Computer Networks 71 (2014) 63–83

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

A game theoretic approach to detect and co-exist with malicious nodes in wireless networks ☆

Wenjing Wang [a], Mainak Chatterjee [b,*], Kevin Kwiat [c], Qing Li [a]

[a] Blue Coat Systems, Inc., Sunnyvale, CA, United States
[b] EECS, University of Central Florida, Orlando, FL, United States
[c] Air Force Research Laboratory, Rome, NY, United States

ARTICLE INFO

Article history:
Received 12 February 2013
Received in revised form 19 May 2014
Accepted 14 June 2014
Available online 26 June 2014

Keywords:
Malicious node
Game theory
Coexistence
Bayesian games
Markov Bayes–Nash Equilibrium

ABSTRACT

Identification and isolation of malicious nodes in a distributed system is a challenging problem. This problem is further aggravated in a wireless network because the unreliable channel hides the actions of each node from one another. Therefore, a regular node can only construct a belief about a malicious node through monitoring and observation. In this paper, we use game theory to study the interactions between regular and malicious nodes in a wireless network. We model the malicious node detection process as a Bayesian game with imperfect information and show that a mixed strategy perfect Bayesian Nash Equilibrium (also a sequential equilibrium) is attainable. While the equilibrium in the detection game ensures the identification of the malicious nodes, we argue that it might not be profitable to isolate the malicious nodes upon detection. As a matter of fact, malicious nodes can co-exist with regular nodes as long as the destruction they bring is less than the contribution they make. To show how we can utilize the malicious nodes, a post-detection game between the malicious and regular nodes is formalized. Solution to this game shows the existence of a subgame perfect Nash Equilibrium and reveals the conditions that are necessary to achieve the equilibrium. Further, we show how a malicious node can construct a belief about the belief held by a regular node. By employing the belief about the belief system, a Markov Perfect Bayes–Nash Equilibrium is reached and the equilibrium postpones the detection of the malicious node. Simulation results and their discussions are provided to illustrate the properties of the derived equilibria. The integration of the detection game and the post-detection is also studied and it is shown that the former one can transit into the latter one when the malicious node actively adjusts its strategies.

1. Introduction

In a distributed wireless system where multiple nodes work towards individual or common goals, cooperative behavior among the nodes (such as controlling the transmit power level, reducing interference for each other, revealing private information, adhering to network policies) is highly desired for increasing system capacity. Though this desirable property makes it easy to analyze a system due to state space reduction; in reality, this assumption might be too strong. For example, there might be entities in the network (also called nodes) which might act in a selfish manner. These selfish nodes, governed by their utility function, care about their own payoffs and

☆ Approved for Public Release; Distribution Unlimited: 88ABW-2014-2735 dated 05 June 2014.

* Corresponding author. Tel.: +1 4078235793.
E-mail addresses: wenjing.wang@bluecoat.com (W. Wang), mainak@eecs.ucf.edu (M. Chatterjee), kevin.kwiat@rl.af.mil (K. Kwiat), qing.li@bluecoat.com (Q. Li).

http://dx.doi.org/10.1016/j.comnet.2014.06.008
1389-1286/

choose corresponding strategies to maximize them. Usually, the payoffs are the benefits a node can derive from other nodes or the network. However, it is possible that there are some nodes whose objective is to cause harm and bring disorder to the network. These nodes, referred as *malicious* nodes, do not reveal their identities while disrupting network services. The objective of such malicious nodes is to maximize the damage before they are detected and isolated. They are also rational, and their payoff is determined by the amount of damage they cause to the network.

In order to minimize the impact of the malicious nodes, detection mechanisms need to be in place. Thus, a regular node should monitor its surroundings and distinguish a malicious node from a regular one. However, the detection process has challenges. First, active monitoring can be costly. To identify malicious behaviors, a regular node has to listen to the channel and/or process the information sent by the nodes being monitored. These monitoring activities consume resources and hence, an "always on" monitoring scheme is not efficient even if plausible. Second, the malicious node can disguise itself. To reduce the probability of being detected, a malicious can behave like a regular node and choose longer intervals between attacks. Third, the randomness and unreliability of the wireless channel bring more uncertainty to the monitoring and detection process.

In spite of the above challenges, mechanisms to detect malicious nodes can always be designed. However, the important question is 'what should the regular node do upon detecting a malicious node?' Though the reasonable response would be to immediately isolate the malicious node, there might be situations where malicious nodes can be kept and made use of. The most straightforward reason for the coexistence is that a malicious node has no idea whether it has been identified or not, and it will continue to operate like a regular node to avoid detection. During this time, i.e., when the malicious node cooperates in disguise, it can be exploited for normal network operations. This "involuntary" help from the malicious node may be valuable, especially when the network resource is limited. As a matter of fact, from the perspective of the malicious nodes, coexistence gives them a longer lifetime in the network and the opportunity to launch future attacks. As far as the regular nodes are concerned, they have a criteria to evaluate the benefit from the malicious nodes. The criteria also determine when to terminate the coexistence and isolate the malicious nodes.

To make the process of detection even more difficult, the malicious nodes do not act passively and wait to be detected. Instead, they also study the interaction they have with the rest of the network and adjust their subsequent actions accordingly. It is also possible that a malicious node is wise enough to learn and predict the actions of the regular nodes to assist itself in making its own decisions on how to behave. The options available to the malicious nodes complicate the solution space and most traditional control theoretic approaches fail to find the equilibrium strategies for both the regular and malicious nodes. In particular, these problems fall more appropriately in the domain of static and dynamic distributed games and thus the application of game theory is an elegant way to tackle such problems. It is important that solution concepts from game theory are used to guide the protocol design process such that nodes working in a distributed manner can co-exist, even with different intents.

Game theory [7,24] has been successfully applied to solve various problems in wireless networks including cooperation enforcement [5,6,11,21,26], routing protocols [10,22,30,34] and other system design issues [2,13,17,20,25,32,33]. Recently, much work has been done that investigates the interactions between the regular and malicious nodes using game theory. Kodialam et al. formally propose a game theoretic framework to model how a service provider detects an intruder [14]. However, their assumptions of zero-sum game and complete, perfect knowledge have limitations. Agah et al. study the non-zero-sum intrusion detection game in [1]; their results infer the optimal strategies in one-stage static game with complete information. In [19], Liu et al. propose a Bayesian hybrid detection approach to detect intrusion in wireless ad hoc networks. They design an energy efficient detection procedure while improving the overall detection power. The intrusion detection game with networked devices are investigated in [35], where Zhu et al. introduce an N-person non-cooperative game to study incentive compatibility of the collaborative detection. [18] models the intention and strategies of a malicious attacker through an incentive-based approach. The importance of the topology on the payoffs of the malicious nodes is investigated in [28]. An interesting flee option for the malicious node is proposed in [16]. In that analysis, a malicious node decides to flee when it believes it is too risky to stay in the network. While the approach focuses on how the flee action affects the result of the game, it does not consider the noises in observation.

There have been some recent researches that focus on the effects of imperfect and/or incomplete information in networking and communications security. In [23], the attacker defender game is modeled as a fictitious play (FP) game, and the authors study the effect of observation errors on the convergence of Nash Equilibrium when the error probability in the channel is unknown. They showed that in a stochastic FP game, the attacker can conceal its true strategy by including an entropy term in the payoff functions. The authors in [8] propose an interesting application of physical layer security game, where the source node pays the surrounding friendly jammer nodes to interfere the eavesdropper, so that the eavesdropper can be masked. The focus is on how to apply game theory to set the price charged by the friendly jammers. The research in [3] deals with malicious jammers when the user does not know how the jamming efforts are distributed among sub-carriers or the fading gains with certainty. The equilibrium strategies in closed form are derived and the range of sub-carriers where the transmitter can expect the jamming attack is specified. The jamming game in multi-band covert timing networks is considered in [25], where the camouflaging resources in the covert time network introduce uncertainty. In their modeling, a sensing game is played so that covert timing network can

determine the amount of camouflaging resources to be deployed. In another subsequent game, the malicious attacker finds the optimal transmit powers on each spectral band it chooses to attack. The existence of Nash equilibria in both games leads to a more effective defense mechanism against the attacker.

The research presented in this paper differs from existing literature in two aspects. First, our research presents a systematic analysis not only for the malicious node detection process, but also the interactions among nodes *after* detection. Unlike [19,35], our game theoretical analysis does not stop when the malicious node is detected. Instead, we propose the notion of co-existence with malicious node and extend the games after detection. Second, we empower the malicious node with countermeasures to learn from the games. In our game, the malicious node is intelligent enough to learn from the outcomes of the games and adjusts its strategies accordingly. We integrate the learning process of malicious node into the detection process and post-detection games.

In this paper we use game theory to model and analyze the interactions between a malicious node and a regular node. In particular, the malicious node is the active node (e.g., sending packets) and the regular node is the observer node (e.g., receiving packets). We formalize the interactions into two cascaded games. The first game, namely *malicious node detection game*, is a Bayesian game with imperfect information. The information is hidden because the malicious node can disguise as a regular node and the actions are hidden due to the noise and imperfect observation. The second game, called *post-detection game*, is played when the regular node knows confidently that its opponent is a malicious node. In the latter game, the regular node observes and evaluates the actions of the malicious node, and decides whether to keep it or isolate it. For both games, we show the existence of equilibria and derive the conditions that achieve them. To address the possible countermeasures the malicious node might take, we propose a belief about the belief model. In this model, the malicious node learns from its private observations and predicts if the regular node has accumulated enough information to make the detection. Associated with the belief, we show that a Markov Perfect Bayes–Nash Equilibrium emerges in the detection game. We also provide simulation study to support the efficiency and other properties of the equilibria.

The main contributions in this paper can be categorized into three parts.

- We model the malicious node detection game under unreliable channels as a Bayesian game with imperfect monitoring and show a mixed strategy perfect Bayesian Nash Equilibrium is attainable. The strategy profile is also shown to give a sequential equilibrium solution. As a special case, the equilibrium is applicable in a multihop fashion if no consecutive nodes along the route are malicious. Results show how the equilibrium strategy profiles are affected by parameters like channel noise, successful attack rate, successful detection rate, attack gain, detection gain, and false alarm rate.
- We propose the notion of coexistence after detection in order to utilize the malicious node. A coexistence index is designed to evaluate the helpfulness of a malicious node. We derive the conditions under which a subgame perfect Nash Equilibrium is achieved. Through simulation, we also show how the malicious node can be used to improve the network throughput and extend network lifetime.
- We introduce a novel belief about belief model employed by the malicious node. A Markov Perfect Bayes–Nash Equilibrium is induced when both nodes constantly update their beliefs. This equilibrium is shown to delay the detection of the malicious node and help the malicious node actively adjust its strategy to avoid detection. This model also helps to integrate the detection and post-detection games with effective transition.

The rest of the paper is organized as follows. In Section 2, we introduce and solve the Bayesian game of malicious node detection. Section 3 presents the post-detection game and discusses how malicious and regular nodes can coexist after detection. Section 4 explores the countermeasures available to the malicious node. Simulation results are presented in Section 5 that illustrate our findings on the detection and post-detection games as well as discussions how to integrate two games. The last section concludes the paper.

2. Malicious nodes detection game

2.1. Network model

We consider a wireless network consisting of *Regular* and *Malicious* nodes. By regular node we mean a node that works towards the common goal of the network. Also, it is rational and its actions are governed by an underlying utility function. On the other hand, a malicious node aims to hamper, disturb, and even attack the network. Although the actions of a malicious node are also determined by certain utility functions, such functions are designed to bring damages to the network. In addition, regular nodes are willing to cooperatively detect the malicious nodes in the network, even if the detection process might consume their own resources. On the contrary, malicious nodes do not work with regular nodes to detect other malicious ones.

Despite the two types of nodes, the identity (type) of a malicious node is not directly revealed to others. Instead, the types can only be estimated or conjectured through observing actions. To identify the attacks and malicious nodes in the network, a regular node can monitor the actions of others. However, such monitoring is costly (e.g., consumes the receivers' own resource) and a node cannot afford to monitor all the time. Moreover, the observations might not be accurate because of the noise, e.g., wireless channel loss. Thus, the regular nodes do not monitor the network all the time and during those times, attacks cannot be identified.

Our research focuses on a two-node interaction process in a wireless network. In particular, we are interested in the packet forwarding process. In this model, we consider

a single hop between a source and a destination, or a part of (one hop) a packet forwarding chain (e.g., ad hoc networks). We assume that node i, or the sender node, **has a packet to send** to the next-hop node. Such a packet can be generated by node i itself, or relayed to node i from another node. If the sender node's type is regular, it only takes the action "*Forward*".[1] If the sender node is malicious, it can choose to "*Attack*" with a risk of being identified or "*Forward*" (not attack) to disguise. The action *Attack* refers to a general set of actions that harm the network and disrupt normal network operation, e.g., intentional dropping of packets or altering the payload of packets. It is noted that unlike other research that tries to exploit the techniques of the *Attack* option, we generalize our approach such that it can be applied to a number of different kinds of attack, regardless of the unique features that an attack might have. Discussions on the applicability and attack types will be presented in the last section. The opponent of node i is node j which has the capacity to monitor the actions taken by node i. In the context of wireless networks, it refers to listening on the same channel node i is on. If node j takes *Monitor*, it turns on the radio and listens to what node i has to send. It is noted that for now we assume node j is not the intended recipient of node i's packet. We will relax this assumption later in Section 2.5. We further assume that time is divided into slots and nodes take their actions within each slot, i.e., one hop of packet forwarding takes one slot.

There are a number of well-known attacks that can be abstracted as our network model. *Byzantine attack* [31] is launched by a set of malicious or compromised nodes that behave arbitrarily to disrupt the network. in Byzantine attack, malicious nodes can selectively drop packets, which results in disruption or degradation of the routing services [4]. Byzantine attacks are hard to detect because malicious nodes drop packets selectively. Our model can also be directly applied to analyze the packet dropping in another attack called *Black hole attack* [27]. In this attack, the malicious node attracts the packets from neighboring nodes, intercepts them and consumes without any forwarding. In addition, the malicious node can selectively forward the packets or even modify the packets. Although attracting packets by advertising false routes is not covered in this research, our model can nevertheless characterize the packet forwarding process, where the attack action is a packet failure/drop and the selectiveness of not launching attacks is packet forwarding.

2.2. Game model

To abstract the interactions among the nodes, we consider a two-player non-zero sum game played by the nodes i and j. The types of these nodes, θ_i and θ_j, are private information. Since the type of each player is hidden, and the observation is not accurate, it is a Bayesian game with imperfect information [24].

