Indirect Reciprocity Game Modelling for Secure Wireless Networks

Liang Xiao *[†], W. Sabrina Lin [†], Yan Chen [†], K. J. Ray Liu [†]

* Department of Communication Engineering, Xiamen Univ., 361005 China.

[†] Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742.
E-mail:{lxiao@xmu.edu.cn,wylin@umd.edu,yan@umd.edu,kjrliu@umd.edu}

***Abstract*—We formulate the wireless security problem as an indirect reciprocity game, and propose a security mechanism that applies the indirect reciprocity principle to suppress attacks in wireless networks. In this system, a large number of nodes cooperate to reject the network access requests from attackers during the punishment periods. If the punishment time is so long that the cost due to the loss of network services exceeds the illegal security gains of the attack, rational nodes do not have incentive to attack, and hence our system can reduce the attacking probability in the network. We develop a social norm and reputation updating process to build such an indirect reciprocity mechanism for the network. We evaluate the evolutionarily stable strategy (ESS) in the game, and provide the optimal action strategy and its corresponding stationary reputation distribution. Our system is robust against collusion attacks, and can significantly reduce the attacking rate for a wide range of attacks. Simulation results show that our system has much better security performance than the direct reciprocity mechanism, especially in the large-scale wireless network with terminal mobility. Our system can be applied to many wireless networks including cognitive radio networks to improve their security performance.**

I. Introduction

With the development of cognitive radios, users are gaining autonomous and control in their radio transmissions, and hence obtain more power and freedom to attack the networks. Current wireless networks are threatened by a wide range of attacks, such as jamming, spoofing, Sybil attacks, and many relay-related attacks [1]–[3]. Attackers obtain illegal security advantages if not being caught and punished. Extensive works have been done to investigate their impacts on the network performance, and many detection and localization algorithms have been proposed to identify the attackers. Although most existing works assume that attackers do not concern their throughput performance, in practice, network access is highly desirable for most users, including many potential attackers. Therefore, we exploit the requests for network service to decrease the attacking rate in wireless networks.

With the trust modelling and evaluation method proposed in [4], the trust/reciprocity mechanism is a powerful tool to improve the security and stimulate cooperations in wireless networks [1], [2], [5]–[7]. These works apply the direct reciprocity principle, whose main idea is "I help you because you helped me" [8]. More specifically, each node chooses its actions according to the past interactions with the opponents, and is more likely to decline the requests from those who have ever hurt it. However, in a large-scale network with terminal mobility, many nodes have a small chance to meet their opponents before, and hence only have limited or outdated knowledge on their past actions. In such a case, attackers have small chance to meet their victims again and to be punished by them, and hence nodes have incentive to take risks to attack the network.

This problem can be addressed by the use of the indirect reciprocity principle, first developed in social science and evolutionary biology. The main idea is "I help you and somebody else helps me" [8]. The indirect reciprocity game is promising to stimulate cooperation in cognitive networks [9], and can be used to improve the Sybil-resistance for the accounting of peer contributions in peer-to-peer networks [10].

In this paper, we formulate an indirect reciprocity security game, and propose a security system to counteract 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. The reputation propagation mechanism in the indirect reciprocity system allows attackers to be known and punished by a larger population of nodes in the network. Compared with the direct reciprocity system, our system provides stronger security protection, especially for the large-scale networks with node mobility.

We assume that multiple transmissions can take place at the same time in a large-scale network without interfering with each other. In each transmission period, the intended receiver and other observing nodes evaluate the behavior of each node in this area, reduce the reputations of the attackers, and propagate the new reputations to the whole network through the gossip channels. More specifically, we build a public social norm and reputation updating process to assign to the attackers low (bad) reputations, due to which most nodes reject their requests for network service over a long time. We

The work by Xiao is partly supported by NSFC (No.61001072), the Natural Science Foundation of Fujian Province of China (No.2010J01347), SRF for ROCS, SEM, IBM OCR, Fundamental Research Funds for the Central Universities (2011121050), and Tsinghua-Qualcomm joint research center.

