J. Cryptol. (2013) 26: 655–713
DOI: 10.1007/s00145-012-9134-5

Journal of
CRYPTOLOGY

FLIPIT: The Game of "Stealthy Takeover"*

Marten van Dijk, Ari Juels, and Alina Oprea
RSA Laboratories, Cambridge, MA, USA
marten.vandijk@rsa.com; ari.juels@rsa.com; alina.oprea@rsa.com

Ronald L. Rivest
Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA
rivest@mit.edu

Communicated by Rogaway

Received 17 February 2012
Online publication 26 October 2012

Abstract. Recent targeted attacks have increased significantly in sophistication, undermining the fundamental assumptions on which most cryptographic primitives rely for security. For instance, attackers launching an Advanced Persistent Threat (APT) can steal *full* cryptographic keys, violating the very secrecy of "secret" keys that cryptographers assume in designing secure protocols. In this article, we introduce a game-theoretic framework for modeling various computer security scenarios prevalent today, including targeted attacks. We are particularly interested in situations in which an attacker periodically compromises a system or critical resource *completely*, learns all its secret information and is not immediately detected by the system owner or *defender*. We propose a two-player game between an attacker and defender called FLIPIT or The Game of "Stealthy Takeover." In FLIPIT, players compete to control a shared resource. Unlike most existing games, FLIPIT allows players to move at any given time, taking control of the resource. The identity of the player controlling the resource, however, is not revealed until a player actually moves. To move, a player pays a certain move cost. The objective of each player is to control the resource a large fraction of time, while minimizing his total move cost. FLIPIT provides a simple and elegant framework in which we can formally reason about the interaction between attackers and defenders in practical scenarios. In this article, we restrict ourselves to games in which one of the players (the defender) plays with a *renewal strategy*, one in which the intervals between consecutive moves are chosen independently and uniformly at random from a fixed probability distribution. We consider attacker strategies ranging in increasing sophistication from simple periodic strategies (with moves spaced at equal time intervals) to more complex *adaptive strategies*, in which moves are determined based on feedback received during the game. For different classes of strategies employed by the attacker, we determine *strongly dominant* strategies for both players (when they exist), strategies that achieve higher benefit than all other strategies in a

* Solicited from Crypto 2011.

particular class. When strongly dominant strategies do not exist, our goal is to characterize the residual game consisting of strategies that are not strongly dominated by other strategies. We also prove equivalence or strict inclusion of certain classes of strategies under different conditions. Our analysis of different FLIPIT variants teaches cryptographers, system designers, and the community at large some valuable lessons:

1. Systems should be designed under the assumption of repeated total compromise, including theft of cryptographic keys. FLIPIT provides guidance on how to implement a cost-effective defensive strategy.
2. Aggressive play by one player can motivate the opponent to drop out of the game (essentially not to play at all). Therefore, moving fast is a good defensive strategy, but it can only be implemented if move costs are low. We believe that virtualization has a huge potential in this respect.
3. Close monitoring of one's resources is beneficial in detecting potential attacks faster, gaining insight into attacker's strategies, and scheduling defensive moves more effectively.

Interestingly, FLIPIT finds applications in other security realms besides modeling of targeted attacks. Examples include cryptographic key rotation, password changing policies, refreshing virtual machines, and cloud auditing.

Key words. FLIPIT, Security modeling, Game theory, Advanced persistent threats, Repeated stealthy takeovers.

1. Introduction

The cyber-security landscape has changed tremendously in recent years. Major corporations and organizations, including the United Nations, Lockheed Martin, Google and RSA, have experienced increasingly sophisticated targeted attacks (e.g., Aurora, Stuxnet). These attacks, often referred in industry circles as *Advanced Persistent Threats* (APTs), exhibit several distinctive characteristics compared to more traditional malware. APTs are extremely well funded and organized, have very specific objectives, and have their actions controlled by human operators. They are in general persistent for long periods of time in a system, advancing stealthily and slowly to maintain a small footprint and reduce detection risks. Typically, APTs exploit infrastructure vulnerabilities through an arsenal of zero-day exploits and human vulnerabilities through advanced social-engineering attacks.

When designing security protocols, cryptographers and computer security researchers model adversarial behavior by making simplifying assumptions. Traditional cryptography relies for security on the secrecy of cryptographic keys (or of system state). Modern cryptography considers the possibility of partial compromise of system secrets. *Leakage resilient* cryptography, for instance, constructs secure primitives under the assumption that a system is subject to continuous, partial compromise. But in an adversarial situation, assumptions may fail completely. For instance, attackers launching an APT can steal full cryptographic keys, crossing the line of inviolate key secrecy that cryptographers assume in designing primitives. Assumptions may also fail repeatedly, as attackers find new ways to penetrate systems.

In this paper, motivated by recent APT attacks, we take a new perspective and develop a model in which an attacker periodically compromises a system *completely*, in the sense of learning its entire state, including its secret keys. We are especially interested

Categories	Classes of Strategies	
Non-adaptive (NA)	Exponential	Periodic
	Renewal	
	General non-adaptive	
Adaptive (AD)	Last move (LM)	
	Full history (FH)	

Fig. 1. Hierarchy of strategies in FLIPIT.

in those scenarios where theft is *stealthy* or *covert* and not immediately noticed by the victim. We model these scenarios by means of a two-player game we call FLIPIT, a.k.a. The Game of "*Stealthy Takeovers*".

FLIPIT is a two-player game with a shared resource that both players (called herein *attacker* and *defender*) wish to control. The resource of interest could be, for instance, a secret key, a password, or an entire infrastructure, depending on the situation being modeled. Players take control of the resource by moving, and paying a certain move cost, but unlike most existing games, in FLIPIT players can move at any given time. Most importantly, *a player does not immediately know when the other player moves*! A player only finds out about the state of the system when she moves herself. This stealthy aspect of the game is a unique feature of FLIPIT, which to the best of our knowledge has not been explored in the game theory literature. The goal of each player is to maximize a metric we call *benefit*, defined as the fraction of time the player controls the resource minus the average move cost. A good strategy for a given player, therefore, is one that gives the player control of the resource a large fraction of the time with few moves.

Our main goal in this paper is to find best-choice (ideally *dominant* or *strongly dominant*) playing strategies for different variants of the FLIPIT game that achieve maximum benefit for each player. An instance of FLIPIT is determined by the information players receive before the game, the feedback players obtain while playing the game, and a space of strategies or actions for each player. We have created a hierarchical graphical representation of different classes of strategies for one player in Fig. 1. The strategies are ordered (from top to bottom) by increasing amount of feedback received by a player during the game. As we move down the hierarchy, we encompass broader classes of strategies. In the paper, we prove results about the equivalence of different classes of strategies, (strongly) dominant and dominated strategies within a particular class, and strict inclusion of a class in a broader class. We also characterize, for several instances of FLIPIT, the residual game consisting of strategies that are not strongly dominated. We detail below several FLIPIT instances that we define and analyze, and then describe the theoretical results proven in the paper, as well as the results based on experimental evidence. We also discuss some of our conjectures regarding the relationship among different classes in the hierarchy, leading to some challenging open problems.

FLIPIT *Instances* Broadly, as graphically shown in Fig. 1, we define two main classes of strategies for a player: *non-adaptive strategies* (NA) in which the player does

not receive any feedback during the game; and *adaptive strategies* (AD) in which the player receives certain types of feedback when moving.

In the class of non-adaptive strategies, one subclass of particular interest is that of *renewal strategies*. For a player employing a renewal strategy, the intervals between the player's consecutive moves are independent and identically distributed random variables generated by a *renewal process*. Examples of renewal strategies include: the *periodic* strategy in which the interval between two consecutive moves is fixed at a constant value and the *exponential* strategy in which the intervals between consecutive moves are exponentially distributed. The moves of a player employing a non-adaptive (but possibly randomized) strategy can be determined before the game starts, as there is no feedback received during the game.

In the class of adaptive strategies, we distinguish a subclass called *last move* or *LM*, in which the player finds out upon moving the exact time when his opponent moved last. In a general adaptive strategy, a player receives upon moving complete information about his opponent's moves (this is called *full history* or *FH*). Classes LM and FH collapse in some cases (for instance when the opponent plays with a renewal strategy).

A further dimension to a player's strategy is the amount of information the player receives before the game starts. Besides the case in which the player receives no information about his opponent before the game, other interesting cases are: (1) *rate-of-play* (or RP), in which the player finds out the exact rate of play of his opponent; (2) *knowledge-of-strategy* (KS), in which the player finds out full information about the strategy of the opponent (but not the opponent's randomness).

Contributions and Results Our first contribution in this paper is the definition of the FLIPIT game, and its application to various computer security scenarios, including APTs. We view FLIPIT as a theoretical framework that helps provide a foundation for the science of cyber-security.

We then analyze several instances of the game of increasing sophistication and complexity. Our main theoretical and experimental results, as well as conjectures, and remaining open questions are summarized in Table 1. To simplify the exposition, we present results for the case when the attacker is more powerful than the defender and organize these results according to the strength of the attacker. Since the game is symmetric, though, analogous results apply when the defender is more powerful than the attacker. In this article, we restrict ourselves to games in which the defender plays with a renewal strategy.

To elaborate, we start by analyzing the simple game in which both players employ a periodic strategy with a random phase. (In such a strategy, moves are spaced at equal intervals, with the exception of the first randomly selected move called the *phase*.) We compute the Nash equilibrium point for the periodic game and show that the benefit of the player with higher move cost is always 0 in the Nash equilibrium. When move costs of both players are equal, players control the resource evenly and both achieve benefit 0.

We next consider a variant of FLIPIT in which both players play with either a renewal or periodic strategy with random phase. We call such games *renewal games*. Our main result for the renewal game (Theorem 4 in Sect. 4.3) shows that the periodic strategy with a random phase strongly dominates all renewal strategies of fixed

Table 1. Paper main results and open problems. Notation used in the table is: NA—non-adaptive strategies; AD—adaptive strategies; LM—last-move strategies; RP—NA strategies with rate-of-play information about opponent; KS—NA strategies with knowledge-of-strategy information about opponent.

Attacker category	Strategy classes		Results	Open problems
	Attacker	Defender		
NA	Periodic	Periodic	Full analysis Nash equilibrium (Theorem 1)	
	Renewal	Renewal	Strong dominance of periodic strategies (Theorem 4)	
	General NA	General NA		Which strategies are (strongly) dominant or dominated?
NA + Additional information before the game	Renewal RP	Renewal	Strong dominance of periodic strategies (Theorem 4) Optimal parameters (Theorem 5)	
	Renewal RP	Renewal RP	Strong dominance of periodic strategies (Theorem 4)	
	General NA RP or KS	General NA (RP)		Which strategies are (strongly) dominant or dominated?
AD	LM	Periodic	The periodic strategy is strongly dominant for attacker	
		Exponential	The periodic strategy is strongly dominant for attacker (Theorem 6) Optimal parameters (Theorems 7 and 8)	
		Delayed exponential	Strongly dominant attacker strategy (Theorem 9) Defender's benefit increases compared to exponential (Experimental)	Theoretical analysis for benefits
		Renewal		Which renewal strategies are (strongly) dominant or dominated for the defender? Which LM strategies are (strongly) dominant or dominated for the attacker?
		LM		Which strategies are (strongly) dominant or dominated for attacker and defender? Does cooperation help?
	Greedy	Renewal	Analyzed for several defender distributions Strongly dominant strategy for exponential defender	For which classes of defender strategies is Greedy a dominant strategy?

rate. Therefore, we can completely characterize the residual FLIPIT game when players employ either renewal or periodic strategies with random phases. It consists of all periodic strategies with random phases.

We also analyze renewal games in which the attacker receives additional information about the defender at the beginning of the game, in particular the rate of play of the defender. For this version of the game, we prove that Theorem 4 still holds, and periodic strategies are strongly dominant. Additionally, we determine in Theorem 5 in Sect. 4.4 parameter choices for the rates of play of both attacker and defender that achieve maximum benefit.

Moving towards increased sophistication in the attacker's strategy, we next analyze several FLIPIT instances for the attacker in the broad Last-Move (LM) class. The attacker in this case adaptively determines his next move based on the feedback learned during the game, in particular the exact time of the defender's last move. In this setting, periodic play for the defender is not very effective, as the attacker learning the defender's period and last-move time can move right after the defender and achieve control of the resource most of the time. We demonstrate in Theorem 6 in Sect. 5.3 that by playing with an exponential strategy instead, the defender forces an adaptive attacker into playing periodically. (That is, the periodic strategy is the strongly dominant attacker strategy in the class of all LM strategies against an exponential defender strategy.) Paradoxically, therefore, the attacker's strongly dominant strategy does not make use of the full knowledge the attacker gains from the feedback received during the game. (The full hierarchy in Fig. 1 for the attacker strategy collapses for exponential defender play.) Additionally, we determine optimal parameters (i.e., parameters that optimize both players' benefits) for both the exponential defender and periodic attacker distributions in Theorems 7 and 8 in Sect. 5.3.

For both non-adaptive and fully adaptive attacker strategies, we show that by playing periodically at sufficiently high rate, the defender can force the attacker to drop out of the game. The move cost of the defender needs to be significantly lower than that of the attacker, however, for such a strategy to bring benefit to the defender. This demonstrates that the player radically lowering his move cost relative to his opponent obtains a high advantage in the game and can effectively control the resource at all times.

We propose an enhancement to the exponential distribution employed by the defender, a strategy that we call *delayed exponential*. In this strategy, the defender waits after each move for some fixed interval of time and then waits for an exponentially distributed time interval until making her next move. We show experimentally that with this strategy the benefit of the defender is higher than that for exponential play. This result provides evidence that the exponential strategy is not the strongly dominant renewal defender strategy against an LM attacker. There are still interesting questions for which we cannot yet provide a definitive answer: What other renewal strategies for the defender could further increase her benefit?; Is there a strongly dominant strategy for the defender in the class of all renewal strategies against an LM attacker?

Lastly, we explore a Greedy algorithm in which the attacker's moves are chosen to maximize the *local benefit* achieved in an interval between two consecutive attacker moves. We analyze the Greedy strategy for several renewal defender strategies and demonstrate that it results in the strongly dominant strategy for an exponential defender strategy. We also show one example for which the Greedy strategy does not result in

a dominant strategy for the attacker. We leave open the challenging problem of which strategies are (strongly) dominant for an LM attacker against particular classes of defender strategies (e.g., renewal strategies).

Lessons Derived from FLIPIT Our analysis of FLIPIT teaches cryptographers, system designers, and the community at large some valuable lessons:

1. Systems should be designed under the assumption of repeated total compromise, including theft of cryptographic keys. Many times, attackers cross the line of secrecy cryptographers assume in their protocol designs. This view has already been expressed by the head of the NSA's Information Assurance Directorate [28]: "No computer network can be considered completely and utterly impenetrable—not even that of the NSA. NSA works under the assumption that various parts of their systems have already been compromised, and is adjusting its actions accordingly."

   FLIPIT provides guidance on how and when to implement a cost-effective defense. For instance, our analysis can help defenders determine when they should change cryptographic keys or user credentials, and how often should they clean their machines or refresh virtual machine instances. FLIPIT analysis, conversely, also guides attackers in how to schedule their attacks to have maximal effect, while minimizing attack cost.
2. Aggressive play by the defender can motivate the attacker to drop out of the game (and essentially not to play at all). The best defensive strategy, therefore, is to play fast (for instance by changing passwords frequently, rebooting servers often, or refreshing virtual machines at short intervals) and make the opponent drop out of the game. To be able to move fast, the defender should arrange the game so that her moves cost much less than the attacker's moves.

   An interesting research challenge for system designers is how to design an infrastructure in which refresh/clean costs are very low. We believe that virtualization has huge potential in this respect.
3. As we have shown in our theoretical analysis, any amount of feedback (even limited) received during the game about the opponent benefits a player in FLIPIT. Defenders, therefore, should monitor their systems frequently to gain information about the attacker's strategy and detect potential attacks quickly after take over. Both monitoring and fast detection help a defender more effectively schedule moves, which results in more control of the resource and less budget spent on moves, increasing the defender's benefit.

Conjectures and Open Problems In this paper, we introduce a new framework, reflected in the FLIPIT game, for modeling different cyber-security situations. The FLIPIT game is unique in the game theory literature mostly for its stealthy aspect and continuous-time play. We prove some interesting results about different variants of the game, and the relationships among several classes of strategies. Nevertheless, many interesting questions remain to be answered in order to understand fully all facets of this simple (but challenging-to-analyze) game. We elaborate here on some of the more interesting open problems we have identified, and also state some conjectures for whose truth our results provide strong evidence.

1. For FLIPIT with non-adaptive strategies for both defender and attacker, we have fully analyzed the subclass of renewal games. An analysis of the game in which one or both players employ general non-adaptive strategies is deferred to future work. It would be interesting to determine if periodic strategies are still dominant in the general non-adaptive FLIPIT game.
2. In the case in which the attacker receives additional information about the defender at the beginning of the game (e.g., the rate of play of the defender or extra information about the defender's strategy) and both players employ non-adaptive strategies, which strategies are best choices for both players? From this class of games, we analyzed the renewal game in which one player receives RP information and the other receives no information, RP, or KS. We leave open the analysis of other instances of FLIPIT in this class.
3. We believe that the most interesting open questions arise in the case in which one player (e.g., the attacker) is adaptive. Here we enumerate several of them:
   - When the attacker is LM, which renewal strategies for the defender are (strongly) dominant?
   - Which LM strategies are (strongly) dominant against a renewal defender?
   - For which class of defender strategies is the Greedy algorithm (strongly) dominant for an LM attacker?
   - When the attacker is LM, the relationship between the non-adaptive and adaptive classes for the defender is not completely understood. We have though strong evidence (also deferred to follow-up work) to state the following conjecture:

Conjecture 1. *Consider an instance of* FLIPIT *with an LM attacker. Then the defender receiving feedback during the game and playing with an adaptive strategy can strictly improve her benefit relative to non-adaptive play.*

4. The most challenging instance of FLIPIT is when both players utilize adaptive strategies. In this setting, it would be interesting to determine (strongly) dominant strategies (if any), and to what extent they involve cooperation among players.

2. Game Motivation

We start by defining FLIPIT in an informal way. In the second part of this section, we describe several practical applications that motivated the game definition. In Sect. 3, we give a formal definition of FLIPIT with careful specification of players' action spaces, strategies, and benefits.

2.1. *Game Definition by Example*

Consider a resource that can be controlled (or "owned") by either of two players (attacker or defender). Ownership will change back and forth following a move of either player, with the goal of each player being to maximize the fraction of time that he or she controls the resource. A distinctive feature of FLIPIT is its *stealthy* aspect, that is, the players *do not know* when the other player has taken over. Nor do they know the current

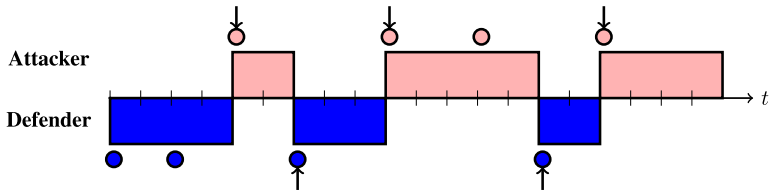

Fig. 2. The FLIPIT game. *Blue* and *red circles* represent defender and attacker moves, respectively. Takeovers are represented by *arrows*. *Shaded rectangles* show the control of the resource—*blue* (*dark gray* in grayscale) for the defender and *red* (*light gray* in grayscale) for the attacker. We assume that upon initialization at time 0, the defender has control.

ownership of the resource unless they perform a move. Also important is the fact that to move, a player must pay a move cost; players thus have a disincentive against moving too frequently.

In an example implementation of a basic version of FLIPIT, each player has a control panel with a single button and a light. The player may push his/her button at any time (in the most general form of FLIPIT, we consider time to be continuous, but we support discrete variants as well). Pushing the button always causes the button-pusher to take ownership of the resource. We assume that players do not push their buttons at exactly the same time (or, if they do, then ownership does not change hands).

If pushing the button causes ownership to change hands, then the light flashes when the button is pushed. We call this a "takeover." If the button-pusher already had ownership of the resource, then the light does not flash and the button-push was wasted.

The players cannot see each other's control panels, and thus the defender does not know when the attacker takes control, and vice versa. The only way a player can determine the state of the game is to push his/her button. Thus a move by either player has two consequences: it acquires the control of the resource (if not already controlled by the mover), but at the same time, it reveals the pre-move state of the resource to the player taking control.

There is always a cost to pushing the button. In this example, pushing the button costs the equivalent of one second of ownership. Thus, at any time t, each player's net score is the number of seconds he has had ownership of the resource, minus the number of times he has pushed his button.

