# Multi-Rate Threshold FlipThem

David Leslie[1], Chris Sherfield[2], and Nigel P. Smart[2]

[1] Department Mathematics and Statistics, University of Lancaster,
d.leslie@lancaster.ac.uk,
[2] Department of Computer Science, University of Bristol, UK,
c.sherfield@bristol.ac.uk, nigel@cs.bris.ac.uk.

**Abstract.** A standard method to protect data and secrets is to apply threshold cryptography in the form of secret sharing. This is motivated by the acceptance that adversaries will compromise systems at some point; and hence using threshold cryptography provides a defence in depth. The existence of such powerful adversaries has also motivated the introduction of game theoretic techniques into the analysis of systems, e.g. via the FlipIt game of van Dijk et al. This work further analyses the case of FlipIt when used with multiple resources, dubbed FlipThem in prior papers. We examine two key extensions of the FlipThem game to more realistic scenarios; namely separate costs and strategies on each resource, and a learning approach obtained using so-called fictitious play in which players do not know about opponent costs, or assume rationality.

## 1 Introduction

The traditional methodology of securing systems is to rely solely on cryptographic means to ensure protection of data (via encryption) and authentication (via signatures and secured-passwords). However, both of these techniques rely on some data being kept secure. The advent of attacks such as Advanced Persistent Threats (APTs) means that exfiltration of the underlying cryptographic keys, or other sensitive data, is now a real threat. To model such long term stealthy attacks researchers have turned to game theory [1, 21, 24, 28], adding to a growing body of work looking at game theory in cybersecurity in various scenarios [7, 15, 34]. Probably the most influential work applying game theory in the security area has been the FlipIt game [33]; with the follow up paper [5] demonstrating applications of FlipIt to various examples including credential management, virtual machine refresh and cloud auditing for service-level-agreement. FlipIt has gained traction and popularity due to the assumption that the adversary can always get into the system; thus aligning with the new rhetoric in the security industry that compromise avoidance is no longer a realistic possibility and it is now about limiting the amount of compromise as quickly and efficiently as possible.

In FlipIt, two players, aptly named the defender and attacker, fight over control of a single resource. Each player has their own button which, when pressed, will give them control over the resource assigning them some form of benefit. Pressing the button has a cost associated to it. The FlipIt paper examines this game as a way of modelling attacks in which a stealthy adversary is trying to control a single resource. For example it could be used to model an adversary which compromises passwords (the adversary's

button press corresponds to a break of a system password), whilst the defender resets passwords occasionally (via pressing their button in the game). FlipIt has of course been generalised in many different directions [10, 16, 18, 25, 23, 26].

However, a standard defence against having a single point of failure for secure data (be it real data or cryptographic secrets), is to use some form of distributed cryptography [9], usually using some form of secret sharing [29]. Such techniques can either be used directly as in [4, 30], or using secret sharing to actually compute some data as in Multi-Party Computation [2, 8]. To capture this and other such situations, Laszka et al. [17] introduce a FlipThem game in which there are multiple resources, each equipped with a button as in FlipIt, and the attacker obtains benefit only by gaining control of every resource in the game. This models the full threshold situation in distributed cryptographic solutions.

The FlipThem game is extended by Leslie et al. [19] to the partial threshold case, where the attacker is not required to gain control of the whole system but only a fraction of the number of resources in order to have some benefit. Assuming both players select the rates of Poisson processes controlling the pressing of buttons, they calculate the proportion of time the attacker will be in control of the system in order to gain some benefit. Nash equilibrium rates of play are calculated which depend on move costs and the threshold of resources required for the attacker to gain any benefit.

A major downside of the analysis in [19] is that the button associated to all resources are given the same cost and move rate for the attacker (and similarly for the defender). So in this current work we introduce the more realistic setting in which a player may have different move rates and costs associated to each resource. For example one resource may be easier to apply patches to than others, or one may be easier to attack due to the operating system it runs. We calculate Nash equilibrium rates of play for this more realistic setting.

We also introduce a framework from the learning in games literature [12] which models the situation where each player responds to the observed actions of the other player, instead of calculating an equilibrium. This adaptive framework is required once we drop the unrealistic assumption that players know their opponent's costs and reward functions, and our version makes considerably weaker assumptions on the information available to the players than the adaptive framework of [33]. In particular we introduce learning into a situation where the FlipThem game is played over a continuing sequence of epochs. We assume players repeatedly calculate the average observed rate of their opponent and respond optimally to that, resulting in a learning rule known as *fictitious play* [6, 12]. Performing multiple experiments, we find that when the costs result in a game with an interior equilibrium point (i.e. one in which all players play non-zero rates of all resources) the fictitious play procedure converges to this equilibrium. On the other hand, when there is not an interior equilibrium, we find unstable behaviour in the learning procedure. This result is important in the real world: the fictitious play formulation assumes that players know only their own benefit functions and can observe the play of others, yet the players still manage to converge to the calculated equilibria. Thus in these situations, even if our players are unable to calculate equilibrium strategies, their naïve optimising play will converge to an equilibrium and players' long term rewards are captured by the equilibrium payoffs.

## 2 Model

Our Multi-Rate Threshold FlipThem game has two players, an attacker and defender, fighting for control over the set of $n$ resources, $\mathcal{R} = \{\mathcal{R}_1, \ldots, \mathcal{R}_n\}$. Both players have buttons in front of each resource that when pressed will give them control over that specific resource. For each resource $\mathcal{R}_i$ the defender and attacker will play an exponential rate $\mu_i$ and $\lambda_i$, respectively, meaning that the times of moves for a player on a given resource will follow a Poisson Process with rate given by $\mu_i$ or $\lambda_i$. Using Markov Chain theory [13] we can construct explicit values for the proportion of time the each player is in control of resource $\mathcal{R}_i$ depending on their rates of play. This stationary distribution is given by $\pi^i = \left(\pi_0^i, \pi_1^i\right) = \frac{1}{\mu_i+\lambda_i}\left(\mu_i, \lambda_i\right)$, and indicates that the defender is in control of the resource $\mathcal{R}_i$ a proportion $\mu_i/(\lambda_i + \mu_i)$ of the time, and the attacker is in control a proportion $\lambda_i/(\lambda_i + \mu_i)$ of the time. We assume that each behaviour on each resource is independent of all the others, and hence the proportion of time that the attacker is control of a particular set of resources $C$, with the defender in control of the other resources, is simply given by the product of the individual resource proportions: $\prod_{i\in C} \frac{\lambda^i}{\lambda_i+\mu_i} \prod_{i\notin C} \frac{\mu^i}{\lambda^i+\mu_i}$.

