When will deception work in `FlipIt` security games?

Do June Min

January 19, 2018

Abstract

In cybersecurity settings, it is often useful to model the interaction between system administrators and malcious operators using game theory, while abstracting away specific threat models and countermeasures. By doing an equilibrium analysis of such games, it is possible to reason about the rational behavior of the competing agents. Our model assumes that the defender receives partial information about the attacker's action. That is, the defender is notified when the attack tries to compromise the resource. Moreover, the attacker is given the option to perform a "deceptive" attack, which is cheaper than compromising attacks but achieves nothing other than notifying the defender. The goal of our study is to see if this modified setting will result in the attacker being able to achieve higher utility, by forcing the defender to selectively respond to the flood of deceptive and real attcks, allowing some real attacks to be undefended. Our analysis of the base game model shows that deceptive strategies not work for the attacker. However, we show that this result does not necessarily hold in some extensions of the game, where there are multiple resources or the defender is contrained in its budget.

1 Introduction

In cybersecurity situations, modelling an adversary might be useful, especially if there is reason to believe that the main threat is from a malicious agent with specific objectives. For instance, a system administrator for a military organization is justified in assuming the existence of an enemy operative whose goal is to gain access to a server with confidential information. If there is a credible threat from adversaries, game theory is a natural choice for analyzing the expected behaviors of the adversaries and the defender[1]. Then, this analysis can be used to inform the defender in making optimal choices, given information about recent attacks and threats from adversaries.

A well-studied application of game theory in cybersecurity is the Stackelberg competition, where the leader commits to a choice, and the follower makes the move after observing the leader's choice. For instance, the defender may consider how to allocate defensive resources, under the assumption that the attacker will figure out this allocation strategy completely. Thus, the defender can form an optimal allocation strategy, expecting that the attacker will also play optimally against this strategy [2]. In the cybersecutiy domain, system administrators can use this model to generate system scanning schedules, for example.

Game theory can also be used to model the prolonged interaction between the attacker and the defender. Recent attacks on organizatiosn including Google, Lockheed Martin, and RSA have shown that *Advanced Persistent Threats*(APTs) have emerged as a formidable means of compromising computer resources[3] . The attackers not only employ zero-day vulnerabilities to attack and control protected resources, but also monitor the defender's response and produce a counter against protective measures[4]. To model such situations, it is necessary to look at stochastic games, where the payoff is dependent of the state of the game[5][6][7].

In this study, we model the interaction between the attacker and the defender using a sequential stochastic game with imperfect information. The game begins with the defender in control of the resource, and the attacker making the first move. Beside the resource-compromising “attack” move, the attacker can play a “deceive” move, which does not compromise the system, but is indistinguishable from a normal attack for the defender. While the attacker does not get any information about the defender’s action, the defender is notified when the attacker makes either a real or fake attack. This creates the possiblity that the attacker might exploit deceive moves to achieve higher payoff by forcing the defender to selectively respond to the flood of deceptive and real attcks, allowing some real attacks to be undefended.

The goal of the paper is to (1) characterize reasonable strategies for the attacker and the defender in this modified game, (2) determine if deceive will be played by the attacker in equilibrium, (3) if so, identify what kind of strategy is deceive played in and how much benefit it provides and (4) finally determine if the ability to deceive allows the attacker to achieve higher payoff in some equilibria. We report that the basic model we propose has a unique equilibrium in which deceive is not played. After analyzing the model, we propose possible extensions of the original model in which deceive is part of equilibrium strategy, and study which properties of the modified model causes deceive to be played. We find that for the base game with one resource, deceptive strategies not work for the attacker. However, we also show that this result does not necessarily hold in some extensions of the game, where there are multiple resources or the defender is contrained in its budget.

