Time-Dependent Strategies in Games of Timing

Jonathan Merlevede[1 [0000-0001-5919-731X]], Benjamin Johnson[2 [0000-0001-8832-9457]], Jens Grossklags[2 (✉) [0000-0003-1093-1282]], and Tom Holvoet[1 [0000-0003-1304-3467]]

[1] imec-DistriNet, KU Leuven
jonathan.merlevede@cs.kuleuven.be
tom.holvoet@cs.kuleuven.be
[2] Chair for Cyber Trust, Technical University of Munich
benjamin.johnson@tum.de
jens.grossklags@in.tum.de

Abstract. Timing, a central aspect of decision-making in security scenarios, is a subject of growing academic interest; frequently in the context of stealthy attacks, or advanced persistent threats (APTs). A key model in this research landscape is FlipIt [1]. However, a limiting simplifying assumption in the FlipIt literature is that costs and gains are not subject to discounting, which contradicts the typical treatment of decision-making over time in most economically relevant contexts.
Our recent work [2] introduces an adaptation of the FlipIt model that applies time-based exponential discounting to the value of a protected resource, while allowing players to choose from among the same canonical strategies as in the original game. This paper extends the study of games of timing by introducing two new classes of strategies that are fundamentally motivated by a time-discounted world view.
Within our game model, we compute player utilities, best responses and give a partial characterization of the game's Nash equilibria. Our model allows us to re-interpret the APT model using a finite total valuation, and a finite time horizon. By applying time-based discounting to the entire decision-making framework, we increase the level of realism as well as applicability to organizational security management.

1 Introduction

Defense against *stealthy* advanced persistent threats (APTs) through perfectly effective preventative investments is an impossible goal in most contexts. In fact, data about reported security incidents reveal that organizations need on average about 200 days to merely detect successful attacks [3]. As such, additional emphasis needs to be placed on the optimization of mitigation strategies against stealthy threats such as the scheduling of investigative in-depth security audits.

To make strategic decisions about the mitigation of stealthy attacks, a central consideration must be the notion of *time*, which has been the focus of the *games of timing* literature stemming from the cold war period (see, for example, [4]). This research field received a new influx of work on the so-called FlipIt game beginning with research by van Dijk *et al.* [1] focused on the competitive dynamics to control a contested resource in a limited information environment.

While the majority of these studies assume that players are indifferent between costs and gains now in comparison to those in the (distant) future, our recent work [2] has begun to apply notions of time-based discounting to the game of FlipIt. These initial efforts have been focused on two canonical classes of timing strategies, the so-called periodic and exponential strategies (introduced in [1]), for which the expected time between strategic actions is constant. In a discounted environment, constant expected time between actions implies a decreasing expected valuation between actions, which calls into question the rational appropriateness of this class of strategies in a discounted setting.

In this work, we conduct game-theoretic analysis of infinite timing games with time-based discounting, using two new classes of strategies (discounted periodic and discounted exponential), constructed so that the strategic timing is aligned with the discounted resource valuation. We determine player utilities for all combinations of strategies within the same class, and provide numerical illustrations for each player's best response strategy, concluding with a partial characterization of the game's Nash equilibria.

Our results differ in several aspects from those of non-discounted games of timing as well as discounted games of timing in which only non-discounted canonical strategies are considered. For example, enacting strategies from the revised classes always results in a finite total number of actions; and the cumulative effect of discounting the cost of action at lower rates is limited to at most doubling the total cost, implying that costs and resources may be time-discounted at substantially different rates without affecting the structure of results.

2 Related work

Our discussion of related work focuses on games of timing, and specifically on research designed to capture key aspects of stealthy APT attacks, related to the FlipIt game [1], [5]. This literature has grown considerably, such that there now exist many adapted and extended versions of the original game. The exponential discounting extension in our previous work [2] represents one such adaptation; and serves as the primary motivation for the current analysis. In the following, we briefly review additional literature.

Laszka *et al.* [6], [7] have investigated the influence of including non-targeted attackers in the FlipIt model. Feng *et al.* [8] and Hu *et al.* [9] modified the game by considering insider threat actors. Feng *et al.* [8] accomplish this by adding a third player, an insider, to the model. The insider derives gains from the resource, when it is under control of the defender, by selling information to the attacker who will learn about ways to decrease the cost of attacks.

In the basic FlipIt game, moves by both the attacker and defender are assumed to be instantaneous and always successful. Farhang and Grossklags [10] introduce the idea of imperfect defensive moves with a quality level $\alpha \in [0, 1]$ that expresses the fraction of the resource that remains under the control of the attacker after a flip by the defender. Zhang *et al.* [11] and Laszka *et al.* [6], [7] capture the realistic notion that attacks are complex and take a random amount of time before taking effect. Johnson *et al.* [12] redefine the probability of success

of an attack as a function of time. They also consider that the cost of flipping may be time-dependent.

In the basic FlipIt game, the game has an infinite time horizon and players compete for the resource forever. Zhang *et al.* [11] and Johnson *et al.* [12] assume that the game ends at a fixed pre-defined point in time. Pham and Cid [13] propose a variation of the FlipIt game in which each action makes it more costly for the opponent to take over the resource again; effectively reducing the game to a finite version.

Laszka *et al.* [14] consider two ways of composing resources: one where the attacker receives gain when she is in control of at least one resource (OR-model) and one where she receives gain only when in control of all of the resources (AND-model). Leslie *et al.* [15] generalize this to a model where the attacker has to compromise a threshold fraction of the defender's resources before receiving any gain. Zhang *et al.* [11] also consider multiple resources, but model no interaction between them except through a resource constraint imposed on players in the form of a maximum play frequency that is shared across resources.

Much of the follow-up work on FlipIt has made changes to the assumption of perfect stealthiness. Often the defender is assumed to be completely overt [6], [7], [10], [11]. Besides the conceptual difference, overtness also allows for a different characterization of the FlipIt game as a convex optimization problem [11]. Pham and Cid [13] add a new audit action to the game, which allows a player to query the current owner of the resource. The insider player introduced by Hu *et al.* [9] is at risk of being caught when selling information, which is integrated into her utility function.

Johnson *et al.* [12] consider a discretized version of a timing game similar to FlipIt, in which players are only allowed to make decisions at discrete points in time. Discretization of time is especially relevant for defender moves, which often have to be performed according to some schedule so as not to interrupt business operations (e. g. only at night) [16]. Zhang *et al.* [11] impose budget constraints on players that limit the maximum flip frequency, a practical consideration that is ignored in other treatments of FlipIt. Pawlick *et al.* [17] define a meta-game that consists of a signalling game and a FlipIt game. The parameters of the FlipIt game are defined by the outcome of the signalling game and vice versa.