At any point in time, the attacker will have compromised $k$ resources, whilst the defender is in control of the remaining $n - k$ resources, for some $k$. In order for the attacker to have some gain, she must compromise $t$ or more resources. The value $t$ is called the threshold, as used in a number of existing threshold security situations discussed in the introduction. From a game theory point of view, whenever $k \geq t$ the attacker obtains benefit, whilst when $k < t$ the defender obtains benefit.

From this we can construct benefit functions for both players. For the attacker it is the proportion of time she is in control of a number of resources over the threshold, such that $k \geq t$, penalised by a cost for moving. For the defender it is the proportion of time that she is in control of at least $n - t + 1$ resources, again penalised by a cost of moving. Thus, the benefit functions for attacker and defender respectively are given by

$$\begin{aligned}
\beta_D(\boldsymbol{\mu}, \boldsymbol{\lambda}) &= 1 - \sum_{\substack{C\subseteq\{1,\ldots,n\} \\ |C|\geq t}} \left[\prod_{i\in C} \frac{\lambda_i}{\lambda_i + \mu_i}\right] \cdot \left[\prod_{i\notin C} \frac{\mu_i}{\lambda_i + \mu_i}\right] - \sum_i d_i \cdot \mu_i \\
\beta_A(\boldsymbol{\mu}, \boldsymbol{\lambda}) &= \sum_{\substack{C\subseteq\{1,\ldots,n\} \\ |C|\geq t}} \left[\prod_{i\in C} \frac{\lambda_i}{\lambda_i + \mu_i}\right] \cdot \left[\prod_{i\notin C} \frac{\mu_i}{\lambda_i + \mu_i}\right] - \sum_i a_i \cdot \lambda_i
\end{aligned} \tag{1}$$

where the $a_i$ and $d_i$ are the (relative) move costs on resource $i$ for attacker and defender, which are assumed fixed throughout the game, and $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ are the vectors of rates over all resources for the defender and attacker, respectively, constrained to be non-negative. The benefit functions in (1) show that the game is non-zero-sum, meaning we are unable to use standard zero-sum results found in the literature [**?**].

## 3 Finding the equilibria of Multi-Rate $(n, t)$-FlipThem

We begin by finding the equilibria of the the multi-rate version of the FlipThem game with $n$ resources. This represents a more realistic scenario than previous studies, by

allowing players to favour certain resources based on differential costs of attacking or defending, for example when a company owns multiple servers located in different areas and with different versions of operating systems. We want to find a stationary point or equilibrium of the two benefit functions (1) in terms of purely the costs of moving on each resource. This would mean both players are playing at a rate that maximises their own benefit function with respect to their opponents play and move costs. Thus, they would not wish to deviate from their current playing strategy or rates. This is known as a Nash Equilibrium [22]. Our challenge in this article, compared with previous works such as [19], is that neither benefit function is trivial to optimise simultaneously across the vector of rates.

### 3.1 Full Threshold: Multi-rate $(n, n)$-FlipThem

We begin with the full threshold case in which the attacker must control all resources in order to obtain some benefit. The algebra is easier in this case; the partial threshold version is addressed in Section 3.2. For this full threshold case, the general benefit functions (1) simplify to

$$\begin{aligned}\beta_D(\boldsymbol{\mu}, \boldsymbol{\lambda}) &= 1 - \prod_{i=1}^{n} \frac{\lambda_i}{\mu_i + \lambda_i} - \sum_{i=1}^{n} d_i \cdot \mu_i \\ \beta_A(\boldsymbol{\mu}, \boldsymbol{\lambda}) &= \prod_{i=1}^{n} \frac{\lambda_i}{\mu_i + \lambda_i} - \sum_{i=1}^{n} a_i \cdot \lambda_i.\end{aligned} \tag{2}$$

Note that these benefit functions reduce to those of [19] if we set $\mu_i = \mu, \lambda_i = \lambda$, $a_i = \frac{a}{n}$ and $d_i = \frac{d}{n}$ for all $i$.

We start by finding the best response function of the defender, which is a function $\text{br}^D$ mapping attacker rates $\boldsymbol{\lambda}$ to the the set of all defender rates $\boldsymbol{\mu}$ which maximise defender payoff $\beta_D$ when the attacker plays rates $\boldsymbol{\lambda}$. A necessary, though not sufficient, condition for $\boldsymbol{\mu}$ to maximise $\beta_D$ is that each $\mu_i$ maximises $\beta_D$ conditional on the other values of $\boldsymbol{\mu}$. Furthermore, maxima with respect to $\mu_i$ occur either when the partial derivative $\frac{\partial \beta_D}{\partial \mu_i}$ is 0, or at a boundary of the parameter space. Equating this partial derivative to zero to gives

$$\frac{\partial \beta_D}{\partial \mu_i} = 0 \Rightarrow -\prod_{j=1}^{n} \lambda_j + d_i \cdot (\lambda_i + \mu_i)^2 \cdot \prod_{j=1, j\neq i}^{n} (\mu_j + \lambda_j) = 0.$$

This is a quadratic in $\mu_i$, meaning that fixing the defender's benefit function shown in (2) for all attacker rates $\boldsymbol{\lambda}$ and all defender rates $\mu_j$ where $j \neq i$, gives only two turning points. Since $\beta_D$ decreases to negative infinity as $\mu_i$ gets large, the two candidates for a maximising $\mu_i$ are at the upper root of this equation, or at $\mu_i = 0$. A non-0 $\mu_i$ must therefore satisfy

$$\mu_i = -\lambda_i + \sqrt{\frac{\lambda_i}{d_i} \cdot \prod_{j=1, j\neq i}^{n} \frac{\lambda_j}{(\mu_j + \lambda_j)}}. \tag{3}$$

Of course, a $\mu_i$ satisfying this equation could be negative and thus inadmissible as a rate, but all we claim for now is that any non-zero $\mu_i$ must be of the this form.