2 Related Work

FlipIt is a stochastic game in which two players compete to control a resource[8]. Each player, the attacker and the defender, can perform a move called ‘flip” that results in the agent controlling of the resource. The game is unique in that model cyber security situations in that it assumes stealthy compromise. Stealthy compromise means that neither the defender nor the attacker knows with certainty who controls the resource at any time(except the beginning), and that any move by either player is unobserved by the other. This condition of no observability makes the model both interesting and intractable. Each player has to plan his action when there is absolutely no information about the state of the game, except for the movement made by the agent. In this study, we relaxed this assumption and allowed the defender to observe when the attacker makes a move.

Also, FlipIt is usually studied as an infinite game with a discount factor. A discount factor is usually set to a value between 0 and 1, to reflect the assumption that an agent prefers an immediate reward over a faraway one [9]. FlipIt and our model belong to a class of model called *Partially Observable Stochastic Game*, in which each agent’s action is unseen by each other, and the state of the game(the control of the resource) dynamically changes as a result of each player’s move, which in turn, determines the payoff function. As of now, there is no known algorithm or method to compute equilibria of infinite POSGs, and analysis of the game relies on limiting strategies. into a few classes. There are also approximation algorithms to solve POSGs [10] [11].

An extension of FlipIt which is of particular interest to us is FlipThem[12], which generalizes to multiple resources. Two control models, AND and OR, are studied. In the AND model, the attacker has to compromise all the resources to derive utility. In the OR model, compromising only one resource is enough. In FlipThem, each player cannot treat a resource independently from others, meaning different strategies will arise as a result. Although the game does not allow deceive move, we look at the concept of non-independence between resources as a possible condition for deceptive strategies in our research.

3 Game Model

In this section, we define our game model formally. The formal definition is not necessary for the first result, but will be used for the extensions of the base model. This section also covers and justify some of the modelling assumptions we made. Finally, we explain our choice of solution concept, which is used in the equilibrium analysis.

3.1 Formal Definition

This section gives a formal definition of the modified `FlipIt` game, and introduces necessary notations.

Players There are two players, the attacker(A) and the defender(D). It is important to note that unlike in the original `FlipIt` game, the game is not symmetric between the two players, as will be explained.

Time The game begins at time $t = 0$ and continues indefinitely as $t \to \infty$. The model assumes discrete time steps. Thus, it can be thought of as an infinitely repeated sequential game, where the payoff for each player varies dependent on the control state of the game. Since we treat time as discrete, the word "round" is used interchangeably with time.

Resources The players compete to control some valuable resources, collectively denoted as a set R. Each resource $r \in R$ has costs and payoffs associated with it.

Game State The time-dependent variable $K_r(t)$ denotes the current player controlling the resource r at time t; $K_r(t)$ is either 0 or 1 at any time t. The Attacker controls the resource if $K_r(t) = 0$, and the Defender controls if $K_r(t) = 1$. The game begins with the Defender in control: $K_r(0) = 1$ for all $r \in R$.

Costs There are three types of costs, the Attack cost, the Deceive cost and the Reboot cost. For both players, we assume that NoOp is of cost 0 for any resource r. We define that at time step t, $c_{i_r}(t)$ denotes the cost of the action made by each player i. Player i's total cost C_i in a given game is just the sum of c_i over t:

$$C_i(t) = \sum_{j=1}^{t} c_i(j)$$

where $c_i(j)$ denotes the cost of the move made by player i at round t, where $c_i(j) = \sum_{r=1}^{R} c_{i_r}(j)$.

Gains We define that at time step t, $g_{D_r}(t) = 1$ if player D controls the resource r, and $g_{A_r}(t) = 0$ if not. Then, player i's total gain G_i in a given game is just the sum of g_i over t:

$$G_i(t) = \sum_{j=1}^{t} g_i(j)$$

where $g_i(j) = \sum_{r=1}^{R} g_{i_r}(j)$.

