

Game Theoretic Modeling of Malicious Users in Collaborative Networks

George Theodorakopoulos, *Member, IEEE*, and John S. Baras, *Fellow, IEEE*

Abstract—If a network is to operate successfully, its users need to collaborate. Collaboration takes the form of following a network protocol and involves some resource expenditure on the part of the user. Therefore, users cannot automatically be expected to follow the protocol if they are not forced to. The situation is exacerbated by the presence of malicious users whose objective is to damage the network and increase the cost incurred by the legitimate users. The legitimate users are, at least initially, unaware of the type (legitimate or malicious) of the other users.

Our contribution is a model for the strategic interaction of legitimate and malicious users as described above. The model is based on repeated graphical games with incomplete information. We describe and analyze two specific instantiations, aiming to demonstrate the model's expressive power and tractability. The main benefit we see from using game theory for this essentially security problem is the ability to bound the damage caused by the malicious users.

Index Terms—collaborative networks, incomplete information, repeated games, security

I. INTRODUCTION

TODAY'S technology has enabled individuals to communicate and collaborate with one another to a much greater extent than ever before. Of course, successful collaboration needs user resources to be dedicated, whether these are the (human) user's time and effort, or a device's battery, or an infrastructure's bandwidth. Even assuming that all collaborating parties willingly share these resources, malicious individuals or organizations have much to gain from exploiting the willingness of others to collaborate. To aggravate things, malicious entities are usually much more motivated, knowledgeable, and sophisticated than the average legitimate user. Any strategy, best practice, or protection used by the legitimate users is sooner or later compromised by the malicious ones, and exploited.

Manuscript received August 15, 2007; revised March 10, 2008. This paper was presented in part at the IEEE/CreateNet International Conference on Security and Privacy in Communication Networks, Baltimore, MD, USA, August 2006, and in part at the IEEE International Conference on Computer Communications, Anchorage, AK, USA, May 2007. This research was supported through collaborative participation in the Communications and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The research was also supported by the U.S. Army Research Office under grant No DAAD19-01-1-0494.

George Theodorakopoulos was with the Institute for Systems Research and the Department of Electrical and Computer Engineering, University of Maryland College Park, USA. He is now with the School of Computer and Communication Sciences at the Ecole Polytechnique Fédérale de Lausanne, Switzerland (e-mail: george.theodorakopoulos@epfl.ch).

John S. Baras is with the Institute for Systems Research and the Department of Electrical and Computer Engineering, University of Maryland College Park, USA (e-mail: baras@isr.umd.edu).

Digital Object Identifier 10.1109/JSAC.2008.080928.

The ideal thing would be to provide a recommendation to the legitimate users that would at the same time be simple, and give provable guarantees against any malicious user strategy. We believe that game theory is a promising model to examine such situations. In this paper, we make a first attempt at a model that incorporates these characteristics.

We model networks consisting of users who try to maximize a personal benefit through their participation in the network. Associated with the network, there is a protocol that describes the available actions that each user has at his disposal. The protocol also defines a behavior that is expected of each user; this behavior involves some cost on the user's part. However, there is no centralized control of the users' decisions, so each user is free to choose his actions in a selfish or even malicious way. If we call "collaboration" the act of following the protocol, we can use the term "collaborative networks" to collectively address networks where these circumstances hold. Examples of such networks are wireless ad hoc networks, peer-to-peer networks, but also social networks or online marketplace communities, as long as there is a protocol or a pre-specified behavior expected from the users, which cannot be enforced and costs something to adhere to.

Each user receives a payoff that reflects the benefit he derives from being part of the network, typically in terms of the services that the rest of the network provides to him through the operation of the protocol. In general, the level of service that a user receives may depend on many or all other users of the network; we model these dependencies as a graphical game. For example, if the protocol is about packet forwarding, then all users on the path from the source to the destination can affect the level of the service. However, the users may not want to follow the protocol, since that will entail resource consumption of some sort, such as energy consumption in the case of packet transmission in wireless networks. The cost associated with following the protocol is also reflected in the payoff, as a decrease in the derived benefit.

The users can have varying degrees of selfishness. Selfishness can be quantified as the level of service that a user desires from the network before he decides to follow the protocol himself. In the examples that we will analyze later, we will only consider users who are equally selfish, but we will point out how different degrees of selfishness can be incorporated. Nevertheless, we will make an important distinction between selfish users, who care about the level of service they receive, and malicious users. More precisely, we model malicious users as users whose payoff does not depend at all on the level of service that they receive. Our choice is to model them as playing a zero-sum game against legitimate players, i.e., their

objective is directly opposite of whatever the objective of the legitimate users is.

A crucial modeling decision is to make the payoff matrix of each user known only to himself. This reflects the lack of knowledge that users have for each other in a network without centralized control. To enhance the capabilities of the malicious users, we will assume that they know the payoffs of everybody. That is, they know who is malicious and who is only selfish. This is a game of incomplete information, and the information is asymmetric, since malicious users know the payoffs of everyone, but each selfish user only knows his own payoff.

When solving the models we do not attempt to find the set of all equilibria. Instead, we assume that the legitimate users follow simple processes that do not place excessive demands on users' hardware in terms of memory or computing power. So, we only look for equilibria that arise through these processes. Also, the malicious users know what simple processes the legitimate users will follow, and take this knowledge into account when choosing a strategy to maximize their payoff. This extra knowledge gives the game the flavor of a Stackelberg game, with the Bad users being leaders and the Good users followers, but the Good user strategies are limited so we choose not to cast the game as a Stackelberg one.

Since the network operates for a long time, the users choose actions repeatedly. We assume that time is discrete, and users choose their actions simultaneously in rounds. At each round, users remember the past history of their own actions and of the actions of the users that affect their payoff. They can use this history to help them choose their action in the next round. Given the incomplete information assumption, repeated interactions can help increase the knowledge of non-malicious users.

II. FORMAL DESCRIPTION OF THE MODEL

The network is modeled as an undirected graph $G = (V, L)$, where each node in V corresponds to one user. An edge $(i, j) \in L$ means that there is a communication link between the users corresponding to nodes i and j . The set of one-hop neighbors of user i is denoted by N_i , $N_i = \{j \in V | (i, j) \in L\}$.

Since the graph is undirected, the neighbor relationship is symmetric: $j \in N_i \Leftrightarrow i \in N_j$. The assumption for an undirected graph can be dropped, in order to model asymmetric links, but we believe the extension to be straightforward. We denote the set of Bad users by V_B , and the set of Good users by V_G . It holds that $V_B \cap V_G = \emptyset$ and $V_B \cup V_G = V$. The *type* $t_i \in \{G, B\}$ of a user i denotes whether he is Good or Bad. The Good users correspond to the selfish users mentioned in the Introduction.

The graph of the game is identified with the graph of the network. Therefore, the payoff of a user is only affected by his one-hop network neighbors, whose actions are also directly observable. As mentioned in the introduction, payoff dependencies on the actions of remote nodes can also be modeled. But that would require a mechanism for each user to learn which user (neighbor in the game graph) played which action. Since those users would be more than one hop away, it

seems that the existence of such a mechanism is a rather strong assumption, especially considering that there are malicious users in the network who would try to interfere.

Users have a choice between two actions: C (for Cooperate), and D (for Defect). When all users choose their actions, each user receives a payoff that depends on his own and his neighbors' actions, and his own and his neighbors' types.

