforming.

Fig. 10 evaluates the performance of the proposed scheme via the receiver operating characteristic (ROC) curves in 10000 times Monte-Carlo simulations. When the intelligent attack uses the beamforming with optimal transmit power allocation, the detection performance reduces more than 5% than the case with random transmit power allocation. Moreover, although more sensors can cooperation to improve detection performance, it is hard to offset the influence

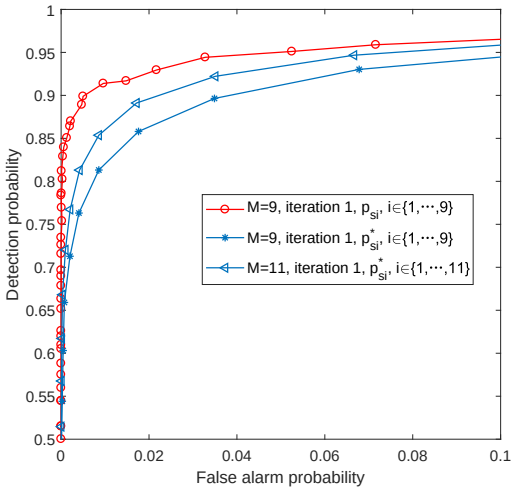

Figure 10: The ROC of the proposed scheme with different transmit power allocation. The optimal transmit power allocation vs. the random transmit power allocation.

that the optimal transmit power allocation brings. Thus, the optimal intelligent attack which uses beamforming with optimal transmit power allocation is the worst case in our analysis.

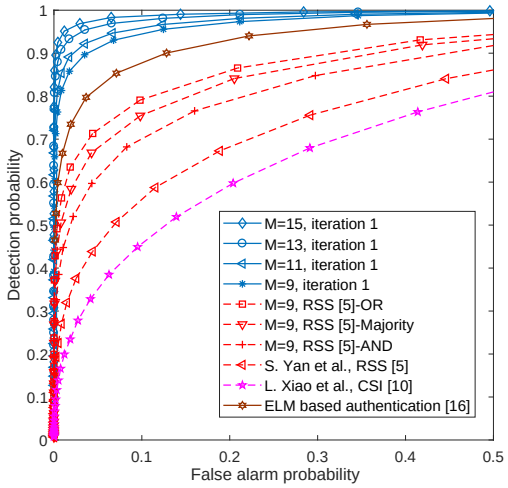

Figure 11: The ROC of the proposed scheme vs. the benchmark methods: ELM based authentication [16], CSI [10], RSS [5], and RSS in OR, Majority, AND.

505

Fig. 11 illustrates the detection performance of the proposed scheme and the benchmark methods under the threat of the optimal intelligent attack. In the figure, the dotted line represents the traditional methods and the solid line represents the machine learning inspired methods. It is shown that the proposed scheme can significantly improve the detection performance. For example, the performance gain of our scheme over traditional CSI based method [10] more than 50% with $P_{f,i} = 0.1$ and over traditional RSS based method [5] more than 40% with $P_{f,i} = 0.1$. It suggests that both CSI and RSS based method are hardly detect such an attack. The performance of the CSI based method is the lowest because the intelligent attack imitates the channel via beamforming. In addition, we leverage the RSS method as a benchmark and further compare our method with RSS in three fusion rules, i.e., RSS-OR, RSS-Majority, RSS-AND. The results show that our method obviously outperforms the RSS in three fusion rules. As a machine learning inspired method, the extreme learning machine (ELM) based authentication [16] has a better detection performance than the traditional methods. It is because that the machine learning inspired method can obtain the deep channel characteristics from training data, thus can identify the attacker more accurately. However, our proposed method is superior to the ELM based authentication. Since the sensor cooperation and the history information consideration are important factors in detecting the proposed intelligent location spoofing attack. In addition, when the number of sensors increases from 9 to 15, the detection performance has obviously improved. It suggests that the more sensors cooperation, the better performance the scheme will be.

