Indirect Reciprocity Security Game for Large-Scale Wireless Networks

Liang Xiao, *Member, IEEE*, Yan Chen, *Student Member, IEEE*, W. Sabrina Lin, *Student Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

***Abstract*—Radio nodes can obtain illegal security gains by performing attacks, and they are motivated to do so if the illegal gains are larger than the resulting costs. Most existing direct reciprocity-based works assume constant interaction among players, which does not always hold in large-scale networks. In this paper, we propose a security system that applies the indirect reciprocity principle to combat attacks in wireless networks. Because network access is highly desirable for most nodes, including potential attackers, our system punishes attackers by stopping their network services. With a properly designed social norm and reputation updating process, the aim is to incur a cost due to the loss of network access to exceed the illegal security gain. Thus rational nodes are motivated to abandon adversary behavior for their own interests. We derive the optimal strategy and the corresponding stationary reputation distribution, and evaluate the stability condition of the optimal strategy using the evolutionarily stable strategy concept. This security system is robust against collusion attacks and can significantly reduce the attacker population for a wide range of attacks when the stability condition is satisfied. Simulation results show that the proposed system significantly outperforms the existing direct reciprocity-based systems, especially in the large-scale networks with terminal mobility. This technique can be extended to many wireless networks, including cognitive radio networks, to improve their security performance.**

***Index Terms*—Collusion attacks, evolutionarily stable strategy (ESS), indirect reciprocity, large-scale network, wireless security.**

I. Introduction

With the development of cognitive radios [1], users are gaining autonomy and control in their radio transmissions, and thus are capable of causing great damage to the networks. Since users are rational and thus naturally selfish, they will make any efforts including launching attacks to maximize their payoffs. Current wireless networks are threatened by a wide range of attacks, such as jamming [2], spoofing [3], Sybil attacks [4], [5], and other relay-related attacks [6], [7]. Attackers can obtain illegal security advantages, if not being caught and punished. Extensive works have been done to investigate the impacts of attacks on the network performance, and many detection and localization algorithms have been proposed to identify the adversaries [7]. On the other hand, we notice that most potential attackers also desire better network services, and this fact can be exploited to prevent nodes from attacking wireless networks. For instance, rational nodes would hesitate to conduct adversary behavior, if the cost due to the deprivation of network services exceeds the illegal security gain.

With the trust modeling and evaluation method proposed in [8], the trust/reciprocity mechanism has become a powerful tool to improve security and stimulate cooperation in wireless networks [6], [9]–[14]. These works apply the direct reciprocity principle, where the main idea is “I help you because you helped me” [15]. More specifically, each node chooses its actions according to the interaction history with its opponents, and is more likely to decline the requests from those who have ever attacked it before. Unfortunately, in a large-scale wireless network with terminal mobility, most nodes have a small chance to meet their opponents again, and thus usually have limited or outdated knowledge on the histories of their current opponents. In other words, attackers are unlikely to meet their victims again and thus are rarely punished by them in the future. Consequently, the direct reciprocity game can effectively defend the network security only when the node population is small and limited without too much mobility.

We have found that this problem can be addressed by the use of the indirect reciprocity principle [16]. First developed in social science and evolutionary biology, the main idea of the indirect reciprocity game is “I help you and somebody else helps me” [15]. This strategy is promising to stimulate cooperation in large cognitive networks [17], and can be used to improve the Sybil-resistance for the accounting of peer contributions in peer-to-peer networks [18]. In Sybil attacks, one node claims to have multiple identities other than itself, in hopes of obtaining illegal advantages in various aspects such as receiving more network services or more weights in the network voting [4], [5].

In this paper, we formulate the security problem as an indirect reciprocity game, and propose a security system[1] to counteract

Manuscript received January 10, 2012; revised May 14, 2012; accepted May 15, 2012. Date of publication June 01, 2012; date of current version July 09, 2012. The work of L. Xiao was supported in part by NSFC (61001072, 6100114), in part by the Natural Science Foundation of Fujian Province of China (2010J01347), in part by NCETFJ, in part by Fundamental Research Funds for the Central Universities (2012121028, 2011121050, 2010121066), in part by SRF for ROCS, SEM, and in part by Tsinghua-Qualcomm Joint Research Center. This work was presented in part at the IEEE International Conference on Communications (ICC’12). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Robert H. Deng.

L. Xiao was with was with the University of Maryland, College Park, MD 20742 USA. She is now with the Department of Communication Engineering, Xiamen University, Xiamen, Fujian, 361005 China (e-mail: lxiao@xmu.edu.cn).

Y. Chen and K. J. R. Liu are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: yan@umd.edu; kjrliu@umd.edu).

W. S. Lin is with Intel, Hillsboro, OR 97124-5961 USA (e-mail: w.sabrina.lin@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2012.2202228

[1] We mainly focus on the modeling aspects in this paper. Due to the page limitations, we do not provide in-depth discussions on the implementations or systematic issues of the wireless security “system.” Instead of indicating the related “systematic” work, the word “system” here only implies that multiple nodes in the network cooperate to improve the security performance by following a certain social norm in the presence of gossip channels.

a wide range of attacks in wireless networks, including jamming, spoofing, Sybil, collusion attacks, relay-related attacks such as the packet dropping attacks, and many others [2]–[7], [19]. The reputation propagation mechanism in this system allows attackers to be recognized and punished by a much larger node population in the network, compared with the direct reciprocity system. Consequently, our system can provide a stronger security protection, especially for the large-scale wireless networks with node mobility.

We assume that multiple transmissions can take place simultaneously in a large-scale network without interfering with each other. During each transmission, the intended receiver and other observing nodes evaluate the behavior of the neighboring nodes in this area, update their reputations, and propagate the new reputations to the whole network through gossip channels. More specifically, we build a public social norm and reputation updating process to assign low (bad) reputations to the attackers. As a result, most nodes in the network reject the future requests by these "bad" nodes for network service over a long punishment time. We apply the attack classification technique, where different reputations are assigned to different attackers according to their impacts on the network performance. By punishing the attackers in different manners, we can more efficiently address the most dangerous attacks against the network. The system is designed to incur the cost to each attacker in the punishment duration to exceed its illegal security gain of attacks, so that rational nodes in the network are stimulated to deviate from the adversary behavior for their own interests. Our system is promising to improve the security performance of many wireless networks, including cognitive radio networks.

Our main contributions can be summarized as follows:

1) We provide a game theoretic analysis regarding the security issues of the large-scale wireless network with possible node mobility, and propose a solution based on the indirect reciprocity principle to counteract a wide range of wireless attacks, including collusion attacks.
2) We present the stationary reputation distribution of our desirable node behavior in the formulated game, and provide the condition for our proposed solution to effectively suppress adversary behaviors.
3) Simulations are performed to show that our desirable node action is an evolutionarily stable strategy and the attacker population diminishes at a fast speed. Simulation results also verify the security gain of the attack classification and that of the discriminated forgetting factors in the reputation updating process.

The remainder of the paper is organized as follows. In Section II, we present the network model and game model. In Section III, we describe our security system for wireless networks based on indirect reciprocity. Next, we analyze its performance in Section IV, and present the simulation results in Section V. Finally, we conclude in Section VI.

II. System Model

A. Network Model

We consider a homogenous wireless network, consisting of N randomly located nodes. Each node leaves the network with probability δ and (re)enters it with probability δ. For simplicity, we assume that each node is assigned a unique identity that cannot be changed by itself, and knows the identity of its neighbors via local information exchange. In the network, N_s nodes are allowed to simultaneously send messages to their intended receivers, if they do not interfere with each other.

As shown in Fig. 1, each transmission scenario includes a transmitter, an intended receiver, and n neighboring nodes, including Player 1 to n. Without loss of generality, in this work, we assume that n is constant for each transmission, and $N_s = N\rho$, where the transmission probability, ρ, decreases with the network geographical density. For simplicity, we also assume that each neighboring node is within the coverage area of a single transmitter.[2]

In each transmission, the transmitter sends a message to the receiver, possibly with the help of some relay nodes. Instead of restricting to a specific relay selection algorithm, in this study, we simply use a relay indicator ϖ to denote the relay selection results that depend on factors such as the network topology, radio channel conditions, and nodes' reputations. More specifically, $\varpi_i = 1$ if Player i is selected by the transmitter to relay, and $\varpi_i = 0$ if otherwise. We mainly focus on how each node reacts to the transmission.

In this study, we consider a wide range of attacks, such as jamming, spoofing, Sybil attacks, malicious packet dropping, and collusion attacks [6], [7]. Each node is able to launch any type of attacks at each time. Extensive work has been done in the literature to detect attacks and/or to identify the attackers, and various algorithms have been proposed for each type of attack, such as jamming [19]. Instead of being restricted to a specific algorithm, our security system can incorporate most existing attack detection/identification algorithms.

To improve the security performance, we split attacks into different categories, according to their impacts on the network performance and their costs to the attackers. Without loss of generality, let Level-g attacks be more dangerous to the network than Level-h attacks, with $g < h$. For example, a jammer can be labelled with a lower level than a spoofer, if the latter is considered to be less dangerous to a given network.

B. Game Model

Each transmission process can be formulated into one round of the indirect reciprocity game with $n+1$ rational players: the transmitter and n nodes in the neighboring area. The transmitter selects a subset of neighboring nodes as its relays. In response, Player i chooses an action at time k, denoted as $A_i[k]$, from the action set $\{1, 2, \ldots, L\}$. As shown in Table I, the action j ($< L-1$) corresponds to Level-j attack. The action $L-1$ is to disobey the request from the transmitter, while the action L is to follow the request from the transmitter. Note that each action can actually correspond to a different communication action, according to the relay indicator. For example, a nonrelay node with action L has to keep silence, while a relay node with action L actually transmits.

