Optimal Penalties for Misbehavior Deterrence in Communication Networks

Mohamad Khattar Awad[†] *Member, IEEE,* Bashar Zogheib[‡], Hamed M.K. Alazemi[†] *Member, IEEE*
[†]Computer Engineering Department, Kuwait University, Kuwait City, Kuwait
[‡]Department of Mathematics and Natural Sciences, American University of Kuwait, Salmiya, Kuwait
Email: mohamad@ieee.org, bzogheib@auk.edu.kw, hamed@eng.kuniv.edu.kw

***Abstract*—The communication among entities in any network is administered by a set of rules and technical specifications detailed in the communication protocol. All communicating entities adhere to the same protocol to successfully exchange data. Most of the rules are expressed in an algorithm format that computes a decision based on a set of inputs provided by communicating entities or collected by a central controller. Due to the increasing number of communicating entities and large bandwidth required to exchange the set of inputs generated at each entity, distributed implementations have been favorable to reduce the control overhead. In such implementations, each entity self-computes crucial protocol decisions; therefore, can alter these decisions to gain unfair share of the resources managed by the protocol. Misbehaving users degrade the performance of the whole network in-addition to starving well-behaving users. In this work we develop a framework to derive the optimal penalty strategy for penalizing misbehaving users. The proposed framework considers users learning of the detection mechanism techniques and the detection mechanism tracking of the users behavior and history of protocol offenses. Analysis indicate that escalating penalties are optimal for deterring repeat protocol offenses.**

***Index Terms*—Computer Networks Security, Penalty Scheme, Resource Allocation.**

I. Introduction

The performance of communication protocols highly depends on entities cooperation and adherence to the protocol rules and specifications. For instance, in the well known multiple access protocol, slotted ALOHA, communicating entities contend for a channel with an optimal probability computed by the protocol. The calculated optimal probability maximizes the overall network throughput and performance. This does not coincide with the objective of selfish entities that is maximizing their own throughput by contending with higher probability to increases their chance of getting the channel. This is particularly possible in distributed and software-based implementations of the protocol as opposed to centralized and hardware-based implementations [1], [2], [3]. As a result, misbehaving entities throughput increases and that of other well-behaving entities decreases in addition to the overall network performance degradation. Similar scenario arises when entities choose their back-off window in the Transmission Control Protocol (TCP) where choosing a smaller window than the one computed by the protocol reduces the cheating entity transmission delay and improves its Quality of Service (QoS). Such selfish behavior can be observed in any distributed protocol where entities are involved in resource access related decision.

978-1-4799-7492-4/15/$31.00

Research efforts have been focusing on preventing selfishness and misbehavior in communication protocols in addition to mitigating their effect on the network. Most efforts addressing the selfishness and cheating behavior adopted one or more of the following approaches: First, designing a mechanism to identify deviating behavior from the protocol. Second, developing a penalty scheme to penalize misbehaving entities. Third, developing a monitoring scheme to monitor the behavior of convicted entities. In the following we give a brief summary of selected research activities.

Authors in [4] provide a diagnosis scheme to detect abnormal behavior through selecting a parameter of interest and observing them applying it over a short period of time. Non-complying entities are identified as misbehaving ones and penalized accordingly. In [5], authors propose a scheme that monitors the upper and lower limits of a critical protocol parameter. These limits vary as communicating entities succeed or fail in transmitting their data. Entities choosing parameters out of the bound are identified as misbehaving ones and are isolated for a certain duration. The work in [6] gives the trusted and well behaving entities the leverage to enforce a random value of the protocol parameters allowing them to detect non-cooperating entities. These entities are reported to a reputation management system. Similarly, work in [7] follows the same approach in detecting misbehaving entities but proposes a penalty scheme. The penalty scheme assigns the lowest value of the altered parameter (i.e. worst performance) to misbehaving entities.