6. Future Work

In this paper, we mainly focus on the static WSN scenario. It has many important applications in mission critical internet of things, such as border patrol. In such a scenario, the sensor are fixed deployment on the border to monitor border security. The mobile WSN also has widely applications in practice. Thus, in the future works, we will investigate the cooperative PHY-layer authentication

scheme in mobile WSN. In addition, the non-Gaussian noise generally exists in nonlinear stochastic models [40, 41], i.e., wireless channel model. We should also consider the effect of non-Gaussian noise on measurements in the practical systems.

7. Conclusion

In this paper, we have proposed a cooperative PHY-layer authentication scheme to defend against an intelligent location spoofing attack in WSN. To attack, we have analyzed the feasibility of it and found the optimal intelligent attack. To protect, we have modeled the networks as a MRF and have designed the local function and the compatibility function. We have obtained the expressions for the detection probability and false alarm probability. We have obtained the belief bounds and the closed-form expression for belief threshold. We have proposed the cooperative neighbour selection algorithm to accelerate belief convergence and reduce the overhead. The simulations have validated the theoretical analysis and compared with five benchmark methods. Some insightful remarks have obtained, for example: 1) The optimal transmit power allocation is irrelevant to the location of the attack. 2) We should be more alert on such an attack in time-vary complex communication environments, i.e., crowed urban areas.

Appendix A. Proof of Theorem 1

According to the necessary and sufficient KKT conditions, the optimal transmit power allocation can be $p_{si}^* = \lceil \bar{\mathcal{R}}_{si}^{-1}(\tau) \rceil$ or take a value in the open interval $(\lceil \bar{\mathcal{R}}_{si}^{-1}(\tau) \rceil, +\infty)$. This gives rise to the following two cases:

Case 1: Suppose $p_{si}^* = \lceil \bar{\mathcal{R}}_{si}^{-1}(\tau) \rceil$, in this case, (22) requires that $\alpha \geq 0$. Substituting this into (23), we have $\frac{dD_{KL}}{dp_{si}} / \frac{d\lceil \bar{\mathcal{R}}_{si} \rceil}{dp_{si}} \geq 0$. Therefore, equation (19) satisfying

$$\frac{dD_{KL}[f(\psi(\mathbf{y}_i)|p_{si}, \mathcal{H}_1)||f(\psi(\mathbf{y}_i)|\mathcal{H}_0)]}{dp_{si}} \geq 0. \tag{A.1}$$

Note that (24) satisfies equation (A.1).

Case 2: Suppose $\lceil \bar{\mathcal{R}}_{si}^{-1}(\tau) \rceil < p_{si}^* < +\infty$, in this case, (22) requires that $\alpha = 0$. Substituting this into (23), we have $\frac{dD_{KL}}{dp_{si}} / \frac{d\lceil \bar{\mathcal{R}}_{si} \rceil}{dp_{si}} = 0$. Therefore, equation (19) satisfying

$$\frac{dD_{KL}[f(\psi(\mathbf{y}_i)|p_{si}, \mathcal{H}_1)||f(\psi(\mathbf{y}_i)|\mathcal{H}_0)]}{dp_{si}} = 0. \tag{A.2}$$

However, p_{si}^* does not satisfy the supposing constraint, then the proof is completed.

Appendix B. Proof of Theorem 2

Lemma 1. *Let $(Q, \| \cdot \|)$ be a normed space (a normed finite dimensional real vector space) and $f : Q \to Q$ a differentiable mapping. Then, for $q_1, q_2 \in Q$*

$$\| f(q_1) - f(q_2) \| \leq \| q_1 - q_2 \| \cdot \sup_{q_3 \in [q_1, q_2]} \| f'(q_3) \|, \tag{B.1}$$

where denote $[q_1, q_2]$ as the segment $\{\xi q_1 + (1-\xi)q_2 : \xi \in [0,1]\}$ joining q_1 and q_2 [42].

