PHY-Layer Authentication with Multiple Landmarks with Reduced Communication Overhead

Xiaoyue Wan*, Liang Xiao*, Qiangda Li*, Zhu Han†

*Dept. of Communication Engineering, Xiamen University, Xiamen, China. Email: lxiao@xmu.edu.cn
†Dept. of Electrical and Computer Engineering, University of Houston, Houston, TX. Email: zhan2@uh.edu

***Abstract*—In this paper, we propose a physical (PHY)-layer authentication system that exploits the channel state information of radio transmitters to detect spoofing attacks in wireless networks. By using multiple landmarks and multiple antennas in the channel estimation, this authentication system enhances the spatial resolution of the channel information and thus improves the spoofing detection accuracy. Unlike most existing hypothesis test based PHY-layer authentication schemes that rely on the known radio channel model, our proposed authentication system uses the logistic regression to remove the assumption on the known channel model and is applicable to more generic wireless networks. The Frank-Wolfe algorithm is then used to estimate the parameters of the logistic regression model, which solves the convex problem under a ℓ_1-norm constraint for weight sparsity to avoid over-fitting in the learning process. The distributed Frank-Wolfe algorithm can further reduce the communication overhead between the landmarks and the security agent while keeping the spoofing detection accuracy. Simulation results can validate the accuracy of the proposed PHY-layer authentication with multiple landmarks and show the performance gain regarding the overall communication overhead.**

***Index Terms*—PHY-layer authentication, learning, distributed Frank-Wolfe algorithm, spoofing detection, multiple landmarks.**

I. Introduction

Wireless networks are vulnerable to spoofing attacks, in which an attacker Eve impersonates another user Alice by claiming her higher-layer identity such as her media access control (MAC) address, which cannot be fully addressed by the traditional higher-layer authentication techniques. Therefore, various physical (PHY)-layer authentication techniques have been proposed to exploit the spatial decorrelation property of the PHY-layer radio channel information, such as received signal strength indicators (RSSIs) [1]–[4], received signal strength (RSS) [5], channel impulse responses (CIR) [6], and channel state information (CSI) [7], [8], to distinguish radio transmitters and thus detect spoofing attacks with low overhead. As the channel estimation by multiple landmarks each with multiple antennas at different locations increases the spatial resolution of the radio channel regarding the radio transmitter, a PHY-layer authentication system can utilize multiple landmarks each equipped with multiple antennas to improve the spoofing detection accuracy.

Most existing PHY-layer authentication systems are based on the hypothesis tests and assume some specific radio channel models, e.g., the frequency-selective Rayleigh channel model as assumed in [9]. However, the assumed channel model is not always accurate, especially in dynamic wireless networks in which the actual channel model is complicated and change over time. Therefore, the logistic regression technique [10] can be used to discriminate radio transmitters according to their channel information and remove the assumptions on the radio channel distribution. In other words, with coefficients being estimated by the maximum log-likelihood (ML) estimation over the training channel data, the logistic regression based authentication is applicable to a broader range of wireless networks.

The ML-based coefficient estimation of the regression can be formulated as a convex optimization problem, and thus can be efficiently addressed by the Frank-Wolfe (FW) algorithm, which uses linear subproblems to approximate the convex and constrained optimization problem at each iteration [11]. The authentication system based on FW requires the transmission of the channel information from each landmark to the agent that can be either the legal receiver Bob or the special security agent. The resulting communication overhead is forbidden, if the number of the landmarks and the antennas are large. Therefore, the PHY-layer authentication can apply the distributed Frank-Wolfe (dFW) algorithm, which performs the FW updates and converges faster with a small amount of computational overhead at each iteration and a lower communication cost, to solve the convex constraint optimization problem in a distributed system.

In this paper, we propose a logistic regression based PHY-layer authentication, which exploits the RSSIs collected at multiple antennas by multiple landmarks to detect spoofing attacks, and estimates the coefficients of the regression model with the FW algorithm. The communication costs of the PHY-layer authentication is further reduced with dFW. Simulations based on the stochastic channel model are performed to verify the spoofing detection performance of our proposed systems. The main contributions of our work can be summarized as follows:

1) We present a PHY-layer authentication system that ex-

This work was supported in part by National Natural Science Foundation of China under Grant 61671396, CCF-Venustech Hongyan Research Initiative (2016-010), and the US NSF CNS-1646607, ECCS-1547201, CCF-1456921, CNS-1443917, and ECCS-1405121.

ploits both multiple landmarks and multiple antennas to improve the spatial resolution of the transmitter, and thus enhance the spoofing detection accuracy. By applying the FW algorithm to estimate the coefficient of the logistic regression model, this authentication system dose not rely on the knowledge of the radio channel model, and thus is applicable to more generic networks without over-fitting the training channel data.