Beyond our previous research (Merlevede *et al.* [2]), we are unaware of any studies that investigate the impact of discounting in FlipIt game models.

3 Model definition

In this section, we introduce our model for stealthy timing-based security games with discounted costs, discounted resource valuations, and discounting-inspired strategies. In this two-player game, a defender (D) and an attacker (A) are vying for control over a central resource. To obtain control, either player $i \in \{A, D\}$ can choose to pay a fixed instantaneous cost c_i to immediately assume control of the resource. The resource is always controlled exclusively by the last player to execute such a move.

The controlling player accrues utility at a rate which decreases exponentially over time. The cost to execute a move is also time-discounted, albeit at a potentially-different discount rate. Control is always stealthy in the sense that neither player knows who controls the resource until the moment that they initiate an instantaneous 'flip'. The remainder of this section formalizes the game.

3.1 Player strategies

For player $i \in \{D, A\}$, define

$$\boldsymbol{t}_i = (t_{i,0},\ t_{i,1},\ t_{i,2},\ \ldots)$$

to be a strictly increasing sequence of times at which player i moves. (A player can move at most once in a particular instance of time.) The length of $\boldsymbol{t}_i$ can be finite or infinite. A player strategy in this game is defined completely by a probability distribution over a set of possible $\boldsymbol{t}_i$.

3.2 Player control

An outcome of the game is a pair of move sequences $(\boldsymbol{t}_D, \boldsymbol{t}_A)$. Times that occur in both vectors complicate a smooth analysis, so for this and other reasons[3], we assume that $\boldsymbol{t}_D \cap \boldsymbol{t}_A = \varnothing$.

Let

$$\boldsymbol{t} = \boldsymbol{t}_D \cup \boldsymbol{t}_A = (t_0,\ t_1,\ t_2,\ \ldots)$$

be the strictly increasing sequence of player move times. Then for any time $t \geq t_0$, we may define the *latest flip time* function by

$$\mathrm{LFT}(t) = \max\{t_k \in \boldsymbol{t} : t_k \leq t\}.$$

From time $t = 0$ until the time of the first flip t_0, the defender has control of the resource. We may thus define the player control function for any time $t > 0$ by

$$\mathrm{PC}\colon t \mapsto \begin{cases} D & \text{if } t < t_0 \text{ or } \mathrm{LFT}(t) \in \boldsymbol{t}_D \\ A & \text{if } \mathrm{LFT}(t) \in \boldsymbol{t}_A. \end{cases}$$

The asymmetry of the player control function shows that the defender has an advantage due to starting off the game in control of the resource.

We also define a player control indicator function $\mathrm{PC}_i(t) = \mathbf{1}_{\mathrm{PC}(t)=i}$, which can be used for integrating. This function tells us, for a given player $i \in \{D, A\}$ and time $t > 0$, whether that player controls the resource at that time.

[3] For each of the strategy distributions that we analyze, the set of all outcomes with non-disjoint strategy vectors has probability zero, so in our setting this assumption is benign. We could also address overlapping sequences by defining the resource control function so that any occurrence of simultaneous moves leaves control of the resource unaffected; and so doing would accomplish the same effect as our assumption, but with a more complicated logical underpinning.

3.3 Player gains

Players achieve gains by controlling the resource. Gains initially accrue value at some rate of V dollars per unit of time. Meanwhile, the resource decreases in value over time at a discount rate of ρ. The gain for player i may be determined by computing the expected exponentially weighted integral of $\mathrm{PC}_i(t)$ over all of time, normalized with respect to the total (discounted) value of the resource:

$$\mathscr{G}_i = \frac{\mathbb{E}\left[\int_{\tau=0}^{+\infty} \mathrm{PC}_i(\tau) V \mathrm{e}^{-\rho\tau}\,\mathrm{d}\tau\right]}{\int_{\tau=0}^{+\infty} V\mathrm{e}^{-\rho\tau}\,\mathrm{d}\tau} = \rho \cdot \mathbb{E}\left[\int_{\tau=0}^{+\infty} \mathrm{PC}_i(\tau)\mathrm{e}^{-\rho\tau}\,\mathrm{d}\tau\right]. \tag{1}$$

The expectation is taken over the distributions involved in defining the player strategies. Normalization allows comparing player gains for different discount rates and the interpretation of total gain as a fraction of total achievable gain.

3.4 Player costs

When players perform a move, this comes at a fixed *instantaneous cost* of $c_i > 0$. Costs are exponentially discounted with discount rate ${}_c\rho_i$, which can be different from the discount rate for the resource, ρ.

A player's *(total) costs* are defined as the expected sum of exponentially discounted instantaneous costs, normalized with respect to the total (discounted) value of the resource:

$$\mathscr{C}_i = \frac{\mathbb{E}\left[\sum_{\tau\in\boldsymbol{t}_i} c_i \mathrm{e}^{-\,{}_c\rho_i\tau}\right]}{\int_{\tau=0}^{+\infty} V\mathrm{e}^{-\rho\tau}\,\mathrm{d}\tau} = \frac{\rho}{V} \cdot \mathbb{E}\left[\sum_{\tau\in\boldsymbol{t}_i} c_i \mathrm{e}^{-\,{}_c\rho_i\tau}\right]. \tag{2}$$

As with gains, the expectation is taken with respect to the distribution used to define $\boldsymbol{t}_i$. Scaling gains and costs by the same factor makes the normalization operation neutral with respect to the behavior of rational players. Finally, since we only deal with normalized costs and gains and since c_i and V are both free parameters, we can without loss of generality assume that

$$V = 1.$$

This assumption allows us to think of the instantaneous cost $c_i = \frac{c_i}{V}$ as a unitless value, expressing a fraction of the initial rate at which the resource accrues value per unit of time.

3.5 Player utilities

Player utilities are equal to the difference of player gains and player costs:

$$u_i = \mathscr{G}_i - \mathscr{C}_i.$$

3.6 Discounted strategies

Because the full strategy space for this game involves probability distributions over countable sequences of real numbers, a desire for a useful analysis requires us to restrict our attention to "reasonable" sub-classes of such strategies. A review of research into timing games has identified two canonical classes of strategies for this purpose, the class of periodic strategies, in which the time between moves is constant, (with the first move being randomized), and the class of exponential strategies, for which the time between moves is exponentially distributed, (as is the time of the first move).