978-1-4577-2053-6/12/$31.00 ©2012 IEEE

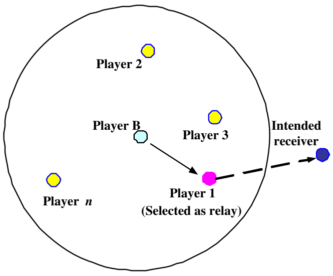


Fig. 1. A communication topology in the game formulation, including a transmitter called Player B, an intended receiver, and n nodes in the communication region, called Type-A players (including Player 1 to Player n), with some Type-A players selected to relay.

can assign different values of reputation to the attackers and punish them in different manners, according to their impacts on the network performance. The system is designed to make the cost to each attacker in the punishment duration exceed its illegal security gain of attacks. Thus rational nodes in the network are stimulated to deviate from the attacking behavior for their own interests. Our system is promising to improve the security performance of many wireless networks including cognitive radio networks.

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. We then analyze its performance in Section IV, and show the simulation results in Section V. Finally, we conclude in Section VI.

II. System Model

We consider a homogenous wireless network, consisting of N randomly located nodes. Each node leaves the network with probability $1-\delta$ and (re)enters it with probability δ. We assume that N_s out of the N nodes are allowed to send messages to their intended receivers at the same time, without interfering with each other. As shown in Fig. 1, each transmission includes a transmitter called Player B, an intended receiver, and n Type-A nodes in the area (called Player i, with $1 \leq i \leq n$). For simplicity of presentation, we assume a constant n for each transmission and let $N_s = N\rho$, where the transmission probability of each node, ρ, decreases with the network geographical density. Note that our system is not restricted to the case with fixed N_s and n.

In each transmission, Player B sends a message to the receiver, while the latter and the observing nodes monitor the transmission and evaluate the behavior of each Type-A node. The transmitter can request a node to relay the message and others to keep silence, according to the network topology, radio channel conditions, and node reputations. The indicator function $Ind_i[k]$ denotes the decision of the relay selection algorithm, where $Ind_i[k] = 1$, if Player i is selected to relay at time k, and zero otherwise. In other words, to follow the request, Player i has to relay the message at time k as $Ind_i[k] = 1$, and keep silence as $Ind_i[k] = 0$. For simplicity, we assume perfect radio propagation in this area, i.e., the transmission is successful if all the Type-A nodes follow the requests from Player B.

Action ID	Physical action
1	Type 1 attacks, e.g., jamming
2	Type 2 attacks, e.g., spoofing
$\cdots$	...
L	Type L attacks
$L+1$	Decline the request by Player B
$L_t(=L+2)$	Follow the request by Player B

TABLE I
Action Set of Type-A players.

We consider a wide range of attacks, such as jamming, spoofing, Sybil attacks, deliberate packet dropping, and collusion attacks. Each Type-A node can perform one of them in each time slot. In the literature, multiple attack detection and attacker identification algorithms have been proposed for each type of attacks. Our goal is to design a security system that can incorporate most existing related algorithms. All the attacks are classified into L levels, according to their impacts on the network performance and their costs to the attackers. Without loss of generality, assume that Level-g attacks are more dangerous to the network than Level-h attacks, if $g < h$. For example, jamming can be labelled with a lower attack level than spoofing. We use $p_D(i)$ to denote the probability that Level-i attack is correctly detected and the attacker is accurately identified.

In this way, each transmission process is formulated into one round of the indirect reciprocity game with $n+1$ players: Player B, and Player 1 to Player n. Player B sends a message and selects a subset of the Type-A nodes as its relay. The actions of Player i at time k is denoted with $A_i[k]$, with $1 \leq i \leq n$. As shown in Table I, its action set has $L_t(=L+2)$ actions, i.e., $A_i[k] \in \{1, 2, \cdots, L_t\}$, where the element $j \in [1, L]$ represents Level-j attacks. The node follows Player B's request as $A_i[k] = L_t$, and declines it when $A_i[k] = L+1$.