2) By replacing the FW algorithm with the distributed FW algorithm to build the logistic regression, the authentication system reduces the overall communication overhead between the landmarks and the security agent and provides a similar detection accuracy.

The rest of this paper is organized as follows. We review the related work in Section II and describe the system model in Section III. We present the PHY-layer authentication based on FW in Section IV and the dFW-based detection in Section V. We present simulation results in Section VI and conclude in Section VII.

II. Related Work

The channel-based spoofing detection in the single antenna system developed in [9] uses the generalized likelihood ratio test and is applicable to the frequency-selective Rayleigh radio channel model. The RSS-based authentication system proposed in [5] applies the cluster analysis method and multiple landmarks to improve the spoofing detection accuracy. An energy ratio based spoofing detection system developed in [12] uses multiple antennas to improve the detection accuracy by exploring the asymmetry of received signal power levels at the transmitter and the legitimate receiver if there exists a pilot spoofing attack. In this paper, we apply both multiple antennas and multiple landmarks to improve the detection performance. The spoofing detection proposed in [13] applies the Q-leaning technique to achieve the optimal test threshold in the hypothesis test for the receiver that is unaware of the radio channel model, but converges much more slower than this work.

III. System Model

As shown in Fig. 1, we consider the spoofing detection in the Alice-Bob-Eve model, assisted by M landmarks (or radio monitors) each equipped with N antennas. Alice as the legal transmitter sends signals with a single antenna, and Eve impersonates Alice by claiming her MAC address. Bob who is either a legitimate receiver or security agent verifies whether the received message under test is indeed sent by Alice. More specifically, upon receiving the i-th message, landmark m estimates the received signal strength indicator of the message at antenna j, denoted by h^i_{mj}, with $1 \leq m \leq M$, and $1 \leq j \leq N$. Each landmark can process the channel information and send some information to Bob without interfering with the transmissions of Alice, if needed.

Although our proposed authentication system does not rely on any specific channel model, for evaluation purposes, the simulations in Section VI assume that the path loss of message

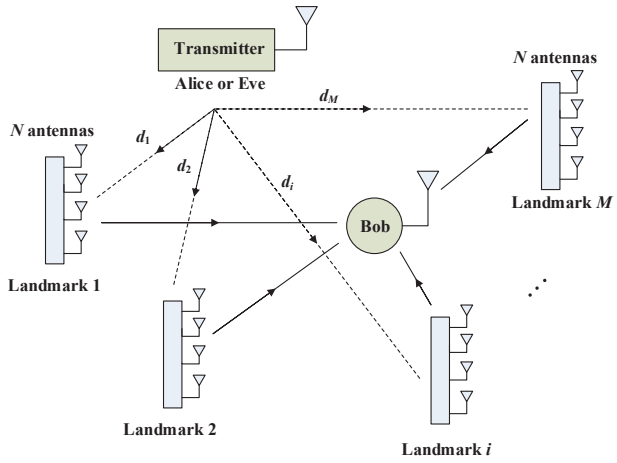


Fig. 1. Illustration of a PHY-layer authentication system consisting of a transmitter that is either Alice or Eve, and a security agent Bob, based on the channel information estimated by M landmarks each with N antennas.

i between the transmitter and antenna j on landmark m, denoted by L^i_{mj}, follows the generic model as specified in [14],

$$L^i_{mj}[dB] = A + 10\gamma \log_{10}(d_m/d_0) + s^i_{mj}, \qquad (1)$$

where d_m is the distance between landmark m and the transmitter, s^i_{mj} is the shadow fading (in dB) along that path, d_0 is a reference distance, the intercept A is the decibel path loss at distance d_0, γ is path loss exponent and s^i_{mj} is a Gaussian random variable with zero mean and standard deviation σ. The channel estimation error is assumed to be zero-mean, complex Gaussian distributed and independent across frequencies and receive antennas, with a common variance denoted by σ_I^2. Important symbols and notations are summarized in Table I.