We show a graphical representation of the game in Fig. 2. The control of the resource is graphically depicted through shaded rectangles, a blue rectangle (dark gray in grayscale) representing a period of defender control, a red rectangle (light gray in grayscale) one of attacker control. Players' moves are graphically depicted with shaded circles. A vertical arrow denotes a takeover, when a player (re)takes control of the resource upon moving.

The main focus of the paper is in analyzing "optimal play" for both attacker and defender in this simple game. We wish to explore this question under various assumptions about the details of the rules, or about the strategies the players employ. In what follows, we motivate the choice of this game through several practical applications, and describe various extensions of the basic game needed to model these applications.

2.2. *Motivating Applications*

A prime motivation for FLIPIT is the rise of Advanced Persistent Threats (APTs), which often play out as a protracted, stealthy contest for control of computing resources. FLIPIT, however, finds many other natural applications in computer security, and even in other realms of trust assurance. Most of these applications correspond to slightly modified versions of the basic FLIPIT game. To motivate exploration of FLIPIT, we describe some of these other applications here, along with the extensions or modifications of the basic game required to model them.

2.2.1. *Advanced Persistent Threats (APTs)—A Macro-level Game*

An APT is a concerted, stealthy, and long-lived attack by a highly resourced entity against a critical digital resource. Publicly acknowledged APTs include disablement of an Iranian uranium enrichment facility by means of the Stuxnet worm [7] and the breach of security vendor RSA, an attack whose ultimate objective was reportedly the theft of military secrets or intellectual property from organizations relying on RSA's authentication tokens [25].

Stealth is a key characteristic: attackers rely on extended reconnaissance and human oversight to minimize the risk of detection by the defender, which can quickly set back or thwart an APT. Similarly, a defender looks to conceal its knowledge of a detected APT, to avoid alerting the attacker and allowing it to achieve renewed stealth by changing its strategy.

In the case where the attacker looks to exercise persistent control over a set of target resources (as opposed to achieving a one-time mission), FLIPIT can serve as a global model for an APT. The defender's sensitive assets may include computers, internal networks, document repositories, and so forth. In such a macro-level view, we treat all of these assets as an aggregate FLIPIT resource that the defender wishes to keep "clean." The goal of the attacker is to compromise the resource and control it for a substantial fraction of time.

In this macro-level model, a move by the attacker is a campaign that results in control over essential target resources, that is, a thorough breach of the system. A move by the defender is a system-wide remediation campaign, e.g., patching of all machines in a network, global password refresh, reinstallation of all critical servers, etc.

The macro-level view of a defender's critical resources as a single FLIPIT resource is attractively simple. A more realistic and refined model, however, might model an APT as a series of stages along a path to system takeover, each stage individually treated as a micro-level FLIPIT game. The attacker's global level of control at a given time in this case is a function of its success in each of the component games. We now give examples of some micro-level games.

2.2.2. *Micro-level Games*

Host Takeover In this version of the game, the target resource is a computing device. The goal of the attacker is to compromise the device by exploiting a software vulnerability or credential compromise. The goal of the defender is to keep the device clean through software reinstallation, patching, or other defensive steps.

An action by either side carries a cost. For the attacker, the cost of host compromise may be that of, e.g., mounting a social-engineering attack that causes a user to open an infected attachment. For the defender, cleaning a host may carry labor and lost-productivity costs.

FLIPIT provides guidance to both sides on how to implement a cost-effective schedule. It helps the defender answer the question, "How regularly should I clean machines?" and the attacker, "When should I launch my next attack?"

A Variant/Extension There are many ways to compromise or clean a host, with varying costs and criteria for success. In a refinement of the game, players might choose among a set of actions with varying costs and effectiveness.

For example, the attacker might choose between two types of move: (1) Use of a published exploit or (2) Use of a zero-day exploit, while the defender chooses either to: (1) patch a machine or (2) reinstall its software. For both players, (1) is the less expensive, but (2) the more effective. Action (2) results in takeover for either player, while action (1) will only work if the opponent's most recent move was action (1). (For instance, patching a machine will not recover control from a zero-day exploit, but software reinstallation will.)

Refreshing Virtual Machines (VMs) Virtualization is seeing heavy use today in the deployment of servers in data centers. As individual servers may experience periods of idleness, consolidating multiple servers as VMs on a single physical host often results in greater hardware utilization. Similarly, Virtual Desktop Infrastructure (VDI) is an emerging workplace technology that provisions users with VMs (desktops) maintained in centrally managed servers. In this model, users are not bound to particular physical machines. They can access their virtual desktops from any endpoint device available to them, even smart phones.

While virtualization exhibits many usability challenges, one key advantage is a security feature: VMs can be periodically refreshed (or built from scratch) from "clean" images, and data easily restored from backups maintained at the central server.

Takeover of a VM results in a game very similar to that for a physical host. Virtualization is of particular interest in the context of FLIPIT, though, because FLIPIT offers a means of measuring (or at least qualitatively illustrating) its security benefits. Refreshing a VM is much less cumbersome than rebuilding the software stack in a physical host. In other words, virtualization lowers the move cost for the defender. Optimal defender play will therefore result in resource control for a higher proportion of the time than play against a comparable attacker on a physical host.

An Extension Virtualization also brings about an interesting question at the macro-level. How can refreshes of individual VMs be best scheduled while maintaining service levels for a data center as a whole? In other words, how can refresh schedules best be crafted to meet the dual goals of security and avoidance of simultaneous outage of many servers/VDI instances?

Password Reset When an account or other resource is password-protected, control by its owner relies on the password's secrecy. Password compromise may be modeled as a FLIPIT game in which a move by the attacker results in its learning the password. (The attacker may run a password cracker or purchase a password in an underground

forum.) The defender regains control by resetting the password, and thus restoring its secrecy.

This game differs somewhat from basic FLIPIT. An attacker can detect the reset of a password it has compromised simply by trying to log into the corresponding account. A defender, though, does not learn on resetting a password whether it has been compromised by an attacker. Thus, the defender can only play *non-adaptively*, while the attacker has a second move option available, a probe move that reveals the state of control of the resource.

2.2.3. *Other Applications*

Key Rotation A common hedge against the compromise of cryptographic keys is *key rotation*, the periodic generation and distribution of fresh keys by a trusted key-management service. Less common in practice, but well explored in the research literature is *key-evolving cryptography*, a related approach in which new keys are generated by their owner, either in isolation or jointly with a second party. In all of these schemes, the aim is for a defender to change keys so that compromise by an attacker in a given epoch (interval of time) does not impact the secrecy of keys in other epochs.

Forward-secure protocols [12] protect the keys of past epochs, but not those of future epochs. Key-insulated [6] and intrusion-resilient [6] cryptography protect the keys of both past and future epochs, at the cost of involving a second party in key updates. Cryptographic tamper evidence [11] provides a mechanism for detecting compromise when both the valid owner and the attacker make use of a compromised secret key.

Key updates in all of these schemes occur at regular time intervals, i.e., epochs are of fixed length. FLIPIT provides a useful insight here, namely the potential benefit of *variable-length epochs*.

The mapping of this scenario onto FLIPIT depends upon the nature of the target key. An attacker can make use of a decryption key in a strongly stealthy way, i.e., can eavesdrop on communications without ready detection, and can also easily detect a change in key. In this case, the two players knowledge is asymmetric. The defender must play non-adaptively, while the attacker has the option of a *probe* move at any time, i.e., can determine the state of the system at low cost.

Use of a signing key, on the other hand, can betray compromise by an attacker, as an invalidly signed message may appear anomalous due to its content or repudiation by the key owner.

Variant/Extension To reflect the information disclosed by use of a signing key, we might consider a variant of FLIPIT in which system state is revealed *probabilistically*. Compromise of a signing key by an attacker only comes to light if the defender actually intercepts a message signed by the attacker and determines from the nature of the signed message or other records that the signature is counterfeit. Similarly, the attacker only learns of a key update by the defender when the attacker discovers a signature by the defender under a new key. (In a further refinement, we might distinguish between moves that update/compromise a key and those involving its actual use for signing.)

Cloud Service Auditing Cloud computing is a recent swing of the pendulum away from endpoint-heavy computing toward centralization of computing resources [16]. Its

benefits are many, including, as noted above, certain forms of enhanced security offered by virtualization, a common element of cloud infrastructure. Cloud computing has a notable drawback, though: it requires users (often called "tenants") to rely on the trustworthiness of service providers for both reliability and security.

To return visibility to tenants, a number of *audit* protocols have been proposed that enable verification of service-level agreement (SLA) compliance by a cloud service provider [1–4,13,26]. The strongest of these schemes are challenge-response protocols. In a Proof of Retrievability (PoR), for instance, a tenant challenges a cloud service provider to demonstrate that a stored file is remotely retrievable, i.e., is fully intact and accessible via a standard application interface. Other protocols demonstrate properties such as quality-of-service (QoS) levels, e.g., retention of a file in high-tier storage, and storage redundancy [4], i.e., distribution of a file across multiple hard drives.

Execution of an audit protocol carries a cost: A PoR, for instance, requires the retrieval of some file elements by the cloud and their verification by a tenant. It is natural then to ask: *What is the best way for a tenant to schedule challenges in an audit scheme?*. (Conversely, what is the best way for a cloud service provider to cheat an auditor?) This question goes unaddressed in the literature on cloud audit schemes. (The same question arises in other forms of audit, and does see some treatment [5,20].)

Auditing for verification of SLA compliance is particularly amenable to modeling in FLIPIT. A move by the defender (tenant) is a challenge/audit, one that forces the provider into compliance (e.g., placement of a file in high-tier storage) if it has lapsed. A move by the attacker (cloud) is a downgrading of its service level in violation of an SLA (e.g., relegation of a file to a low storage tier). The metric of interest is the fraction of time the provider meets the SLA.

3. Formal Definition and Notation

This section gives a formal definition of the stealthy takeover game, and introduces various pieces of useful notation.

Players There are two players: the defender is the "good" player, identified with 0 (or Alice). The attacker is the "bad" player, identified with 1 (or Bob). It is convenient in our development to treat the game as symmetric between the two players.

Time The game begins at time $t = 0$ and continues indefinitely as $t \to \infty$. In the general form of the game, time is viewed as being *continuous*, but we also support a version of the game with *discrete* time.

Game State The time-dependent variable $C = C(t)$ denotes the current player controlling the resource at time t; $C(t)$ is either 0 or 1 at any time t. We say that the game is in a "good state" if $C(t) = 0$, and in a "bad state" if $C(t) = 1$.

For $i = 0, 1$ we also let

$$C_i(t) = I\big(C(t) = i\big)$$

denote whether the game is in a good state for player i at time t. Here I is an "indicator function": $I(\cdot) = 1$ if its argument is true, and 0 otherwise. Thus, $C_1(t) = C(t)$ and

$C_0(t) = 1 - C_1(t)$. The use of C_0 and C_1 allows us to present the game in a symmetric manner.

The game begins in a good state: $C(0) = 0$.

Moves A player may "move" (push his/her button) at any time, but is only allowed to push the button a finite number of times in any finite time interval. The player may not, for example, push the button at times $1/2, 2/3, 3/4, \ldots$, as this means pushing the button an infinite number of times in the time interval $[0, 1]$. (One could even impose an explicit lower bound on the time allowed between two button-pushes by the same player.)

A player cannot move more than once at a given time. We allow different players to play at the same time, although with typical strategies this happens with probability 0. If it does happen, then the moves "cancel" and no change of state happens. (This tie-breaking rule makes the game fully symmetric, which we prefer to alternative approaches such as giving a preference to one of the players when breaking a tie.) It is convenient to have a framework that handles ties smoothly, since discrete versions of the game, wherein all moves happen at integer times, might also be of interest. In such variants ties may be relatively common.

We denote the sequence of move times, for moves by both players, as an infinite non-decreasing sequence:

$$\mathbf{t} = t_1, t_2, t_3, \ldots.$$

The sequence might be non-decreasing, rather than strictly increasing, since we allow the two players move at the same time.

We let p_k denote the player who made the kth move, so that $p_k \in \{0, 1\}$. We let $\mathbf{p}$ denote the sequence of player identities:

$$\mathbf{p} = p_1, p_2, p_3, \ldots.$$

We assume that $t_1 = 0$ and $p_1 = 0$; the good player (the defender) moves first at time $t = 0$ to start the game.

For $i = 0, 1$ we let

$$\mathbf{t}_i = t_{i,1}, t_{i,2}, t_{i,3}, \ldots$$

denote the infinite increasing sequence of times when player i moves.

The sequences $\mathbf{t}_0$ and $\mathbf{t}_1$ are disjoint subsequences of the sequence $\mathbf{t}$. Every element t_k of $\mathbf{t}$ is either an element $t_{0,j}$ of $\mathbf{t}_0$ or an element $t_{1,l}$ of $\mathbf{t}_1$.

The game's state variable $C(t)$ denotes the player who has moved most recently (not including the current instant t), so that

$$C(t) = p_k \quad \text{for } t_k < t \leq t_{k+1} \text{ and for all } k \geq 1.$$

When $C(t) = i$ then player i has moved most recently and is "in control of the game", or "in possession of the resource." We assume, again, that $C(0) = 0$.

Note that $C(t_k) = p_{k-1}$; this is convenient for our development, since if a player moves at time t_k then $C(t_k)$ denotes the player who was previously in control of the game (which could be either player).

For compatibility with our tie-breaking rule, we assume that if $t_k = t_{k+1}$ (a tie has occurred), then the two moves at times t_k and t_{k+1} have subscripts ordered so that no net change of state occurs. That is, we assume that $p_k = 1 - C(t_k)$ and $p_{k+1} = 1 - p_k$. Thus, each button-push causes a change of state, but no net change of state occurs.

We let $n_i(t)$ denote the *number of moves made by player i* up to and including time t, and let

$$n(t) = n_0(t) + n_1(t)$$

denote the *total number of moves made by both players* up to and including time t.

For $t > 0$ and $i = 0, 1$, we let

$$\alpha_i(t) = n_i(t)/t$$

denote the *average move rate by player i* up to time t.

We let $r_i(t)$ denote the time of the most recent move by player i; this is the largest value of $t_{i,k}$ that is less than t (if player i has not moved since the beginning of the game, then we define $r_i(t) = -1$). Player 0 always moves at time 0, and therefore $r_0(t) \geq 0$. We let $r(t) = \max(r_0(t), r_1(t)) \geq 0$ denote the time of the most recent move by either player.

Feedback During the Game We distinguish various types of feedback that a player may obtain during the game (specifically upon moving).

It is one of the most interesting aspects of this game that the players do *not* automatically find out when the other player has last moved; moves are *stealthy*. A player must move himself to find out (and reassert control).

We let $\phi_i(t_k)$ denote the feedback player i obtains when the player moves at time t_k. This feedback may depend on which variant of the game is being played.

- *Non-adaptive* [NA]. In this case, a player does not receive any useful feedback whatsoever when he moves; the feedback function is constant.

$$\phi_i(t_k) = 0.$$

- *Last move* [LM]. The player moving at time $t_k > 0$ finds out the exact time when the opponent played last before time t_k. That is, player i learns the value:

$$\phi_i(t_k) = r_{1-i}(t_k).$$

- *Full history* [FH]. The mover finds out the complete history of moves made by both players so far:

$$\phi_i(t_k) = \big((t_1, t_2, \ldots, t_k), (p_1, p_2, \ldots, p_k)\big).$$

We abbreviate these forms of feedback as NA, LM, and FH, and we define other types of feedback in Sect. 7. We consider "non-adaptive" (NA) feedback to be the default (standard) version of the game. When there is feedback and the game is adaptive, then players interact in a meaningful way and therefore cooperative (e.g.,"tit-for-tat") strategies may become relevant.

Views and History A *view* is the history of the game from one player's viewpoint, from the beginning of the game up to time t. It lists every time that player moved, and the feedback received for that move.

For example, the view for player i at time t is the list:

$$\mathbf{v}_i(t) = \big(\big(t_{i,1}, \phi(t_{i,1})\big), \big(t_{i,2}, \phi(t_{i,2})\big), \ldots, \big(t_{i,j}, \phi(t_{i,j})\big)\big)$$

where $t_{i,j}$ is the time of player i's jth move, her last move up to time t, and $\phi(t_{i,j})$ is the feedback player i obtains when making her jth move.

A *history* of a game is the pair of the players' views.

Strategies A *strategy* for playing this game is a (possibly randomized) mapping S from views to positive real numbers. If S is a strategy and v a view of length j, then $S(v)$ denotes the time for the player to wait before making move $j+1$, so that $t_{i,j+1} = t_{i,j} + S(v)$.

The next view for the player following strategy S will thus be

$$\big(\big(t_{i,1}, \phi(t_{i,1})\big), \big(t_{i,2}, \phi(t_{i,2})\big), \ldots, \big(t_{i,j+1}, \phi(t_{i,j+1})\big)\big).$$

We define now several classes of strategies, and we refer the reader to Fig. 1 for a hierarchy of these classes.

Non-adaptive Strategies We say that a strategy is *non-adaptive* if it does not require feedback received during the game, and we denote by $\mathcal{N}$ the class of all non-adaptive strategies. A player with a non-adaptive strategy plays in the same manner against every opponent. A player with a non-adaptive strategy can in principle generate the time sequence for all of his moves in advance, since they do not depend on what the other player does. They may, however, depend on some independent source of randomness; non-adaptive strategies may be randomized.

Renewal Strategies Renewal strategies are non-adaptive strategies for which the intervals between consecutive moves are generated by a renewal process (as defined, for instance, by Feller [8]). Therefore, the inter-arrival times between moves are independent and identical distributed random variables chosen from a probability density function f. As the name suggests, these strategies are "renewed" after each move: the interval until the next move only depends on the current move time and not on previous history.

Periodic Strategies An example of a simple renewal strategy is a *periodic* strategy. We call a strategy *periodic* if there is a δ such that the player always presses his button again once exactly δ seconds have elapsed since his last button-push. We assume that the periodic strategy has a *random phase*, i.e., the first move is selected uniformly at random from interval $[0, \delta]$ (if the strategy is completely deterministic an adaptive opponent can find out the exact move times and schedule his moves accordingly).

Exponential Strategies We call a strategy *exponential* or *Poisson* if the player pushes his button in a Poisson manner: there is some rate λ such that in any short time increment Δ the probability of a button-push is approximately $\lambda \cdot \Delta$. In this case, $S(v)$ has an exponential distribution and this results in a particular instance of a renewal strategy.

Adaptive Strategies The class of adaptive strategies encompasses strategies in which players receive feedback during the game. In the LM class of strategies, denoted $\mathcal{A}_{\mathsf{LM}}$, a player receives last-move feedback, while in the FH-adaptive class, denoted $\mathcal{A}_{\mathsf{FH}}$, a player receives full-history feedback, as previously defined.

No-play Strategies We denote by Φ the strategy of not playing at all (effectively dropping out of the game). We will show that this strategy is sometimes the best response for a player against an opponent playing extremely fast.

Information Received Before the Game Starts Besides receiving feedback during the game, sometimes players receive additional information about the opponent before the game starts. We capture this with $\phi_i(0)$, which denotes the information received by player i before the game starts. There are several cases we consider:

- *Rate of Play* [RP]. In this version of the game, player i finds out the limit of the rate of play $\alpha_{1-i}(t)$ of its opponent at the beginning of the game (assuming that the rate of play converges to a finite value):

$$\phi_i(0) = \lim_{t\to\infty} \alpha_{1-i}(t).$$

  No additional information, however, is revealed to the player about his opponent's moves during the game.
- *Knowledge of Strategy* [KS]. Player i might find additional information about the opponent's strategy. For instance, if the opponent (player $1-i$) employs a renewal strategy generated by probability density function f, then KS information for player i is the exact distribution f:

$$\phi_i(0) = f.$$

  Knowledge of the renewal distribution in this case does not uniquely determine the moves of player $1-i$, as the randomness used by player $1-i$ is not divulged to player i. In this paper, we only use KS in conjunction with renewal strategies, but the concept can be generalized to other classes of strategies.

RP and KS are meaningfully applicable only to a non-adaptive attacker. An adaptive attacker from class LM or FH can estimate the rate of play of the defender, as well as the defender's strategy during the game from the information received when moving. But a non-adaptive attacker receiving RP or KS information before the game starts can adapt his strategy and base moves on pre-game knowledge about the opponent's strategy.

Pre-game information received at the beginning of the game, hence, could be used, in conjunction with the feedback received while moving, to determine the strategy of playing the game. We can more formally extend the definition of views to encompass this additional amount of information as

$$\mathbf{v}_i(t) = \big(\phi_i(0), \big(t_{i,1}, \phi(t_{i,1})\big), \big(t_{i,2}, \phi(t_{i,2})\big), \ldots, \big(t_{i,j}, \phi(t_{i,j})\big)\big).$$

Gains and Benefits Players receive benefit equal to the number of time units for which they are the most recent mover, minus the cost of making their moves. We denote the

cost of a move for player i by k_i; it is important for our modeling goals that these costs could be quite different.

