Prospect Theoretic Study of Cloud Storage Defense Against Advanced Persistent Threats

Dongjin Xu*, Yanda Li*, Liang Xiao* †, Narayan B. Mandayam‡, H. Vincent Poor§
* Dept. of Communication Engineering, Xiamen University, China. Email: lxiao@xmu.edu.cn
† Key Laboratory of Underwater Acoustic Communication and Marine Information Technology Ministry of Education, Xiamen University, China.
‡ WINLAB, Dept. of ECE, Rutgers University, Piscataway, NJ. Email: narayan@winlab.rutgers.edu
§ Dept. of EE, Princeton University, Princeton, NJ. Email: poor@princeton.edu

***Abstract*—Cloud storage is vulnerable to Advanced Persistent Threats (APTs), which are stealthy, continuous, well funded and targeted. In this paper, prospect theory is applied to study the interactions between a subjective cloud storage defender and a subjective APT attacker. Two subjective APT games are formulated, in which the defender chooses its interval to scan the storage device and the attacker decides its duration between launching two attacks under uncertain APT attack durations and action of the opponent, respectively. The Nash equilibria of the static subjective APT games are derived. We also study the dynamic APT game and propose a Q-learning based APT defense strategy for cloud storage. Simulation results show that the APT defense benefits from the subjective view of the attacker and the proposed defense strategy can improve detection performance with a higher utility.**

***Index Terms*—Cloud storage, advanced persistent threat, game theory, prospect theory.**

I. Introduction

Cloud storage services provide considerable resources and ubiquitous access, and data privacy is critical for their further development [1]. Cloud storage is vulnerable to Advanced Persistent Threats (APTs), in which the attacker applies sophisticated methods to hack the target system continuously and stealthily to steal information over a long term instead of crashing the data immediately [2]. The seminal game theoretic framework proposed in [3] captures the stealthy-takeover property of APT, in which both the attacker and the defender make decisions to maximize their expected utilities according to the Expected Utility Theory (EUT).

However, according to the Allais Paradox [4], subjective decisions made by end-users sometimes deviate from EUT and prospect theory (PT) models these deviations via probability weighting functions and value functions. For example, PT can describe how most people underweight the outcomes with high probabilities, and overweight outcomes with low probabilities [5].

In this paper, we apply prospect theory to investigate the defense of cloud storage against APT, in which a subjective attacker chooses the interval between two trials to compromise the storage device while the defender chooses its scan interval to recapture the compromised device. We formulate two subjective APT storage games, for both uncertain attack durations and unknown action of the opponent. We derive the Nash equilibria (NEs) of the subjective games and investigate the impact of the end-user subjectivity on the APT defense of cloud storage. We also propose an APT defense strategy based on Q-learning to derive the scan interval via trials in dynamic APT games without being aware of the system parameters such as the attack cost and the gain from a longer scan interval. Simulation results show that the subjectivity of the APT attacker decreases its attack rate while a subjective defender tends to inspect the cloud storage more frequently to address APT.

The contributions of this work can be summarized as:

- We formulate two subjective APT games under uncertain attack durations and actions of the opponent, respectively, derive their NEs, and provide conditions under which the equilibria exist.
- We propose a Q-learning based defense scheme to detect APT for cloud storage and investigate its performance against subjective APT attackers in dynamic subjective APT games via simulations.

The remainder of the paper is organized as follows. We review related work in Section II and present the system model in Section III. We present a subjective APT game with uncertain attack durations in Section IV and that with unknown action of the opponent in Section V. We investigate a dynamic APT game in Section VI. We provide simulation results in Section VII and conclude in Section VIII.

II. Related Work

Game theory has been applied to investigate network security against APT attackers. A cyber-physical signaling game among an attacker, a cloud defender and a cloud-connected device was formulated in [6], in which the device judiciously decides whether to trust the command from the cloud threatened by an attacker. The dynamic programming algorithm proposed in [7] provides a nearly optimal defense strategy against stealthy attacks. The interaction between an overt

The work was supported in part by National Science Foundation of China (61271242), and the U. S. National Science Foundation under Grants CMMI-1435778, ECCS-1549881, CNS-1421961 and ACI-1541069.

978-1-5090-1328-9/16/$31.00 ©2016 IEEE

defender and a stealthy attacker investigated in [8] shows that a defender with a periodic strategy has a best response against a non-adaptive attacker. The two-layer game model proposed in [9] investigated the joint threat from an APT attacker and insiders.