Observe that each user is playing the same action against all neighbors. If the user can play different actions against different neighbors, then the game on each link is decoupled from all the other games. Each of these games can be solved independently using our framework, only the graph will be a two node, one edge graph. Which user model is more appropriate depends on the granularity of the user's control over his hardware. For instance, if all the user can do is switch his wireless connection ON or OFF, then he should be modeled as only being able to play the same action against all neighbors.

The payoff is decomposed as a sum of payoffs, one term for each adjacent link. Each term of the sum depends on the user's own action and type, and the action of his neighbor along that link. It may also depend on the type of his neighbor along that link, but in that case the separate terms of the sum are not revealed to him. Only the total payoff (sum of all link payoffs) is revealed to the user, otherwise he would be able to discover immediately the types of his neighbors.

The payoff of user i is denoted by $R_i(a_i|t_i)$, when i 's action is a_i and i 's type is t_i . We extend this notation to denote by $R_i(a_i a_j | t_i t_j)$ the payoff for i when j is a neighbor of i and j 's action and type are a_j and t_j . So, the decomposition of i 's payoff along each adjacent link can be written as

$$R_i(a_i|t_i) = \sum_{j \in N_i} R_i(a_i a_j | t_i t_j), \quad (1)$$

or, when i 's payoff does not depend on the types of his neighbors,

$$R_i(a_i|t_i) = \sum_{j \in N_i} R_i(a_i a_j | t_i). \quad (2)$$

We assume there are no links between any two Bad users. The Bad users are supposed to be able to communicate and coordinate perfectly; hence, there is no need to restrict their interaction by modeling it as a game theoretic interaction. Moreover, the Bad users are assumed to know exactly both the topology and the type of each user in the network. Good users only know their local topology, i.e., how many neighbors they have and what each one of them plays, but not their types. These assumptions enhance the power of the Bad users, as is standard practice in security research.

Since there are no links between Bad users, the two games that can appear are Good versus Good, and Good versus Bad. The payoff of a Bad user is the opposite of the payoff of the Good user. That is, the Good versus Bad game is a zero-sum game. Also, the Good versus Good game obviously needs to be symmetric. These two observations are encoded in the payoffs shown in Fig. 1.

The game is played repeatedly with an infinite horizon, and time proceeds in rounds $t = 1, 2, 3, \dots$. The action and payoff of user i at round t are denoted with superscript t : a_i^t and R_i^t . Each Good user i remembers his own payoff and action at

	<i>C</i>	<i>D</i>	
<i>C</i>	$r_1, -r_1$	$r_2, -r_2$	
<i>D</i>	$r_3, -r_3$	$r_4, -r_4$	

	<i>C</i>	<i>D</i>
<i>C</i>	s_1, s_1	s_2, s_3
<i>D</i>	s_3, s_2	s_4, s_4

Fig. 1. The two games that can take place on a link: Good versus Bad and Good versus Good. The payoffs reflect that the Good versus Bad game is zero-sum, and the Good versus Good game is symmetric. The row player is assumed to be Good in both cases.

each past round, and the actions of his neighbors at each past round. The n -round *history* available to user i is defined to be the collection of all this information from round 1 up to and including round n :

$$\mathcal{H}_i^{1\dots n} = \{R_i^t\} \cup \{a_j^t, j \in N_i \cup \{i\}\}, t = 1, \dots, n. \quad (3)$$

User i then uses the n -round history to decide what his action will be in round $n + 1$. User i 's strategy is a function that, for each n -round history and for all n , determines the probability with which user i will play C . The history is used to form estimates about actions expected of i 's neighbors at round $n + 1$. The estimates can either be directly about the neighbors' actions, or indirectly through first estimating their types and then the probability that, given their type, they will play a certain action.

The quantity $R_i^{n+1}(C)$ is the expected round $n + 1$ payoff for user i , if he plays C at round $n + 1$.

$$\begin{aligned} R_i^{n+1}(C) = & \sum_{j \in N_i} R_i(CC|GB) \cdot \Pr(a_j^{n+1} = C|t_j = B) \cdot \Pr(t_j = B) \\ & + \sum_{j \in N_i} R_i(CC|GG) \cdot \Pr(a_j^{n+1} = C|t_j = G) \cdot \Pr(t_j = G) \\ & + \sum_{j \in N_i} R_i(CD|GB) \cdot \Pr(a_j^{n+1} = D|t_j = B) \cdot \Pr(t_j = B) \\ & + \sum_{j \in N_i} R_i(CD|GG) \cdot \Pr(a_j^{n+1} = D|t_j = G) \cdot \Pr(t_j = G) \end{aligned} \quad (4)$$

We can similarly calculate $R_i^{n+1}(D)$. After all players choose their actions, they receive their payoffs and observe their neighbors' actions. They incorporate that information in their history, and update the probability estimates accordingly. The optimal strategy maximizes the game payoff R_i , which is a function of the per-round payoffs $R_i^n, n = 1 \dots \infty$. In this paper, we consider the limit-of-means payoff $R_i = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{n=1}^T R_i^n$, and the δ -discounted payoff $R_i = (1 - \delta) \sum_{n=1}^{\infty} \delta^{n-1} R_i^n$.

In Sections IV and V, we will give specifics of two models that instantiate the general model we just described. The difference will be, first and foremost, in the way that Good users treat collaboration with Bad users. In Section IV, collaboration with Bad users will be identical to collaboration with Good ones, that is, the costs and benefits will be the same in both cases. In Section V, the assumption will be that a Good user prefers to collaborate with Good users, but receives a negative payoff when collaborating with Bad users. So, he will be able to detect that Bad users are among his neighbors, but he will not immediately know which are the Bad ones.

There, the interest will be in how soon the Good user discovers which the Bad neighbors are. Additionally, upon discovering a Bad user, a Good user will be able to break the link that joins them, thus altering the game graph.

III. RELATED WORK

Repeated games with incomplete information were introduced in [1]. The 2-person case is also discussed in [2], and, at greater depth, in [3].

The most relevant piece of game theoretic literature is [4]. There, the authors deal with the case of tree games with incomplete information, i.e., the graph of the game is a tree, the players have private information, but there is no history (the game is not repeated). They provide algorithms for finding approximate Bayes-Nash equilibria. Their algorithm does not generalize to non-tree graphs in an obvious way.

In the literature for inducing cooperation among network users, mostly selfish users have been studied, where incentives are provided for users to cooperate [5], [6]. However, they are modeling Malicious users as "Never Cooperative", without any further sophistication, since their main focus was discouraging selfish free-riders. There is no degree of selfishness that can approximate the behavior of our Malicious users. For example, in [7] the authors assume that the payoff function of a user is non-decreasing in the throughput experienced by the user. Our Bad users do not care about their data being transmitted. For the same reason, the model proposed in [8] does not apply (as the authors themselves point out).

In other related work [9] a modified version of Generous Tit for Tat is used (for an early famous paper in the history of Tit for Tat see [10]), but they have no notion of topology and, consequently, of neighborhoods. In their setting, each user is comparing his own frequency of cooperation to the aggregate frequency of cooperation of the rest of the network. In [11], a scheme is proposed for punishing users whose frequency of cooperation is below the one dictated by a certain Nash equilibrium. Aimed particularly against free-riding in wireless networks is [12], and also [13].