Player i's total *gain* G_i in a given game (before subtracting off the cost of moves) is just the integral of C_i:

$$G_i(t) = \int_0^t C_i(x)\,dx.$$

Thus $G_i(t)$ denotes the total amount of time that player i has owned the resource (controlled the game) from the start of the game up to time t, and

$$G_0(t) + G_1(t) = t.$$

The *average gain rate* for player i is defined as

$$\gamma_i(t) = G_i(t)/t;$$

so that $\gamma_i(t)$ is the fraction of time that player i has been in control of the game up to time t. Thus, for all $t > 0$:

$$\gamma_0(t) + \gamma_1(t) = 1.$$

We let $B_i(t)$ denote *player i's net benefit* up to time t; this is the gain (total possession time) minus the cost of player i's moves so far:

$$B_i(t) = G_i(t) - k_i n_i(t).$$

We also call $B_i(t)$ the *score* of player i at time t. The maximum benefit or score $B_0(t) = t - k_0$ for player 0 would be obtained if neither player moved again after player 0 took control at time $t = 0$.

We let $\beta_i(t)$ denote *player i's average benefit rate* up to time t:

$$\begin{aligned}\beta_i(t) &= B_i(t)/t \\ &= \gamma_i(t) - k_i\alpha_i(t);\end{aligned}$$

this is equal to the fraction of time the resource has been owned by player i, minus the cost rate for moving.

In a given game, we define player i's *asymptotic benefit rate* or simply *benefit* as

$$\beta_i = \liminf_{t\to\infty} \beta_i(t).$$

We use lim inf since $\beta_i(t)$ may not have limiting values as t increases to infinity.

An alternative standard and reasonable approach for summarizing an infinite sequence of gains and losses would be to use a *discount rate* $\lambda < 1$ so that a unit gain at future time t is worth only λ^t. Our current approach is simpler when at least one of the players is non-adaptive, so we shall omit consideration of the alternative approach here.

While we have defined the benefit for a given game instance, it is useful to extend the notion over strategies. Let S_i be the strategy of player i, for $i \in \{0, 1\}$. Then the benefit

of player i for the FLIPIT game given by strategies (S_0, S_1) (denoted $\beta_i(S_0, S_1)$) is defined as the expectation of the benefit achieved in a game instance given by strategies (S_0, S_1) (the expectation is taken over the coin flips of S_0 and S_1).

Game-Theoretic Definitions We denote by $\mathsf{FlipIt}(\mathcal{C}_0, \mathcal{C}_1)$ the FLIPIT game in which player i chooses a strategy from class $\mathcal{C}_i$, for $i \in \{0, 1\}$. For a particular choice of strategies $S_0 \in \mathcal{C}_0$ and $S_1 \in \mathcal{C}_1$, the benefit of player i is defined as above. In our game, benefits are equivalent to the notion of *utility* used in game theory.

Using terminology from the game theory literature:

- A strategy $S_0 \in \mathcal{C}_0$ is *strongly dominated* for player 0 in game $\mathsf{FlipIt}(\mathcal{C}_0, \mathcal{C}_1)$ if there exists another strategy $S'_0 \in \mathcal{C}_0$ such that
$$\beta_0(S_0, S_1) < \beta_0\big(S'_0, S_1\big), \quad \forall S_1 \in \mathcal{C}_1.$$
- A strategy $S_0 \in \mathcal{C}_0$ is *weakly dominated* for player 0 in game $\mathsf{FlipIt}(\mathcal{C}_0, \mathcal{C}_1)$ if there exists another strategy $S'_0 \in \mathcal{C}_0$ such that
$$\beta_0(S_0, S_1) \leq \beta_0\big(S'_0, S_1\big), \quad \forall S_1 \in \mathcal{C}_1,$$
with at least one S_1 for which the inequality is strict.
- A strategy S_0 is *strongly dominant* for player 0 in game $\mathsf{FlipIt}(\mathcal{C}_0, \mathcal{C}_1)$ if
$$\beta_0(S_0, S_1) > \beta_0\big(S'_0, S_1\big), \quad \forall S'_0 \in \mathcal{C}_0,\ \forall S_1 \in \mathcal{C}_1.$$
- A strategy S_0 is *weakly dominant* for player 0 in game $\mathsf{FlipIt}(\mathcal{C}_0, \mathcal{C}_1)$ if
$$\beta_0(S_0, S_1) \geq \beta_0\big(S'_0, S_1\big), \quad \forall S'_0 \in \mathcal{C}_0,\ \forall S_1 \in \mathcal{C}_1.$$

Similar definitions can be given for player 1 since the game is fully symmetric.

There is an implicit assumption in the game theory literature that a rational player does not choose to play a strategy that is strongly dominated by other strategies. Therefore, iterative elimination of strongly dominated strategies for both players is a standard technique used to reduce the space of strategies available to each player (see, for instance, the book by Myerson [18]). We denote by $\mathsf{FlipIt}^*(\mathcal{C}_0, \mathcal{C}_1)$ the *residual* FLIPIT game consisting of surviving strategies after elimination of strongly dominated strategies from classes $\mathcal{C}_0$ and $\mathcal{C}_1$. A rational player will always choose a strategy from the residual game.

A *Nash equilibrium* for the game $\mathsf{FlipIt}(\mathcal{C}_0, \mathcal{C}_1)$ is a pair of strategies $(S_0, S_1) \in \mathcal{C}_0 \times \mathcal{C}_1$ such that

$$\begin{aligned}
\beta_0(S_0, S_1) &\geq \beta_0\big(S'_0, S_1\big), \quad \forall S'_0 \in \mathcal{C}_0; \\
\beta_1(S_0, S_1) &\geq \beta_1\big(S_0, S'_1\big), \quad \forall S'_1 \in \mathcal{C}_1.
\end{aligned}$$

4. Renewal Games

In this section, we analyze FLIPIT over a simple but interesting class of strategies, those in which a player "renews" play after every move, in the sense of selecting each

interval between moves independently and uniformly at random from a fixed distribution. Such a player makes moves without regard to his history of previous moves, and also without feedback about his opponent's moves. The player's next move time depends only on the time he moved last.

In this class of strategies, called *renewal strategies*, the intervals between each player's move times are generated by a *renewal process*, a well-studied type of stochastic process. (See, for example, the books of Feller [8], Ross [23] and Gallager [9] that offer a formal treatment of renewal theory.) We are interested in analyzing the FLIPIT game played with either renewal or periodic strategies (with random phases as defined in Sect. 4.1), finding Nash equilibria for particular instances of the game, determining strongly dominated and dominant strategies (when they exist), and characterizing the residual FLIPIT game.

These questions turn out to be challenging. While simple to characterize, renewal strategies lead to fairly complex mathematical analysis. Our main result in this section is that renewal strategies are strongly dominated by periodic strategies (against an opponent playing also a renewal or periodic strategy), and the surviving strategies in the residual FLIPIT game are the periodic strategies. In addition, in the subclass of renewal strategies with a fixed rate of play, the periodic strategy is the strongly dominant one.

We also analyze in depth the FLIPIT game in which both players employ a periodic strategy with a random phase. We compute the Nash equilibria for different conditions on move costs, and discuss the choice of the rate of play when the attacker receives feedback according to NA and RP definitions given in Sect. 3.

4.1. *Playing Periodically*

We start this section by analyzing a very simple instance of the FLIPIT game. We consider a non-adaptive continuous game in which both players employ a periodic strategy with a random phase. If the strategy of one player is completely deterministic, then the opponent has full information about the deterministic player's move times and therefore can schedule his moves to control the resource at all times. We introduce phase randomization into the periodic strategy for this reason, selecting the time of the first move uniformly at random from some interval.

More specifically, a *periodic strategy with random phase* is characterized by the fixed interval between consecutive moves, denoted δ. We assume that the first move called the *phase* move is chosen uniformly at random in interval $[0, \delta]$. The average play rate (excluding the first move) is given by $\alpha = 1/\delta$. We denote by P_α the periodic strategy with random phase of rate α and $\mathcal{P}$ the class of all periodic strategies with random phases:

$$\mathcal{P} = \{P_\alpha | \alpha > 0\}.$$

An interesting game to consider is one in which player i employs strategy P_{α_i}. Let $\delta_i = 1/\alpha_i$ be the period, and k_i the move cost of player i, for $i \in \{0, 1\}$. This game is graphically depicted in Fig. 3.

As an initial exercise, and to get acquainted to the model and definitions, we start by computing benefits for both players in the periodic game.

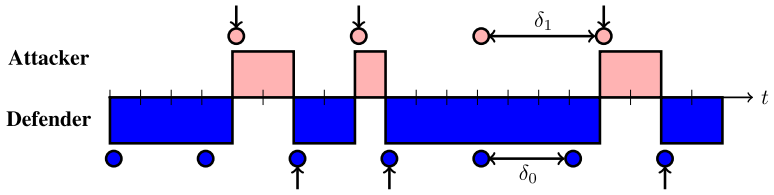

Fig. 3. The FLIPIT game with both players playing periodically with respective periods δ_0 and δ_1.

Computing Benefits For the periodic game defined above, the benefits of both players depend on rates of play α_0 and α_1. We thus denote the benefit of player i by $\beta_i(\alpha_0, \alpha_1)$ (this is the expected benefit computed over the random phase selections of the two players as in the definition given in Sect. 3). To compute both players' benefits, we consider two cases:

Case 1: $\alpha_0 \geq \alpha_1$ (The defender plays as least as fast as the attacker.)
Let $r = \delta_0/\delta_1$. The intervals between two consecutive defender's moves have length δ_0. Consider a given defender move interval. The probability over the attacker's phase selection that the attacker moves in this interval is r. Given that the attacker moves within the interval, he moves exactly once within the interval (since $\delta_0 \leq \delta_1$) and his move is distributed uniformly at random. Thus the expected period of attacker control within the interval is $r/2$. We can therefore express the players' benefits as

$$\beta_0(\alpha_0, \alpha_1) = 1 - \frac{r}{2} - k_0\alpha_0 = 1 - \frac{\alpha_1}{2\alpha_0} - k_0\alpha_0;$$
$$\beta_1(\alpha_0, \alpha_1) = \frac{r}{2} - k_1\alpha_1 = \frac{\alpha_1}{2\alpha_0} - k_1\alpha_1.$$

Case 2: $\alpha_0 \leq \alpha_1$ (The defender plays no faster than the attacker.)
Similar analysis yields benefits

$$\beta_0(\alpha_0, \alpha_1) = \frac{\alpha_0}{2\alpha_1} - k_0\alpha_0;$$
$$\beta_1(\alpha_0, \alpha_1) = 1 - \frac{\alpha_0}{2\alpha_1} - k_1\alpha_1.$$

Nash Equilibria As a second step, we are interested in finding Nash equilibria, points for which neither player will increase his benefit by changing his rate of play. More formally, a Nash equilibrium for the periodic game is a point (α_0^*, α_1^*) such that the defender's benefit $\beta_0(\alpha_0, \alpha_1^*)$ is maximized at $\alpha_0 = \alpha_0^*$ and the attacker's benefit $\beta_1(\alpha_0^*, \alpha_1)$ is maximized at $\alpha_1 = \alpha_1^*$.

To begin with, some useful notation. We denote by $\mathsf{opt}_0(\alpha_1)$ the set of values (rates of play α_0) that optimize the benefit of the defender for a fixed rate of play α_1 of the attacker. Similarly, we denote by $\mathsf{opt}_1(\alpha_0)$ the set of values (rates of play α_1) that optimize the benefit of the attacker for a fixed rate of play α_0 of the defender. The following theorem specifies Nash equilibria for the periodic game and is proven in Appendix A.

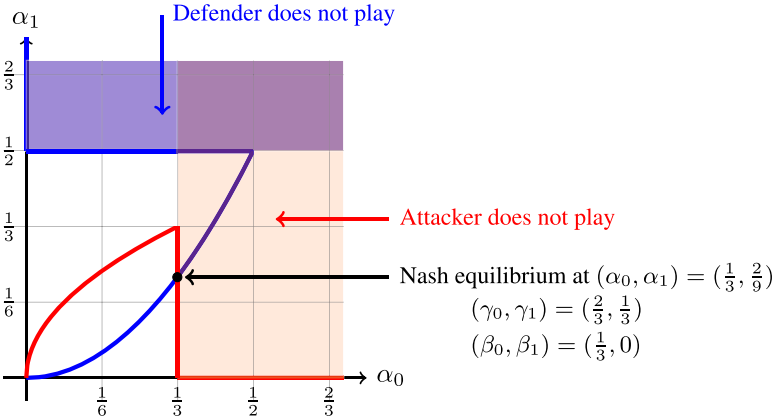

Fig. 4. Nash equilibrium for game with periodic attacker and periodic defender and move costs $k_0 = 1$ and $k_1 = 1.5$. The optimal strategy opt_0 of the defender is depicted in *blue* (*dark gray* in grayscale), and the optimal strategy of the attacker opt_1 is depicted in *red* (*light gray* in grayscale).

Theorem 1. *The* FLIPIT *game* $\mathsf{FlipIt}(\mathcal{P}, \mathcal{P})$ *in which both players employ periodic strategies with random phases has the following Nash equilibria*:

$$k_0 < k_1 \quad \alpha_0^* = \frac{1}{2k_1}, \quad \alpha_1^* = \frac{k_0}{2k_1^2};$$

$$k_0 = k_1 \quad \alpha_0^* = \alpha_1^* = \frac{1}{2k_0};$$

$$k_0 > k_1 \quad \alpha_0^* = \frac{k_1}{2k_0^2}, \quad \alpha_1^* = \frac{1}{2k_0}.$$

We illustrate the case for $k_0 = 1$ and $k_1 = 1.5$ in Fig. 4. We depict in the figure the two curves opt_0 and opt_1 and the unique Nash equilibrium at their intersection. (See the proof of Theorem 1 for an explicit formula for opt_0 and opt_1.) The highlighted blue (dark gray in grayscale) and red (light gray in grayscale) regions correspond to cases for which it is optimal for the defender, and the attacker, respectively, not to move at all (when $\mathsf{opt}_0(\alpha_1) = 0$ or $\mathsf{opt}_1(\alpha_0) = 0$, respectively).

Parameter Choices Finally, we would like to show the impact the parameter choice has on the benefits achieved by both players in the game. The periodic strategy with random phase for player i is uniquely determined by choosing rate of play α_i. Again, in the non-adaptive version of the game neither player has any information about the opponent's strategy (expect for the move costs, which we assume are known at the beginning of the game). Therefore, players need to decide upon a rate of play based only on knowledge of the move costs.

We show in Fig. 5 the benefits of both players as a function of play rates α_0 and α_1. We plot these graphs for different values of move costs k_0 and k_1. Darker shades correspond to higher values of the benefits, and white squares correspond to negative or zero benefit.

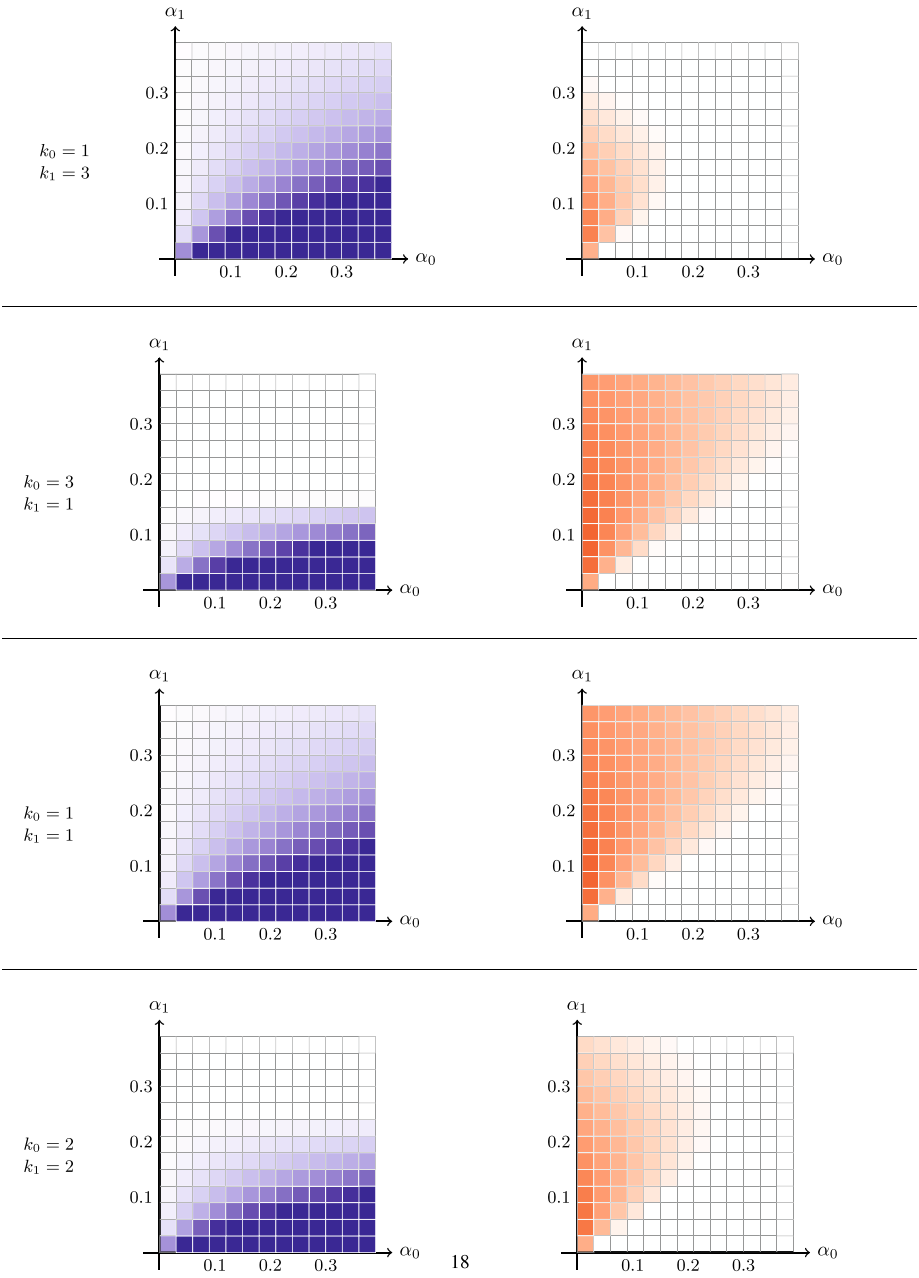

Fig. 5. Defender's (*left*, depicted in *blue*, *dark gray* in gray scale) and attacker's (*right*, depicted in *red*, *light gray* in gray scale) benefit for various move costs. Benefits are represented as a function of rates of play α_0 and α_1. *Darker shades* correspond to higher values of the benefits, and *white squares* correspond to negative or zero benefit.

The top four graphs clearly demonstrate the advantage the player with a lower move cost has on its opponent. Therefore, an important lesson derived from FLIPIT is that by lowering the move cost a player can obtain higher benefit, no matter how the opponent plays! The bottom four graphs compare the benefits of both players for equal move costs (set at 1 and 2, respectively). The graphs show that when move costs are equal, both players can achieve similar benefits, and neither has an advantage over the other. As expected, the benefits of both players are negatively affected by increasing the move costs. We also notice that in all cases playing too fast results eventually in negative benefit (depicted by white squares), as the cost of the moves exceeds that of the gain achieved from controlling the resource.

Achieve Good Benefits in Any Game Instance We notice that if both players play periodically with random phase, the choice of the random phase could result in gain 0 for one of the players (in the worst case). This happens if the opponent always plays "right after" and essentially controls the resource at all times. The benefits computed above are averages over many game instances, each with a different random phase.

We can, nevertheless, guarantee that expected benefits are achieved in *any game instance* by "re-phasing" during the game. The idea is for a player to pick a new random phase at moves with indices $a_1, a_2, \ldots, a_n, \ldots$. It turns out that for $a_j = j^2$, the expectation of $\beta_i(t)$ (the benefit up to time t) over all random phases converges to $\beta_i(\alpha_0, \alpha_1)$ as $t \to \infty$ and by the law of large numbers its standard deviation converges to 0 as $t \to \infty$. This can be used to show that each game instance achieves benefits $\beta_0(\alpha_0, \alpha_1)$ and $\beta_1(\alpha_0, \alpha_1)$ computed above.