IV. PHY-layer Authentication with FW

The PHY-layer authentication system exploits the spatial decorrelation of the radio channel information with M landmarks each equipped with N antennas to authenticate a message sent by the radio transmitter under test that is either Alice or Eve. The MN-dimensional column channel vector for the i-th message received by Bob denoted by $\mathbf{H}_i$ consists of the RSSIs from all the M landmarks, with $\mathbf{H}_i = [h^i_{mj}]_{MN}$. Let the spoofing detection result denoted by y_i represent the actual transmitter of the message, i.e., Alice or Eve. If the signal is sent by Alice, we set $y_i = 1$; otherwise if the message is a spoofing signal sent by Eve, $y_i = 0$.

Let an MN-dimensional row vector $\boldsymbol{\beta}$ indicate the importance of each feature in channel vector $\mathbf{H}_i$ and be the coefficients of the logistic regression model. The logistic regression presents the predicted probability for each class with the softmax function, which is the gradient-log-normalizer of the categorical probability distribution. In the authentication of message i, the predicted probability for $y_i = 0$ or $y_i = 1$ given

TABLE I
SUMMARY OF SYMBOLS AND NOTATIONS

M	Number of landmarks
N	Number of receive antennas
T	Number of training data in the regression
h_{mj}^i	RSSI at antenna j on landmark m
$\mathbf{H}_i^m$	Local channel vector of message i on landmark m
$\mathbf{H}_i$	Channel vector of message i
L_{mj}^i	Path loss of the message i at antenna j on landmark m
$\boldsymbol{\beta}/\hat{\boldsymbol{\beta}}$	Actual/estimated coefficient vector in the regression model
$\hat{\boldsymbol{\beta}}_m$	Local estimated coefficient vector of landmark m
$y_i/\hat{y}_i$	Actual/estimated transmitter of the message i
χ	Flag to stop the iteration
ψ_m	Index of the largest local gradient of landmark m
S_m	Partial sum for stopping criterion of landmark m
ρ	Node index with the largest overall gradient
ψ	Index of the largest overall gradient
ϵ	Approximation quality

the channel vector $\mathbf{H}_i$ is presented with the softmax function as described in [10] as follows

$$\Pr(y_i = 1|\mathbf{H}_i) = \frac{e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_i}}{1+e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_i}}, \tag{2}$$

$$\Pr(y_i = 0|\mathbf{H}_i) = \frac{1}{1+e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_i}}. \tag{3}$$

If $\Pr(y_i = 1|\mathbf{H}_i) > \Pr(y_i = 0|\mathbf{H}_i)$, we have $\beta_0 + \boldsymbol{\beta}\mathbf{H}_i > 0$ by (2) and (3) after simplification, and assume that message i is sent by Alice with $\hat{y}_i = 1$. Otherwise, message i is assumed to be sent by Eve with $\hat{y}_i = 0$ if $\beta_0 + \boldsymbol{\beta}\mathbf{H}_i \leq 0$, i.e.,

$$\hat{y}_i = \begin{cases} 1, & \text{if } \beta_0 + \boldsymbol{\beta}\mathbf{H}_i > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Without knowing the distribution of the RSSI vector, Bob can use the logistic regression technique to authenticate message i according to channel vector $\mathbf{H}_i$, and apply the previous T channel vectors, (i.e., $\{\mathbf{H}_k\}_{i-T-1\leq k\leq i-1}$) as the training data to estimate the coefficients of the logistic regression model. More specifically, the minus log-likelihood of the T training data denoted by $f(\boldsymbol{\beta})$ is given by

$$\begin{aligned} f(\boldsymbol{\beta}) &= -\ln\left(\prod_{k=i-T-1}^{i-1} \Pr(y_k|\mathbf{H}_k)\right) \\ &= -\sum_{k=i-1-T}^{i-1} \left(y_k\left(\beta_0+\boldsymbol{\beta}\mathbf{H}_k\right) - \ln\left(1+e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_k}\right)\right). \end{aligned} \tag{5}$$

The coefficient vector $\boldsymbol{\beta}$ is estimated by the maximum log-likelihood method following the ℓ_1-norm regularization. The cost function $f(\boldsymbol{\beta})$ given by (5) is minimized under the ℓ_1-norm penalty to restrict the complexity of the logistic regression model. Thus, the estimated coefficient vector denoted by $\hat{\boldsymbol{\beta}}$ is given by

$$\begin{aligned} \hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} & \sum_{k=i-1-T}^{i-1} \left(\ln(1+e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_k}) - y_k(\beta_0+\boldsymbol{\beta}\mathbf{H}_k)\right) \\ s.t. \quad & ||\hat{\boldsymbol{\beta}}||_1 \leq C, \end{aligned} \tag{9}$$

where the positive parameter C restricts the model complexity and makes a tradeoff between the training error and the testing error of the regression algorithm.