To model the process of detecting the malicious nodes in the network, we apply a special category of Bayesian game called the signaling game. A *signaling game* is played between a sender and a receiver. The sender has a certain type and a set $\mathcal{M}$ of available messages to be sent. Based on its knowledge on its own type, the sender chooses a message from $\mathcal{M}$ and sends it to the receiver.[2] However, the receiver does not know the type of the sender and can only observe the message but not the type. Through observation, the receiver then takes an action in response to the message it observed. In the malicious node detection game, the sender, node i can be either regular $\theta_i = 0$ or malicious $\theta_i = 1$. The receiver node can also be regular or malicious. However, because the receiver node is the observer, a malicious node as receiver will not help to detect other malicious nodes in the network. Therefore, for now, we assume the receiver, node j is always a regular node, i.e., $\theta_j = 0$. We will revisit this assumption and relax it in Section 2.5.

The action profiles a_i available to node i are based on its type. For $\theta_i = 0, a_i = \{Forward\}$. For $\theta_i = 1, a_i \in \{Attack, Forward\}$. The receiver node j has the option to monitor if node i is attacking or not, thus $a_j \in \{Monitor, Idle\}$.

To further construct the game, we define the following values. Let g_A be the payoff of a malicious node if it successfully attacks. The cost associated with such an attack is c_A. For the receiver node j, the cost of monitoring is c_M and 0 if it is idle. Hence, for the action profile $(a_i, a_j) = (Attack, Idle)$, the net utility for a successful attacking node i is $g_A - c_A$, the loss for node j is $-g_A$ due to the attack. Similarly, if the action profile is $(a_i, a_j) = (Attack, Monitor)$, the attacking malicious node i loses $g_A + c_A$, and the net gain for node j is $g_A - c_M$. However, if a malicious node chooses not to attack, the cost to forward a packet is c_F, which is the same cost to a regular sender node. We make $-g_A - c_A < -c_F$ so that the malicious node has incentive to forward packet when it's under monitoring. Based on the types of node i and node j, the payoffs matrices are presented in Table 1. For quick reference, the notations used in this paper are tabulated in Table A.1 in Appendix A.

In addition, in our model, we introduce p_e as the channel loss rate. The channel unreliability implies that monitoring can be accurate with probability $1 - p_e$. We also denote γ as the attack success rate for a malicious node. γ represents the rate at which attack can be successfully launched when a malicious node intends to. γ also considers the physical limitation of the malicious node.[3] It is noted that with this model, unsuccessful attacks can be accurately monitored with reliable channel. When the channel is unreliable, the monitoring node cannot tell *Forward* from unsuccessful attacks.

[1] *Forward* only applies to the case when the sender node is forwarding others' packets. If the sender node is the origin of a packet, such action is called *Send*. For the simplicity of presentation, we use *Forward* to refer to this case.

[2] It's the receiver of the game signal, not the packet.

[3] We could have defined *monitor success rate* (ψ) to show the probability of a regular node to successfully monitor an attack considering its physical limitations. However, ψ has a similar effect as p_e, and we omit this variable for the sake of simplicity. If both ψ and p_e need to be considered, we can define a nominal channel loss rate $p'_e = 1 - (1 - p_e)\psi$ to replace p_e in our model. Such linear transform will not alter the current form of our analysis and results.

Table 1
Payoff matrix of two player malicious node detection game.

		Node j			
		Monitor		Idle	
(a) $\theta_i = 1$, *Malicious sender*					
Node i	Attack	$-g_A - c_A$	$g_A - c_M$	$g_A - c_A$	$-g_A$
	Forward	$-c_F$	$-c_M$	$-c_F$	0
(b) $\theta_i = 0$, *regular sender*					
Node i	Forward	$-c_F$	$-c_M$	$-c_F$	0

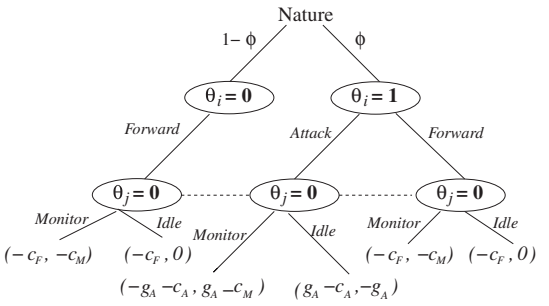


Fig. 1. Stage malicious node detection game tree.

2.3. Equilibrium analysis for the stage game

We begin our analysis on the malicious node detection game from the extensive form of the static Bayesian game as illustrated in Fig. 1. To solve this game, we are interested in finding the possible *Bayesian Nash Equilibrium* (BNE). In a static Bayesian game, the BNE is the Nash Equilibrium given the beliefs of both nodes. In our case, node i knows for sure that for node j, $\theta_j = 0$. However, node j is not clear of node i's type. Although for node i, its type is deterministic and not probabilistic, to consider both games in Table 1 and the hidden nature of its identity to node j, in Fig. 1, we illustrate that node j's belief about node i's type $\theta_i = 1$ happens with probability ϕ.

First, let us consider pure strategies only. Based on θ_i, the pure strategies (σ_i) available for node i are dependent on its type. If $\theta_i = 1$, node i can either play *Attack* or *Forward*. However when $\theta_i = 0$, node i can only play *Forward*. To categorize the actions available based on node type, we summarize the strategies available to node i as σ_i={(*Attack if malicious, Forward if regular*), *Always Forward*}. For node j, the strategy set is σ_j={*Monitor, Idle*}. To find the BNE, we let σ_i and σ_j play with each other and derive the conditions under which neither node can increase its utility by unilaterally changing its strategy.

Lemma 1. *In the malicious node detection game, there is a malice belief threshold ϕ_0, such that no pure strategy BNE exists if $\phi > \phi_0$.*

Proof. We start by eliminating a trivial pure strategy pair (σ_i, σ_j)=(*Always Forward, Monitor*). From Table 1(a), we know that for both nodes, they can improve their payoffs by deviating from the strategy pair. We further analyze the following two cases.

Case 1: σ_i = (*Attack if malicious, Forward if regular*). For node j, if σ_j = *Monitor*, the expected payoff is

$$\begin{aligned} u_j(Monitor) &= \phi(1-p_e)\gamma(g_A - c_M) + \phi(1-p_e)(1-\gamma) \\ &\quad \times (g_A - c_M) + \phi p_e \gamma(-g_A - c_M) + \phi p_e(1 \\ &\quad - \gamma)(-c_M) - (1-\phi)c_M, \end{aligned} \tag{1}$$

where each term represents perfect monitoring of the successful attack, perfect monitoring of the unsuccessful attack, failing to monitor the successful attack, failing to monitor the unsuccessful attack and node i is regular respectively. If σ_j = *Idle*, the expected payoff is

$$u_j(Idle) = -\phi\gamma g_A. \tag{2}$$

If (2) > (1), the dominant strategy for node j is *Idle*. Correspondingly, for node i, the best response would be (*Attack if malicious, Forward if regular*). Thus (σ_i, σ_j) = {(*Attack if malicious, Forward if regular), Idle*} is a BNE under the condition that $\phi < \frac{c_M}{(1-p_e)(1+\gamma)g_A}$. If (2) < (1), or $\phi > \frac{c_M}{(1-p_e)(1+\gamma)g_A}$, the dominant strategy for node j is *Monitor*, however, the best response to *Monitor* for node i is *Always Forward*. Hence (σ_i, σ_j) = {(*Attack if malicious, Forward if regular), Monitor*} is not a BNE under the condition that $\phi > \frac{c_M}{(1-p_e)(1+\gamma)g_A}$.

Case 2: (σ_i, σ_j) = {*Always Forward, Idle*}. If node j chooses not to monitor, the best response for node i is to *Attack if malicious*. This will lead to the previous case when $\phi < \frac{c_M}{(1-p_e)(1+\gamma)g_A}$. Therefore, there is no BNE if (σ_i, σ_j) = {*Always Forward, Idle*}.

To sum up, the pure strategy BNE exists if and only if $\phi < \frac{c_M}{(1-p_e)(1+\gamma)g_A}$. The equilibrium strategy profile is (σ_i, σ_j)={(*Attack if malicious, Forward if regular), Idle*}. In other words, we can find $\phi_0 = \frac{c_M}{(1-p_e)(1+\gamma)g_A}$, such that no pure strategy BNE exists if $\phi > \phi_0$. □

Although pure strategy BNE exists, it is not practical because the equilibrium requires node j to be *Idle* at all times, and hence the malicious nodes cannot be detected. It is also called Pooling Equilibrium [24] in which the receiver has no clue about sender's type. Therefore, it is desirable to seek a mixed-strategy BNE, and obviously, such BNE exists when $\phi > \phi_0$.

Let us denote p as the probability with which node i of type $\theta_i = 1$ plays *Attack* and q as the probability with which node j plays *Monitor*. To find the mixed strategy BNE of this game, we need to find the values of p and q such that neither node i nor j can increase payoff by altering the actions.

Lemma 2. *The malicious node detection game has a mixed strategy BNE when* $\sigma_i, \sigma_j = \left\{\left(Attack\ with\ \frac{c_M}{\phi(1+\gamma)(1-p_e)g_A}\ if\ \theta_i = 1,\ Forward\ if\ \theta_i = 0\right),\ Monitor\ with\ \frac{\gamma g_A - c_A + c_F}{(1-p_e)(1+\gamma)g_A}\right\}$.

Proof. For the mixed strategy played by node i, the payoff of node j playing *Monitor* is

$$\begin{aligned} u_j(Monitor) &= \phi p[(1-p_e)\gamma(g_A - c_M) \\ &\quad + (1-p_e)(1-\gamma)(g_A - c_M) \\ &\quad - p_e(1-\gamma)c_M - p_e\gamma(g_A + c_M)] - \phi(1-p)c_M \\ &\quad - (1-\phi)c_M = \phi[1 - p_e(1+\gamma)]pg_A - c_M. \end{aligned} \tag{3}$$

If node j plays *Idle*,

$$u_j(Idle) = -\phi\gamma p g_A. \quad (4)$$

Thus, in the mixed BNE strategy, $u_j(Monitor) = u_j(Idle)$. Thus $p = \frac{c_M}{\phi(1-p_e)(1+\gamma)g_A}$. Similarly, when node j plays the mixed strategy, the payoff of node i playing *Attack* is

$$\begin{aligned} u_i(Attack) &= -(1-p_e)q(g_A + c_A) + \gamma(1-q)(g_A - c_A) \\ &\quad + p_e\gamma q(g_A - c_A) - p_e(1-\gamma)qc_A - (1-\gamma) \\ &\quad \times (1-q)c_A \\ &= (p_e - 1)(1+\gamma)qg_A + \gamma g_A - c_A. \end{aligned} \quad (5)$$

When node i plays *Forward*,

$$u_i(Forward) = -c_F. \quad (6)$$

Hence, to obtain $q, u_i(Attack) = u_i(Forward)$, and $q = \frac{\gamma g_A - c_A + c_F}{(1-p_e)(1+\gamma)g_A}$. □

Lemmas 1 and 2 provide us with the conditions under which BNE can be attained. One of the conditions is the belief of malice threshold ϕ_0. As suggested in Lemma 1, this threshold is related to the channel reliability $(1-p_e)$, attack success rate (γ) and detection gain (g_A/c_M). In the pure strategy BNE, node i always attacks if it is malicious. The belief of node j on node i's malice is very low since the detection gain is usually very large as $p_e, \gamma \in [0,1]$. However, when the belief grows and eventually exceeds the threshold, the mixed strategy BNE requires node i to be less aggressive in attacking. In other words, the equilibrium implies node i should know about node j's belief when making the decision. When node j is absolutely sure about node i's type, node i's equilibrium attack probability drops to the value of the belief threshold.

2.4. Belief update and dynamic Bayesian games

So far, the analysis on the malicious node detection stage game has shown that the equilibrium is associated with node j's belief on node i's type. However, the difficulty lies in the assignment of the belief as a priori information available to node j. Thus, it is desirable that this belief can be accurately presented and dynamically updated. We apply dynamic Bayesian game theory to discuss how the belief is updated.

We assume that the static malicious node detection game is repeatedly played at every time slot, and we consider the infinite repeated game without discounting (i.e., payoffs in every stage/slot have equal weight). In addition to the notation defined in the stage game, we introduce $\mu_j^{(t)}(\theta_i = \bar{\theta}_i)$ as the belief node j holds about $\theta_i = \bar{\theta}_i$ at the tth stage of the subgame. Since node j is always a regular node, $\mu_i^{(t)}(\theta_j = 0)$=1 for all $t > 0$. We further define $a_i(t)$ as the action node i plays at tth stage. Node j may monitor node i's actions through the observed signal $\hat{a}_i(t)$. The reasons for the discrepancy between $a_i(t)$ and $\hat{a}_i(t)$ are the observation error caused by the channel unreliability and the false alarm rate (α) caused by the inaccuracy and limitation in the detection of node j.

Based on Bayes' theorem, we construct our belief update rule. If node j is continuously monitoring, its belief on θ_i can be calculated with the belief it holds at the immediate previous stage and the actions it observed. We write the belief at the $(t+1)$th stage as:

$$\mu_j^{(t+1)}(\theta_i) = \frac{\mu_j^{(t)}(\theta_i)P(\hat{a}_i(t)|\theta_i)}{\sum_{\tilde{\theta}_i \in \Theta}\mu_j^{(t)}(\tilde{\theta}_i)P(\hat{a}_i(t)|\tilde{\theta}_i)}, \quad (7)$$

where Θ is the space of all possible values θ_i can take; in our case $\Theta = \{0,1\}$.[4]

For each of the terms in Eq. (7), we have the following equations.

$$P(\hat{a}_i(t) = Attack|\theta_i = 1) = (1-p_e)p + \alpha(1-p), \quad (8)$$
$$P(\hat{a}_i(t) = Attack|\theta_i = 0) = \alpha, \quad (9)$$
$$P(\hat{a}_i(t) = Forward|\theta_i = 1) = p_e p + (1-\alpha)(1-p), \quad (10)$$
$$P(\hat{a}_i(t) = Forward|\theta_i = 0) = 1-\alpha. \quad (11)$$

Since node j does not monitor node i's actions at every stage, but rather with probability q. When node j is not monitoring, its belief remains the same at the next stage. Thus, Eq. (7) is revised as:

$$\mu_j^{(t+1)}(\theta_i) = q\frac{\mu_j^{(t)}(\theta_i)P(\hat{a}_i(t)|\theta_i)}{\sum_{\tilde{\theta}_i \in \Theta}\mu_j^{(t)}(\tilde{\theta}_i)P(\hat{a}_i(t)|\tilde{\theta}_i)} + (1-q)\mu_j^{(t)}(\theta_i). \quad (12)$$

The concept of *belief system* is hence introduced to describe the aforementioned belief building and updating process. A belief system is a function that assigns each information set[5] a probability distribution over the histories (i.e., past moves and states) in that information set [24]. Although in our discussions above, we did not explicitly state how history is accounted for in Eqs. (7) and (12), it is easy to observe that every updated belief is determined by the actions node j observes in the current stage and the belief it holds. The beliefs are further determined by the actions in the previous stages and it can be backtracked to the initial belief and the subsequent actions. Thus, the current belief and observed action can fully represent the histories in the information sets, and those information sets can be reached with positive probabilities if the strategies are carefully designed.

