Trust–based Strategies for Wireless Networks Under Partial Monitoring

Konstantinos Ntemos* Nicholas Kalouptsidis* Nicholas Kolokotronis†

*Dept. of Informatics and Telecommunications
University of Athens, 15784 Athens, Greece
Emails: {kdemos, kalou}@di.uoa.gr

†Dept. of Informatics and Telecommunications
University of Peloponnese, 22100 Tripolis, Greece
Email: nkolok@uop.gr

***Abstract*—In modern wireless networks autonomous agents may exhibit selfish or malicious behavior which can compromise the performance of the network. For this reason, Intrusion Detection Systems (IDS) have been proposed to monitor the agents' behavior, along with the deployment of Trust/reputation Management Systems (TMS) to enforce cooperation among the agents. IDS may not continuously monitor agents' behavior to avoid excessive deployment costs. In this work we consider agents that exhibit both selfish and malicious behavior and study their pairwise interactions when they participate in a packet–forwarding task, in the scenario of partial monitoring of their actions by the IDS. We investigate the decision–making process of the agents and derive conditions that if satisfied, the trust–based strategy proposed by the TMS constitutes an optimal strategy for the agents.**

***Index Terms*—Trust; Game Theory; Monitoring; Cooperation.**

I. Introduction

In modern wireless networks, autonomous agents, that do not necessarily belong to the same authority, interact to pursue individual goals. In order for these networks to work efficiently, agents must cooperate and forward packets to intended destinations. However, it is unrealistic to take for granted that agents will always cooperate in an altruistic fashion. The main types of "misbehavior" are *selfish* and *malicious* behavior. Selfish agents are interested in receiving cooperation benefits from others while contributing as little as possible to the forwarding task to reduce transmission cost. In contrast, malicious agents aim at harming other agents by launching various kinds of attacks such as packet/content modification, Byzantine and Sybil attacks, amongst others.

Game Theory (GT) has been widely utilized for studying interactions among autonomous self-interested agents participating in a packet–forwarding task in wireless networks [1], [2], [5]. It is shown in [5] that network–wide cooperation among autonomous self-interested is unlikely to occur without providing *incentives*. For this reason, *trust/reputation models* [1]–[4] have been proposed to stimulate cooperation. In [4] the *watchdog* mechanism monitors the agents' misbehavior, while the *pathrater* mechanism selects routes consisting of trustworthy nodes. In [2], [3] reputation systems based on the indirect reciprocity principle are designed.

A Trust Management System (TMS) categorizes agents into *trustworthy* and *untrustworthy* based on their behavior. Apart from updating agents' trust values, an effective TMS aims at enforcing a desired behavior pattern. This behavior helps the trustworthy agents and ignores the untrustworthy ones.

Agents' actions are usually assumed to be observable. An exception is [1] which considers interactions among selfish agents and an agent's packet forwarding action may be perceived by another agent as packet drop due to not overhearing the transmission. The monitoring mechanisms deployed for detecting agents' actions and especially misbehavior patterns are called *Intrusion Detection Systems* (IDS). The most commonly used IDS technique is *overhearing* an agent's transmissions [1], [4], [8]. However, low energy budget networks, such as WSNs, may employ partial monitoring IDS schemes that periodically sample the agents' packets to decide on their behavior as discussed in [9].

In this paper, we consider agents that can act both selfishly by not forwarding others' packets and maliciously by launching packet modification attacks, a subject less considered in the literature. We assume that an IDS performs overhearing periodically and then based on the IDS action detection outcome, the TMS updates the agents' trust values. We then study the impact of the sampling probability of the IDS along with the properties of the TMS on the decision–making process of the agents. We highlight the implications in the analysis caused by considering malicious behavior and study the effectiveness and limitations of *trust–based strategies* when the agents condition their actions on the publicly observed trust values. Finally, we derive conditions that if satisfied an agent is incentivized to follow a desired strategy which is beneficial for the network, despite its myopic temptation to misbehave. In doing so, we formulate the agents interactions as a *Stochastic Game* (SG) and the solution concept of *Perfect Public Equilibrium* (PPE) is utilized.

The rest of the paper is organized as follows. Section II outlines the system model and the trust update mechanism. The decision making process is analyzed in Section III. The results are discussed in Section IV while conclusions are given in section V.