We can use the same method in differentiating the attacker's benefit with respect to her rate $\lambda_i$ for resource $\mathcal{R}_i$, equating this to zero and manipulating to give

$$\lambda_i = -\mu_i + \sqrt{\frac{\mu_i}{a_i} \cdot \prod_{j=1, j\neq i}^{n} \frac{\lambda_j}{(\mu_j + \lambda_j)}}. \tag{4}$$

Any Nash equilibrium in the interior of the strategy space (i.e. with strictly positive rates on all resources) must be a simultaneous solution of (3) and (4). Note that both equations can be rearranged to express $\lambda_i + \mu_i$ in terms of a square root, and then we can equate the square root terms to find that $\frac{\lambda_i}{\mu_i} = \frac{d_i}{a_i}$. Substituting this relationship back in to (3) and (4) gives

$$\lambda_i^* = \frac{d_i}{(a_i + d_i)^2} \cdot \prod_{\substack{j=1 \\ j\neq i}}^{n} \frac{d_j}{(a_j + d_j)}, \quad \mu_i^* = \frac{a_i}{(a_i + d_i)^2} \cdot \prod_{\substack{j=1 \\ j\neq i}}^{n} \frac{d_j}{(a_j + d_j)}. \tag{5}$$

If a company was reviewing its defensive systems and was able to calculate these equilibria, they can see the rate at which they'd have to move in order to defend their system optimally, under the assumption that the attacker is also playing optimally. (Of course, if the attacker was playing sub-optimally, the defender would be able to gain a larger pay off.) From these equilibria the parties are able to calculate the long run proportion of time each of their resources within the system would be compromised by looking at the stationary distribution shown in section 2. From this information, the company can also calculate the value of the benefit functions for each player when playing these rates. This is useful in the real world as it can be used by companies to make strategic decisions on the design of their systems. We do this by substituting the two equilibria back into the benefit functions (2). We can express these rewards in terms of the ratio between the costs of each resource, $\rho_j = \frac{a_j}{d_j}$, giving the dimensionless expression

$$\begin{aligned} \beta_D^* &= 1 - \prod_{i=1}^{n} \frac{1}{\rho_i + 1} \left[ 1 + \sum_{j=1}^{n} \frac{\rho_j}{\rho_j + 1} \right], \\ \beta_A^* &= \prod_{i=1}^{n} \frac{1}{\rho_i + 1} \left[ 1 - \sum_{j=1}^{n} \frac{\rho_j}{\rho_j + 1} \right]. \end{aligned} \tag{6}$$

We have expressed the payoffs to both players at the putative interior equilibrium. If this is an equilibrium, neither player will wish to deviate from these equilibrium rates. However we have thus far ignored the potential maximising $\mu_i$ at 0. While no individual $\mu_i$ or $\lambda_i$ will deviate unilaterally to zero (recall that the partial derivatives have only two zeroes, and thus payoff decreases as $\mu_i$ decreases to zero), if several rates simultaneously switch the payoff to a player could increase. We therefore consider what happens when rates can switch to zero, starting with the attacker.

By considering the attacker's benefit function (2), we can see that if the attacker plays a zero rate on any resource, then in order to maximise the payoff, she should play zero rates on the rest of the system. In other words, she should drop out of the game and receive zero reward. Thus by comparing the attacker payoff in (6) to zero we can see quickly whether this is indeed a point at which the attacker will be content.

For the defender, things are more complicated. However, we can see from the benefit function (2) that a zero payoff by withdrawing from the game is again an option, and so we compare the equilibrium payoff (6) to zero as a partial check on whether the fixed point (5) is an equilibrium. We do not not explicitly discount partial dropout of the defender, but note that by dropping out the defender is effectively reducing the game to a smaller number of servers, and hence should not have invested in the additional servers in the first place. Comparing the benefits in (6) to zero it is easy to see in this full threshold case that the defender's benefit $\beta_D^*$ is always non-negative when playing (5) meaning dropping out is not an equilibrium point for the defender, whereas $\beta_A^*$ can drop below 0. We use this to find a condition which must be satisfied for the point (5) to be an equilibrium. In particular, we require $\beta_A^*$ in (6) to be positive, and therefore

$$1 - \sum_{j=1}^{n} \frac{\rho_j}{\rho_j + 1} > 0. \tag{7}$$

Thus, we have a condition (7) that, if satisfied, the attacker will not drop out of the game. If the condition is not satisfied, the attacker will prefer to drop out of the game entirely than to play at the interior equilibrium. Ensuring condition (7) is met can thus be viewed as an design criterium for system engineers when designing defensive systems.

### 3.2 Partial Threshold: Multi-Rate $(n, t)$-FlipThem

So far we have extended the full threshold FlipThem game [17] by obtaining the equilibria of the benefit functions constructed from the proportion of time the attacker is in control of the whole system. In order to generalise this further, we return to our partial threshold case in which the attacker gains benefit from controlling only $t < n$ resources. Our general benefit functions for both players are written in (1); the analysis is analogous to methods demonstrated above in section 3.1. In this $(n, t)$-threshold case, the analogous best response conditions to (3) and (4) are

$$\mu_i = -\lambda_i + \sqrt{\frac{\lambda_i \cdot S_i}{d_i}} \quad \text{and} \quad \lambda_i = -\mu_i + \sqrt{\frac{\mu_i S_i}{a_i}}, \tag{8}$$