Combining contracting mapping principle [33] and Lemma 1, we have the following lemma.

Lemma 2. *Let $(Q, \| \cdot \|)$ be a normed space and $f : Q \to Q$ be a differentiable mapping. If $\sup_{q \in Q} \| f'(q) \| < 1$, then f is a $\| \cdot \|$ contraction. Moreover, we deduce that for any $q \in Q$, the sequence $q, f(q), f^2(q), \cdots$ converges to a unique fixed point $q_\infty \in Q$ with a convergence rate that is at least linear.*

The belief forecasting propagation update (31) can be rewritten as [33]

$$\tanh \tilde{\nu}_{j \to i} = \tanh(J_{ij}) \tanh \left(\theta_j + \sum_{v \in \partial j \setminus i} \nu_{v \to j} \right), \tag{B.2}$$

where $\tilde{\nu}_{j \to i}$ is the information sent from variable j to i, θ_j is the "local fields" with respect to ϕ_j and $\partial j = \{v \in \mathcal{V} : \{j, v\} \in \mathcal{F}_2\}$ are the variables that

interact with i via a compatibility function. Defining the set of ordered pairs $D \triangleq \{i \to j : \{i,j\} \in \mathcal{F}_2\}$, we see that the parallel propagation update is a mapping $f : \mathbb{R}^D \to \mathbb{R}^D$, (B.2) specifies the component $f(\nu)_{i\to j} \triangleq \tilde{\nu}_{i\to j}$ in terms of the components of ν. Denote the ℓ_1 norm of a linear mapping $\mathbf{A}$ on $\mathbb{R}^D$ as

$$\| \mathbf{A} \|_1 = \max_{j\in\{1,\cdots,N\}} \sum_{i=1}^{N} |\mathbf{A}_{ij}|, \tag{B.3}$$

where $\mathbf{A}_{ij} \triangleq (\mathbf{A}\mathbf{e}_j)_i$, $\mathbf{e}_j$ is the j-th canonical basis vector. The derivative of f is calculated from (B.2) and is given by

$$f'(\nu)_{j\to i,k\to l} = \frac{\partial \tilde{\nu}_{j\to i}}{\partial \nu_{k\to l}} = \mathbf{A}_{j\to i,k\to l}\mathbf{B}_{j\to i}(\nu), \tag{B.4}$$

where

$$\mathbf{B}_{j\to i}(\nu) \triangleq \frac{1-\tanh^2(\theta_j + \Sigma_{v\in\partial j\setminus i}\nu_{v\to j})}{1-\tanh^2(\tilde{\nu}_{j\to i}(\nu))}\mathrm{sgn}J_{ij} \tag{B.5}$$

and the linear mapping $\mathbf{A}_{j\to i,k\to l} \triangleq \tanh|J_{ij}|\delta_{j,l}\mathbf{1}_{\partial j\setminus i}(k)$.

Since $\sup_{\nu\in Q}|\mathbf{B}(\nu)_{j\to i}| = 1$ and $\mathbf{A}_{j\to i,k\to l}$ are nonnegative and independent of ν. For everywhere on Q, we obtain $\left|\frac{\partial\tilde{\nu}_{j\to i}}{\partial\nu_{k\to l}}\right| \le \mathbf{A}_{j\to i,k\to l}$. Choosing the ℓ_1 norm on $\mathbb{R}^D$, we obtain

$$\begin{aligned}
\| f'(\nu)_{j\to i,k\to l} \|_1 &= \max_{k\to l}\sum_{j\to i}\left|\frac{\partial\tilde{\nu}_{j\to i}}{\partial\nu_{k\to l}}\right| \\
&\le \max_{k\to l}\sum_{i\to j}\tanh|J_{ij}|\delta_{j,l}\mathbf{I}_{\partial j\setminus i}(k) \\
&= \max_{j\in\mathcal{V}}\max_{k\in\partial j}\sum_{i\in\partial j\setminus k}\tanh|J_{ij}|.
\end{aligned} \tag{B.6}$$

Apply Lemma 2, we have

$$\max_{j\in\mathcal{V}}\max_{k\in\partial j}\sum_{i\in\partial j\setminus k}\tanh|J_{ij}| < 1. \tag{B.7}$$

For random nodes $j \in \mathcal{V}$, substituting the maximum dimension of its neigh-

bouring random nodes, (33) is obtained and the proof is completed.