II. System Model

A. Agents' interactions and modeling

We study the interactions of two agents i, j which want to forward their packets to their intended destinations d_i, d_j,

ISBN 978-0-9928626-7-1 © EURASIP 2017

respectively. Node d_i (resp. d_j) is outside the transmission range of i (resp. j). Thus, there is need for i (resp. j) to forward the packets of j (resp. i) to d_j (resp. d_i). Both agents may exhibit selfish or malicious behavior. The set of admissible transmission decisions (i.e. actions) for an agent is

$$A = \{-1, 0, 1\}. \tag{1}$$

The values $-1, 0, 1$ correspond to malicious, selfish and honest action, meaning the agent chooses to launch an attack (i.e. modify the packets and then forward them), not to forward the packets and forward them to the destination, respectively.

The actions of the agents can be observed by *overhearing* agents' transmissions, which is possible due to the broadcast nature of wireless transmissions [4], [8]. In this work we assume that an IDS periodically performs the overhearing process and then the TMS updates the trust values of the agents and disseminates the new ones to them.

We assume a slotted time structure. At every time slot t the IDS monitors the agents' actions with a *sampling probability* P_a. We use $I_d = 1$ to denote that the IDS performs sampling, while $I_d = 0$ denotes that IDS is idle during current time slot. The interactions of two transmitting agents i, j during a time slot t are summarized as follows:

1. TMS informs agents about their trust values s_i, s_j.
2. If IDS performs sampling (i.e. $I_d = 1$)
   - The agents (i, j) exchange the packets to be forwarded to the corresponding destinations.
   - IDS overhears the original packets.
   - Agents select their actions a_i, a_j.
   - IDS overhears the transmitted packets if any (i.e. $a_i, a_j \neq 0$), then compares them with the original packets to check for content modifications. If it detects $a_i = -1$ (resp. $a_j = -1$), it informs destination d_j (resp. d_i) to discard the modified packets, else d_j (resp. d_i) uses the packets.
   - Based on current trust values s_i, s_j and detected actions a_i, a_j, the TMS updates trust values to s_i', s_j' for the next time slot.
3. If IDS does not perform sampling (i.e. $I_d = 0$)
   - Agents exchange their packets to be forwarded to the corresponding destinations.
   - Agents select their actions (a_i, a_j).
   - Receivers d_i, d_j use the packets received.

Note that agents i, j do not know at the beginning of the time slot if the IDS will perform sampling but they know the value of the sampling probability P_a. Moreover, due to the existence of malicious agents a distributed IDS cannot rely on exchange of detected actions as in the selfish agents case [1]. Moreover, the source can not directly notify its intended destination to disregard modified packets. For these reasons, a centralized IDS/TMS is employed in this study.

B. Instantaneous Reward

Each agent wants its packets to reach the desired destination. This interest is captured by a forwarding benefit $f > 0$ if the other agent forwards its packets (i.e. $a = 1$). Forwarding the other agent's packets incurs a transmission cost $c > 0$. Finally, if an agent launches an undetected attack (i.e. the IDS does not perform sampling), then an illegal gain $e > 0$ would be acquired expressing the gain of a successful attack. In this case, the other agent would suffer a loss $\ell > 0$ from the attack. Agent i does not know whether the IDS will be active in the current time slot, so it uses the sampling probability P_a to form the *instantaneous expected reward*

$$R_i(a_i(t), a_j(t)) = R_i^x(a_i(t)) + R_i^r(a_j(t)), \tag{2}$$

where

$$R_i^x(a_i(t)) = \begin{cases} -c_i, & \text{if } a_i(t) = 1, \\ 0, & \text{if } a_i(t) = 0, \\ (1 - P_a)e_i - c_i & \text{if } a_i(t) = -1 \end{cases} \tag{3}$$

$$R_i^r(a_j(t)) = \begin{cases} f_i, & \text{if } a_j(t) = 1, \\ 0, & \text{if } a_j = 0 \\ -(1 - P_a)\ell_i & \text{if } a_j(t) = -1 \end{cases}. \tag{4}$$

R_i^x and R_i^r are the expected rewards resulting from the transmission actions of agents i and j. R_j is defined accordingly.

