Flip the virus: A gametheoretic approach to cybersecurity

Sophie Marien

***Samenvatting*—Recently the attacks on Belgacom and other high profile targeted attacks shows us that even the most secure companies can still be compromised. It also shows that these attacks are not immediately detected. FlipIt is a framework that can model these stealthy takeovers and is proposed by a group of researchers at RSA. It is a 2-players game composed of a single attacker, a single defender and a single shared resource. These players will try to gain control over the shared resource and they do this in a stealthy way. In this paper we try to adapt FlipIt in a way that we can use it to model the game of defending a company network that is attacked by a virus. The FlipIt formulas are adapted for this virus propagation. Through analytic results we can provide useful information for the defender to defend his network.**

I. Introduction

IN THIS ERA where digitalization becomes prominent in every aspect of our lives, where technology is growing fast and where businesses are always under attack, security becomes an issue of increasing complexity. Without security, there is no protection to keep somebody out of a system. It is the same as leaving the door of your house wide open for everyone to come in.

Why is it so important to keep a system secure? Many businesses store confidential information on clients, which can be lost and possible be abused by competitors through data leakage. Also, disruption caused by DOSS attacks, may restuls in businesses failing to meet their service-level agreements. Ultimately, system and network security helps protecting a business's reputation, which is one of its most important assets.

A particular kind of frequently occurring threaths are Advanced Persistent Threats (APT). An APT is a targeted cyber attack that targets organisations in a stealthy way and that can stay undetected for a long period. This makes it so hard to protect a network or a system against an APT. Bruce Schneier describes an APT as something different and stronger than a conventional threat: "*A conventional hacker or criminal isn't interested in any particular target. He wants a thousand credit card numbers for fraud, or to break into an account and turn it into a zombie, or whatever. Security against this sort of attacker is relative; as long as you're more secure than almost everyone else, the attackers will go after other people, not you. An APT is different; it's an attacker who - for whatever reason - wants to attack you. Against this sort of attacker, the absolute level of your security is what's important. It doesn't matter how secure you are compared to your peers; all that matters is whether you're secure enough to keep him out*" - Bruce Schneier: APT is a Useful Buzzword [].

II. Stealthy takeover model

Since it is so difficult to protect a system or a network against APT's, researchers have been looking for effective ways to predict in advance which defence strategy might be the better one. Game theory is gaining increasing interest as an effective technique to model and study Cyber Security. Game theory analyses the security problem as a game where the players are an attacker and a defender of a system, and where both players have to make decisions. In particular, both players will aim for the strategy that results in a maximal benefit for them. Researchers at RSA made a game theoretic framework to model targeted attacks. They study the specific scenario where a system or network is repeatedly taken over completely by an attacker and this attack is not immediately detected by the defender of the system or network. In game theory, such a game is known as "FlipIt" [2]. This is a two players game where the attacker and the defender are competing to get control over a shared resource. Both players do not know who is currently in control of the resource until they move. In FlipIt every move gives them immediately control over the resource. But what if the attacker moves and it takes a while before the attacker gets full control over the resource? FlipIt does not take into account that a move may not be instantaneous, but has a certain delay. Consider for example a network with different nodes (laptops, datacenters) as a resource. The attacker drops a virus on one of the nodes and then wait till this virus infects the whole network. The attacker will only be in control of the resource once the whole network is infected.

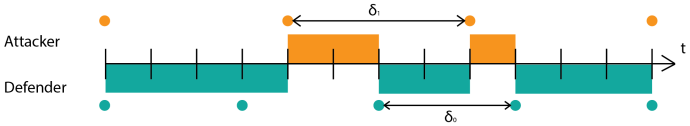
III. The FlipIt game

FlipIt is a game introduced by van Dijk et al. To understand how to model a FlipIt game with virus propagation it is important to get familiar with the concepts of the normal FlipIt game and its notations. Therefore, we first explain the framework of FlipIt and introduce the most important formulas that will be used throughout the paper.