Prospect theory was applied in [10] to study a two-user random access game. The subjective wireless random access and data pricing game was identified in [11]. The spectrum investment of a subjective secondary operator was investigated in [12], yielding an optimal sensing and leasing decision. An energy exchange game among microgrids and a power plant formulated in [13] provides the criteria to design the energy price in the local energy market for subjective users. PT-based pricing proposed in [14] improves the revenue of service provider and maintains the radio resource allocation over subjective users. An anti-jamming transmission game formulated in [15] investigated the user subjectivity on data throughput in cognitive radio networks.

III. System Model

We consider an APT attacker (A) against a storage device, which is protected by a storage defender (D), as shown in Fig. 1. The defender chooses the time interval between the k-1-th scan and the k-th scan of the storage device, denoted by x^k. The defender is assumed to restore the storage once it detects an attack. The APT attacker observes the defense strategies of the storage defender and takes stealthy actions. Specifically, the APT attacker knows whether the compromised storage device has been recaptured while the defender is unaware of whether the device is compromised unless it scans the storage device. The attacker waits a time interval y^k to start launching the k-th APT attack after the defender detects and restores the storage device. The duration for the attacker to complete the k-th attack, denoted by z^k, is in general a random variable that is difficult to predict. The defender takes charge of the storage device at the beginning.

Prospect theory captures the behavior deviation from that of EUT, such as risk seeking and loss aversion. The probability weighting function proposed in [4] models the subjective decision-making, mapping from an objective probability to a subjective one. As in [16], we will use Prelec's probability weighting function, denoted by $w_r(p)$, to investigate subjective decision-making of player r, with $r = A, D$, in the APT games, with

$$w_r(p) = \exp\left(-\left(-\ln p\right)^{\alpha_r}\right), \tag{1}$$

where the objective probability weight $\alpha_r \in (0, 1]$ represents the subjective level of player r in making decisions.

IV. Subjective APT game with pure strategy

We formulate the interaction between an APT attacker and a defender who compete to take charge of a storage device at time k as a static subjective APT game $\mathbb{G}$, in which the defender chooses $x \in (0, 1]$, the interval between its k-1-th and k-th scans, while the attacker decides $y \in [0, 1]$, the interval between its k-1-th and k-th attacks. We assume $x > 0$ for simplicity. As shown in Fig. 1, the normalized time during which the device is safe is given by $\min\left(\left(y+z\right)/x, 1\right)$. The gain of the defender to wait unit time to scan is denoted by G, and the cost for the attacker to launch APT is C. Similarly to [7], the utility of the defender denoted by u_D and that of the attacker denoted by u_A are defined respectively, as

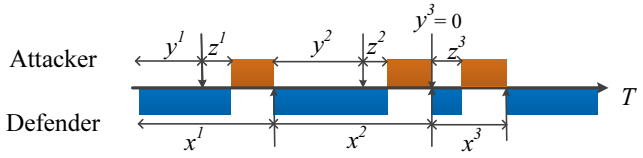


Fig. 1. Illustration of an APT game, in which the defender scans the storage device after interval x^k, the attacker launches APT after interval y^k and the attack duration is z^k, where k is the index of the interaction.

$$u_D(x, y) = \min\left(\frac{y+z}{x}, 1\right) + xG \tag{2}$$

$$u_A(x, y) = -\min\left(\frac{y+z}{x}, 1\right) - \mathrm{I}(y < x)C, \tag{3}$$

where $\mathrm{I}(\cdot)$ is the indicator function that takes 1 if the event is true and 0 otherwise.

The time that the attacker takes to control the device follows the distribution $[P_l]_{0\le l\le L}$, where $P_l = Pr(z = l/L)$. By definition, we have $P_l \ge 0$ and $\sum_{l=0}^{L} P_l = 1$. According to (2) and (3), the expected utilities of the defender and the attacker over the realizations of z, denoted by U_D^{EUT} and U_A^{EUT}, respectively, are given by

$$U_D^{EUT}(x, y) = \sum_{l=0}^{L} P_l \min\left(\frac{yL+l}{xL}, 1\right) + xG \tag{4}$$

$$U_A^{EUT}(x, y) = -\sum_{l=0}^{L} P_l \min\left(\frac{yL+l}{xL}, 1\right) - \mathrm{I}(y < x)C. \tag{5}$$