To the best of our knowledge, there has been no game theoretic modeling of malicious users as we describe them here. Malicious users are modeled in [14] in a game theoretic setting, but in a different way. The game they are considering is a virus inoculation game, in which selfish users decide whether to pay the cost for installing anti-virus software (inoculation), or not pay and risk getting infected. The malicious users declare that they have been inoculated, when in fact they have not, so as to mislead the selfish ones. After the selfish users have made their decisions, the attacker chooses an uninoculated user, uniformly at random, and infects him. The infection propagates to all unprotected users that can be reached from the initially infected users on paths consisting of unprotected users (the malicious ones are equivalent to unprotected). One major difference is that in this model the selfish users are assumed to know the topology of the network (a grid, in particular), whereas in our model they only know their local neighborhood topology. Another difference is that the notion of Byzantine Nash Equilibrium that they consider is restricted to the strategies of the selfish users alone.

	C	D
C	$N - E, E - N$	$-E, E$
D	$0, 0$	$0, 0$

	C	D
C	$N - E, N - E$	$-E, 0$
D	$0, -E$	$0, 0$

Fig. 2. The two games that can take place on a link: Good versus Bad and Good versus Good.

IV. COLLABORATION IN THE FACE OF MALICE

There are two salient points introduced and used in this instantiation of the general model. First, Good users do not care if they cooperate with other Good users or with Bad users. They get the same costs and benefits in either case. Second, the Good users' indifference to the other players' type implies that Good users only care about estimating their neighbors' potential future actions, and not their type. A preliminary analysis of this model appeared in [15].

The main conclusion is that the higher the cost-to-benefit ratio for the Good users, the lower the achievable payoff for the Bad ones. Intuitively, a high cost-to-benefit ratio makes the Good users more demanding against their neighbors (Good or Bad, indifferently), so the Bad users cannot get away with very "bad" behavior.

A. Malicious and Legitimate User Model

For the purposes of this section, we define a game with the payoffs shown in table form in Fig. 2 for the two pairs of types that can arise (Good versus Good, and Good versus Bad). We explain the payoffs as follows: In the example of a wireless network, a C means that a user makes himself available for communication, that is, forwarding traffic of other nodes through himself. A link becomes active (i.e., data is exchanged over it) only when the users on both endpoints of the link cooperate, that is, play C . Playing C is in line with what Good users want to achieve – good network operation – but it costs energy, since it means receiving and forwarding data. So, when both players on a link play C , the Good player (or both players, if they are both Good) receives N (for Network) minus E (for Energy) for a total of $N - E$. We assume $N > E > 0$, otherwise no player would have an incentive to play C . On the other hand, when a Good player plays C and the other player D , then the Good player only wastes his energy since the other endpoint is not receiving or forwarding any data. For this reason, the payoff is only $-E$. In general, the values of E and N provide a way to quantify the selfishness of Good users, if we allow these values to vary across users. That is, values E^i and N^i , instead of E and N , would signify user i 's *personal* cost of spending energy, and his *personal* benefit from activating an adjacent link. The larger the ratio $\frac{E^i}{N^i}$, the more selfish the user.

The Bad user's payoff is always the opposite of the Good user's payoff. In particular, we do not assume any energy expenditure when the Bad users play C .

In a peer-to-peer network, a C would mean uploading high quality content (as well as, of course, downloading), and a D

would be the opposite (e.g. only downloading). The benefit of cooperation is the increased total availability of files. The cost of cooperation E could be, for instance, the hassle and possible expense associated with continuing to upload after one's download is over. In a general social network, edges would correspond to social interactions, a C would mean cooperating with one's neighbors toward a socially desirable objective (like cleaning the snow from the sidewalk in front of one's house), and a D would mean the opposite of a C . The cost E and benefit N are also obvious here.

As we have discussed, we allow the Bad users to have all information about the past (their own moves, as well as everybody else's moves since the first round). On the other hand, the Good users follow a *fictional play* process [16], that is, they assume that each of their neighbors chooses his actions independently at each round and identically distributed according to a probability distribution with unknown parameters (Bernoulli in this case, since there are only two actions available: C and D). So, at each round they are choosing the action that maximizes their payoff given the estimates they have for each of their neighbors' actions. For example, if player i has observed that player j has played c C 's and d D 's in the first $c + d$ rounds, then i assumes that in round $t = c + d + 1$, j will play C with probability $\frac{c}{c+d}$ and D with probability $\frac{d}{c+d}$. We denote by q_j^t the estimated probability that j will play C in round $t + 1$, which is based on j 's actions in rounds $1, \dots, t$.

We choose the limit-of-means payoff as our repeated game payoff function:

$$R_i = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T R_i^t. \quad (5)$$

Let us now calculate the expected payoff for each of the two actions of a Good user. We assume that t rounds have been completed, and Good user i is contemplating his move in round $t + 1$. The following equation is an instantiation of the generic payoff equation (4) for the specific payoffs shown in Fig. 2.

$$\begin{aligned} R_i^{t+1}(C|G) &= \sum_{j \in N_i} \{q_j^t R_i(CC|G) + (1 - q_j^t) R_i(CD|G)\} \\ &= N \cdot \sum_{j \in N_i} q_j^t - |N_i| \cdot E \\ R_i^{t+1}(D|G) &= 0. \end{aligned} \quad (6)$$

So, in order to decide what to play, user i has to compare the expected payoffs that the available actions will bring. Action C will be chosen if and only if $R_i(C|G) \geq R_i(D|G)$, that is, if and only if

$$\sum_{j \in N_i} q_j^t \geq |N_i| \frac{E}{N}. \quad (7)$$

We can see why the ratio $\frac{E}{N}$ can encode the selfishness of a user. It is the empirical frequency of cooperation per neighbor that the user expects to see before cooperating himself. The simplicity of (7) makes the implementation of the Good user behavior particularly easy, since they only need to keep track of one number for each neighbor, as opposed to the whole

history of actions observed. It remains to see what the Bad users can do against this strategy. On the one hand, perhaps they can exploit its simplicity to gain the upper hand against the Good users. On the other hand, the fact that the Bad users have full knowledge of the history and can achieve perfect coordination in their actions may not be very useful here. The reason is that Good users only care about the frequencies of the actions that they observe, and the Bad users may achieve nothing by more elaborate strategies.

B. Searching for a Nash Equilibrium

In game theory, the main solution concept is the Nash Equilibrium. In our case, we have already restricted the Good players' strategies to fictitious play, so the Nash Equilibrium will be restricted to be obtainable through fictitious play. More formally, the Nash Equilibrium in our case would be a vector $\vec{s} \in [0, 1]^{|V|}$, where s_i is the strategy of user i , i.e., the function of the n -round history that determines the probability with which user i plays C at round $n + 1$. What a Good user plays will be determined by the choice of strategies by all the Bad users. Given a choice of strategies \vec{s}_B by the Bad users, the responses of the Good users are determined by the fictitious play process. It is then easy to determine the payoffs for every user (Good or Bad).

In general, a change in the strategy of a Bad player i will affect the payoffs of other Bad players, too, since it will change the optimal responses of Good users that could, e.g., be common neighbors of i and other Bad users. So, the definition of the Nash Equilibrium we gave earlier could be expanded to regard the Bad players as a team that aims to maximize the sum of their payoffs, rather than each Bad user trying to maximize his own individual payoff. In Section IV-C, we will see what happens in the "Uncoupled Configurations": the set of configurations (i.e., values for E and N , graphs, and associated Good/Bad node placements) where local payoff maximization by each Bad user is equivalent to maximization of the sum of all the Bad users' payoffs. In Section IV-D we will examine in an example exactly how a general configuration differs, and hint at a heuristic solution.

C. The Uncoupled Configurations

Let us start with the case of a single Bad player in the whole network. Since no other Bad players exist, the choice of the Bad user will only affect his own payoff. We will see what he has to do in order to maximize his payoff in a tree topology where the Bad player is at the root of the tree.