Note that each action in Table I can correspond to different communication states of a node, based on whether it is selected as a relay. For example, the action that Player i stops transmission corresponds to $A_i[k] = L_t$, for a non-relay node with $Ind_i[k] = 0$; while the same communication state becomes $A_i[k] = L+1$, as $Ind_i[k] = 1$. Similarly, the payoff of an action to Player i also depends on Ind_i. For example, the action $A_i[k] = L_t$ costs Player i more energy as $Ind_i[k] = 1$, compared with the non-relay case.

Let $C_{Ind_i}[j]$ (or $G[j]$) denote the instant payoff of the action $A_i[k] = j$ to Player i (or Player B), with $1 \leq j \leq L_t$ and $1 \leq i \leq n$, ignoring the actions of the other Type-A nodes. A positive payoff value indicates that the node actually gains from its action, while a negative value represents a loss. Due to the efforts to attack the network or to decline the request, a rational node never chooses them unless receiving non-negative payoffs, and hence $C_{Ind}[j] \geq 0$ for $j < L_t$. In such a case, Player B does not benefit due to the security or throughput loss, indicating $G[j] \leq 0$ for $j < L_t$. For the

N	Number of nodes in the network
n	Number of Type-A players in a transmission
N_s	Number of transmitters at each time
L_t	Size of the action set
L	Number of attack levels
$Q = [Q_{i,j}]_{L_t \times L_t}$	Social norm
$\mathbf{p}_l = [p_1, \cdots, p_{L_t}]^T$	Reputation vector of Player l
a^*	Optimal action strategy
Λ	Forgetting factor vector
Φ	Reputation propagation matrix
$\mathbf{p}^* = [p_1^*, \cdots, p_{L_t}^*]^T$	Stationary reputation distribution
p_D	Prob. to successfully identify an attacker
$C_{Ind}[i]$	Payoff of action i to Player A
$G[i]$	Payoff of action i to Player B

TABLE II
SUMMARY OF SYMBOLS AND NOTATIONS.

case $j = L_t$, the transmission of Player B benefits, indicating $G[L_t] > 0$. In this case, a relay node consumes transmission energy while a non-relay node does not. Thus the cost to the former is higher, with $C_1[L_t] < C_0[L_t] \leq 0$.

The transmission performance of Player B actually depends on the actions of all its n neighbors. Thus the payoff to Player B at time k, denoted as $U_B[k]$, is a function of the actions of all the n nodes at that time. For simplicity, we assume it as the minimum instant payoffs due to their actions, i.e.,

$$U_B[k] = \min_{i=1,\cdots,n} G[A_i[k]]. \tag{1}$$

That means, the transmission performance is determined by the worst neighbor. For example, the transmission is assumed to fail, if any node refuses to follow its request by dropping the message as $Ind = 1$, or transmitting as $Ind = 0$. In another example, an attacker can ruin the whole transmission even if all the other nodes follow the requests.

Let $U_i[k]$ denote the payoff of $A_i[k]$ to Player i, and assume it to be independent of $A_j[k]$, for any $j \neq i$. Hence we have

$$U_i[k] = C_{Ind_i[k]}[A_i[k]]. \tag{2}$$

For ease of reference, the commonly used notations are summarized in Table II.

III. A Security System Based on Indirect Reciprocity

We design a security system that applies the indirect reciprocity principle to reduce the attacker population in wireless networks. This system maintains a reputation mechanism, where each node obtains a reputation vector according to its past and current actions. Each time slot consists of the message transmission stage and the performance evaluation stage. In the second stage, the reputation of each Type-A player is updated and broadcast to the other nodes via the gossip channels. The network determines whether to accept its future transmission requests based on its reputation. In general, the reputation of a node decreases, if it attacks the network or declines the request from a transmitter with a good reputation. In order to punish attackers, the reputation of the node decreases, if it helping a node with a bad reputation. Its reputation improves in other cases.