Most of the above-mentioned contributions focused on detecting greedy and selfish misbehaviors, however, penalization was limited to either isolating the entity for a short period of time, reducing its QoS, or adding it to the set of monitored misbehaving entities. The literature schemes isolate entities to recover network performance loss incurred by modifying the protocol rather than deterring future offenses. Therefore, misbehaving nodes might recurrently repeat the activity provided the gain is still desirable. This calls for a penalty scheme that reduces the likelihood of an entity committing a future protocol offense in addition to punishing the entity for the

present offense to recover the incurred resource loss.

The objective of this work is to develop a mathematical framework under which a penalty scheme can be derived not only to penalize misbehaving entities but also to deter future deviation from the protocol. Misbehaving entities learn about the detection mechanisms employed and avoid them. This study will provide insights into the relation between the misbehaving entities learning of the detection mechanism and the severity of the penalty.

The rest of the paper is organized as follows. Section II introduces the considered generic network model. The users behavior model is presented in Section III and followed by their decision-making model in Section IV. Optimal penalties are derived in Section V and conclusions are drawn in Section VI.

II. Network Model

We consider a network of communication entities participating in a communication process according to a semi-distributed protocol. Both the central controller and communicating entities contribute to the resource access decision process. Furthermore, the entities self-compute one or more of the protocol parameters or implement parameters computed by the central controller. This setting is very common in communication networks where some entities are able to deviate from the protocol rules or alter computed parameters to improve their performance.

The considered network collects information for misbehavior detection and penalty enforcement [8]. The service provider or network operator announces penalties for protocol offenses in the service level agreement. Therefore, entities are aware of the consequences of their misbehavior and committed protocol offenses. Entities are operated by rational users who might commit a protocol offense if the potential benefit in terms of network performance outweighs the expected penalty.

To capture the change of users behaviors over time, three types of users are considered: users complying with the protocol (type $\mathcal{C}$); clean record users who committed at-least one protocol offense and were not detected (type $\mathcal{CR}$), and misbehaving users who were detected for committing at least one protocol offense (type $\mathcal{MB}$). The type of users change over time based on their behavior and compliance with the protocol, and the system performance in detecting misbehaving users.

III. Users Behavior Model

The misbehaving users manipulate the protocol parameters to their advantage in order to improve their performance. In various networks, several QoS performance metrics can be considered for performance improvement; for example, delay, blocking probability, packet dropping, bandwidth [9]. Let the gain or benefit, b, be the motivation behind committing a protocol offense. The benefit derived from committing a protocol offense varies from one user to the other. This variation is modeled by the gain probability distribution function $Z(b)$ and $b \in [0, \widehat{b}]$; $\widehat{b}$ is the maximum benefit can be achieved by not complying with the protocol. When all users comply with the protocol, each one of them receives resource w which represents the user's share of the resources under normal network operations.

We are interested in analyzing repeat protocol offenses, thus we model users behavior over two equal time periods. The general approach adopted for modeling student behavior is based on previous work [10] that models law-breaking behavior and decision making. In the first period, all users join the network as type $\mathcal{C}$ users. A misbehaving user committing a protocol offense in the first period is either detected with probability $p^{\mathcal{C}}$ and its type changes to $\mathcal{MB}$, or not detected with probability $1 - p^{\mathcal{C}}$ and its type changes to $\mathcal{CR}$. Users who don't commit any protocol offense remain of type $\mathcal{C}$. In the second period, type $\mathcal{C}$ users may chose not to misbehave again and remain of type $\mathcal{C}$. Alternatively, users of this type may chose to commit an offense for the first time in the second period and be either detected with probability $p^{\mathcal{C}}$ or not detected with probability $1 - p^{\mathcal{C}}$. On the other hand, type $\mathcal{CR}$ may commit a protocol offense and not get detected with probability $1 - p^{\mathcal{CR}}$ or get detected with probability $p^{\mathcal{CR}}$ which changes its type to $\mathcal{MB}$. Misbehaving users who were detected in the first period, remain of the same type if they don't commit an offense or be identified as repeat offenders if they are detected in the second period with probability $p^{\mathcal{MB}}$.