In this section, we present adaptations of these two classes of strategies that are specifically-motivated by our game's time-based discounting factor ρ: instead of defining strategies in terms of inter-arrival times, we define them in terms of the value generated by the resource between subsequent moves. To do this, we first introduce the concept of a compressed timeline, where the compression is adapted to the rate of exponential discounting of the resource.

Compressed timeline To begin, we define a transformation $\mathscr{T}$ that maps time $t \in [0, \infty)$ onto a compressed time $x \in [0, 1)$, such that the total value of the resource up until t equals x:

$$\mathscr{T} : \begin{cases} [0, +\infty) \to [0, 1) \\ t \mapsto \rho \cdot \displaystyle\int_{\tau=0}^{t} \mathrm{e}^{-\rho\tau} \, \mathrm{d}\tau = 1 - \mathrm{e}^{-\rho \cdot t}. \end{cases}$$

We can map compressed time x onto *real times* t using the inverse transformation $\mathscr{T}^{-1}$:

$$\mathscr{T}^{-1} : \begin{cases} [0, 1) \to [0, +\infty) \\ x \mapsto -\dfrac{\ln(1-x)}{\rho}. \end{cases}$$

Discounted exponential strategy A *discounted exponential strategy* is a strategy for which all compressed inter-arrival times, as well as the time of the first move, are drawn from the same exponential distribution:

$$\begin{aligned} \Delta_{i,n} &\sim \mathrm{Exp}(\nu_i) \text{ and} \\ x_{i,0} &\sim \mathrm{Exp}(\nu_i). \end{aligned}$$

Here $\Delta_{i,n} = x_{i,n+1} - x_{i,n}$ is the difference between the timing of the nth and $(n+1)$th move as observed on a compressed timeline. Parameter ν is the flip rate, move rate or *play rate* of the discounted exponential strategy. The expected compressed time between two moves is constant and equal to $1/\nu$. The expected real time between moves is therefore time-dependent and increasing. The expected total number of moves is finite and equal to ν.

Discounted periodic strategy A *discounted periodic strategy* is a strategy where the compressed inter-arrival times are equal to a constant value δ. We refer to δ as the *period* of a discounted periodic strategy and to the inverse of the period, $\nu = 1/\delta$, as a strategy's *play rate*. Discounted periodic strategies with random phase are those periodic strategies where the compressed time of the first move or *phase* φ is drawn from the positive values smaller than δ:

$$\begin{aligned} \Delta_{i,n} &= \delta_i \text{ and} \\ x_{i,0} &= \varphi_i \sim U[0, \delta_i]. \end{aligned}$$

If δ is chosen to be greater than one, and the randomly-drawn phase also turns out to be greater than one, the realized move time is not on the compressed timeline, implying that the player never moves. Constant compressed inter-arrival times imply increasing real inter-arrival times. The total number of flips that a player performs when executing a discounted periodic strategy is always between $\lfloor \nu_i \rfloor$ and $\lfloor \nu_i \rfloor + 1$. See Fig. 1 for an illustration.

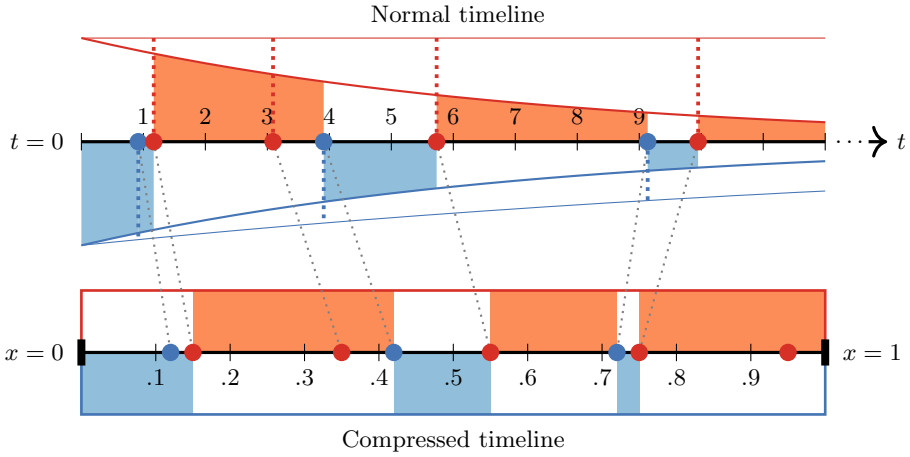


Fig. 1: A game outcome in which each player uses a discounted periodic strategy. Dots (●, ●) indicate defender and attacker moves. Shaded areas (■, ■) are proportional to the gain obtained. Thick dotted lines (┈┈, ┈┈) are proportional to the cost of moving.

Interpretation Fig. 1 provides an illustration for the discounted periodic strategy. On the normal timeline, moves occur less and less frequently over time, which makes strategic sense because the resource is valued less and less over time. While the strategy has a simple periodic representation only on the compressed timeline, the value generated by the resource between two moves (the area under the curve) is constant on both timelines. In moving to a discounting-aware strategy, we add some complexity to our automation processes because the strategy must be implemented in normal time; but we gain in exchange an improved alignment with our valuation.

4 Analysis

4.1 Player gains

We begin our analysis by determining the player gains in our model.

Anonymous gains We first note that due to the normalization, the total value of the resource is exactly 1. Since at any point in time one of the two players is gaining revenue from the resource, the sum of the gains of the two players must also equal to one.

Let us define the *anonymous gain of player i* to be the (total expected) gain for that player after the first flip (by either player) has occurred. We use the word anonymous because, for both (discounted) exponential and periodic strategies, the calculation of this quantity is symmetric with respect to the two player identities (A and D). We sometimes refer to the *total anonymous gain* by which we mean the sum of the anonymous gains of the two players.

Note that the anonymous gain of the attacker is the total expected gain of the attacker, because if any value accrues at all for the attacker, it does so after his first flip. Since the gain of the defender is one minus the gain of the attacker, we may easily convert expressions involving the anonymous gains into expressions for player gains. This is useful because it allows us to express player gains in a uniform notation even though the structure of gains is identity-dependent.