Remark 1. We note that in the case of multi-hop communications when agent i wants to send packets to a destination d_i that is more than two hops away, then all the agents which lie in the path from i to d_i are needed to forward the packets of i and thus the reward for agent i would depend on the actions of all the agents on that path.

C. Trust

TMS categorizes the agents into *trustworthy* and *untrustworthy*. Thus, $S = \{0, 1\}$ is the trust value set and a trust value $s_i = 0$ (resp. $s_i = 1$) means that the agent i is considered as untrustworthy (resp. trustworthy). Apart from disseminating and updating the agents' trust values, an effective TMS would ideally like to enforce a desired behavior to the agents. This desired behavior, henceforth called as *honest policy* and denoted as π_{H}, promotes helping the trustworthy agents (i.e. $a_i = 1$ if $s_j = 1$) and ignoring the untrustworthy ones (i.e. $a_i = 0$ if $s_j = 0$). Thus

$$\pi_{\text{H}}(s_i, s_j) = s_j \quad \forall s_i \in S. \tag{5}$$

In this paper we focus on *Markovian* TMS where the updated trust values s_i', s_j' depend on the current trust values s_i, s_j, on the actions a_i, a_j and on whether the IDS performed sampling during the current time slot. If the IDS does not perform sampling (i.e. $I_d = 0$), the trust values remain unchanged

$$\Pr(s_i', s_j' | s_i, s_j, a_i, a_j, I_d = 0) = \mathbb{1}_{\{(s_i', s_j') = (s_i, s_j)\}}, \tag{6}$$

otherwise the trust values are updated based on the following principle. The TMS rewards the agents if they follow the desired honest policy π_{H} and punishes them if they do not.

If both agents are trustworthy (i.e. $(s_i, s_j) = (1, 1)$) and i follows π_{H} (i.e. $a_i = \pi_{\text{H}}(1, 1) = 1$), then i remains as trustworthy (i.e. $s_i' = 1$). If i does not cooperate ($a_i = 0$) or modifies the packets ($a_i = -1$), then it is punished (i.e.

ISBN 978-0-9928626-7-1 © EURASIP 2017

transits to $s_i'=0$) with a punishment rate q and 1, respectively (modifying packets is more harmful than not forwarding packets). If agent i is trustworthy and j untrustworthy (i.e. $(s_i,s_j)=(1,0)$), and i selects $a_i=\pi_{\text{H}}(1,0)=0$, then it remains at $s_i'=1$, while if i helps the untrustworthy j ($a_i=1$) or attacks ($a_i=-1$), then i is punished and transits to $s_i'=0$ with probability 1. If both agents are untrustworthy (i.e. $(s_i,s_j)=(0,0)$) and $a_i=\pi_{\text{H}}(0,0)=0$, then i could be forgiven and transit to $s_i'=1$ with a redemption probability ϕ, while by selecting $a_i=1,-1$, i remains untrustworthy (i.e. $s_i'=0$). Finally, if i is untrustworthy and j is trustworthy (i.e. $(s_i,s_j)=(0,1)$) and $a_i=\pi_{\text{H}}(0,1)=1$, then agent i is forgiven with a redemption rate p, while by selecting $a=0$, or $a=-1$ it remains untrustworthy (i.e. $s_i'=0$). Because trust should be more difficult to earn than to lose, we assume $q\geq\max\{p,\phi\}$. In summary

$$\Pr(s_i'=1|s_i=0,s_j=0,a_i=0,I_d=1)=\phi, \tag{7}$$
$$\Pr(s_i'=1|s_i=1,s_j=0,a_i=0,I_d=1)=1, \tag{8}$$
$$\Pr(s_i'=1|s_i=0,s_j=1,a_i=1,I_d=1)=p, \tag{9}$$
$$\Pr(s_i'=1|s_i=1,s_j=1,a_i=1,I_d=1)=1, \tag{10}$$
$$\Pr(s_i'=1|s_i=1,s_j=1,a_i=0,I_d=1)=1-q. \tag{11}$$