In the absence of the other nodes, an individual node taking the action j ($1 \le j \le L$) can receive an instant payoff, denoted

[2] Our analysis can be extended to the other cases.

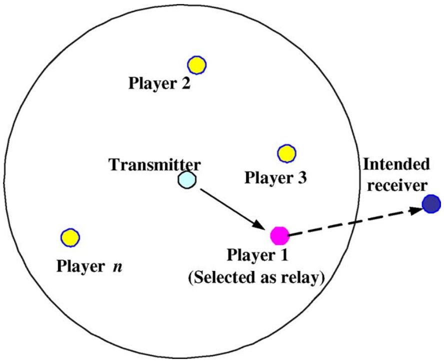


Fig. 1. Communication topology in the game formulation, including a transmitter, an intended receiver, and n nodes in the communication region (including Player 1 to Player n), with some neighboring nodes selected to relay (Player 1 in this example).

TABLE I
ACTION SET OF EACH NODE

Action ID	Physical action
1	Level-1 (most dangerous) attacks, e.g., jamming
2	Level-2 attacks, e.g., collusion attacks
$\cdots$	...
$L-3$	Level-($L-3$) attacks, e.g., Sybil attacks
$L-2$	Level-($L-2$) (mildest) attacks, e.g., spoofing
$L-1$	Disobey the request from the transmitter
L	Follow the request from the transmitter

as $C_j(w)$, which is the security gain minus the related cost. At the same time, with such an action j, the transmitter receives an instant payoff, denoted as G_j, which can be the transmission gain or the security loss. A node with a positive payoff gains from that action, while a negative payoff indicates a loss to the node. Note that the payoff of an action to the player itself also depends on the relay indicator ϖ. For instance, the action L, i.e., to follow the request, costs more energy to the relay node, compared with the nonrelay node with $\varpi = 0$. For simplicity of notation, we use $C(\varpi) = [C_1(\varpi), \ldots, C_L(\varpi)]$ (or $G = [G_1, \ldots, G_L]$).

In this game, if a node follows the request by the transmitter, the latter benefits (i.e., $G_L > 0$), while a relay node has to consume energy to transmit and thus takes a higher cost compared with a nonrelay node, i.e., $C_L(\varpi = 1) < C_L(\varpi = 0) \leq 0$. Note that our system is designed to punish the nodes that attack the network or reject the request by a good node. Therefore, a rational node never launches any of those actions unless obtaining a positive instant payoff, i.e., $C_j(\varpi) \geq 0$, for an action $j < L$. In this case, the transmitter suffers from the security or throughput loss, implying $G_j \leq 0$, for $j < L$. In addition, the action with a lower label is more dangerous to the network and brings more (illegal) security advantages to the player itself, and thus $G_L \geq G_{L-1} \geq \cdots \geq G_1$ and $C_L(\varpi) \leq C_{L-1}(\varpi) \leq \cdots \leq C_1(\varpi)$.

For simplicity, we assume perfect radio propagation in this area. That is, the transmission is successful if all the nodes in the area follow the requests. The performance of the transmitter depends on its worst neighbor, or the worst action taken by its n neighbors. For instance, the transmission fails if any neighbor disobeys the request. Another example is that a single attacker can ruin the whole transmission. Therefore, the payoff to the transmitter at time k, denoted as $U_T[k]$, is assumed to be the minimum instant payoff as follows:

$$U_T[k] = \min_{1 \leq i \leq n} G_{A_i[k]} \tag{1}$$

where $A_i[k]$ is the action taken by Player i at time k. In addition, we denote the payoff of the action i to the player itself as $U_i[k]$ and assume it to be independent of other nodes, i.e.,

$$U_i[k] = C_i(\varpi) \tag{2}$$

where ϖ is the current relay indicator of the player. The intended receiver and the observing nodes monitor the transmission and evaluate the behavior of each node. For ease of reference, the commonly used notations are summarized in Table II.

III. Security System Based on Indirect Reciprocity

We design a security system that applies the indirect reciprocity principle to reduce the potential attacker population in a wireless network, where attackers are not only punished by their direct victims, but also by most other nodes in the network. This system does not require a centralized process or a central unit. Each node in the network checks the action of its neighbors, updates their reputations, and broadcasts the new reputations to the network via gossip channels.

In this reputation-based system, the reputation vector that is allocated to each node according to its action history determines its probability to receive the network services. More specifically, the transmitter with a higher reputation value is more likely to obtain the node cooperation in the network. On the other hand, the reputation of a node decreases if it attacks the network or disobeys the transmitter with a good reputation. Meanwhile, the reputation also decreases if the node helps a "bad" node. In this way, our system motivates nodes not to attack.

Each transmission consists of two stages: the message transmission stage and the reputation evaluation stage. In the latter stage, a node's reputation is updated and broadcast via gossip channels, based on its current reputation and the instant reputation due to its action in the first stage. A forgetting factor, which can be either fixed or related to the value of the instant reputation, is introduced to weight the current reputation in this calculation. Assuming that most nodes are rational and requiring network services, our system forgives a former "bad" node and allows it to regain the network resources, if it follows what is required by the network social norm during the punishment period. The punishment duration is determined by the forgetting factors in the reputation updating process.

With a reputation set $R = \{1, \ldots, L\}$, our system assigns to each node a scalar reputation $j \in R$ and a reputation vector $\mathbf{p} = [p_1, p_2, \ldots, p_L]^T$, where p_l is the probability for the node to have a scalar reputation l. Clearly, we have $0 \leq p_l \leq 1$ and $\sum_{l=1}^{L} p_l = 1$. The scalar reputation j is a realization of an integer random variable whose probability mass function (PMF) is the corresponding reputation vector $\mathbf{p}$. In general, a node whose reputation vector has a larger

TABLE II
SUMMARY OF SYMBOLS AND NOTATIONS

N	Number of nodes in the network
n	Number of neighboring nodes in the area
N_s	Number of transmitters in the network at a given time
L	Size of the action set
$Q = [Q_{i,j}]_{L \times L}$	Social norm
$\mathbf{p} = [p_1, \cdots, p_L]^T$	Reputation vector of a player
$\mathbf{p}^* = [p_1^*, \cdots, p_L^*]^T$	Stationary reputation distribution
$\mathbf{e}_x$	Standard basis vector
σ	Prob. to successfully identify an attacker
p_s	Prob. to be selected to relay
ϖ	Relay selection indicator
$\mathbf{a}^*$	Optimal action strategy
Λ (λ)	Forgetting factor vector (value)
Φ	Reputation propagation matrix
δ	Prob. to (re)enter a network
ρ	Prob. to be a transmitter
$C_i(\varpi)$	Payoff of action i to the node itself
G_i	Payoff of action i to the transmitter
$U_i[k]$	Payoff to the player with action i at time k
$U_{i,j}^l$	Expected long-term payoff to the player with reputation i who takes action l towards a node with reputation j
$U_{i,j}$	Average expected long-term payoff to the player with reputation i when meeting a node with reputation j
$r_{i,j}^l[k]$	Prob. the reputation of the node changing from i to k after an action l when meeting a node with reputation j

mean value is more likely to have a higher scaler reputation. Compared with the scalar reputation that only describes the instant or average value, the reputation vector includes the likelihood of each reputation, and thus contains more information on the past actions of the node. Therefore, the reputation vector can provide better performance.

Each node chooses its action $a_{i,j} \in \{1, \ldots, L\}$ according to its own scalar reputation i and the transmitter's reputation j. During each transmission, the nodes' reputations are updated based on the same social norm, denoted as $Q = [Q_{i,j}]_{L \times L}$, where $Q_{i,j} \in \{1, \ldots, L\}$ is the instant scalar reputation assigned to the node who takes action i towards a transmitter with a scalar reputation j.

In this reciprocity system, the social norm is designed to guide the node behavior and to suppress attacks in the network. In general, nodes can receive high reputations by helping the good nodes or disobeying the transmitters with bad reputations. On the other hand, in order to maintain a healthy network, the system reduces the reputation of each identified attacker even in the presence of a "bad" transmitter. For simplicity, we assign an instant reputation i to the node that launches Level-i attacks, with $1 \leq i \leq L-2$.