If the defender and the attacker distort probabilities of events with their objective probability weights to make decisions, prospect theory can capture the subjective decision-making processes. The utilities of the defender and the attacker under prospect theory, denoted by U_D^{PT} and U_A^{PT}, are given respectively by

$$U_D^{PT}(x, y) = \sum_{l=0}^{L} w_D(P_l) \min\left(\frac{yL+l}{xL}, 1\right) + xG \tag{6}$$

$$U_A^{PT}(x, y) = -\sum_{l=0}^{L} w_A(P_l) \min\left(\frac{yL+l}{xL}, 1\right) - \mathrm{I}(y < x)C. \tag{7}$$

A Nash equilibrium of the subjective APT game $\mathbb{G}$ is denoted by (x^*, y^*), in which both the attacker and the defender choose their best-response strategies given that the

opponent chooses the NE strategy, i.e.,

$$U_D^{PT}(x^*, y^*) \geq U_D^{PT}(x, y^*), \quad 0 < x \leq 1 \tag{8}$$
$$U_A^{PT}(x^*, y^*) \geq U_A^{PT}(x^*, y), \quad 0 \leq y \leq 1. \tag{9}$$

Theorem 1. The NE of the subjective APT game $\mathbb{G}$ with $L = 2$ non-zero attack time quantization levels is given by

$$(x^*, y^*) = \begin{cases} (1/2, 0), & I_1 \quad (10a) \\ (1, 0), & I_2 \quad (10b) \\ (1, 1), & I_3, \quad (10c) \end{cases}$$

where

$$I_1: \ G < \exp\left(-\left(-\ln P_1\right)^{\alpha_D}\right) \tag{11}$$
$$C < \exp\left(-\left(-\ln P_0\right)^{\alpha_A}\right) \tag{12}$$
$$I_2: \ G > \exp\left(-\left(-\ln P_1\right)^{\alpha_D}\right) \tag{13}$$
$$C < \exp\left(-\left(-\ln P_0\right)^{\alpha_A}\right) + \frac{1}{2}\exp\left(-\left(-\ln P_1\right)^{\alpha_A}\right) \tag{14}$$
$$I_3: \ C > \exp\left(-\left(-\ln P_0\right)^{\alpha_A}\right) + \frac{1}{2}\exp\left(-\left(-\ln P_1\right)^{\alpha_A}\right). \tag{15}$$

Proof: By (1) and (7), we have

$$\frac{\partial U_A^{PT}}{\partial y} = \begin{cases} -\frac{1}{x}\exp\left(-\left(-\ln P_0\right)^{\alpha_A}\right) < 0, & 0 < x - y \leq \frac{1}{2} \\ -\frac{1}{x}\left(\exp\left(-\left(-\ln P_0\right)^{\alpha_A}\right)\right. & \\ \left. + \exp\left(-\left(-\ln P_1\right)^{\alpha_A}\right)\right) < 0, & \frac{1}{2} < x - y \leq 1 \\ 0, & o.w. \end{cases}$$

Thus, U_A^{PT} decreases with $0 \leq y < x$ and is constant if $x \leq y \leq 1$. It is clear that $U_A^{PT}(x, 0) \geq U_A^{PT}(x, \hat{y})$, for $x \leq \hat{y} < 1$, if (12) holds with $0 < x \leq 1/2$, or (14) holds with $1/2 < x \leq 1$. Thus (9) holds for $(x, 0)$. Otherwise, if (10c) holds, we have (9) holds for $(x, \hat{y})$, $\forall x \leq \hat{y} \leq 1$. By (1) and (6), we have

$$\left.\frac{\partial U_D^{PT}}{\partial x}\right|_{y=0} = \begin{cases} G, & 0 < x \leq \frac{1}{2} \\ -\frac{\exp(-(-\ln P_1)^{\alpha_D})}{2x^2} + G, & \frac{1}{2} < x \leq 1 \end{cases}$$
$$\left.\frac{\partial^2 U_D^{PT}}{\partial x^2}\right|_{y=0} = \begin{cases} 0, & 0 < x \leq \frac{1}{2} \\ \frac{1}{x^3}\exp\left(-\left(-\ln P_1\right)^{\alpha_D}\right), & \frac{1}{2} < x \leq 1 \end{cases},$$

indicating that $U_D^{PT}(x, 0)$ is maximized at $x = 1/2$ or 1. By (6), if (11) holds, we have $U_D^{PT}(1/2, 0) > U_D^{PT}(1, 0)$ and thus (8) holds for $(x^*, y^*) = (1/2, 0)$. Thus, if both (11) and (12) hold, $(1/2, 0)$ is an NE of the game.