For the combinations of a_i,s_i,s_j not defined in (7)-(11), it is $\Pr(s_i'=1|s_i,s_j,a_i,I_d=1)=0$. Note finally that, $\Pr(s_i'=0|s_i,s_j,a_i,I_d=1)=1-\Pr(s_i'=1|s_i,s_j,a_i,I_d=1)$. Thus

$$\Pr(s_i',s_j'|s_i,s_j,a_i,a_j)=P_a\Pr(s_i'|s_i,s_j,a_i,I_d=1) \tag{12}$$
$$\times\Pr(s_j'|s_j,s_i,a_j,I_d=1)+(1-P_a)\mathbb{1}_{\{(s_i',s_j')=(s_i,s_j)\}}.$$

III. Optimal action selection

In the static case where agents interact only once, an agent wants to maximize its instantaneous expected reward (2). From (2)–(4) it can be seen that choosing to forward the other agent's packets ($a_i=1$) is dominated by either not forwarding ($a_i=0$) or by attacking ($a_i=-1$), regardless of a_j. In particular, if $(1-P_a)e>c$ (resp. if $(1-P_a)e<c$) then $a=-1$ (resp. $a_i=0$) is the optimal myopic choice. If agents interact repeatedly for an infinite (or unknown) time interval, then an agent's goal is to maximize the sum of discounted expected long–term rewards

$$\max_{\{a_i(t)\}_{t=0}^{t=\infty}}\sum_{t=0}^{\infty}\delta_i^t\mathbb{E}\big[R_i(a_i(t),a_j(t))\big]. \tag{13}$$

where the expectation is over other agent's actions and $\delta\in(0,1)$ is the discount factor which expresses the foresightedness of the agent. Agent i decides on its *strategy* using the available information up to time t (history). A (pure) strategy π_i is a sequence of maps π_i^t from histories h_i^t to actions $a_i(t)$. At instant t, the agents' histories are

$$h_i^t=\{s_i^{(0:t-1)},s_j^{(0:t-1)},a_i^{(0:t-1)},s_i(t),s_j(t)\},$$
$$h_j^t=\{s_i^{(0:t-1)},s_j^{(0:t-1)},a_j^{(0:t-1)},s_i(t),s_j(t)\},$$

where $x^{(0:t-1)}=\{x(0),\dots,x(t-1)\}$ for a variable x. Note that $s_i(\tau),s_j(\tau)$ belong to both histories h_i^t and h_t^t for all $t,\tau\leq t$. Thus, the history $h_c^t=\{s_i(0),s_j(0),\dots,s_i(t),s_j(t)\}$ is *public history*. An appropriate solution concept for a game with imperfect action monitoring is *Perfect Public Equilibrium* (PPE) [6] which is an extension of the idea of Subgame Perfect Equilibrium (a refinement of Nash Equilibrium (NE) that eliminates non-credible threats).

Definition 1. The strategy profile (π_i,π_j) is a PPE if (π_i,π_j) are *public strategies* and for each time slot t and history h_t, (π_i,π_j) yields a NE from that time on.

A strategy $\pi=\{\pi^t\}_{t=0}^{t=\infty}$ is called public if for every t, π^t depends only on public history h_c^t and not on agent's private information (in this case the actions of the agents). If a strategy depends only on current history (state) (i.e. $a_i(t)=\pi_i^t(s_i(t),s_j(t)),a_j(t)=\pi_j^t(s_j(t),s_i(t))$ for all t) and not on the previous public history up to t, then it is called *Markovian* public strategy. If additionally it does not depend on t, then it is called *Markovian stationary* public strategy. We note that if an agent employs π_{H} over all time slots, it constitutes a Markovian stationary public strategy. Henceforth we drop subscripts i,j,t whenever it is clear from the context.

Theorem 1. *Suppose $P_a,p,\phi>0$ and $q>=\max\{p,\phi\}$. Suppose further that $p>\phi$. Then the honest strategy profile $(\pi_{\text{H}},\pi_{\text{H}})$ is PPE if and only if the following hold for both agents*

$$f\geq\max\{k_1(1-P_a)e-k_2c,k_3(1-P_a)e,k_3c\}, \tag{14}$$

where

$$k_1=\frac{(1-\delta+\delta P_ap)(1-\delta+2\delta P_a\phi-\delta P_a\phi^2)}{\delta P_a\phi(1-\delta+\delta P_ap+\delta P_a\phi-\delta P_ap\phi)},$$
$$k_2=\frac{(1-\delta+\delta P_a\phi)(1-\delta+\delta P_ap+\delta P_a\phi-\delta P_a\phi^2)}{\delta P_a\phi(1-\delta+\delta P_ap+\delta P_a\phi-\delta P_ap\phi)},$$
$$k_3=\frac{1-\delta+\delta P_ap}{\delta P_ap}.$$