We design a reputation updating process to compute the reputation vector for each node, based on its last reputation and the instant reputation resulting from its current action. The forgetting factor is used to weight the last reputation in the calculation, and is related to the instant reputation. In this system, attackers are not only rejected by their direct victims, but also by most other nodes in the network during a long time. The punishment period is determined by the forgetting factors. The nodes are forgiven and regains the network access, if following the network social norm during the punishment period.

We use a reputation set $R = \{1, \cdots, L_t\}$. Nodes are encouraged to help the transmitters with higher reputation values. Denote the reputation vector assigned to Player i as $\mathbf{p}_i = [p_1, p_2, \cdots, p_{L_t}]^T$, whose l-th element can be approximately viewed as its probability to have a scalar reputation l. We have $0 \leq p_l \leq 1$ and $\sum_l p_l = 1$. Actually, Player i also has a scaler reputation for given time slot, $j \in [1, L_t]$, as a realization of an integer random variable whose probability mass function (PMF) is $\mathbf{p}_i$. In general, Player i is more likely to have a higher scaler reputation, if the random variable whose PMF is $\mathbf{p}_i$ has a larger expectation.

During each transmission, the reputations of all Type-A nodes are updated based on the same criterion, called social norm. Denote the social norm as $Q = [Q_{i,j}]_{L_t \times L_t}$, whose element $Q_{i,j}$ is the instant scalar reputation assigned to the node who takes action $i \in \{1, \cdots, L_t\}$ towards a transmitter whose scalar reputation is $j \in \{1, \cdots, L_t\}$.

In order to suppress attacks, the system assigns a high reputation value to the nodes who help the good nodes and to those who refuse to help the nodes with bad reputations. In addition, to avoid the network performance downgrade or collapse, the system gives a low reputation to each attacker, even if the transmitter has a bad reputation. In this system, our desirable action towards a transmitter with a low reputation is to refuse its relay requests and keep silence.

Following the above principles, we build the social norm. For the relay node (i.e., $Ind = 1$), the social norm is given by

$$Q\mid_{Ind=1} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L & L & \cdots & L & L \\ L_t & L_t & \cdots & L_t & L+1 \\ 1 & 2 & \cdots & L+1 & L_t \end{bmatrix}. \tag{3}$$

For a non-relay node, however, the action $i = L+1$ means that this node and Player B transmit at the same time, and can lead to the undesirable packet collision. Thus we assign $L+1$ as its reputation, even if Player B has a bad reputation. Similarly, our desirable action for the non-relay nodes is always $i = L_t$ to avoid packet collision, and the social norm is chosen as

$$Q\mid_{Ind=0} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L & L & \cdots & L & L \\ L+1 & L+1 & \cdots & L+1 & L+1 \\ L_t & L_1 & \cdots & L_t & L_t \end{bmatrix}. \tag{4}$$

Without loss of generality, we assume that Player l chooses its action at time k based on its own scalar reputation i and Player B's scalar reputation j, with $A_l[k] = a_{i,j}$, $1 \leq l \leq n$. As mentioned, the scalar reputations $i, j \in \{1, \cdots, L_t\}$ are derived from $\mathbf{p}_l[k]$ and $\mathbf{p}_B[k]$, respectively. There are $L_t^{L_t^2}$ possible action strategies for each node under all the scenarios. Denote our desirable action strategy as $a^* = [a^*_{i,j}]_{L_t \times L_t}$. Following the same principles as the social norm, we hope each relay node to relay for the good nodes, and drop packets otherwise, and hence have

$$a^* \mid_{Ind=1} = \begin{bmatrix} L+1 & L+1 & \cdots & L+1 & L_t \\ L+1 & L+1 & \cdots & L+1 & L_t \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L+1 & L+1 & \cdots & L+1 & L_t \end{bmatrix}. \quad (5)$$