Similarly, we see that $(1, 0)$ and $(1, 1)$ are another two NEs of the subjective APT game for I_2 and I_3, respectively. ■

Remark: Under a low attack cost (i.e., (12) and (14)), the attacker launches APT attacks immediately and the defender maximizes its scan interval to save energy (i.e., (13)). Otherwise, if the attack cost is high (i.e., (15)), a subjective APT attacker has no motivation to launch APT.

V. Subjective APT Game with mixed strategy

We consider a subjective APT game denoted by $\mathbb{G}'$, in which the defender chooses the scan interval, with $x \in \{m/M\}_{1 \leq m \leq M}$, while the attacker quantizes its attack interval, with $y \in \{n/N\}_{0 \leq n \leq N}$. In this mixed-strategy game, the scan interval x is chosen according to the mixed strategy $\mathbf{p} = [p_m]_{1 \leq m \leq M}$, where $p_m = Pr(x = m/M)$ is the probability to detect the storage device after x, while the attack interval y is selected based on the strategy $\mathbf{q} = [q_n]_{0 \leq n \leq N}$, where $q_n = Pr(y = n/N)$ is the probability for the attacker to wait y to launch APT. By definition, we have $p_m \geq 0, q_n \geq 0, \sum_{m=1}^{M} p_m = 1$ and $\sum_{n=0}^{N} q_n = 1$. For simplicity, we assume that an attacker takes a constant time to control the device in this game. The expected utility of the defender and that of the attacker are given by (2) and (3) as

$$U_D^{EUT}(\mathbf{p}, \mathbf{q}) = \sum_{m=1}^{M}\sum_{n=0}^{N} p_m q_n \times \left(\min\left(\frac{nM + zMN}{mN}, 1\right) + \frac{mG}{M}\right) \tag{16}$$
$$U_A^{EUT}(\mathbf{p}, \mathbf{q}) = \sum_{m=1}^{M}\sum_{n=0}^{N} p_m q_n \times \left(-\min\left(\frac{nM + zMN}{mN}, 1\right) - I\left(\frac{n}{N} < \frac{m}{M}\right)C\right). \tag{17}$$

If both the defender and the attacker hold subjective views under uncertain actions of their opponents, their strategies are chosen to maximize their PT-based utilities, given by

$$U_D^{PT}(\mathbf{p}, \mathbf{q}) = \sum_{m=1}^{M}\sum_{n=0}^{N} p_m w_D(q_n) \times \left(\min\left(\frac{nM + zMN}{mN}, 1\right) + \frac{mG}{M}\right) \tag{18}$$
$$U_A^{PT}(\mathbf{p}, \mathbf{q}) = \sum_{m=1}^{M}\sum_{n=0}^{N} w_A(p_m) q_n \times \left(-\min\left(\frac{nM + zMN}{mN}, 1\right) - I\left(\frac{n}{N} < \frac{m}{M}\right)C\right). \tag{19}$$

By definition, the NE of the subjective mixed-strategy APT

Algorithm 1 Q learning-based APT defense strategy.

Initialize $\gamma, \delta, y^0 + z^0, Q(s, x) = 0, V(s) = 0, \forall s, x$.
For $k = 1, 2, 3, ...$
  Update the state $s^k = y^{k-1} + z^{k-1}$
  Choose x^k with an ϵ-greedy policy
  Detect the storage after time x^k
  Observe utility u_D and $y^k + z^k$
  Update $Q(s^k, x^k)$ via (32)
  Update $V(s^k)$ via (33)
End for

game $\mathbb{G}'$, denoted by $(\mathbf{p}^*, \mathbf{q}^*)$ is given by

$$\begin{cases} \mathbf{p}^* = \arg\max_{\mathbf{p}} U_D^{PT}(\mathbf{p}, \mathbf{q}^*) & \text{(20a)} \\ \mathbf{q}^* = \arg\max_{\mathbf{q}} U_A^{PT}(\mathbf{p}^*, \mathbf{q}) & \text{(20b)} \\ \sum_{m=1}^{M} p_m = 1, \mathbf{p} \succeq \mathbf{0} & \text{(20c)} \\ \sum_{n=0}^{N} q_n = 1, \mathbf{q} \succeq \mathbf{0}. & \text{(20d)} \end{cases}$$