When the transmitter has the highest reputation L, we encourage the other nodes to help it. More specifically, nodes that follow the request by the transmitter whose reputation is L obtain the highest instant reputation (L), while nodes that refuse to cooperate receive the reputation $L-1$. Otherwise, when the reputation of the transmitter is less than L, our desirable action is to keep silence, and hence relay nodes receive the highest reputation by taking the action $L-1$. Based on the above principles, we build the following social norm for relay nodes:

$$Q(\varpi = 1) = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 2 & 2 & \cdots & 2 & 2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L-2 & L-2 & \cdots & L-2 & L-2 \\ L & L & \cdots & L & L-1 \\ 1 & 2 & \cdots & L-1 & L \end{bmatrix}.$$

On the other hand, nonrelay nodes with the action L do not transmit and thus receive the highest reputation. Meanwhile, the action $L-1$ of the nonrelay nodes results in packet collisions. Therefore, the node with $\varpi = 0$ receives an instant reputation i by taking the action i, with $i \in \{L-1, L\}$, regardless of the transmitter's reputation. In summary, the social norm can be written as

$$Q(\varpi) = \varpi \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 2 & 2 & \cdots & 2 & 2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L-2 & L-2 & \cdots & L-2 & L-2 \\ L & L & \cdots & L & L-1 \\ 1 & 2 & \cdots & L-1 & L \end{bmatrix} + (1-\varpi) \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 2 & 2 & \cdots & 2 & 2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L-2 & L-2 & \cdots & L-2 & L-2 \\ L-1 & L-1 & \cdots & L-1 & L-1 \\ L & L & \cdots & L & L \end{bmatrix}. \quad (3)$$

Similarly, our desirable action strategy, denoted as $\hat{\mathbf{a}}$, can be written as

$$\hat{\mathbf{a}}(\varpi) \triangleq [\hat{a}_{i,j}]_{L \times L} = \varpi \begin{bmatrix} L-1 & L-1 & \cdots & L-1 & L \\ L-1 & L-1 & \cdots & L-1 & L \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L-1 & L-1 & \cdots & L-1 & L \\ L-1 & L-1 & \cdots & L-1 & L \end{bmatrix} + (1-\varpi) \begin{bmatrix} L & L & \cdots & L \\ L & L & \cdots & L \\ \cdots & \cdots & \cdots & \cdots \\ L & L & \cdots & L \\ L & L & \cdots & L \end{bmatrix}. \quad (4)$$

Each node newly entering to the network obtains a high initial reputation, $\mathbf{p}[0] = [0, \ldots, 0, 1]^T$. During the transmission at time k, the observing nodes in this area monitor the action of each node, and assign an instant scalar reputation $x = Q_{a,j}(\varpi)$ to the node with a relay indicator ϖ, who takes an action a to a transmitter with a reputation j.

The reputation updating process is illustrated in Fig. 2, where the instant vector reputation $\mathbf{e}_x$ is the standard basis vector based on the scalar reputation. The new vector reputation at time $k+1$ relies on both the instant reputation $\mathbf{e}_x$ and the current reputation $\mathbf{p}[k]$. The latter is weighted by Λ_x, the xth element of the forgetting factor vector $\Lambda = [\Lambda_1, \Lambda_2, \ldots, \Lambda_L]$. The current action a has less impacts on $\mathbf{p}[k+1]$, under a larger forgetting factor value Λ_x. The nodes' reputations are then broadcast to the network via gossip channels, with the reputation propagation matrix denoted as Φ. As shown in

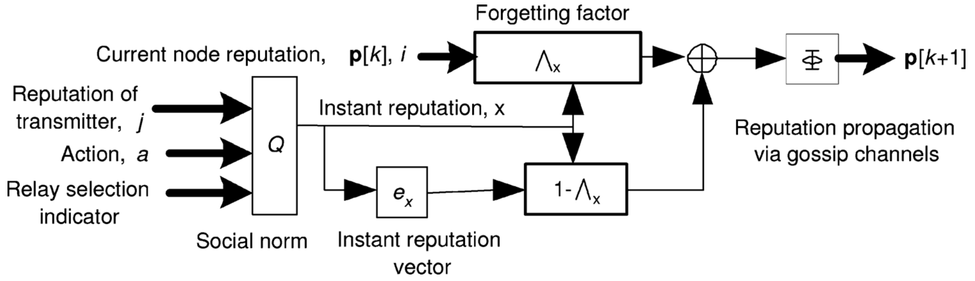

Forgetting factor
Current node reputation, p[k], i
Reputation of transmitter, j
Action, a
Relay selection indicator
Q
Social norm
Instant reputation, x
Λ_x
e_x
Instant reputation vector
$1-\Lambda_x$
Φ
p[k+1]
Reputation propagation via gossip channels

Fig. 2. Reputation updating process in our security system.

Fig. 2, the new reputation of the node propagated via gossip channels is given by

$$\mathbf{p}[k+1] = \Phi\left(\Lambda_x \mathbf{p}[k] + (1-\Lambda_x)\,\mathbf{e}_x\right). \tag{5}$$

The reputation propagation matrix $\Phi \triangleq [\Phi_{l,m}]_{L\times L}$, where $\Phi_{l,m}$ is the probability for the reputation m to be taken as l, due to both the behavior detection error and the gossip channel propagation error. Denoting the probability for the action l to be correctly recognized by the whole network with p_{Dl}, we can model the reputation propagation process as follows:

$$\Phi = \begin{bmatrix} p_{D1} & \frac{1-p_{D1}}{L-1} & \frac{1-p_{D1}}{L-1} & \cdots & \frac{1-p_{D1}}{L-1} \\ \frac{1-p_{D2}}{L-1} & p_{D2} & \frac{1-p_{D2}}{L-1} & \cdots & \frac{1-p_{D2}}{L-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1-p_{DL}}{L-1} & \frac{1-p_{DL}}{L-1} & \frac{1-p_{DL}}{L-1} & \cdots & p_{DL} \end{bmatrix}. \tag{6}$$

Note that our system is not restricted to (6) and can be easily extended to other models. In addition, each node can be a transmitter, a receiver, a neighboring node for a transmitter (either a relay or a nonrelay), or an isolated node that is far away from all the other nodes and is not inspected at the time slot. If a node is not inspected, its reputation does not change during this time slot.

In our system, all the nodes in the network know the social norm, the forgetting factor vector, and the reputation updating algorithm as shown in Fig. 2. In addition, each node that updates the reputation for its neighbor is assumed to know the current reputation vector, the relay indicator and the action of the node under study, as well as the reputation vector of the corresponding transmitter. More specifically, as a heuristic method, the reputation vectors of all the nodes in the network are stored in a central server. Each node observes the behaviors of its neighboring nodes and updates their reputations according to the reputation updating process, after retrieving their current reputation vectors and the reputation of the transmitter if they are absent. An in-depth study on the implementation will be carried out in the future.

IV. Performance Analysis

Our proposed system provides the "good" nodes that have higher reputations with the benefits of receiving better network accesses in the future. On the other hand, attackers are labelled with very low reputations and lose their network services during the punishment time, whose duration depends on the forgetting factor in the reputation updating process. If the cost to be punished is larger than the illegal security gain, rational nodes have incentives to choose the desirable actions and abandon adversary behaviors.

In this section, we evaluate whether the desirable "good" action (i.e., to follow the social norm) cannot be invaded by any "bad" strategy (such as a type of attacks) that is initially rare, if the "good" one is adopted by a large number of nodes. To this end, we apply the evolutionarily stable strategy concept [18] to investigate the stable equilibrium of the proposed security system, and use the Wright–Fisher model [20] to study how the action rules spread over the network. We derive the stationary reputation distribution of the optimal action, and investigate the stability condition for this system to efficiently counteract attacks. This section also covers related issues such as the robustness against collusion attacks.

A. Evolutionarily Stable Strategy

An evolutionarily stable strategy (ESS) cannot be invaded by any alternative strategy that is initially rare, and natural selection alone is sufficient to prevent alternative strategies from invading [20]. To evaluate the stability property, we apply the Wright–Fisher model [20] for the action spreading, where the probability for a node to choose a strategy is proportional to the expected payoff to the population using that strategy. More specifically, the probability for a node to choose action strategy i at time $k+1$, $y_i[k+1]$, is given by

$$y_i[k+1] = \frac{y_i[k]U_i[k]}{\sum\limits_l y_l[k]U_l[k]}. \tag{7}$$

In this way, we can perform simulations to investigate the ESS property of our desirable action strategy in the formulated game. More details are given in the next section.

B. Optimal Action and Stationary Reputation Distribution

In what follows, we evaluate the optimal action strategy in our security game and provide the corresponding stationary reputation distribution $\mathbf{p}^*$. The optimal strategy, $\mathbf{a}^* \triangleq [a^*_{i,j}]_{L\times L}$, is defined to maximize the expected long-term payoff for the player under various scenarios, where $a^*_{i,j}$ is the optimal strategy for the node with a scalar reputation i when meeting a node with a reputation j. Let $U_{i,j}$ denote the expected maximal reward to a node with a reputation i in presence of node with a reputation j. By definition, we have

$$U_{i,j} = \max_{1\le a_{i,j}\le L} U_{i,j}^{a_{i,j}} = U_{i,j}^{a^*_{i,j}} \tag{8}$$

where $U^l_{i,j}$ is the expected reward to a node with a reputation i against a node with a reputation j, when the action l is taken. The optimal action here $a^*_{i,j}$ satisfies

$$a^*_{i,j} = \arg\max\nolimits_{1\le a_{i,j}\le L} U^{a_{i,j}}_{i,j}. \tag{9}$$

For simplicity, we assume in this section that each node is randomly and independently selected to relay with a probability p_s, and that the forgetting factor $\Lambda_x = \lambda$ is constant for each reputation level x. Let $\mathbf{r}^l_{i,j}$ denote the reputation transfer vector, which provides the new reputation vector for a node with initial scalar reputation i, which takes the action l towards a transmitter with a reputation j. According to the reputation updating process described in (5), the reputation transfer vector can be written as

$$\begin{aligned}\mathbf{r}^l_{i,j} &\triangleq \begin{bmatrix} r^l_{i,j}[1] \\ r^l_{i,j}[2] \\ \cdots \\ r^l_{i,j}[L] \end{bmatrix} \\ &= \Phi\left(\lambda \mathbf{e}_i + (1-\lambda)\left(p_s \mathbf{e}_{Q_{l,j}(1)} + (1-p_s)\mathbf{e}_{Q_{l,j}(0)}\right)\right)\end{aligned} \tag{10}$$

where $r^l_{i,j}[k]$ is the probability for the reputation of the player that takes the action l towards a transmitter with a reputation j to change from i to k. The first term in the parentheses $\lambda\mathbf{e}_i$ results from the current reputation, whereas the second term depends on the current action and is averaged over all the relay selection results.