The detection mechanism can not differentiate between type $\mathcal{C}$ and type $\mathcal{CR}$ students; thus, we refer to them as well-behaving $\mathcal{WB}$ users. The network applies penalty $f^{\mathcal{WB}}$ when detecting $\mathcal{WB}$ user committing an offense while applies penalty $f^{\mathcal{MB}}$ when detecting $\mathcal{MB}$ user. In communication networks, penalties can take various forms. For example, network isolation [8] or performance degradation are possible approaches to penalize misbehaving users. The penalties $f^{\mathcal{WB}}$ and $f^{\mathcal{MB}}$ are of the same form but imposed on different type of users. Finding the optimal penalties $f^{\mathcal{WB}}$ and $f^{\mathcal{MB}}$ is the focus of this work.

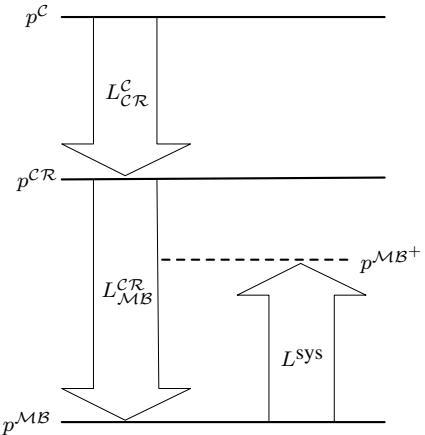

Figure 1. The detection probability for different types of users varying with the learning of the user, $L^{\mathcal{C}}_{\mathcal{CR}}$ or $L^{\mathcal{CR}}_{\mathcal{MB}}$, and learning of the detection system, L^{sys}.

Misbehaving users not only improve their skills over time but also learn how to avoid the detection mechanism. $\mathcal{CR}$ users

find their approach successful in deviating from the protocol without being detected; thus, their detection probability $p^{\mathcal{CR}}$ is less than first time offenders of type $\mathcal{C}$ who are less experienced. The difference is proportional to the learning of type $\mathcal{C}$ users which is denoted by $L^{\mathcal{C}}_{\mathcal{CR}}$. Furthermore, detected $\mathcal{MB}$ users learn more than other users about the detection mechanism of the network which further reduces their detection probability, $p^{\mathcal{MB}}$, for future offenses. The $\mathcal{CR}$ user learning is denoted by $L^{\mathcal{CR}}_{\mathcal{MB}}$. Therefore, the detection probabilities of the three type of users are related by

$$p^{\mathcal{MB}} \leq p^{\mathcal{CR}} \leq p^{\mathcal{C}}. \tag{1}$$

At the same time, the detection mechanism blacklists $\mathcal{MB}$ users and monitors their adherence to the protocol after detections. This increases their detection probability $p^{\mathcal{MB}}$ to $p^{\mathcal{MB}^+}$. The increase is proportional to the detection mechanism tracking of users behavior, types and history denoted by L^{sys}. Figure 1 shows the change in detection probability for different types of users as the learning magnitude changes.

The magnitude of decay in the detection probability from $p^{\mathcal{CR}}$ to $p^{\mathcal{MB}}$ and the magnitude of raise back to $p^{\mathcal{MB}^+}$ play a major role in setting the optimal penalties $f^{\mathcal{WB}}$ and $f^{\mathcal{MB}}$ as will be shown in Section V.

IV. Users Decision-Making

Rational users of communicating entities commit a protocol offense if the expected performance gain is higher than the expected penalty. In other words, knowing the penalties, $f^{\mathcal{WB}}$ and $f^{\mathcal{MB}}$, users evaluate their expected benefit and decide on committing a protocol offense if it is beneficial.

Figure 2 shows users decision alternatives, detection probability, penalties and payoffs. In the first period, all users are of type $\mathcal{C}$ and commit an offense if their benefit is greater than their expected penalty, $b > p^{\mathcal{C}} f^{\mathcal{WB}}$. Therefore, the first timer expected payoff is $[b - p^{\mathcal{C}} f^{\mathcal{WB}}]^+$ where $[.]^+$ denotes $max(., 0)$.