FlipIt is a two-players game with a shared single resource that the players want to control as long as possible. The shared resource can be a password, a network or a secret key depending on the setting being modelled. In the remainder of the paper we name the two players the attacker, denoted by the subscript A and the defender, denoted by subscript D.

The game begins at $t = 0$ and continues indefinitely ($t \to \infty$). The time in the game is assumed as being continuous, but a discrete time could also be considered. To get control over the resource, the players i, with $i \in \{A, D\}$, can flip the resource at any given time. A flip will be regarded as a move from a player *i*. Each move will imply a certain cost k_i and the cost can vary for each player. Both players will try to minimize their cost. Adding a cost will prevent players to move too frequently.

The unique feature of FlipIt is that every move will happen in a stealthy way, meaning that the player has no clue that the other player (his adversary) has flipped the resource. For instance, the defender will not find out if the resource has already been compromised by the attacker until he flips the resource himself. The goal of the player is to maximize the time that he or she has control over the resource while minimizing the total cost of the moves. A move can also result in a "wasted move", called a flop. It may happen that the resource was already under control by the player. If the player moves when he or she has already control over the resource, he or she would have wasted a move since it does not result in a change of ownership, so the cost is wasted.



Figuur 1: A representation of a FlipIt game where both players are playing periodically and at discrete time intervals. Every move or flip is indicated by a blue or orange circle. The attacker is represented in orange and plays with a period of $\delta_A = 4$. The defender is represented in blue and plays with a period of $\delta_D = 3$. The blue and orange rectangles represent the amount of time the respective player is in control of the resource.

The state of the resource is denoted as a time-dependent variable $C = C_i(t)$. $C_D(t)$ is 1 if the game is under control by the defender and 0 if the game is under control by the attacker. Reversely, $C_A(t)$ will be 1 if the game is under control by the attacker and 0 if under controle by the defender. So, $C_A(t) = 1 - C_D(t)$. The game starts with the defender being in control: $C_D(0) = 1$.

The players receive a benefit equal to the time units they were in possession of the resource minus the cost of making their moves. The cost of a player *i* is denoted by k_i. The total gain of player *i* is equal to the total amount of time that a player *i* has owned the resource from the beginning of the game up to time *t*. It is expressed as follows:

$$G_i(t) = \int_0^t C_i(x)dx. \tag{1}$$

If we add up the gain of the defender and the gain of the attacker it should sum up to t:

$$G_D(t) + G_A(t) = t \tag{2}$$

The average gain rate of player *i* is defined as:

$$\gamma_i(t) = G_i(t)/t. \tag{3}$$

And thus for all $t > 0$:

$$\gamma_D(t) + \gamma_A(t) = 1 \tag{4}$$

Let $\beta_i(t)$ denote player's *i* average benefit upto time *t*:

$$\beta_i(t) = \gamma_i(t) - k_i\alpha_i. \tag{5}$$

This is equal to the fraction of time the resource has been owned by player *i*, minus the cost of making the moves. α_i defines the average move rate by player *i* up to time *t*. In a given game, the asymptotic benefit rate (or simply benefit) will be defined as the lim inf of the average benefit because time t will increase to infinity and the average benefit may not have limiting values.

$$\beta_i(t) = \lim_{t\to\infty} inf\beta_i(t) \tag{6}$$

1) strategies: Because the players move in a stealthy way, there are different types of feedback that a player can get while moving. These types of feedback can be divided into two groups of strategies. The non-adaptive strategies and the adaptive strategies. These are described in table I.