Theorem 2. The NE of the subjective APT game $\mathbb{G}'$ is given by

$$\begin{cases} \left[u_D(\frac{m}{M}, \frac{n}{N})\right]_{1\leq m\leq M, 0\leq n\leq N} \left[w_D\left(q_k^*\right)\right]_{0\leqslant k\leqslant N}^T = \lambda_D \mathbf{1}_{N+1} & \text{(21a)} \\ \left[u_A(\frac{m}{M}, \frac{n}{N})\right]^T_{1\leq m\leq M, 0\leq n\leq N} \left[w_A\left(p_k^*\right)\right]_{1\leqslant k\leqslant M}^T = \lambda_A \mathbf{1}_M & \text{(21b)} \\ \sum_{m=1}^{M} p_m^* = 1, \mathbf{p} \succeq \mathbf{0} & \text{(21c)} \\ \sum_{n=0}^{N} q_n^* = 1, \mathbf{q} \succeq \mathbf{0} & \text{(21d)} \\ \lambda_D \geq 0, \lambda_A \leq 0, & \text{(21e)} \end{cases}$$

if its solution exists, where $\mathbf{1}_\eta$ represents the η-dimensional all-1 column vector.

Proof: We define L_D as

$$L_D = U_D^{PT}\left(\mathbf{p}, \mathbf{q}^*\right) - \varphi\left(\sum_{m=1}^{M} p_m - 1\right) + \sum_{m=1}^{M} \mu_m p_m. \quad (22)$$

The Karush-Kuhn-Tucker (KKT) conditions of (20) are given by

$$\begin{cases} \frac{\partial L_D}{\partial p_m} = 0 \\ -p_m \leq 0, \mu_m \geq 0, \mu_m p_m = 0, \ \ 1 \leq m \leq M \\ \sum_{m=1}^{M} p_m - 1 = 0 \end{cases}. \quad (23)$$

According to (18) and (22), we apply the complementary slackness for (23) and obtain

$$\begin{cases} \sum_{n=0}^{N} u_D\left(\frac{k}{M}, \frac{n}{N}\right) w_D(q_n^*) - \lambda_D = 0, \ \ 1 \leq k \leq M \\ \sum_{m=1}^{M} p_m = 1 \\ \lambda_D \geq 0 \end{cases}, \quad (24)$$

yielding (21a). Similarly, we obtain (21b). ■

Corollary 1. If $M = 2$, $N = 1$ and

$$\begin{cases} \dfrac{u_D(1/2,0) - u_D(1,0)}{u_D(1,1) - u_D(1/2,1)} > 1 & \text{(25a)} \\ \dfrac{u_A(1/2,1) - u_A(1/2,0)}{u_A(1,0) - u_A(1,1)} > 1, & \text{(25b)} \end{cases}$$

the subjective APT game $\mathbb{G}'$ has a unique NE, which is given by

$$\ln\left(\frac{u_A(1/2,1) - u_A(1/2,0)}{u_A(1,0) - u_A(1,1)}\right) + \left(-\ln\left(1 - p_1^*\right)\right)^{\alpha_A} - \left(-\ln\left(p_1^*\right)\right)^{\alpha_A} = 0 \quad (26)$$

$$\ln\left(\frac{u_D(1/2,0) - u_D(1,0)}{u_D(1,1) - u_D(1/2,1)}\right) + \left(-\ln\left(1 - q_0^*\right)\right)^{\alpha_D} - \left(-\ln\left(q_0^*\right)\right)^{\alpha_D} = 0. \quad (27)$$

Proof: According to (1), (21a) and (25a), we have (27). Similarly, we can obtain (26) by (1), (21b) and (25b). Next, we prove the uniqueness of q_0^*. Note that $f(x) = \left(-\ln(x)\right)^\alpha$ monotonically decreases with x. By (25a) and (27) we have $f(q_0^*) > f(q_1^*)$, yielding $0 < q_0^* < q_1^* < 1$. As $q_0^* + q_1^* = 1$, we have $0 < q_0^* < 1/2$. If $0 < x < 1/2$, we have

$$\frac{d(f(1-x) - f(x))}{dx} = f'(1-x) - f'(x) > 0, \quad (28)$$

indicating that $f(1-x) - f(x)$ increases with x. Therefore, (27) has a unique solution. Similarly, p_1^* is unique. ■

According to Corollary 1, the NE of the EUT-based APT game is given by (26) and (27) as

$$p_1^* = \frac{1 - z - C}{\min(2z, 1) - z} \quad (29)$$

$$q_0^* = \frac{G}{2\min(2z, 1) - 2z}. \quad (30)$$

VI. Dynamic Subjective APT game

In dynamic subjective APT games, the defender and attacker are usually unaware of the system parameters such as the attack cost (C) and the gain from the scan interval (G). As a simple and widely-used reinforcement learning algorithm, Q-learning [17] enables the storage defender to derive its optimal strategy via trial-and-error without knowing the APT attack model and parameters in advance.