Assume that the Bad user – labeled user 0 – has k neighbors, labeled $1, \dots, k$. We also assume that all the Good users will start by playing C , and will only change to D if they are forced by the Bad user. Applying (7) for each neighbor, we see that each expects to see a different sum of empirical frequencies from his own neighbors in order to keep playing C . User i expects to see a sum of empirical frequencies that is at least $\frac{E}{N}|N_i|$. Since all of i 's neighbors except user 0 are Good, they will at least start by playing C , so user i will see a sum of frequencies equal to $|N_i| - 1$ from them. So, in order to make user i play C at round $t + 1$, the Bad user should, in the first t

rounds, play C with a large enough empirical frequency. We call the minimum such empirical frequency the *threshold* t_i :

Definition 1. *The threshold t_i of a Good user i is the empirical frequency of C s that his Bad neighbor needs to play, so that i keeps playing C . We assume that all the Good neighbors of i play C .*

$$t_i = \max \left\{ \frac{E}{N}|N_i| - (|N_i| - 1), 0 \right\} \quad (8)$$

Observe that t_i is decreasing as $|N_i|$ increases, since $E < N$. Also, t_i increases as E increases towards N , that is, as the selfishness of user i grows.

Without loss of generality, we assume that the Bad user's neighbors are labeled in increasing order of t_i . So, $t_1 < t_2 < \dots < t_k$. We will now describe a strategy that provably achieves optimal limit-of-means payoff for the Bad user. The proof proceeds in two steps. First, we assume the Bad user only has one Good neighbor, whose threshold is θ . We describe and prove the optimality of a strategy as a function of θ in this case, and then show how it changes when the Bad user has more than one Good neighbor.

1) *One Bad user with one Good neighbor:* The Bad user can choose any strategy s he wants, where the space of all strategies is the set of infinite-length vectors of C s and D s: $S = \{C, D\}^\infty$. In choosing a strategy, his objective is to maximize his limit-of-means payoff, knowing that the Good user will play C in the first round, and then play according to fictitious play (FP). Strategies can be constructed that result in non-convergent payoffs, but we do not concern ourselves with them. Let θ be the threshold to be exceeded by the Bad user's empirical frequency of C s $q_0^t = \sum_{s=1}^t \mathbf{1}_{\{a_s^0=C\}}$ in order for the Good user to play C . Since θ and q_0^t are known to the Bad user, he can predict after the end of round t what the Good user will play at round $t + 1$.

The limit-of-means payoff for the Bad user is:

$$\begin{aligned} R_0 &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T R_0^t \quad (9) \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \left(E - N \mathbf{1}_{\{a_0^t=C\}} \right) \mathbf{1}_{\{q_0^{t-1} \geq \theta \}}. \end{aligned}$$

We now prove that no strategy can achieve a payoff higher than $(1 - \theta)E$. Then we will describe a strategy that achieves this payoff, and is, as a consequence, optimal. So we can conclude that the more selfish the Good user is, the lower the maximum payoff for the Bad user. Selfishness makes the Good users more demanding, and more likely to stop participating in the network, which would make the payoff of the Bad user 0.

Theorem 1. *For any strategy, $R_0 \leq (1 - \theta)E$.*

Proof: We will assume $R_0 > (1 - \theta)E$, and reach a contradiction:

$$\begin{aligned} R_0 &> (1 - \theta)E \Rightarrow \\ \Rightarrow \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \left(1 - \frac{N}{E} \mathbf{1}_{\{a_0^t=C\}} \right) \mathbf{1}_{\{q_0^{t-1} \geq \theta \}} &> 1 - \theta \end{aligned}$$

$$\begin{aligned}
&\Rightarrow \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbf{1}_{\{a_0^t = C\}} < \theta & (10) \\
&\Rightarrow \lim_{T \rightarrow \infty} q_0^T < \theta \\
&\Rightarrow \exists T_0 \text{ such that } \forall T > T_0 : q_0^T < \theta \\
&\Rightarrow \exists T_0 \text{ such that } \forall T > T_0 : \text{Good plays } D \text{ at } T + 1 \\
&\Rightarrow \exists T_0 \text{ such that } \forall T > T_0 : R_0^{T+1} = 0 \\
&\Rightarrow R_0 = 0,
\end{aligned}$$

where we have used the facts that $N \geq E$, and $\mathbf{1}_{\{\cdot\}} \leq 1$. So we have reached a contradiction, which proves the theorem. ■

Theorem 2. *The strategy $s^*(\theta)$:*

$$a_0^t = \begin{cases} D & t = 1, \\ C & t > 1, \text{ and } q_0^{t-1} < \theta, \\ D & t > 1, \text{ and } q_0^{t-1} \geq \theta \end{cases}$$

achieves payoff $R_0 = (1 - \theta)E$.

Proof: The first step is proving that the empirical frequency of C s of this strategy converges to θ , so the empirical frequency of D s converges to $1 - \theta$.

We will omit the subscript 0 from q_0^t for convenience. We partition the set of all time instants (rounds) into four subsets:

- 1) $T_1 = \{t | q^t \geq \theta, a^{t+1} = D, q^{t+1} \geq \theta\}$
- 2) $T_2 = \{t | q^t \geq \theta, a^{t+1} = D, q^{t+1} < \theta\}$
- 3) $T_3 = \{t | q^t < \theta, a^{t+1} = C, q^{t+1} \geq \theta\}$
- 4) $T_4 = \{t | q^t < \theta, a^{t+1} = C, q^{t+1} < \theta\}$

Lemma 1. *The set $T_2 \cup T_3$ has an infinite number of elements. Also, $\sup T_2 \cup T_3 = \infty$.*

Proof: Assume $|T_2 \cup T_3| < \infty$. Then, $\exists K < \infty$ such that $(\forall t > K : q_t < \theta)$ or $(\forall t > K : q_t \geq \theta)$. But, according to the strategy, $q_t \geq \theta \Rightarrow a^{t+1} = D \Rightarrow \exists K'$ such that $q_{K'} < \theta$, and similarly for $q_t < \theta$. So we have reached a contradiction. The second part of the lemma follows immediately, since the set $T_2 \cup T_3$ is a subset of the positive integers, and as such has no finite accumulating points. ■

Lemma 2. *Along the infinite subsequence of time instants in $T_2 \cup T_3$, the empirical frequency of C s under the strategy $s^*(\theta)$ converges to θ .*

Proof:

$$\begin{aligned}
t \in T_2 &\Rightarrow q^{t+1} < \theta \leq q^t \\
&\Rightarrow \frac{tq^t}{t+1} = q^{t+1} < \theta \leq q^t < \frac{tq^t + 1}{t+1} & (11)
\end{aligned}$$

$$\begin{aligned}
t \in T_3 &\Rightarrow q^t < \theta \leq q^{t+1} \\
&\Rightarrow \frac{tq^t}{t+1} < q^t < \theta \leq q^{t+1} = \frac{tq^t + 1}{t+1} & (12)
\end{aligned}$$

So, $t \in T_2 \cup T_3 \Rightarrow \frac{tq^t}{t+1} \leq \theta \leq \frac{tq^t + 1}{t+1}$. Subtracting q^t from the three parts of the inequality, and recalling that $q^t \leq 1$, we get that $-\frac{1}{t+1} \leq \theta - q^t \leq \frac{1}{t+1}$. Letting $t \in T_2 \cup T_3$ go to infinity (which we can do because of Lemma 1), we reach the desired conclusion. ■

Lemma 3. *The empirical frequency of C s under the strategy $s^*(\theta)$ converges to θ : $\lim_{t \rightarrow \infty} q_t = \theta$.*