On the other hand, no matter what the reputation of the transmitter has, we hope the non-relay nodes to keep silence, i.e., to follow the request, in order to avoid packet collision. Thus we have

$$a^* \mid_{Ind=0} = \begin{bmatrix} L_t & L_t & \cdots & L_t \\ \cdots & \cdots & \cdots & \cdots \\ L_t & L_t & \cdots & L_t \end{bmatrix}. \quad (6)$$

During this time slot, the reputation vectors of Player B, the intended receiver and the nodes outside this area do not change, with $\mathbf{p}[k+1] = \mathbf{p}[k]$. On the other hand, the observing nodes in this area monitor the action of each Type-A player and assign an instant reputation to it accordingly. The new reputation of Player l, $\mathbf{p}_l[k+1]$ is updated, according to both the instant reputation and the last reputation $\mathbf{p}_l[k]$. As mentioned, we have derived the scale reputation, i and j, for Player l and Player B, respectively, at time k. Based on the public social norm and reputation updating process, all the nodes in the network agree with the reputation of the nodes.

More specifically, we assign for each new node a high initial reputation, $\mathbf{p}[0] = [0, 0, \cdots, 0, 1]^T$. The forgetting factor Λ_x, which is used to weight the last reputation vector $\mathbf{p}[k]$, is a function of x, the immediate scalar reputation due to the current action. Letting $0 < \Lambda_x < 1$ for each reputation level (x), we use the following process to update the reputation of Player l,

$$\begin{aligned} \mathbf{p}_l[k+1] = {} & \Phi \Lambda_{Q_{a_{i,j},j|Ind}} \mathbf{p}_l[k] \\ & + \Phi \left(1 - \Lambda_{Q_{a_{i,j},j|Ind}}\right) \mathbf{e}_{Q_{a_{i,j},j|Ind}}, \end{aligned} \quad (7)$$

where $\mathbf{e}_x$ is the standard basis vector. The propagation error matrix of reputation, Φ, includes both the detection error and the propagation error due to imperfect gossip channels. For simplicity, we use p_{Di} to denote the probability for the i-th action to be correctly identified and broadcast to the network, and model Φ as

$$\Phi = \begin{bmatrix} p_{D1} & \frac{1-p_{D1}}{L_t-1} & \frac{1-p_{D1}}{L_t-1} & \cdots & \frac{1-p_{D1}}{L_t-1} \\ \frac{1-p_{D2}}{L_t-1} & p_{D2} & \frac{1-p_{D2}}{L_t-1} & \cdots & \frac{1-p_{D2}}{L_t-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1-p_{DL_t}}{L_t-1} & \frac{1-p_{DL_t}}{L_t-1} & \frac{1-p_{DL_t}}{L_t-1} & \cdots & p_{DL_t} \end{bmatrix}. \quad (8)$$

Actually, our system is not restricted to (8), and can be easily extended to the other models.

IV. Performance Analysis

A. Evolutionarily Stable Strategy

An evolutionarily stable strategy (ESS) cannot be invaded by any alternative strategy that is initially rare, and the natural selection alone is sufficient to prevent alternative strategies from invading [11]. To evaluate the ESS property of our desirable action strategy, we assume the Wright-Fisher model [11] 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

$$y_i[k+1] = \frac{y_i[k] U_i[k]}{\sum_l y_l[k] U_l[k]}. \quad (9)$$

B. Optimal Action & Stationary Reputation Distribution

Now we evaluate the optimal action strategy and the corresponding stationary reputation distribution. In a simplified scenario, Player A who has a scalar reputation i and a vector reputation e_i meets a transmitter whose vector reputation is e_j (thus its scalar reputation is j) in a time slot, and we assume for now a fixed forgetting factor $\Lambda_x = \lambda$ for any reputation x.