In the dynamic APT game with an objective storage defender and a subjective attacker under uncertain duration to successfully compromise the storage device, the defense interval can be determined by assuming that the attack interval is chosen to maximize the PT utility of the attacker given in Eq. (7) according to the attack history in the last time slot, i.e.,

$$Pr(\hat{y} = \tilde{y}) = \begin{cases} 1 - N\nu, & \tilde{y} = \arg\max_{\mathbf{y}} U_A^{PT}(\hat{x}, y) \\ \nu, & o.w. \end{cases}, \quad (31)$$

where ν is a small positive value, with $0 < \nu < 1$.

At time k, the defender observes the total attack duration in the last slot, i.e., $y^{k-1} + z^{k-1}$, and the system state is given by $s^k = z^{k-1} + y^{k-1}$. The defender applies the ε-greedy policy to determine its detection interval x^k based on the state s^k.

Let $Q(s, x)$ denote the quality function of the scan interval x and state s. $V(s)$ is the value function of state s. The defender updates its Q-function based on its utility u_D and the value

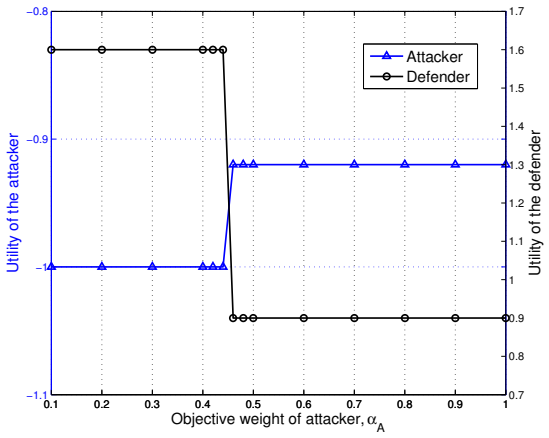


Fig. 2. Performance of the subjective APT game with uncertain attack durations G, with $C = 0.62$, $G = 0.6$, $\alpha_D = 1$ and $L = 2$.

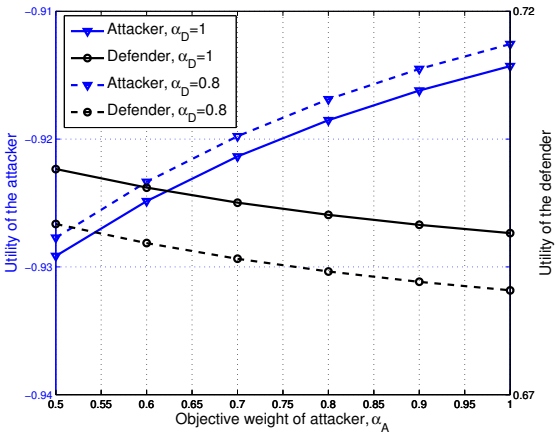


Fig. 3. Performance of the subjective APT game with unknown action of the opponent G', with $C = 0.5$, $G = 0.1$ and $z = 0.2$.

function as follows:

$$Q(s^k, x^k) \leftarrow (1-\gamma)Q(s^k, x^k) + \gamma\left(u_D\left(s^k, x^k\right) + \delta V(s^{k+1})\right) \quad (32)$$

$$V(s^k) = \max_{x \in \mathbf{x}} Q(s^k, x), \quad (33)$$

where $\delta \in [0, 1]$ is the discount factor in the learning, $\gamma \in (0, 1]$ is the learning rate of the defender. The algorithm is summarized in Algorithm 1.

VII. Simulation Results

Simulations were performed to evaluate the subjectivity of the attacker with $C = 0.62$, $G = 0.6$, $\alpha_D = 1$ and $L = 2$ in the APT game $\mathbb{G}$ and $C = 0.5$, $G = 0.1$ and $z = 0.2$ in the game $\mathbb{G}'$. As shown in Fig. 2, the utility of the attacker has a sudden increase with its objective weight while that of the defender decreases. For instance, the utility of the attacker increases sharply from -1 to -0.92, as α_A changes at around 0.42. The reason is that a subjective APT attacker tends to overweigh its loss of being detected by the defender and thus reduces its attack rate if its subjective level is high. As shown in Fig. 3, the subjectivity of the attacker reduces its own utility and improves the defender's utility. For example, if the objective weight of the attacker increases from 0.5 to 1 against an objective defender, its utility increases from -0.929 to -0.914. In addition, the defender's utility increases with its objective weight while the attacker's utility decreases. For instance, if a defender whose objective weight increases from 0.8 to 1 detects an objective attacker, the utility of the defender increases by 1.1%, because an objective defender detects the APT attack more frequently to suppress attack motivation.