Furthermore, if $p\leq\phi$, $(\pi_{\text{H}},\pi_{\text{H}})$ is PPE if and only if the following hold for both agents

$$f\geq\max\{k_3(1-P_a)e,k_3c\}. \tag{15}$$

Proof. If the honest policy π_{H} constitutes the optimal strategy for both agents at every instant and state (s_i,s_j), then the *Bellman Equation* (BE) [7] for i,

$$V_i(s_i,s_j)=\max_{a_i}\Big\{R_i(a_i,a_j)$$
$$+\delta\sum_{s_i',s_j'}\Pr(s_i',s_j'|s_i,s_j,a_i,a_j)V_i(s_i',s_j')\Big\}, \tag{16}$$

must be satisfied for $a_i=\pi_{\text{H}}(s_i,s_j)$, given that $a_j=\pi_{\text{H}}(s_j,s_i)$. Then the corresponding value function $V_i^{\text{H}}(s_i,s_j)$ for agent i when both agents follow π_{H} is

$$V_i^{\text{H}}(0,0)=\frac{\delta P_a\phi(1-\delta+2\delta P_ap-\delta P_ap\phi)}{(1-\delta)(1-\delta+\delta P_ap)k_4}(f_i-c_i), \tag{17}$$

ISBN 978-0-9928626-7-1 © EURASIP 2017

$$V_i^{\mathrm{H}}(0,1) = \frac{1}{(1-\delta)k_3} f_i - \frac{1}{1-\delta} c_i, \quad (18)$$

$$V_i^{\mathrm{H}}(1,0) = \frac{1}{1-\delta} f_i - \frac{1}{(1-\delta)k_3} c_i, \quad (19)$$

$$V_i^{\mathrm{H}}(1,1) = \frac{f_i - c_i}{1-\delta}, \quad (20)$$

where $k_4 = 1 - \delta + 2\delta P_a \phi - \delta P_a \phi^2$.

The strategy profile $(\pi_{\mathrm{H}}, \pi_{\mathrm{H}})$ is a PPE if and only if the *one shot deviation principle* (OSDP) holds, namely no profitable one shot deviations exist, for every possible public history. OSDP states that if there is a better strategy than the one under consideration, then it is profitable to deviate once and then use the considered strategy for the rest of the game [6]. The only relevant part of public history is the vector of binary trust values (s_i, s_j) giving rise to four possibilities. Thus optimality in BE for agent i gives

$$R(a_i^*, a_j^*) + \delta \sum_{s_i', s_j'} \Pr(s_i', s_j' | s_i, s_j, a_i^*, a_j^*) V_i^{\mathrm{H}}(s_i', s_j') \geq R(a_i, a_j^*) + \delta \sum_{s_i', s_j'} \Pr(s_i', s_j' | s_i, s_j, a_i, a_j^*) V_i^{\mathrm{H}}(s_i', s_j') \quad (21)$$

where $a_i^* = \pi_{\mathrm{H}}(s_i, s_j)$, $a_j^* = \pi_{\mathrm{H}}(s_j, s_i)$. Eq. (21) must hold for every possible state (s_i, s_j) and action $a_i \neq a_i^*$. The same reasoning follows for agent j. Next we consider one shot deviations from the honest profile at each state. Let $(s_i, s_j) = (0,0)$. Then $\pi_{\mathrm{H}}(0,0) = 0$. The possible deviations are $a = 1$ and $a = -1$. We thus evaluate the optimality inequality (21) by utilizing (2), (12), (17)-(20) to obtain

$$f \geq -k_2 c \quad (22) \qquad f \geq k_1(1-P_a)e - k_2 c \quad (23)$$

Next we consider the remaining deviations from π_{H} at the remaining states. Proceeding along similar lines, deviations from π_{H} at $(0,1)$ by selecting $a = 0$, $a = -1$ lead to

$$f \geq k_3 c \quad (24) \qquad f \geq k_3(1-P_a)e \quad (25)$$

respectively. Deviations from π_{H} at $(1,0)$ by selecting $a = 1$, $a = -1$ lead to