Appendix C. The Derivation of Belief Bounds

Supposing the number of neighbours of sensor i is m, the messages passing from neighbour j to i are independent. Substituting (31) into (32), we derive the belief with attack in (C.1) at the top of next page. Then, substituting (29) and (27) into (C.1), (C.1) can be transformed to (C.2). After algebraic manipulations, we find that (C.2) is a function of θ_{ij} with $0 < \theta_{ij} < 1, j = 1, \cdots, m$. Taking logarithm to (C.2), we have (C.3). Note that each term of (C.3) is a convex function. By maximizing/minimizing the log-function of each term, we obtain the bounds on the belief of (C.2) in (37), (38). Similarly, from (C.4), (C.5), we obtain the bounds on the belief of (C.4) in (39), (40). In (39), (40), the bounds are piecewise functions, which depend on the one-step ahead belief prediction.

$$
\begin{aligned}
&b_i(1) = c_i\phi_i(1)m_1(1)\cdots m_m(1)\\
&= c_i\phi_i(S_i=1)[P(S_1=1|\Lambda_1)P(S_i=1|S_1=1)+P(S_1=0|\Lambda_1)P(S_i=1|\mathcal{S}_1=0)]\\
&\cdots[P(S_m=1|\Lambda_m)P(S_i=1|S_m=1)+P(S_m=0|\Lambda_m)P(S_i=1|S_m=0)].
\end{aligned}
\tag{C.1}
$$

$$
\begin{aligned}
&b_i(1)\\
&= c_i\varsigma^{m+1}\left[\delta(O_i=1)+\kappa\hat{b}_{i,t}\right]\left[\left(\delta(O_1=1)+\kappa\hat{b}_{1,t}\right)\theta_{i1}+\left(\frac{1}{\Omega}+\kappa\hat{b}_{1,t}\right)\frac{1-\theta_{i1}}{\Omega-1}\right]\\
&\cdots\left[\left(\delta(O_m=1)+\kappa\hat{b}_{m,t}\right)\theta_{im}+\left(\frac{1}{\Omega}+\kappa\hat{b}_{m,t}\right)\frac{1-\theta_{im}}{\Omega-1}\right]\\
&= c_i\varsigma^{m+1}\left[\delta(O_i=1)+\kappa\hat{b}_{i,t}\right]\prod_{j=1}^{m}\left[\left(\delta(O_j=1)+\kappa\hat{b}_{j,t}\right)\theta_{ij}+\left(\frac{1}{\Omega}+\kappa\hat{b}_{j,t}\right)\frac{1-\theta_{ij}}{\Omega-1}\right].
\end{aligned}
\tag{C.2}
$$

$$\begin{aligned}\ln b_i(1) &= \ln c_i \varsigma^{m+1} + \ln[\delta(O_i = 1) + \kappa \hat{b}_{i,t}] \\ &+ \ln\left[\left(\delta(O_1 = 1) + \kappa \hat{b}_{1,t}\right)\theta_{i1} + \left(\frac{1}{\Omega} + \kappa \hat{b}_{1,t}\right)\frac{1-\theta_{i1}}{\Omega - 1}\right] \\ &+ \cdots + \ln\left[\left(\delta(O_m = 1) + \kappa \hat{b}_{m,t}\right)\theta_{im} + \left(\frac{1}{\Omega} + \kappa \hat{b}_{m,t}\right)\frac{1-\theta_{im}}{\Omega - 1}\right]. \end{aligned} \tag{C.3}$$