With the belief system, the games are played in a sequential manner. These games are independent of each other and as the game evolves, neither nodes can stick to the very same strategy at every stage to yield the most payoffs. Thus, the best response strategies are dependent on the current beliefs held by the nodes. Perfect Bayesian Equilibrium (PBE) can be applied to characterize the aforementioned dependency. In PBE, the belief system is updated by Bayes' rule. PBE also demands the optimality of subsequent play given the belief. Next, we show how to construct a PBE in the dynamic malicious node detection game.

[4] An alternative way to represent this process is using Naive Bayes estimation. Although the calculation complex is less, the representation is less intuitive and illustrative to the dynamics of the subgames.

[5] An information set is a set of all the possible moves that could have taken place in the game so far, for a particular player, given what that player has observed. In an imperfect information game, an information set contains all possible states in the history, e.g., in Fig. 1, the dotted lines show the information set available to node j.

We first show the existence of a mixed strategy equilibrium and then argue the unfeasibility of the pure strategy equilibrium. Consider an arbitrary stage k of the game; we denote $p^{(k)}$ as the probability node i of type $\theta_i = 1$ plays *Attack*, $q^{(k)}$ as the probability node j plays *Monitor*. In the equilibrium, $u_i^{(k)}(Attack) = u_i^{(k)}(Forward)$ and $u_j^{(k)}(Monitor) = u_j^{(k)}(Idle)$. The analysis is in similar form as the proof of Lemma 2. In particular,

$$\begin{aligned} u_i^{(k)}(a_i^{(k)} = Attack, a_j^{(k)} = Monitor) &= -(1-p_e)q^{(k)}(g_A + c_A) \\ &+ \gamma(1-q^{(k)})(g_A - c_A) + \gamma q^{(k)} p_e (g_A - c_A) \\ &- p_e(1-\gamma)q^{(k)}c_A - (1-\gamma)(1-q^{(k)})c_A, \end{aligned} \tag{13}$$

$$u_i^{(k)}(a_i^{(k)} = Forward, a_j^{(k)} = Monitor) = -c_F. \tag{14}$$

$$\begin{aligned} u_j^{(k)}(a_j^{(k)} &= Monitor, a_i^{(k)} = Attack) \\ &= \mu_j^{(k)}(\theta_i = 1)p^{(k)}[(1-p_e)\gamma(g_A - c_M) \\ &+ (1-p_e)(1-\gamma)(g_A - c_M) \\ &- p_e(1-\gamma)c_M - p_e\gamma(g_A + c_M)] \\ &- \mu_j^{(k)}(\theta_i = 1)(1-p^{(k)})c_M - \mu_j^{(k)}(\theta_i = 0)c_M, \end{aligned} \tag{15}$$

$$u_j^{(k)}(a_j^{(k)} = Idle, a_i^{(k)} = Attack) = -\mu_j^{(k)}(\theta_i = 1)\gamma p^{(k)} g_A. \tag{16}$$

The solutions to the above equations are

$$p^{(k)} = \frac{c_M}{\mu_j^{(k)}(\theta_i = 1)(1-p_e)(1+\gamma)g_A}, \tag{17}$$

$$q^{(k)} = \frac{\gamma g_A - c_A + c_F}{(1-p_e)(1+\gamma)g_A}. \tag{18}$$

What $p^{(k)}$ and $q^{(k)}$ suggest is an equilibrium profile $(\sigma_i^{(k)}, \sigma_j^{(k)})$. This profile shows the sequential rationality [7,24], that is, each node's strategy is optimal whenever it has to move, given its belief and the other node's strategy. In other words, at any stage k, for any alternative strategies $\sigma_{i'}{}^{(k)}$ and $\sigma_{j'}{}^{(k)}$,

$$\begin{aligned} &u_i^{(k)}((\sigma_i^{(k)}, \sigma_j^{(k)})|\theta_i, a_i(t), \mu_j^{(k)}(\theta_i)) \geqslant \\ &u_i^{(k)}((\sigma_{i'}{}^{(k)}, \sigma_j^{(k)})|\theta_i, a_i(t), \mu_j^{(k)}(\theta_i)), \end{aligned} \tag{19}$$

$$\begin{aligned} &u_j^{(k)}((\sigma_i^{(k)}, \sigma_j^{(k)})|\theta_i, \hat{a}_i(t), \mu_j^{(k)}(\theta_i)) \geqslant \\ &u_j^{(k)}((\sigma_i^{(k)}, \sigma_{j'}{}^{(k)})|\theta_i, \hat{a}_i(t), \mu_j^{(k)}(\theta_i)). \end{aligned} \tag{20}$$

Besides sequential rationality, a PBE also demands that the belief system satisfies the Bayesian conditions [7].

Definition 1 [7], pp. 331–332. The Bayesian conditions defined for PBE are

- B(i): Posterior beliefs are independent. For history $h^{(k)}$, $\mu_i(\theta_{-i}|\theta_i, h^{(k)}) = \prod_{j\neq i}\mu_i(\theta_j|h^{(k)})$.
- B(ii): Bayes' rule is used to update beliefs whenever possible.
- B(iii): Nodes do not signal what they do not know.
- B(iv): Posterior beliefs are consistent for all nodes with a common joint distribution on θ given $h^{(k)}$.

Our proposed belief system satisfies the Bayesian conditions. B(i) is satisfied because $\theta_j = 0$ all the time. Eq. (7) is derived from Bayes' rule, and hence B(ii) is also satisfied. B(iii) is fulfilled because node i's signal is determined by its action and if $a_i(k) = \hat{a}_i(k), \mu_j(\theta_i|a_i(k), h_j^{(k)}) = \mu_j(\theta_i|\hat{a}_i(k), h_j^{(k)})$. B(iv) is trivial in our game because no third player exists.

The analysis on Bayesian conditions and sequential rationality serves as the proof of the following theorem.

Theorem 1. *The dynamic malicious node detection game has a perfect Bayesian equilibrium that can be attained with strategy profile* $(\sigma_i^{(k)}, \sigma_j^{(k)}) = (p^{(k)}, q^{(k)})$.

Remark 1. The infeasibility of pure strategy PBE is proved as follows: If node i attacks, the best response for node j is *Monitor*, which makes node i non-profitable to play *Attack*. If node i plays *Forward*, $p^{(k)} = 0$, the best response for node j is *Idle* (i.e., $q^{(k)} = 0$). However, the sequential rationality requires $q^{(k)} \geqslant \frac{\gamma g_A - c_A + c_F}{(1-p_e)(1+\gamma)g_A}$, which leads to a contradiction. Therefore, no pure strategy PBE exists in the dynamic malicious node detection game. It is noted that the infeasibility of the pure strategy PBE in the dynamic settings should not be confused with the existence of a pure strategy BNE in a static game because the pure strategy BNE in a static game is always an artifact.

Remark 2. The proved PBE can be further refined to *Sequential Equilibrium* [15]. In the sequential equilibrium, the Bayesian conditions are extended as *belief sensibility* and *consistency*. The belief sensibility requires the information sets can be reached with positive probabilities (μ) given the strategy profile σ. The consistency demands an assessment (σ, μ) should be a limit point of a sequence of the mixed strategies and associated sensible beliefs, i.e., $(\sigma, \mu) = \lim_{n\to\infty}(\sigma^n, \mu^n)$. In our game, belief sensibility is satisfied because our proposed belief system updates the beliefs according to Bayes' rule and it assigns a positive probability to each of the information set. Theorem 8.2 in [7] states that in incomplete information multi-stage games, if neither player has more than two types, Bayesian condition is equivalent to belief consistency requirement. In our game, $\theta_i = 0, 1, \theta_j = 0$, and hence consistency is fulfilled. Together with the sequential rationality, the PBE in our game is also a sequential equilibrium. Since every finite extensive-form game has at least one sequential equilibrium, which is a refinement to PBE, it also implies the existence of PBE in our game.

2.5. Detection game beyond one hop

We now extend our discussion to beyond one hop. As an illustrative example, we consider a multi-hop packet forwarding chain which consists of multiple one hop sending processes. A series of malicious node detection game take place as the packet is relayed from one node to another along a pre-defined route. We describe the malicious node detection in a multi-hop scenario as follows.

(1) The source node generates the packet and sends it to the first forwarding node. In the event that the source node is malicious, as long as it does not play *Attack*, we still regard the packet generated from a malicious node as useful and harmless.[6]
(2) For an intermediate node, it is node j at nth hop, and it becomes the node i of $n+1$th hop game as long as node i in nth hop does not successfully play *Attack*. The series of cascaded detection game will terminate once an *Attack* is successful or packet is lost due to channel unreliability.
(3) Once packet reaches the destination node, no matter what type of node the destination is, multi-hop packet forwarding is regarded as successful and complete. There will be no more detection game, even if the destination node is malicious.
(4) Although when multiple nodes are in the game, it is possible for them to have differentiated payoffs, for mathematical tractability, in this research, we apply the same payoff structure to all nodes.

In multi-hop forwarding, the receiver at nth hop is the sender at $n+1$th hop. This requires us to relax the limit on the types of node i and j at single shot detection game. In particular, there are be four different combinations.

(1) Node i is malicious, node j is regular. This is the original case we discussed when θ_i=1.
(2) Both nodes i and j are regular. This is the case we already discussed when θ_i=0.
(3) Both nodes i and j are malicious. In this case, if node i does not forward packet, the packet forwarding ends. Since no packets can be forwarded from either node, the rest of the network can treat nodes i and j as a wormhole [9].
(4) Node i is regular, node j is malicious. This is an interesting setup as the game at this stage does not provide any useful results. When the malicious node is receiving, it cannot launch attack and the sender node cannot observe either. The implication of this case really depends on the next hop game. If the next node is malicious, then packet forwarding will end and create a wormhole. If the next node is regular, then detection game will resume at next hop.

By studying the detection game at nth hop, we conclude that we can apply the analysis of single hop detection game to an individual hop at multi-hop malicious node detection game. The cascade games can be played sequentially as long as no two malicious nodes play with each other at any hop. In addition, one node j can play the detection games with different nodes is on different packet forwarding paths. However, if only one packet forwarding path is considered, whether there is a detection game in the $n+1$th hop is determined by whether the sending node in the nth hop plays *Attack*. If *Attack* is played, the series of detection game on that path are terminated; nonetheless, detection games on other paths are not affected.

To abstract the dynamic malicious node detection games beyond one hop, and clarify the applicability of Theorem 1 on such games, we provide the following conclusion.

Corollary 1. *The dynamic malicious node detection game can be extended to the format of spacial cascaded two-player games in multiple hops. Perfect Bayesian equilibrium is attainable in each of the hops until attack action is taken or observed.*

3. Post-detection game and coexistence

In the previous section, we have discussed how to update node j's belief system based on Bayes' rule. It is natural that through observation, although imperfect at every stage game, node j can accumulate a better estimation about θ_i. Eventually, after repeated monitoring, there will be a stage at which node j can predict with confidence whether node i is regular or malicious.

3.1. Game model

Traditionally speaking, after node j has identified node i as a malicious node, it will try to report and isolate node i from the rest or the network immediately to prevent future attacks. However, there are also situations where "isolation" may not be a good choice. Let us consider a wireless network which operates on a limited resource budget. In order to prolong the lifetime of the network, every regular node has to be economical towards packet forwarding. Hence, if a malicious node can be used to handle some of the traffic, it is beneficial not to isolate it.

Although the idea of "making malicious node beneficial" might sound counter-intuitive, it is backed by the following reasoning. In the malicious node detection game, we explained that the malicious node needs to be *cooperative* and not attack in order to camouflage. Furthermore, the malicious node is not aware of the outcome of the detection process employed by the regular node. Therefore, the regular nodes can exploit the fact that the malicious node sometimes are involuntarily cooperative to avoid detection. In the context of packet forwarding as the underlying application, when a malicious node plays the cooperative strategy of *Forward*, the packet drop is less compared to the *Attack* action. Lower packet drop means network output improves, and it comes from the "helpfulness" of a malicious node, even if its intention is to camouflage, the action is indeed useful and helpful.

However, there is a trade-off between how much benefit a malicious node can bring and what damage it can do. We denote n_F and n_A as the number of successful forwarding actions and number of attacks taken by a malicious node. Recall the cost of forwarding is c_F and the loss due to an attack to the network is g_A.[7] Thus, for a regular node,

[6] Technically, the malicious sender may choose to generate some packets with false or ill-intended data; in our modeling, such packets are regarded as *Attack*.

[7] Due to different settings of the network, g_A need not be a constant. It changes with time, topology, and the network traffic pattern. In order to keep the analysis tractable, we regard g_A as an average value of loss due to an attack to the network.

if it observes that the total saving due to forwarding ($n_F c_F$) a malicious node contributes is greater than the total cost due to its attack ($n_A g_A$), then keeping that node in the network is profitable. It is also worthwhile to mention that although the values of c_F and g_A vary from one application to another, for a given application, the values are constant and measurable.

To further analyze the conditions under which a malicious node can be kept and coexist with the regular ones, we formally define the post detection game. The game has two players: node i and node j. Unlike the Bayesian nature of the detection game in Section 2, in post-detection game, both nodes know the types of their opponent, i.e., node j knows that node i is malicious but has not taken any action to isolate it. Thus, $\theta_i = 1, \theta_j = 0$. The actions available for node i is $a_i \in \{Attack, Forward\}$, while the actions for node j is $a_j \in \{Monitor, Idle\}$. When node j monitors, it keeps a record of what node i has done since the beginning of the game. It also calculates a coexistence index $\mathcal{E}_i = \mathcal{E}_i^{(0)} + \hat{n}_F c_F - \hat{n}_A g_A$ for node i, where $\mathcal{E}_i^{(0)}$ is an initial value of the index, $\hat{n}_F$ is the observed number of forwarding actions and $\hat{n}_A$ is the observed number of attacks. If $\mathcal{E}_i$ falls under a certain threshold τ, node j will isolate node i and terminate the post-detection game because keeping node i is no longer beneficial. If $\mathcal{E}_i \geqslant \tau$, the game will be played in a repeated manner. The payoff matrix for the post-detection game is the same as the detection game for $\theta_i = 1$ as was shown in Table 1(a).