Compressed Time Note that the time compression used in the presentation of our discounted strategies was defined so as to preserve player gains over time. (For example, the areas of each shaded region in Fig. 1 are matched across the timelines.) Because the strategies are defined in a compressed timeline, it is easier to compute the gains by integrating over compressed time. To do this, we need to first apply the appropriate transformations so that an integration over the player control function in normal time can be performed on the compressed coordinate system. To accomplish this, let $\widetilde{\mathrm{PC}}_i := \mathrm{PC}_i \circ \mathscr{T}^{-1}$. We can then write:

$$\begin{aligned}
\mathscr{G}_i =& \mathbb{E}\left[\lim_{\overline{T}\to+\infty} \rho \cdot \int_{t=0}^{\overline{T}} \mathrm{PC}_i(t) \cdot \mathrm{e}^{-\rho\tau}\,\mathrm{d}t\right] \\
=& \mathbb{E}\left[\lim_{\overline{T}\to+\infty} \rho \cdot \int_{x=\mathscr{T}(0)}^{\mathscr{T}(\overline{T})} \mathrm{PC}_i(\mathscr{T}^{-1}(x)) \cdot \mathrm{e}^{-\rho\cdot\mathscr{T}^{-1}(x)} \cdot \frac{\mathrm{d}t}{\mathrm{d}x}\,\mathrm{d}x\right] \\
=& \mathbb{E}\left[\rho \cdot \int_{x=0}^{1} \widetilde{\mathrm{PC}}_i(x) \cdot \mathrm{e}^{\ln(1-x)} \cdot \frac{1}{(1-x)\rho}\,\mathrm{d}x\right] \\
=& \mathbb{E}\left[\int_{x=0}^{1} \widetilde{\mathrm{PC}}_i(x)\,\mathrm{d}x\right].
\end{aligned}$$

Exponential play Now, we assume that both players are playing a discounted exponential strategy, using play rates ν_D and ν_A.

Lemma 1. *For exponential play, each player i obtains a fraction of the total anonymous gain equal to:*

$$\frac{\nu_i}{\nu_i + \nu_j}.$$

Proof. At any moment in time, the probability that player i is the next player to move is equal to

$$p_i = \int_{\tau_i=0}^{+\infty} \nu_i \mathrm{e}^{-\nu_i \tau_i} \int_{\tau_j=\tau_i}^{+\infty} \nu_j \mathrm{e}^{-\nu_i \tau_j} \, \mathrm{d}\tau_j \, \mathrm{d}\tau_i = \frac{\nu_i}{\nu_i + \nu_j}.$$

Probability p_i is, therefore, also equal to the probability that any flip by either player after time t_0 is made by player i.

Consider the set of all intervals between flips together with the interval between the last flip and the end of the game. For each such interval, with probability p_i player i is the player who receives gain over the entire interval; and with probability $1 - p_i$ she receives nothing. Her expected gain over each interval is, therefore, p_i times the length of the interval. By linearity of expectation, her total expected gain over this set of intervals is therefore p_i times the total combined duration of the intervals.

□

Lemma 2. *If $\nu_i + \nu_j = 0$, then the total anonymous gain is zero. Otherwise, the expected total anonymous gain for exponential play is:*

$$1 - \frac{1 - \mathrm{e}^{-(\nu_i + \nu_j)}}{\nu_i + \nu_j}. \tag{3}$$

Proof. Let X_i be the time of i's first flip, and define random variable

$$Z = \min\{X_i, X_j\}$$

as the time until the first flip by either player. Then

$$\begin{aligned} F_Z(z) &= \Pr[Z \leq z] = \Pr[X_i \leq z \text{ or } X_j \leq z] \\ &= 1 - \Pr[X_i \geq z \text{ and } X_j \geq z] \\ &= 1 - \Pr[X_i \geq z] \cdot \Pr[X_j \geq z] \\ &= 1 - (1 - (1 - \mathrm{e}^{-\nu_i z}))(1 - (1 - \mathrm{e}^{-\nu_j z})) \\ &= 1 - \mathrm{e}^{-(\nu_i + \nu_j) z}, \end{aligned}$$

that is, Z is distributed exponentially with rate parameter $\nu_i + \nu_j$.

We can then express the expected total anonymous gain as

$$(\nu_i + \nu_j) \cdot \int_{z=0}^{1} \mathrm{e}^{-(\nu_i + \nu_j) z} \int_{\tau=z}^{1} \mathrm{d}\tau \, \mathrm{d}z = (\nu_i + \nu_j) \cdot \int_{z=0}^{1} \mathrm{e}^{-(\nu_i + \nu_j) z} (1 - z) \, \mathrm{d}z,$$

which evaluates to Eq. (3). □

Note that the expected total anonymous gain never quite reaches one. There are two reasons:

- The expected gain before t_0 is not part of the anonymous gain.
- There is a probability of $\mathrm{e}^{-\nu_i} + \mathrm{e}^{-\nu_j} - \mathrm{e}^{-(\nu_i+\nu_j)}$ that neither player ever flips.

An expression for the anonymous gain of player i now follows easily from Lemmas 1 and 2.

Lemma 3 (Anonymous gain for discounted exponential play). *Player i's anonymous gain for discounted exponential play is:*

$$\overline{\mathscr{G}}_i = \frac{\nu_i}{\nu_i + \nu_j} - \nu_i \cdot \frac{1 - \mathrm{e}^{-(\nu_i+\nu_j)}}{(\nu_i + \nu_j)^2}.$$

Periodic play Next, we assume that both players are playing a discounted periodic strategy using play rates ν_D and ν_A.

Lemma 4 (Anonymous gain for discounted periodic play). *Player i's anonymous gain for discounted periodic play is:*

$$\overline{\mathscr{G}}_i = \begin{cases} 1 - \frac{1+\nu_j}{2\nu_i} + \frac{\nu_j}{3\nu_i^2} & \text{if } \nu_i \geq 1 \text{ and } \nu_i \geq \nu_j, \\ \frac{\nu_i}{2} - \frac{\nu_i \nu_j}{6} & \text{if } \nu_i \leq 1 \text{ and } \nu_j \leq 1, \text{ and} \\ \frac{\nu_i}{2\nu_j} - \frac{\nu_i}{6\nu_j^2} & \text{otherwise.} \end{cases} \tag{4}$$

Proof outline. By linearity of expectation, player i's total anonymous gain equals the sum of the expected gains over the following intervals:

- The time before player i's first flip.
- The time after player i's last flip.
- The time in between.

Player i's expected anonymous gain before her first flip is always zero. Appendix A lists derivations for the expected gains over the other two intervals.

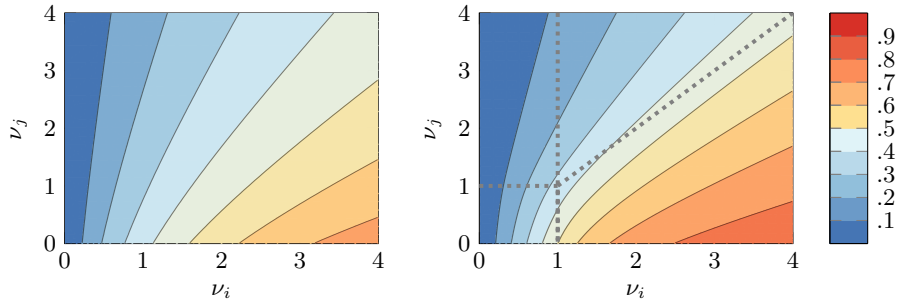
Illustrations of anonymous gain Fig. 2 shows a contour plot of the anonymous gain as a function of the players' play rates. Note that the anonymous gain is monotone in both play rates. For both types of strategy configurations, the anonymous gain of player i is increasing in player i's own play rate, and decreasing in player j's play rate.