$$f \geq k_5 c \quad (26) \qquad f \geq k_6(1-P_a)e + k_5 c \quad (27)$$

respectively. Deviations from π_{H} at $(1,1)$ by selecting $a = 0$, $a = -1$ lead to

$$f \geq \frac{p}{q} k_3 c \quad (28) \qquad f \geq p k_3 (1-P_a)e \quad (29)$$

respectively, where

$$k_5 = \frac{\delta^2 P_a^2 (1-p)(p-\phi) - (1-\delta+\delta P_a p)k_4}{\delta P_a k_7},$$

$$k_6 = \frac{(1-\delta+\delta P_a p)k_4}{\delta P_a k_7},$$

$$k_7 = 1 - \delta + \delta P_a p + \delta P_a \phi - \delta P_a p^2 - \delta P_a \phi^2 + \delta P_a p \phi.$$

For given $\delta, P_a, p, q, \phi, e$, the right hand side of each inequality (22)-(29) defines a line. Let these lines be denoted as $b_1(c)$-$b_8(c)$, respectively and let λ_1-λ_8 be the respective slopes. It is $\lambda_3 = k_3 > 1 > \lambda_2$, as $\lambda_2 = -k_2 < 0$, $\lambda_3 \geq \lambda_7$, because $q \geq p$ and $\lambda_3 > \lambda_5$. Moreover, $b_1(0) = b_3(0) = b_5(0) = b_7(0) = 0$. Thus, $b_3(c) \geq \max\{b_1(c), b_5(c), b_7(c)\}$ for all $c \geq 0$.

Furthermore, it is $b_4 \geq b_8$ (as $0 < p \leq 1$, note that b_4, b_8 are independent of c). Moreover, $0 < b_6(0) \leq b_4$ and $b_6(c)$ intersects b_4 at $E = (c_E, f_E)$ with

$$c_E = \frac{(1-\delta+\delta P_a p)(1-p)k_8}{\delta^2 P_a^2 p(1-p)(p-\phi) - p(1-\delta+\delta P_a p)k_4}(1-P_a)e$$

where

$$k_8 = 1 - \delta + \delta P_a p + \delta P_a \phi - \delta P_a \phi^2. \quad (30)$$

If $\lambda_6 \leq 0$, then $c_E \leq 0$ and thus $b_4 \geq b_6(c)$ for all $c \geq 0$. Let $B = (c_B, f_B)$ denote the intersection of $b_3(c)$ with b_4 with $c_B = (1-P_a)e$. If $\lambda_6 > 0$, then it is $c_B \leq c_E$. Thus, if (24) and (25) hold, then (27) holds. Finally,

$$b_2(0) > b_4(0) \Leftrightarrow p > \phi. \quad (31)$$

Since $\lambda_2 < 0$, if $p \leq \phi$, then $b_4 \geq b_2(c)$ for all $c \geq 0$. On the other hand, if $p > \phi$, $b_2(c)$ intersects b_4 at $A = (c_A, f_A)$, with $c_A = \frac{(p-\phi)(1-\delta+\delta P_a p)}{p k_8}(1-P_a)e$ and it is

$$c_A < c_B \Leftrightarrow \frac{(p-\phi)(1-\delta+\delta P_a p)}{p(1-\delta+\delta P_a p + \delta P_a \phi - \delta P_a \phi^2)} < 1, \quad (32)$$

which holds. Thus, if $p > \phi$ (resp. $p \leq \phi$) then (23), (24) and (25) (resp. (24) and (25)) suffice to represent the region over which the honest policy π_{H} prevails and the rest of inequalities are redundant. The proof is complete. ■

Next, we consider the behavior of the set defined by (14) if $p > \phi$ and (15) if $p \leq \phi$ as the illegal gain e varies. This set describes the structural properties of reward-related parameters f, c, e that are consistent with the optimality of the honest policy π_{H} for given values of δ, P_a, p, ϕ, q (*see* Fig. 1).

Proposition 1. *The set*

$$K(e) = \{(c,f) \in \mathbb{R}^2 : c, f \geq 0 \text{ and (14) holds}\}, \text{ if } p > \phi$$
$$K(e) = \{(c,f) \in \mathbb{R}^2 : c, f \geq 0 \text{ and (15) holds}\}, \text{ if } p \leq \phi$$

is decreasing in e, i.e. $K(e') \subset K(e)$ *for* $e' > e$.