Forcing the Attacker to Drop Out We make the observation that if the defender plays extremely fast (periodic with rate $\alpha_0 > 1/2k_1$), the attacker's strongly dominant non-adaptive strategy is to drop out of the game. The reason is that in each interval between defender's consecutive moves (of length $\delta_0 = 1/\alpha_0 < 2k_1$), the attacker can control the resource on average at most half of the time, resulting in gain strictly less than k_1. However, the attacker has to spend k_1 for each move, and therefore the benefit in each interval is negative.

We can characterize the residual FLIPIT game in this case:

$$\mathsf{FlipIt}^*(P_{\alpha_0}, \mathcal{N}) = \mathsf{FlipIt}(P_{\alpha_0}, \Phi), \quad \forall \alpha_0 > \frac{1}{2k_1}.$$

4.2. *General Renewal Strategies*

We now analyze the general case of non-adaptive renewal games. In this class of games both players's strategies are non-adaptive and the inter-arrival times between each player's moves are produced by a renewal process (characterized by a fixed probability distribution). We start by presenting some well-known results from renewal theory that will be useful in our analysis and then present a detailed analysis of the benefits achieved in the renewal game.

Renewal Theory Results Let $\{X_j\}_{j \geq 0}$ be independent and identically distributed random variables chosen from a common probability density function f. Let F be the corresponding cumulative distribution function. $\{X_j\}_{j \geq 0}$ can be interpreted as inter-arrival

times between events in a renewal process: the nth event arrives at time $S_n = \sum_{j=0}^{n} X_j$. Let $\mu = E[X_j]$, for all $j \geq 0$.

A renewal process generated by probability density function f is called *arithmetic* if inter-arrival times are all integer multiples of a real number d. The span of an arithmetic distribution is the largest d for which this property holds. A renewal process with $d = 0$ is called *non-arithmetic*.

For a random variable X given by a probability density function f, corresponding cumulative distribution F, and expected value $\mu = E[X]$, we define the *size-bias density function* as

$$f^*(z) = \frac{1 - F(z)}{\mu}$$

and the *size-bias cumulative distribution function* as

$$F^*(z) = \frac{\int_{x=0}^{z} (1 - F(x))\, dx}{\mu}.$$

The *age function* $Z(t)$ of a renewal process is defined at time interval t as the time since the last arrival. Denote by $f_{Z(t)}$ and $F_{Z(t)}$ the age density and cumulative distribution functions, respectively. The following lemma [8] states that as t goes to ∞, the age density and cumulative distribution functions converge to the size-bias density and cumulative distribution functions, respectively:

Lemma 1. *For a non-arithmetic renewal process given by probability density function f, the age density and cumulative distribution functions converge as*

$$\lim_{t \to \infty} f_{Z(t)}(z) = f^*(z);$$
$$\lim_{t \to \infty} F_{Z(t)}(z) = F^*(z).$$

Playing FLIPIT *with Renewal Strategies* A *renewal strategy*, as explained above, is one in which a player's moves are generated by a renewal process. In a non-arithmetic renewal strategy, the player's moves are generated by a non-arithmetic renewal process. We denote by R_f the renewal strategy generated by a non-arithmetic renewal process with probability density function f, and by $\mathcal{R}$ the class of all non-arithmetic renewal strategies:

$$\mathcal{R} = \{R_f \mid f \text{ is a non-arithmetic probability density function}\}.$$

Here we consider the FLIPIT game in which both players employ non-arithmetic renewal strategies: player i uses strategy R_{f_i}. Let us denote the intervals between defender's moves as $\{X_j\}_{j \geq 0}$ and the intervals between attacker's moves as $\{Y_j\}_{j \geq 0}$. $\{X_j\}_{j \geq 0}$ are identically distributed random variables chosen independently from probability density function f_0 with average μ_0, while $\{Y_j\}_{j \geq 0}$ are independent identically distributed random variables chosen from probability density function f_1 with average μ_1. Let $\alpha_i = 1/\mu_i$ be the rate of play of player i.

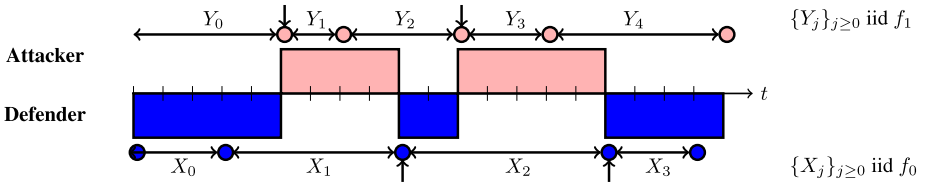

Fig. 6. The FLIPIT game with both players playing with renewal strategies. The defender and attacker play with inter-arrival times iid from probability density functions f_0 and f_1, respectively.

We denote the corresponding cumulative distribution functions as F_0 and F_1. Since both renewal processes are non-arithmetic, we can apply Lemma 1 and infer that the age density and cumulative distribution functions of both processes converge. Denote by f_0^*, f_1^* and F_0^*, F_1^* the size-bias density and cumulative distribution functions for the two distributions.

A graphical representation of this game is given in Fig. 6.

In this context, the following theorem (proven in Appendix A) gives a formula for the players' benefits:

Theorem 2. *In the non-adaptive renewal* FLIPIT *game* $\mathsf{FlipIt}(\mathcal{R}, \mathcal{R})$, *the players' benefits for strategies* $(R_{f_0}, R_{f_1}) \in \mathcal{R} \times \mathcal{R}$ *are*

$$\beta_0(R_{f_0}, R_{f_1}) = \int_{x=0}^{\infty} f_1^*(x) F_0^*(x)\, dx - k_0\alpha_0;$$
$$\beta_1(R_{f_0}, R_{f_1}) = \int_{x=0}^{\infty} f_0^*(x) F_1^*(x)\, dx - k_1\alpha_1.$$

Interestingly, Theorem 2 can be generalized to strategies from class $\mathcal{R} \cup \mathcal{P}$. We observe that for strategy $P_\alpha \in \mathcal{P}$, the density and cumulative distribution of the underlying periodic distribution are

$$f(x) = \begin{cases} 1, & x = \frac{1}{\alpha}, \\ 0, & x \neq \frac{1}{\alpha}, \end{cases} \qquad F(x) = \begin{cases} 0, & x < \frac{1}{\alpha}, \\ 1, & x \geq \frac{1}{\alpha}. \end{cases}$$

The size-bias density and cumulative distribution functions for the periodic distribution are

$$f^*(x) = \begin{cases} \alpha, & x < \frac{1}{\alpha}, \\ 0, & x \geq \frac{1}{\alpha}, \end{cases} \qquad F^*(x) = \begin{cases} \alpha x, & x < \frac{1}{\alpha}, \\ 1, & x \geq \frac{1}{\alpha}. \end{cases}$$

Based on these observations, we obtain the following theorem, whose proof is given in Appendix A.

Theorem 3. *In the non-adaptive* FLIPIT *game* $\mathsf{FlipIt}(\mathcal{R} \cup \mathcal{P}, \mathcal{R} \cup \mathcal{P})$, *the players' benefits for strategies* $(S_0, S_1) \in (\mathcal{R} \cup \mathcal{P}) \times (\mathcal{R} \cup \mathcal{P})$ *are*

$$\beta_0(S_0, S_1) = \int_{x=0}^{\infty} f_1^*(x) F_0^*(x)\, dx - k_0\alpha_0;$$
$$\beta_1(S_0, S_1) = \int_{x=0}^{\infty} f_0^*(x) F_1^*(x)\, dx - k_1\alpha_1.$$

4.3. *The Residual Renewal Game*

We are now able to completely analyze the FLIPIT game with strategies in class $\mathcal{R} \cup \mathcal{P}$. We show first that for a fixed rate of play α_i of player i the periodic strategy with random phase is the strongly dominant strategy for player i (or, alternatively, all non-arithmetic renewal strategies are strongly dominated by the periodic strategy with the same rate). As a consequence, after both players iteratively eliminate the strongly dominated strategies, the surviving strategies in the residual game $\mathsf{FlipIt}^*(\mathcal{R} \cup \mathcal{P}, \mathcal{R} \cup \mathcal{P})$ are the periodic strategies.

For a fixed $\alpha > 0$, we denote by $\mathcal{R}_\alpha$ the class of all non-arithmetic renewal strategies of fixed rate α, and by $\mathcal{P}_\alpha$ the class of all periodic strategies of rate α. As there exists only one periodic strategy of rate α, $\mathcal{P}_\alpha = \{P_\alpha\}$. The proof of the following theorem is given in Appendix A.

Theorem 4.

1. *For a fixed rate of play* α_0 *of the defender, strategy* P_{α_0} *is strongly dominant for the defender in game* $\mathsf{FlipIt}(\mathcal{R}_{\alpha_0} \cup \mathcal{P}_{\alpha_0}, \mathcal{R} \cup \mathcal{P})$. *A similar result holds for the attacker.*
2. *The surviving strategies in the residual* FLIPIT *game* $\mathsf{FlipIt}(\mathcal{R} \cup \mathcal{P}, \mathcal{R} \cup \mathcal{P})$ *are strategies from class* $\mathcal{P}$:

$$\mathsf{FlipIt}^*(\mathcal{R} \cup \mathcal{P}, \mathcal{R} \cup \mathcal{P}) = \mathsf{FlipIt}(\mathcal{P}, \mathcal{P}).$$

Remark. While we have proved Theorem 4 for the non-adaptive case, the result also holds if one or both players receive information at the beginning of the game according to the RP definition. This follows immediately from the proof of Theorem 4 given in Appendix A.

4.4. *Renewal Games with an RP Attacker*

In the case in which one player (assume the attacker) receives feedback according to RP (i.e., knows the rate of play of the defender), we can provide more guidance to both players on choosing the rates of play that achieve maximal benefit. In particular, we consider a scenario in which the attacker finds out the rate of play of the defender before the game starts, and we assume that the attacker plays rationally in the sense that he tries to maximize his own benefit: for a fixed rate of play α_0 of the defender, the attacker chooses to play with rate $\alpha_1^* = \mathsf{argmax}\ \beta_1(\alpha_0, \cdot)$. Under this circumstance, the

defender can also determine for each fixed rate of play α_0, the rate of play α_1^* of a rational attacker. In a pre-game strategy selection phase, then, a rational defender selects and announces the rate of play α_0^* that maximizes her own benefit: $\alpha_0^* = \mathsf{argmax}\, \beta_0(\cdot, \alpha_1^*)$. The following theorem provides exact values of α_0^* and α_1^*, as well as maximum benefits achieved by playing at these rates.

Theorem 5. *Consider the periodic* FLIPIT *game with NA defender and RP attacker. Assume that the attacker always chooses his rate of play to optimize his benefit*: $\alpha_1^* = \mathsf{argmax}\, \beta_1(\alpha_0, \cdot)$ *given a fixed rate of play of the defender* α_0, *and the defender chooses her rate of play* α_0^* *that achieves optimal benefit*: $\alpha_0^* = \mathsf{argmax}\, \beta_0(\cdot, \alpha_1^*)$. *Then*:

1. *For* $k_1 < (4 - \sqrt{12})k_0$ *the rates of play that optimize each player's benefit are*

$$\alpha_0^* = \frac{k_1}{8k_0^2}; \qquad \alpha_1^* = \frac{1}{4k_0}.$$

The maximum benefits for the players are

$$\beta_0\big(\alpha_0^*, \alpha_1^*\big) = \frac{k_1}{8k_0}; \qquad \beta_1\big(\alpha_0^*, \alpha_1^*\big) = 1 - \frac{k_1}{2k_0}.$$

2. *For* $k_1 \geq (4 - \sqrt{12})k_0$ *the rates of play that optimize each player's benefit are* ($\downarrow$ *denotes convergence from above*):

$$\alpha_0^* \downarrow \frac{1}{2k_1}; \qquad \alpha_1^* = 0.$$

The maximum benefits for the players are ($\uparrow$ *denotes convergence from below*):

$$\beta_0\big(\alpha_0^*, \alpha_1^*\big) \uparrow 1 - \frac{k_0}{2k_1}; \qquad \beta_1\big(\alpha_0^*, \alpha_1^*\big) = 0.$$

5. LM Attacker

In the previous section, we analyzed a non-adaptive instance of FLIPIT in which both players employ (non-arithmetic) renewal or periodic strategies with a random phase. We consider now a FLIPIT game in which the defender is still playing with a renewal or periodic strategy, but the attacker is more powerful and receives feedback during the game. In particular, the attacker finds out the exact time of the defender's last move every time he moves, and is LM according to our definition in Sect. 3. It is an immediate observation that for a defender playing with a renewal strategy, an LM attacker is just as powerful as an FH one.

We believe that this models a realistic scenario in which attackers have access to more information than defenders. For instance, attackers can determine with high confidence the time intervals when the defender changes her password by trying to use a compromised password at different time intervals and observe its expiration time. In this model, our goals are two-fold: first, we would like to understand which renewal strategies chosen by the defender achieve high benefit against an adaptive attacker; second, we are

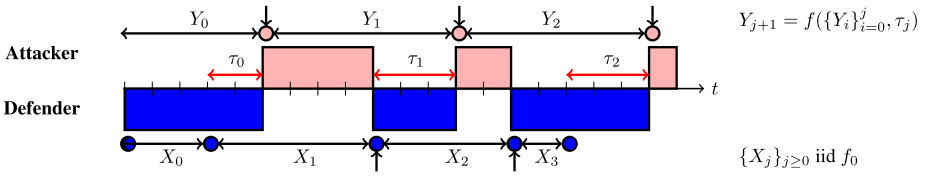

Fig. 7. The FLIPIT game with renewal defender and LM attacker. The defender plays with inter-arrival times iid from distribution f_0. The attacker finds out upon moving the interval since defender's last move τ_j.

interested in finding good adaptive strategies for the attacker, given a fixed distribution for the defender.

We analyze this game for the defender playing with periodic and exponential distributions. The periodic defender strategy is a poor choice against an LM attacker, as the defender's benefit is negative (unless the defender can afford to play extremely fast and force the attacker to drop out of the game). For the defender playing with an exponential distribution, we prove that the attacker's strongly dominant strategy is periodic play (with a rate depending on the defender's rate of play). We also define a new distribution called *delayed exponential*, in which the defender waits for a fixed interval of time before choosing an exponentially distributed interval until her next move. We show experimentally that for some parameter choices, the delayed-exponential distribution results in increased benefit for the defender compared to exponential play.

5.1. *Game Definition*

In this version of the game, the defender plays with either a non-arithmetic renewal strategy from class $\mathcal{R}$ or with a periodic strategy with random phase from class $\mathcal{P}$. Hence, the inter-arrival times between defender's consecutive moves are given by independent and identically distributed random variables $\{X_j\}_{j\geq 0}$ from probability density function f_0. (In case the defender plays with a periodic strategy with a random phase, X_0 has a different distribution that $\{X_j\}_{j\geq 1}$.)

The attacker is LM and his inter-arrival move times are denoted by $\{Y_j\}_{j\geq 0}$. At its jth move, the attacker finds out the exact time since the defender's last move, denoted τ_j. The attacker determines the interval until his next move Y_{j+1} taking into consideration the history of his own previous moves (given by $\{Y_i\}_{i=0}^{j}$), as well as τ_j. Without loss of generality, the attacker does not consider times $\{\tau_i\}_{i<j}$ when determining interval Y_{j+1} for the following reason. The defender "restarts" its strategy after each move, and thus knowledge of all the history $\{\tau_i\}_{i<j}$ does not help the attacker improve his strategy. This game is depicted graphically in Fig. 7.

Observation We immediately observe that an LM attacker is as powerful as an FH attacker against a renewal or periodic defender:

$$\mathsf{FlipIt}^*(\mathcal{R}\cup\mathcal{P}, \mathcal{A}_{\mathsf{LM}}) = \mathsf{FlipIt}^*(\mathcal{R}\cup\mathcal{P}, \mathcal{A}_{\mathsf{FH}}).$$

The reason is that an FH attacker, while receiving more feedback during the game than an LM one, will only make use of the last defender's move in determining his

strategy. Information about previous defender's moves does not provide any additional advantage, as the defender's moves are independent of one another.

5.2. *Defender Playing Periodically*

We start by analyzing the simple game in which the defender plays periodically with period δ_0 and a random phase. The attacker finds out after his first move after δ_0 the exact phase of the distribution (the time of the defender's first move denoted X_0). As such, the attacker knows the exact time of all future moves by the defender: $X_0 + \delta_0, X_0 + 2\delta_0, \ldots$. The strongly dominant strategy for the attacker is then to play "immediately after" the defender, and control the resource all the time after the first move. However, when the defender plays sufficiently fast, it can force the attacker to drop out of the game. We distinguish three cases based on the defender's rate of play.

Case 1: $\alpha_0 < 1/k_1$

It is quite easy to see that in this case the strongly dominant attacker strategy is to play right after the defender. Therefore, the defender only controls the resource for a small time interval at the beginning of the game. We can prove that the gain of the defender is asymptotically 0 ($\gamma_0 = 0$) and that of the attacker converges to 1 ($\gamma_1 = 1$). The benefits of the players are

$$\beta_0 = -k_0/\delta_0; \qquad \beta_1 = 1 - k_1/\delta_0.$$

Case 2: $\alpha_0 > 1/k_1$

In this case, the attacker would have to spend at least as much budget on moving as the gain obtained from controlling the resource, resulting in negative or zero benefit. Therefore, the defender forces the attacker to drop out of the game. The following observation is immediate:

Forcing the Attacker to Drop Out If the defender plays periodically with rate $\alpha_0 > 1/k_1$, an LM attacker's strongly dominant strategy is not playing at all. Therefore all LM attacker strategies are strongly dominated by the no-play strategy against a defender playing with strategy P_{α_0}, with $\alpha_0 > 1/k_1$, and we can characterize the residual FLIPIT game in this case:

$$\mathsf{FlipIt}^*(P_{\alpha_0}, \mathcal{A}_{\mathsf{FH}}) = \mathsf{FlipIt}(P_{\alpha_0}, \Phi), \quad \forall \alpha_0 > 1/k_1.$$

The benefit achieved by the defender is $\beta_0 = 1 - k_0/k_1$.

It is important to note that this holds for FH attackers, as well, since we have discussed that the residual games for LM and FH attackers against a $\mathcal{R} \cup \mathcal{P}$ defender result in the same set of surviving strategies.

Case 3: $\alpha_0 = 1/k_1$

Two attacker strategies are (weakly) dominant in this case: play right after the defender and no play at all. Both strategies achieve zero benefit for the attacker. The defender yields benefit either $\beta_0 = -k_0/k_1$ (for the attacker playing after the defender) or $\beta_0 = 1 - k_0/k_1$ (for the no-play attacker strategy).

This demonstrates that the periodic strategy with a random phase is a poor choice for the defender (unless the defender plays extremely fast), resulting always in negative benefit. This motivates us to search for other distributions that will achieve higher benefit for the defender against an LM attacker.

5.3. *Defender Playing Exponentially*

We now consider the game in which the defender plays with an exponential distribution with rate λ, denoted E_λ. The exponential distribution is *memoryless*: in a renewal process with inter-arrival time between events distributed exponentially, the time of the next event is independent of the time elapsed since the last event. Intuitively, an LM attacker that knows the time of the defender's last move has no advantage over a non-adaptive attacker in predicting the time of the defender's next move. Accordingly, we can prove that strategies from the LM class are strongly dominated by periodic strategies (and, hence, the attacker does not benefit from the additional feedback received during the game).

Theorem 6. *Suppose the defender plays exponentially with rate λ. Then the strongly dominant LM strategy is either*

1. *P_α for some $\alpha > 0$, the associated game* $\mathsf{FlipIt}(E_\lambda, P_\alpha)$ *yielding benefits*

$$\beta_0(E_\lambda, P_\alpha) = 1 - \frac{1 - e^{-\lambda\delta}}{\lambda\delta} - \lambda k_0,$$

$$\beta_1(E_\lambda, P_\alpha) = \frac{1 - e^{-\lambda\delta}}{\lambda\delta} - \frac{k_1}{\delta}, \text{ or}$$

2. *No play, the associated game* $\mathsf{FlipIt}(E_\lambda, \Phi)$ *yielding benefits $\beta_0(E_\lambda, P_\alpha) = 1 - \lambda k_0$ and $\beta_1(E_\lambda, P_\alpha) = 0$.*

The proof of Theorem 6 is given in Appendix B. The following theorem, also proven in Appendix B, provides a relationship between the period δ of the dominant LM attacker strategy and the rate λ of the defender's exponential distribution.