We do not penalize the intercept term β_0 and standardize the channel vectors to make a meaningful penalty. By [10], β_0 is estimated by

$$\hat{\beta}_0 = \frac{1}{T}\sum_{k=i-T-1}^{i-1} y_k. \tag{10}$$

The Frank-Wolfe algorithm described in [11] is used to estimate $\boldsymbol{\beta}$, as the objective function $f(\boldsymbol{\beta})$ in (5) is convex and continuously differentiable with respect to $\boldsymbol{\beta}$. By (5), the gradient $\nabla f(\boldsymbol{\beta})$ as an MN-dimensional vector consists of the local gradient of landmark m, with the j-th element given by

$$\nabla f(\boldsymbol{\beta})_j = \sum_{k=i-T-1}^{i-1}\left(\frac{e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_k}H_{kj}}{1+e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_k}} - y_k H_{kj}\right), \tag{11}$$

where H_{kj} is the j-th element of channel vector $\mathbf{H}_k$ and $1 \leq j \leq MN$. The largest absolute value of the overall gradient $\nabla f(\boldsymbol{\beta})$ is chosen as the optimal descent direction of the cost function. The corresponding index of the direction denoted by ψ is given by

$$\psi = \arg\max_{1\leq j\leq MN} |\nabla f(\boldsymbol{\beta})_j|. \tag{12}$$

The optimal descent direction of the cost function in terms of $\boldsymbol{\beta}$ is given by $\text{sgn}(-\nabla f(\boldsymbol{\beta})_\psi)C\mathbf{e}^\psi - \boldsymbol{\beta}$. The iteration step size at iteration n denoted by γ is set to be $\gamma = \frac{2}{n+2}$ by [11]. Therefore, the subsequent iteration point after moving step size γ along the optimal the descent direction at the current iteration point is updated with

$$\boldsymbol{\beta} \leftarrow (1-\gamma)\boldsymbol{\beta} + \gamma\text{sgn}\left(-\nabla f(\boldsymbol{\beta})_\psi\right)C\mathbf{e}^\psi, \tag{13}$$

where $\mathbf{e}^\psi$ is the all zeros MN-dimensional vector except 1 at the ψ-th entry. The flag to stop the iteration is denoted by χ and given by

$$\chi = \sum_{j=1}^{MN}\beta_j\nabla f(\boldsymbol{\beta})_j + C\left|\nabla f(\boldsymbol{\beta})_\psi\right|. \tag{14}$$

The approximation quality of the above FW algorithm to solve the convex problem in (9), denoted by ϵ, means that the estimated $\hat{\boldsymbol{\beta}}$ is at least ϵ-close to the optimal $\boldsymbol{\beta}$. If $\chi < \epsilon$, Bob obtains the ϵ-approximate of $\boldsymbol{\beta}$ and then uses (4) to detect message i. By (4), if $\hat{y}_i = 1$ Bob receives message i. Otherwise, Bob sends the spoofing alarm, if $\hat{y}_i = 0$.

In summary, the FW-based authentication calculates the local gradient of each landmark to find the largest overall

Proof. The derivation of the equation (5) is as follows

$$
\begin{aligned}
f(\boldsymbol{\beta}) &= -\ln\left(\prod_{k=i-T-1}^{i-1}\Pr(y_k|\mathbf{H}_k)\right) \\
&= -\ln\left(\prod_{\substack{k=i-T-1\\ y_k=1}}^{i-1}\Pr(y_k=1|\mathbf{H}_k)\prod_{\substack{k=i-T-1\\ y_k=0}}^{i-1}\Pr(y_k=0|\mathbf{H}_k)\right) \qquad (6)\\
&= -\sum_{k=i-T-1}^{i-1}\left(y_k\ln\Pr(y_k=1|\mathbf{H}_k)+(1-y_k)\ln\Pr(y_k=0|\mathbf{H}_k)\right) \qquad (7)\\
&= -\sum_{k=i-1-T}^{i-1}\left(y_k\ln\frac{e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_k}}{1+e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_k}}+(1-y_k)\ln\frac{1}{1+e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_k}}\right) \qquad (8)\\
&= -\sum_{k=i-1-T}^{i-1}\left(y_k(\beta_0+\boldsymbol{\beta}\mathbf{H}_k)-\ln\left(1+e^{\beta_0+\boldsymbol{\beta}\mathbf{H}_k}\right)\right),
\end{aligned}
$$