Proof: From Lemma 2, $\lim_{t \rightarrow \infty, t \in T_2 \cup T_3} q^t = \theta$. By the definition of q_t , and of the sets T_1 and T_4 , the distance $|q^t - \theta|$ is decreased by actions a_0^t taken whenever $t \in T_1 \cup T_4$. Given $\epsilon > 0$, we pick $K(\epsilon) = \min\{t | t \in T_2 \cup T_3 \wedge \frac{1}{t+1} < \epsilon\}$. Then, $\forall t > K(\epsilon) : |q^t - \theta| < \epsilon$, so $\lim_{t \rightarrow \infty} q_t = \theta$. ■

Having proven that $q^t \rightarrow \theta$, it follows that the empirical frequency of the D s that the Bad player plays converges to $1 - \theta$. But, by construction of the strategy $s^*(\theta)$, each D brings a round-payoff of E and each C brings a round-payoff of 0, so from (9) the game-payoff of $s^*(\theta)$ is $(1 - \theta)E$. ■

Corollary 1. *The strategy $s^*(\theta)$ is optimal for the Bad user.*

Proof: Follows from Theorems 1 and 2. ■

The description of strategy $s^*(\theta)$ is equivalent to ‘‘Play C when the Good user will play D , and play D when the Good user will play C ’’. Then, it follows directly that the frequency of the C s played by the Good user converges to $1 - \theta$, and the payoff that the Good user receives from the Bad user is $(1 - \theta) \cdot (-E)$.

2) *One Bad user with many Good neighbors:* Each Good neighbor $i, 1 \leq i \leq k$ compares q_0^t with a different threshold $t_i, 1 \leq i \leq k$ before deciding whether to cooperate at round $t + 1$. The Bad user will choose the strategy $s^*(t_i^*)$ that gives him the highest limit-of-means payoff. In particular, out of the k Good neighbors, the $i^* - 1$ with lower thresholds than t_i^* will play always C , the $k - i^*$ with higher thresholds will play always D , and the i^* th neighbor will behave as the single Good neighbor in the preceding section. So, the payoff for the Bad user will be:

$$\begin{aligned}
R_0(s^*(t_i^*)|B) &= (i^* - 1) \cdot \{t_i^*(E - N) + (1 - t_i^*)E\} \\
&\quad + (k - i^*) \cdot \{t_i^* \cdot 0 + (1 - t_i^*) \cdot 0\} \\
&\quad + (1 - t_i^*)E & (13) \\
&= (i^* - 1)(E - Nt_i^*) + (1 - t_i^*)E
\end{aligned}$$

Again, we see that the more selfish the Good users (higher thresholds t_i), the lower the payoff for the Bad user.

Since the thresholds $t_i, 1 \leq i \leq k$ are known to the Bad user, he can pick the value for i^* that maximizes his payoff. He then follows $s^*(t_i^*)$ augmented with an initial stage: Play n_C initial C s and then n_D D s, so that $t_{i^*-1} < \frac{n_C}{n_C + n_D} < t_{i^*}$. In this way, the Bad user makes sure that all Good users with thresholds below t_{i^*} play C forever. This initial stage is of finite duration so it does not affect the game payoff computed above. After the initial stage, the Bad user reverts to the strategy $s^*(t_{i^*}^*)$. The optimality proof of the strategy then follows the one in the preceding section. The intuition behind the proof is that any strategy with a higher payoff would have to have an empirical frequency of C s lower than the chosen optimal threshold, thus making more Good users start playing D , thus reducing the (claimed) higher payoff.

When the Bad user plays the augmented $s^*(t_{i^*}^*)$ strategy, as we have said $i^* - 1$ of his neighbors will play C , $k - i^*$ will play D , and one will play both C and D . The ones who play at least some D may cause their own neighbors to start playing D , and so on. However, the D s cannot, by propagating, influence the other Good neighbors of the Bad user: That is a consequence of the tree topology that we have assumed, since the only path that joins two one-hop neighbors of the Bad user, goes

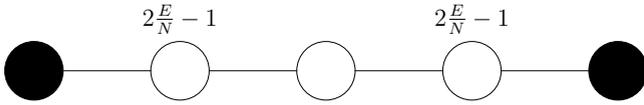


Fig. 3. Example to contrast the uncoupled with the general configuration. Shown next to each node is its threshold. The black nodes are the Bad users. Depending on the value of $\frac{E}{N}$, the Bad users may or may not need to coordinate to maximize the sum of their payoffs. In particular, if $\frac{E}{N} > \frac{2}{3}$, they need to coordinate.

through the Bad user. So, tree topologies with a single Bad node are uncoupled configurations. What happens when there are multiple Bad users in a general topology? Can we have an uncoupled configuration then?

If so, the Nash equilibrium consists of the Bad users playing their augmented $s^*(t_i^*)$ strategy for a suitable t_i^* each, and the Good users following fictitious play as they have been constrained to do. It is then straightforward to compute the payoffs of the Good and Bad users. In what follows, we will see these considerations coming into play in the more general case.

D. Example and General Configuration

We now examine a specific example, shown in Fig. 3, to see the difference between the uncoupled and the general configurations. There are two Bad users, labeled B_1 and B_2 , and three Good users. Shown in the picture are the thresholds of the Good users. If $\frac{E}{N} \leq \frac{1}{2}$, the thresholds are non-positive, so both Bad users can play D forever without causing the Good users ever to play D . That would be the payoff maximizing strategy for the Bad users, given that the Good users play according to fictitious play. The payoff for each Bad user would be E , the payoff for the Good user in the middle would be $2(N - E)$, and the payoff for each of the other two Good users would be $N - 2E$. Since the Bad players optimize separately, but the resulting strategy also optimizes the sum of their payoffs, this configuration is an uncoupled one.

For any values of $\frac{E}{N} > \frac{1}{2}$, the strategy maximizing each Bad user’s payoff separately is $s^*(2\frac{E}{N} - 1)$. That strategy causes their two Good neighbors to play C with frequency $1 - (2\frac{E}{N} - 1) = 2 - 2\frac{E}{N}$. So the Good user in the middle sees a total frequency of $4 - 4\frac{E}{N}$, which needs to be at least $2\frac{E}{N}$ for him to cooperate. So, as long as $\frac{E}{N} \leq \frac{2}{3}$, he will play C forever, and the thresholds of the other two Good users will hold. The payoffs will be $2(1 - \frac{E}{N})E$ for each Bad user, and accordingly for the Good users. So we are still in an uncoupled configuration. However, if $\frac{E}{N} > \frac{2}{3}$, then the Good user in the middle will not cooperate, so the other two Good users will not see the total empirical frequency that they require. As a consequence, they will all play D , and the Bad users’ payoff will be reduced to zero. So, the Bad users can no longer optimize separately.

A heuristic solution would then be for one of the Bad users to make a “sacrifice” and choose a strategy that gives him a payoff that is lower than the one he would get if he were alone in the network. The rationale would be that the other Bad user would gain in payoff enough to cover the difference lost, and even more.

V. DETECTING MALICIOUS USERS

In this part, the general scenario that we want to capture is the following: Good users want to cooperate with other Good users, but not with Bad users. Bad users, on the other hand, want to cooperate with Good users. The Good are unaware of who is Good and who is Bad, but since the game is played repeatedly they can gradually detect the Bad ones. We will explore strategies that the Good users can follow to detect and isolate the Bad ones.