4.2 Player costs

Next, we determine the total player costs in our model.

Lemma 5. *The cost of performing an exponential or periodic strategy with rate parameter ν_i is*

$$\mathscr{C}_i = \frac{c_i \cdot \nu_i \cdot \rho^2}{\rho + {}_c\rho_i}.$$



(a) Exponential play (b) Periodic play

Fig. 2: Contour plot of the anonymous gain of player i for periodic and exponential play. The dotted line (⋯) illustrates where the different cases of Eq. (4) apply.

Proof. For both the periodic and the exponential strategy, the compressed probability density of flipping at any specific moment in time on the compressed timeline is constant and equal to ν_i. With respect to real time, the probability density of a move at time t is therefore:

$$\nu_i \frac{\mathrm{d}x}{\mathrm{d}t} = \nu_i \frac{\mathrm{d}\mathscr{T}(t)}{\mathrm{d}t} = \nu_i \frac{\mathrm{d}}{\mathrm{d}t}\left(1 - \mathrm{e}^{-\rho t}\right) = \nu_i \cdot \rho \cdot \mathrm{e}^{-\rho t}.$$

The discounted instantaneous cost of performing a flip at time t is $c_i \cdot \mathrm{e}^{-\,{}_c\rho_i t}$. The total cost of all flips becomes:

$$\rho \cdot \int_{t=0}^{+\infty} (\nu_i \cdot \rho \cdot \mathrm{e}^{-\rho t}) \cdot (c_i \cdot \mathrm{e}^{-\,{}_c\rho_i t})\,\mathrm{d}t = \frac{c_i \cdot \nu_i \cdot \rho^2}{\rho + {}_c\rho_i}. \qquad \square$$

4.3 Player utilities

The utility of a player is simply her expected gains minus her expected costs. The costs are provided above. The total gain for the attacker is the same as the anonymous gain, and the total gain for the defender is 1 minus that. Therefore, this section is just an exercise in translating the results from the previous two sections. Here, we provide the explicit formulation of player utilities for exponential play. The expression of utilities for periodic play may be determined similarly, but has been omitted due to space considerations.

Theorem 1 (Utility for discounted exponential play). *Player utilities for discounted exponential play are:*

$$u_A = \frac{\nu_A}{\nu_A + \nu_D} - \nu_A \cdot \frac{1 - \mathrm{e}^{-(\nu_A+\nu_D)}}{(\nu_A + \nu_D)^2} - \frac{c_A \cdot \nu_A \cdot \rho^2}{\rho + {}_c\rho_A}$$
$$u_D = 1 - \left(\frac{\nu_A}{\nu_A + \nu_D} - \nu_A \cdot \frac{1 - \mathrm{e}^{-(\nu_A+\nu_D)}}{(\nu_A + \nu_D)^2}\right) - \frac{c_D \cdot \nu_D \cdot \rho^2}{\rho + {}_c\rho_D}.$$

Proof. This follows from Lemma 3 and Lemma 5 and the fact that the total gain for the attacker and defender are $\overline{\mathscr{G}}_A$ and $1 - \overline{\mathscr{G}}_A$.

4.4 Player incentives

With player utilities in hand, we may now ask what players want to do. For both discounted exponential strategies and discounted periodic strategies, each player i has the choice of selecting one real parameter ν_i. We are especially interested in the behavior of the partial derivative of a player's utility with respect to her own play rate, which, for a player i, may be expressed in standard notation as

$$\frac{\partial u_i}{\partial \nu_i},$$

and to which we refer in the following discussion as the *incentive* of player i.

The following lemmas provide us with information that we need about incentives to determine the best response strategies.

Lemma 6. *For discounted exponential play, each player's incentive is strictly decreasing in her own play rate.*

Lemma 7. *For discounted periodic play, each player's incentive is independent of her own play rate if she is the slower player or if her play rate is smaller than 1. It is strictly decreasing in her own play rate otherwise.*

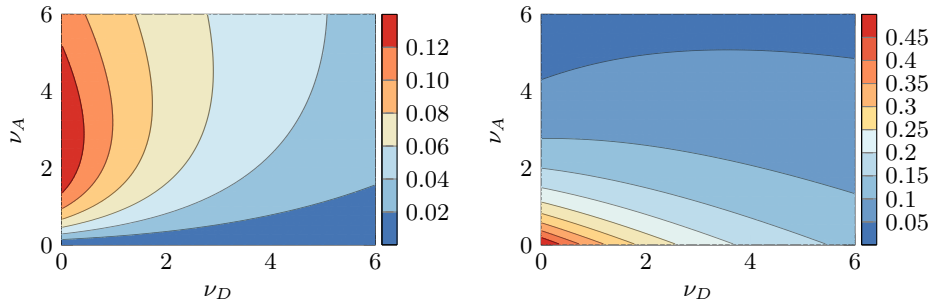
Lemmas 6 and 7 tell us that given an opponent's play rate, a player's incentive is upper-bounded by her incentive when not playing. We will refer to this incentive as her *base incentive*.

Figures 3 and 4 display player incentives with each cost parameter c_i set to zero. With zero costs, increasing play rate ν_i implies increasing utility u_i, so that the incentives are always strictly positive. Increasing the cost c_i decreases the incentive, but only by a constant amount, since the derivative of a player's costs with respect to her own play rate is constant as a function of play rates.

This means that adding cost to the figures will only change the color labels of the graphs. We can verify that ν_i and ν_j are the only variables impacting the rate of change of the incentive with respect to ν_i. Figures 3 and 4, therefore, provide strong support for the claims made in Lemmas 6 and 7.