where we have introduced $S_i$ to denote

$$\sum_{\substack{C \subseteq \mathcal{A}_i \\ |C| \geq t}} \left[ \prod_{j \in C} \frac{\lambda_j}{\lambda_j + \mu_j} \right] \cdot \left[ \prod_{j \notin C} \frac{\mu_j}{\lambda_j + \mu_j} \right]$$

and $\mathcal{A}_i = \{1, \ldots, i-1, i+1, \ldots, n\}$. Interestingly, equating the square root terms results in the same relationship $\frac{\mu_i}{a_i} = \frac{\lambda_i}{d_i}$ as in Section 3.1. Finally, substituting this

relationship back into the best response functions (8) gives us

$$\begin{aligned}\mu_i^* &= \frac{a_i}{(a_i+d_i)^2}\cdot\sum_{\substack{C\subseteq\mathcal{A}_i\\|C|\geq t}}\left[\prod_{j\in C}\frac{d_j}{a_j+d_j}\right]\cdot\left[\prod_{j\notin C}\frac{a_j}{a_j+d_j}\right],\\ \lambda_i^* &= \frac{d_i}{(a_i+d_i)^2}\cdot\sum_{\substack{C\subseteq\mathcal{A}_i\\|C|\geq t}}\left[\prod_{j\in C}\frac{d_j}{a_j+d_j}\right]\cdot\left[\prod_{j\notin C}\frac{a_j}{a_j+d_j}\right].\end{aligned} \tag{9}$$

As in Section 3.1, (9) is a Nash equilibrium for the game, unless one or other player can improve their payoff by dropping out of one or more resources. We can substitute these rates back into the players' benefit functions as we did in the full threshold case in Section 3.1 to check that the payoffs at this putative equilibrium are non-negative. However, the formulae become very complicated to write down explicitly in general and we leave this to Section 4, when we deal with specific examples. Note this also means we do not have a clean condition analogous to (7) to test whether (9) is an equilibrium.

## 4 Introducing Fictitious Play into Multi-Rate $(n, t)$-FlipThem

While the equilibrium analysis above offers useful insight into the the security game Multi-Rate Threshold FlipThem, it can be viewed as an unrealistic model of real world play. In particular it is extremely unlikely the players have full knowledge of the payoff and move costs of their opponent, and therefore cannot calculate the equilibrium strategies. We now introduce game-theoretical learning, in which the only knowledge a player has of their opponent is the actions that they take. When the game is repeatedly played through time, players respond to their observations and attempt to improve their payoff. In this article we focus on a method of learning known as *fictitious play* [6, 3, 12].

We break the game up into periods of fixed length of time. At the end of period $\tau$ each player observes the number of times the button of each resource $i$ was pressed by their opponent in that period. Denote by $\lambda_i^\tau$ and $\mu_i^\tau$ the actual rate played by attacker and defender in period $\tau$, and use $\widetilde{\lambda_i}^\tau$ and $\widetilde{\mu_i}^\tau$ to denote the number of button presses by the attacker and defender respectively, normalised by the length of the time interval. After $\mathcal{T}$ plays of the game, each player averages the observations he has made of the opponent, resulting in estimates for each resource

$$\widehat{\lambda}_i^{\mathcal{T}} = \frac{1}{\mathcal{T}}\sum_{\tau=1}^{\mathcal{T}}\widetilde{\lambda}_i^\tau, \qquad \widehat{\mu}_i^{\mathcal{T}} = \frac{1}{\mathcal{T}}\sum_{\tau=1}^{\mathcal{T}}\widetilde{\mu}_i^\tau.$$

The players select their rates for time period $\mathcal{T}+1$ by playing a best response to their estimates;

$$\mu_i^{\mathcal{T}+1} = \mathrm{br}_i^D(\widehat{\boldsymbol{\lambda}}^{\mathcal{T}}), \qquad \lambda_i^{\mathcal{T}+1} = \mathrm{br}_i^A(\widehat{\boldsymbol{\mu}}^{\mathcal{T}}).$$

where $\widehat{\boldsymbol{\mu}}^{\mathcal{T}}$ and $\widehat{\boldsymbol{\lambda}}^{\mathcal{T}}$ are the defender and attacker's vector of rates played on each resource. If it were the case that opponent rates were constant, the averaging of observations over time would be an optimal estimation of those rates. Since both players are learning, the rates are not constant, and averaging uniformly over time does not result in statistically optimal prediction. However lack of a better informed model of rate evolution means that averaging is retained as the standard mechanism in fictitious play; see [20, 32] for attempts to move beyond this assumption.

```
Set maximum periods to be played;
Randomly assign move costs for each player;
Randomly choose initial starting rates for both players;
One period is played;
while Number of periods has not reached the maximum do
    Each player receives last period's observation of their opponent's play;
    Each player averages their opponent's history to form a belief;
    Each player uses their best response function to calculate rate for this period;
    Each player plays this rate;
end
```

**Algorithm 1:** Fictitious Play algorithm for multi-rate $(n, t)$-FlipThem

This fictitious play process is described in Algorithm 1, where we see the simplicity of the learning process and the sparsity of the information required by the players. The only challenging step is in calculating the best response function; as observed in Section 3 the best response of each player is not in general a simple analytical function of the rates of the opponent. From the defender's point of view we consider all subsets of the resources; setting the rates of these resources to zero, and solving (8) for the non-zero rates, we find a putative best response; the set of rates with the highest payoff given the fixed belief $\widehat{\boldsymbol{\lambda}}^{\mathcal{T}}$ is the best response. The attacker's best response is calculated analogously. An interesting question, which we address, is whether this simple learning process converges to the equilibria calculated previously in Section 3.