If there is no feedback for neither of the players, we have a non-adaptive strategy. Because a player does not receive any feedback during the game he will play in the same manner against every opponent. The strategy is called non-adaptive because the playing strategy is not dependent on the opponents movements. An interesting subclass of the non-adaptive strategies is the one where the time intervals between two consecutive moves are generated by a renewal process. An example of such renewal strategy is the periodic strategy where the time between two consecutive moves of the players are a fixed interval. An exponential strategy is a renewal strategy in which the interval between two consecutive moves is exponentially distributed.
In case there is feedback, a player can adapt his strategy to the information received about the opponent's moves. Depending on the amount of information received, two subclasses of adaptive strategies can be identified. The Last Move (LM) strategies represent the class where whenever a player flips he will find out the exact time that the opponent played the last time. In the second class, called Full History (FH), whenever a player flips he will find out the whole history of the opponent's move.
In this paper we will focus on the non-adaptive strategies. This choice is motivated by the fact that in a security game a player (defender or attacker) rarely has information about the moves (last move or full history) of his opponent.

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

Tabel I: Hierarchy of Classes of strategies in FlipIt

The study of the different strategies by means of FlipIt framework allows to derive a number of interesting results:

- periodic games dominate the other renewal strategies, meaning that it is always advantagous to play periodically against an opponent with a renewal strategy;
- periodic games are disadvantagous against players following a Last Move adaptive strategy;
- if the defender plays with a periodic rate that is fast enough he'll force the attacker to drop out;
- any amount of feedback about the opponent received during the game, benefits to a player.

IV. FlipIt with virus propagation

A FlipIt game consists of a single resource. To represent the security problem, the game now defines its single resource as a computer network with multiple nodes. One of the players, the defender, will try to defend his network. The defender will do this by flipping all the nodes of the network (i.e. the entire resource) in every move he plays. The attacker, the other player, will try to infect all the nodes in the network. The attacker will do this by flipping the node in the graph that can infect all the nodes in the shortest possible time. After dropping a virus on the first node, it takes a while for the virus to infect the entire network. However, since the original FlipIt game works with a single resource that is always flipped entirely, the assumption is made that the attacker is considered to gain immediate full control over the resource when the network has been infected, even it is only one node that has been infected.

In reality however, after dropping a virus on the first node, it takes a while for the virus to infect the entire network. So, the assumption that the attacker has full control over the resource as soon as a node has been infected, is not realistic. The attacker has only control of the network once all or a sufficient large number of nodes are infected. The time that it takes for the virus to infect every node (or a sufficient number of nodes) will be denoted as an infection-delay variable *d* (called 'delay' for short in the remainder of this paper). If we want to measure how long it takes for the virus to infect all the nodes in the network, we have to calculate the shortest path from the first infected node to the farthest node. Rather than denoting the time needed for infecting *all* the nodes, the variable d can also be used to denote the time needed to infect *a sufficient large number* of nodes.

Assume that an attacker attacks at time *t*, he doesn't get immediate control over the resource, but he only gains control at time *t* + *d*, with d denoting the time needed to infect a sufficiently large number (or all) nodes. If the defender flips the network before the period d has elapsed (so, somewhere between t and $t + d$), then the attacker will never gain full control over the resource. This implies that the mathematical formulas for gain and benefit need to be adapted to the fact that the attacker loses part of its benefit because of this delay. In the remainder of this paper, we will adapt the formalization of the FlipIt game using the variable d.

The formalization starts from the model of the non-adaptive continuous basic FlipIt game where players use a periodic strategy with a random phase. This choice is motivated by the assumption that in most organisations, the defence strategy is to periodically defend the network. This corresponds to a periodic defender strategy. A periodic attacker strategy is assumed as well, as this also corresponds to a common real life strategy. Further research can investigate the effect of relaxing this assumption.

Similarly as in [], we split the formalization in two cases. The first case is where the defender plays at least as fast as the attacker, the second case is where the attacker plays at least as fast as the defender. For each of these cases, first the benefit formula of the basic case without delay is presented, and then the delay is introduced.