Benefits Players receive utility equal to the accumulated control payoff they received over time, subtracted by the cost of the moves they made. Thus, for each player i, the *net benefit* $B_i(t)$ is defined as follows:

$$B_i(t) = G_i(t) - C_i(t)$$

The average benefit of player i is defined

$$\beta_i(t) = B_i(t)/t = G_i(t)/t - C_i(t)/t$$

One Round of the Game At each round t,

(1) The Attacker can play Attack, Deceive, and NoOp.
(2) Attack will compromise the resource, while Deceive does not.
(3) If the Attacker played Attack or Deceive, the Defender receives a signal that there has been a Probe or not.
(4) After receiving the signal, the Defender either Reboots or NoOp.
(5) Payoff computed according to the control status and move cost is accumulated. This is specified in the next section.

Table 1: Player Actions and Resulting Control State for Each Resource

Attacker Move	**Defender Move**	**Control State**
Attack	Reset	Defender
Deceive		
NoOp		
Attack	NoOp	Attacker
Deceive		Previous State
NoOp		

3.2 Assumptions

Before we present our first result, it is important to note that our model abstracts real cybersecurity situations by making the following assumptions.

First, we assume the game to be infinitely repeated. At first glance, it may seem more natural to model cybersecurity situations as a finite game, since all such situations happen on a finite horizon in real life. However, analyzing a finite game with standard equilibrium solution concepts would result in an implicit assumption that each player is aware of precisely when the interaction will terminate. This is a highly unrealistic assumption, and will result in backward induced equilibrium that does not generalize well. On the other hand, infinitely repeated games do not suffer from this problem, and can model situations where each player is uncertain of the end of the game. Also, our model uses discrete time frame, mainly an abstraction to simplify analysis.

Another important assumption made is maximum patience by each player. Maximum patience means that the discount factor is 1. The discount factor is multiplied when calculating expected future reward of each player. In infinitely repeated games, having non-1 discount factor would generally result in different equilibrium behavior. Thus, after deriving results about the equilibria of the game, we consider the effect of having different values of discount factor.

Moreover, we suppose complete information. That is, each player is aware of what strategy the other is using. However, this is different from perfect information, which means that every action is fully observed by all players. Our model is an imperfect information game, since each player gets only part of the full information. If the model is to be studied further, the model may be extended to a Bayesian game with probability distribution over possible player types. However, we limit our focus to complete information set-up in this paper.

3.3 Solution Concept

To study the equilibria of the game, we use *Nash equilibrium* as solution concept.

Definition. *Nash Equilbrium Let $S = S_1 \times S_2 \times ... \times S_n$ be the set strategy profiles for players $1, 2, ..., n$ and $f(x) = (f_1(x), f_2(x), ..., f_n(x))$ be the payoff function where each $f_i(x)$ denotes the utility of player i where $x \in S$. Then $x = x_1 \times x_2 .... \times x_n \in S$ is a Nash equilibrium if it satifies the following condition.*

$$\forall i, f_i(x_i^*, x_{-i}^*) \geq f_i(x_i, x_{-i}^*)$$

Intuitively, this means that no player i can achieve higher payoff by unilaterly deviating from x. In the rest of the paper, equilibrium means Nash equilibrium unless stated otherwise. Since mixed strategy includes probablistic choice over possible actions, the payoff function computes the *expected payoff* for each player.

Other relevant solution concepts are *Markov perfect equilibrium*, *perfect Bayesian equilbrium*, and *sequential equilibrium.* Markov perfect equilibrium is a refinement of subgame perfect equilbrium, and considers strategies which are conditioned only on the current state of the game. However, in our model, each player does not fully observe each other's actions, and therefore cannot determine the current state. Perfect Bayesian equilibrium, also a refinement of subgame perfect equilibrium seems more related to other model for it does not assume full knowledge of the game. However, our model assumes imperfect information, not incomplete information.

Finally, sequential equilibrium specifies a belief for each player, which represents the probability distribution over nodes for each information set. This distribution corresponds to the belief held by the player, computed by the application of Bayes rule. When the strategy profile S maximizes the expected reward and the belief assessment b is consistent with the observations by each player, we say that s, b is in sequential equilibrium. Sequential equilibrium has the advantage that it guarantees each players plays a best move at any given information set. However, we did not analyze the game using sequential equilibrium, since the infinitely repeated natured of the game makes the computation of the belief assessment intractable.