Assume that a node with a reputation i would take an action $a_{i,j}$ towards a node with a reputation j, with the latter following the optimal strategy. First, the average one-shot reward of the action to the node itself can be written as $p_s C_{a_{i,j}(1)}(1) + (1 - p_s)C_{a_{i,j}(0)}(0)$. Assuming that the probability for a node to stay or (re)enter the network, denoted as δ, is the discounting factor of the future, as the stationary reputation distribution $\mathbf{p}^*$ does not change over time, we obtain the reward to the nontransmitter node as $p_s C_{a_{i,j}}(1)+(1-p_s)C_{a_{i,j}}(0)+\delta\sum_k\sum_l r^{a_{i,j}}_{i,j}[k]p^*_l U_{k,l}$.

On the other hand, the expected maximal one-shot reward for the optimal action to the transmitter is $(1 - p_s)G_{a^*_{j,i}(0)} + p_s G_{a^*_{j,i}(1)}$. As the reputation of a transmitter does not change, nor does the stationary reputation distribution $\mathbf{p}^*$, the reward to the transmitter is given by $(1 - p_s)G_{a^*_{j,i}(0)} + p_s G_{a^*_{j,i}(1)} + \delta\sum_l p^*_l U_{i,l}$.

Let ρ be the probability for a node to transmit and $1-\rho$ be its probability to be a neighboring node. Then, according the above discussion, the Bellman equation of the expected reward to the node can be given by

$$\begin{aligned}U^{a_{i,j}}_{i,j} =& (1-\rho)\Bigg(p_s C_{a_{i,j}}(1) + (1-p_s)C_{a_{i,j}}(0) \\ &\qquad + \delta\sum_k\sum_l r^{a_{i,j}}_{i,j}[k]p^*_l U_{k,l}\Bigg) \\ &+ \rho\left((1-p_s)G_{a^*_{j,i}(0)} + p_s G_{a^*_{j,i}(1)} + \delta\sum_l p^*_l U_{i,l}\right).\end{aligned} \tag{11}$$

According to the reputation updating process given by (5), the stationary reputation distribution corresponding to the optimal action $\mathbf{a}^*$ averaged over both $\varpi = 1$ and $\varpi = 0$, can be written as (12), shown at the bottom of the page, where $\sum_{i=1}^L\sum_{l:Q_{a^*_{i,l},l}(\varpi)=k} p^*_i p^*_l$ is the probability for the node with a relay selection indicator ϖ to obtain a reputation k by taking the optimal action $\mathbf{a}^*$. In this way, the optimization of the optimal actions can be formulated as a Markov Decision Process (MDP).

According to (10)–(12), we can compute the optimal action $\mathbf{a}^*$ and the stationary reputation distribution $\mathbf{p}^*$ alternatively by iteratively fixing one and solving the other. More specifically, similar to [17], we first apply the gradient descent algorithm to find the stationary reputation distribution for a given (initial) optimal action, and then use the dynamic programming technique to derive the optimal action under this stationary reputation. The iteration continues unless the solution converges.

There might be multiple stationary reputation points in this system. Our system motivates users to follow our desirable result by setting an appropriate social norm and reputation propagation mechanism. We will show later that the desired action strategy is evolutionarily stable, i.e., the system can reach our desired point if a large number of the nodes follow the desirable action strategy and the costs satisfy some conditions presented in the following. The stationary reputation distribution for a^* is unique and depends on the social norm Q and the reputation propagation matrix Φ, as shown by (12).

C. Stability Condition

In this subsection, we analyze the condition that our desired strategy $\hat{\mathbf{a}}$, in (4), is evolutionarily stable, i.e., each node is motivated to adopt the desired strategy, $\hat{\mathbf{a}} = \mathbf{a}^*$. As indicated by (4),

$$\mathbf{p}^* = \Phi\left(\lambda\mathbf{p}^* + (1-\lambda)\left(p_s\begin{bmatrix}\sum_{i=1}^L\sum_{l:Q_{a^*_{i,l},l}(1)=1} p^*_i p^*_l \\ \sum_{i=1}^L\sum_{l:Q_{a^*_{i,l},l}(1)=2} p^*_i p^*_l \\ \cdots \\ \sum_{i=1}^L\sum_{l:Q_{a^*_{i,l},l}(1)=L} p^*_i p^*_l\end{bmatrix} + (1-p_s)\begin{bmatrix}\sum_{i=1}^L\sum_{l:Q_{a^*_{i,l},l}(0)=1} p^*_i p^*_l \\ \sum_{i=1}^L\sum_{l:Q_{a^*_{i,l},l}(0)=2} p^*_i p^*_l \\ \cdots \\ \sum_{i=1}^L\sum_{l:Q_{a^*_{i,l},l}(0)=L} p^*_i p^*_l\end{bmatrix}\right)\right) \tag{12}$$

nodes select their actions regardless of their own reputations. Therefore, the optimal action in this game can be expressed as

$$\mathbf{a}^*(\varpi) = \begin{bmatrix} a_1^*(\varpi) \\ a_2^*(\varpi) \\ \cdots \\ a_{L-1}^*(\varpi) \\ a_L^*(\varpi) \end{bmatrix} = \begin{bmatrix} a_{i,1}^*(\varpi) \\ a_{i,2}^*(\varpi) \\ \cdots \\ a_{i,L-1}^*(\varpi) \\ a_{i,L}^*(\varpi) \end{bmatrix} = \varpi \begin{bmatrix} L-1 \\ L-1 \\ \cdots \\ L-1 \\ L \end{bmatrix} + (1-\varpi) \begin{bmatrix} L \\ L \\ \cdots \\ L \\ L \end{bmatrix} \quad (13)$$

where $a_i^*(\varpi)$ is the optimal action for the node with relay indicator ϖ against a transmitter with a reputation i. In addition, we assume that the probability for each reputation to be accurately broadcast is $P_{Di} \equiv \sigma$, where σ is the probability to successfully identify an attacker, and thus can rewrite the reputation propagation matrix (6) as

$$\Phi = \begin{bmatrix} \sigma & \frac{1-\sigma}{L-1} & \cdots & \frac{1-\sigma}{L-1} \\ \frac{1-\sigma}{L-1} & \sigma & \cdots & \frac{1-\sigma}{L-1} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1-\sigma}{L-1} & \frac{1-\sigma}{L-1} & \cdots & \sigma \end{bmatrix}. \quad (14)$$

Proposition 1: Given any scalar reputations, $1 \le j \le L$ and $1 \le m \le L-1$, we have

$$U_{L,j} - U_{m,j} = \frac{\rho p_s(G_L - G_{L-1})(L-1)}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)}. \quad (15)$$

Proof: According to (10), (13), and (14), the difference between two reputation transfer vectors in terms of their optimal strategies is given by

$$\begin{aligned} \mathbf{r}_{m,j}^{a^*_{m,j}} - \mathbf{r}_{L,j}^{a^*_{L,j}} &= \mathbf{r}_{m,j}^{a^*_j} - \mathbf{r}_{L,j}^{a^*_j} = \Phi\lambda(\mathbf{e}_m - \mathbf{e}_L) \\ &= \lambda\Phi\Omega = \lambda\frac{L\sigma-1}{L-1}\Omega \end{aligned} \quad (16)$$

where the vector $\Omega = \mathbf{e}_m - \mathbf{e}_L$. For convenience, let $\mathbf{x}[k]$ denote the kth element of a vector $\mathbf{x}$ in the following proof.

By combining (11), (13), and (16), we obtain (17), shown at the bottom of the page.

As shown in (17), $U_{m,j} - U_{L,j}$ is independent of j. Therefore, (17) can be further simplified into

$$\begin{aligned} U_{m,j} - U_{L,j} = {} & \rho p_s(G_{L-1} - G_L) \\ & + \delta\left(\rho + (1-\rho)\lambda\frac{L\sigma-1}{L-1}\right)(U_{m,j} - U_{L,j}) \end{aligned} \quad (18)$$

which leads to (15). ■

The above proposition indicates that by choosing our desirable strategy, (13), each node whose reputation m is less than the highest, can obtain an expected reward, i.e.,

$$U_{L,j} - \frac{\rho p_s(G_L - G_{L-1})(L-1)}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)}$$

no matter how low its current reputation is, or who its opponent is. In this way, our system can motivate the nodes to behave nicely and not to attack the network.

Theorem 1: The sufficient condition for the desired strategy to be an ESS in our indirect reciprocity-based security system as described in Section III is

$$\begin{cases} C_1(\varpi) - C_L(\varpi) < \frac{\delta(1-\lambda)(L\sigma-1)\rho p_s(G_L - G_{(L-1)})}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)}, & \varpi = 0, 1 \\ \sigma > \frac{1}{L} \end{cases}. \quad (19)$$

Proof: According to the one shot deviation principle for MDP [21] and the definition of ESS [22], a^* is an ESS if the following inequality holds:

$$U_{i,j}^{a^*_{i,j}} > U_{i,j}^{a_{i,j}} \quad (20)$$

for all $a_{i,j}$ that is not equal to $a^*_{i,j}$.