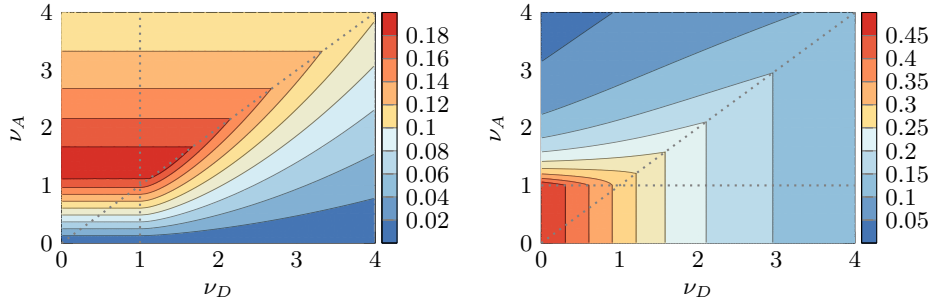
4.5 Player best responses

This subsection characterizes the best-response strategies for the attacker and defender. We begin with a discussion of non-participatory responses and characterize when they are optimal. These results apply equally to both strategy regimes. We then discuss properties of participatory best responses for the exponential and periodic strategy regimes.



(a) Defender incentive (b) Attacker incentive

Fig. 3: Player incentives for exponential play and $c_i = 0$.



(a) Defender incentive (b) Attacker incentive

Fig. 4: Player incentives for periodic play and $c_i = 0$.

Non-participatory best responses The following results apply generally, without restricting players to specific strategy classes.

Lemma 8. *The unique best response by a defender to a non-participatory attacker is not to play.*

Lemma 9. *If the attacker's best response to a non-participatory defender includes not playing, then not playing is a best response to any participatory defender.*

These two lemmas follow directly from the properties of the incentive functions (Lemmas 6 and 7).

Lemma 10. *For exponential and periodic discounted play, not playing is a best response to a non-participatory defender iff*

$$c_A \geq \frac{\rho + {}_c\rho_i}{2\rho^2}.$$

If the inequality is strict, then there cannot be any other best responses.

Proof. We can verify that equation $\lim_{\nu_A \to 0} \lim_{\nu_D \to 0} \frac{\partial u_A}{\partial \nu_A}$ has a single root at $c_A = \frac{\rho +_c \rho_i}{2\rho^2}$ for both exponential and periodic discounted play. The attacker's incentive is strictly non-increasing in her own play rate (Lemmas 6 and 7) and strictly decreasing in c_A. It is, therefore, sufficient to show that the attacker's incentive is non-increasing in ν_D provided that the defender does not move. We do this by showing that the limits of the second order partial derivatives are negative.

For exponential play, we can compute

$$\lim_{\nu_D \to 0} \frac{\partial^2 u_A}{\partial \nu_A^2} = \frac{2\mathrm{e}^{-\nu_A}}{\nu_A^3} \left(1 + \nu_A + \frac{\nu_A^2}{2} - \mathrm{e}^{-\nu_A} \right) = -\frac{2\mathrm{e}^{-\nu_A}}{\nu_A^3} \sum_{i=3}^{+\infty} \frac{x^i}{i!}.$$

For periodic play, the attacker's incentive is independent of her own play rate if $\nu_A < 1$. For $\nu_A > 1$, we can compute

$$\lim_{\nu_D \to 0} \left. \frac{\partial^2 u_A}{\partial \nu_A^2} \right|_{\substack{\nu_D \leq \nu_A \\ \nu_A > 1}} = \frac{-1}{\nu_A^3}. \qquad \square$$

Participatory best-responses Our first two results for participating players are deemed corollaries because they follow immediately from properties of the players' incentive functions (Lemmas 6 and 7), and the definition of a best response.

Corollary 1. *Best-responses for exponential play are single-valued.*

Corollary 2. *Player i's best response for periodic play to an opponent play rate $\bar{\nu}_j$ can be characterized in terms of her base incentive as follows.*

- *If her base incentive is strictly negative, then her unique best response is not to play.*
- *If her base incentive is zero, then moving at any play rate $\alpha_i \in [0, \max\{\bar{\nu}_j, 1\}]$ is a best response.*
- *If her base incentive is strictly positive, then her unique best response is to play at the rate $\alpha_i > \max\{\bar{\alpha}_j, 1\}$ for which her incentive is zero.*

While an algebraic characterization of the best response functions is cumbersome to present, a numerical computation of best responses is straightforward. Moreover, because each player's incentive is non-increasing, each player is playing a best response precisely when her incentive is zero. This gives rise to an alternative interpretation of Figs. 3 and 4 as best-response curves. This interpretation is valid exactly when the value (color) of the incentive function is exactly equal to the effective cost $\frac{\partial \mathscr{C}_i}{\partial \nu_i} = \frac{c_i \cdot \rho^2}{\rho +_c \rho_i}$. (An expression for total costs was provided in Lemma 5.)

4.6 Nash equilibria

Finally, with basic properties of our game's best responses characterized, we may extend this characterization to important properties of the game's Nash equilibria.

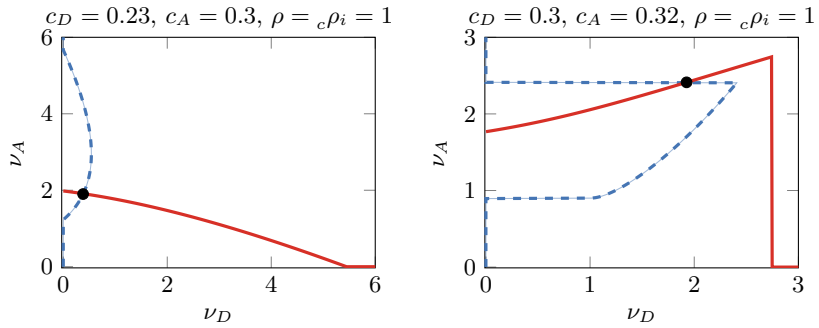
Non-participatory Nash equilibra Our first two results follow directly from the characterization of non-participatory best responses given in Lemmas 8 and 9.

Theorem 2. *There is never a Nash equilibrium in which the defender plays, but the attacker does not.*

Theorem 3. *If the attacker is playing a discounted periodic or exponential strategy, there is a Nash equilibrium in which neither player moves iff*

$$c_A \geq \frac{\rho + {}_c\rho_i}{2\rho^2}.$$

Participatory Nash equilibra Our last result describes the necessary conditions for there to be a Nash equilibrium in the discounted periodic regime. This result follows from the best response characterization for periodic play provided by Corollary 2. Nash equilibria for the exponential and periodic cases are exemplified numerically in Fig. 5.



(a) Exponential play (b) Periodic play

Fig. 5: Defender (- - -) and attacker (—) best response curves and Nash equilibria (•) for exponential and periodic play.

Corollary 3. *In any Nash equilibrium where both players play discounted periodic strategies at non-zero rates, the faster player f plays at a rate α_f that is a root of the slower player's base incentive function. Player s is then indifferent between playing at any rate in $[0, \alpha_f]$.*

5 Discussion

This section discusses results from the previous sections and their practical impacts. Specifically, we consider the total number of player actions (Section 5.1) and the limited impact of time-based discounting of costs (Section 5.2).