4 Will deceive be played by the Attacker in an equilibrium?

4.1 1-Resource Model

First, we look at an instance of the game where $|R| = 1$. That is, there is only one resource A and D are competing for. In this case, we show that the attacker will not play deceive in equilibrium by showing that in the only equilibrium of the game, the attacker does not play deceive.

Definition. *NoOp strategy is an attacker and defender strategy where every information set is mapped to the action no-op.*

Definition. *AutoReboot strategy is a defender strategy which maps every attacker probe to the action reboot and attacker no-op to no-op.*

Lemma 1. *NoOp vs AutoReboot is an equilibrium of the game.*

Proof. We show this is a Nash equilibrium of the game by showing that there is no unilateral deviation by either one of the player that results in higher payoff for the deviator.

For the attacker, any deviation means playing some number of probes. Since all the probes will be rebooted right after, there is no control benefit gain. Thus, this results in negative average gain due to the costs of probe, whereas in the original state the attacker receives zero average gain.

For the defender, the original gain of 1 is the maximum attainable payoff. Thus, there is no beneficial deviation. ■

Lemma 2. *In equilibrium, the only way for the attacker to have positive control benefit is to at least attack twice.*

Proof. For the defender best response to allow some attacker control of the resource, it must be that rebooting after each attack is too costly. If the attacker plays attack only once, the defender can reboot immediately and compensate for any arbitrary cost of reboot by ensuring infinite control of the resource.

Note that this is a very weak condition for the attacker to have positive control benefit. ■

Theorem 3. *There are no other equilibria of the game other than the one shown in Lemma 1.*

Proof. First, consider any equilibrium in which the attacker has positive control benefit. By Lemma 2, we know that the attacker attacks at least twice. Because the defender cannot distinguish attack from deceive, the defender's strategy must have the same response to deceive. Therefore, the second attack can be replaced by deceive for higher benefit, since deceive has a cheaper cost. This is a contradiction, since the original strategy is not a best response to the defender strategy.

Next, consider the case in which the attacker has zero control benefit. In this case, the attacker's best response is to minimize move cost, hence the best response is NoOp strategy. Against this attacker strategy, the defender best responds by playing AutoReboot, as shown in Lemma 1.

Therefore, NoOp vs AutoReboot is the only equilibrium of the game ■

Corollary 4. *The attacker never plays deceive in equilibrium.*

Proof. This follows from Lemma 1, Theorem 3, and the fact that deceive is never played by the attacker in the equilibrium. ■

Lemma 5. *When $0 < \delta < 1$, Corollary 4 still holds.*

Proof. Lemma 1 still holds, since the ordering of the expected returns of each action is preserved. Also, in Lemma 2, unless $\delta = 0$, the expected return of infinite control outweighs the cost of reboot. Thus, Lemma 2 also holds. Therefore, Theorem 3 also holds. ■

4.2 Discussion

There are opposing intuitions that motivated our research. First, intuitively, deception does not provide any cost for the attacker while it incurs some cost. Thus, a rational attacker would play deceive only when it guarantees future gain. On the other hand, deception might be reasonable under certain circumstances; if mixing deceiving moves with real attacks may force the opponent to selectively respond, allowing some attacks as a result.

In the end, our analysis shows that playing deceive is not a rational choice for the attacker in this game. Looking back, this is not surprising because deception offers no immediate reward while incurring cost for the attacker. For deception to be useful for the attacker, it must move the defender to an information set in which the defender expects to achieve higher reward by not rebooting right after a probe. Our proof of Theorem 3 effectively shows that there can be no such information set in any play of the game. This conclusion is sweeping because there are no restraints on the costs of the moves other than that NoOp is cheaper than Deceive, which is cheaper than Attack.

What is more surprising, which we will show next, is that having the option do Deceive seems to result in a worse outcome for the Attacker.