We first consider the case that the opponent has a reputation L, and $\varpi = 1$. By (13), $a_L^*(\varpi = 1) = L$ is the desirable strategy here, and thus $U_{i,L}^{a_{i,L}(1)=L} - U_{i,L}^{a_{i,L}(1)=m}$ should be positive, for any action $m < L$. Therefore, according to (11), we have (21), shown at the bottom of the next page. Thus, according to (3), (10), and (21), we can obtain

$$\begin{aligned} & p_s(C_L(1) - C_m(1)) \\ & > \delta\sum_k\sum_l (r_{i,L}^{a_{i,L}(1)=m}[k] - r_{i,L}^{a_{i,L}(1)=L}[k])p_l^* U_{k,l} \\ & = \delta\sum_k\sum_l \Phi(1-\lambda)\left(p_s\left(\mathbf{e}_{Q_{m,L}(1)} - \mathbf{e}_{Q_{L,L}(1)}\right)\right)[k]p_l^* U_{k,l} \\ & = \delta(1-\lambda)p_s\left(\frac{L\sigma-1}{L-1}\right)\sum_l p_l^*\left(U_{m,l} - U_{L,l}\right). \end{aligned} \quad (22)$$

$$\begin{aligned} & U_{m,j} - U_{L,j} U_{m,j}^{a^*_{m,j}} - U_{L,j}^{a^*_{L,j}} = U_{m,j}^{a^*_j} - U_{L,j}^{a^*_j} \\ & = (1-\rho)\delta\sum_k\sum_l \left(r_{m,j}^{a^*_j}[k] - r_{L,j}^{a^*_j}[k]\right)p_l^* U_{k,l} \\ & \quad + \rho\left((1-p_s)\left(G_{a^*_m(0)} - G_{a^*_L(0)}\right) + p_s\left(G_{a^*_m(1)} - G_{a^*_L(1)}\right) + \delta\sum_l p_l^*\left(U_{m,l} - U_{L,l}\right)\right) \\ & = (1-\rho)\delta\sum_k\sum_l \lambda\frac{L\sigma-1}{L-1}\Omega[k]p_l^* U_{k,l} + \rho\left(p_s\left(G_{L-1} - G_L\right) + \delta\sum_l p_l^*\left(U_{m,l} - U_{L,l}\right)\right) \\ & = \rho p_s(G_{L-1} - G_L) + \delta\left(\rho + (1-\rho)\lambda\frac{L\sigma-1}{L-1}\right)\sum_{l=1}^{L} p_l^*\left(U_{m,l} - U_{L,l}\right) \end{aligned} \quad (17)$$

By incorporating (15) into (22), we have

$$\begin{aligned}C_m(1) - C_L(1) &< \delta(1-\lambda)\left(\frac{L\sigma - 1}{L-1}\right)(U_{L,l} - U_{m,l})\\ &= \frac{\delta(1-\lambda)(L\sigma-1)\rho p_s(G_L - G_{L-1})}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)}\end{aligned} \tag{23}$$

and hence

$$C_1(1) - C_L(1) < \frac{\delta(1-\lambda)(L\sigma-1)\rho p_s(G_L - G_{L-1})}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)}. \tag{24}$$

As $G_L \geq G_{L-1}$ and $C_L(1) \leq C_1(1)$, (24) implies that

$$\sigma > \frac{1}{L}. \tag{25}$$

Next, we study the second case, where the opponent has a reputation L and $\varpi = 0$. It is indicated by (13) that $a_L^*(0) = L$ is the desirable strategy here. That is, $U_{i,L}^{a_{i,L}(0)=L} > U_{i,L}^{a_{i,L}(0)=m}$, for $m < L$, which can be rewritten by (11) as

$$\begin{aligned}&U_{i,L}^{a_{i,L}(0)=L} - U_{i,L}^{a_{i,L}(0)=m} = (1-\rho)\\ &\times\Bigg((1-p_s)(C_L(0) - C_m(0))\\ &\qquad + \delta\sum_k\sum_l (r_{i,L}^{a_{i,L}(0)=L}[k] - r_{i,L}^{a_{i,L}(0)=m}[k])p_l^* U_{k,l}\Bigg) > 0.\end{aligned} \tag{26}$$

By combining (3), (10), and (26), we obtain

$$\begin{aligned}&(1-p_s)(C_L(0) - C_m(0))\\ &> \delta\sum_k\sum_l (r_{i,L}^{a_{i,L}(0)=m}[k] - r_{i,L}^{a_{i,L}(0)=L}[k])p_l^* U_{k,l}\\ &= \delta\sum_k\sum_l (1-\lambda)\left((1-p_s)\Phi(\mathbf{e}_{Q_{m,L}}(0) - \mathbf{e}_{Q_{L,L}}(0))\right)[k]p_l^* U_{k,l}\\ &= \delta(1-\lambda)(1-p_s)\left(\frac{L\sigma-1}{L-1}\right)\sum_l p_l^*\left(U_{m,l} - U_{L,l}\right).\end{aligned} \tag{27}$$

As $1 \leq m \leq L-1$, by (15), (27) becomes

$$C_1(0) - C_L(0) < \delta(1-\lambda)\left(\frac{L\sigma-1}{L-1}\right)(U_{L,j} - U_{1,j}) \tag{28}$$

$$= \frac{\delta(1-\lambda)\left(L\sigma-1\right)\rho p_s\left(G_L - G_{L-1}\right)}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)}. \tag{29}$$

In the third case, the opponent has a reputation $j < L$, and $\varpi = 1$. By (13), $a_j^*(1) = L-1$ is the desirable strategy, which means $U_{i,j}^{a_{i,j}(1)=L-1} - U_{i,j}^{a_{i,j}(1)=m} > 0$, for any action $m < L-1$. We have, by (11), (30), shown at the bottom of the page, which can be simplified into

$$\begin{aligned}&p_s(C_{L-1}(1) - C_m(1))\\ &> \delta\sum_k\sum_l (r_{i,j}^{a_{i,j}(1)=m}[k] - r_{i,j}^{a_{i,j}(1)=L-1}[k])p_l^* U_{k,l}\\ &= \delta\sum_k\sum_l \Phi(1-\lambda)\left(p_s\left(\mathbf{e}_{Q_{m,j}(1)} - \mathbf{e}_{Q_{L-1,j}(1)}\right)\right)[k]p_l^* U_{k,l}\\ &= \delta(1-\lambda)p_s\left(\frac{L\sigma-1}{L-1}\right)\sum_l p_l^*\left(U_{m,l} - U_{L,l}\right).\end{aligned} \tag{31}$$

Therefore,

$$\begin{aligned}C_m(1) - C_{L-1}(1) &< \delta(1-\lambda)\left(\frac{L\sigma-1}{L-1}\right)(U_{L,l} - U_{m,l})\\ &= \frac{\delta(1-\lambda)(L\sigma-1)\rho p_s(G_L - G_{L-1})}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)}.\end{aligned} \tag{32}$$

As $1 \leq m \leq L-2$, we have

$$C_1(1) - C_{L-1}(1) < \frac{\delta(1-\lambda)(L\sigma-1)\rho p_s(G_L - G_{L-1})}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)}. \tag{33}$$

On the other hand, the fact that $a_j^*(1) = L-1$, for $j < L$, also indicates $U_{i,j}^{a_{i,j}(1)=L-1} - U_{i,j}^{a_{i,j}(1)=L} > 0$. Similar to the

$$\begin{aligned}&U_{i,L}^{a_{i,L}(1)=L} - U_{i,L}^{a_{i,L}(1)=m}\\ &\quad= (1-\rho)\left(p_s(C_L(1) - C_m(1)) + \delta\sum_k\sum_l\left(r_{i,L}^{a_{i,L}(1)=L}[k] - r_{i,L}^{a_{i,L}(1)=m}[k]\right)p_l^* U_{k,l}\right) > 0\end{aligned} \tag{21}$$

$$\begin{aligned}&U_{i,j}^{a_{i,j}(1)=L-1} - U_{i,j}^{a_{i,j}(1)=m}\\ &\quad= (1-\rho)\left(p_s(C_{L-1}(1) - C_m(1)) + \delta\sum_k\sum_l\left(r_{i,j}^{a_{i,j}(1)=L-1}[k] - r_{i,j}^{a_{i,j}(1)=m}[k]\right)p_l^* U_{k,l}\right) > 0\end{aligned} \tag{30}$$

previous discussion, we obtain (34), shown at the bottom of the page, which can be simplified into

$$\begin{aligned} & p_s(C_{L-1}(1) - C_L(1)) \\ & > \delta \sum_k \sum_l (r_{i,j}^{a_{i,j}(1)=L}[k] - r_{i,j}^{a_{i,j}(1)=L-1}[k]) p_l^* U_{k,l} \\ & = \delta \sum_k \sum_l \Phi(1-\lambda) \left(p_s \left(\mathbf{e}_{Q_{L,j}(1)} - \mathbf{e}_{Q_{L-1,j}(1)}\right)\right) [k] p_l^* U_{k,l} \\ & = \delta(1-\lambda) p_s \left(\frac{L\sigma - 1}{L-1}\right) \sum_l p_l^* \left(U_{j,l} - U_{L,l}\right). \end{aligned} \tag{35}$$

Consequently

$$\begin{aligned} C_L(1) - C_{L-1}(1) & < \delta(1-\lambda) \left(\frac{L\sigma-1}{L-1}\right) (U_{L,l} - U_{j,l}) \\ & = \frac{\delta(1-\lambda)(L\sigma-1)\rho p_s(G_L - G_{L-1})}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)}. \end{aligned} \tag{36}$$

Now we consider the last case with $j < L$ and $\varpi = 0$. Similar to the third case, we have

$$C_1(0) - C_{L-1}(0) < \frac{\delta(1-\lambda)\left(L\sigma-1\right)\rho p_s\left(G_L - G_{L-1}\right)}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)} \tag{37}$$

and

$$C_L(0) - C_{L-1}(0) < \frac{\delta(1-\lambda)(L\sigma-1)\rho p_s(G_L - G_{L-1})}{(L-1)(1-\delta\rho) - \delta(1-\rho)\lambda(L\sigma-1)}. \tag{38}$$

As $C_L(\varpi) \le C_{L-1}(\varpi) \le C_1(\varpi)$, $\varpi = 0, 1$, by summarizing (24), (25), (29), (33), (36), (37), and (38), we can obtain (19) as the condition for the desirable solution to be an ESS in our security game. ■

D. Discussions on Security Issues

First, the stability condition, (19), provides the lower performance bound of the gossip channels. When $\sigma < 1/L$, due to the low attack detection rate and/or the deteriorated gossip channels, our reputation mechanism and hence the proposed security system collapse. In addition, another condition, as shown in (19), is that the payoff to the most dangerous attackers does not far exceed that to the good nodes with the highest reputations, and the threshold relies on the payoff to the transmitter with the highest reputation. Otherwise, if an attacker receives a high enough illegal payoff, rational nodes would have incentives to take risks to launch attacks.