3.2. Searching for a coexistence equilibrium

Let us explore the strategies that both nodes can take to reach the equilibrium of coexistence. To avoid confusion, we denote $p^*(t)$ and $q^*(t)$ as the probability node i plays *Attack* and node j plays *Monitor* respectively with time. It is noted that these probabilities are different from the ones we obtained in Section 2.4. Also, since this game is no longer Bayesian, we are more interested in obtaining a subgame perfect Nash Equilibrium.

We first derive the Nash Equilibrium using indifference conditions. Suppose the post-detection game is played at kth repetition, i.e., subgame k. The expected payoff for player j playing *Monitor* is

$$\begin{aligned} u_j^{(k)}(Monitor) &= \{p^*(k)[(1-p_e)\gamma(g_A - c_M) + (1-p_e) \\ &\quad \times (1-\gamma)(g_A - c_M) - p_e(1-\gamma)c_M \\ &\quad - p_e\gamma(g_A + c_M)]\}\Pr(\mathcal{E}_i \\ &\geqslant \tau) + (1-p_e)p^*(k)(g_A - c_M)\Pr(\mathcal{E}_i \\ &< \tau) - (1 - p^*(k))c_M \\ &= [(1 - p_e + \gamma p_e)g_A - c_M]p^*(k)\Pr(\mathcal{E}_i \\ &\geqslant \tau) + (1-p_e)p^*(k)(g_A - c_M)\Pr(\mathcal{E}_i \\ &< \tau) - (1 - p^*(k))c_M. \end{aligned} \tag{21}$$

If node j plays *Idle*, the expected payoff is always

$$u_j^{(k)}(Idle) = -\gamma p^*(k) g_A. \tag{22}$$

Thus, the indifference condition require $u_j^{(k)}(Monitor) = u_j^{(k)}(Idle)$, and hence $p^*(k)$ is obtained as in Eq. (23) on next page.

Similarly, we can apply the indifference condition to node i as:

$$\begin{aligned} u_i^{(k)}(Attack) &= q^*(k)\{-(1-p_e)(g_A + c_A)\Pr(\mathcal{E}_i < \tau) \\ &\quad + (1-p_e)[\gamma(g_A - c_A) - (1-\gamma)c_A]\Pr(\mathcal{E}_i \geqslant \tau) \\ &\quad + p_e\gamma(g_A - c_A) - p_e(1-\gamma)c_A\} \\ &\quad - (1 - q^*(k))[(1-\gamma)c_A - \gamma(g_A - c_A)] \\ &= q^*(k)\{-(1-p_e)(g_A + c_A)\Pr(\mathcal{E}_i < \tau) \\ &\quad + (\gamma g_A - c_A)[(1-p_e)\Pr(\mathcal{E}_i \geqslant \tau) + p_e]\} \\ &\quad + (1 - q^*(k))(\gamma g_A - c_A). \end{aligned} \tag{24}$$

$$u_i^{(k)}(Forward) = -c_F. \tag{25}$$

Therefore, $q^*(k)$ can be expressed as Eq. (26) above.

The problem is then reduced to obtaining the probability distribution of $\mathcal{E}_i$. Let us assume at the beginning of the post-detection game $\mathcal{E}_i^{(0)} = c_0 \geqslant \tau$. For the sake of discussion, we also assume that node j is constantly monitoring. Hence, if we consider l subgames, in each of the subgame, $\mathcal{E}_i$ is updated.

We denote a random variable $y = \mathcal{E}_i = c_0 + \hat{n}_F c_F - \hat{n}_A g_A$. Since the mixed strategy profile requires node i to choose *Attack* with probability $p^*(t)$, $\hat{n}_F$ and $\hat{n}_A$ are binomially distributed as:

$$p^*(k) = \frac{c_M}{[(1 - p_e + \gamma p_e)g_A - c_M]\Pr(\mathcal{E}_i \geqslant \tau) + (1-p_e)(g_A - c_M)\Pr(\mathcal{E}_i < \tau) + c_M + \gamma g_A}. \tag{23}$$

$$q^*(k) = \frac{c_A - \gamma g_A - c_F}{-(1-p_e)(g_A + c_A)\Pr(\mathcal{E}_i < \tau) + (1-p_e)(\gamma g_A - c_A)(\Pr(\mathcal{E}_i \geqslant \tau) - 1)}. \tag{26}$$

$$\Pr(\hat{n}_F = \widehat{N}_F) = C_l^{\widehat{N}_F}[(1-p^*(t))(1-p_e)]^{\widehat{N}_F} \times [1-(1-p^*(t))(1-p_e)]^{l-\widehat{N}_F}, \quad (27)$$

$$\Pr(\hat{n}_A = \widehat{N}_A) = C_l^{\widehat{N}_A}[p^*(t)(1-p_e)]^{\widehat{N}_F}[1-p^*(t)(1-p_e)]^{l-\widehat{N}_F}. \quad (28)$$

Since $y = c_0 + \hat{n}_F c_F - \hat{n}_A g_A = c_0 + \hat{n}_F c_F - (l-\hat{n}_F)g_A = (c_F + g_A)\hat{n}_F - lg_A + c_0$ and l, c_F, g_A, c_0 are constants, to get the distribution of y, we first get the distribution of $w = y + lg_A - c_0$.

We use the probability generation function (pgf). For discrete random variable x, its pgf is defined as

$$G_X(z) = E[z^X] = \sum_{x=0}^{\infty} z^x \Pr(X = x). \quad (29)$$

The pgf for w is

$$\begin{aligned} G_W(z) &= E[z^W] = E\left[z^{\widehat{N}_F(c_F+g_A)}\right] \\ &= \sum_{\hat{n}_f=0}^{l}\left\{z^n C_l^{\hat{n}_f}[(1-p^*(t))(1-p_e)]^{\hat{n}_f} \times [1-(1-p^*(t))(1-p_e)]^{l-\hat{n}_f}\right\}^{(c_F+g_A)} \\ &= \{(1-p^*(t))(1-p_e) + [1-(1-p^*(t))(1-p_e)]z\}^{(c_F+g_A)l}. \end{aligned} \quad (30)$$

Let $f^{(n)}(x) = \frac{\partial^n f(x)}{\partial x^n}$,

$$\Pr(w = \omega) = \frac{G_W^{(\omega)}(0)}{\omega!}. \quad (31)$$

Therefore, we could obtain the probability terms in Eqs. (23) and (26) as,

$$\Pr(y = \mathcal{E}_i \geqslant \tau) = \Pr(w \geqslant lg_A + \tau - c_0) = \sum_{n \geqslant lg_A+\tau} \frac{G_W^{(n)}(0)}{n!}, \quad (32)$$

$$\Pr(\mathcal{E}_i < \tau) = 1 - \sum_{n \geqslant lg_A+\tau-c_0} \frac{G_W^{(n)}(0)}{n!}. \quad (33)$$

To relax the assumption of node j's constant monitoring, the current stage t for the analysis is $\lceil t = l/q^*(t) \rceil$. Therefore, we have obtained the equilibrium strategy parameter $p^*(t)$ and $q^*(t)$ for every subgame.

So far, we have shown that for the mixed strategy profile, attaining a Nash Equilibrium is feasible. As a matter of fact, every game has a mixed strategy Nash Equilibrium. To further refine the equilibrium, we apply the One-Shot Deviation Property to derive the condition for subgame perfect Nash Equilibrium. The property states:

Definition 2. One-Shot Deviation Property (OSDP) [24]: No player can increase her payoff by changing her action at the start of any subgame in which she is the first-mover, given the other player's strategies and the rest of her own strategy.

We take node j as an example and assume the repeated game has no discount. In our previous equilibrium analysis using the indifference condition, we have proved that deviation from $p^*(t)$ or $q^*(t)$ will not increase the payoffs. Hence, in the following derivation, we show the deviation strategy is related to $\mathcal{E}_i$.

From Eqs. (21) and (22), we can express the expected payoff for node j as:

$$\begin{aligned} U_j = \sum_{t=0}^{T} q^*(t)\{&[(1-p_e+\gamma p_e)g_A - c_M]p^*(t)\Pr(\mathcal{E}_i \geqslant \tau) \\ &+ (1-p_e)p^*(t)(g_A - c_M)\Pr(\mathcal{E}_i < \tau) - (1-p^*(t))c_M\} \\ &- \gamma(1-q^*(t))p^*(t)g_A. \end{aligned} \quad (34)$$

Suppose node j deviates at r^{th} stage and $r \leqslant T$. The deviation can be either of the following two cases.

Case 1: Isolate node i while $\mathcal{E}_i \geqslant \tau$. In this case, if node i attacks and is successfully observed, it will be isolated. The expected payoff at this stage for node j is

$$\begin{aligned} U_{j,dev,1}^{(r)} = \{&q^*(r)\{(1-p_e)p^*(r)(g_A - c_M) - p_e\gamma p^*(r)(g_A + c_M) \\ &- [p_e(1-\gamma)p^*(r) + (1-p^*(r))]c_M\} \\ &- \gamma(1-q^*(r))p^*(r)g_A\}\Pr(\mathcal{E}_i \geqslant \tau). \end{aligned} \quad (35)$$

Case 2: Keep node i while $\mathcal{E}_i < \tau$. Since node j only deviates one stage, node i will be isolated in the next stage. The expected payoff for node j at this stage is the same as above except for the last probability term.

$$\begin{aligned} U_{j,dev,2}^{(r)} = \{&q^*(r)\{(1-p_e)p^*(r)(g_A - c_M) - p_e\gamma p^*(r)(g_A + c_M) \\ &- [p_e(1-\gamma)p^*(r) + (1-p^*(r))]c_M\} \\ &- \gamma(1-q^*(r))p^*(r)g_A\}\Pr(\mathcal{E}_i < \tau). \end{aligned} \quad (36)$$

In this way, the total expected payoff for node j under deviation is

$$U_{j,dev} = \sum_{t=0}^{r-1} U_j^{(t)} + U_{j,dev,1}^{(r)} + U_{j,dev,2}^{(r)} + \sum_{t=r+1}^{T} U_j^{(t)}. \quad (37)$$

OSDP require $U_{j,dev} \leqslant U_j$. After algebraic manipulation, we have

$$\begin{aligned} &\gamma g_A(q^*(t)p_e + 1) + q^*(t)p_e(\gamma c_M + 1 - \gamma) \geqslant \gamma(1-q^*(t))g_A \\ &\quad + q^*(t)[\gamma g_A \Pr(\mathcal{E}_i < \tau) + p_e c_M \Pr(\mathcal{E}_i \geqslant \tau)], \end{aligned} \quad (38)$$

or

$$\begin{aligned} \gamma g_A[p_e + 1 - \Pr(\mathcal{E}_i < \tau)] \geqslant p_e[&c_M \Pr(\mathcal{E}_i \geqslant \tau) \\ &+ \gamma - 1 - \gamma c_M]. \end{aligned} \quad (39)$$

To sum up, for the equilibrium on the post-detection game, we state the following theorem.

Theorem 2. *The post-detection game has a mixed strategy Nash Equilibrium when node i attacks with $p^*(t)$ and node j monitors with $q^*(t)$. This strategy is also subgame perfect if $g_A\gamma[p_e + 1 - \Pr(\mathcal{E}_i < \tau)] \geqslant p_e[c_M \Pr(\mathcal{E}_i \geqslant \tau) + \gamma - 1 - \gamma c_M]$.*

3.3. Convergence of the coexistence equilibrium

The post-detection game described above ends when $\mathcal{E}_i < \tau$. Since $\Pr(\mathcal{E}_i < \tau) > 0$, the game is of finite stages. In this subsection, we try to derive the expected length (number of stages) of the game.

We focus on the random variable $\mathcal{E}_i$. As we mentioned earlier, $\mathcal{E}_i = c_0 + \hat{n}_F c_F - \hat{n}_A g_A$. Again, we assume node j is

constantly monitoring. After one stage game, the probability of $\hat{n}_F = \hat{n}_F + 1$ is $(1-p^*(t))(1-p_e)$, and the probability of $\hat{n}_A = \hat{n}_A + 1$ is $p^*(t)(1-p_e)$. Thus, we model the evolution of $\mathcal{E}_i$ as a random process similar to a 1-dimensional random walk, where the value of $\mathcal{E}_i$ increases by c_F with probability $(1-p^*(t))(1-p_e)$, and decreases by g_A with probability $p^*(t)(1-p_e)$. The $1-p_e$ term comes from the unreliability of the channel. To obtain the expected length of the post-detection game, it is equivalent to calculating the expected first hitting time of the random process with the absorbing boundary $\mathcal{E}_i = \tau$.

Theorem 3. *The expect length of the post-detection game is*

$$\sum_{\eta>0}\eta\frac{\binom{\eta}{\hat{n}_F}-\sum_{d}^{\hat{n}_F-1}\binom{\frac{(c_0-\tau)/c_F+d}{g_A/c_F}}{d}\binom{\eta-\frac{(c_0-\tau)/c_F+d}{g_A/c_F}}{\hat{n}_F-d}}{2^\eta}.$$

Proof. Please refer to Appendix A. □

4. Countermeasures for the malicious node

Let us revisit the malicious node detection game. In our discussions so far, we haven shown that it is feasible to design strategies in order to achieve the proposed PBE in the malicious node detection game. However, there are still some issues that must be resolved before the equilibrium strategies can be applied and followed by practitioners. These issues can be categorized into two aspects. First, the PBE requires the malicious node perfectly know the belief held by the regular node. However, in practice, the belief information is never shared. Second, the malicious node may not remain passive in the detection game; instead, it can also form its belief about the current status in the game and adjust its strategy accordingly.

It is natural that not only the regular node but also the malicious node (node i) study the game through observation. In particular, node i understands that although the unreliable channel makes the observations inaccurate, the more often it attacks, the quicker node j can form a correct belief about its malicious type. Thus, node i should take different strategies when different beliefs are held by node j. These strategies (e.g., the PBE strategy in Eq. (17)) are Markovian when we view the beliefs as a set of states. The Markovian strategy adopted by node i is only determined by the current state of the belief, i.e., when the belief update process takes place. Therefore, if we regard the strategy taken by node i as decision making, it is similar as a Markovian Decision Process (MDP). However, the belief held by node j is its private information, and by no means can node i access this information. Therefore, it is essential for node i to construct its own belief system, which is the belief on the belief node j holds towards node i and we call this belief developed by node i *belief about belief*.