This scenario applies to intruder detection in the following sense: Assume there is a pre-existing mechanism that detects the existence of malicious users in a user’s neighborhood, but cannot identify exactly which the malicious users are. They may be eavesdroppers who monitor traffic (through cooperating in the network operation, i.e., playing C) and try to learn and then leak sensitive information. The mechanism is then supposed to be able to detect the leak. As a result, the existence of malicious users would be deduced, but not their identity. Alternatively, they may inject malformed packets in the network (worms, etc.), which cannot be traced back to them.

Note that we will not be assuming collusion among the Bad users, although this can be an extension of our model. Also, our model for the Bad users means that they benefit from cooperating for as long as possible without getting caught. So, we do not cover situations where a single cooperation between a Bad and a Good user is enough, e.g., to destroy the whole network.

Our main conclusion is that the optimal policy for the Bad users is to always cooperate, aiming to maximize their short-term payoff, even though this may result in a quick detection. The intuition is that, even when the Bad user is hiding, some of the Good users are eliminated as potential candidates for being Bad. As a consequence, the Bad users still come closer to being detected, but do not gain any payoff when hiding.

Again, playing C corresponds to making oneself available for communication (e.g. sending/receiving data). Playing D corresponds to shutting down all communications to and from the user. After all users have chosen an action, each user learns his neighbors’ actions (i.e. which neighbor played which action), and his own payoff for that round, which depends only on his own action and these of his neighbors. Note that a user’s payoff is known only to him, and is never reported to others. If a Good user is able to tell that a particular neighbor of his is Bad, then he can sever the link that joins them, so as not to be affected by that neighbor’s actions in the future. In Section V-A, we will discuss how Good users can detect Bad ones.

At round t , the payoff R_i^t of a Good user i who played C equals the number of Good neighbors who played C minus the number of Bad neighbors who played C . This reflects the preference of Good nodes to cooperate with other Good nodes and not with Bad ones. A Good user who played D receives a zero payoff regardless of the actions of the neighbors. This means that he risks no losses, but he has no gain, he learns nothing about his neighbors, and his neighbors learn nothing about him. The payoff of a Bad user who plays C is equal to the number of his Good neighbors who played C (remember

	<i>C</i>	<i>D</i>	
<i>C</i>	-1, +1	0, 0	
<i>D</i>	0, 0	0, 0	

	<i>C</i>	<i>D</i>
<i>C</i>	+1, +1	0, 0
<i>D</i>	0, 0	0, 0

Fig. 4. The two games that can take place on a link: Good versus Bad and Good versus Good.

that a Bad user has only Good neighbors). So, it is the opposite of R_i^t . We can translate the above considerations to payoffs, shown in Fig. 4 for the two possible games (Good versus Good and Good versus Bad).

A. Analysis

To simplify the analysis, we will concentrate on a star topology network, where the central node is a Good user, and his neighbors are N Good users and 1 Bad. We assume that the central node knows that he has exactly one Bad neighbor, but he does not know who that is. We will see that the assumption that only one Bad user exists can be removed without significant conceptual change in the analysis.

Note from the previous discussion that Good users only learn the total payoff they receive after each round, and not the per-link payoffs they receive due to their interactions with individual neighbors. So, they cannot immediately tell which of their neighbors are Good and which are Bad, but they do get some information about the types of their neighbors. In what follows, we will describe strategies for the Good and Bad players that form a Nash equilibrium. For the most part, we will be seeing things either from the central Good user's point of view, or from the point of view of the Bad user. Since in a general network all Good nodes will see themselves in the role of a central node in a (local) star topology, we are looking for strategies that are symmetrical with respect to Good nodes. That is, we want all the Good nodes to follow the same rules when choosing what to play.

Assume that the central Good user i has memory of the past history (own and neighbor moves, as well as received payoffs). Let CN_i^t (resp. DN_i^t) be the subset of i 's neighbors that play C (resp. D) at round t . We assume that i plays C at round t , so i 's payoff at round t is $|CN_i^t|$ if the Bad user played D , or $|CN_i^t| - 2$ if the Bad user played C (Remember that a C from a Good user gives +1, whereas from a Bad user it gives -1). So, just by looking at his payoff, the central Good user i can deduce whether the Bad user played C or D at round t . The Bad user is then known to be either in the set CN_i^t or in DN_i^t . Without loss of generality, let's assume that the Bad user played C .

In the next round ($t + 1$), if the Bad user plays C again, then i can deduce that he is in the intersection $CN_i^t \cap CN_i^{t+1}$. If he plays D , then he is in $CN_i^t \cap DN_i^{t+1}$. This sequence of sets (the sequence of *hiding sets*, the initial of which is N_i) is non-increasing, but the Bad user will only be detected if the sequence converges to a singleton set. If the behavior of the Good users is deterministic, then the Bad user can imitate a Good user, and he will never be discovered. However, if the Good nodes choose their actions in a randomized manner, they are no longer predictable.

We will look at the simplest possible randomization: each Good user plays C with probability p independently at each round. We will use the δ -discounted payoff: Given an infinite sequence of round payoffs $\{R_i^t, t = 1, 2, \dots\}$, the game payoff for user i is $R_i = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} R_i^t$. In our case, δ could correspond to how long the users think that the network will keep operating. To be precise, δ could be seen as the probability that the network will stop operating at time $t + 1$, given that it has been operating up to and including time t . We do not claim that playing such a strategy is optimal among all strategies, but we will prove that, for δ sufficiently close to 1, such a simple randomization strategy (for any value of p) is indeed optimal among all strategies. For a given value of δ , we will find the probability p that maximizes the central Good user's payoff among all iid randomization strategies.

To compute the payoff, we split the network evolution into two stages: pre-detection and post-detection of the Bad user. In the pre-detection stage, the Good users start by playing C with probability p , and we assume that the Bad user always plays C . We will see that this is the best that the Bad user can do. We again use the generic payoff equation (4) with the specific payoffs shown in Fig. 4. Moreover, the Bad users always play C , so $\Pr(a_j^{t+1} = C | t_j = B) = 1$. We know that there is exactly one Bad neighbor, so for each j the probability that j is Bad is $\Pr(t_j = B) = \frac{1}{|HS^t|}$ if j is in the round- t hiding set HS^t and zero otherwise. The expected round t payoff for the central Good user i is

$$\begin{aligned}
R_i^t(C) &= \sum_{j \in N_i} R_i(CC|GB) \Pr(a_j^t = C | t_j = B) \Pr(t_j = B) \\
&\quad + \sum_{j \in N_i} R_i(CC|GG) \Pr(a_j^t = C | t_j = G) \Pr(t_j = G) \\
&= \sum_{j \in HS^t} (-1) \cdot 1 \cdot \frac{1}{|HS^t|} + (+1) \cdot p \cdot \frac{|HS^t| - 1}{|HS^t|} \\
&\quad + \sum_{j \in N_i \setminus HS^t} (+1) \cdot p \cdot 1 \\
&= p(N_i - 1) - 1, \tag{14}
\end{aligned}$$

when he plays C , and

$$R_i^t(D) = 0, \tag{15}$$

when he plays D . Recalling that $N_i - 1 = N$ is the number of Good neighbors, the overall expected payoff for user i at round t is

$$\begin{aligned}
R_i^t &= \Pr(a_i^t = C) R_i^t(C) + \Pr(a_i^t = D) R_i^t(D) \\
&= p(pN - 1) \tag{16}
\end{aligned}$$