Theorem 7. *Assume that the defender plays with an exponential distribution E_λ with fixed rate λ. Then*:

1. *If $\lambda < 1/k_1$, the strongly dominant LM strategy is P_α with period $\delta = 1/\alpha$ satisfying equation*:

$$e^{-\lambda\delta}(1 + \lambda\delta) = 1 - \lambda k_1.$$

2. *If $\lambda \geq 1/k_1$, then the strongly dominant LM strategy for the attacker is not playing at all.*

Choosing the Defender's Rate of Play We have proved that by playing exponentially, the defender induces a strongly dominant attacker strategy consisting of periodic play (or no play at all). But what is the rate λ at which the defender should play to maximize her own benefit, assuming the defender plays exponentially?

An LM attacker can determine the rate of play of the defender since it has information about the exact times of the defender's moves. A rational attacker, one that aims to maximize his benefit, responds by playing the strongly dominant strategy, i.e., periodically with interval size δ defined by Theorem 7. A rate λ chosen by the defender therefore induces a period δ for the attacker that the defender can compute in advance. In a pre-game strategy selection, then, a defender committed to exponential play can determine the rate of play λ that maximizes her own benefit $\beta_0(E_\lambda, P_{1/\delta})$. The following theorem (proven in Appendix B) gives a formula for the rate λ that maximizes the defender's benefit assuming such pre-game rate selection against an attacker that chooses an optimal periodic strategy:

Theorem 8. *Assume that players' move costs are k_0 and k_1, the defender plays exponentially at rate λ, and an LM attacker chooses period δ as given by Theorem 7. Then*:

1. *If $k_0 \geq 0.854 \cdot k_1$, the maximum defender's benefit is achieved by playing at rate*:

$$\lambda = \frac{1-(1+z)e^{-z}}{k_1},$$

where z is the unique solution of equation

$$\frac{k_0}{k_1} = \frac{e^z - 1 - z}{z^3}.$$

For this choice of λ, the attacker's maximum benefit is achieved for period $\delta = \frac{z}{\lambda}$.

The benefits achieved by both players are

$$\beta_0(E_\lambda, P_{1/\delta}) = 1 - \frac{1-e^{-z}}{z} - \frac{k_0}{k_1}\left[1-(1+z)e^{-z}\right] \geq 1 - k_0/k_1,$$

$$\beta_1(E_\lambda, P_{1/\delta}) = \frac{1-e^{-z}}{z} - \frac{1-(1+z)e^{-z}}{z} = e^{-z}.$$

2. *If $k_0 < 0.854 \cdot k_1$, the maximum defender's benefit is achieved by playing at rate $\lambda = 1/k_1$ and the attacker's best response is not playing at all. The benefits achieved by both players are*

$$\beta_0(E_\lambda, P_{1/\delta}) = 1 - k_0/k_1; \qquad \beta_1(E_\lambda, P_{1/\delta}) = 0.$$

Remark. While we proved Theorems 6, 7 and 8 for a powerful LM attacker, they also hold for an RP attacker playing against an exponential defender (since the feedback received by an LM attacker is not helpful in improving his strategy). In addition, Theorem 6 holds for an NA or KS attacker (but Theorems 7 and 8 require that the attacker knows the rate of play of the defender). In conclusion, if the defender plays exponentially with a fixed rate, the strongly dominant strategy for an NA, RP, KS or LM attacker is periodic play. For an RP or LM attacker, the period δ of the dominant attacker strategy is given by Theorem 7, and the defender maximizes her benefit by choosing her rate of play according to Theorem 8.

5.4. *Defender Playing with Delayed-Exponential Distribution*

A natural question to consider now is if the exponential distribution is a dominant renewal strategy for the defender playing against an LM attacker. We provide experimental evidence that the defender's benefit achieved when playing exponentially can be further improved using a *delayed-exponential* distribution. When playing with a delayed-exponential distribution, the defender's inter-move times X_j are computed as $X_j = \Delta + X'_j$, for $j \geq 0$ where Δ is a fixed "wait" time interval and $\{X'_j\}_{j\geq 0}$ are iid from an exponential distribution of rate λ'.

Intuitively, we might expect suitably parameterized delayed-exponential play by the defender to achieve higher benefits for both players than a game pitting an exponential defender against a periodic attacker (as analyzed in the previous section). Thanks to the fixed interval Δ, a delayed-exponential defender playing with rate λ' will move less frequently than an exponential defender playing with rate $\lambda = \lambda'$. Conversely, if the attacker moves at time $t < \Delta$ after the defender's move, it is best off waiting for $\Delta - t$ before moving next (as the defender's moves are always separated by time at least Δ). As such, both players play less frequently and pay less for move costs than in the exponential vs. periodic case, resulting in higher benefits for both.

Theorem 9 (proven in Appendix B) characterizes the dominant LM strategy for the attacker when the defender plays with a delayed-exponential distribution. The next subsection provides experimental results affirming for a range of game cost structures (and suggesting as a general fact) that there is a delayed-exponential strategy yielding higher defender benefit than exponential play. Still the question of the existence and form of a dominant renewal strategy for the defender against an LM attacker remains open.

Theorem 9. *Assume that the defender plays with a delayed-exponential distribution given by parameters Δ and λ'. There exists a period δ' such that the strongly dominant LM strategy for the attacker is given by inter-move times $\{Y_j\}_{j\geq 0}$ with*

$$Y_{j+1} = \begin{cases} \Delta - \tau_j + \delta', & \text{if } \tau_j < \Delta, \\ \delta', & \text{if } \tau_j \geq \Delta. \end{cases}$$

5.5. *Simulation of Different Defender Strategies*

We conclude this section by presenting simulations of several defender strategies against attackers receiving different types of feedback during the game. The graph on the left in Fig. 8 depicts the *maximum* defender's benefit against a non-adaptive attacker with knowledge of both the strategy and the rate of play of the defender. We denote such an attacker by KS + RP. (It receives feedback according to the KS and RP notions defined in Sect. 3.) The graph on the right shows the maximum defender's benefit against an LM attacker. For the periodic and exponential renewal strategies, we plot the maximum benefit rates derived by our theoretical analysis (in Theorems 5 and 8). As we have not analytically derived exact benefit rates for the defender playing with a delayed-exponential distribution, we present experimental results for this distribution. (Precise analytic characterization remains an open problem.)

We have implemented the FLIPIT game in Python with time discretized at integer intervals of length one. Each round of the game is played for 100,000 time units. To

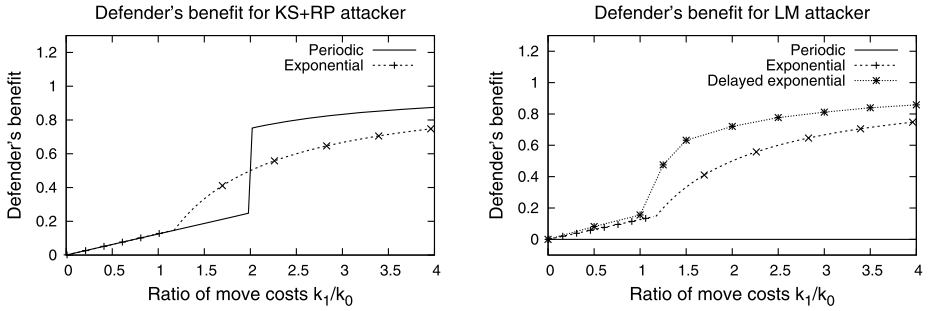

Fig. 8. Defender's benefit by playing different strategies for KS + RP and LM attackers.

compute the benefit for fixed attacker and defender strategies, we run 10 instances of the game and report the average benefit rates. For the results presented in Fig. 8, we have fixed the defender's move cost k_0 at 5, and varied k_1 to achieve a particular ratio k_1/k_0 shown on the horizontal axis (the exception is the point with $k_1/k_0 = 0.5$, for which we used $k_0 = 10$ and $k_1 = 5$).

The delayed-exponential distribution employed by the defender is completely characterized by two parameters: the wait interval Δ and the rate of play λ'. We exhaustively search the optimal parameters in the following space: $\Delta \in [0, 50]$ and $\lambda' \in [0.01, 0.2]$ (discretized at a step of 0.01). We determined the search space for the rate of play by looking at the optimal rate for the exponential distribution and varying it within at least a factor of two both above and below. As for the wait interval, we noticed that in virtually all cases, waiting for longer than 50 time units results in very low benefit for the defender. For fixed parameters Δ and λ', we experimentally search for an attacker strategy as given by Theorem 9. While the theorem gives us the strongly dominant attacker strategy against a delayed-exponential distribution, it does not specify the exact value of parameter δ'. We exhaustively search for δ' in a sufficiently large interval (at least 8–10 larger than the move cost k_1).

As we see in Fig. 8, by playing exponentially the defender achieves the same benefit for both KS + RP and LM attackers, as shown in Sect. 5.3. The periodic strategy, in contrast, is a poor choice for an LM attacker, as the maximum defender's benefit converges over time to 0.

Interestingly, the exponential strategy outperforms the periodic one against an KS + RP attacker, for cases in which $k_1/k_0 < 2$. This clearly demonstrates that Theorem 4 does not hold when one player has additional information about his opponent. (In this case, the attacker knows the distribution of the defender renewal strategy.) Another counterintuitive observation is that it is sometimes in a player's best interest to release information about his own strategy. In this example, if the attacker knows that the defender plays with an exponential distribution, the attacker chooses a different rate of play than in a renewal game in which the attacker has no additional information about the defender's strategy. This additional knowledge released to the attacker results in increased defender benefit (for particular choices of move costs).

We did not plot the benefit achieved by playing a delayed-exponential distribution against an KS + RP attacker, as we were not able to analyze this case and determine a

Table 2. Parameters and benefits for optimal exponential and delayed-exponential strategies in our experiments.

k_1/k_0	Exponential			Delayed exponential			
	λ	β_0	β_1	Δ	λ'	β_0	β_1
0.5	0.0062	0.062	0.76	11	0.007	0.082	0.75
1	0.0258	0.1262	0.5368	4	0.045	0.1558	0.4188
1.25	0.16	0.2	0	4	0.18	0.475	0
1.5	0.13	0.33	0	8	0.18	0.6324	0
2	0.1	0.5	0	12	0.17	0.7203	0
2.5	0.08	0.6	0	14	0.12	0.776	0
3	0.066	0.66	0	18	0.12	0.8113	0
3.5	0.057	0.72	0	20	0.09	0.839	0
4	0.05	0.75	0	24	0.09	0.858	0

dominant attacker strategy. Our intuition, though, is that if the attacker only knows the rate of play and the defender's distribution, but receives no additional feedback during the game, he cannot do better than playing with a periodic strategy (as he cannot predict the exact move times of the defender).

We also see in Fig. 8 that for an LM attacker, the delayed-exponential distribution outperforms the exponential strategy for carefully chosen parameters. The observed increase in the defender's benefit ranges between 15 % and 140 %. For a more thorough comparison, we present in Table 2 the exact parameters and benefits achieved for both the exponential and delayed-exponential strategies. We observe that the optimal wait intervals Δ increase with the attacker's move cost k_1, and that the optimal rate λ' is always higher than the optimal rate λ of the optimal exponential strategy with the same cost structure.

6. The Greedy Strategy for an LM Attacker

We now propose and analyze a natural strategy that we call *Greedy*. This is an LM attacker strategy for play against non-adaptive defender employing a known renewal strategy. We show that against an exponential defender strategy, Greedy results in the periodic strategy for the attacker, which is the strongly dominant one. We analyze the Greedy strategy for other renewal strategies of interest (in particular the periodic strategy with random phase and the uniform strategy). We also give an example of a renewal defender strategy for which Greedy *is not* a dominant strategy for the attacker.

6.1. *Algorithm Definition*

At a high level, the Greedy strategy optimizes the local benefit the attacker achieves between two consecutive moves. Assume that the defender employs a renewal strategy given by probability density function f_0 and cumulative distribution function F_0. The LM attacker receives information upon moving (at time t) about the time τ passed since the defender's last move. The Greedy strategy with inputs f_0 and τ outputs a time interval $\hat{z}$ until the attacker's next move that maximizes the local benefit achieved in interval $[t, t + \hat{z}]$.

In more detail, the first step of the Greedy strategy is to estimate the probability density $\hat{f}_0$ of the time till the defender's next move. Assume that the next attacker's move after time t is at time $t+z$. If the defender's next move after time t is at time $t+x$ (this happens with probability $\hat{f}_0(x)$) and $t+x < t+z$, then the attacker controls the resource a fraction x/z in interval $[t, t+z]$. On the other hand, if the defender's next move after time t is at time $t+x > t+z$, then the attacker controls the resource the entire interval $[t, t+z]$. The local benefit $L(z)$ in interval $[t, t+z]$ can then be computed based on these two cases. The interval $\hat{z}$ until the next attacker's move is the point that maximizes the local benefit $L(z)$.

The Greedy strategy proceeds as follows:

1. Compute $\hat{f}_0(x)$ defined as the probability density of the time until the defender's next move:

$$\hat{f}_0(x) = \frac{f_0(\tau + x)}{1 - F_0(\tau)}.$$

2. Define the local benefit:

$$L(z) = \frac{1}{z}\left[\int_{x=0}^{z} x\,\hat{f}_0(x)\,dx + z\int_{z}^{\infty} \hat{f}_0(x)\,dx - k_1\right].$$

3. Compute $\hat{z}$ that maximizes the local benefit $L(z)$.
4. If $L(\hat{z}) \geq 0$, move after $\hat{z}$ time units; otherwise do not move.

We now analyze the Greedy strategy for several defender strategies, including exponential, uniform and periodic renewal distributions. Our goal is to determine the attacker strategies given by the Greedy strategy and compare them with the dominant ones (in the case of periodic and exponential defender strategies, the strongly dominant LM attacker strategies are given by our analysis in Sect. 5).

6.2. *Defender Playing Exponentially Against Greedy*

Assume that f_0 is an exponential distribution of rate λ: $f_0(x) = \lambda e^{-\lambda x}$, $x > 0$. In this case,

$$\hat{f}_0(x) = \frac{\lambda e^{-\lambda(x+\tau)}}{e^{-\lambda\tau}} = \lambda e^{-\lambda x} = f_0(x).$$

Then $L(z)$ does not depend on the time since the defender's last move τ, and as such the optimum $\hat{z}$ is always the same when the Greedy strategy is invoked (after each attacker's move). The exception is the first attacker move, for which the attacker does not have information about the defender's previous move. Assuming that the attacker picks the first move uniformly at random from some fixed interval, the strategy given by the Greedy strategy is periodic with random phase.

We can now compute $L(z)$:

$$\begin{aligned} L(z) &= \frac{1}{z}\int_0^z x\lambda e^{-\lambda x}\,dx + \int_z^\infty \lambda e^{-\lambda x}\,dx - \frac{k_1}{z} \\ &= \frac{1-e^{-\lambda z}}{\lambda z} - e^{-\lambda z} + e^{-\lambda z} - \frac{k_1}{z} \\ &= \beta_1(E_\lambda, P_{1/z}). \end{aligned}$$

We observe that $L(z)$ is equal to the attacker's benefit in the game $\mathsf{FlipIt}(E_\lambda, P_{1/z})$. Therefore, the point $\hat{z}$ that optimizes $L(z)$ in step 3 of the Greedy strategy is the period δ of the strongly dominant periodic strategy given by Theorem 7. Thus the Greedy strategy finds the strongly dominant LM strategy against an exponential defender.

6.3. *Defender Playing Periodically Against Greedy*

Suppose that the defender plays periodically with random phase P_α given by period $\delta = 1/\alpha$. The attacker is LM and his move cost is $k_1 < \delta/2$.

Assume that the first move of the attacker is at time $t = \delta$, and the interval since the defender's last move is τ. Since the defender plays periodically, the next defender's move is scheduled at time $t + (\delta - \tau)$. To compute the local benefit $L(z)$ in interval $[t, t+z]$, we consider two cases:

1. $z \geq \delta - \tau$: The attacker controls the resource for time $\delta - \tau$ (until the defender's next move) in interval $[t, t+z]$. Therefore, benefit $L(z)$ can be written as

$$L(z) = \frac{\delta - \tau - k_1}{z}.$$

2. $z < \delta - \tau$: The attacker controls the resource the entire interval $[t, t+z]$ since the defender does not move in this interval. We can, hence, express $L(z)$ as

$$L(z) = \frac{z - k_1}{z} = 1 - \frac{k_1}{z}.$$

The local benefit $L(z)$ is strictly increasing on interval $[0, \delta - \tau]$. However, on the interval $[\delta - \tau, \infty]$, $L(z)$ could be either strictly increasing or strictly decreasing depending on whether $\tau > \delta - k_1$ or $\tau < \delta - k_1$. Since τ is uniformly distributed in interval $[0, \delta]$, the probability that $L(z)$ is strictly increasing on $[\delta - \tau, \infty]$ is equal to $k_1/\delta < 1/2$. Hence, with probability less than $1/2$ the optimum point $\hat{z}$ in step 3 of the Greedy strategy is ∞, and the attacker will stop playing after the first move.

However, with probability greater than $1/2$, $L(z)$ is strictly decreasing on $[\delta - \tau, \infty]$, and, therefore the maximum $L(z)$ is achieved at point $\hat{z} = \delta - \tau$. If we restrict the defender strategy space to the set of strategies against which Greedy does not stop playing (with size at least half of the original strategy space for the defender), then Greedy yields the strongly dominant attacker strategy specified in Sect. 5.2.

6.4. *Defender Playing by the Uniform Distribution Against Greedy*

We analyze now the Greedy strategy against a uniform defender distribution f_0 centered at δ and of support u, denoted $U(\delta - u/2, \delta + u/2)$. We assume that $3u/2 < \delta$. The density and cumulative probability functions of the uniform distribution are

$$f_0(x) = \begin{cases} \frac{1}{u}, & \delta - \frac{u}{2} \le x \le \delta + \frac{u}{2}, \\ 0, & \text{otherwise}, \end{cases} \qquad F_0(x) = \begin{cases} 0, & x < \delta - \frac{u}{2}, \\ \frac{x - (\delta - \frac{u}{2})}{u}, & \delta - \frac{u}{2} \le x \le \delta + \frac{u}{2}, \\ 1, & x < \delta + \frac{u}{2}. \end{cases}$$

When the attacker moves at time t, he learns the time τ since the defender's last move and the Greedy strategy is invoked to compute the time $\hat{z}$ until his next move. The following theorem gives a formula for $\hat{z}$ depending on the parameters of the uniform defender distribution, time since last defender's move, and the attacker's move cost:

Theorem 10. *Assume the defender plays with a renewal strategy given by probability density function $U(\delta - u/2, \delta + u/2)$, and the attacker finds out when moving at time t the interval τ since defender's last move. Assume that $3u/2 < \delta$ and that the attacker's move cost is $k_1 < \delta - u$. Then the time until the next attacker's move given by the Greedy strategy is $\hat{z} = \sqrt{(\delta - u/2 - \tau)^2 + 2k_1/c}$, where c is a constant defined as*

$$c = \begin{cases} \frac{1}{u}, & \tau \le u, \\ \frac{1}{\delta + u/2 - \tau}, & \tau \ge \delta - u/2. \end{cases}$$

In addition the time since the defender's last move is always either less than u or greater than $\delta - u/2$.

The proof of this theorem is given in Appendix C.

The uniform distribution is an example of a renewal defender strategy for which we cannot easily find the (strongly) dominant LM strategy. Theorem 10, however, specifies the LM strategy given by the Greedy algorithm.

6.5. *Greedy Does Not Always Result in Dominant Strategies!*

The analysis in Sect. 6.3 showed that for a periodic defender, there is a certain probability that the Greedy strategy will stop playing after the first move, but this probability can be made arbitrarily small. With high probability in that case, the Greedy strategy finds the strongly dominant LM strategy. We give now an example of a renewal strategy for which Greedy outputs a non-dominant strategy (with high probability).

Consider the defender renewal strategy given by density function f_0 defined as

$$f_0(x) = \begin{cases} 1, & \text{with probability } 0.01; \\ 1000, & \text{with probability } 0.99. \end{cases}$$

Assume that the move costs for both players are $k_0 = k_1 = 5$. We first compute the strategy found by the Greedy algorithm, and then define another strategy that achieves higher benefit.