With simulation results against various wireless attacks provided in Section V, we here focus on the collusion attacks, which are highly dangerous to the reputation-based networks. First, we consider a typical scenario, where several nodes collude to spread the false information regarding a third-party node, in hopes of ruining its reputation and thus its future transmission. In such a case, when noticing the reputation loss, the victim node sends alarm to initiate investigations on the colluders.

In another typical example, several colluders report faked internal cooperations, aiming to improve the reputation of each other. Although the gaining nodes never complain, the system reduces the reputation of the nodes that have obtained abnormal amounts of internal helps.

In both cases, our system assigns the lowest reputation values to the detected colluders, and the low reputations punish the nodes with stopped network accesses. The punishment duration is set to be long enough to counteract the illegal collusion gain. Therefore, our system can also resist these collusion attacks. Note that the above discussions only provide initial study on the collusion resistance and we will further investigate this issue in the future.

V. Simulation Results

Simulations have been performed to evaluate the performance of the proposed security system under different network scenarios. More specifically, we provide the evolutionarily stable property of different action strategies in a network with N nodes. In the simulations, we can assume large-scale mobile wireless networks with $N = 5000$ nodes whose locations change randomly over time, and $\delta = 0.1$ of the nodes newly entering the network and $\delta = 0.1$ of the nodes leaving the network at each time slot.

Suppose that $\rho = 0.2$ of the nodes are far away from each other during a time slot and thus are allowed to transmit simultaneously during each time slot. The system has a probability of $\sigma = 0.99$ to successfully identify an attacker. Each transmitter selects a neighboring node as relay with $p_s = 1$ and $\varpi \equiv 1$. Thus in each run of the simulation, there are ρN transmitters, ρN neighboring nodes that are chosen as relays and are inspected, and $(1-2\rho)N$ nodes whose reputations do not change over the time slot.

The action spreading is assumed to follow the Wright–Fisher model, (7). In addition, at the beginning of each simulation process, 70% of the inspected nodes choose the optimal strategy while the remaining nodes randomly select from the whole strategy set. In the following, we first investigate the efficiency of our system in the simplified scenarios without attack classification, and compare its performance with the direct reciprocity-based security scheme. Then we provide the security gain of the attack classification, and validate our derived stability condition of our security system.

$$\begin{aligned} & U_{i,j}^{a_{i,j}(1)=L-1} - U_{i,j}^{a_{i,j}(1)=L} \\ & = (1-\rho)\left(p_s(C_{L-1}(1) - C_L(1)) + \delta \sum_k \sum_l \left(r_{i,j}^{a_{i,j}(1)=L-1}[k] - r_{i,j}^{a_{i,j}(1)=L}[k]\right) p_l^* U_{k,l}\right) > 0 \end{aligned} \tag{34}$$

A. Case 1: No Attack Classification

In a case without attack classification, $L = 3$ and the action set is $\{1, 2, 3\}$, whose elements represent attacks, request rejection, and to follow the request by the transmitter, respectively. According to (3) and (13), the social norm and our desirable action rule are given, respectively, by

$$Q^{3\times 3}(\varpi = 1) = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix} \qquad (39)$$

and

$$\hat{\mathbf{a}}^3(\varpi = 1) = [2, 2, 3]^T. \qquad (40)$$

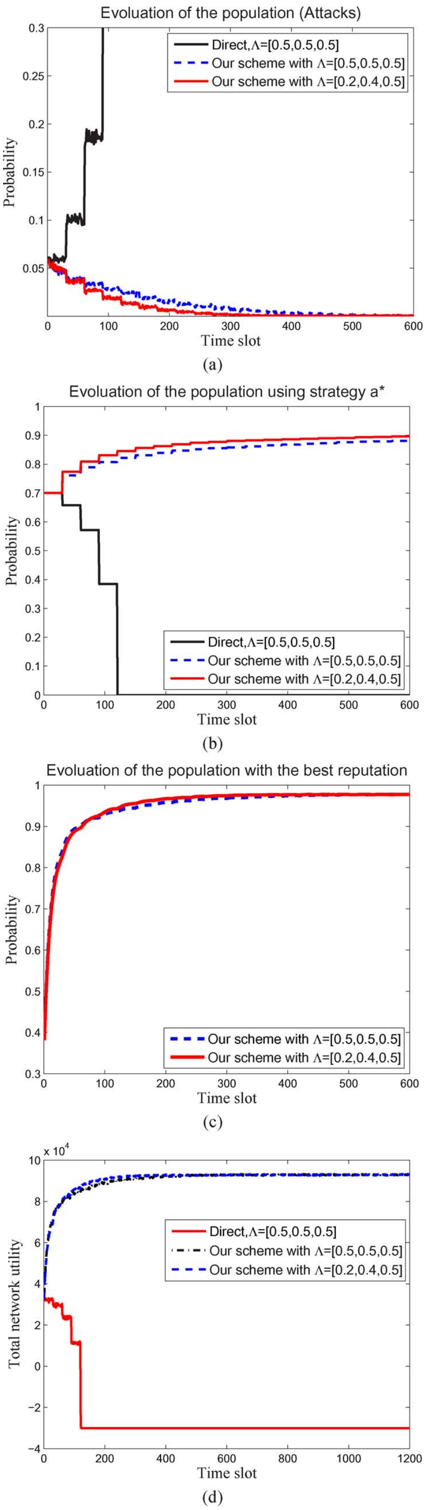
Simulation results first show that the stationary reputation distribution corresponding to $\mathbf{a}^* = \hat{\mathbf{a}}^3$ is $\mathbf{p}^* = [0.0128, 0.0128, 0.9744]^T$, with the forgetting vector $\Lambda = [0.5, 0.5, 0.5]$. This result indicates that most users can obtain the highest scalar reputation $L = 3$.

As comparison, we also consider a security system based on the direct reciprocity principle, where each node chooses its actions according to its past direct interaction with its current opponents. As shown in Fig. 3, this system fails to work in the networks with $N = 5000$ nodes, and the network corrupts shortly after the start of the process with an eruption of attacks. That is because the long-term punishment cost to an attacker in the direct reciprocity-based approach is small, compared with its illegal security gain in this game, as nodes are unlikely to meet each other again very soon in the large-scale network.

Fortunately, our indirect reciprocity-based system can address this problem and efficiently combat attacks in the large-scale network. For example, our system reduces the attacker population from around 5% to less than 0.05% shortly after 400 time slots, as shown in Fig. 3(a). It is shown in Fig. 3(b) that more than 90% of the population chooses the desirable strategy, shortly after the start of the process in our system. In addition, we see in Fig. 3(c) that most nodes obtain the highest reputation in the network soon after the start of the game, indicating that our system can maintain a healthy reputation mechanism.

We then evaluate the impacts of the differentiated forgetting factor vector in the reputation updating process of the indirect reciprocity game. As shown in Fig. 3, our scheme that applies $\Lambda = [0.2, 0.4, 0.5]$ has a better security performance than the case with $\Lambda = [0.5, 0.5, 0.5]$. The reason for the significant drop of the attacker population is that our system takes into account the attacking behavior with more weights in the reputation updating, and thus punishes attackers with longer punishment duration. This figure exhibits marked stepwise behavior, suggesting that the simulation progresses by a series of one shot games that are used to update the instantaneous reputations and system-wide reputation updates. The results verify the discussions in Section IV.

Fig. 4 verifies that our system has a better performance in the network with a smaller scale. For instance, the attacking rate here with $N = 200$ nodes decreases in a much faster speed than the case with $N = 5000$ as in Fig. 3(a). On the other hand, the performance gain of our system over the direct reciprocity system increases with the network size. In summary, our security system can efficiently improve the security performance of wireless networks.



(a)

(b)

(c)

(d)

Fig. 3. Population evolution of our security system in a network with $N = 5000$ nodes, whose transmission probability $\rho = 0.2$, for both the direct reciprocity system and the indirect reciprocity system, with different forgetting factor vectors (Λ) in the reputation updating process, the size of the action set $L = 3$. and the attacker identification rate $\sigma = 0.99$. (a) Percentage of the attacker population. (b) Percentage of the population using our desirable strategy. (c) Percentage of the population with the highest reputation. (d) Total network utility.

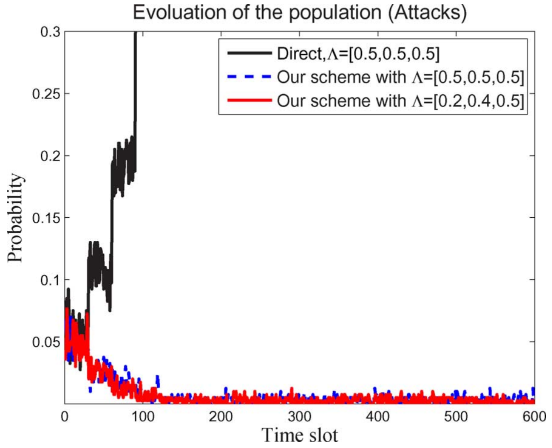


Fig. 4. Population evolution of attackers in our security system in a network with $N = 200$ nodes, and the other network settings are the same as Fig. 3.