After the Bad user has been detected, the link to him is severed and the Good nodes are free to play C forever. So, the central Good user's payoff is N from then on (+1 from each one of the N Good neighbors). We now compute the total game payoff for the central Good user. Define the random variable M to be the number of observations needed to be made by the Good user until detection. Define the random variable R to be the number of rounds until detection. Obviously, $M \leq R$, since observation rounds are only the rounds when the central Good user plays C . To compute the distribution of M observe that each of the Good neighbors

remains in the hiding set for as long as he plays C at each observation round. Each C is played independently at each round with probability p , so each Good neighbor remains in the hiding set for a geometrically distributed number of observation rounds.

$$\begin{aligned}
 \Pr(M \leq m) &= \\
 &= \Pr(\text{After } m \text{ obs. rounds, } |HS| = 1) \\
 &= \Pr(\text{After } m \text{ obs. rounds, all Good nbr have played } \geq 1D) \\
 &= (1 - p^m)^N \\
 \Rightarrow \Pr(M = m) &= (1 - p^m)^N - (1 - p^{m-1})^N \quad (17)
 \end{aligned}$$

To compute the distribution of R – the number of rounds until detection – observe that, given the number of observation rounds M , the conditional distribution $\Pr(R = r | M = m)$ is a negative binomial. It is the number of rounds needed to have m “successes”, each “success” happening with probability p .

$$\begin{aligned}
 \Pr(R = r) &= \\
 &= \sum_{m=1}^r \Pr(R = r, M = m) \\
 &= \sum_{m=1}^r \Pr(R = r | M = m) \Pr(M = m) \quad (18) \\
 &= \sum_{m=1}^r \binom{r-1}{m-1} p^m (1-p)^{r-m} ((1-p^m)^N - (1-p^{m-1})^N)
 \end{aligned}$$

For the total game payoff, we have:

$$\begin{aligned}
 R_i(\delta, p, N) &= \\
 &= (1 - \delta) \mathbf{E} \left\{ \sum_{t=1}^R \delta^{t-1} p(pN - 1) + \sum_{t=R+1}^{\infty} \delta^{t-1} N \right\} \\
 &= \frac{(1 - \delta)}{\delta} \left(p(pN - 1) \mathbf{E} \left[\sum_{t=1}^R \delta^t \right] + N \mathbf{E} \left[\sum_{t=R+1}^{\infty} \delta^t \right] \right) \quad (19)
 \end{aligned}$$

We define

$$S(p, \delta) = \sum_{r=1}^{\infty} \delta^r \Pr(R = r) \quad (20)$$

and after some calculations we can see that the payoff can be rewritten as:

$$R_i(\delta, p, N) = p(pN - 1)(1 - S(p, \delta)) + NS(p, \delta). \quad (21)$$

Corollary 2. *Given any value for p , the strategy “Play C independently with probability p ” is optimal for the Good users among all possible strategies for δ sufficiently close to 1, i.e., $\delta \in [\delta^*(p), 1]$.*

Proof: For $\delta \rightarrow 1$, we can see that $S(p, \delta) \rightarrow 1$, so the payoff $R_i(\delta, p, N) \rightarrow N$, which is the maximum achievable by any strategy, since it corresponds to receiving the maximum per-round payoff N at every round. Another way of proving the same result is observing that, for any p , detection is achieved in a finite number of rounds with probability 1 ($\Pr(R < \infty) = 1$), so for sufficiently patient players ($\delta \rightarrow 1$) the pre-detection payoff becomes irrelevant. ■

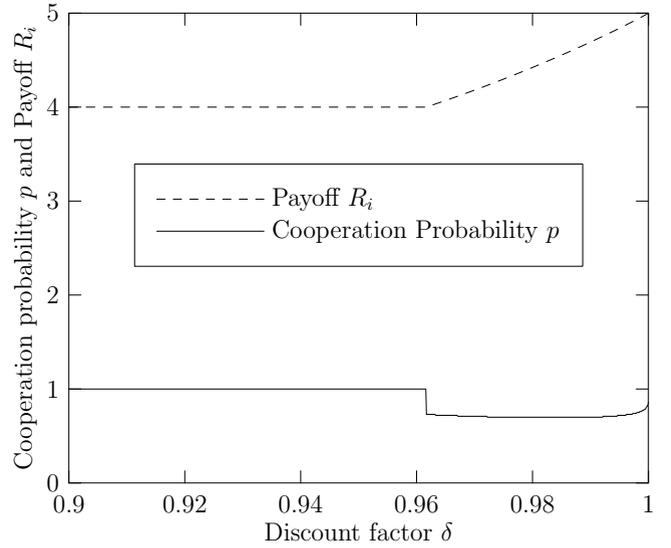


Fig. 5. Payoff R_i and optimal probability of cooperation p versus discount factor δ . When δ is close enough to 1, it becomes worthwhile for the Good users to aim for detection. Otherwise, the best thing they can do is always cooperate, and not try to detect the Bad user.

For a given value of δ , the optimal value of p maximizing (21) can be computed numerically. We show the resulting optimal values of p and corresponding payoffs in Fig. 5 when the number of Good neighbors (N) is 5. For values of δ less than 0.9 the optimal p is 1, and the achieved payoff is equal to 4. The optimal p being 1 means that the Bad user is never discovered, as a consequence of the impatience (small δ) of the Good users. After δ crosses a threshold, it becomes worthwhile for the Good users to try and detect the Bad one, resulting in a finite detection period and larger payoff. For various given values of δ , it may be interesting to observe how the payoff changes as a function of p . Due to lack of space, we point the interested reader to [17], where a preliminary analysis of the current model appeared.

Why does the Bad user have to play C all the time? If he plays what most Good play, he prolongs the time of detection. If he plays C , he gains payoff. If p was chosen by the Good to be larger than $\frac{1}{2}$, then these two considerations of the Bad user would both concur to playing C . However, the maximizing p may be less than $\frac{1}{2}$ for some values of δ , so it might make sense for the Bad user to play D once in a while so as to hide a bit longer.

We will see that playing D never increases the Bad user’s payoff, and it can even decrease it. Suppose the current hiding set is X , and the Bad plays D . If the central Good plays D , nothing changes. If the central Good plays C , he observes who plays D and who plays C , so by looking at the payoff he can tell what the Bad played. The new hiding set is $X \cap Y \subseteq X$, i.e. smaller than X , the Bad has gained nothing in the current period, and because of the discount factor δ the payoff of a C has become smaller. So, in effect, the Bad player is facing the exact same situation he was facing before, only he is in a smaller hiding set, and the benefit of a C is smaller. This allows us to conclude that the best thing the Bad can do is play C from the first round until he is detected. Hence, the strategies “always C ” for the Bad and “ C with probability

p " (for the maximizing value of p) for the Good form a Nash equilibrium, as long as the Good user strategies are constrained to be of the type "Play C with some probability p ". No one can do better by unilateral deviation within the allowed strategies.

Observe that the "Always C " Bad user strategy is optimal regardless of the value of p in the "Play C with probability p " Good user strategy. Therefore, as long as the maximizing p is unique, the Nash equilibrium is unique. In any case, all maximizing values of p would give the same payoff to the Good user, so the payoff will definitely be unique across all Nash equilibria.

B. Discussion of the General Case

In the general case of the star topology, the central Good user will have more than one Bad neighbors, and will not know exactly how many he has. Then, the concept of the hiding set has to be amended. At each observation, the central Good user will know how many Bad neighbors played C at that round. As a heuristic, the Good users can play C in the first round, which would disclose immediately how many Bad neighbors each Good user has. Later, by suitable randomization, the Bad users will definitely be detected: At the very least, in some round the central Good user will play C and all other Good users will play D (this is a positive probability event), so the Bad users who play D at that round will be detected.