5.1 Finite number of player actions

One interesting feature of our model with compressed strategies is that the total number of player actions is always finite. This contrasts with the non-discounted models of APTs in which the total number of enacted actions in an outcome of a strategy is generally infinite.

Our model exhibits this feature because we apply the periodic or exponential paradigm to a time line which has been exponentially compressed. If we think about our resources and expenditures in a discounted sense, it makes sense that we would not want to keep playing forever. At some point the value of the resource will be extremely small, in which case there comes a point in time where further attacks and further expenditures on security are largely pointless. We consider it a useful feature that our modeling framework captures this dynamic.

5.2 Limited impact of cost discounting

Our illustrations of best response strategies involve costs which are discounted at the same rate as the value of the resource. However, the formula for costs (see Lemma 1) is expressed in terms of notation that can apply different cost-discounting rates for each player. If we consider any fixed play rate, the impact of changing the cost-discounting rate from the resource-discounting rate down to zero is to double the total cost.

This effect is substantially different from the regime in which players choose strategies from an exponential or periodic strategy on a non-discounted time line. In evaluating a non discounted strategy, the discount factor for costs could be infinitely important. But for the strategies considered in this paper, varying the cost-discounting factor has an effect more similar to a rounding error.

An implication of this result is that it offers an additional interpretation to time-based discounting that involves a shorter duration of time. Here, we might presume that a resource being protected were discounted not merely because of economic considerations (which also apply to costs), but rather because the very nature of the resource was short-lived. For example, the resource could involve a private key or token, with a fixed duration of validity. Such a resource would become less valuable the closer to its expiration time, although the costs to attack or defend the same resource might not decrease at all in this short time frame. The fact that our model exhibits a relatively small effect for cost discounting (compared to discounting the value of the resource) means that it could be applied in cases where time-based discounting of resource valuation were justified even though time-based discounting of costs were not.

6 Conclusion

The timing of security decisions is an aspect of policy-making that is generally under-appreciated. The overwhelming majority of existing research that does investigate the timing of offensive and defensive actions does not consider how the passing of time can affect the value of the resource. Of the very small number of studies that do consider this aspect of timing, each allows players to choose strategies from classes that are only well-motivated in an environment in which security artifacts retain the same value over time. In this paper, we consider the full gamut of time-dependent considerations for making strategic security-based investment decisions for stealthy resources. Our costs and resources are valued in a time-dependent manner. The individual strategies employed by our players involve making investments over time; and the strategy spaces from which players may choose are motivated by a time-discounted worldview.

Discounting the value of a resource and the costs of defending it already have important implications for interpreting the security landscape involving persistent threats. When applying exponential discounting to a resource and its defense costs, its total value over time becomes finite, as does the total cost of implementing a given strategy. This fact already provides significantly more realism over less time-sensitive models, because the costs and valuation for any real world security decision is of non-infinite magnitude. The time-discounted regime also allows for the possibility of achieving perfect security by raising the costs of an attack – something that would not be possible if the resource were considered to have infinite value. This important consequence of exponential discounting was discussed extensively in our previous work [2] and also holds true in the regime of restricted strategies used in this paper.

When we focus our attention on revised canonical strategies that are motivated by a discounted time horizon, we move further toward reality. More than simply having a finite valuation for our resource, we now consider only reasonable strategies that exhibit a finite number of actions to attack or defend it. This finite time window offers a simpler framework for understanding good responses, which can be incredibly useful in communicating policy decisions.

With this work, we advance the study of time-based aspects of security decisions; but it still has a long way to go. Future work might further extend time-based discounting to methods beyond exponential discounting, and additional classes of attack and defense strategies could be useful to consider. In the meantime, advanced persistent security threats will remain, tempered only by the certainty that the value of every protected resource is bounded; the time horizon for each new threat is finite; and every strategy with an infinite number of plans will cease to be well-motivated long before its completion.

Acknowledgments

We thank the anonymous reviewers for their constructive comments and feedback. This work was partially supported by the German Institute for Trust and Safety on the Internet (DIVSI) and the Research Fund KU Leuven.

References

[1] M. van Dijk, A. Juels, A. Oprea, and R. L. Rivest, "FlipIt: The Game of "Stealthy Takeover"," *Journal of Cryptology*, vol. 26, Oct. 2012.

[2] J. Merlevede, B. Johnson, J. Grossklags, and T. Holvoet, "Exponential Discounting in Security Games of Timing," in *Workshop on the Economics of Information Security (WEIS)*, Jun. 2019.

[3] S. Farhang and J. Grossklags, "When to Invest in Security? Empirical Evidence and a Game-Theoretic Approach for Time-Based Security," in *Workshop on the Economics of Information Security (WEIS)*, Jun. 2017.

[4] T. Radzik, "Results and Problems in Games of Timing," *Lecture Notes-Monograph Series*, 1996.

[5] K. D. Bowers, M. van Dijk, R. Griffin, A. Juels, A. Oprea, R. L. Rivest, and N. Triandopoulos, "Defending against the Unknown Enemy: Applying FlipIt to System Security," in *Decision and Game Theory for Security*, Nov. 2012.

[6] A. Laszka, B. Johnson, and J. Grossklags, "Mitigating Covert Compromises," in *Web and Internet Economics*, Dec. 2013.

[7] A. Laszka, B. Johnson, and J. Grossklags, "Mitigation of Targeted and Non-targeted Covert Attacks as a Timing Game," in *Decision and Game Theory for Security*, Nov. 2013.

[8] X. Feng, Z. Zheng, P. Hu, D. Cansever, and P. Mohapatra, "Stealthy Attacks Meets Insider Threats: A Three-Player Game Model," in *2015 IEEE Military Communications Conference (MILCOM)*, Oct. 2015.

[9] P. Hu, H. Li, H. Fu, D. Cansever, and P. Mohapatra, "Dynamic Defense Strategy against Advanced Persistent Threat with Insiders," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, Apr. 2015.

[10] S. Farhang and J. Grossklags, "FlipLeakage: A Game-Theoretic Approach to Protect Against Stealthy Attackers in the Presence of Information Leakage," in *Decision and Game Theory for Security*, Nov. 2016.

[11] 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*, Nov. 2015.

[12] B. Johnson, A. Laszka, and J. Grossklags, "Games of Timing for Security in Dynamic Environments," in *Decision and Game Theory for Security*, Nov. 2015.