A. Formalization the benefit formula including the infection-delay

A Periodic strategy is a non-adaptive renewal strategy where the time intervals between consecutive moves are a fixed period, denoted by δ. Moreover it has a random phase, that is chosen uniformly and random in the interval $[0, \delta]$ for the first move. The average rate of play of a player is denoted by $\alpha_i = \frac{1}{\delta_i}$.

Case 1: $\delta_D \leq \delta_A$ *(The defender plays at least as fast as the attacker.)*

Let $r = \frac{\delta_D}{\delta_A}$. The intervals between two consecutive defender's moves have length δ_D. 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 $\delta_D \leq \delta_A$) and his move is distributed uniformly at random.

The expected period of attacker control within the interval would be r/2, without considering the delay by a virus. Therefore the benefit for the attacker, without considering the delay, can be expressed as follows:

$$\beta_A(\alpha_D, \alpha_A) = \frac{r}{2} - k_A\alpha_A = \frac{\delta_D}{2\delta_A} - k_A\alpha_A \qquad (7)$$

Correspondeingly, the benefit for the defender can be expressed as:

$$\beta_D(\alpha_D, \alpha_A) = 1 - \frac{r}{2} - k_D\alpha_D = 1 - \frac{\delta_D}{2\delta_A} - k_D\alpha_D \quad (8)$$

However, because of the delay required for virus propagation, the maximal time of control is reduced to $\delta_D - d$. There is a probability of *r* that the attacker will move in the interval of the defender. However, the gain will not be half of the interval. Indeed, the attacker has to play soon enough to gain control, meaning that the attacker has to play during the period of $\delta_D - d$ during the interval of the defender. The probability that the attacker plays soon enough is $\frac{\delta_D - d}{\delta_D}$ and this will give the attacker an avarage gain of $\frac{\delta_D - d}{2}$. If the attacker moves after the period of $\delta_D - d$, the gain of the attacker will be zero. The probability that this happens is $\frac{d}{\delta_D}$. The average gain rate of the attacker can then be expressed as follows if we look at one interval of the defender:

$$\gamma_A(\alpha_D, \alpha_A) = \frac{1}{\delta_D}[\frac{\delta_D}{\delta_A} \cdot \frac{\delta_D - d}{\delta_D} \cdot \frac{\delta_D - d}{2} + \frac{\delta_D}{\delta_A} \cdot \frac{d}{\delta_D} \cdot 0] \quad (9)$$

To derive the benefit, the cost of moving is substracted from the average gain.

$$\beta_A(\alpha_D, \alpha_A) = \frac{(\delta_D - d)^2}{2 \cdot \delta_D\delta_A} - k_A\alpha_A \quad (10)$$

$$\beta_A(\alpha_D, \alpha_A) = \frac{\delta_D}{2 \cdot \delta_A} - k_A\alpha_A + \frac{d}{\delta_A} + \frac{d^2}{2 \cdot \delta_A\delta_D} \quad (11)$$

The benefit of the defender is expressed as follows:

$$\beta_D(\alpha_D, \alpha_A) = 1 - \frac{(\delta_D - d)^2}{2 \cdot \delta_D\delta_A} - k_D\alpha_D \quad (12)$$

We can easily see that when *d*=0, we obtain the formula of the original FlipIt game.

Case 2: $\delta_A \leq \delta_D$ *(The attacker plays at least as fast as the defender.)*

First let $r = \frac{\delta_D}{\delta_A}$. The intervals between two consecutive attacker's moves have length δ_A. Consider a given attackers move interval. The probability over the attacker's phase selection that the defender moves in this interval is $\frac{\delta_D}{\delta_A} = (1/r)$. Given that the defender moves within the interval, he moves exactly once within the interval (since$\delta_A \leq \delta_D$) and his move is distributed uniformly at random.