where (6) is obtained by (4), and (8) is derived by (2) and (3). □

gradient of the cost function of the training data and determine the subsequent iteration point to estimate $\boldsymbol{\beta}$ in the logistic regression model. Therefore, the resulting communication overhead between the security agent Bob and the M landmarks depends on the landmark topology and can be prohibitive under a large number of landmarks and antennas. The PHY-layer authentication with FW is summarized in Algorithm 1.

Algorithm 1 PHY-layer Authentication based on FW

1: **Initialize**: $\boldsymbol{\beta}=\mathbf{0}, n=0, \chi, \epsilon$, and C.
2: Channel estimation to form $\mathbf{H}_i$
3: Calculate $\hat{\beta}_0$ via (10)
4: **While** $\chi > \epsilon$ **do**
5: Compute ψ via (12)
6: Set $\gamma = \frac{2}{n+2}$
7: Update $\boldsymbol{\beta}$ via (13)
8: Compute χ via (14)
9: $n \leftarrow n+1$
10: **End while**
11: Obtain $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}$
12: Calculate $\hat{y}_i$ via (4)
13: **If** $\hat{y}_i = 1$ **then**
14: Accept message i
15: **Else**
16: Send spoofing alarm for message i
17: **End if**

V. PHY-layer Authentication Based on dFW

In order to reduce the communication overhead of the FW-based authentication, the PHY-layer authentication can apply the dFW algorithm as described in [15] to solve the convex optimization problem with the ℓ_1-norm regularization in (9). As an advanced FW algorithm designed for distributed systems, the dFW-based authentication requires each landmark to send a few channel information instead of the entire channel vector. More specifically, we use $\mathcal{A}_m = \{(m-1)N+1, \cdots, mN\}$, i.e., the set of the column indices in both $\nabla f(\boldsymbol{\beta})$ and $\boldsymbol{\beta}$ that associated with landmark m. Thus the local gradient of landmark m is given by the elements in $\nabla f(\boldsymbol{\beta})$ with indices belonging to $\mathcal{A}_m$. At iteration n, similar to (12) in FW, landmark m identifies the largest component of its local gradient in absolute value, whose indice in the gradient is denoted by ψ_m and given by

$$\psi_m = \arg\max_{j\in\mathcal{A}_m} |\nabla f(\boldsymbol{\beta})_j|. \qquad (15)$$

Landmark m calculates the local flag to stop the iteration denoted by S_m and given by [15] as

$$S_m = \sum_{j\in\mathcal{A}_m} \beta_j \nabla f(\boldsymbol{\beta})_j. \qquad (16)$$

Then the information set, $\{\nabla f(\boldsymbol{\beta})_{\psi_m}, \psi_m, S_m\}$ is sent by landmark m to Bob. Upon receiving the information set from all the M landmarks, Bob identifies the landmark with the largest overall gradient, whose index is denoted by ρ and given by

$$\rho = \arg\max_{m\in[1,M]} |\nabla f(\boldsymbol{\beta})_{\psi_m}|. \qquad (17)$$

Bob computes the flag to stop the iteration χ by

$$\chi = \sum_{m=1}^{M} S_m + C|\nabla f(\boldsymbol{\beta})_{\psi_\rho}|. \qquad (18)$$

If $\chi > \epsilon$, Bob broadcasts $\{\psi_\rho, \nabla f(\boldsymbol{\beta})_{\psi_\rho}\}$ to the M landmarks, and then each landmark updates the coefficient vector accordingly by

$$\boldsymbol{\beta} \leftarrow (1-\gamma)\boldsymbol{\beta} + \gamma \text{sgn}(-\nabla f(\boldsymbol{\beta})_{\psi_\rho}) C\mathbf{e}^{\psi_\rho}. \qquad (19)$$