[13] V. Pham and C. Cid, "Are We Compromised? Modelling Security Assessment Games," in *Decision and Game Theory for Security*, Nov. 2012.

[14] A. Laszka, G. Horvath, M. Felegyhazi, and L. Buttyán, "FlipThem: Modeling Targeted Attacks with FlipIt for Multiple Resources," in *Decision and Game Theory for Security*, Nov. 2014.

[15] D. Leslie, C. Sherfield, and N. P. Smart, "Threshold FlipThem: When the Winner Does Not Need to Take All," in *Decision and Game Theory for Security*, Nov. 2015.

[16] S. Rass, S. König, and S. Schauer, "Defending Against Advanced Persistent Threats Using Game-Theory," *PLOS ONE*, vol. 12, Jan. 2017.

[17] 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*, Nov. 2015.

A Anonymous gains for periodic play

Lemma 11. *Player i expects gain after her last move:*

- *(Case 1) If $\delta_i \leq 1$ and $\delta_i \leq \delta_j$, then she expects $\delta_i/2 - \delta_i^2/6\delta_j$.*
- *(Case 2) If $\delta_i \leq 1$ and $\delta_i \geq \delta_j$, then she expects $\delta_j/2 - \delta_j^2/6\delta_i$.*
- *(Case 3) If $\delta_i \geq 1$ and $\delta_j \leq 1$, then she expects $\delta_j/2\delta_i - \delta_j^2/6\delta_i$.*
- *(Case 4) If $\delta_i \geq 1$ and $\delta_j \geq 1$, then she expects $1/2\delta_i - 1/6\delta_i\delta_j$.*

Proof. For every case, we first fix player i's strategy and take expectations over player j' strategy. We then compute player i's expected gain by taking expectations over player i's strategy. Let T be the game time that remains after i's last move. After this move, player i remains in control until a move by j or the end of the game.

(Case 1; $i = f$, $j = s$) Player s moves before the end of the game with probability T/δ_s at a time distributed uniformly between $1-T$ and T. Player f then expects a gain of $T/2$. With probability $1 - T/\delta_s$, player s does not move and player f receives a gain of T. Summarizing, for a specific T, player f can expect to receive

$$\frac{T}{\delta_s}\frac{T}{2} + \left(1 - \frac{T}{\delta_s}\right)T = \frac{T^2}{2\delta_s} + T - \frac{T^2}{\delta_s} = T - \frac{T^2}{2\delta_s}.$$

Taking expectations over T yields the stated result: $\frac{1}{\delta_f}\int_{T=0}^{\delta_f} T - \frac{T^2}{2\delta_s}\,\mathrm{d}T$.

(Case 2: $i = s$, $j = f$) With probability δ_f/δ_s, we have $T \leq \delta_f$. The analysis is then the same as for Case 1, and player s expects to receive $T - T^2/2\delta_f$. With probability $1-\delta_f/\delta_s$, we have $T > \delta_f$. Player f then always regains control before the gain ends, at a time distributed uniformly between $1 - T$ and $1 - T + \delta_f$, yielding an expected gain of $\delta_f/2$ for player f. Taking expectations over T yields the stated result: $\delta_f/\delta_s \int_{T=0}^{\delta_f} 1/\delta_f (T - T^2/2\delta_f)\,\mathrm{d}T + (1 - \delta_f/\delta_s)\delta_f/2$.

(Case 3) Player i is the slower player. With probability $1/\delta_i$, i moves once, in which case she expects $\delta_j/2 - \delta_j^2/6\delta_i$ as argued in Case 2. With probability $1-1/\delta_i$, i never moves and she receives nothing. Taking expectations over player i's strategy yields the stated result.

(Case 4) There are four possible outcomes:

- Neither player moves. Player i receives no gain.
- Player i moves, player j does not. Player i receives an expected gain of $1/2$. This outcome occurs with probability $1/\delta_i(1 - 1/\delta_j)$.

- Player j moves, player i does not. Player i receives no gain.
- Both players move. Player i receives the same expected gain as a player with strategy $\delta_i = 1$ against a player with strategy $\delta_j = 1$. From any of the previous cases, we know that the expected gain in this scenario is $1/2 - 1/6 = 1/3$. This outcome occurs with probability $1/\delta_i\delta_j$.

Taking expectations yields the stated result: $1/\delta_i(1 - \frac{1}{\delta_j})1/2 + 1/\delta_i\delta_j 1/3$. □

Lemma 12. *Player i expects gain after her first and before her last move:*

- *(Case 1) If $\delta_i \geq 1$, then she expects 0.*
- *(Case 2) If $\delta_i \leq 1$ and $\delta_i \geq \delta_j$, then she expects $\delta_j/2\delta_i - \delta_j/2$.*
- *(Case 3) If $\delta_i \leq 1$ and $\delta_i \leq \delta_j$, then she expects $(1 - \delta_i)(1 - \delta_i/2\delta_j)$.*

Proof. Let I_i be an arbitrary instance of the interval between player i's first and last move. The duration of interval I_i equals the duration of the entire game, minus the time before the first and the time after the last flip. These both have an expected duration of $\delta_i/2$, so the expected duration of interval I_i is equal to $1 - \delta_i$ (assuming $\delta_i \leq 1$).

(Case 1) Player i flips either once or not at all, implying that I_i is always empty. Player i's expected gain over an empty interval is zero.

(Case 2; $i = s$, $j = f$). Partition I_s into sub-intervals of length δ_s. At the beginning of any such sub-interval, player s is in control of the resource. In expectation (over player f's strategy), player s remains in control for a duration of $\delta_f/2$ at the beginning of every sub-interval. It follows that player s's gain over the course of interval I_s is $\delta_f/2\delta_s$ times its duration. Taking expectations over player s's strategy yields $(1 - \delta_s)\delta_f/2\delta_s = \delta_f/2\delta_s - \delta_f/2$.

(Case 3; $i = f$, $j = s$). Partition I_f into sub-intervals of length δ_f. Consider any sub-interval. With probability δ_f/δ_s, the slower player moves once over the course of the sub-interval at a time that is uniformly distributed over the sub-interval, yielding an expected gain of $\delta_f/2$. With probability $1-\delta_f/\delta_s$, the slower player does not move, yielding a gain of δ_f. Player f's expected gain over the sub-interval is therefore $\delta_f/\delta_s\delta_f/2 + (1 - \delta_f/\delta_s)\delta_f = \delta_f(1 - \delta_f/2\delta_s)$, and player f's expected gain over the course of interval I_f is $(1 - \delta_f/2\delta_s)$ times its duration. Taking expectations over player f's strategy yields the result stated above. □