Fig. 4 presents the performance of the proposed APT defense scheme, in which the defender chooses its scan interval based on Q-learning, with $L = 5$, $C = 0.4$, $G = 0.6$, $\alpha_A = 0.8$ and $\alpha_D = 1$. As shown in Fig. 4 (a), the attack rate decreases during learning process, e.g., it decreases from 32% to 12% after 2500 time slots since the start of the game, which is 62.5% lower than that of the random detection strategy. As shown in Fig. 4 (b), the utility of the defender increases over time and converges to a high value, e.g., it increases by 7.7% and converges to 1.4 after 2500 time slots. The reason is that the defender learns the total attack durations and adjusts its scan interval via trials.

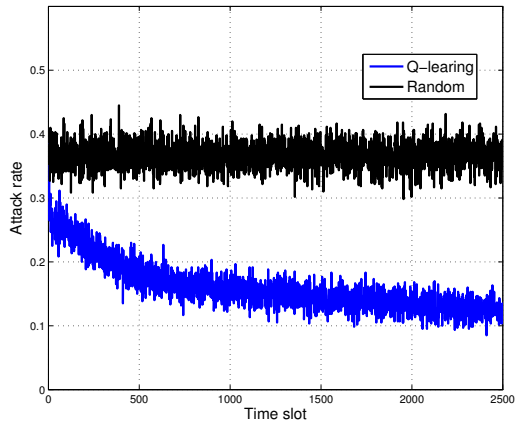
As shown in Fig. 5 (a), the average attack rate increases with the objective weight of the attacker, e.g., it increases from 0.05 to 0.35, if α_A changes from 0.2 to 1. The reason is that a subjective attacker tends to overweigh the attack durations and thus reduces its attack frequency. Compared with the random strategy, the proposed scheme decreases the average attack rate by 30% with $\alpha_A = 1$. Consequently, as shown in Fig. 5 (b), the average utility of the storage defender decreases from 1.43 to 1.28, if α_A changes from 0.2 to 1.

VIII. Conclusion

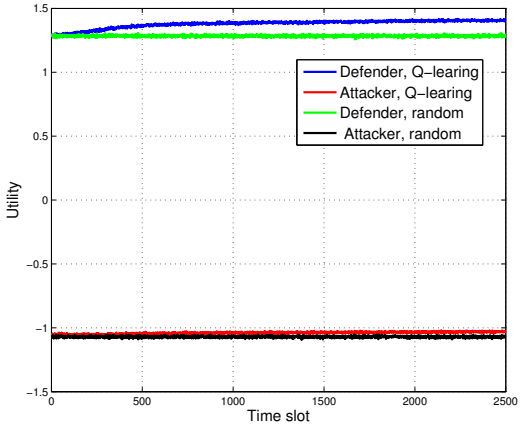
In this work, we have formulated two subjective APT static games for cloud storage, in which the storage defender determines its interval to scan the storage device, while the subjective attacker chooses its interval to launch APT under uncertain attack durations in the pure-strategy game, or unknown scan interval in the mixed-strategy game. We have derived the NEs of the two static games and provided conditions under which the NEs exist, showing that a subjective attacker tends to overweigh its attack cost and thus increases its attack interval, improving the utility of the defender. We have also investigated a dynamic game and proposed a Q-learning based APT defense scheme. Simulation results show that the proposed scheme can improve the performance of the APT game, e.g., the utility of the defender increases by 7.7% and the attack rate decreases by 62.5%, compared with the benchmark strategy.

References

[1] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," *IEEE Trans. Parallel and Distributed Systems*, vol. 25, no. 1, pp. 222–233, 2014.