Each landmark also determines whether to broadcast its local channel information to other landmarks. For simplicity, let

Algorithm 2 Procedure at landmark m

Initialize: $\boldsymbol{\beta} = \mathbf{0}$, $n = 0$, flag=0, χ, and C.
Channel estimation to form $\mathbf{H}_m^i$
While flag=0 **do**
  Compute $\nabla f(\boldsymbol{\beta})_{j\in\mathcal{A}_m}$ via (11)
  Compute ψ_m via (15)
  Compute S_m via (16)
  Send $\{\nabla f(\boldsymbol{\beta})_{\psi_m}, \psi_m, S_m\}$ to Bob
  Receive $\{\psi_\rho, \nabla f(\boldsymbol{\beta})_{\psi_\rho}\}$ from Bob
  if $\psi_m = \psi_\rho$ **do**
    Broadcast the ψ_m-th column of $\mathbf{H}$
  End if
  Set $\gamma = \frac{2}{n+2}$
  Update $\boldsymbol{\beta}$ via (19)
  $n \leftarrow n + 1$
End while
Send $\hat{\boldsymbol{\beta}}_m \mathbf{H}_i^m$ to Bob

the $T \times MN$ matrix $\mathbf{H} = [\mathbf{H}_k]_{i-T-1\le k\le i-1}$ represent the training channel data matrix of message i. If finding $\psi_m = \psi_\rho$, landmark m broadcasts the ψ_m-th column of $\mathbf{H}$, which is used by the other $M-1$ landmarks to update their local gradients in the next iteration.

Otherwise, if $\chi \le \epsilon$, Bob stops the iteration and notifies all the landmarks about it. At this time, landmark m obtains the estimated $\hat{\boldsymbol{\beta}}$. Landmark m the determines $\hat{\boldsymbol{\beta}}_m$, which is the locally estimated coefficient vector with $\hat{\boldsymbol{\beta}}_m = \{\hat{\beta}_j\}_{j\in\mathcal{A}_m}$. Let the N-dimensional vector $\mathbf{H}_i^m$ denote the local channel vector of message i at landmark m, which sends $\hat{\boldsymbol{\beta}}_m \mathbf{H}_i^m$ to Bob. Bob classifies message i with

$$\hat{y}_k = \begin{cases} 1, & \text{if } \hat{\beta}_0 + \sum_{m=1}^{M} \hat{\boldsymbol{\beta}}_m \mathbf{H}_i^m > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

The procedure of landmark m and that of Bob are summarized in Algorithms 2 and 3, respectively.

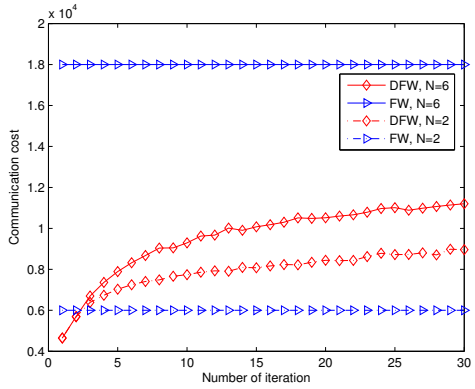
VI. Simulation Results

Simulations are performed based on the channel model as shown in (1) to evaluate the spoofing performance for the proposed PHY-layer authentication system with M landmarks. The false alarm rate of the spoofing detection denoted by P_f is the probability of rejecting a legal message, i.e., $P_f = \Pr(\hat{y}_i = 0 | y_i = 1)$. The miss detection rate denoted by P_m is the probability to accept a spoofing massage, i.e., $P_m = \Pr(\hat{y}_i = 1 | y_i = 0)$. In the simulations, we set $A = 20\log_{10}(4\pi d_0/\lambda)$, where λ is wavelength and $d_0 = 100$m, $\tau = 4, \sigma = 0.2$dB. If not specified otherwise, we set $M = 6$, $N = 6$, $T = 100$, $C = 10$, and $\epsilon = 0.1$.