We denote $\mu_i(\mu_j(\theta_i))$ as the belief node i holds about node j's belief about node i, i.e., $\mu_i(\mu_j(\theta_i))$ is the belief about $\mu_j(\theta_i)$. It is noted that for node j, its belief $\mu_j(\theta_i)$ is used to determine whether the node i is malicious and accordingly switch to the post-detection game. Hence, the belief does not change its strategy. For node i, its belief $\mu_i(\mu_j(\theta_i))$ is used to initiate the countermeasure in order to reach the equilibrium strategy. For the game we presented in Table 1(a), depending on the actions nodes i and j take, the payoff of node i, u_i, can be one of the three different values: $-g_A - c_A, g_A - c_A$ or $-c_F$. While the observations of the payoffs are node i's private information, given a specific observation o_i, node i can predict the actions taken by node j, despite the prediction may be inaccurate. For example, when $o_i = -g_A - c_A$, node i knows for sure $a_j = Monitor$. However, when $o_i = -c_F$, node i cannot tell what node i has done. Further, based on the prediction of the actions node j takes, node i can update its belief $\mu_i(\mu_j(\theta_i))$ on how node j's belief $\mu_j(\theta_i)$ has changed due to a_j. Continuing with the same examples, when $o_i = -g_A - c_A, a_j = Monitor$, so node j observes the *Attack* launched by node i and it will update $\mu_j(\theta_i)$ according to Eq. (7). Similarly, when $o_i = g_A - c_A$, node i knows that node j is idle and $\mu_j(\theta_i)$ will not change. However, the uncertainty comes when $o_i = -c_F$, where node i cannot accurately update its belief about $\mu_j(\theta_i)$.

To construct the belief update system for node i, we employ the Bayes' Theorem. At stage t of the game, based on the observation $o_i^{(t)}$, node i's belief $\mu_i(\theta_i)$ is updated as:

$$\mu_i^{(t+1)}(\mu_j(\theta_i)) = \frac{\mu_i^{(t)}(\mu_j(\theta_i))P(o_i^{(t)}|\theta_i)}{\sum_{\tilde{\theta}_i\in\Theta}\mu_i^{(t)}(\mu_j(\tilde{\theta}_i))P(o_i^{(t)}|\tilde{\theta}_i)}, \tag{40}$$

where Θ={0,1}.

The conditional probabilities of observing o_i given its type θ_i can be calculated as follows. To distinguish from the strategy profiles we used previously, we denote $\tilde{p}$ as the probability node i launches attacks, and $\tilde{q}$ as the probability node j monitors. Therefore, the probabilities that arise due to the different observations and node i's type are:

$$P(o_i^{(t)} = -g_A - c_A|\theta_i = 1) = (1-p_e)\tilde{p}\tilde{q} + \alpha(1-\tilde{p})\tilde{q}, \tag{41}$$
$$P(o_i^{(t)} = -g_A - c_A|\theta_i = 0) = \alpha\tilde{q}, \tag{42}$$
$$P(o_i^{(t)} = g_A - c_A|\theta_i = 1) = \tilde{p}[p_e\tilde{q} + (1-\tilde{q})], \tag{43}$$
$$P(o_i^{(t)} = g_A - c_A|\theta_i = 0) = 0, \tag{44}$$
$$P(o_i^{(t)} = -c_F|\theta_i = 1) = (1-\tilde{p})[(1-\alpha)\tilde{q} + (1-\tilde{q})], \tag{45}$$
$$P(o_i^{(t)} = -c_F|\theta_i = 0) = (1-\alpha)\tilde{q}. \tag{46}$$

With the above equations, for each of the observations $o_i \in \mathbf{O}$, where $\mathbf{O} = \{-g_A - c_A, g_A - c_A, -c_F\}$, $\mu_i^{(t+1)}(\theta_i)$ is updated independently. Since for the malicious node i, its type $\theta_i = 1$ is known to itself, the overall belief is hence updated considering each of the possible observations with $P(o_i^{(t)}|\mathbf{1})$ representing $P(o_i^{(t)}|\theta_i = 1)$.

$$\mu_i^{(t+1)}(\mu_j(\theta_i)) = \sum_{o_i\in\mathbf{O}}P(o_i^{(t)}|\mathbf{1})\frac{\mu_i^{(t)}(\mu_j(\theta_i))P(o_i^{(t)}|\theta_i)}{\sum_{\tilde{\theta}_i\in\Theta}\mu_i^{(t)}(\mu_j(\tilde{\theta}_i))P(o_i^{(t)}|\tilde{\theta}_i)}. \tag{47}$$

Further, with the belief system of node i, the malicious node detection game can be solved again to obtain the sequential rationality. The derivation process is similar to what we have presented in Eqs. 13 and 16, with the exception that $\mu_i^{(t)}(\mu_j(\theta_i))$ will be considered. The equilib-

rium strategy profiles that reaches sequential rationality are,

$$\tilde{p}^{(t)} = \frac{c_M}{\mu_i^{(t)}(\mu_j(\theta_i = 1))(1 - p_e)(1 + \gamma)g_A}, \quad (48)$$

$$\tilde{q}^{(t)} = \frac{g_A\gamma - c_A + c_F}{(1 - p_e)(1 + \gamma)g_A}. \quad (49)$$

Moreover, it is easy to justify that the belief update process for node i also satisfies the Bayesian condition in Definition 1. In addition, Eq. (48) suggests that node i's strategy is purely dependent on the current belief it holds. Thus, we can further refine the PBE in malicious detection game.

Theorem 4. *With the belief about belief system for node i, the dynamic malicious node detection game has a Markov Perfect Bayes–Nash Equilibrium (MPBNE) when the strategy profiles are* $(\sigma_i^{(t)}, \sigma_j^{(t)}) = (\tilde{p}^{(t)}, \tilde{q}^{(t)})$.

The equilibrium is called Markov because the strategies associated are Markovian based on the beliefs. It is noted that the PBE obtained in Theorem 1 is also a MPBNE, however, the strategy profile has limited applicability because the equilibrium profile for node i requires the knowledge of node j's state (belief). On the contrary, the profiles in Theorem 4 only rely on the private information available to the nodes themselves. Our analysis of node i's belief system can be also interpreted as an Interactive Partially Observed Markovian Decision Process (IPOMDP) solution to a Partially Observed Stochastic Game (POSG) [12].

A special case for the strategy profile σ_i is "Always attack when $\mu_i^{(t)}(\mu_j(\theta_i = 1)) < \bar{\mu}$ and forward otherwise, for a predefined threshold $\bar{\mu} \in (0, 1)$". In this strategy, when $\mu_i^{(t)}(\mu_j(\theta_i = 1)) < \bar{\mu}$, node i will attack with $\tilde{p} = 1$. In this way, node j will progressively update its belief when it monitors because node i is always behaving maliciously. However, when the belief threshold is reached, node i will refrain from launching attacks, and hence its payoff will decrease. It is clear that the strategy deviates from the MPBNE because $\tilde{p}$ does not adhere to the equilibrium. As a result, node i will be identified quickly and it will be dormant for the rest of the time. While this strategy is favorable to node j and the network, from node i's perspective, this strategy will limit its attacks and hence it is not desirable.

5. Simulation model and results

In this section, we study the properties of the perfect Bayesian Nash equilibrium in the malicious node detection game and the post-detection subgame perfect Nash equilibrium through simulations. In our simulator, two players play the games repeatedly; the payoffs and strategy profiles for each of the subgames are recorded to analyze the properties of the equilibria. To set up the simulation environment, unless otherwise redefined, the default values of the parameters are $p_e = 0.01, \gamma = 0.95$ and $\alpha = 0.01$. The goal of the simulation is to show how certain parameter will effect the property of the equilibrium strategies given the rest of the parameters fixed.

5.1. Malicious node detection game

We first present the simulation results on the malicious node detection game. In Fig. 2, we show how the monitoring probability in PBE strategy increases with the malicious node attack success rate. The plots infer that the equilibrium require node j to increase its monitoring frequency as the attack success rate increases. Also, as the channel becomes more unreliable, node j must play *Monitor* more frequently.

Fig. 3 compares the convergence of node j's belief system when different attack gains are presented. In Fig. 3(a), we show how the belief system forms a correct belief on the type of node i when only *Attack* is observed. The convergence of the belief system under PBE is illustrated in Fig. 3(b). The plots suggest that the lower the attack gain is, the quicker the belief system converges. This property can be explained as follows. A smaller attack gain requires node i to attack more often in order to get more payoff, and increasing the attack frequency also increases the risk of being successfully observed. With more observations, the belief is updated more frequently and accurately. Belief system converges slower in Fig. 3(b) than in Fig. 3(a) because in the PBE, instead of constantly monitoring, node j only monitors with probability q.

A more complete study on the convergence of the belief system is shown in Fig. 4. Plots in Fig. 4(a) indicate the larger the disguise cost c_F/c_A is, the less time it takes to converge. This is because, with a larger disguise cost, it is unprofitable for node i to disguise by forwarding packets. Instead, it will launch more attacks, thus increasing the chances to be identified. Fig. 4(b) shows a quicker converged belief system for a smaller detection gain because node j needs to monitor more often to be profitable. Figs. 4(c) and (d) relate the convergence with less errors and uncertainties in the system. As expected, with errors and uncertainties (i.e., low channel loss, high attack success rate and low false alarm rate), the belief system converges quickly.

Finally, the parameters affecting the PBE attack probability p are investigated in Fig. 5. The attack gain is a very important factor in determining the value of p as shown

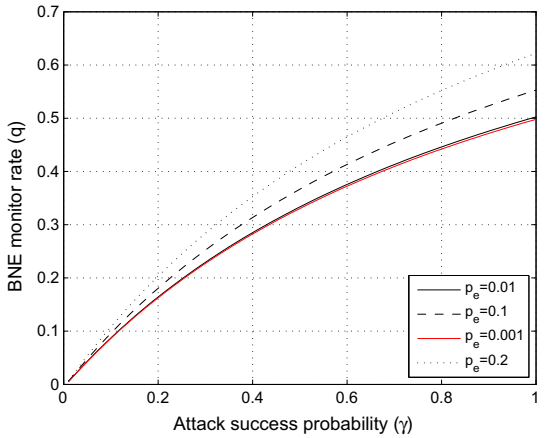
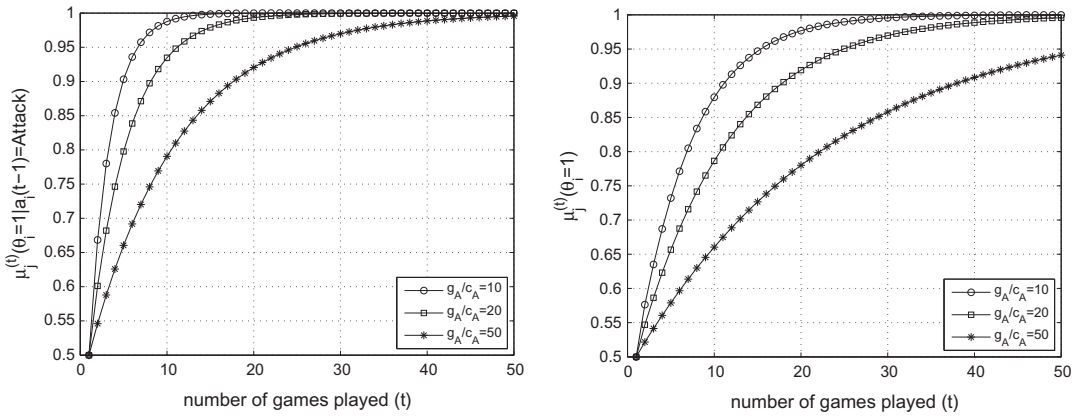
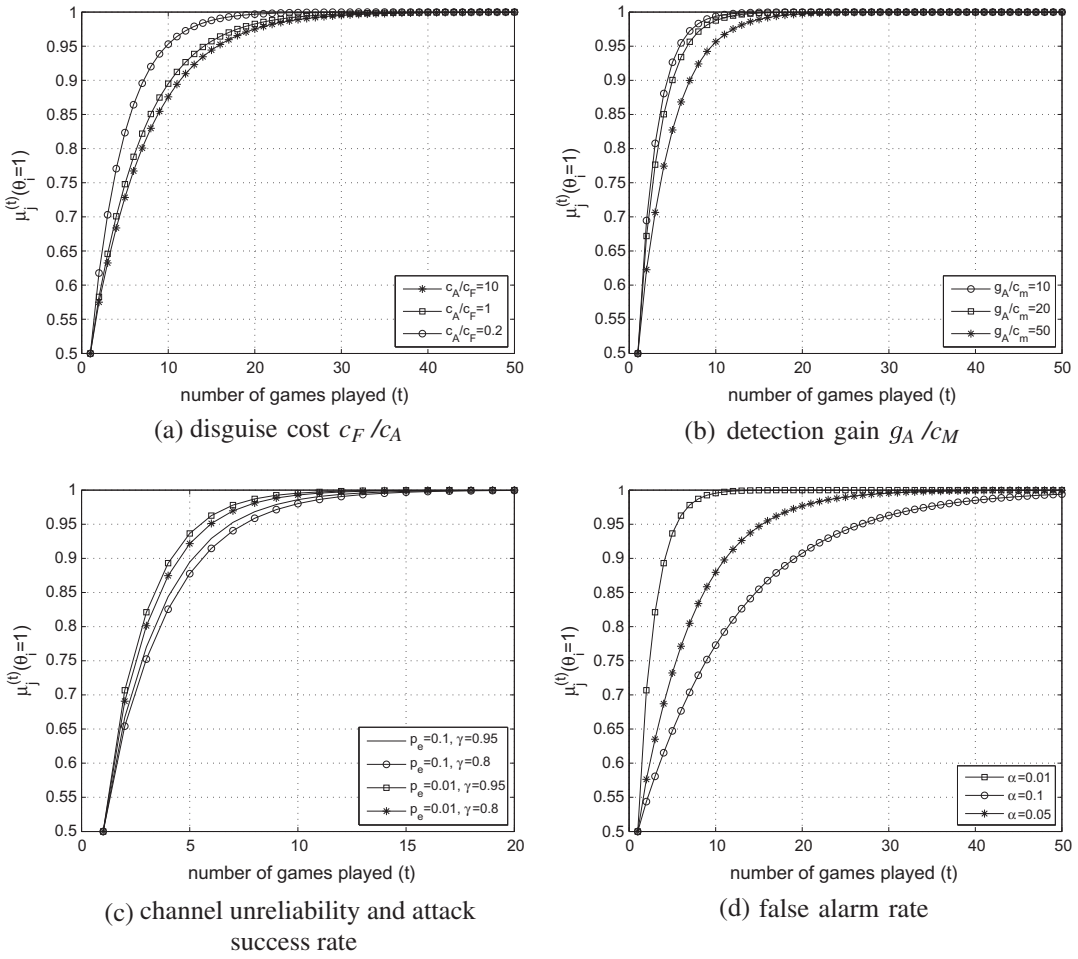


Fig. 2. Equilibrium strategy q vs. the attack success rate in malicious node detection game.