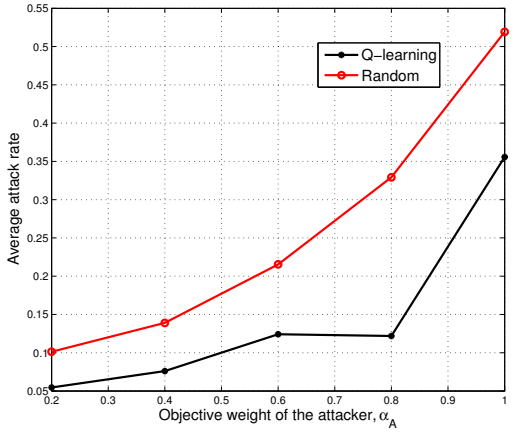


(a) Attack rate

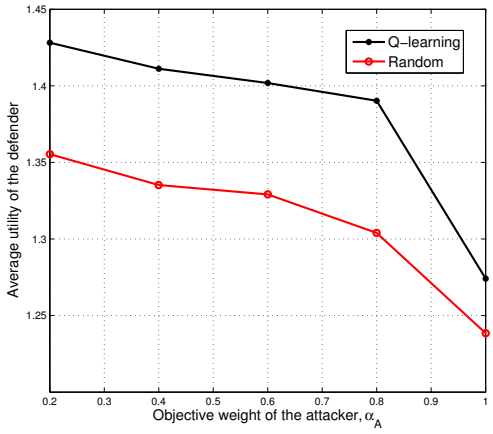


(b) Utility

Fig. 4. Performance of the dynamic APT game averaged over 1000 runs, with $L = 5$, $C = 0.4$, $G = 0.6$, $\alpha_A = 0.8$ and $\alpha_D = 1$.



(a) Average attack rate



(b) Average utility of defender

Fig. 5. Performance of the dynamic APT game averaged over 2500 time slots, with $L = 5$, $C = 0.4$ and $G = 0.6$.

[2] E. Cole, *Advanced Persistent Threat: Understanding the Danger and How to Protect Your Organization*. Syngress, 2013.

[3] M. van Dijk, A. Juels, A. Oprea, and R. L. Rivest, "Flipit: The game of stealthy takeover," *J. Cryptology*, vol. 26, no. 4, pp. 655–713, 2013.

[4] A. Tversky and D. Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *J. Risk Uncertainty*, vol. 5, no. 4, pp. 297–323, 1992.

[5] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica*, pp. 263–291, 1979.

[6] J. Pawlick, S. Farhang, and Q. Zhu, "Flip the cloud: Cyber-physical signaling games in the presence of advanced persistent threats," in *Decision and Game Theory for Security*, pp. 289–308, Springer, 2015.

[7] M. Zhang, Z. Zheng, and N. B. Shroff, "A game theoretic model for defending against stealthy attacks with limited resources," in *Decision and Game Theory for Security*, pp. 93–112, Springer, 2015.

[8] M. Zhang, Z. Zheng, and N. B. Shroff, "Stealthy attacks and observable defenses: A game theoretic model under strict resource constraints," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, pp. 813–817, 2014.

[9] P. Hu, H. Li, H. Fu, D. Cansever, and P. Mohapatra, "Dynamic defense strategy against advanced persistent threat with insiders," in *Proc. IEEE Conf. Comput. Commun.*, pp. 747–755, 2015.

[10] T. Li and N. B. Mandayam, "Prospects in a wireless random access game," in *Proc. 46th Annu. Conf. Inf. Sci. Syst.*, pp. 1–6, Mar. 2012.

[11] T. Li and N. B. Mandayam, "When users interfere with protocols: Prospect theory in wireless networks using random access and data pricing as an example," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 1888 – 1907, April 2014.

[12] J. Yu, M. H. Cheung, and J. Huang, "Spectrum investment with uncertainty based on prospect theory," in *Proc. IEEE Int. Conf. Commun.*, pp. 1620 –1625, 2014.

[13] L. Xiao, N. B. Mandayam, and H. V. Poor, "Prospect theoretic analysis of energy exchange among microgrids," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 63–72, 2015.

[14] Y. Yang, L. Park, N. B. Mandayam, I. Seskar, A. Glass, and N. Sinha, "Prospect pricing in cognitive radio networks," *IEEE Trans. Cognitive Commun. Netw.*, vol. 1, no. 1, pp. 56–70, 2015.

[15] L. Xiao, J. Liu, Q. Li, N. B. Mandayam, and H. V. Poor, "User-centric view of jamming games in cognitive radio networks," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 12, pp. 2578–2590, 2015.

[16] D. Prelec, "The probability weighting function," *Econometrica*, pp. 497–528, 1998.

[17] C. J. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3-4, pp. 279–292, 1992.