In the second period, type $\mathcal{C}$, $\mathcal{CR}$ and $\mathcal{MB}$ users violate the protocol if $b > p^{\mathcal{C}} f^{\mathcal{WB}}$, $b > p^{\mathcal{CR}} f^{\mathcal{WB}}$ and $b > p^{\mathcal{MB}^+} f^{\mathcal{MB}}$, respectively. Therefore, repeat offenders in the second period have either of the following expected payoffs:

- $[b - p^{\mathcal{CR}} f^{\mathcal{WB}}]^+$ if it was not detected in the first period with probability $1 - p^{\mathcal{C}}$, and
- $[b - p^{\mathcal{MB}^+} f^{\mathcal{MB}}]^+$ if it was detected in the first period with probability $p^{\mathcal{C}}$.

Considering payoffs over both periods, repeat offenders have the following payoff,

$$\begin{aligned} O^r = {} & [b - p^{\mathcal{C}} f^{\mathcal{WB}}]^+ + p^{\mathcal{C}} [b - p^{\mathcal{MB}^+} f^{\mathcal{MB}}]^+ \\ & + (1 - p^{\mathcal{C}}) [b - p^{\mathcal{CR}} f^{\mathcal{WB}}]^+. \end{aligned} \tag{2}$$

However, first timers in the second period have the same payoff as first timers in the first period, $[b - p^{\mathcal{C}} f^{\mathcal{WB}}]^+$. Thus, first timers over both periods have the following payoff,

$$O^f = [b - p^{\mathcal{C}} f^{\mathcal{WB}}]^+. \tag{3}$$

V. Optimal Penalties

Based on the above users behavior model, network operators can set penalties to deter deviation from the protocol. A comparison between equation (2) and (3) implies that deterrence can be achieved by sitting the expected penalty to be larger than or equal the largest expected benefit such that the payoff is eliminated, i.e.,

$$p^{\mathcal{C}} f^{\mathcal{WB}} \geq \widehat{b} \tag{4}$$

$$p^{\mathcal{MB}^+} f^{\mathcal{MB}} \geq \widehat{b} \tag{5}$$

$$p^{\mathcal{CR}} f^{\mathcal{WB}} \geq \widehat{b}. \tag{6}$$

Since $p^{\mathcal{CR}} < p^{\mathcal{C}}$, sufficient deterrence conditions are given by,

$$p^{\mathcal{MB}^+} f^{\mathcal{MB}} = \widehat{b} \tag{7}$$

$$p^{\mathcal{CR}} f^{\mathcal{WB}} = \widehat{b}. \tag{8}$$

To maintain equality, a decrease in detection probability enforces an increase in penalty and vice versa.

It was established in section III that the decrease from $p^{\mathcal{CR}}$ to $p^{\mathcal{MB}}$ by an amount proportional to misbehaving users learning of the detection mechanism $L^{\mathcal{CR}}_{\mathcal{MB}}$. The probability $p^{\mathcal{MB}}$ increases to $p^{\mathcal{MB}^+}$ by an amount proportional to L^{sys}. In communication networks, $\mathcal{CR}$ users improve their skills and learn the detection holes of the system more than what the detection system is able to learn about users, i.e. $L^{\mathcal{CR}}_{\mathcal{MB}} > L^{\text{sys}}$. Thus, the detection probability $p^{\mathcal{CR}}$ is larger than $p^{\mathcal{MB}^+}$, and based on equation (8) the optimal penalty for repeat protocol offenses $f^{\mathcal{MB}}$ is larger than the penalty for the first time offenses $f^{\mathcal{WB}}$. Hence, adopting escalating penalties strategy is optimal for protocol offenses deterrence in communication networks.

Misbehavior detection and penalty enforcement mechanisms can be designed to assign higher penalty to $\mathcal{MB}$ users for every repeated protocol offense. The result is applicable in any network where users learn about the detection mechanism more than what the detection mechanism can learn about the users. In other words, the result applies in all networks with static detection mechanisms.