(a) constant monitoring (b) PBE

Fig. 3. Belief system update with different attack gains.

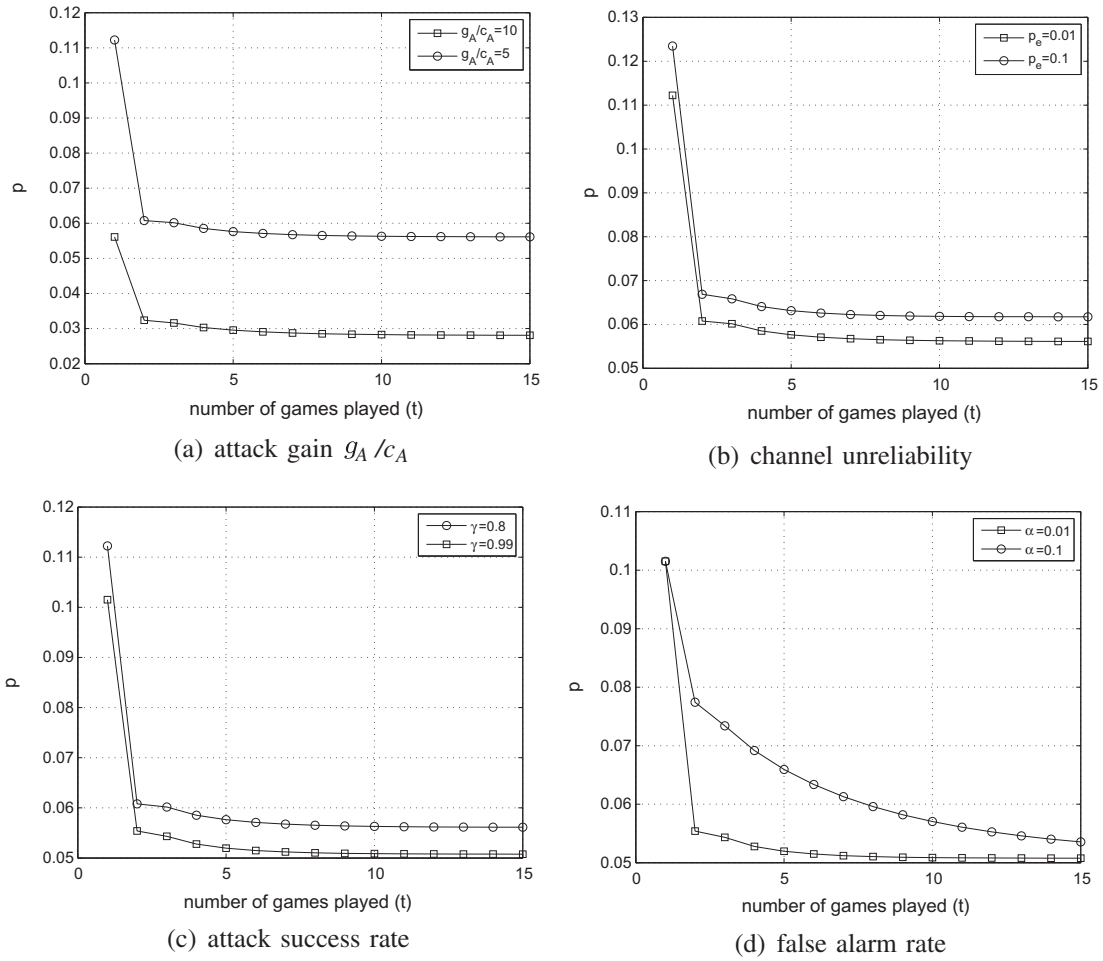


(a) disguise cost c_F/c_A (b) detection gain g_A/c_M

(c) channel unreliability and attack success rate (d) false alarm rate

Fig. 4. Effects of parameters on belief system update.

in Fig. 5(a). A large attack gain means more payoff gained from an attack, which implies less number of attacks are needed. Hence p should be smaller. Figs. 5(b) and (c) indicate that node i should attack less frequently under a reliable channel as every attack is more likely to be successful. However, as suggested in Fig. 4(d), if the false alarm rate is high for the regular node, the malicious node can take advantage of it and attack more often.



(a) attack gain g_A/c_A (b) channel unreliability

(c) attack success rate (d) false alarm rate

Fig. 5. Effects of parameters on the PBE strategy.

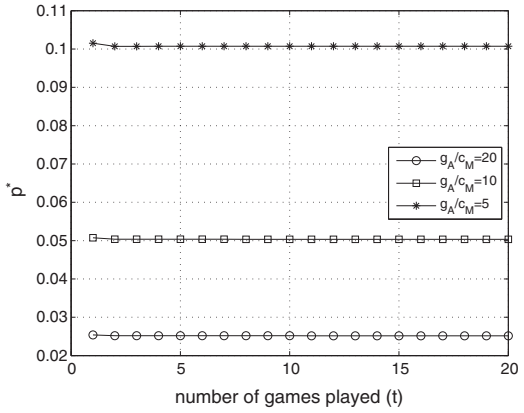
5.2. Post-detection game

After the belief system of node j converges ($\theta_i \geqslant 0.99$), we can safely conclude that node j has detected the malicious node. Therefore, the post-detection game starts. To show the continuity, at the beginning of the post detection game, node i sticks to its PBE strategy.

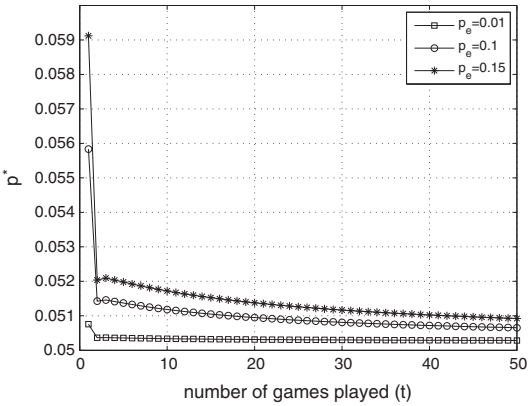
Fig. 6 presents how the attack probability $p^*(t)$ evolves to the SPNE strategy from the PBE. It is clear in the plots that in the SPNE, node i should decrease its attack probability to avoid isolation. Fig. 6(a) shows a larger detection gain that corresponds to a smaller attack rate; thus in the equilibrium, the payoffs for node j will not increase due to the large detection gain. Fig. 6(b) states that if the channel is lossy, node i should attack more often. The reason behind this claim is that the more unreliable the channel is, the less probable node j can accurately observe an attack. Plots in Fig. 6(c) are obtained from detection gain equals to 5. This figure shows that the equilibrium is not sensitive to the initial value and threshold of the coexistence index $\mathcal{E}_i$.

The expected length of the post-detection game is shown in Fig. 7. First, the figure states that the less errors (i.e., less channel loss and more successful attack) in the system, the longer the post-detection games can be played. Second, the length of the game grows with the attack gain. This interesting phenomenon can be explained in the following way. The larger attack gain enables the malicious node to attack less while keeping its payoff high. Thus, more often, the malicious node will play as a regular node to avoid isolation. This will increase the time for the regular and malicious nodes to coexist. This property can be used to extend the lifetime of the network.

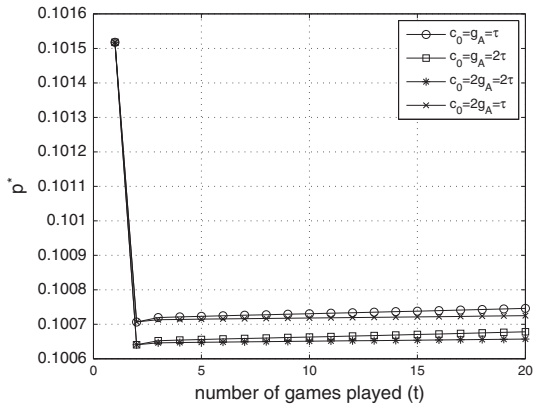
Last but not the least, we show how the network throughput can benefit from coexistence in Fig. 8. We use the throughput of no co-existence as the baseline and define throughput gain as the ratio of added network throughput over the baseline when co-existence strategy is played. Similar observations can be made as the game length property. With a larger attack gain, the malicious node decreases its attack rate and does more packet forwarding as a regular node. Therefore, the malicious nodes can be utilized to increase the throughput more often as the attack gain grows. The throughput gain property illustrates clearly that malicious and regular nodes can coexist, and the coexistence equilibria improve the throughput of the network.



(a) detection gain g_A/c_M



(b) channel unreliability



(c) coexistence index

Fig. 6. Effects of parameters on the SPNE strategy.

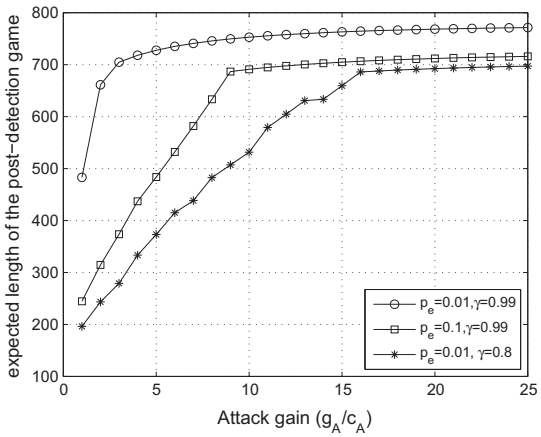


Fig. 7. Expected length of the post-detection game.

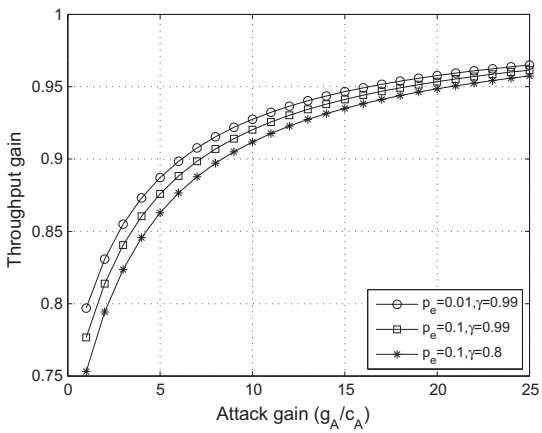
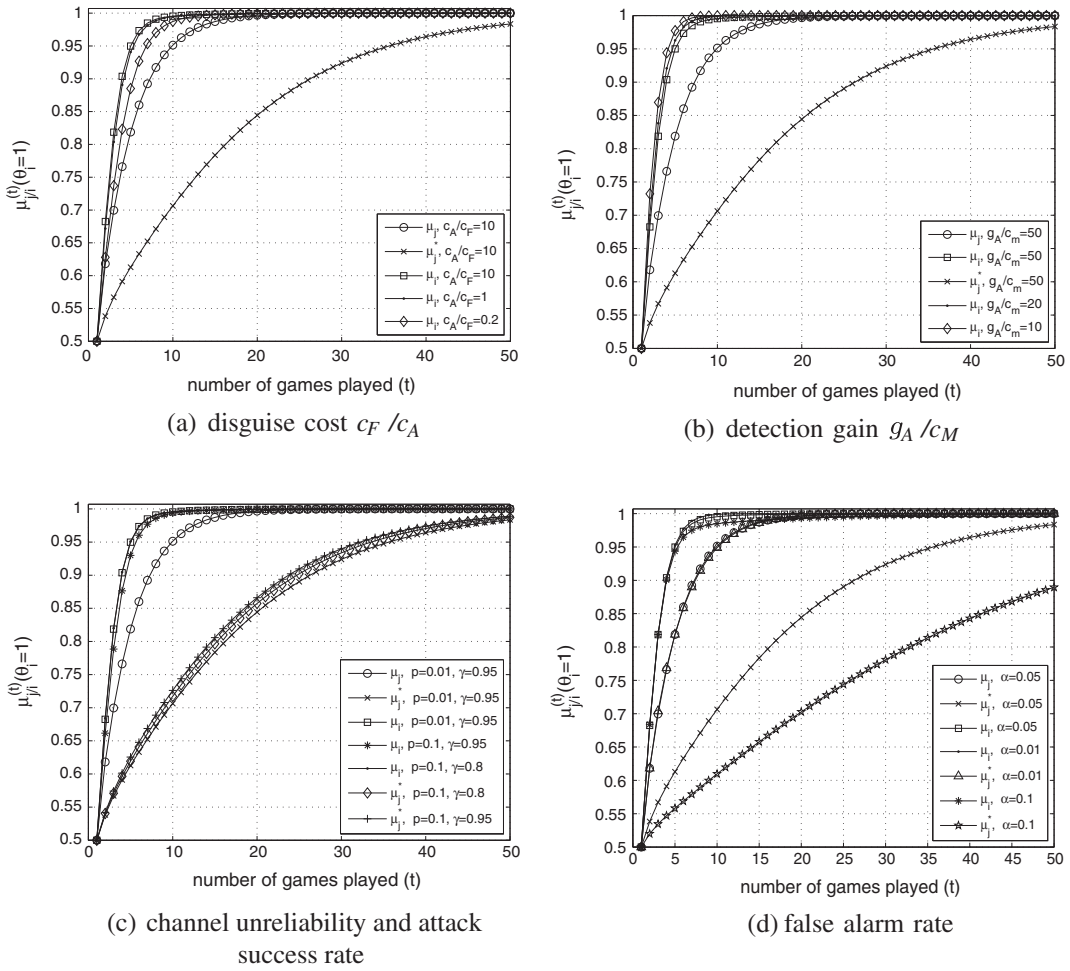


Fig. 8. Throughput gain.

5.3. Characteristics of MPBNE

We study the characteristics of the Markov Perfect Bayes–Nash Equilibrium. In particular, we are interested in the properties of node *i*'s belief update system (i.e., belief about belief) and how the introduction of node *i*'s belief would affect the results we obtained in Section 5.1.