The optimal strategy $a^* = [a^*_{i,j}]_{L_t \times L_t}$ is defined as the strategy that maximizes $U_{i,j}$, the long-term payoff for Player A. Let $r_{i,j}(k)$ denote the probability that the scalar reputation of Player A changes from i to k, when meeting Player B with reputation j. Assume for simplicity that each node other than Player B and the intended receiver is randomly and independently selected as relay with probability p_s. By (7), following the derivation similar to that in [9], we have

$$\begin{aligned} \mathbf{r}_{i,j} & = [r_{i,j}(1), \cdots, r_{i,j}(L_t)]^T \\ & = \Phi \left(\lambda \mathbf{e}_i + (1-\lambda) \left(p_s \mathbf{e}_{Q_{a_{i,j},j|1}} + (1-p_s) \mathbf{e}_{Q_{a_{i,j},j|0}}\right)\right). \end{aligned} \quad (10)$$

As each node initiates a transmission with probability ρ, we can write the Bellman equation of $U_{i,j}$ as (11), where the discounting factor of the future is set to be δ, the probability for a node to stay or (re)enter the network, and the optimal action $a^*_{i,j}$ is given by

$$a^*_{i,j} = \arg\max_{a_{i,j}} U_{i,j}. \quad (12)$$

We now consider the stationary reputation distribution $\mathbf{p}^* = [p^*_1, \cdots, p^*_{L_t}]^T$ for a given optimal action strategy a^*. Since $\mathbf{p}^*$ does not change over time k, by (7), we have (13). Hence the optimization of the optimal actions can be formulated as a Markov Decision Process (MDP) and we can apply the modified value iteration algorithm in [9] and use (10)-(13) to obtain the optimal action and the stationary reputation distribution.

$$U_{i,j} = \max_{a_{i,j}} \left[(1-\rho) \left(p_s C_1[a_{i,j|1}] + (1-p_s) C_0[a_{i,j|0}] + \delta \sum_k \sum_l r_{i,j}(k) p_l^* U_{k,l} \right) + \rho \left(G[a^*_{j,i}] + \delta \sum_l p_l^* U_{i,l} \right) \right]. \quad (11)$$

$$\mathbf{p}^* = \Phi \left(\lambda \mathbf{p}^* + (1-\lambda) \left(p_s \left[\begin{array}{l} \sum_{i=1}^{L_t} \sum_{l:Q_{a^*_{i,l},l|1}=1} p_i^* p_l^* \\ \cdots \\ \sum_{i=1}^{L_t} \sum_{l:Q_{a^*_{i,l},l|1}=L_t} p_i^* p_l^* \end{array} \right] + (1-p_s) \left[\begin{array}{l} \sum_{i=1}^{L_t} \sum_{l:Q_{a^*_{i,l},l|0}=1} p_i^* p_l^* \\ \cdots \\ \sum_{i=1}^{L_t} \sum_{l:Q_{a^*_{i,l},l|0}=L_t} p_i^* p_l^* \end{array} \right] \right) \right). \quad (13)$$

C. Security Analysis

In our system, a rational node has incentive to help good nodes and punish attackers, in order to obtain a higher reputation and hence better network services. On the other hand, an attacker receives a bad reputation and loses its network access during a long time. The punishment strength/duration is adjusted by the forgetting factors in the reputation updating process. If the punishment cost is large enough, the rational nodes are motivated to abandon attacks.

Since collusion attacks are highly dangerous to the reputation-based networks, we take it as an example. First, we consider a case where several nodes collude to spread false information on another node, in hopes of ruining its reputation and hence its future transmission. In such a case, the victim node can find its reputation loss and hence sends alarm to initiate investigations on the colluders. We set a very low reputation value to the colluders, and assign a small forgetting factor to enable a longer punishment. In this way, the reputations of the attackers significantly drop, and the network stops serving them over a very long time. Thus the colluders cannot benefit by this attack.

Next, we consider another case, where several nodes try to improve each others' reputation by reporting faked internal cooperations. Although the gaining nodes never complain, the system reduces the reputation of the nodes who have obtained abnormal amounts of internal helps. Hence the incentive to falsely promote each others' reputations is reduced, and our system can resist this type of collusion attacks.