$$b_i(0) = c_i \varsigma^{m+1}\left[\frac{1}{\Omega} + \kappa \hat{b}_{i,t}\right]\prod_{j=1}^{m}\left[\left(\delta(O_j = 1) + \kappa \hat{b}_{j,t}\right)\frac{1-\theta_{ij}}{\Omega - 1} + \left(\frac{1}{\Omega} + \kappa \hat{b}_{j,t}\right)\theta_{ij}\right]. \tag{C.4}$$

$$\begin{aligned}\ln b_i(0) &= \ln c_i \varsigma^{m+1} + \ln\left[\frac{1}{\Omega} + \kappa \hat{b}_{i,t}\right] \\ &+ \ln\left[\left(\delta(O_1 = 1) + \kappa \hat{b}_{1,t}\right)\frac{1-\theta_{i1}}{\Omega - 1} + \left(\frac{1}{\Omega} + \kappa \hat{b}_{1,t}\right)\theta_{i1}\right] \\ &+ \cdots + \ln\left[\left(\delta(O_m = 1) + \kappa \hat{b}_{m,t}\right)\frac{1-\theta_{im}}{\Omega - 1} + \left(\frac{1}{\Omega} + \kappa \hat{b}_{m,t}\right)\theta_{im}\right]. \end{aligned} \tag{C.5}$$

Appendix D. The Derivation of Belief Threshold

From (37)-(40), we have $b_i(0) \cap b_i(1) \neq \emptyset$, then from (35) (36), we have $P_{f,i} \neq 0$ and $P_{d,i} \neq 1$ with probability one. We derive that $\forall\ (b_i > b_T|\mathcal{H}_0) \in b_i(0) \cap b_i(1)$, $\exists\ (b_i > b_T|\mathcal{H}_1) \in b_i(0) \cap b_i(1)$, and then we get $\lfloor b_i(1) \rfloor \leq b_T \leq \lceil b_i(0) \rceil$. Since $P_{f,i} \neq 1$, we can further obtain $\lfloor b_i(0) \rfloor \leq b_T \leq \lceil b_i(0) \rceil$.

In this case, the closed-form expression for belief threshold b_T can be calculated by

$$\frac{\lceil b_i(0) \rceil - b_T}{\lceil b_i(0) \rceil - \lfloor b_i(0) \rfloor} = P_{f,i}. \tag{D.1}$$

After some algebraic manipulations, we obtain (41) and the proof is completed.

Appendix E. Proof of Theorem 4

Define the weight sum of graph G as $\mathcal{S} = \sum_{d=1}^{\mathcal{U}} w_d$, then the weight complement of edge e_d is $w_d^c = \mathcal{S} - w_d$. By exploiting greedy algorithm, the prob-

lem can be transformed to minimize spanning tree of graph G^c, subject to $\forall e_d \in E, e_d \cup \mathbb{T}$ contains a loop, which is given by

$$\min_{w_d} \mathbb{T}^c = \sum_{e_d} (\mathcal{S} - \max_{d=1,\cdots,\mathcal{U}} w_d). \tag{E.1}$$

After some algebraic manipulations, we can get

$$\max_{w_d} \mathbb{T} = (M-1)\mathcal{S} - \min_{w_d} \mathbb{T}^c. \tag{E.2}$$

References

[1] N. Sastry, U. Shankar, D. Wagner, Secure verification of location claims, in: Proc. ACM Workshop on Wireless Security, 2003, pp. 1–10.

[2] S. Capkun, M. Cagalj, Integrity regions: Authentication through presence in wireless networks, IEEE Trans. Mobile Comput. 9 (11) (2010) 1–10.

[3] Y. S. Shiu, S. Y. Chang, H. C. Wu, C. H. Huang, H. H. Chen, Physical layer security in wireless networks: a tutorial, IEEE Wireless Commun. 18 (2) (2011) 66–74.