Proposition 1. *If the attacker moves at time t, and receives feedback $0 < \tau < 1000 - k_1$, then the Greedy strategy outputs $\hat{z} = 1000 - \tau$.*

Proof. We consider several cases:

1. $0 < \tau < 1$: We can compute $L(z)$ as

$$L(z) = \begin{cases} \frac{z-k_1}{z}, & \tau + z \leq 1, \\ \frac{0.01(1-\tau)+0.99z-k_1}{z}, & 1 < \tau + z < 1000, \\ \frac{0.01(1-\tau)+0.99(1000-\tau)-k_1}{z}, & 1000 \leq \tau + z. \end{cases}$$

   For $\tau + z < 1$, $L(z)$ is negative. For $1 < \tau + z < 1000$, $L(z)$ is strictly increasing, and for $1000 \leq \tau + z$, $L(z)$ is strictly decreasing, hence, the maximum is $\hat{z} = 1000 - \tau$.
2. $1 \leq \tau < 1000$: We can compute $L(z)$ as

$$L(z) = \begin{cases} \frac{z-k_1}{z} = 1 - \frac{k_1}{z}, & z \leq 1000 - \tau, \\ \frac{1000-\tau-k_1}{z}, & z \geq 1000 - \tau. \end{cases}$$

   Therefore $L(z)$ is strictly increasing when $z \leq 1000 - \tau$, strictly decreasing when $z \geq 1000 - \tau$ and $\tau \leq 1000 - k_1$, and strictly increasing when $z \geq 1000 - \tau$ and $\tau \geq 1000 - k_1$. The maximum is achieved at $\hat{z} = 1000 - \tau$ or $\hat{z} = \infty$. □

We define now a strategy that achieves higher benefit than the Greedy strategy. At time t, assume that the attacker moves and receives feedback τ. With this strategy, the attacker always moves after time 1, if $\tau < 1$. On the other hand, if $1 < \tau < 1000$, then the attacker moves after time $1000 - \tau$.

Let $\beta_1 = \gamma_1 - k_1\alpha_1$ be the benefit of the attacker in the above strategy. Also, denote by $\beta_1^g = \gamma_1^g - k_1\alpha_1^g$ the benefit achieved by playing with the Greedy strategy from Proposition 1.

With the newly defined strategy, the attacker moves right after all defender's moves, virtually controlling the resource at all times. Compared to the Greedy strategy, the attacker controls the resource at least an additional 0.001 fraction of time, resulting in $\gamma_1 - \gamma_1^g \geq 0.001$. The increase in move cost for this strategy is at most $0.01k_1\alpha_0$.

We can infer that

$$\beta_1 - \beta_1^g \geq 0.001 - 0.01k_1\alpha_0 = 0.001 - 0.01\frac{5}{990.01} > 0.$$

Therefore the benefit of the attacker is increased compared to the Greedy strategy, demonstrating that the Greedy strategy does not result in a dominant attacker strategy against this example renewal defender distribution.

6.6. *Other Greedy Variants*

One refinement to the Greedy strategy explored above is to optimize the attacker's local benefit across an interval that includes two or more attacker moves (rather than just looking at the interval until the next attacker's move). For instance, we can define a

Greedy strategy that looks ahead at two attacker moves and optimizes at time t the local benefit $L(z_0, z_1)$ in interval $[t, t + z_1]$, given exactly two moves of the attacker at times $z_0 < z_1$ following t. After maximizing function $L(z_0, z_1)$ and finding optimum values $\hat{z}_0$ and $\hat{z}_1$, the attacker moves next at time $t + \hat{z}_0$ (assuming $L(\hat{z}_0, \hat{z}_1) > 0$) and then evaluates the new local benefit in order to compute his next move.

We can prove that this Greedy strategy results in the strongly dominant strategy against a defender playing with the periodic strategy with a random phase (assuming $k_1 < \delta/2$). We believe it would be interesting to investigate under which conditions such strategies that simultaneously optimize across two or more attacker moves are strongly dominant against renewal defender strategies.

7. General Variations and Extensions

We have defined and analyzed a basic version of the FlipIt game. The motivating applications we discussed in Sect. 2 raise a number of extensions and variations of likely interest in practical settings.

Other Types of Feedback In Sect. 3 we defined three different types of feedback in FlipIt: NA, RP and LM. All results proven in this paper are related to one of these three versions of the game. We could define different FlipIt instances in which players receive other types of feedback during the game, for example:

- *Previous mover* [PM]. The player who moves at time $t_k > 0$ learns the identity of the most recent previous mover—that is, player p_k learns the value

$$\phi_i(t_k) = C(t_k) = p_{k-1}.$$

  For $k > 1$, if $p_k \neq p_{k-1}$, then move k at time t_k is a *takeover* or a *turnover*, control of the game changes hands from player p_{k-1} to player p_k.
- *Score* [SC]. The mover finds out the score of both players. Note that this information is sufficient to figure out when your opponent regained control after your last move, if he had done so.

$$\phi_i(t_k) = \big(B_0(t_k), B_1(t_k)\big)$$

  where $B_i(t_k)$ is the score of player i up to time t_k.

Cost Variations Our FlipIt models above assume that the move cost k_i at time t is constant for each player. In variants of this game, k_i may depend on other factors, such as time t, $C(t)$ (who has possession of the resource at time t) and the amount of time a player has controlled the resource. For instance, the cost of moving might go up with the amount of time that the opponent has had control.

We can also imagine other extensions with a "discount rate" for benefits and costs, e.g., a move cost k_i at time $t - d$ is only ρ^d of that at time t, for some $\rho < 1$.

Budgets A natural option to consider is one in which one or both players have a move budget they cannot exceed. For instance, an enterprise's security operations may have an operating budget capping the security investment over the course of a year. (Alternatively, upper bounds might be imposed on the players' move rate.)

Effective/Ineffective Moves In reality, moves do not always succeed. In our host takeover example, cleaning of user machines are effective only with some probability.

We may therefore consider a model in which only with some probability $q < 1$ is a move *effective*, in the sense that it gives control to the player that has made it (if she does not already have control). A move is *ineffective*, then, i.e., does not effect a change of control, with probability $1 - q$.

Paying for More Information The players in FlipIt might pay for more information. That is, a basic "move" would cost k_i, but to find out the previous mover would cost an additional amount. Password reset offers an example application of this variant. Some systems display to users the time of last login to their accounts, information that can reveal to an attentive user the compromise and use of her account by an attacker. Getting users to pay attention to this extra information, though, carries a cost. It might involve, for instance, a pop-up window displaying the account's login history and requiring verification by the user—a valuable but burdensome security feature.

Extra discovery options might have tiered pricing: To find out the current score would cost even more than the current state of control. To find out the complete history would cost more yet.

Other Variants Of course, countless other variants or extensions of FlipIt may have practical application. Natural ones include:

- *Finite-time games*: Our experiments, as explained above, are simulations (in a discrete-time model) with a bounded time horizon T. Such games are of interest given the finite lifetime of most system deployments, and may be combined with other variant games features (e.g., bounded player budgets). For small T, optimal strategies can be exactly computed.
- *Refractory periods*: In this variant, a player that has moved at time t is not allowed to move again until at least time $t + \rho$. Here, ρ is what we call "refractory period," a time delay that reflects the resource depletion caused by a move (in, e.g., financial budget, personnel time, etc.).
  Note that in the basic (continuous) form of the game with $k_0 = k_1 = 1$ and no feedback, a refractory period of $\rho = 1$ is implicit: Since the move cost equals the benefit of control for an interval of length 1, it is always suboptimal to move twice in one such interval.
- *Multiplayer games*: Straightforward extensions of FlipIt can model competition for resource control among three or more players.

8. Related Work

FlipIt was first presented at an invited talk by Rivest at CRYPTO 2011[22].

The prior work most relevant to FlipIt seems to be in the field of game theory. In particular, work on "repeated games" seems especially relevant. (See, for example, the excellent text by Mailath and Samuelson [15].)

However, FlipIt does not meet the usual criteria for a repeated game, since:

- In FlipIt, time is typically continuous, not discrete. A repeated game has a sequence of stages, and a "stage game" is played again each stage. Thus, time is normally not continuous for a repeated game.
- In FlipIt, players do not know when the other player moves. In a repeated game, each player moves within each stage.
- FlipIt contains a hidden "state variable" that specifies which player has control at any given time. In a repeated game, though, there are no such "state variables" propagating information from one stage game to the next; each stage game starts afresh.

Nonetheless, at a higher level, FlipIt does share some qualitative characteristics with repeated games. For example, one does need to exercise some care in defining appropriate measures of payoff for each player, since the game is infinite in duration. The discussion by Myerson [18] on the use of a discount rate in defining the "payoff" for a player in an infinite repeated game is quite relevant. Similarly, if both players of FlipIt play adaptively, then FlipIt acquires the rich complexity of repeated games such as repeated Prisoner's Dilemma, where players may choose to cooperate for their mutual benefit, defect from a cooperation strategy, retaliate against a defector, etc. For our purposes it is helpful that in FlipIt *non-adaptive* play is meaningful, interesting, and worth analyzing.

When one player (e.g. the attacker) is adaptive, and the other is non-adaptive, then the problem for the attacker decomposes naturally into two parts: (1) learning what non-adaptive (renewal) strategy is being employed by the defender, and (2) playing optimally against the defender's strategy. Viewed as a problem in machine learning, this might be addressable using reinforcement learning [27]. However, since FlipIt is an infinite game, the learning (or estimation) problem is trivial to do asymptotically, and one might as well (without loss of generality) assume that the attacker is initially given a complete specification of the defender's non-adaptive strategy, rather than forcing the attacker to learn this specification.

Conventional game theory has a long history of application to and enhancement by cryptography; see Katz [14] for a survey. More pertinent to FlipIt are games modeling system security. Roy et al. [24] offer a taxonomy and survey of game-theoretic models in network security in particular. They note a preponderance of games differing from FlipIt in two regards. The games assume perfect information, meaning that players know the full game state. They assume synchronous moves by players.

The asynchronicity of FlipIt addresses a pervasive limitation highlighted by Hamilton et al. [10] in a broad study of information-security games. They observe that in most games once a move occurs, it occurs instantaneously. In information warfare this is not true. In fact, it can take a variable amount of time to carry out a move. (They also enumerate other directions for model enrichment, such as variation across time in action spaces and in players' goals and resources.)

Some recent information-security modeling has made use of extensive forms, which permit complex modeling, strictly within a framework of synchronous moves. This approach gives rise to security games with imperfect information, as in a game devised by Moore et al. to model zero-day disclosure by competing entities [17]. Nguyen et al. [19] consider an abstract, repeated security game with imperfect information and

also incomplete information, in the sense that players do not know one another's payoffs. (As we have shown, FlipIt can similarly be studied as a game of incomplete information in which strategy spaces and/or payoffs are unknown.) Also of relevance to FlipIt is a synchronous territorial game of incomplete, perfect information proposed by Pavlovic [21]. It aims to model two-player cyber-security scenarios and, in particular, information gathering via deception.

Acknowledgements

We thank members of RSA Laboratories and RSA Office of the CTO for many helpful suggestions regarding the practical applications of the framework proposed here. We also thank Alex Rivest for suggesting the name "FLIPIT".

Appendix A. Analysis of Renewal Games

This section is dedicated to the analysis of renewal FlipIt games. We start by proving Theorem 1 that gives the Nash equilibria for the FlipIt game with periodic strategies with random phases. We then prove Theorem 2 that gives an explicit formula of the players' benefits in a game with non-arithmetic renewal strategies. We can generalize Theorem 2 to a class that includes both non-arithmetic renewal strategies, as well as periodic strategies with random phases, resulting in Theorem 3. As a consequence, we can characterize in Theorem 4 the residual renewal game and the surviving strategies remaining after elimination of strongly dominated strategies.

Proof of Theorem 1. Nash equilibria are points with the property that neither player benefits by deviating in isolation from the equilibrium. We can compute Nash equilibria for the periodic game as intersection points of curves opt_0 and opt_1, as defined in Sect. 4.1.

To determine $\mathsf{opt}_0(\alpha_1)$, we need to compute the derivative of $\beta_0(\alpha_0, \alpha_1)$ for fixed α_1. We consider two cases:

Case 1: $\alpha_0 \geq \alpha_1$
Since

$$\frac{\partial \beta_0(\alpha_0, \alpha_1)}{\partial \alpha_0} = \frac{\alpha_1}{2\alpha_0^2} - k_0,$$

it follows that $\beta_0(\cdot, \alpha_1)$ is increasing on $[0, \sqrt{\frac{\alpha_1}{2k_0}}]$, decreasing on $[\sqrt{\frac{\alpha_1}{2k_0}}, \infty]$ and thus has a maximum at $\alpha_0 = \max\{\alpha_1, \sqrt{\frac{\alpha_1}{2k_0}}\}$.

Case 2: $\alpha_0 \leq \alpha_1$
Since

$$\frac{\partial \beta_0(\alpha_0, \alpha_1)}{\partial \alpha_0} = \frac{1}{2\alpha_1} - k_0,$$

it follows that (a) $\beta_0(\cdot, \alpha_1)$ is increasing if $\alpha_1 < \frac{1}{2k_0}$, and (b) decreasing if $\alpha_1 > \frac{1}{2k_0}$.

Based on that, we distinguish three cases for different values of α_1:

- If $\alpha_1 < \frac{1}{2k_0}$, then $\sqrt{\frac{\alpha_1}{2k_0}} \geq \alpha_1$. From case 1 and case 2(a) above, it follows that the optimal benefit for the defender is achieved at rate $\alpha_0 = \sqrt{\frac{\alpha_1}{2k_0}}$.
- If $\alpha_1 = \frac{1}{2k_0}$, then $\beta_0(\alpha_0, \alpha_1) = 0$, for all $\alpha_0 \in [0, \frac{1}{2k_0}]$. For $\alpha_0 \geq \alpha_1$ case 1 applies, hence, the benefit $\beta_0(\alpha_0, \alpha_1)$ decreases in α_0. In this case, the defender's maximum benefit is achieved for any α_0 in $[0, \frac{1}{2k_0}]$ and has value 0.
- If $\alpha_1 > \frac{1}{2k_0}$, it follows from case 1 and 2(b) that the defender's benefit is always non-positive, and as such the defender's optimal strategy is not playing at all.

From this analysis we can compute $\mathsf{opt}_0(\alpha_1)$ as

$$\mathsf{opt}_0(\alpha_1) = \begin{cases} \sqrt{\frac{\alpha_1}{2k_0}}, & \alpha_1 < \frac{1}{2k_0}, \\ [0, \sqrt{\frac{\alpha_1}{2k_0}}], & \alpha_1 = \frac{1}{2k_0}, \\ 0, & \alpha_1 > \frac{1}{2k_0}. \end{cases}$$

Similarly, we can also compute $\mathsf{opt}_1(\alpha_0)$:

$$\mathsf{opt}_1(\alpha_0) = \begin{cases} \sqrt{\frac{\alpha_0}{2k_1}}, & \alpha_0 < \frac{1}{2k_1}, \\ [0, \sqrt{\frac{\alpha_0}{2k_1}}], & \alpha_0 = \frac{1}{2k_1}, \\ 0, & \alpha_0 > \frac{1}{2k_1}. \end{cases}$$

Intersection points of curves opt_0 and opt_1 result in Nash equilibria. We distinguish several cases based on the relationship between players' move costs:

Case 1: $k_0 < k_1$

The Nash equilibrium is obtained for rates:

$$\alpha_0^* = \frac{1}{2k_1}; \qquad \alpha_1^* = 2k_0\alpha_0^2 = \frac{k_0}{2k_1^2}$$

and the players' benefits achieved in the Nash equilibrium are

$$\beta_0 = 1 - \frac{k_0}{k_1}; \qquad \beta_1 = 0.$$

Case 2: $k_0 > k_1$

The Nash equilibrium is obtained for rates:

$$\alpha_0^* = \frac{k_1}{2k_0^2}; \qquad \alpha_1^* = \frac{1}{2k_0}$$

and the players' benefits achieved in the Nash equilibrium are

$$\beta_0 = 0; \qquad \beta_1 = 1 - \frac{k_1}{k_0}.$$

Case 3: $k_0 = k_1$

The Nash equilibrium is given by

$$\alpha_0^* = \alpha_1^* = \frac{1}{2k_0}$$

and both players' benefits for the Nash equilibrium are $\beta_0 = \beta_1 = 0$. □

Proof of Theorem 2. We consider the two renewal strategies R_{f_0} and R_{f_1}. At time interval t, we denote the age of the first process (defined as the time lapsed since player 0's last move) as $x = Z_0(t)$ and the age of the second process as $y = Z_1(t)$. Let $f_{Z_0(t),Z_1(t)}$ denote the joint probability density function of the ages $Z_0(t)$ and $Z_1(t)$. The first player (defender) is in control at time t if $x \leq y$. As such, the probability that the defender is in control at time interval t can be computed as

$$C_0(t) = \int_{y=0}^{\infty} \int_{x=0}^{y} f_{Z_0(t),Z_1(t)}(x, y)\, dx\, dy.$$

Since both players use non-adaptive strategies, ages $Z_0(t)$ and $Z_1(t)$ are uncorrelated and statistically independent. We may write $f_{Z_0(t),Z_1(t)}(x, y) = f_{Z_0(t)}(x) f_{Z_1(t)}(y)$ and derive

$$\begin{aligned} C_0(t) &= \int_{y=0}^{\infty} f_{Z_1(t)}(y) \int_{x=0}^{y} f_{Z_0(t)}(x)\, dx\, dy \\ &= \int_{y=0}^{\infty} f_{Z_1(t)}(y) F_{Z_0(t)}(y)\, dy. \end{aligned}$$

Our goal is to prove that $C_0(t)$ converges to

$$G_0 = \int_{y=0}^{\infty} f_1^*(y) F_0^*(y)\, dy$$

as $t \to \infty$. As an immediate consequence we will show that γ_0, the gain of player 0, is equal to G_0.

We start by writing:

$$\begin{aligned} C_0(t) &= \int_{y=0}^{N} f_{Z_1(t)}(y) F_{Z_0(t)}(y)\, dy + \int_{y=N}^{\infty} f_{Z_1(t)}(y) F_{Z_0(t)}(y)\, dy \\ &= \int_{y=0}^{N} g_t(y)\, dy + C_0(N, t) \end{aligned}$$

and

$$\begin{aligned} G_0 &= \int_{y=0}^{N} f_1^*(y) F_0^*(y)\, dy + \int_{y=N}^{\infty} f_1^*(y) F_0^*(y)\, dy \\ &= \int_{y=0}^{N} g^*(y)\, dy + G_0(N) \end{aligned}$$

where for all $N > 0$ and all $t > 0$, we have defined

$$g_t(y) = f_{Z_1(t)}(y) F_{Z_0(t)}(y),$$
$$C_0(N,t) = \int_{y=N}^{\infty} f_{Z_1(t)}(y) F_{Z_0(t)}(y)\, dy,$$
$$g^*(y) = f_1^*(y) F_0^*(y),$$
$$G_0(N) = \int_{y=N}^{\infty} f_1^*(y) F_0^*(y)\, dy.$$

Our proof proceeds in several steps. First, we fix an $\epsilon > 0$. Then, we show that there exist an N and a T', such that for all $t \geq T'$, the tail $C_0(N,t)$ is bounded by 2ϵ. We also prove that $G_0(N)$ is bounded by ϵ. Finally, we show that there exists a T such that $|\int_{y=0}^{N}(g_t(y) - g^*(y))\, dy| \leq 2\epsilon$, for all $t \geq T$.