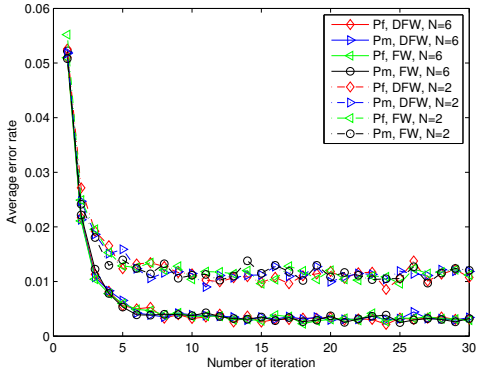
As shown in Fig. 2(a), the dFW-based authentication reduces the communication overhead compared with the FW algorithm, e.g., the overhead is reduced by 44.0% at iteration 15 with $N = 6$, because only a small portion of the data, $\{\nabla f(\boldsymbol{\beta})_{\psi_m}, \psi_m, S_m, \psi_\rho, \nabla f(\boldsymbol{\beta})_{\psi_\rho}\}$ are transmitted between Bob and the landmarks at each iteration and dFW requires a number of iterations as small as 15. The communication cost for FW is a constant with iteration because landmarks send all

Algorithm 3 PHY-layer authentication with dFW

Initialize: flag=0, χ, ϵ, and C.
Calculate $\hat{\beta}_0$ via (10)
While $\chi > \varepsilon$ **do**
  Receive $\{\nabla f(\boldsymbol{\beta})_{\psi_m}, \psi_m, S_m\}$ from landmark m.
  Compute ρ via (17)
  Broadcast $\{\psi_\rho, \nabla f(\boldsymbol{\beta})_{\psi_\rho}\}$ to each landmark
  Compute χ via (18)
End while
Set flag=1 and notify the M landmarks
Receive $\hat{\boldsymbol{\beta}}_m \mathbf{H}_i^m$ from the M landmarks
Compute $\hat{\beta}_0 + \sum_{m=1}^{M} \hat{\boldsymbol{\beta}}_m \mathbf{H}_i^m$
Estimate $\hat{y}_i$ via (20)
If $\hat{y}_i = 1$ **then**
  Accept message i
Else
  Send spoofing alarm for message i
End if



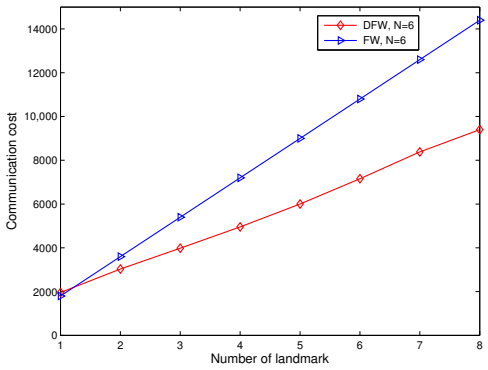
(a) Communication cost



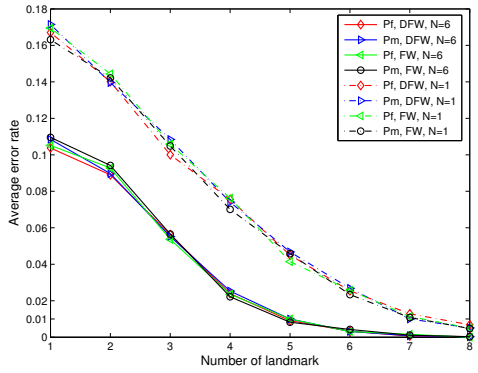
(b) Average error rate

Fig. 2. Comparison between dFW and FW in authentication performance and communication cost with $M = 6$, $T = 100$, and $C = 10$ in the simulation with channel vector generated by (1).

the collected RSSIs to Bob before iteration. As shown in Fig. 2(b), both authentication schemes provide accurate spoofing detection, e.g., their average error rate is less than 3% at iteration 15.



(a) Communication cost



(b) Average error rate

Fig. 3. Spoofing detection performance of the PHY-layer authentication system with N antennas at each landmark with $C = 10$, $T = 100$, and $\epsilon = 0.1$.

As shown in Fig. 3, the communication overheads of both authentication schemes increase with the number of landmarks and the performance gain of dFW over FW regarding the overhead increases with M. For example, the performance gain increases from 15.8% to 34.7%, if M changes from 2 to 8 and $N = 6$. The detection accuracy of the dFW-based authentication decreases with both the number of landmarks and antennas, e.g., P_f and P_m reduce by 89.8% and 88.9%, respectively, as M changes from 2 to 5, if $N = 6$. In another example, P_f and P_m reduce by 95.6% and 96.1%, as N changes from 1 to 6, if $M = 8$.