The process we have defined is a discrete time stochastic process. It is in actual fact a stochastic fictitious play process [11]; since the number of button presses in a period is Poisson with expected value the played rate multiplied by the length of the time interval, the observations can be seen to satisfy

$$\widetilde{\mu}_i^{\mathcal{T}+1} = \mathrm{br}_i^D(\widehat{\boldsymbol{\lambda}}^{\mathcal{T}}) + M_{\mu,i}^{\mathcal{T}+1}, \qquad \widetilde{\lambda}_i^{\mathcal{T}+1} = \mathrm{br}_i^A(\widehat{\boldsymbol{\mu}}^{\mathcal{T}}) + M_{\lambda,i}^{\mathcal{T}+1},$$

where $\mathbb{E}(M_{\cdot,\cdot}^{\tau+1}|\mathcal{F}^{\tau}) = 0$ if $\mathcal{F}^{\tau}$ denotes the history of the process up to time $\tau$. The methods of [3] then apply directly to show that the convergence (or otherwise) of the discrete stochastic process is governed by the continuous deterministic differential equations

$$\frac{\mathrm{d}\boldsymbol{\lambda}}{\mathrm{d}t} = \mathrm{br}(\boldsymbol{\mu}) - \boldsymbol{\lambda}, \qquad \frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = \mathrm{br}(\boldsymbol{\lambda}) - \boldsymbol{\mu}. \tag{10}$$

In standard fictitious play analyses, one might show that solutions of (10) are globally convergent to the equilibrium set. This is commonly achievable only in some classes of games, and since we do not have a zero-sum game we have not been able to show the required global convergence of (10).

### 4.1 Original FlipIt

The game of FlipIt can be considered a special case of our game of multi-rate $(n, t)$-FlipThem, seen by setting $n = t = 1$. This has the advantage that the best responses can be written in closed form, and we can use (10) to set up a two dimensional ordinary differential equation in terms of the players' rates and time. We start by writing the benefit functions for this special case

$$\beta_D(\mu, \lambda) = 1 - \frac{\lambda}{\mu + \lambda} - d\mu, \qquad \beta_A(\mu, \lambda) = \frac{\lambda}{\mu + \lambda} - a\lambda. \tag{11}$$

Differentiating the players' benefit functions (11) in terms of their respective resource rates and then solving for these gives the best response functions

$$\text{br}^D(\lambda) = \left(-\lambda + \sqrt{\frac{\lambda}{d}}\right)^+, \qquad \text{br}^A(\mu) = \left(-\mu + \sqrt{\frac{\mu}{a}}\right)^+$$

where $(x)^+ = \max(x, 0)$. The ordinary differential equation (10) becomes

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t} = \left(-\mu + \sqrt{\frac{\mu}{a}}\right)^+ - \lambda, \qquad \frac{\mathrm{d}\mu}{\mathrm{d}t} = \left(-\lambda + \sqrt{\frac{\lambda}{d}}\right)^+ - \mu. \tag{12}$$

This is a two dimensional ordinary differential equation in terms of the players' rates and time. The plot of the phase portrait of this is shown in Figure 1. where we have used $d = 0.1, a = 0.3$ as the move costs. It easy to see that the arrows demonstrating the direction of the rates over time converge upon a single point. This point is the equilibrium we can calculate easily from the more general equilibria derived in Section 3.2. We can also use Algorithm 1 to plot trajectories of the system in order to view convergence of the system; the convergence is monotonic and uninteresting so we omit the plots.

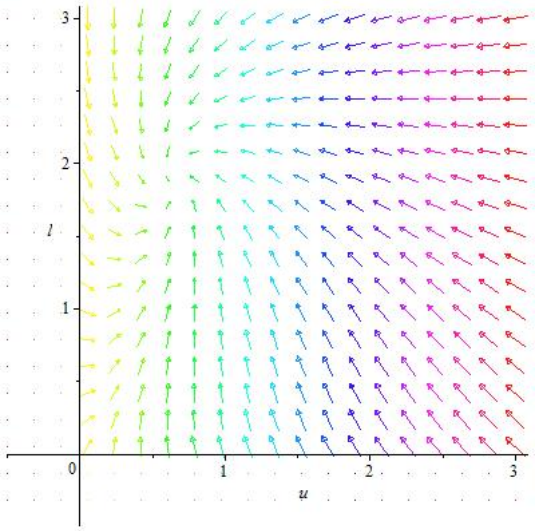


**Fig. 1.** Phase Portrait of (12) with $d = 0.1, a = 0.3$

### 4.2 $(n, t)$-FlipThem

A multi-resource example in which we retain one rate per player (as opposed to one rate per resource for each player) is given by the situation in [17] and [19]; each player chooses a rate at which to play all resources. Whilst this allows us to retain a two-dimensional system when considering the multiple resource case, unfortunately obtaining explicit best response functions is extremely difficult. Therefore, we revert to Algorithm 1 using time intervals of length 100; we fix this time interval for all further experiments in this article.

As in those previous works, we consider the stationary distribution of the system, with the defender playing with rate $\mu$ and attacker rate $\lambda$. This results in a stationary distribution for the whole system, given by

$$\pi = \frac{1}{(\mu+\lambda)^n}\left(\mu^n, n\cdot\lambda\cdot\mu^{n-1}, \ldots, \binom{n}{k}\cdot\mu^{n-k}\cdot\lambda^k, \ldots, n\cdot\mu\cdot\lambda^{n-1}, \lambda^n\right),$$

where the states correspond to the number of compromised resources, ranging from from 0 to $n$. Benefit functions for both players are given by

$$\beta_D(\mu,\lambda) = 1 - \sum_{i=t}^{n}\pi_i - d\cdot\mu = 1 - \frac{1}{(\mu+\lambda)^n}\cdot\sum_{i=t}^{n}\binom{n}{i}\cdot\mu^{n-i}\cdot\lambda^i - d\cdot\mu,$$

$$\beta_A(\mu,\lambda) = \sum_{i=t}^{n}\pi_i - a\cdot\lambda = \frac{1}{(\mu+\lambda)^n}\cdot\sum_{i=t}^{n}\binom{n}{i}\cdot\mu^{n-i}\cdot\lambda^i - a\cdot\lambda,$$

and best responses are calculated by differentiating these benefit functions, as in [19] and Section 3. In Figure 2 we plot the rate of the attacker and defender respectively by

applying Algorithm 1 to random starting rates for both players, with $(3, 2)$-threshold and costs $a = 0.65$ and $d = 0.5$. The straight horizontal lines represents the players' Nash equilibrium rates that can be calculated as a special case of the general equilibrium from (9). Note that we have chosen these costs in order to produce positive benefits for both players whilst playing the calculated Nash equilibrium. We see that both the defender's and attacker's mean rate converges to these lines.