Proof. If $p > \phi$, then $K(e)$ is defined by $b_2(c), b_3(c), b_4$. The slopes $\lambda_2 = -k_2$, $\lambda_3 = k_3$, $\lambda_4 = 0$ are independent of e. The offsets of $b_2(c), b_4$ (i.e. $k_1(1-P_a)e, k_3(1-P_a)e$, respectively) are both increasing in e, as the coefficients of e are always positive for $0 < \delta < 1$, $0 < P_a, p, \phi$, while the offset of $b_3(c)$ is equal to 0. Let $A' = (c_{A'}, f_{A'})$, $B' = (c_{B'}, f_{B'})$, $D' = (c_{D'}, f_{D'})$ be the intersection points defining $K(e')$ for $e' > e$. Then $c_{B'} > c_B$, $f_{B'} > f_B$, $c_{D'} = c_D = 0$, $f_{D'} > f_D$, $f_{A'} > f_A$ and

$$c_{A'} > c_A \Leftrightarrow k_1 > k_3 \Leftrightarrow p > \phi$$

and according to (32), it is $c_{A'} < c_{B'}$.

ISBN 978-0-9928626-7-1 © EURASIP 2017

On the other hand, if $p \leq \phi$, then $K(e)$ is defined only by $b_3(c), b_4$, according to Theorem 1. Let $Z = (c_Z, f_Z)$ be the intersection of b_4 with the line $c = 0$. The points defining $K(e)$ are Z, B. For $e' > e$ the points defining $K(e')$ are B' and $Z' = (c_{Z'}, f_{Z'})$, with $c_{Z'} = c_Z = 0$ and $f_{Z'} > f_Z$, because the coefficient of e (i.e. $k_3(1 - P_a)$) is positive. ■

Proposition 1 states that as the illegal gain increases, attacking becomes more tempting and thus the region where π_{H} constitutes the optimal strategy becomes smaller.

IV. Discussion

Theorem 1 states that if it is not profitable for an agent to deviate from π_{H} and select $a = -1$ when being at states $(0, 0)$, $(0, 1)$ (*see* (23), (25)) and $a = 0$ at state $(0, 1)$ (*see* (24)), then there will be no profitable deviations from π_{H} at other states, as well, given that the other agent follows π_{H}.

Assuming only selfish agents that can not take the malicious action, $K(e)$ is actually independent of e and it is defined only by (24) (given that the punishment rate q is at least as big as the redemption rate p). This means that if a selfish agent has enough incentives at state $(0, 1)$ to follow $\pi_{\text{H}}(0, 1) = 1$ and pay the transmission cost c to transit to $(1, 1)$ with probability $P_a p$ in order to receive forwarding benefits f, then it will be beneficial to follow π_{H} at every other state.

Malicious behavior changes the above conditions. Eq. (25) is analogous to (24) and ensures that deviations $a = -1$ at state $(0, 1)$ are not profitable. Eq. (23) ensures that $a = -1$ at $(0, 0)$ is not profitable. This is needed because malicious behavior could have benefits by this deviation as it could result in a positive instantaneous reward as opposed to the selfish case where the myopic optimal choice coincides with $\pi_{\text{H}}(0, 0) = 0$. Thus, there is need to provide agents with incentives not to opt to $a = -1$ at state $(0, 0)$. This is the purpose of the redemption rate ϕ. For instance, if $\phi = 0$, then $(0, 0)$ is an *absorbing state* and $a = -1$ (resp. $a = 0$) is the optimal choice if $(1-P_a)e > c$ (resp. $(1 - P_a)e \leq c$).

The individual impact of the remaining parameters on the honest optimality region $K(e)$ can be assessed by differentiating the respective lines on the boundary. Suppose $p \leq \phi \leq q$. Then analysis of the partial derivatives of $b_3(c)$ and b_4 with respect to P_a and p shows that the set $K(e)$ gets larger as P_a, p increase. This confirms intuition because for higher values of the IDS sampling probability P_a, misbehaving will be detected and punished with higher probability. Moreover, for higher values of the redemption rate p the incentives to comply with the honest policy become larger and misbehaving becomes less tempting.