VII. Conclusions

In this paper, we proposed a PHY-layer authentication system that applies multiple landmarks each equipped with multiple antennas to estimate the received signal strength of the message to improve the spoofing detection accuracy. The logistic regression based authentication does not being restricted to the known radio channel model. The FW algorithm estimates the coefficients of the logistic regression model by maximizing the posterior probability under the ℓ_1-norm type constraint to reduce the computational complexity and avoid over-fitting in the learning process. The overall communication overhead between the landmarks and the security agent is further reduced by applying the distributed FW algorithm to replace FW. Simulation results show that the proposed authentication system can improve the spoofing detection accuracy, e.g., both the false alarm rate and miss detection rate are 95% less than the benchmark system that locates 2 landmarks equipped with single antenna, if using 8 landmarks each equipped with 6 antennas. Moreover, the communication cost is further reduced by the dFW algorithm, e.g., the dFW based system with 6 landmarks and 6 antennas takes only 66.2% of the communication overhead.

References

[1] V. Bhargava and M. L. Sichitiu, "Physical authentication through localization in wireless local area networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, vol. 5, pp. 2658–2662, St. Louis, MO, Dec. 2005.

[2] L. Xiao, Q. Yan, W. Lou, and G. Chen, "Proximity-based security techniques for mobile users in wireless networks," *IEEE Trans. Information Forensics and Security*, vol. 8, no. 12, pp. 2089–2100, Dec. 2013.

[3] J. Liu, L. Xiao, G. Liu, and Y. Zhao, "Active authentication with reinforcement learning based on ambient radio signals," *Springer Multimedia Tools and Applications*, pp. 2–22, Oct. 2015.

[4] L. Xiao, T. Chen, G. Han, W. Zhuang, and L. Sun, "PHY-game theoretic study on channel-based authentication in MIMO systems," *IEEE Trans. Vehicular Technology*, Jan. 2017.

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

[6] F. J. Liu, X. Wang, and S. L. Primak, "A two dimensional quantization algorithm for cir-based physical layer authentication," in *Proc. IEEE Int'l Conf. Commun. (ICC)*, pp. 4724–4728, Budapest, Jun. 2013.

[7] H. Liu, Y. Wang, J. Liu, J. Yang, and Y. Chen, "Practical user authentication leveraging channel state information (CSI)," in *Proc. ACM Symp. Inform., Computer and Commun. Security*, pp. 389–400, Kyoto, Jun. 2014.

[8] L. Xiao, Y. Li, G. Han, G. Liu, and W. Zhuang, "PHY-layer spoofing detection with reinforcement learning in wireless networks," *IEEE Trans. Vehicular Technology*, vol. 65, no. 12, pp. 10037–10047, Dec. 2016.

[9] L. Xiao, L. J. Greenstein, N. B. Mandayam, and W. Trappe, "Channel-based spoofing detection in frequency-selective Rayleigh channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5948–5956, Jan. 2009.

[10] T. J. Hastie, R. J. Tibshirani, and J. H. Friedman, *The elements of statistical learning: Data mining, inference, and prediction*. Springer, 2011.

[11] M. Jaggi, "Revisiting Frank-Wolfe: Projection-free sparse convex optimization," in *Proc. Int'l Conf. Machine Learning (ICML)*, pp. 427–435, Atlanta, Jun. 2013.

[12] Q. Xiong, Y. Liang, K. H. Li, and Y. Gong, "An energy-ratio-based approach for detecting pilot spooing attack in multiple-antenna systems," *IEEE Trans. Information Forensics and Security*, vol. 10, no. 5, pp. 932–940, May 2015.

[13] L. Xiao, Y. Li, G. Liu, Q. Li, and W. Zhuang, "Spoofing detection with reinforcement learning in wireless networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Diego, CA, Dec. 2015.

[14] V. Erceg, L. J. Greenstein, and et al., "An empirically based path loss model for wireless channels in suburban environments," *IEEE J. Select. Areas Commun.*, vol. 17, no. 7, pp. 1205–1211, Jul. 1999.

[15] A. Bellet, Y. Liang, and et al., "A distributed Frank-Wolfe algorithm for communication-efficient sparse learning," in *Proc. SIAM Int'l Conf. Data Mining*, pp. 478–486, Vancouver, Apr. 2015.