A similar analysis as in case 1 for a FlipIt game without virus propagation yields the following benefits:

$$\beta_D(\alpha_D, \alpha_A) = \frac{1}{2r} - k_D\alpha_D = \frac{\delta_A}{2\delta_D} - k_D\alpha_D \quad (13)$$

$$\beta_A(\alpha_D, \alpha_A) = 1 - \frac{1}{2r} - k_A\alpha_A = 1 - \frac{\delta_A}{2\delta_D} - k_A\alpha_A \quad (14)$$

For the case with a virus we consider two cases, Case a and Case b, to compute the benefit of both players.

Case a: $d + \delta_A \leq \delta_D$*:*
If $d + \delta_A$ is a timespan for the attacker, the defender will never move twice during this timespan because $d + \delta_A \leq \delta_D$. With a probability of $\frac{\delta_A}{\delta_D}$ the defender move will during the interval of the attacker. When this happens the defender will end with being in control at the end of the interval. In the next interval the attacker will move in the beginning but will not gain control because of the delay, see figure 2 cases (1) and (2). This means that the defender will keep the control over the resource in the next interval over a period of the delay, namely *d*. Because $d + \delta_A \leq \delta_D$ the next move of the defender in this second interval will not be during the delay.

Every time the defender plays, he will get a gain rate of $\frac{\delta_A}{2\delta_A}$ and the next interval will always receive a gain rate of $\frac{d}{\delta_A}$. To derive the total gain formula for the defender we can calculate the probability that the defender will move during an interval of the attacker and multiply it by the gain rate of the defender plus the gain rate that it will give to the next interval.

$$\gamma_D(\alpha_D, \alpha_A) = \frac{\delta_A}{\delta_D} \cdot \frac{(d + \frac{\delta_A}{2})}{\delta_A} \quad (15)$$

$$\gamma_D(\alpha_D, \alpha_A) = \frac{\delta_A}{2\delta_D} + \frac{d}{\delta_D} \quad (16)$$

This yields in the following benefit formula:

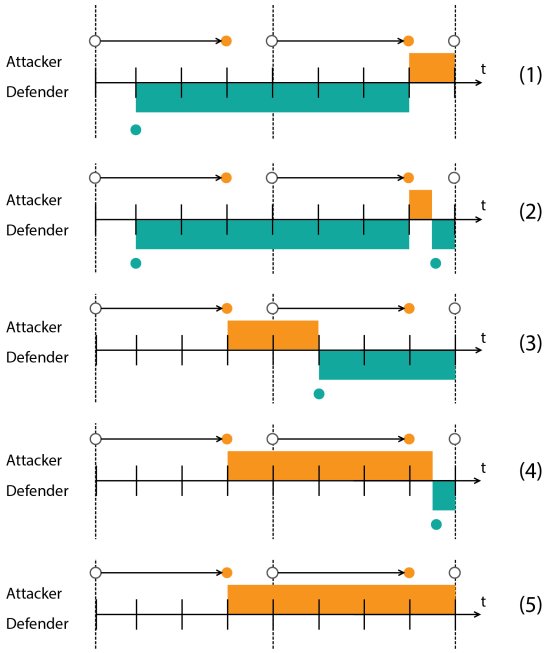
$$\beta_D(\alpha_D, \alpha_A) = \frac{\delta_A}{2\delta_D} + \frac{d}{\delta_D} - k_D\alpha_D \quad (17)$$

The benefit for the attacker will be as follows:

$$\beta_A(\alpha_D, \alpha_A) = 1 - \frac{\delta_A}{2\delta_D} - \frac{d}{\delta_D} - k_A\alpha_A \quad (18)$$

It is crucial that the $d + \delta_1$ is not greater than δ_0. If it is greater that means that the defender can move during a delay after an interval of the attacker where the defender already has moved. If we calculate the benefit formula in the same manner, to much gain is added. Namely the overlap during the delay.