An example of such networks is the IEEE 802.11 wireless networks. The broadcast nature of the wireless channel makes it accessible to all entities in the network. Access to this valuable resource is abused if communicating entities do not adhere to the protocol rules. Entities trying to maximize their transmission rate, by manipulating the standard protocol, keep the channel busy and prevent other entities from transmitting. This results in suspending or interrupting services provided by the network and can be considered as a form of Denial-of-Service attacks (DoS) [2], [11], [12].

In IEEE 802.11 networks, entities sense the channel and transmit only if the channel is free using Carrier Sense Multiple Access - Collision Avoidance (CSMA-CA). In case the channel is not idle, the transmission is delayed for random back-off value and entities start counting down this value when the channel is sensed idle. Misbehaving nodes violate the IEEE 802.11 protocol and uses shorter and non-random back-off value in order to have a higher chance in accessing the

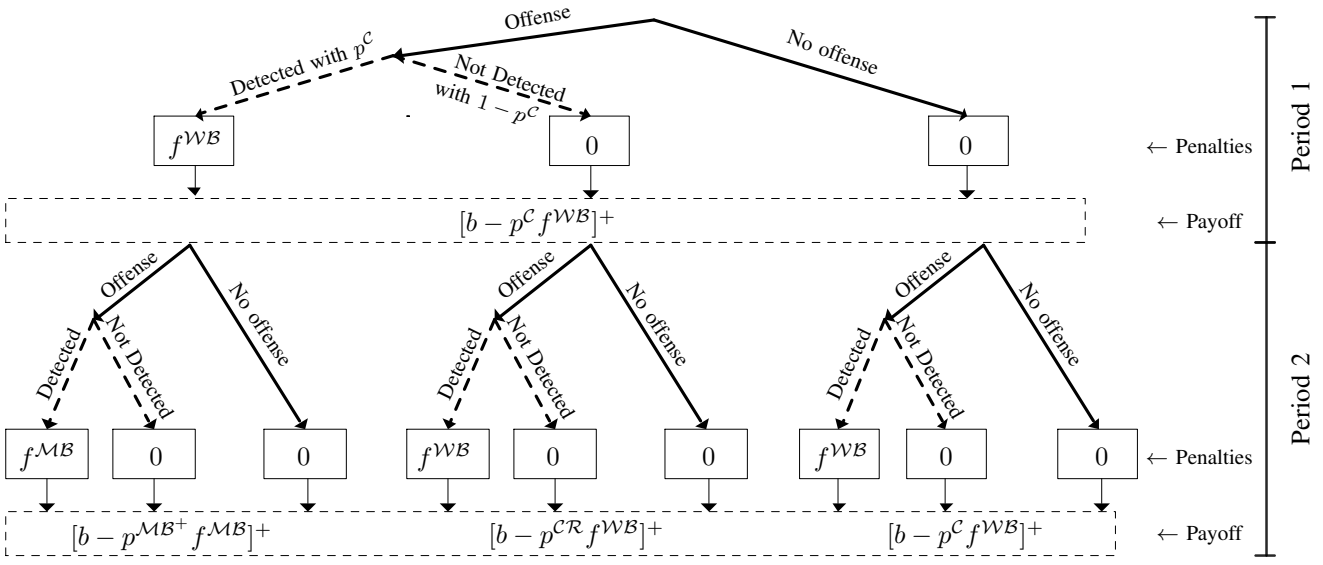

Figure 2. Users decision alternatives and consquences over two periods.

medium. Several research works have focused on detecting and preventing such behavior [6], [13], [14], [9]. In [13], authors proposed a modification to the protocol where the back-off value is assigned by the receiver not the sender. The receiver is able to detect deviation from the protocol by counting the number of idle slots between consecutive transmissions. Detected misbehaving senders are assigned larger back-off values than conforming senders as penalties; however, the repeat offenses are not considered. In light of the proposed analysis, escalating penalties are necessary to prevent deviation from the protocol. For example, the scheme presented in [13] can be improved by tracking misbehaving users offenses and assigning an escalating back-off value for repeat offenses.