[4] Y. Chen, J. Yang, W. Trappe, R. P. Martin, Detecting and localizing identity-based attacks in wireless and sensor networks, IEEE Trans. Veh. Tech. 59 (5) (2010) 2418–2434.

[5] S. Yan, R. Malaney, I. Nevat, G. W. Peters, Optimal information-theoretic wireless location verification, IEEE Trans. Veh. Tech. 63 (7) (2014) 3410–3422.

[6] S. Yan, I. Nevat, G. W. Peters, R. Malaney, Location verification systems under spatially correlated shadowing, IEEE Trans. Wireless Commun. 15 (6) (2016) 4132–4144.

[7] S. Yan, R. Malaney, I. Nevat, G. W. Peters, Location verification systems for VANETs in rician fading channels, IEEE Trans. Veh. Tech. 65 (7) (2016) 5652–5664.

[8] J. K. Tugnait, Wireless user authentication via comparison of power spectral densities, IEEE J. Sel. Areas Commun. 31 (9) (2013) 1791–1802.

[9] H. Liu, Y. Wang, J. Liu, J. Yang, Y. Chen, Practical user authentication leveraging channel state information (CSI), in: Proc. ACM Symp. Inf., Comput. Commun. Security, 2014, pp. 389–400.

[10] L. Xiao, L. J. Greenstein, N. B. Mandayam, W. Trappe, Channel-based spoofing detection in frequency-selective rayleigh channels, IEEE Trans. Wireless Commun. 8 (12) (2009) 5948–5956.

[11] P. Baracca, N. Laurenti, S. Tomasin, Physical layer authentication over MIMO fading wiretap channels, IEEE Trans. Wireless Commun. 11 (7) (2012) 2564–2573.

[12] X. Li, Y. Chen, J. Yang, X. Zheng, Designing localization algorithms robust to signal strength attacks, in: Proc. IEEE INFOCOM, 2011, pp. 341–345.

[13] L. Xiao, X. Wan, Z. Han, PHY-layer authentication with multiple landmarks with reduced overhead, IEEE Trans. Wireless Commun. 17 (3) (2018) 1676–1687.

[14] L. Xiao, Y. Li, G. Han, G. Liu, W. Zhuang, PHY-layer spoofing detection with reinforcement learning in wireless networks, IEEE Trans. Veh. Tech. 65 (12) (2016) 10037–10047.

[15] L. Xiao, T. Chen, G. Han, W. Zhuang, L. Sun, Game theoretic study on channel-based authentication in MIMO systems, IEEE Trans. Veh. Tech. 66 (8) (2017) 7474–7484.

[16] N. Wang, T. Jiang, S. Lv, L. Xiao, A physical layer authentication based on extreme learning machine, IEEE Commun. Lett. 21 (7) (2017) 1557–1560.

[17] N. Nedic, D. Prsic, C. Fragassa, V. Stojanovic, A. Pavlovic, Simulation of hydraulic check valve for forestry equipment, International Journal of Heavy Vehicle Systems 24 (3) (2017) 260–276.

[18] D. Prsic, N. Nedic, V. Stojanovic, A nature inspired optimal control of pneumatic-driven parallel robot platform, Proceedings of the Institution of Mechanical Engineers Part C Journal of Mechanical Engineering Science 231 (1) (2016) 59–71.

[19] N. Nedic, D. Prsic, L. Dubonjic, V. Stojanovic, V. Djordjevic, Optimal cascade hydraulic control for a parallel robot platform by PSO, International Journal of Advanced Manufacturing Technology 72 (5-8) (2014) 1085–1098.

[20] L. Xiao, X. Wan, X. Lu, Y. Zhang, D. Wu, IoT security techniques based on machine learning: How do IoT devices use AI to enhance security?, IEEE Signal Process. Mag. 35 (5) (2018) 41–49.