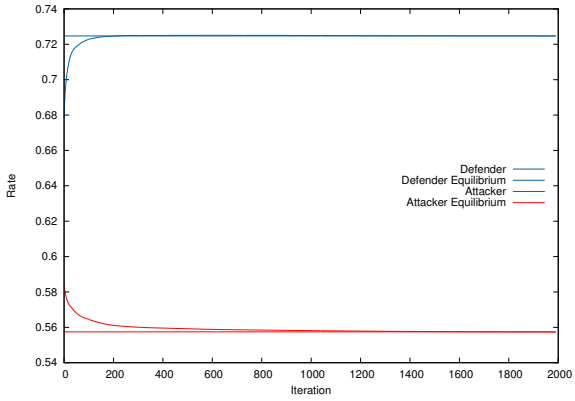


**Fig. 2.** Mean of the defender's and attacker's rate with $(3, 2)$-threshold and ratio $\rho = \frac{a}{d} = 1.3$.

### 4.3 Multi-Rate $(n, t)$-FlipThem

Finally, we come to the most general case of the paper, Multi-Rate $(n, t)$-FlipThem. As observed previously, depending on the player's belief of their opponents rates on each resource, they may choose to drop out of playing on a certain resource, or perhaps even all of them. Our best response functions must iterate through all possibilities of setting the resource rates to zero and chooses the configuration with the highest benefit as the best response. This solution is then used as $\mu^{\mathcal{T}+1}$ or $\lambda^{\mathcal{T}+1}$ for the following period $\mathcal{T} + 1$.

We want to find a condition such that we can gain some insight as to whether in this most complicated setting our iterative learning rule will converge to the equilibria we calculated analytically in Section 3.2. We experimented with multiple combinations of $n$ and $t$, randomly simulating costs of both players. From these empirical results, we observe that convergence occurs whenever the internal fixed point (9) results in non-negative benefits to both players.

**Specific case ($n = 3, t = 2$):** In order to illustrate the outcomes of our fictitious play algorithm we specify our threshold $(n, t) = (3, 2)$, and choose two representative cases of our randomly simulated examples. These particular examples were selected for ease of display, allowing us to illustrate their properties of convergence (or divergence)

clearly on just one figure. The first, when the benefits are positive and the internal equilibrium is not ruled out, we term 'success'. The second is an example in which the internal fixed point is not an equilibrium, and we term this case 'failure'.

*Success:* Our first example is with ratios $(\rho_1, \rho_2, \rho_3) \approx (0.7833, 0.7856, 0.7023)$ and attacker costs $(a_1, a_2, a_3) \approx (0.6237, 0.5959, 0.5149)$. Thus, the benefit values at equilibrium are $\beta_D^* \approx 0.0359$ and $\beta_A^* \approx 0.2429$. Since we have positive payoff for both players we expect convergence of the learning algorithm. This is exactly what we observe in Figure 3; convergence of rates to the lines representing the equilibria calculated in Section 3.2.

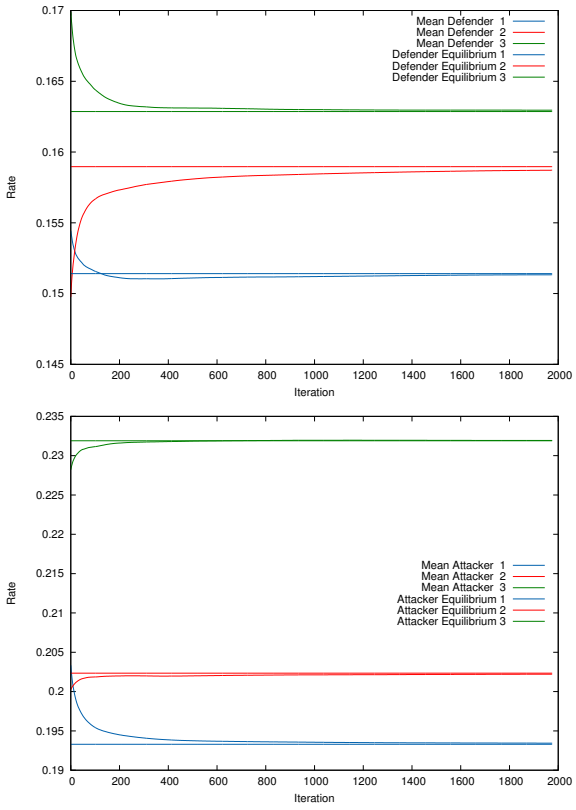


**Fig. 3.** Mean of the defender's rates (top) and attacker's rates (bottom) on all resources with $(3, 2)$-threshold and ratios $(\rho_1, \rho_2, \rho_3) \approx (0.7833, 0.7856, 0.7023)$

*Failure:* Our second example shows a lack of convergence when conditions are not met. Therefore, we choose ratios to be $(\rho_1, \rho_2, \rho_3) \approx (0.5152, 0.5074, 0.5010)$ and attacker costs $(a_1, a_2, a_3) \approx (0.2597, 0.2570, 0.2555)$. This gives 'equilibrium' benefits for the

defender and attacker of $\beta_D^* \approx -0.0354$ and $\beta_A^* \approx 0.4368$. The defender's benefit in this situation is negative, meaning we expect a lack of convergence. Figure 4 shows the development of both players' rates over time. We can see the rates do not approach anywhere near the equilibrium. Intuitively, this makes sense as the defender will not choose to end up playing in a game in which she receives negative payoff when she can drop out of the game completely in order to receive zero payoff.