Finally, it is indicated in Fig. 3 that compared with the direct reciprocity-based security scheme that actually collapses in such a network, our system significantly reduces the attacker population, increases the population using our desirable strategy, and improves the total network utility. Our system is more robust, because the indirect reciprocity game can better stimulate cooperation in networks, especially in the large-scale networks with mobility.

B. Case 2: Attack Classification

We now consider a more interesting case, where each node is able to perform jamming, Sybil, and spoofing attacks. In this case, $L = 5$ and the action set is $\{1, 2, 3, 4, 5\}$, corresponding to jamming, Sybil, spoofing, rejection of the request, and to follow the request.[3] The corresponding instant payoffs to the node itself and the transmitter are $G = [-10, -8, -6, -1, 10]$ and $C(\varpi = 1) = [7, 5, 3, 1, -0.5]$, respectively. Using (3) and (13), we obtain

$$Q^{5\times 5}(\varpi = 1) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 \\ 5 & 5 & 5 & 5 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} \tag{41}$$

and

$$\hat{\mathbf{a}}^5(\varpi = 1) = [4, 4, 4, 4, 5]^T. \tag{42}$$

We study the performance of attack classification, where the system evaluates the impacts of each type of attack and sets the value of the forgetting factor in the reputation updating process as $\Lambda = [0.1, 0.2, 0.3, 0.4, 0.5]$. As a benchmark, simulation results are provided for the strategy with the same forgetting factor for all the attacks, $\Lambda = [0.2, 0.2, 0.2, 0.4, 0.5]$. It is indicated in Fig. 5 that the attack classification can significantly reduce the population that launches the most serious attacks, just as expected. In addition, simulation results in Fig. 6 show that the proposed scheme can significantly increase the population that take our desirable actions, and reach the desirable network performance with a much faster speed.

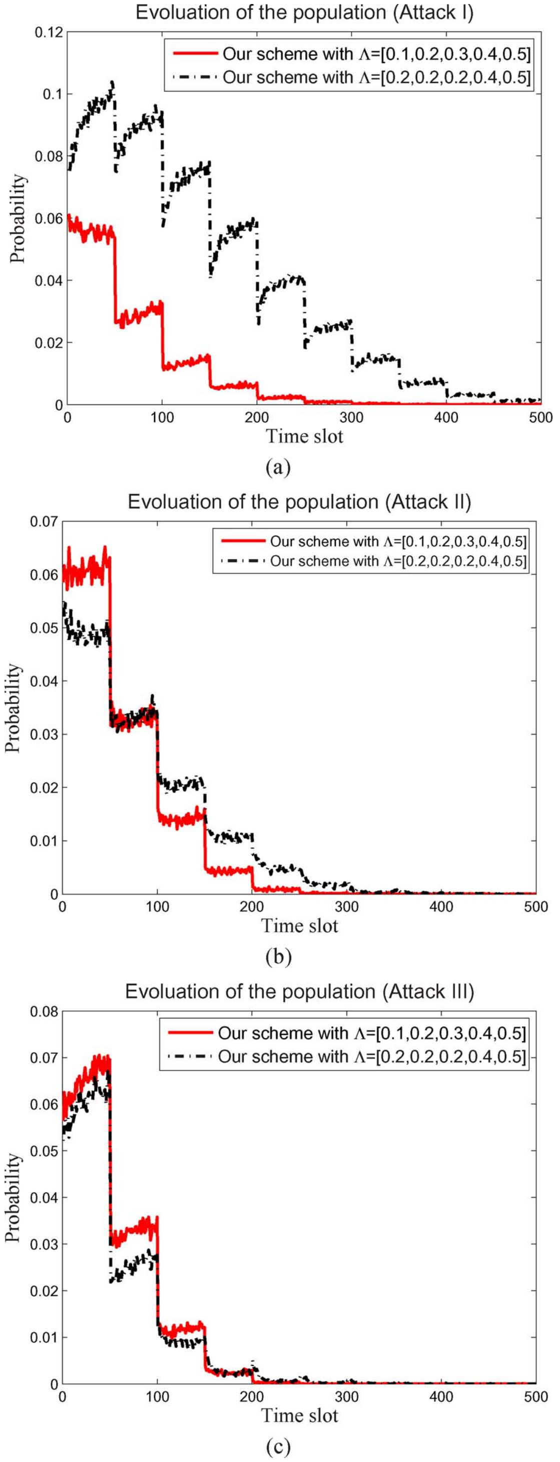
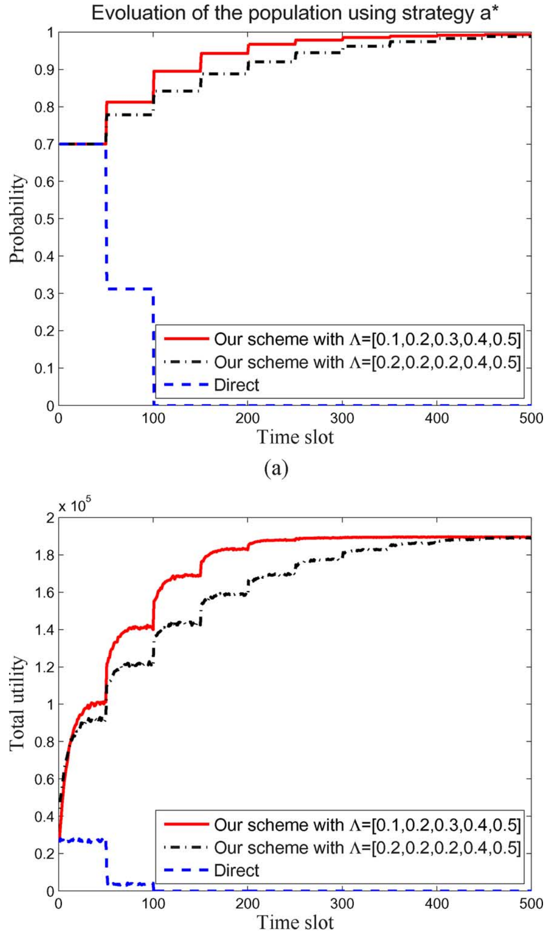


Fig. 5. Security performance of our system in a network with $N = 1000$ nodes, whose transmission probability $\rho = 0.2$, against three types of attacks, with $L = 5$ and $\sigma = 0.999$. Our scheme applies attack classification with forgetting factor vector $\Lambda = [0.1, 0.2, 0.3, 0.4, 0.5]$ in the reputation updating process, whereas the benchmark scheme does not apply attack classification with $\Lambda = [0.2, 0.2, 0.2, 0.4, 0.5]$. (a) Percentage of the jamming population. (b) Percentage of the Sybil-attacking population. (c) Percentage of the spoofing population.

Finally, we have also done simulations to validate the stability condition given by Theorem 1, with $G = [-12, -8, -6, -1, 10]$, $C(\varpi = 1) = [C_1, 5, 3, 1, -0.5]$,

[3] These attacks are very different from one another, and can be detected with different probabilities over different timescales. In reality, the model is completely agnostic to these distinctions. These are only representative labels that can be used to exemplify the kinds of behaviors that can be defended against.



(a)

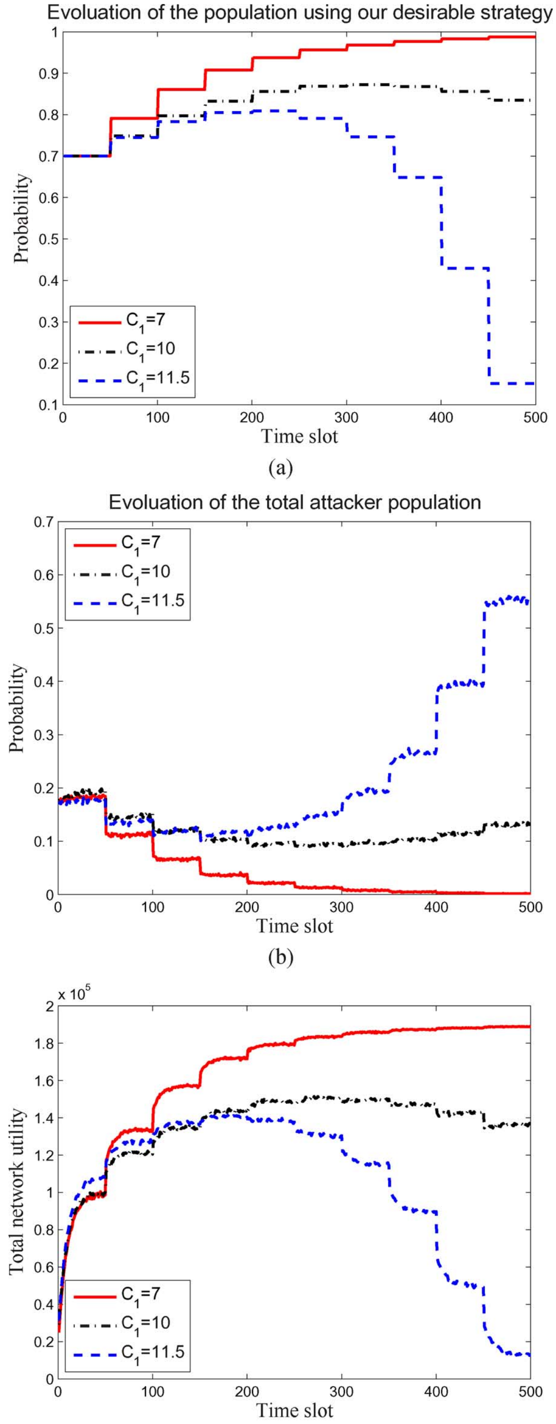
(b)

Fig. 6. Game-theoretic performance of our scheme in the same network scenario with Fig. 5. (a) Percentage of the population with the desired strategy. (b) Total network utility.