The impact of the sampling probability P_a and redemption rate p on the honest optimality region can be studied in a similar manner. Suppose $p \leq \phi \leq q$. We express (24) and (25) in terms of P_a and p for given δ, ϕ, q and for a given set of payoff parameters, assuming $f > c, e$. The research region has smooth boundaries that form conic sections. Differentiation shows that p is decreasing in P_a. Thus for higher values of IDS sampling probability P_a, smaller values of the redemption rate

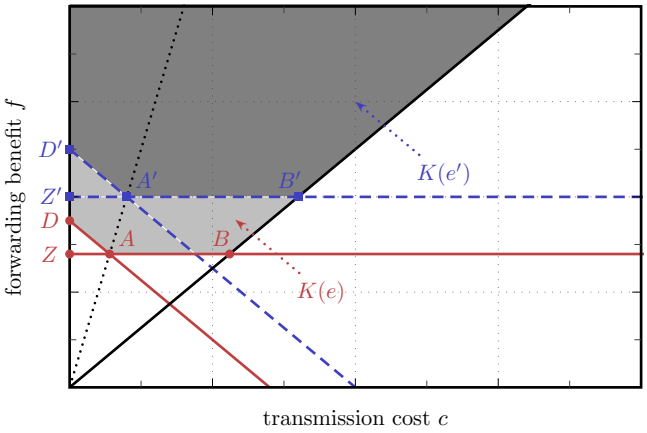

Fig. 1. The regions $K(e)$ and $K(e')$ for $e' > e$ are illustrated with the light gray area being equal to their difference $K(e) \setminus K(e')$.

p suffice to sustain optimality for π_{H}. This confirms intuition because for higher values of P_a, compliance with π_{H} is more likely to be detected and thus the agent to be forgiven and transit to $s' = 1$. Moreover, for higher values of P_a attacking is more likely to be detected.

V. Conclusion

In this paper the impact of action monitoring and trust update mechanisms on the decision making process of autonomous agents able to exhibit both selfish and malicious behavior was investigated. Conditions leading the agents to follow a desired trust–based strategy were derived when the agents condition their actions to the publicly observed trust values. Such conditions can be utilized to jointly design cost-effective IDS and TMS mechanisms for detecting misbehaving patterns and incentivizing agents to follow a desired behavior to enhance network security and other performance measures.

References

[1] J.J. Jaramillo, and R. Srikant, "DARWIN: distributed and adaptive reputation mechanism for wireless ad-hoc networks," In Proc. of *13th annual ACM international conference on Mobile computing and networking, ACM*, pp. 87–98, 2007.

[2] C. Tang, A. Li, and X. Li, "When reputation enforces evolutionary cooperation in unreliable MANETs," *IEEE transactions on cybernetics* vol. 45, no. 10, pp. 2190-2201, 2015.

[3] L. Xiao, et al. "Indirect reciprocity security game for large-scale wireless networks," *IEEE Transactions on Information Forensics and Security* vol. 7, no. 4 , pp. 1368-1380, 2012.

[4] S. Marti, et al. "Mitigating routing misbehavior in mobile ad hoc networks," *In Proc. of the 6th annual international conference on Mobile computing and networking. ACM*, 2000.

[5] M. Felegyhazi, J. P. Hubaux, and L. Buttyan, "Nash equilibria of packet forwarding strategies in wireless ad hoc networks," *IEEE Transactions on Mobile Computing* vol. 5, no. 5, pp. 463-476, 2006.

[6] D. Fudenberg, and J. Tirole, "Game theory," Cambridge, 1991.

[7] D. Bertsekas, "Dynamic programming and optimal control," Vol. 1,2. No. 2. Belmont, MA: Athena Scientific, 1995.

[8] K.F. Ssu, C. H. Chou, and L. W. Cheng, "Using overhearing technique to detect malicious packet-modifying attacks in wireless sensor networks," *Computer Communications* vol. 30, no. 11, pp. 2342-2352, 2007.

[9] B. Subba, S. Biswas, and S. Karmakar, "Intrusion detection in Mobile Ad-hoc Networks: Bayesian game formulation," *Engineering Science and Technology, an International Journal* vol. 19, no. 2, pp. 782-799, 2016.

ISBN 978-0-9928626-7-1 © EURASIP 2017