In Fig. 9, we study node *i*'s belief system in the MPBNE. To better show the properties of node *i*'s belief system in the MPBNE, we also present node *j*'s belief system. In



(a) disguise cost c_F /c_A

(b) detection gain g_A /c_M

(c) channel unreliability and attack success rate

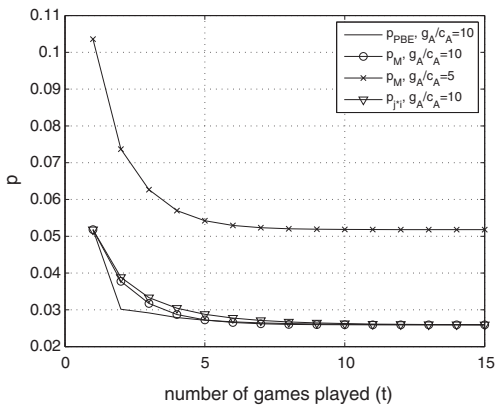
(d) false alarm rate

Fig. 9. Node *i*'s belief system update in the Markov Perfect Bayes–Nash Equilibrium.

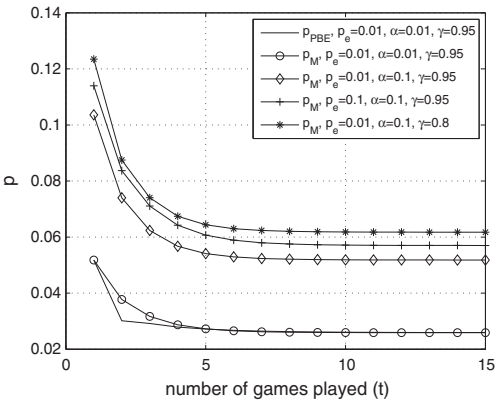
particular, we plot μ_i as node *i*'s belief system in MPBNE according to Eq. (47), μ_j as node *j*'s belief system in PBE as stated in Eq. (12) and μ_j^* as node *j*'s belief system update in the MPBNE as a result of node *i*'s actions in the belief about belief model. A common observation is that node *i*'s belief μ_i converges much faster than the belief μ_j in PBE, which means that node *i* holds a false belief that node *j* can identify its malice quicker than node *j* actually could. As a result of the inaccuracy in node *i*'s belief, it takes longer time for node *j* to form a belief on node *i*. This is evident from the plots that show μ_j^* converges much slowly than it does in PBE, when node *i* does not employ any belief system.

In addition, Fig. 9 shows some similar properties of node *i*'s belief system to what we have observed in Fig. 4. For example, Fig. 9(b) indicates a larger detection gain will force node *i*'s belief system converge quicker. Fig. 9(c) and (d) infer that reliable channel, high attack success rate and accurate detection (low false alarm rate) will also induce a fast convergence of μ_i. However, the only discrepancy is with the disguise cost; for node *i*, a high disguise cost makes update of μ_i slow, while for node *j*, a high disguise cost helps μ_i converge faster. The reason lies in the inaccuracy of node *i*'s belief system. From our previous discussion, it is stated that when the observed payoff is $-c_F$, node *i* cannot predict what node *j*'s action is. Thus, an internal error resides in node *i*'s belief system, and this error is amplified when c_F is large (i.e., c_F, takes a high weight in the payoff), which corresponds to a large disguise cost.

The properties of the MPBNE strategy are further investigated in Fig. 10. Once again, similarity is found between Figs. 10(a) and 5(a), as well as Fig. 10(b) and Fig. 5(b)–(d). Both the MPBNE strategy attack probability, denoted as p_{M} and the PBE strategy attack probability in Fig. 5 will increase with smaller attack gain and attack success rate, as well as larger channel error rate and false alarm rate. Moreover, it is noted that p_{M} is smaller than what node *j* believes it would be (denoted as p_{j*i} in the Figs. 10(b)). In addition, p_{M} is larger than the PBE strategy attack probability p_{PBE} in the first several stage games, however, as the games repeat, p_{M} drops below p_{PBE}. This interesting observation implies that when node *i* implements the belief system, it attacks more aggressively (than without the belief about belief model, i.e., in PBE) in the first several games, because it believes node *j* is far from reaching a successful



(a) attack gain g_A/c_A



(b) channel unreliability, attack success rate and false alarm rate

Fig. 10. Effect of parameters on the Markov Perfect Bayes–Nash Equilibrium strategy.

detection. As the game unfolds, node i adjusts its attack rate to prevent from detection. The difference between p_M and p_{PBE} also explains why node j's belief system alters in the MPBNE as shown in Fig. 9.

5.4. Transition from detection game to post-detection games

Our discussions above are focused on how the involvement of node i's belief system would make the MPBNE different from the PBE. However, since the detection of the malicious node is not the only aim of this research, we are also motivated to find the link between the MPBNE and/or PBE in the detection game and the SPNE in the post detection game. Fig. 11 shows the equilibrium strategy profiles in terms of attack probability. It is clearly evident from the plot that although in MPBNE, node i attacks less often in PBE, in order to reach SPNE, node i still needs to further lower its attack probability. As a matter of fact, the post-detection game is initialized by node j when its belief about node i's malice reaches a threshold value ($\mu_j(\theta_i) > 0.99$ in our setting). However, this information is never revealed to node i, so that node i has no idea if the post-detection game has started or not. When node i is also equipped with the belief system, it can make a prediction on when the post-detection game starts based on its belief about node j's belief. For example, if node i's belief $\mu_i(\mu_j(\theta_i)) > 0.99$, node i might assume that the post-detection game has begun and adjust its strategy profile accordingly.

Fig. 12 examines the transition process through simulation. In our simulation, once $\mu_i(\mu_j(\theta_i))$ reaches 0.99, node i's attack probability is set to be the same as the probability in SPNE (denoted as p_{SPNE}). Despite the change, node j still sticks to its criteria and does not start the post-detection game until its belief crosses the threshold. In other words, in this setting, node i deviates from the MPBNE and plays the SPNE strategy even when node j still plays the detection game. Fig. 12(a) shows although node i deviates from MPBNE, the attack probability node j believes node i takes ($p_{i\to j}$) and p_{PBE} are very close. Furthermore, Fig. 12(b) indicates that when node i adheres to p_{SPNE}, node j's belief updates are slightly slower than that in the PBE. The simulation results suggest that with node i's belief system, malicious node detection game and the post-detection game can be integrated with an effective transition process.

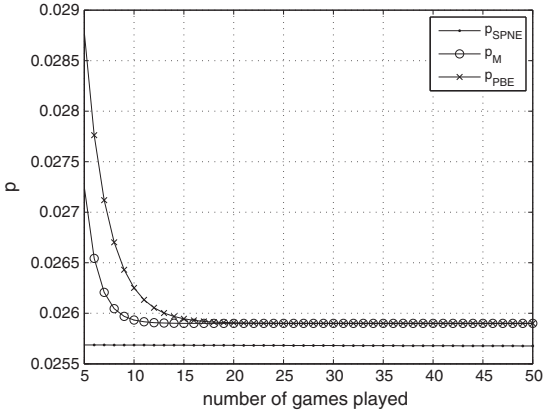
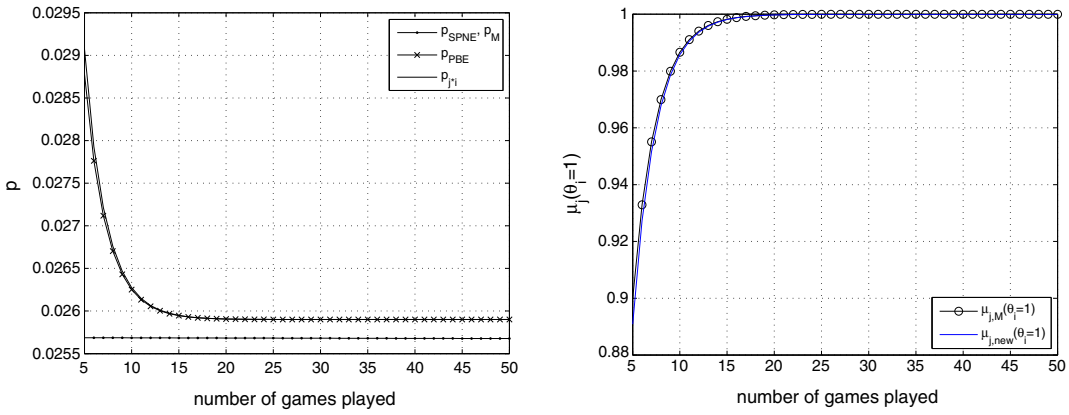


Fig. 11. Comparison of node i's strategy profiles.

6. Discussions and conclusions

6.1. Discussions

There are many types of attacks in wireless networks. Each attack may have its unique features and require different techniques to detect and defend. One of the goals of this research is to present a general model and approach to analyze the detection process. In particular, in order to apply the detection game model, the monitoring node j must be able to observe and identify attacks. The ability to observe means the physical capability to sense the occurrence of potential attacks. In the context of wireless networks, such ability generally translates to the location proximity falling within the communication range of malicious node. However, being close to the attack does not necessarily lead to identification of the attack. In order to identify the attack, the node j must be able to analyze



(a) Node i's strategy profiles. (b) Node j's belief update.

Fig. 12. Effects of the integration of the detection and post-detection games.

the observed signal (i.e., the potential attacking action) and deduce if it is an attack or not. An example of analyzing the observed signal is observing data dropping and alternation in the context of data forwarding. In this regard, the proposed detection model cannot be applied directly to a number of attacks that are not identifiable at the scene, e.g., Sybil attack, man-in-the-middle attack and traffic hijacking attack (advertising false routes). In order to use the model, the node j should be furnished with appropriate mechanism to verify certain attacks.

As far as the post-detection game is concerned, it requires that the attacker has the ability to camouflage itself as a benign node. According to the game model, if every action the attacker takes harms the utility of its opponent, the post-detection game will be terminated soon, because the attacker is not useful at all and cannot be exploited. A lot of the attack types fall into this category and cannot apply the post-detection game model, including the aforementioned Sybil attack, man-in-the-middle attack and traffic hijacking attack (advertising false routes).

6.2. Conclusions

In this paper, we apply game theory to study the coexistence of malicious and regular nodes in a wireless network with unreliable channels. We formulate a malicious node detection game and a post-detection game played by the regular and malicious nodes. While both games are of imperfect information type, we show that the former game has a mixed strategy perfect Bayesian Nash equilibrium and provide a solution to achieve that equilibrium. For the latter game, a coexistence index is proposed. We also prove that keeping the coexistence index above a threshold, the post-detection game has a subgame perfect Nash Equilibrium which is also the coexistence equilibrium for malicious and regular nodes. We also propose a belief about belief system that can be used by the malicious node to predict if it has been detected. We also prove the existence of a Markov Perfect Bayes–Nash Equilibrium when both nodes constantly update their beliefs. The properties of the equilibrium are studied and it is shown that how the detection of the malicious node can be delayed. Simulations are provided to illustrate the properties of the equilibria. In particular, we show how the system parameters such as attack gain, attack success rate, detection gain, and channel loss affect the convergence of the games and the equilibrium strategies. Simulation results also state that the coexistence equilibrium helps to extend the length of the games and improves the throughput of the network. With the help of the proposed belief about belief system, the malicious node is able to adjust its strategy in the game and finally the detection game and post-detection game are integrated with effective transition.

Acknowledgements

This research was sponsored by the Air Force Office of Scientific Research (AFOSR) under the Federal Grant No. FA9550-07-1-0023. A preliminary version of this work appeared in International Conference on Game Theory for Networks (GameNets) 2009 [29]. Approved for Public Release; Distribution Unlimited: 88ABW-2014-2735 dated 05 June 2014.

Appendix A

Table A.1.

Theorem 3. *The expect length of the post-detection game is*

$$\sum_{\eta>0}\eta\frac{\binom{\eta}{\hat{n}_F}-\sum_{d}^{\hat{n}_F-1}\binom{\frac{(c_0-\tau)/c_F+d}{g_A/c_F}+d}{d}\binom{\eta-\frac{(c_0-\tau)/c_F+d}{g_A/c_F}-d}{\hat{n}_F-d}}{2^{\eta}}.$$

Proof. Let η be a random variable representing the first hitting time. We assume that time is divided into slots and each slot represent a stage game. It is easy to see that $\eta=\hat{n}_F+\hat{n}_A$. At every slot, the random process has 2 possible evolution directions, i.e., $\hat{n}_F+1$ or $\hat{n}_A+1$. Therefore, for η slots, there are 2^{η} possible realizations.

We try to calculate how many paths hit the boundary exactly on the η^{th} slot. The following notations are made. Let $m=\frac{g_A}{c_F}, s=(c_0-\tau)/c_F$ and m,s are integers. In Fig. A.1,

Table A.1
Notations used.

Node i	(potential) malicious node, attacker
Node j	regular node, monitor
θ_i	type of node i, 1 for malicious, 0 for regular
u_i, u_j	payoff of node i or j in the stage game
g_A	gain of successfully attack for node i
c_A	cost of any attack for node i
c_F	cost of forward (not attack) for node i
c_M	cost of monitoring for node j
ϕ	the belief of node i being malicious in stage game
γ	attack success rate for node i
p_e	channel error rate
α	false alarm rate for node j
a_i, a_j	action profile for node i or j
$\hat{a}_i(t)$	node i's action observed by node j in stage t
o_i	node i's observation of its payoffs
$\mu_j(\theta_i)$	belief node j holds about node i's type in the dynamic game
$\mu_i(\mu_j(\theta_i))$	belief node i holds about node j's belief in the dynamic game (used to derive the MPBNE)
σ_i, σ_j	node i or j's strategy profile
p, q	random variable of probability node i attacks or node j monitors in the malicious node detection game
$p^*(t), q^*(t)$	random variable of probability node i attacks or node j monitors in the post-detection game
$\tilde{p}, \tilde{q}$	random variable of probability node i attacks or node j monitors in the malicious node detection game with node i's belief about belief model
C_i	coexistence index
μ_j^*	node j's belief about node i's type when node i uses belief about belief model
p_{j*i}	node j's belief about node i's attack probability when node i uses belief about belief model
p_{M}	the value of node i's attack probability in MPBNE
p_{PBE}	the value of node i's attack probability in PBE
p_{SPNE}	the value of node i's attack probability in SPNE

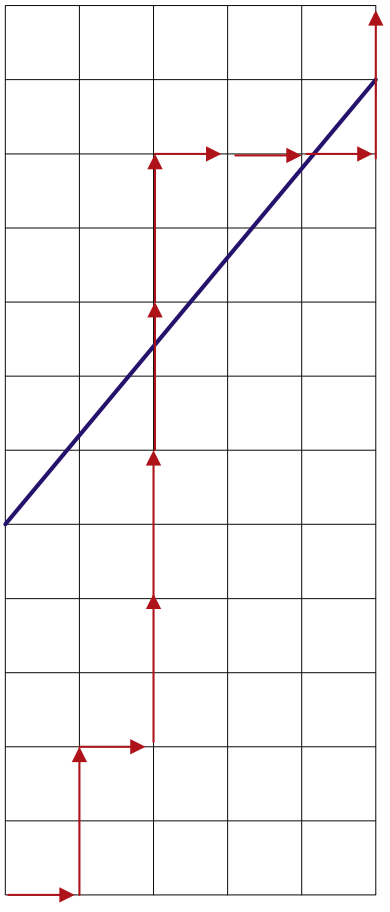


Fig. A.1. Realizations of the random walk.

we interpret how to move on a grid according to a random process. Consider a random walk from the left bottom point. If $\hat{n}_F$ increases, move one block right. If $\hat{n}_A$ increases, move m blocks up. While each block is a squarelet with the length c_F, the width of the grid is $\hat{n}_F c_F$, the height is $g_A \hat{n}_A$, and diagonal line represents $\mathcal{E}_i = \tau$. Each walk consists of η moves and must end on or beyond the upper rightmost corner. What we are interested in is the number of monotonically increasing paths that wholly fall under the diagonal line, because each of those paths is a realization of the random process which hits the boundary for the first time at the η^{th} slot.