$\Lambda = [0.5, 0.5, 0.5, 0.5, 0.5]$, and the probability for a node to be the transmitter is 0.5. It is given by (19) that the system has to satisfy $C_1 \leq 10$. If the condition holds, e.g., $C_1 = 7$, our security system can efficiently improve the network performance and reduce the attacker population, as indicated in Fig. 7. Otherwise, if the condition does not hold, e.g., $C_1 = 11.5$, our desirable strategy no longer dominates over time and the network collapses in the end. In general, the network security performance improves, when the attacker obtains less payoff, C_1.

VI. Conclusion

In this paper, we have proposed an indirect reciprocity-based security system for large-scale wireless networks, and developed a social norm and a reputation updating process to address a wide range of attacks. The main idea is to suppress the attacker population by exploiting the requirement for network services by most users. The optimal action and the corresponding stationary reputation distribution are presented in the formulated game. We also provided the condition that our security can effectively defend the network, and showed that our system is robust against collusion attacks. In addition, simulation results show that our desirable action strategy is evolutionarily stable, i.e., the natural selection itself is sufficient to prevent adversary behaviors from invading. Our system can significantly



(a)

(b)

(c)

Fig. 7. Game-theoretic performance of our scheme in the same network scenario with Fig. 5, except the payoff to the transmitter, $G = [-12, -8, -6, -1, 10]$, the payoff to the node itself, $C(\varpi = 1) = [C_1, 5, 3, 1, -0.5]$, and $\Lambda = [0.5, 0.5, 0.5, 0.5, 0.5]$. (a) Percentage of the population with the desired strategy. (b) Percentage of the total attacker population. (c) Total network utility.

suppress a wide range of attacks in various network scenarios, and is much more robust than the system based on the direct reciprocity principle, especially in the large-scale networks with node mobility. For instance, our system can reduce the attacking rate from more than 5% to less than 0.05% in a network with 5000 nodes, whereas the system based on direct-reciprocity collapses in this case. Moreover, the discriminated forgetting factors in the reputation updating process has been shown to improve the efficiency to counteract attacks. Finally, we have verified the security gain of the attack classification and the derived stability condition for our security system via simulations.

REFERENCES

[1] K. J. R. Liu and B. Wang, *Cognitive Radio Networking and Security: A Game Theoretical View*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[2] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," in *Proc. 6th ACM Int. Symp. Mobile Ad Hoc Networking and Computing (MobiHoc '05)*, New York, 2005, pp. 46–57, ACM Press.

[3] D. Faria and D. Cheriton, "Detecting identity-based attacks in wireless networks using signalprints," in *Proc. ACM Workshop on Wireless Security*, 2006, pp. 43–52.

[4] J. R. Douceur, "The sybil attack," in *Proc. First Int. Workshop on Peer-to-Peer Systems (IPTPS)*, Mar. 2002, pp. 251–260.

[5] J. Newsome, E. Shi, D. Song, and A. Perrig, "The sybil attack in sensor networks: Analysis and defenses," in *Proc. Int. Symp. Information Processing in Sensor Networks (IPSN)*, Apr. 2004, pp. 259–268.

[6] W. Yu and K. J. R. Liu, "Attack-resistant cooperation stimulation in autonomous ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 12, pp. 2260–2271, Dec. 2005.

[7] D. Djenouri, L. Khelladi, and A. N. Badache, "A survey of security issues in mobile ad hoc and sensor networks," *IEEE Commun. Surveys Tutorials*, vol. 7, no. 4, pp. 2–28, 4th Quarter, 2005.

[8] Y. Sun, W. Yu, Z. Han, and K. J. R. Liu, "Information theoretic framework of trust modelling and evaluation for ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 2, pp. 305–317, Feb. 2006.

[9] Y. L. Sun, Z. Han, and K. J. R. Liu, "Defense of trust management vulnerabilities in distributed networks," *IEEE Commun. Mag.*, vol. 46, no. 2, pp. 112–119, Feb. 2008.

[10] Y. Sun, Z. Han, W. Yu, and K. J. R. Liu, "A trust evaluation framework in distributed networks: Vulnerability analysis and defense against attacks," in *Proc. IEEE INFOCOM*, Barcelona, Spain, 2006.

[11] W. Yu and K. J. R. Liu, "Game theoretic analysis of cooperation stimulation and security in antonomous mobile ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 5, pp. 459–473, May 2007.

[12] N. Zhang, W. Yu, X. Fu, and S. K. Das, "Maintaining defender's reputation in anomaly detection against insider attacks," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 3, pp. 597–611, Jun. 2010.

[13] B. Niu, H. V. Zhao, and H. Jiang, "A cooperation stimulation strategy in wireless multicast networks," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2355–2369, May 2011.

[14] E. Ayday, H. Lee, and F. Fekri, "Trust management and adversary detection for delay tolerant networks," in *Proc. IEEE Military Commun. Conf.*, San Jose, CA, 2010.

[15] M. Nowak and K. Sigmund, "Evolution of indirect reciprocity," *Nature*, vol. 437, pp. 1291–1298, Oct. 2005.

[16] L. Xiao, W. S. Lin, Y. Chen, and K. J. R. Liu, "Indirect reciprocity game modelling for secure wireless networks," in *Proc. Int. Conf. Communications (ICC)*, Ottawa, Canada, Jun. 2012.

[17] Y. Chen and K. J. R. Liu, "Indirect reciprocity game modelling for cooperation stimulation in cognitive networks," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 159–168, Jan. 2011.

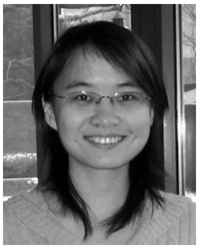
[18] R. Landa, D. Griffin, R. Clegg, E. Mykoniati, and M. Rio, "A sybilproof indirect reprociy mechansim for peer-to-peer networks," in *Proc. IEEE INFOCOM*, 2009, pp. 343–351.

[19] W. X. Z. Liu, H. Liu, and Y. Chen, "Exploiting jamming-caused neighbor changes for jammer localization," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 3, pp. 547–555, Mar. 2011.

[20] R. Fisher, *The Genetical Theory of Natural Selection*. Cambridge, U.K.: Cambridge Univ. Press, 1930.

[21] D. Blackwell, "Discounted dynamic programming," *Ann. Math. Statist.*, vol. 36, pp. 226–235, 1965.

[22] J. M. Smith, *Evolutionary and the Theory of Games*. Cambridge, U.K.: Cambridge Univ. Press, 1982.



Liang Xiao (M'09) received the B.S. degree in communication engineering from Nanjing University of Posts and Telecommunications, China, in 2000, the M.S. degree in electrical engineering from Tsinghua University, China, in 2003, and the Ph.D. degree in electrical engineering from Rutgers University, NJ, in 2009.

She is currently an Associate Professor in the Department of Communication Engineering, Xiamen University, Fujian, China. She was a visiting professor with the University of Maryland, College Park in 2011. From 2003 to 2004, she was with North Carolina State University, NC. Her research interests include wireless security, smart grids, and wireless communications.

Yan Chen (S'06) received the Bachelor's degree from University of Science and Technology of China in 2004, the M.Phil. degree from Hong Kong University of Science and Technology (HKUST) in 2007, and the Ph.D. degree from the University of Maryland College Park in 2011.

He is currently a research associate in the Department of Electrical and Computer Engineering at the University of Maryland College Park. His current research interests are in social learning and networking, smart grid, cloud computing, crowdsourcing, network economics, and multimedia signal processing.

Dr. Chen received the University of Maryland Future Faculty Fellowship in 2010, Chinese Government Award for outstanding students abroad in 2011, University of Maryland ECE Distinguished Dissertation Fellowship Honorable Mention in 2011, and was the Finalist of A. James Clark School of Engineering Dean's Doctoral Research Award in 2011.

W. Sabrina Lin (S'06) received the B.S. and M.S. degrees in electrical engineering from National Taiwan University in 2002 and 2004, respectively, and the Ph.D. degree with the Electrical and Computer Engineering Department, University of Maryland, College Park, in 2009.

Her research interests are in the area of information security and forensics, multimedia signal processing, and multimedia social network analysis.

Dr. Lin received the University of Maryland Innovation Award in 2011 and has coauthored the book *Behavior Dynamics in Media-Sharing Social Networks*, Cambridge Press, 2011.

K. J. Ray Liu (F'03) was named a Distinguished Scholar-Teacher of University of Maryland, College Park, in 2007, where he is the Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group conducting research encompassing broad areas of signal processing and communications with recent focus on cooperative communications, cognitive networking, social learning and networks, and information forensics and security.

Dr. Liu is the recipient of numerous honors and awards including IEEE Signal Processing Society Technical Achievement Award and Distinguished Lecturer. He also received various teaching and research recognitions from University of Maryland including university-level Invention of the Year Award; and Poole and Kent Senior Faculty Teaching Award and Outstanding Faculty Research Award, both from A. James Clark School of Engineering. An ISI Highly Cited Author in Computer Science, he is a Fellow of AAAS. He is President of IEEE Signal Processing Society where he has served as Vice President – Publications and on the Board of Governors. He was the Editor-in-Chief of *IEEE Signal Processing Magazine* and the founding Editor-in-Chief of EURASIP JOURNAL ON ADVANCES IN SIGNAL PROCESSING.