Bounding $C_0(N,t)$: For all $N > 0$ and $t > 0$,

$$\begin{aligned} C_0(N,t) &= \int_{y=N}^{\infty} f_{Z_1(t)}(y) F_{Z_0(t)}(y)\, dy \\ &\leq \int_{y=N}^{\infty} f_{Z_1(t)}(y)\, dy \leq 1 - F_1^*(N) + \left|F_1^*(N) - F_{Z_1(t)}(N)\right|. \end{aligned} \tag{A.1}$$

Since $\lim_{N\to\infty} F_1^*(N) = 1$, we choose N such that $1 - F_1^*(N) \leq \epsilon$. By Lemma 1, $\lim_{t\to\infty} F_{Z_1(t)}(N) = F_1^*(N)$. This implies that there exists T' such that for all $t \geq T'$: $|F_1^*(N) - F_{Z_1(t)}(N)| \leq \epsilon$. From this and (A.1) it follows that

$$\exists N \text{ and } T' \text{ such that } F_1^*(N) \geq 1 - \epsilon \text{ and } C_0(N,t) \leq 2\epsilon,\ \forall t \geq T'. \tag{A.2}$$

Bounding $G_0(N)$: By using similar arguments we can bound the tail $G_0(N)$. For the value of N chosen above,

$$G_0(N) = \int_{y=N}^{\infty} f_1^*(y) F_0^*(y)\, dy \leq \int_{y=N}^{\infty} f_1^*(y)\, dy = 1 - F_1^*(N) \leq \epsilon. \tag{A.3}$$

Bounding $|\int_{y=0}^{N}(g_t(y) - g^*(y))\, dy|$: Denote $s_t = \int_{y=0}^{N}(g_t(y) - g^*(y))\, dy$. By the definition of Riemann integrability, it follows that there exists an integer $k \geq N/\sqrt{\epsilon}$ such that for $y_i = Ni/k$, $0 \leq i \leq k$:

$$\left| s_t - \sum_{i=0}^{k-1} \big(g_t(y_i) - g^*(y_i)\big)(y_{i+1} - y_i) \right| \leq \epsilon.$$

By Lemma 1, it follows that $\lim_{t\to\infty} g_t(y) = g^*(y)$, and as such there exists $\{T_i\}_{i=0}^{k-1}$ such that

$$\left|g_t(y_i) - g^*(y_i)\right| \leq \sqrt{\epsilon}/k \quad \text{for all } t \geq T_i.$$

For $t \geq \max_i\{T_i\}$, we have

$$\begin{aligned} |s_t| &\leq \epsilon + \left| \sum_{i=0}^{k-1} \big(g_t(y_i) - g^*(y_i)\big)(y_{i+1} - y_i) \right| \\ &\leq \epsilon + \sum_{i=0}^{k-1} \sqrt{\epsilon}/k \cdot N/k \\ &\leq \epsilon + \sum_{i=0}^{k-1} \sqrt{\epsilon}/k \cdot \sqrt{\epsilon} = 2\epsilon. \end{aligned} \tag{A.4}$$

Finally, we can combine (A.2), (A.3) and (A.4) and infer that

$$\begin{aligned} &\exists N, T = \max\big\{T', \{T_i\}_{i=0}^{k}\big\} \text{ such that} \\ &\big|C_0(t) - G_0\big| \leq C_0(t, N) + G_0(N) + |s_t| \leq 5\epsilon, \ \forall t \geq T, \end{aligned}$$

which proves that

$$\lim_{t\to\infty} C_0(t) = G_0.$$

Since the limit of $C_0(t)$ is finite, it follows immediately (using standard calculus techniques) that

$$\gamma_0 = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} C_0(t) = G_0.$$

In a similar way we can derive the expression for player 1's gain. The formula for both players' benefits follows immediately since the average move cost of player i is $k_i\alpha_i$. □

Proof of Theorem 3. From Theorem 2, we can immediately infer that the result holds if both players use strategies from class $\mathcal{R}$.

We prove that the theorem holds for the defender playing with a strategy in class $\mathcal{R}$ and the attacker playing with a strategy in class $\mathcal{P}$. Let f_0 and F_0 be the probability density and cumulative distribution function of the non-arithmetic renewal process of the defender. Let δ be the period of the attacker's distribution, and θ the phase randomly chosen in $[0, \delta]$.

Let t denote an arbitrary time interval, and we would like to compute the probability that the defender is in control at time t. Denote by x the time lapsed since the last move of the defender and by y the time lapsed since the last move of the attacker (both computed at time t). Then we can express:

$$C_0(t) = \int_{x=0}^{y} f_{Z_0(t)}(x)\, dx = F_{Z_0(t)}(y) \quad \text{with } y = t - (\theta + k\delta)$$

where $k \geq 0$ is an integer such that $t - (\theta + k\delta) \geq 0 > t - (\theta + (k+1)\theta)$.

We separate time into intervals between attacker's moves. These have fixed length (except the first one): $[0,\theta]$, $[\theta,\theta+\delta]$, $[\theta,\theta+2\delta]$, $\ldots$, $[\theta,\theta+k\delta]$, $\ldots$. We compute the defender's gain γ_0^k in an interval $[\theta+k\delta,\theta+(k+1)\delta]$:

$$\begin{aligned}\gamma_0^k &= \frac{1}{\delta}\int_{\theta+k\delta}^{\theta+(k+1)\delta} C_0(t)\,dt = \frac{1}{\delta}\int_{\theta+k\delta}^{\theta+(k+1)\delta} F_{Z_0(t)}\big(t-(\theta+k\delta)\big)\,dt \\ &= \frac{1}{\delta}\int_0^{\delta} F_{Z_0(x+\theta+k\delta)}(x)\,dx. \end{aligned} \tag{A.5}$$

From Lemma 1 we know that $F_{Z_0(t)}(x)$ converges as $t\to\infty$ and

$$F_0^*(x) = \lim_{t\to\infty} F_{Z_0(t)}(x).$$

We can now use this limit in (A.5). Since the integral in (A.5) is finite, the limit and the integral commute and we can infer

$$\lim_{k\to\infty}\gamma_0^k = \frac{1}{\delta}\int_0^{\delta} F_0^*(x)\,dx.$$

If we denote the latter limit by L, we can use the definition of limits and derive that

$$\forall\epsilon,\ \exists k_\epsilon \text{ such that } \left|\gamma_0^k - L\right| < \epsilon,\ \forall k \geq k_\epsilon. \tag{A.6}$$

Our final goal is to compute the defender's gain, which by definition is

$$\gamma_0 = \lim_{t\to\infty}\frac{1}{T}\int_{t=0}^{T} C_0(t)\,dt.$$

Let $\epsilon > 0$ be fixed, and k_ϵ be as in (A.6). Then

$$\begin{aligned} A_T &= \frac{1}{T}\int_0^T C_0(t)\,dt = \frac{1}{T}\int_0^{\theta+k_\epsilon\delta} C_0(t)\,dt + \frac{1}{T}\int_{\theta+k_\epsilon\delta}^{T} C_0(t)\,dt \\ &= \frac{1}{T}\int_0^{\theta+k_\epsilon\delta} C_0(t)\,dt \\ &\quad + \frac{1}{T}\left[\int_{\theta+k_\epsilon\delta}^{\theta+(k_\epsilon+1)\delta} C_0(t)\,dt + \int_{\theta+(k_\epsilon+1)\delta}^{\theta+(k_\epsilon+2)\delta} C_0(t)\,dt + \cdots + \int_{\theta+n\delta}^{T} C_0(t)\,dt\right] \\ &= \frac{1}{T}\int_0^{\theta+k_\epsilon\delta} C_0(t)\,dt + \frac{1}{T}\int_{\theta+n\delta}^{T} C_0(t)\,dt + \frac{\delta}{T}\left[\gamma_0^{k_\epsilon} + \cdots + \gamma_0^{n-1}\right], \end{aligned} \tag{A.7}$$

where n is such that $\theta+n\delta \leq T < \theta+(n+1)\delta$.

Both integrals $\int_0^{\theta+k_\epsilon\delta} C_0(t)\,dt$ and $\int_{\theta+n\delta}^{T} C_0(t)\,dt \leq \delta\gamma_0^n$ are finite, and therefore

$$\lim_{T\to\infty}\frac{1}{T}\int_0^{\theta+k_\epsilon\delta} C_0(t)\,dt + \frac{1}{T}\int_{\theta+n\delta}^{T} C_0(t)\,dt = 0. \tag{A.8}$$

From (A.6) we can bound the latter sum in (A.7):

$$\delta(L-\epsilon)\frac{n-k_\epsilon}{T} \le \frac{\delta}{T}\left[\gamma_0^{k_\epsilon}+\cdots+\gamma_0^{n-1}\right] \le \delta(L+\epsilon)\frac{n-k_\epsilon}{T}.$$

Since $\lim_{T\to\infty}\frac{n-k_\epsilon}{T}=\frac{1}{\delta}$, it follows immediately that

$$\lim_{T\to\infty}\frac{\delta}{T}\left[\gamma_0^{k_\epsilon}+\cdots+\gamma_0^{n-1}\right]=L. \tag{A.9}$$

Combining (A.7), (A.9), and (A.8), we prove that $\lim_{T\to\infty}A_T=L$, implying that the gain of player 0 is L:

$$\gamma_0=\lim_{T\to\infty}A_T=\frac{1}{\delta}\int_0^{\delta}F_0^*(x)\,dx.$$

The last step is to prove that this last integral is equal to the expression given in the theorem statement.

The size-bias density function for the periodic strategy of the attacker is uniform and can be expressed as

$$f_1^*(x)=\begin{cases}\frac{1}{\delta}, & x\in[0,\delta),\\ 0, & x\ge\delta.\end{cases}$$

Then

$$\int_0^{\infty}f_1^*(x)F_0^*(x)\,dx=\frac{1}{\delta}\int_0^{\infty}F_0^*(x)\,dx=\gamma_0.$$

To complete the proof, we have to show that the statement of the theorem also holds for both strategies in class $\mathcal{P}$. Denote by δ_0 and δ_1 the periods of the defender's and attacker's strategies. Then

$$f_1^*(x)=\begin{cases}\frac{1}{\delta_1}, & x<\delta_1,\\ 0, & x\ge\delta_1,\end{cases}\qquad F_0^*(x)=\begin{cases}\frac{x}{\delta_0}, & x<\delta_0,\\ 1, & x\ge\delta_0.\end{cases}$$

We can now compute $\int_{x=0}^{\infty}f_1^*(x)F_0^*(x)\,dx$ by considering two cases:

1. $\delta_0\le\delta_1$:

$$\begin{aligned}\int_{x=0}^{\infty}f_1^*(x)F_0^*(x)\,dx&=\int_{x=0}^{\delta_0}\frac{x}{\delta_0\delta_1}\,dx+\int_{\delta_0}^{\delta_1}\frac{1}{\delta_1}\,dx\\ &=\frac{\delta_0}{2\delta_1}+\frac{1}{\delta_1}(\delta_1-\delta_0)=1-\frac{\delta_0}{2\delta_1}.\end{aligned}$$

2. $\delta_0>\delta_1$:

$$\int_{x=0}^{\infty}f_1^*(x)F_0^*(x)\,dx=\int_{x=0}^{\delta_1}\frac{x}{\delta_0\delta_1}\,dx=\frac{\delta_1}{2\delta_0}.$$

From the analysis in Sect. 4.1, it follows that this last expression indeed equals the gain γ_0 of player 0. This completes our proof. □

Proof of Theorem 4. We prove the first part of the theorem. Consider two strategies from class $\mathcal{R} \cup \mathcal{P}$ and denote by f_0 and f_1 the probability densities of the two processes (if the strategy of player i is in class $\mathcal{R}$, then f_i is the non-arithmetic probability density of the renewal process; otherwise if player i employs a strategy in class $\mathcal{P}$, then f_i is the probability density of the periodic distribution).

We also denote by F_0 and F_1 the corresponding cumulative distribution functions, and by $f_0^*, f_1^*, F_0^*, F_1^*$ the size-bias density and cumulative distributions for the two strategies. From the definition of the size-bias density and cumulative distribution function, it follows immediately that f_0^* and f_1^* are non-increasing.

By applying Theorem 3 the average benefit rate of player 0 is

$$\begin{aligned}\beta_0 &= \int_0^\infty f_1^*(x) F_0^*(x)\,dx - k_0\alpha_0 = \int_0^\infty f_1^*(x)\left(\int_0^x f_0^*(y)\,dy\right)dx - k_0\alpha_0 \\ &= \int_{y=0}^\infty \left(\int_{x=y}^\infty f_1^*(x)\,dx\right) f_0^*(y)\,dy - k_0\alpha_0 \\ &= \int_{y=0}^\infty f_0^*(y)\big(1 - F_1^*(y)\big)\,dy - k_0\alpha_0. \end{aligned} \tag{A.10}$$

Assume that the rate of play α_0 of player 0 is fixed. Then $f_0^*(0) = \alpha_0$. In (A.10) functions $f_0^*(y)$ and $1 - F_1^*(y)$ are non-increasing. Accordingly, the function $f_0^*(y)$ that maximizes β_0 (for any strategy of player 1) is one that takes the largest possible values for small values for y:

$$f_0^*(y) = \begin{cases} \alpha_0, & 0 \le y \le 1/\alpha_0, \\ 0, & y > 1/\alpha_0. \end{cases} \tag{A.11}$$

But this size-bias probability density corresponds to the periodic distribution. This shows that in the class of strategies $\mathcal{R}_{\alpha_0} \cup \mathcal{P}_{\alpha_0}$ (with fixed rate of play α_0) for player 0, the strategy P_{α_0} is the strongly dominant strategy (assuming player 1 plays with any strategy in $\mathcal{R} \cup \mathcal{P}$). A similar result can be proven for player 1.

We make the observation that the same argument holds if a player knows the rate of play of his opponent. For instance, if player 0 knows the rate of play α_1 of player 1, function $1 - F_1^*(y) = 1 - F_1^*(\alpha_1, y)$ (in variable y) is still decreasing, and using (A.10) player 0's benefit is maximized for f_0^* as in (A.11).

In addition, the argument can be extended to the case in which one player (player 0) knows the distribution F_1 of his opponent and the other player (player 1) is renewal and knows the rate of play of his opponent α_0 or has no information about his opponent before the game starts. Function $1 - F_1^*(y) = 1 - F_1^*(\alpha_0; y)$ (in variable y) is still decreasing (for fixed α_0), and using (A.10) player 0's benefit is maximized for f_0^* as in (A.11) regardless of additional knowledge about the exact distribution $F_1(\alpha_0; y)$.

The second part of the theorem is an immediate consequence: Since all non-arithmetic renewal strategies are strongly dominated by periodic strategies, the residual FlipIt game consists of only periodic strategies. □

Proof of Theorem 5. From the proof of Theorem 1, we know that if α_0 is fixed by the defender, the attacker's rate of play α_1^* that maximizes his benefit can be chosen according to $\mathsf{opt}_1(\alpha_0)$:

$$\mathsf{opt}_1(\alpha_0) = \begin{cases} \sqrt{\frac{\alpha_0}{2k_1}}, & \alpha_0 < \frac{1}{2k_1}, \\ [0, \sqrt{\frac{\alpha_0}{2k_1}}], & \alpha_0 = \frac{1}{2k_1}, \\ 0, & \alpha_0 > \frac{1}{2k_1}. \end{cases}$$

We consider several cases:

Case 1: $\alpha_0 < \frac{1}{2k_1}$

Then, $\alpha_1^* = \sqrt{\frac{\alpha_0}{2k_1}} \geq \alpha_0$. We can write the defender's benefit as

$$\beta_0(\alpha_0, \alpha_1^*) = \frac{\alpha_0}{2\alpha_1^*} - k_0\alpha_0 = \sqrt{\frac{\alpha_0 k_1}{2}} - k_0\alpha_0.$$

The maximum for $\beta_0(\cdot, \alpha_1^*)$ is achieved at $\alpha_0 = \frac{k_1}{8k_0^2}$. As such, function β_0 is increasing on $[0, \frac{k_1}{8k_0^2}]$ and decreasing on $[\frac{k_1}{8k_0^2}, \infty]$. We further distinguish two cases:

- $k_1 < 2k_0$: Then $\beta_0(\cdot, \alpha_1^*)$ is maximized at point $\frac{k_1}{8k_0^2} \in [0, \frac{1}{2k_1}]$. It follows immediately that the defender's rate of play that maximizes his benefit is $\alpha_0^* = \frac{k_1}{8k_0^2}$. The attacker's rate of play that maximizes his benefit is then $\alpha_1^* = \sqrt{\frac{\alpha_0^*}{2k_1}} = \frac{1}{4k_0}$. The players' benefits are

$$\beta_0(\alpha_0^*, \alpha_1^*) = \frac{k_1}{8k_0}; \qquad \beta_1(\alpha_0^*, \alpha_1^*) = 1 - \frac{k_1}{2k_0}.$$

- $k_1 > 2k_0$: Then $\frac{1}{2k_1} \in [0, \frac{k_1}{8k_0^2}]$. Since function $\beta_0(\cdot, \alpha_1^*)$ is increasing on $[0, \frac{k_1}{8k_0^2}]$, the optimal defender strategy is to play with maximum rate $\alpha_0^* \uparrow \frac{1}{2k_1}$. The attacker's optimal rate of play is $\alpha_1^* \uparrow \sqrt{\frac{\alpha_0^*}{2k_1}} = \frac{1}{2k_1}$. This results in benefits

$$\beta_0(\alpha_0^*, \alpha_1^*) = \frac{1}{2} - \frac{k_0}{2k_1}; \qquad \beta_1(\alpha_0^*, \alpha_1^*) = 0.$$

Case 2: $\alpha_0 = \frac{1}{2k_1}$

The attacker's rate of play that optimizes his benefit is any $\alpha_1^* \in [0, \frac{1}{2k_1}]$. The defender's benefit $\beta_0(\alpha_0, \alpha_1^*)$ ranges in interval $[\frac{1}{2} - \frac{k_0}{2k_1}, 1 - \frac{k_0}{2k_1}]$.

Case 3: $\alpha_0 > \frac{1}{2k_1}$

The attacker optimizes his benefit by not playing at all, resulting in $\alpha_1^* = 0$. The defender's benefit is $\beta_0(\alpha_0, \alpha_1^*) = 1 - k_0\alpha_0$, which is maximized for $\alpha_0^* \downarrow \frac{1}{2k_1}$. With

these rates of play, the players' benefits are

$$\beta_0 = 1 - \frac{k_0}{2k_1}; \qquad \beta_1 = 0.$$

Based on this analysis, we can distinguish several cases of interest:

Case 1: $k_1 < 2k_0$

We proved previously that for $\alpha_0 \in (0, \frac{1}{2k_1})$, the maximum benefit is $\frac{k_1}{8k_0}$ (achieved at rate $\alpha_0 = \frac{k_1}{8k_0^2}$), and for $\alpha_0 \in [\frac{1}{2k_1}, \infty)$, the maximum benefit is $1 - \frac{k_0}{2k_1}$ (achieved at rate $\alpha_0 \downarrow \frac{1}{2k_1}$).

The optimal defender's benefit is then $\beta_0 = \max(\frac{k_1}{8k_0}, 1 - \frac{k_0}{2k_1})$. By analyzing the quadratic function $k_1^2 - 8k_0k_1 + 4k_0^2$ in variable k_1 (for fixed k_0) it turns out that $\frac{k_1}{8k_0} > 1 - \frac{k_0}{2k_1}$ when $k_1 < k_0(4 - \sqrt{12})$. It follows that

$$\alpha_0^* = \begin{cases} \frac{k_1}{8k_0^2}, & k_1 \le k_0(4 - \sqrt{12}), \\ \downarrow \frac{1}{2k_1}, & k_0(4 - \sqrt{12}) < k_1 < 2k_0 \end{cases}$$

and the maximum defender's benefit is

$$\beta_0(\alpha_0^*, \alpha_1^*) = \begin{cases} \frac{k_1}{8k_0}, & k_1 \le k_0(4 - \sqrt{12}), \\ 1 - \frac{k_0}{2k_1}, & k_0(4 - \sqrt{12}) < k_1 < 2k_0. \end{cases}$$

Case 2: $k_1 \ge 2k_0$

The optimal defender's benefit is $1 - \frac{k_0}{2k_1}$ achieved at rate $\alpha_0^* \downarrow \frac{1}{2k_1}$.

The conclusion of the theorem follows. □

Appendix B. Analysis of Renewal Defender Against LM Attacker

Proof of Theorem 6. Let $f_0(x)$ be the probability density function for the defender's exponential distribution:

$$f_0(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

Suppose that the attacker moves at some time interval t, and finds out that the interval since the defender's last move is y. Then, the probability that the interval till the defender's next move is time x is

$$\frac{f_0(x+y)}{\int_{z=y}^{\infty} f_0(z)\,dz} = \frac{\lambda e^{-\lambda(x+y)}}{\int_{z=y}^{\infty} \lambda e^{-\lambda z}\,dz} = \frac{\lambda e^{-\lambda(x+y)}}{\lambda e^{-\lambda y}} = \lambda e^{-\lambda x} = f_0(x).$$

Thus, an attacker moving at time t only knows that the interval till the next defender's move is distributed with probability density function $f_0(x)$. Based on this knowledge, and move costs k_0 and k_1, the attacker can determine an interval of time till his next

move: he decides to move at time $t + \delta$. But at each move, the knowledge of the attacker is exactly the same (it does not depend in any way on the time interval of his move, or the time interval since the defender's last move). As such, we can make the argument that every time he moves, the attacker chooses the same interval δ until his next move. The exception is the first move (when the attacker has no knowledge about the defender), which we may assume is uniformly distributed. This proves that the dominant LM strategy against an exponential defender belongs to class $\mathcal{P}$ of periodic strategies with random phases.

Assume that the attacker's strongly dominant strategy is $P_\alpha \in \mathcal{P}$ and its period is $\delta = 1/\alpha$. We compute players' benefits in game $\mathsf{FlipIt}(E_\lambda, P_\alpha)$.