Also, in a general topology each Good user will have a different number of Good and Bad neighbors, so the optimal cooperation probabilities will be different for each Good user. Even then, however, we claim that, as long as there is randomization in the actions of the Good users, they will eventually be able to detect all the Bad ones.

VI. CONCLUSION

We have presented a game theoretic model for the interaction of legitimate and malicious users in collaborative networks. To show the expressive power of the model we described two particular instantiations: The first is inspired from wireless packet forwarding protocols, and captures the double objective of the attackers to reduce the connectivity of the network, while at the same time depleting the legitimate users' energy. The second resembles an intruder detection framework, where the intruder causes damage while trying to stay hidden for as long as possible.

In both instantiations we have focused on equilibria that result when the legitimate users are following simple procedures with no severe requirements in terms of memory and computing power. We can conclude that useful conclusions can be derived from the model, without delving too deep into intricate analysis. A future objective is to see if more detailed analysis can model more realistic scenarios, or find equilibria that improve the payoffs of the legitimate users. In any case, however, the current model appears to be quite tractable.

From a security perspective, finding Nash equilibria has allowed us to find strategies for legitimate users that enforce upper bounds on the damage that the malicious users can do. We have been assuming that complete knowledge of the network is available to the attackers; if that is not the case, then their payoff will be lower than the equilibrium payoff.

Identifying and quantifying the tradeoff between the malicious users' knowledge and their equilibrium payoff is, we believe, another worthwhile future goal.

REFERENCES

- [1] R. J. Aumann and M. Maschler, "Game theoretic aspects of gradual disarmament," in *Report to the U.S. Arms Control and Disarmament Agency ST-80*. Princeton: Mathematica, Inc., 1966.
- [2] R. B. Myerson, *Game Theory: Analysis of Conflict*. Harvard University Press, 1991.
- [3] R. J. Aumann and S. Hart, Eds., *Handbook of Game Theory*. Elsevier Science Publishers B.V., 1992.
- [4] S. Singh, V. Soni, and M. Wellman, "Computing approximate bayesian equilibria in tree-games of incomplete information," in *Proceedings of the 5th ACM conference on Electronic commerce*. ACM Press, 2004, pp. 81–90.
- [5] A. Blanc, Y.-K. Liu, and A. Vahdat, "Designing incentives for peer-to-peer routing," in *Proceedings IEEE Infocom 2005*, Miami, FL, March 2005.
- [6] L. Buttyan and J.-P. Hubaux, "Stimulating Cooperation in Self-Organizing Mobile Ad Hoc Networks," *ACM/Kluwer Mobile Networks and Applications*, vol. 8, no. 5, 2003.
- [7] M. Félegyházi, J.-P. Hubaux, and L. Buttyán, "Nash Equilibria of Packet Forwarding Strategies in Wireless Ad Hoc Networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 5, pp. 463–476, May 2006.
- [8] A. Urpi, M. Bonuccelli, and S. Giordano, "Modelling Cooperation in Mobile Ad Hoc Networks: A Formal Description of Selfishness," in *Proceedings WiOpt 2003: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, INRIA Sophia-Antipolis, France, Mar. 2003.
- [9] V. Srinivasan, P. Nuggehalli, C.-F. Chiasserini, and R. Rao, "Cooperation in Wireless Ad Hoc Networks," in *Proceedings IEEE Infocom 2003*, San Francisco, March 30–April 3 2003.
- [10] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, no. 4489, pp. 1390–1396, March 2002.
- [11] E. Altman, A. Kherani, P. Michiardi, and R. Molva, "Non-cooperative Forwarding in Ad Hoc Networks," INRIA, Tech. Rep. RR-5116, 2004.
- [12] R. Mahajan, M. Rodrig, D. Wetherall, and J. Zahorjan, "Sustaining cooperation in multi-hop wireless networks," in *Proceedings 2nd Networked Systems Design and Implementation (NSDI)*, Boston, MA, May 2005.
- [13] M. Feldman, C. Papadimitriou, J. Chuang, and I. Stoica, "Free-riding and whitewashing in peer-to-peer systems," in *Proceedings 3rd Annual Workshop on Economics and Information Security (WEIS04)*, May 2004.
- [14] T. Moscibroda, S. Schmid, and R. Wattenhofer, "When Selfish Meets Evil: Byzantine Players in a Virus Inoculation Game," in *Proceedings 25th Annual Symposium on Principles of Distributed Computing (PODC)*, Denver, Colorado, USA, July 2006.
- [15] G. Theodorakopoulos and J. S. Baras, "Malicious users in unstructured networks," in *Proceedings IEEE Infocom 2007*, Anchorage, AK, May 2007.
- [16] G. W. Brown, "Iterative solution of games by fictitious play," in *Activity Analysis of Production and Allocation*, T. Koopmans, Ed. New York: John Wiley and Sons, 1951, pp. 374–376.
- [17] G. Theodorakopoulos and J. S. Baras, "Enhancing benign user cooperation in the presence of malicious adversaries in ad hoc networks," in *Proceedings Second IEEE Communications Society/CreateNet International Conference on Security and Privacy in Communication Networks, SecureComm 2006*, Baltimore, MD, August 2006.



George Theodorakopoulos received the B.S. degree from the National Technical University of Athens, Greece, and the M.S. and Ph.D. degrees from the University of Maryland, College Park, MD, all in Electrical Engineering, in 2002, 2004, and 2007, respectively.

He is the co-recipient of the Best Paper award at the ACM Workshop on Wireless Security, October 2004, and the 2007 IEEE ComSoc Leonard Abraham prize. He is currently a Senior Researcher at the Ecole Polytechnique Fédérale de Lausanne, Switzerland. His research interests include network security, trust and reputation systems, game theory, and algebraic combinatorics.



John S. Baras received the B.S. in Electrical Engineering with highest distinction from the National Technical University of Athens, Greece, in 1970. He received the M.S. and Ph.D. degrees in Applied Mathematics from Harvard University, Cambridge, MA, in 1971 and 1973 respectively.

Since 1973 he has been with the Department of Electrical and Computer Engineering, University of Maryland at College Park, where he is currently Professor, member of the Applied Mathematics and Scientific Computation Program Faculty, and Affiliate Professor in the Department of Computer Science. From 1985 to 1991 he was the Founding Director of the Institute for Systems Research (ISR) (one of the first six NSF Engineering Research Centers). In February 1990 he was appointed to the Lockheed Martin Chair in Systems Engineering. Since 1991 Dr. Baras has been the Director of the Maryland Center for Hybrid Networks (HYPNET), which he co-founded. Dr. Baras has held visiting research scholar

positions with Stanford, MIT, Harvard, the Institute National de Recherche en Informatique et en Automatique, the University of California at Berkeley, Linköping University and the Royal Institute of Technology in Sweden.

Among his awards are: the 1980 George S. Axelby Prize of the IEEE Control Systems Society; the 1978, 1983 and 1993 Alan Berman Research Publication Award from NRL; the 1991 and 1994 Outstanding Invention of the Year Award from the University of Maryland; the 1996 Engineering Research Center Award of Excellence for Outstanding Contributions in Advancing Maryland Industry; the 1998 Mancur Olson Research Achievement Award, from the Univ. of Maryland College Park; the 2002 Best Paper Award at the 23rd Army Science Conference; the 2004 Best Paper Award at the Wireless Security Conference WISE04; the 2007 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communication Systems.

Dr. Baras' research interests include control, communication and computing systems. He is a Fellow of the IEEE and a Foreign Member of the Royal Swedish Academy of Engineering Sciences (IVA).