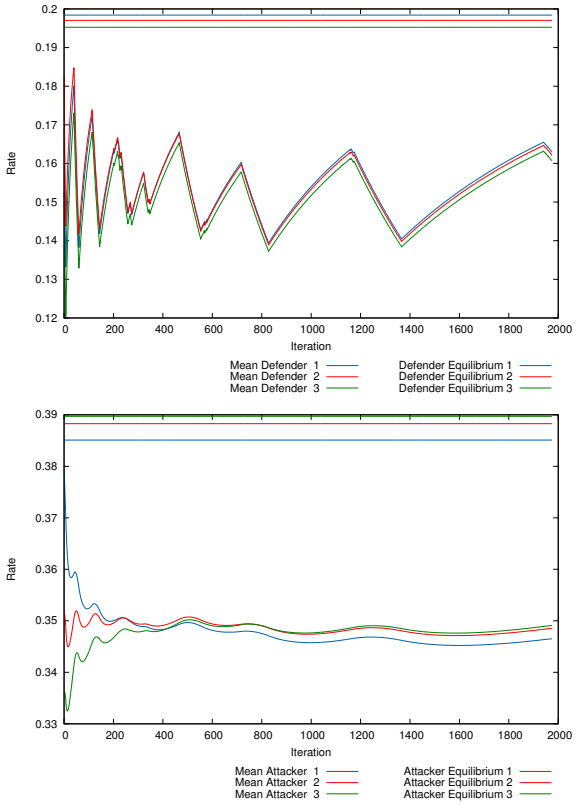


**Fig. 4.** Mean of the defender's rates (top) and attacker's rates (bottom) on all resources with $(3, 2)$-threshold and ratios $(\rho_1, \rho_2, \rho_3) \approx (0.5152, 0.5074, 0.5010)$.

We can see evidence of this dropping out in Fig. 5, which shows the rates the defender is actually playing on the resource (rather than the mean over time). The defender has certain periods where she drops out of the game entirely. The attacker's mean rates then start to drop until a point when the defender decides to re-enter the game.

To see a reason for this volatility, Figure 6 shows the defender's payoff surface for nearby attacker strategies on either side of a 'dropout' event from Fig. 5. We fix the defender's rate on resource 1 and plot the benefit surface as the rates on resources 2 and 3 vary. We also plot the plane of zero reward. It's easy to observe that the maximum

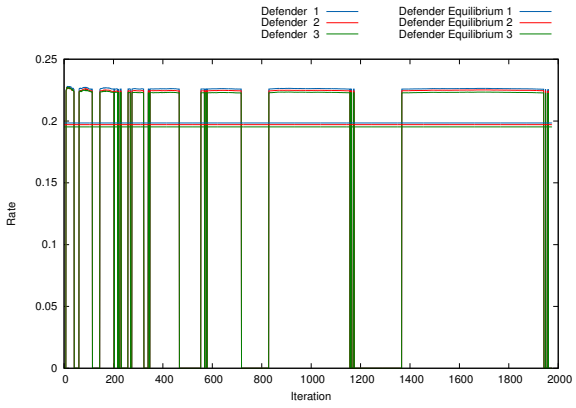


**Fig. 5.** Defender's actual rates on all resources with $(3, 2)$-threshold and ratios $(\rho_1, \rho_2, \rho_3) \approx (0.5152, 0.5074, 0.5010)$.

of this reward surface is above 0 in the left hand plot, but a small perturbation of the attacker rates pushes the maximal defender benefit below zero in the right hand plot, thus forcing the defender to drop out entirely. We conjecture that as the ratios decrease (and therefore become more costly for the defender) the defender drops out of the game more often within this learning environment.

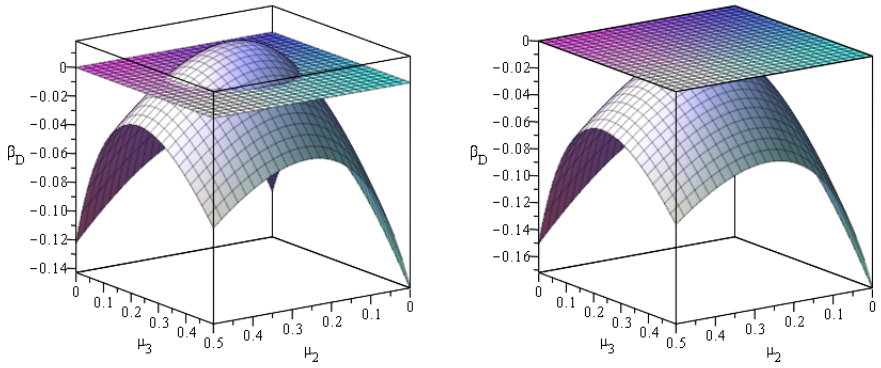


**Fig. 6.** Two snapshots of the defender's reward surface with a slight perturbation in attackers rates. (Left) The Payoff is just above the 0-plane. (Right) The whole payoff surface is below the 0-plane.

## Acknowledgements

The second author was supported by a studentship from GCHQ. This work has been supported in part by ERC Advanced Grant ERC-2015-AdG-IMPaCT and by EPSRC via grant EP/N021940/1.