[21] J. Yan, H. He, X. Zhong, Y. Tang, Q-learning-based vulnerability analysis of smart grid against sequential topology attacks, IEEE Trans. Inf. Forensics Security 12 (1) (2017) 200–210.

[22] C. J. Watkins, P. Dayan, Q-learning, Machine Learning 8 (3-4) (1992) 279–292.

[23] T. Wang, Y. Yang, Analysis on perfect location spoofing attacks using beamforming, in: Proc. IEEE INFOCOM, 2013, pp. 2778–2786.

[24] H. Ye, G. Y. Li, Deep reinforcement learning for resource allocation in V2V communications, in: Proc. IEEE ICC, 2018, pp. 1–6.

[25] Z. Qin, Y. Gao, M. D. Plumbley, C. G. Parini, Wideband spectrum sensing on real-time signals at sub-nyquist sampling rates in single and cooperative multiple nodes, IEEE Trans. Signal Process. 64 (12) (2016) 3106–3117.

[26] Z. Qin, Y. Gao, M. D. Plumbley, Malicious user detection based on low-rank matrix completion in wideband spectrum sensing, IEEE Trans. Signal Process. 66 (1) (2018) 5–17.

[27] Z. Qin, J. Fan, Y. Liu, Y. Gao, G. Y. Li, Sparse representation for wireless communications: A compressive sensing approach, IEEE Signal Process. Mag. 35 (3) (2018) 40–58.

[28] S. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice Hall, 1993.

[29] H. S. Wang, N. Moayeri, Finite-state markov channel a useful model for radio communication channels, IEEE Trans. Veh. Tech. 44 (1) (1995) 163–171.

[30] L. Xiao, Y. Li, C. Dai, H. Dai, H. V. Poor, Reinforcement learning-based NOMA power allocation in the presence of smart jamming, IEEE Trans. Veh. Tech. 67 (4) (2018) 3377–3389.

[31] E. Altman, K. Avrachenkov, A. Garnaev, Jamming in wireless networks under uncertainty, Mobile Netw. Appl. 16 (2) (2011) 246–254.

[32] T. M. Cover, J. A. Thomas, Elements of Information Theory, New York: Wiley, 2006.

[33] Mooij, M. Joris, Kappen, J. Hubert, Sufficient conditions for convergence of the sum-product algorithm, IEEE Trans. Inf. Theory 53 (12) (2007) 4422–4437.

[34] Z. Yin, R. T. Collins, Belief propagation in a 3D spatio-temporal mrf for moving object detection, in: Proc. IEEE CVPR, 2007, pp. 1–8.

[35] S. Srkk, Bayesian Filtering and Smoothing, Cambridge University Press, 2013.

[36] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference., Morgan Kaufmann, 1988.

[37] M. H. Eiza, Q. Ni, An evolving graph-based reliable routing scheme for VANETs, IEEE Trans. Veh. Tech. 62 (4) (2013) 1493–1504.

[38] J. A. Bondy, U. S. R. Murty, Graph Theory with Applications, Vol. 290, Macmillan London, 1976.

[39] V. Erceg, L. J. Greenstein, S. Y. Tjandra, S. R. Parkoff, A. Gupta, B. Kulic, A. A. Julius, R. Jastrzab, An empirically-based path loss model for wireless channels in suburban environments, IEEE J. Sel. Areas Commun. 17 (7) (2002) 1205–1211.

[40] V. Stojanovic, N. Nedic, D. Prsic, L. Dubonjic, Optimal experiment design for identification of ARX models with constrained output in non-gaussian noise, Applied Mathematical Modelling 40 (13-14) (2016) 6676–6689.

[41] V. Stojanovic, N. Nedic, Robust kalman filtering for nonlinear multivariable stochastic systems in the presence of non-gaussian noise, International Journal of Robust and Nonlinear Control 26 (2) (2015) 445–460.

[42] J. Dieudonné, Foundations of Modern Analysis, Eds. New York: Academic Press, 1969.