V. Simulation Results

We evaluate the performance of the proposed security system with simulations for a simple case that views all types of attacks as the same, i.e., $L = 1$. The action set is $\{1, 2, 3\}$, representing attacks, request rejection, and to follow the request of the transmitter, respectively. For simplicity, we assume that all the non-relay nodes are outside the signal coverage area of the transmitters, and hence only consider the relay nodes in each round of the game with $p_s \equiv 1$ and $Ind \equiv 1$. The instant payoff of the actions to Player B and Player A are given by $G = [-8, -1, 10]$ and $C = [5, 1, -0.5]$, respectively. The social norm Q and the desirable action rule a^* are given respectively by

$$Q^{3\times 3} = \left[\begin{array}{ccc} 1 & 1 & 1 \\ 3 & 3 & 2 \\ 1 & 2 & 3 \end{array} \right], \quad (14)$$

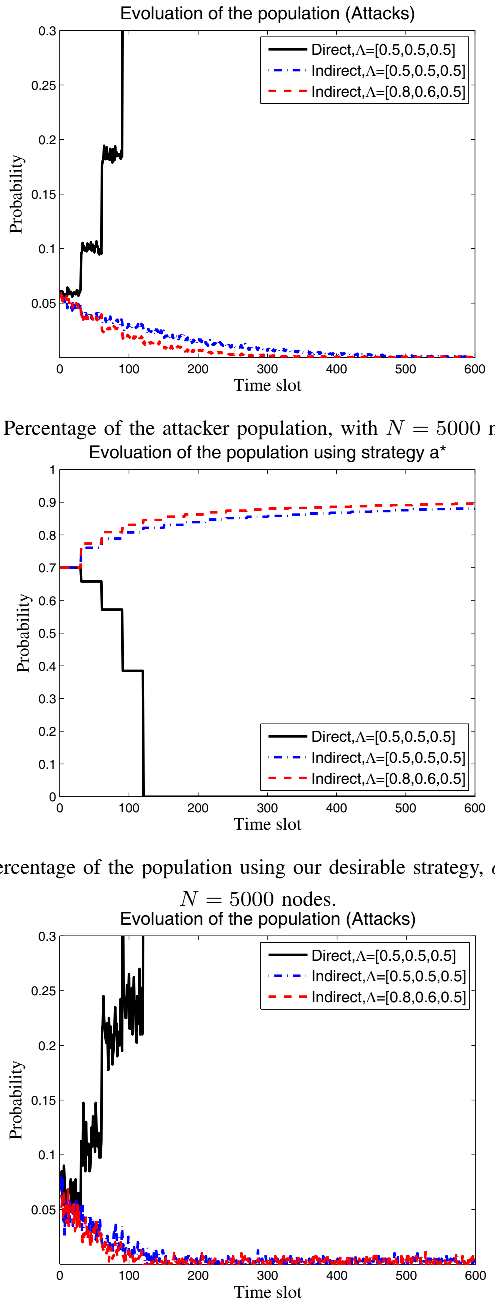
and

$$a^* = \left[\begin{array}{ccc} 2 & 2 & 3 \\ 2 & 2 & 3 \\ 2 & 2 & 3 \end{array} \right]. \quad (15)$$

In the simulations, we assume $N = 5000$, $\rho = 0.2$, $\delta = 0.9$, and $p_D = 0.99$ for each action, if not specified otherwise. Simulation results show that the stationary reputation distribution corresponding to a^* is $x^* = [0.0128, 0.0128, 0.9744]$, as $\Lambda = [0.5, 0.5, 0.5]$, indicating that most users can obtain the best scalar reputation.

Next, we evaluate the evolutionarily stable property of a^*, assuming that at the beginning of each process in the simulation, 70% of the nodes choose a^*, while the other nodes randomly select from the remaining strategies. We also assume that the action spreading follows the Wright-Fisher model, with the population selecting each strategy given by (9). As comparison, we consider another security system based on direct reciprocity, where each node chooses its actions according to its past interaction with its opponents. As shown in Fig. 2, the direct reciprocity-based scheme fails to work in the networks with $N = 5000$ or $N = 200$, and the network corrupts shortly after an eruption of attacks. The reason is that as nodes are unlikely to meet each other again in the large-scale network, the long-term cost to an attacker is small compared with its illegal security gain in the direct reciprocity game.