## References

1. H. S. Bedi, S. G. Shiva, and S. Roy. A game inspired defense mechanism against distributed denial of service attacks. *Security and Communication Networks*, 7(12):2389–2404, 2014.
2. M. Ben-Or, S. Goldwasser, and A. Wigderson. Completeness theorems for non-cryptographic fault-tolerant distributed computation (extended abstract). In Simon [31], pages 1–10.
3. M. Benaim and M. W. Hirsch. Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. *Games and Economic Behaviour*, 29:36–72, 1999.
4. D. Boneh, X. Boyen, and S. Halevi. Chosen ciphertext secure public key threshold encryption without random oracles. In D. Pointcheval, editor, *Topics in Cryptology - CT-RSA 2006, The Cryptographers' Track at the RSA Conference 2006, San Jose, CA, USA, February 13-17, 2006, Proceedings*, volume 3860 of *Lecture Notes in Computer Science*, pages 226–243. Springer, 2006.
5. K. D. Bowers, M. van Dijk, R. Griffin, A. Juels, A. Oprea, R. L. Rivest, and N. Triandopoulos. Defending against the unknown enemy: Applying FlipIt to system security. In Grossklags and Walrand [14], pages 248–263.
6. G. Brown. Iterative solution of games by fictitious play. *Koopmans, T.C. (Ed.), Activity Analysis of Production and Allocation*, pages 374–376, 1951.
7. H. Çeker, J. Zhuang, S. J. Upadhyaya, Q. D. La, and B. Soong. Deception-based game theoretical approach to mitigate DoS attacks. In Zhu et al. [35], pages 18–38.
8. D. Chaum, C. Crépeau, and I. Damgård. Multiparty unconditionally secure protocols (extended abstract). In Simon [31], pages 11–19.
9. Y. Desmedt. Threshold cryptography. *European Transactions on Telecommunications*, 5(4):449–458, 1994.
10. S. Farhang and J. Grossklags. Flipleakage: A game-theoretic approach to protect against stealthy attackers in the presence of information leakage. In Zhu et al. [35], pages 195–214.
11. D. Fudenberg and D. Kreps. Learning mixed equilibria. *Games and Economic Behavior*, 5:320–367, 1993.
12. D. Fudenberg and D. K. Levine. *The Theory of Learning in Games*. The MIT Press, 1st edition edition, 1998.
13. G. Grimmett and D. Stirzaker. *Probability and Random Processes*. Oxford University Press, 3rd edition edition, 2001.
14. J. Grossklags and J. C. Walrand, editors. *Decision and Game Theory for Security - Third International Conference, GameSec 2012, Budapest, Hungary, November 5-6, 2012. Proceedings*, volume 7638 of *Lecture Notes in Computer Science*. Springer, 2012.
15. A. R. Hota, A. A. Clements, S. Sundaram, and S. Bagchi. Optimal and game-theoretic deployment of security investments in interdependent assets. In Zhu et al. [35], pages 101–113.
16. P. Hu, H. Li, H. Fu, D. Cansever, and P. Mohapatra. Dynamic defense strategy against advanced persistent threat with insiders. In *Computer Communications (INFOCOM), 2015 IEEE Conference on*, pages 747–755. IEEE, 2015.

17. A. Laszka, G. Horvath, M. Felegyhazi, and L. Buttyan. FlipThem: Modeling targeted attacks with FlipIt for multiple resources. In Poovendran and Saad [27], pages 175–194.
18. A. Laszka, B. Johnson, and J. Grossklags. Mitigating covert compromises. In *International Conference on Web and Internet Economics*, pages 319–332. Springer, 2013.
19. D. Leslie, C. Sherfield, and N. P. Smart. Threshold FlipThem: When the winner does not need to take all. In M. H. R. Khouzani, E. Panaousis, and G. Theodorakopoulos, editors, *Decision and Game Theory for Security - 6th International Conference, GameSec 2015, London, UK, November 4-5, 2015, Proceedings*, volume 9406 of *Lecture Notes in Computer Science*, pages 74–92. Springer, 2015.
20. D. S. Leslie and E. J. Collins. Generalized weakened fictitious play. *Games and Economic Behavior*, 56:285–298, 2006.
21. B. Z. Moayedi and M. A. Azgomi. A game theoretic framework for evaluation of the impacts of hackers diversity on security measures. *Rel. Eng. & Sys. Safety*, 99:45–54, 2012.
22. J. Nash. Non-cooperative games. *The Annals of Mathematics*, 54:286–295, 1951.
23. A. Nochenson and J. Grossklags. A behavioral investigation of the FlipIt game. In *Proceedings of the 12th Workshop on the Economics of Information Security (WEIS)*, 2013.
24. E. Panaousis, A. Fielder, P. Malacaria, C. Hankin, and F. Smeraldi. Cybersecurity games and investments: A decision support approach. In Poovendran and Saad [27], pages 266–286.
25. J. Pawlick, S. Farhang, and Q. Zhu. Flip the cloud: cyber-physical signaling games in the presence of advanced persistent threats. In *International Conference on Decision and Game Theory for Security*, pages 289–308. Springer, 2015.
26. V. Pham and C. Cid. Are we compromised? modelling security assessment games. In Grossklags and Walrand [14], pages 234–247.
27. R. Poovendran and W. Saad, editors. *Decision and Game Theory for Security - 5th International Conference, GameSec 2014, Los Angeles, CA, USA, November 6-7, 2014. Proceedings*, volume 8840 of *Lecture Notes in Computer Science*. Springer, 2014.
28. S. Rass and Q. Zhu. GADAPT: A sequential game-theoretic framework for designing defense-in-depth strategies against advanced persistent threats. In Zhu et al. [35], pages 314–326.
29. A. Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, 1979.
30. V. Shoup. Practical threshold signatures. In B. Preneel, editor, *Advances in Cryptology - EUROCRYPT 2000, International Conference on the Theory and Application of Cryptographic Techniques, Bruges, Belgium, May 14-18, 2000, Proceeding*, volume 1807 of *Lecture Notes in Computer Science*, pages 207–220. Springer, 2000.
31. J. Simon, editor. *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*. ACM, 1988.
32. M. Smyrnakis and D. S. Leslie. Dynamic opponent modelling in fictitious play. *The Computer Journal*, 53(9):1344–1359, 2010.
33. M. van Dijk, A. Juels, A. Oprea, and R. L. Rivest. FlipIt: The game of "stealthy takeover". *J. Cryptology*, 26(4):655–713, 2013.
34. Z. Zhou, N. Bambos, and P. W. Glynn. Dynamics on linear influence network games under stochastic environments. In Zhu et al. [35], pages 114–126.
35. Q. Zhu, T. Alpcan, E. A. Panaousis, M. Tambe, and W. Casey, editors. *Decision and Game Theory for Security - 7th International Conference, GameSec 2016, New York, NY, USA, November 2-4, 2016, Proceedings*, volume 9996 of *Lecture Notes in Computer Science*. Springer, 2016.