While counting the number of realizations under the diagonal line might be difficult, we calculate the realizations that do cross the line. Let the number of realizations crossing the line be M, the number of realizations under the line is then $C_n^{\hat{n}_F} - M$, where the combinatorial number $C_n^{\hat{n}_F}$ denotes the total number of possible realizations on the grid. Consider a sample realization crossing the line as shown in Fig. A.1. Let d be the number of horizontal steps taken in the path before hitting the diagonal line. To hit the line, at least $\frac{s+d}{m}$ vertical steps should be taken, covering a total height of $(d+s)c_F$. The total number of such paths is $\sum_d C^d_{\frac{s+d}{m}+d}$. After hitting the line, the rest of the path should consist of $\hat{n}_F - d$ vertical steps and the total number of moves left is $\eta - \frac{s+d}{m} - d$. So, the total number of paths that cross the diagonal line is $M = \sum_d^{\hat{n}_F - 1} C^d_{\frac{s+d}{m}+d} C^{\hat{n}_F - d}_{\eta - \frac{s+d}{m} - d}$.

To sum up, out of 2^η realizations, $C_n^{\hat{n}_F} - \sum_d^{\hat{n}_F - 1} C^d_{\frac{s+d}{m}+d} C^{\hat{n}_F - d}_{\eta - \frac{s+d}{m} - d}$ realizations hit the diagonal line for the first time at the η^{th} move. The probability of game length being

η is then $\frac{C_{\eta}^{\hat{n}_F} - \sum_{d}^{\hat{n}_F-1} C_{\frac{s+d}{m}+d}^{d} C_{\eta-\frac{s+d}{m}-d}^{\hat{n}_F-d}}{2^{\eta}}$. Finally, we can express the expected length of the post detection game as

$$E[length] = \sum_{\eta>0} \eta \frac{C_{\eta}^{\hat{n}_F} - \sum_{d}^{\hat{n}_F-1} \binom{\frac{s+d}{m}+d}{d} \binom{\eta-\frac{s+d}{m}-d}{\hat{n}_F-d}}{2^{\eta}}. \quad \square \qquad (1)$$

References

[1] A. Agah, S.K. Das, K. Basu, M. Asadi, Intrusion detection in sensor networks: a non-cooperative game approach, in: Proceedings of IEEE NCA, 2004, pp. 343–346.

[2] E. Altman, A. Kumar, C. Singh, R. Sundaresan, Spatial SINR Games Combining Base Station Placement and Mobile Association, in: Proceedings of IEEE Infocom, 2009.

[3] E. Altman, K. Avrachenkov, A. Garnaev, Jamming in wireless networks under uncertainty, in: Proceedings of Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT), 2009, pp. 1–7.

[4] B. Awerbuch, R. Curtmola, D. Holmer, C. Nita-rotaru, H. Rubens, ODSBR: an on-demand secure Byzantine resilient routing protocol for wireless ad hoc networks, ACM Trans. Inf. Syst. Secur.ty 10 (4) (2008) (no. 6).

[5] L. Buttyán, J.P. Hubaux, Stimulating cooperation in self-organizing mobile ad-hoc networks, ACM/Kluwer Mobile Networks Appl. 8 (5) (2003) 579–592.

[6] J. Crowcroft, R. Gibbens, F. Kelly, S. Ostring, Modelling incentives for collaboration in mobile ad hoc networks, Perform. Eval. 57 (4) (2004) 427–439.

[7] D. Fudenberg, J. Tirole, Game Theory, MIT press, Cambridge, MA, 1991.

[8] Z. Han, N. Marina, M. Debbah, A. Hjørungnes, Physical layer security game: interaction between source, eavesdropper, and friendly jammer, EURASIP J. Wireless Commun. Networking 2009 (2009) (Article ID 452907).

[9] Y.C. Hu, A. Perrig, D. Johnson, Packet leashes: a defense against wormhole attacks in wireless networks, in: Proceedings of IEEE INFOCOM, 2003, pp. 1976–1986.

[10] J.J. Jaramillo, R. Srikant, DARWIN: distributed and adaptive reputation mechanism for wireless ad-hoc networks, in: Proceedings of ACM MobiCom, 2007, pp. 87–97.

[11] Z. Ji, W. Yu, K.J.R. Liu, Cooperation enforcement in autonomous MANETs under noise and imperfect observation, in: Proceedings of IEEE Secon, 2006, pp. 460–468.

[12] L.P. Kaelbling, M.L. Littman, A.R. Cassandra, Planning and acting in partially observable stochastic domains, Artif. Intell. 101 (1998) 99–134.

[13] S. Kim, Multi-leader multi-follower Stackelberg model for cognitive radio spectrum sharing scheme, Comput. Networks 56 (17) (2012) 3682–3692.

[14] M. Kodlialam, T.V. Lakshman, Detecting network intrusion via sampling: a game theoretic approach, in: Proceedings of IEEE Infocom, 2003, pp. 1880–1889.

[15] D.M. Kreps, R. Wilson, Sequential equilibria, Econometrica 50 (4) (1982) 863–894.

[16] F. Li, J. Wu, Hit and Run: a Bayesian game between malicious and regular nodes in MANETs, in: Proceedings of IEEE Secon, 2008, pp. 432–440.

[17] X.-Y. Li, Y. Wu, P. Xu, G. Chen, M. Li, Hidden information and actions in multi-hop wireless ad hoc networks, in: Proceedings of ACM Mobihoc, 2008, pp. 283–292.

[18] P. Liu, W. Zhang, M. Yu, Incentive-based modeling and inference of attacker intent, objectives, and strategies, ACM Trans. Inf. Syst. Secur. 56 (3) (2005) 78–118.

[19] Y. Liu, C. Comaniciu, H. Man, A Bayesian game approach for intrusion detection in wireless ad hoc networks, in: Proceedings of ACM GameNets, 2006.

[20] A.B. Mackenzie, L.A. DaSilva, Game Theory for Wireless Engineers, Morgan & Claypool Publishers, San Rafael, California, 2006.

[21] P. Michiardi, R. Molva, Analysis of coalition formation and cooperation strategies in mobile ad hoc networks, Ad Hoc Networks 3 (2005) 193–219.

[22] F. Milan, J.J. Jaramillo, R. Srikant, Achieving cooperation in multihop wireless networks of selfish nodes, in: Proceedings of ACM GameNets, 2006.

[23] K.C. Nguyen, T. Alpcan, T. Başar, Security games with incomplete information, in: Proceedings of IEEE ICC, 2009, pp. 1–6.

[24] M.J. Osborne, An Introduction to Game Theory, Oxford University Press, New York, NY, 2004.

[25] S. Sengupta, M. Chatterjee, K.A. Kwiat, A game theoretic framework for power control in wireless sensor networks, IEEE Trans. Comput. 59 (2) (2010) 231–242.

[26] V. Srinivasan, P. Nuggehalli, C.F. Chiasserini, R.R. Rao, Cooperation in wireless ad hoc networks, in: Proceedings of IEEE Infocom, 2003, pp. 807-817.

[27] B. Sun, Y. Guan, J. Chen, U.W. Pooch, Detecting black-hole attack in mobile ad hoc networks, in: IEEE Europe Personal Mobile Communications Conference, 2003, pp. 490–495.

[28] G. Theodorakopoulos, J.S. Baras, Malicious users in unstructured networks, in: Proceedings of IEEE Infocom, 2007, pp. 884–891.

[29] W. Wang, M. Chatterjee, K. Kwiat, Coexistence with malicious nodes: a game theoretic approach, in: Proceedings of GameNets, 2009.

[30] W. Wang, S. Eidenbez, Y. Wang, X.-Y. Li, OURS: Optimal unicast routing system in non-cooperative wireless networks, in: Proceedings of ACM Mobicom 2006, pp. 402–413.

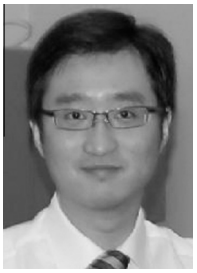
[31] B. Wu, J. Chen, J. Wu, M. Cardei, A Survey on Attacks and Countermeasures in Mobile Ad Hoc Networks, Wireless Network Security, Springer, 2007. pp. 103–135.

[32] P. Xu, X.-Y. Li, S. Tang, Efficient and Strategyproof spectrum allocations in multichannel wireless networks, IEEE Trans. Comput. 60 (4) (2011) 580–593.

[33] J. Zhang, Q. Zhang, Stackelberg Game for utility-based cooperative cognitive radio networks, in: Proceedings of ACM Mobihoc, 2009.

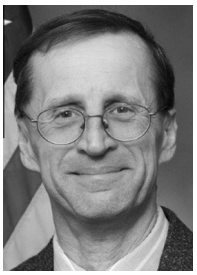
[34] S. Zhong, L. Li, Y. Liu, Y. Yang, On designing incentive-compatible routing and forwarding protocols in wireless ad-hoc networks—an integrated approach using game theoretical and cryptographic techniques, in: Proceedings of ACM Mobicom, 2005, pp. 117–131.

[35] Q. Zhu, C. Fung, R. Boutaba, T. Başar, A game-theoretical approach to incentive design in collaborative intrusion detection networks, in: Proceedings of GameNets, 2009.



Wenjing Wang obtained his Ph.D. from the School of Electrical Engineering and Computer Science, at the University of Central Florida (UCF). He holds MS degree from UCF and BS degree from University of Science and Technology of China (USTC), all in Electrical Engineering. His research interests are in the broad areas of wireless communication and networking, with particular emphasis on network privacy and security, cognitive radio networks, and vehicular networks. He is currently with Advanced and Emerging Technology Group, Blue Coat Systems in Sunnyvale, California.

Mainak Chatterjee is an Associate Professor in the department of Electrical Engineering and Computer Science, University of Central Florida, Orlando. He received the BSc degree in physics (Hons.) from the University of Calcutta, the ME degree in electrical communication engineering from the Indian Institute of Science, Bangalore, and the PhD degree from the Department of Computer Science and Engineering from the University of Texas at Arlington. His research interests include economic issues in wireless networks, applied game theory, cognitive radio networks, dynamic spectrum access, and mobile video delivery. He has published over 125 conferences and journal papers. He got the Best Paper Awards in IEEE Globecom 2008 and IEEE PIMRC 2011. He is the recipient of the AFOSR sponsored Young Investigator Program (YIP) Award. He co-founded the ACM Workshop on Mobile Video (MoVid). He serves on the editorial board of Elsevier's Computer Communications and Pervasive and Mobile Computing Journals. He has served as the TPC Co-Chair of several conferences including IEEE WoWMoM 2011, WONS 2010, IEEE MoVid 2009, Cognitive Radio Networks Track of IEEE Globecom 2009 and ICCCN 2008. He also serves on the executive and technical program committee of several international conferences.



Kevin A. Kwiat has been with the U.S. Air Force Research Laboratory (AFRL) in Rome, New York for over 30 years. Currently is assigned to the Cyber Assurance Branch. He received the BS in Computer Science and the BA in Mathematics from Utica College of Syracuse University, and the MS in Computer Engineering and the Ph.D. in Computer Engineering from Syracuse University. He holds 4 patents. In addition to his duties with the Air Force, he is an adjunct professor of Computer Science at the State University of New York at Utica/Rome, an adjunct instructor of Computer Engineering at Syracuse University, and a Research Associate Professor with the University at Buffalo. He is an advisor for the National Research Council. He has been by recognized by the AFRL Information Directorate with awards for best paper, excellence in technology teaming, and for outstanding individual basic research. His main research interest is dependable computer design.

Qing Li, Chief Scientist and Vice President of Advanced Technologies, an industry veteran with over 20 years of experience, has spent the past 10 years designing and developing industry leading technologies and products at Blue Coat Systems. Qing was fully responsible for the IPv6 secure proxy and WAN optimization technology and product lines at Blue Coat. He produced the industry's first IPv6 Secure Web Gateway product in 2009, and received the IPv6 Application Solution Pioneer Award from the IPv6 Forum in April 2010. Subsequently he produced the industry's first IPv6 WAN Optimization appliance in 2011, and in early 2012 he produced and released the industry's first IPv6 visibility solution. And in March 2014 he lead the effort that produced Blue Coat's first 10 Gbps visibility and QoS solution. He has been an active speaker at industry and academia conferences and is an active voice in the technology media around the world. In the past 3 years Qing's research has concentrated on emerging technologies including advanced application classification algorithms, mobile security, SSL interception and data analytics. His innovations have transformed the Blue Coat technology and product landscape. Qing is a published author, most notably the two-volume reference series on IPv6. Volume I, IPv6 Core Protocols Implementation, and Volume II, IPv6 Advanced Protocols Implementation, were published in October 2006 and in April 2007 respectively, by Morgan Kaufmann Publishers. In 2003 Qing published the embedded systems development book titled Real-Time Concepts for Embedded Systems, which has served as reference text in the industry as well as in universities. Qing was also a contributing author to the first of its kind book entitled Handbook of Networked and Embedded Control Systems, published in June of 2005 by Springer-Verlag. Qing holds 12 US patents with many more pending in the areas of security and networking.