VI. CONCLUSIONS

In this paper, the optimal penalties for repeat protocol offenses were analyzed. A network users behavior model was presented. The model captures users history of offenses, detection probability and expected penalties. Moreover, the users decision making process was analyzed taking into consideration users and system learning in addition to expected penalty and gain from committing a protocol offense. Based on both models, optimal penalties were derived. Results indicate that escalating penalties are optimal for deterring misbehavior in communication networks.

ACKNOWLEDGMENT

The authors thank Murat Mungan for very informative discussions on the approach proposed in [10]. This work was supported by Kuwait University, Research Grant No. [EO03/13].

REFERENCES

[1] M. Raya, J.-P. Hubaux, and I. Aad, "Domino: a system to detect greedy behavior in ieee 802.11 hotspots," in *Proceedings of the 2nd international conference on Mobile systems, applications, and services*. ACM, 2004, pp. 84–97.

[2] J. Bellardo and S. Savage, "802.11 denial-of-service attacks: Real vulnerabilities and practical solutions." in *USENIX security*, 2003, pp. 15–28.

[3] L. Guang and C. Assi, "Mitigating smart selfish MAC layer misbehavior in ad hoc networks," in *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, 2006, pp. 116–123.

[4] S. A. S.G. Gollagi and H. Zinage, "MAC layer's misbehavior handling in wireless network," *International Journal of Computational Intelligence Techniques*, vol. 1, no. 2, pp. pp–18–21, 2010.

[5] L. G. H. M. Alazemi, M. J. Khabbaz and A. Fairouz, "Extending the prb access method for wireless networks with misbehaving nodes: Design and analysis," *submitted*.

[6] A. A. Cardenas, S. Radosavac, and J. S. Baras, "Detection and prevention of MAC layer misbehavior in ad hoc networks," in *Proceedings of the 2nd ACM workshop on Security of ad hoc and sensor networks*, 2004, pp. 17–22.

[7] P. Kyasanur and N. Vaidya, "Selfish mac layer misbehavior in wireless networks," *IEEE Transactions on Mobile Computing*, vol. 4, no. 5, pp. 502–516, 2005.

[8] O. F. G. Duque, A. M. Hadjiantonis, G. Pavlou, and M. Howarth, "Adaptable misbehavior detection and isolation in wireless ad hoc networks using policies," in *IFIP/IEEE International Symposium on Integrated Network Management*, 2009, pp. 242–250.

[9] T. Hayajneh, G. Almashaqbeh, and S. Ullah, "A green approach for selfish misbehavior detection in 802.11-based wireless networks," *Mobile Networks and Applications*, pp. 1–13, 2015.

[10] M. C. Mungan, "Repeat offenders: If they learn, we punish them more severely," *International Review of Law and Economics*, vol. 30, no. 2, pp. 173–177, 2010.

[11] V. Gupta, S. Krishnamurthy, and M. Faloutsos, "Denial of service attacks at the MAC layer in wireless ad hoc networks," in *Proceedings of the IEEE Military Communications Conference*, vol. 2, 2002, pp. 1118–1123.

[12] S. Szott, "Selfish insider attacks in ieee 802.11s wireless mesh networks," *IEEE Communications Magazine*, vol. 52, no. 6, pp. 227–233, 2014.

[13] P. Kyasanur and N. Vaidya, "Diagnosing and penalizing mac layer misbehavior in wireless networks," *University of Illinois at Urbana-Champaign, Technical Report*, 2002.

[14] S. Radosavac, A. A. Cárdenas, J. S. Baras, and G. V. Moustakides, "Detecting ieee 802.11 mac layer misbehavior in ad hoc networks: Robust strategies against individual and colluding attackers," *Journal of Computer Security*, vol. 15, no. 1, pp. 103–128, 2007.