Figuur 2: Case 2 where d + deltaA < deltaD



Case b: $d + \delta_A > \delta_D$*:*

V. Simulations and results

VI. Related work on FlipIt

There a various possible ways to extend FlipIt. Laszka et al. made a lot of additions and extensions on the original game of FlipIt. For instance Laszka et al. extended the basic FlipIt game to multiple resources. The incentive is that for compromising a system in a real case it needs more than just taking over just one resource. An example is that one resource can be gaining access to a system and breaking the password of the system is another resource. The model is called FlipThem [3] . They use two ways to flip the multiple resources: the AND and the OR control model. In the AND model the attacker only controls the system if he controls all the resources of the system, whereas in the OR model the attacker only needs to compromise one resource to be in control of the entire system.

Another addition of Laszka et al. to the game of FlipIt [4] is extending the game to also consider non-targeted attacks by non-strategic players. In this game the defender tries to maintain control over the resource that is subjected to both targeted and non-targeted attacks. Non-targeted attacks can include phishing, while targeted attacks may include threats delivered through zero day attack vulnerabilities.
One of the last important addition from Laszka et al. [5] is to consider a game where the moves made by the attacker are still covert but the moves made by the defender are known to the attacker. This means that the attacker can base his attacks on the defender's moves. Both the targeted and non-targeted attacks don't succeed immediately. For the targeted attack the time till it succeeds is given by an exponential distributed random variable with a known rate. The non-targeted attacks are modelled as a single attacker and the time till it succeeds is given by a Poisson process. The conclusion of this paper is that the optimal strategy for the defender is moving periodically.

Other authors used the FlipIt game to apply it on a specific scenario. To be able to use the FlipIt game, modifications where required for the FlipIt model. One of the scenarios by Pham [6] was to find out whether a resource was compromised or not by the attacker. This could be verified by the defender, who has an extra move "test"beside the flip move. The basic idea is to test with an extra action if the resource has been compromised or not. This move involves also an extra cost. This model is useful if somebody wants to know for example if his or her password has been compromised.

Finally researchers also have investigated the behavioural of humans when playing FlipIt. A Nochenson and Grossklags [7] investigate how people really act when given temporal decisions. Reitter et al. [8] observed continuous games, 20-seconds FlipIt game..

Referenties

[1] "It security risk survey 2014." [Online]. Available: http://media.kaspersky.com/en/IT_Security_Risks_Survey_2014_Global_report.pdf/

[2] M. van Dijk, A. Juels, A. Oprea, and R. Rivest, "Flipit: The game of "stealthy takeover"," *Journal of Cryptology*, vol. 26, no. 4, pp. 655–713, 2013. [Online]. Available: http://dx.doi.org/10.1007/s00145-012-9134-5

[3] A. Laszka, "Flipthem: Modeling targeted attacks with flipit for multiple resources," *5th International Conference, GameSec 2014, Los Angeles, CA, USA, November 6-7, 2014. Proceedings*, vol. 8840, pp. 175–194, 2014.

[4] A. Laszka, B. Johnson, and J. Grossklags, "Mitigating covert compromises," *iets*, vol. 8289, pp. 319–332, 2013. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-45046-4-26

[5] ——, "Mitigation of targeted and non-targeted covert attacks as a timing game," vol. 8252, pp. 175–191, 2013. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-02786-9-11

[6] V. Pham and C. Cid, "Are we compromised? modelling security assessment games," in *Decision and Game Theory for Security*, ser. Lecture Notes in Computer Science, J. Grossklags and J. Walrand, Eds. Springer Berlin Heidelberg, 2012, vol. 7638, pp. 234–247. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-34266-0_14

[7] A. Nochenson, J. Grossklags *et al.*, "A behavioral investigation of the flipit game," in *Proceedings of the 12th Workshop on the Economics of Information Security (WEIS)*, 2013.

[8] D. Reitter, J. Grossklags, and A. Nochenson, "Risk-seeking in a continuous game of timing," in *Proceedings of the 13th International Conference on Cognitive Modeling (ICCM)*, 2013, pp. 397–403.