Fortunately, our system can address this problem and efficiently suppress attacks in the large-scale network. For example, our system reduces the attacker population from more than 5% to less than 0.05% after 400 time slots, as in Fig. 2 (a). It is also shown that the forgetting factor vector $\Lambda = [0.8, 0.6, 0.5]$ has a better security performance than that of $\Lambda = [0.5, 0.5, 0.5]$. The former assigns a larger weight to the instant bad reputation due to attacks in the reputation updating process, and thus punishes the attackers with longer punishment, resulting to the significant drop of the attacker population. It is also shown in Fig. 2 (b) that more than 90% of the population chooses the desirable strategy, shortly after the start of the process in our system. Finally, Fig. 2 (c) shows that our system has a better performance in the network with a smaller scale, where the attacking rate decreases much faster than that in Fig. 2 (a). On the other hand, the performance gain of our system over the system based on direct reciprocity increases with the network size. In summary, our security system can efficiently improve the security performance of wireless networks.



(a) Percentage of the attacker population, with $N = 5000$ nodes.

(b) Percentage of the population using our desirable strategy, a^*, with $N = 5000$ nodes.

(c) Percentage of the attacker population, with $N = 200$ nodes.

Fig. 2. The population evolution of our security system in a network with N 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.

VI. Conclusion & Future Work

We have proposed an indirect reciprocity-based security system for wireless networks, which exploits the requirement of network access by most users to motivate them not to attack. We build a social norm and reputation updating process to punish the attackers and to encourage node cooperation. We have shown that our desirable action strategy is evolutionarily stable, and our system can address a wide range of attacks, including collusion attacks and many others. Simulation results show that our system can significantly suppress attacks and is much more robust than the system based on the direct reciprocity principle, especially in the large-scale networks with node mobility. For example, our system can reduce the attacking rate from more than 5% to less than 0.05% in a network with 5000 nodes, while the system based on direct-reciprocity collapses.

Moving forward, further investigation is needed to evaluate the impacts of the attacking classification on the security performance of our system. Another interesting topic is to study the condition for our desirable strategy to be evolutionarily stable. Finally, we are working to thoroughly evaluate the performance of our system under a wide range of network scenarios against many types of attacks.

References

[1] W. Yu and K.J.R. Liu, "Attack-resistant cooperation stimulation in autonomous ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 2260–2271, Dec. 2005.

[2] 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.*, pp. 459–473, May 2007.

[3] D. Djenouri, L. Khelladi, and A.N. Badache, "A survey of security issues in mobile ad hoc and sensor networks," *IEEE Communications Surveys and Tutorials*, vol. 7, pp. 2 – 28, 2005.

[4] 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, pp. 305–317, Feb. 2006.

[5] N. Zhang, W. Yu, X. Fu, and S.K. Das, "Maintaining defender's reputation in anomaly detection against insider attacks," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 40, pp. 597 – 611, June 2010.

[6] B. Niu, H.V. Zhao, and H. Jiang, "A cooperation stimulation strategy in wireless multicast networks," *IEEE Trans. Signal Processing*, vol. 59, pp. 2355 – 2369, May 2011.

[7] E. Ayday, H. Lee, and F. Fekri, "Trust management and adversary detection for delay tolerant networks," in *Proc. IEEE Military Comm. Conf.*, 2010.

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

[9] Y. Chen and K.J.R. Liu, "Indirect reciprocity game modelling for cooperation stimulation in cognitive networks," *IEEE Trans. Comm.*, vol. 59, pp. 159–168, Jan. 2011.

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

[11] R. Fisher, *The Genetical Theory of Natural Selection*, Cambridge University Press, 1930.