In each interval of length δ between attacker's two moves (assume for simplicity the interval is $[0, \delta]$), the defender moves at some time u with probability $f_0(u)$ and controls the resource for $\delta - u$ time. As such, the defender's gain is

$$\gamma_0 = \int_{u=0}^{\delta} \lambda e^{-\lambda u} \frac{\delta - u}{\delta}\, du = 1 - \frac{1 - e^{-\lambda\delta}}{\lambda\delta}.$$

The gain of the attacker is equal to $\gamma_1 = 1 - \gamma_0$. The cost of players 0 and 1 are λk_0 and k_1/δ. As such, we can derive the formulas for the benefits:

$$\beta_0(E_\lambda, P_\alpha) = \gamma_0 - \lambda k_0 = 1 - \frac{1 - e^{-\lambda\delta}}{\lambda\delta} - \lambda k_0,$$

$$\beta_1(E_\lambda, P_\alpha) = 1 - \gamma_0 - \frac{k_1}{\delta} = \frac{1 - e^{-\lambda\delta}}{\lambda\delta} - \frac{k_1}{\delta}.$$

□

Proof of Theorem 7. For game $\mathsf{FlipIt}(E_\lambda, P_\alpha)$, the attacker's benefit $\beta_1(E_\lambda, P_\alpha)$ is a function of rates of play λ and α. Let $\delta = 1/\alpha$ be the period of the defender's strategy.

We fix λ and analyze the optimum for β_1 as a function of period δ. For that, we compute the derivative of the benefit β_1 with respect to δ:

$$\frac{d\beta_1}{d\delta} = \frac{e^{-\lambda\delta}(1 + \lambda\delta) - (1 - \lambda k_1)}{\lambda\delta^2}.$$

We define function $g(z) = e^{-z}(1 + z)$, for $z > 0$. Since $g'(z) = -ze^{-z} < 0$, function g is decreasing. We consider two cases:

1. If $\lambda < 1/k_1$, then the equation $g(\lambda\delta) = 1 - \lambda k_1$ has a unique solution δ, which satisfies the equation:

$$e^{-\lambda\delta}(1 + \lambda\delta) = 1 - \lambda k_1.$$

   The solution of the above equation is the period δ for which β_1 is maximized.
2. If $\lambda \geq 1/k_1$, then $\frac{d\beta_1}{d\delta} > 0$, and as such β_1 is strictly increasing. In this case, the benefit β_1 is increased by increasing δ arbitrarily, which corresponds to a strategy of not playing at all for the attacker. □

Proof of Theorem 8. From Theorem 7, it follows that if $\lambda < 1/k_1$, then the attacker's strongly dominant strategy is P_α with period $\delta = 1/\alpha$ satisfying:

$$e^{-\lambda\delta}(1 + \lambda\delta) = 1 - \lambda k_1. \tag{B.1}$$

If $\lambda \geq 1/k_1$, then the attacker's strongly dominant strategy is not playing at all.

We would like to find the optimal rate of play λ of the defender that maximizes his benefit $\beta_0(E_\lambda, P_{1/\delta})$. If the defender plays with rate $\lambda = 1/k_1$, then his benefit is constant at $\beta_0(E_\lambda, P_{1/\delta}) = 1 - k_0/k_1$ (by using Theorem 6 and the fact that the attacker's strongly dominant strategy is not playing at all). The question of interest is if there are rates $\lambda < 1/k_1$ for which the defender's benefit is higher than $1 - k_0/k_1$. The answer depends on the defender's and attacker's move costs, k_0 and k_1.

Since we focus on the case $\lambda \leq 1/k_1$, let $\lambda = \frac{1-\epsilon}{k_1}$, for $0 \leq \epsilon \leq 1$, and define $z = \lambda\delta$, as well as function $g(x) = (1 + x)e^{-x}$, for $x > 0$. Function g is strictly decreasing for $x > 0$.

From (B.1), it follows that $g(z) = 1 - \lambda k_1 = \epsilon$, and as such z is the unique solution of this equation. We can now use Theorem 6 and express the defender's benefit as a function of z, k_0 and k_1:

$$\beta_0(z) = 1 - \frac{k_0}{k_1}\big(1 - g(z)\big) - \frac{1 - e^{-z}}{z}. \tag{B.2}$$

We would like to find the point z for which the benefit β_0 is maximized. Once we determine such z, we can find $\epsilon = g(z)$ and then find the optimal rate of play λ. For this, we compute the derivative of β_0 with respect to z:

$$\frac{\mathrm{d}\beta_0}{\mathrm{d}z} = -ze^{-z}\left\{\frac{k_0}{k_1} - \frac{e^z - 1 - z}{z^3}\right\}. \tag{B.3}$$

The number of zeroes of (B.3) for $z \in (0, \infty)$ is equal to the number of solutions of

$$\frac{k_0}{k_1}z = \frac{e^z - 1 - z}{z^2}. \tag{B.4}$$

The Taylor series of $(e^z - 1 - z)/z^2$ around 0 has non-negative coefficients, therefore the right side of (B.4) is convex. The left side is a line $k_0 z/k_1$ through the origin. It intersects the convex function $(e^z - 1 - z)/z^2$ in at most two points.

Let us compute the point at which the tangent to $(e^z - 1 - z)/z^2$ passing through the origin intersects the function. This is determined by equation:

$$\frac{(e^z - 1 - z)/z^2}{z} = \frac{\mathrm{d}}{\mathrm{d}z}\big[\big(e^z - 1 - z\big)/z^2\big],$$

equivalent to

$$e^z z - 3e^z + 2z + 3 = 0,$$

which has a solution for $z = 2.1491$. Function $(e^z - z - 1)/z^3$ evaluated in $z = 2.1491$ has value 0.5468. As such, we distinguish several cases for analyzing the optimum of β_0:

Case 1: $k_0/k_1 > 0.5468$

In this case (B.3) has two zeros, $z_1 < 2.1491$ and $z_2 > 2.1491$. From these, z_1 is a local maximum point. We need to compare the benefit $\beta_0(z_1)$ achieved in the local

maximum z_1 with $\beta_0(0) = 1 - k_0/k_1$ achieved for $\lambda = 1/k_1$. The optimum β_0 is the maximum of the two.

By using (B.2), inequality $\beta_0(z_1) \geq \beta_0(0)$ is equivalent to:

$$\frac{k_0}{k_1} g(z_1) \geq \frac{1 - e^{-z_1}}{z_1}.$$

Since z_1 is a solution of (B.4), this is equivalent to

$$e^{z_1}\left(z_1 + 1 - z_1^2\right) \geq 1 + 2z_1,$$

which is only satisfied if $z_1 \leq 0.77$. From (B.4), this is achieved if and only if $k_0/k_1 \geq 0.854$. We now distinguish two subcases:

(a) $k_0/k_1 \geq 0.854$: the maximum β_0 is achieved for $z_1 \leq 0.77$ solution of (B.4).
(b) $k_0/k_1 < 0.854$: the maximum β_0 is achieved for $\lambda = 1/k_1$.

Case 2: $k_0/k_1 \leq 0.5468$

In this case $(e^z - z - 1)/z^2 \geq (k_0/k_1)z$, and from (B.3) it follows that $d\beta_0/dz \geq 0$ and as such β_0 is increasing. The maximum β_0 is achieved for $z \to \infty$, which implies $\epsilon \to 0$ and $\lambda \to 1/k_1$. The maximum benefit in this case according to Theorem 7 is $\beta_0 = 1 - k_0/k_1$. □

Proof of Theorem 9. Let $p'(x) = \lambda' e^{-\lambda' x}, x > 0$ be the probability density function for the exponential distribution with rate λ'. After each move, the defender waits for fixed interval Δ, and then chooses a random variable with pdf $p'(x)$ until his next move. Assume that at his nth move, the attacker finds out that the time since defender's last move is τ_n. There are two cases we consider:

1. $\tau_n < \Delta$: Then the defender does not move in the next interval of length $[\Delta - \tau_n]$, and thus the attacker controls the resource in this interval. After time $\Delta - \tau_n$, the defender plays with exponential pdf $p'(x)$. With a similar argument as in the proof of Theorem 6, the attacker only knows that at each moment in time after $\Delta - \tau_n$, the probability until defender's next move is given by $p'(x)$. Based on this knowledge and move costs k_0 and k_1, the attacker determines an interval δ' until his next move. The important point here is that δ' is the same in all intervals when the defender move times are determined by exponential play due to the memoryless property of the exponential distribution.
2. $\tau_n \geq \Delta$: In this case, the defender moves with probability density function $p'(x)$. As in the first part, the view of the attacker is the same at each time due to the memoryless property of the exponential distribution. As such, the attacker chooses the same δ' as in the previous case.

It follows that the strongly dominant attacker strategy is given by the strategy defined in the theorem's statement. □

Appendix C. Analysis of the LM Greedy Algorithm Against an Uniform Defender

Proof of Theorem 10. Assume that the attacker moves first at time $\delta + u/2$ and finds out the time τ since the defender's last move. We shall show that if the attacker selects moves according to the Greedy algorithm, then after each move the following invariants hold:

$$\tau \le u \quad \text{or} \quad \tau \ge \delta - u/2,$$
$$\hat{z} \in [\delta - u/2 - \tau, \delta + u/2 - \tau], \quad \text{with } \hat{z} > 0.$$

We first compute density function $\hat{f}_0$ defined as

$$\hat{f}_0(x) = \frac{f_0(\tau + x)}{1 - F_0(\tau)}.$$

Assuming that the invariants above hold, there are two cases to consider:

(1) $0 < \tau + x < \delta - \frac{u}{2}$: $\hat{f}_0(x) = 0$,
(2) $\delta - \frac{u}{2} \le \tau + x \le \delta + \frac{u}{2}$:

- If $\tau \le u$, then $\tau \le \delta - u/2$, and $F_0(\tau) = 0$, $\hat{f}_0(x) = 1/u$.
- If $\tau \ge \delta - u/2$, then $F_0(\tau) = \frac{\tau - (\delta - u/2)}{u}$ and $\hat{f}_0(x) = 1/(\delta + u/2 - \tau)$.

Denote by c the constant value of $\hat{f}_0(x)$:

$$c = \begin{cases} \frac{1}{u}, & \tau \le u, \\ \frac{1}{\delta + u/2 - \tau}, & \tau \ge \delta - u/2. \end{cases}$$

For computing the local benefit $L(z)$, we make some notations:

$$I_1 = \int_0^z x \hat{f}_0(x)\, dx = c \int_{\max(0, \delta - u/2 - \tau)}^{\min(z, \delta + u/2 - \tau)} x\, dx,$$
$$I_2 = \int_z^\infty \hat{f}_0(x)\, dx = c \int_{\max(z, \delta - u/2 - \tau)}^{\delta + u/2 - \tau} dx.$$

We can write the local benefit $L(z)$ as

$$L(z) = \frac{1}{z}(I_1 + zI_2 - k_1).$$

We will compute both integrals I_1 and I_2, as well as the local benefit $L(z)$ for z in different intervals:

(1) $0 < z < \delta - u/2 - \tau$: This case only occurs if $\tau \le u$.
Then $I_1 = 0$ and

$$I_2 = \frac{1}{u} \int_{\delta - u/2 - \tau}^{\delta + u/2 - \tau} dx = 1.$$

We can compute the local benefit as $L_1 = 1 - k_1/z$, which clearly shows that $L(z)$ is strictly increasing on this interval.

(2) $\delta - u/2 - \tau \le z \le \delta + u/2 - \tau$: We compute

$$I_1 = c\int_{\delta-u/2-\tau}^{z} x\,dx = \frac{c}{2}\big[z^2 - (\delta - u/2 - \tau)^2\big],$$

$$I_2 = c\int_{z}^{\delta+u/2-\tau} dx = c[\delta + u/2 - \tau - z].$$

The local benefit $L(z)$ can be expressed as

$$L(z) = \frac{c}{2z}\big[z^2 - (\delta - u/2 - \tau)^2\big] + c[\delta + u/2 - \tau - z] - \frac{k_1}{z}$$

$$= c[\delta + u/2 - \tau] - \frac{cz}{2} - c\frac{(\delta - u/2 - \tau)^2}{2z} - \frac{k_1}{z}.$$

To analyze the function $L(z)$ in interval $[\delta - u/2 - \tau, \delta + u/2 - \tau]$, we compute its derivative:

$$L'(z) = c\frac{(\delta - u/2 - \tau)^2}{2z^2} - \frac{c}{2} + \frac{k_1}{z^2} = \frac{(\delta - u/2 - \tau)^2 + 2k_1/c - z^2}{2z^2/c}.$$

The positive solution of the derivative is $z_0 = \sqrt{(\delta - u/2 - \tau)^2 + 2k_1/c}$. Obviously, $z_0 > \delta - u/2 - \tau$. We shall show that $z_0 \le \delta + u/2 - \tau$ by considering two cases:

- $\tau \le u$: Then $c = 1/u$, $k_1 \le \delta - u \le \delta - \tau$ and

$$z_0 = \sqrt{(\delta - u/2 - \tau)^2 + 2k_1 u}$$

$$\le \sqrt{(\delta - u/2 - \tau)^2 + 2(\delta - \tau)u} = \delta + u/2 - \tau.$$

- $\tau \ge \delta - u/2$: Then $c = 1/(\delta + u/2 - \tau)$, $k_1 \le \delta - u$, $\delta + u/2 - \tau \le u$ and

$$z_0 = \sqrt{(\delta - u/2 - \tau)^2 + 2k_1(\delta + u/2 - \tau)}$$

$$\le \sqrt{(\delta - u/2 - \tau)^2 + 2(\delta - \tau)u} = \delta + u/2 - \tau.$$

In both cases, function $L'(z)$ is positive on $[\delta - u/2 - \tau, z_0]$, and is negative on $[z_0, \delta + u/2 - \tau]$. Therefore, function $L(z)$ is strictly increasing on $[\delta - u/2 - \tau, z_0]$ and strictly decreasing on $[z_0, \delta + u/2 - \tau]$.

Since $z_0 \le \delta + u/2 - \tau$, $L(z_0) = c(\delta + u/2 - \tau - z_0) \ge 0$.

(3) $z > \delta + u/2 - \tau$: Then $I_2 = 0$ and there are two cases for computing I_1:

- $\tau \le u \le \delta - u/2$:

$$I_1 = \frac{1}{u}\int_{\delta-u/2-\tau}^{\delta+u/2-\tau} x\,dx = \frac{1}{2u}\big[(\delta + u/2 - \tau)^2 - (\delta - u/2 - \tau)^2\big] = (\delta - \tau).$$

The local benefit $L(z)$ can be computed as

$$L(z) = \frac{1}{z}[\delta - \tau - k_1].$$

– $\tau \geq \delta - u/2$:

$$I_1 = \frac{1}{\delta + u/2 - \tau} \int_0^{\delta+u/2-\tau} x\,dx = \frac{\delta + u/2 - \tau}{2}.$$

The local benefit $L(z)$ can be computed as

$$L(z) = \frac{1}{2z}[\delta + u/2 - \tau - 2k_1].$$

Therefore, in both cases $L(z) \to 0$ as $z \to \infty$ and $L(z)$ is either strictly decreasing or strictly increasing on this interval.

Based on the three cases, the local benefit $L(z)$ is strictly increasing on $[0, z_0]$ and strictly decreasing on $[z_0, \delta + u/2 - \tau]$ with $L(z_0) \geq 0$. Since $L(z) \to 0$ as $z \to \infty$ and $L(z)$ is either strictly decreasing or strictly increasing on interval $[\delta + u/2 - \tau, \infty)$, the maximum value is achieved at some point $\hat{z} = z_0 \in [\delta - u/2 - \tau, \delta + u/2 - \tau]$.

We have hence proved that if $\tau \leq u$ or $\tau \geq \delta - u/2$, then $\hat{z} \in [\delta - u/2 - \tau, \delta + u/2 - \tau]$. To conclude the proof, we need to show that in the next step of the algorithm (at the next attacker move), the invariant that $\tau \leq u$ or $\tau \geq \delta - u/2$ also holds.

Assume, for simplicity, that the previous defender move before current time t (when the attacker moved) happened at time $t_0 < t$. Then the defender's next move t_1 is in interval $[t_0 + \delta - u/2, t_0 + \delta + u/2]$. The attacker receives as feedback $\tau = t - t_0$ and his next move given by the Greedy algorithm is scheduled at time

$$t + \hat{z} = t_0 + \tau + \hat{z} \in [t_0 + \delta - u/2, t_0 + \delta + u/2].$$

There are two cases to consider:

– $t_1 \leq t + \hat{z}$ (the defender's next move is before the attacker's next move): Then the feedback received by the attacker at the next move is $\tau' = t + \hat{z} - t_1 \leq (t_0 + \delta + u/2) - (t_0 + \delta - u/2) = u$.
– $t + \hat{z} < t_1$ (the attacker's next move is before the defender's next move): Then the feedback received by the attacker at the next move is $\tau' = t + \hat{z} - t_0 \geq \delta - u/2$.

This concludes our proof. □

References

[1] Cloud Security Alliance. www.cloudaudit.org
[2] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, D. Song, Provable data possession at untrusted stores, in *Proc. 14th ACM Conference on Computer and Communication Security (CCS)* (2007)
[3] K.D. Bowers, A. Juels, A. Oprea, HAIL: A high-availability and integrity layer for cloud storage, in *Proc. 16th ACM Conference on Computer and Communication Security (CCS)* (2009)

[4] K.D. Bowers, M. van Dijk, A. Juels, A. Oprea, R. Rivest, Remote assessment of fault tolerance, in *Proc. 18th ACM Conference on Computer and Communication Security (CCS)* (2011)
[5] D.Y. Chan, M.A. Vasarhelyi, Innovation and practice of continuous auditing. *Int. J. Account. Inf. Syst.* **12**(2), 152–160 (2011)
[6] Y. Dodis, J. Katz, S. Xu, M. Yung, Key-insulated public key cryptosystems, in *Proc. IACR EUROCRYPT* (2002)
[7] N. Falliere, L.O. Murchu, E. Chien, W32.stuxnet dossier (2011). Symantec white paper
[8] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd edn. (Wiley, New York, 2011)
[9] R.G. Gallager, *Discrete Stochastic Processes* (Springer, Berlin, 1996)
[10] S.N. Hamilton, W.L. Miller, A. Ott, O.S. Saydjari, Challenges in applying game theory to the domain of information warfare, in *Information Survivability Workshop (ISW)* (2002)
[11] G. Itkis, Cryptographic tamper-evidence, in *Proc. 10th ACM Conference on Computer and Communication Security (CCS)* (2003)
[12] G. Itkis, *Handbook of Information Security* (Wiley, New York, 2006)
[13] A. Juels, B. Kaliski, PORs: Proofs of retrievability for large files, in *Proc. 14th ACM Conference on Computer and Communication Security (CCS)* (2007), pp. 584–597
[14] J. Katz, Bridging game theory and cryptography: recent results and future directions, in *Proc. Theory of Cryptography Conference (TCC)* (2008), pp. 251–272
[15] G.J. Mailath, L. Samuelson, *Repeated Games and Reputations: Long-Run Relationships* (Oxford University Press, London, 2006)
[16] P. Mell, T. Grance, The NIST definition of cloud computing. NIST Special Publication 800-145 (2011)
[17] T. Moore, A. Friedman, A. Procaccia, Would a "cyber warrior" protect us? Exploring trade-offs between attack and defense of information systems, in *Proc. New Security Paradigms Workshop (NSPW)* (2010), pp. 85–94
[18] R.B. Myerson, *Game Theory—Analysis of Conflict* (Harvard University Press, Cambridge, 1997)
[19] K.C. Nguyen, T. Alpcan, T. Basar, Security games with incomplete information, in *Proc. IEEE International Conference on Communications (ICC)* (2009)
[20] J. Pathak, B. Chaouch, R.S. Sriram, Minimizing cost of continuous audit: counting and time dependent strategies. *J. Account. Public Policy* **24**, 61–75 (2005)
[21] D. Pavlovic, Gaming security by obscurity (2011). CoRR abs/1109.5542
[22] R.L. Rivest, Illegitimi non carborundum. Invited keynote talk given at CRYPTO 2011, August 15 (2011). http://people.csail.mit.edu/rivest/pubs.html#Riv11b
[23] S. Ross, *Stochastic Processes* (Wiley, New York, 1996)
[24] S. Roy, C. Ellis, S. Shiva, D. Dasgupta, V. Shandilya, Q. Wu, A survey of game theory as applied to network security, in *Proc. Hawaii International Conference on System Sciences (HICSS)* (2010), pp. 1–10
[25] N.D. Schwartz, C. Drew, RSA faces angry users after breach. *New York Times*, p B1, 8 June 2011
[26] H. Shacham, B. Waters, Compact proofs of retrievability, in *Proc. IACR ASIACRYPT*. LNCS, vol. 5350 (2008), pp. 90–107
[27] R.S. Sutton, A.G. Barto, *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, 1998) (Bradford Book)
[28] Z. Zorz, NSA considers its networks compromised (2010). Referenced at http://http://www.net-security.org/secworld.php?id=10333