4.3 Comparison with No-Deception Game

Since the Attacker cannot achieve any positive utility in this game, it is natural to ask if the option to Deceive is helpful at all. It is easy to think that having more options cannot hurt. Contrary to this intuition, we observe that the option to deceive causes a lower payoff for the attacker. From our proof of Theorem 3, we saw that the option to Deceive created a unilateral deviation that is beneficial for the attacker, effectively checking off any possible equilibrium where the Attacker might achieve a positive payoff. Indeed, it is possible to construct an equilibrium where the attacker achieves positive payoff, if deceive is removed from the possible action set.

Consider the following game in which the Attacker can only NoOP or Attack, and otherwise identical to the model we have just analyze. Assume Attack costs between 0 and 1, Reset between 1 and 2 and NoOp 0 for both players. Also, assume that the control benefit is 1 for both. Then, let the Attacker play a strategy where it plays Attack at any timestep t. Also, let the Defender play NoOp after an Attack from A, otherwise it plays Reset. This strategy profile results in a payoff profile of a, d such that $0 < a < 1$ and $d = 0$. It is easy to check that this is a Nash equilibrium: if the Attacker deviates by playing one or more NoOp, its final utility strictly decreases because the payoff at that round will decrease to 0. Also, if the Defender plays Reset after an Attack, the cost of Reset will result in a lower average payoff.

Thus, we have a seemingly paradoxical result that Deceive actually hurts the Attacker. However, it is important to remind that several assumptions made in the model may not hold in real situations. For instance, system administrators might not be even aware that the Attacker can issue deceptive attacks. In this case, the assumption of complete information is not valid- it is highly likely that the Attacker will be able to use Deceive moves to its advantage.

4.4 Extensions of the model

In the previous seciton, we have seen that the 1-resource version of the game has one trivial equilibrium, independent of move costs and discounting factor. In this section, we explore the extent to which this result holds true, by modifying some aspects of the game.

4.4.1 Multiple resource game with non-linear costs

One natural extension of the game is to model multiple resourcesm i.e. $|R| > 1$. We first show that if each resource is independent of any other resource, there is only one equilibrium where the attacker and the defender each play NoOp and AllReboot on all resources.

Theorem 6. *In a multiple-resource version of the game where each resource is independent, the only equilibrium is NoOp vs AllReboot.*

Proof. Since resources are independent from each other, a game with $|R|$ resources can be analyzed as $|R|$ instances of 1-resource model. Then, by Theorem 3, NoOp vs AllReboot is played on each of the resources. ■

Thus, if the resources are treated as independent, i.e. the state of one resource does not affect others, then a result analogous to that of the base model still holds. However, this independence assumption frequently does not hold in cybersecurity situations. For system administrators, rebooting several resources is often costlier than rebooting just one system, as multiple systems become unusable and user access decreases as a result. Also, it is also possible that the attacker is constrained in the number of attacks he/she can simultaneously execute. Thus, we consider the case when resources are not independent.

4.4.2 An example equilibrium in a 2-resource game

We assume there are 2 resources the players are competing for. The cost of all possible actions at each time step is given in the following tables.

Table 2: Cost Table

Table 3: Defender Cost

Move	Cost
R, N	
R, R	
N, N	

Table 4: Attacker Cost

Move	Cost
A, N	
D, N	
A, D	
N, D	
A, A	
D, D	
N, N	

Now, let s_A be an attacker strategy that picks one resource with probability 0.5, and attacks that resource and plays deceive on the other. Also, let s_D be a defender strategy that at $t = 0$, picks one resource, reboots it and noops on the other, and $\forall t > 0$, reboots the resource nooped on $t - 1$. Now, let $s = (s_D, s_A)$ and b be the belief induced by the strategy. Now, WTS that (s, b) is a sequential equilibrium.

4.4.3 Budget-limited defender game

In real life, a system administrator or security personnel is often confined by practical concerns over resource usage over a fixed period of time. For example, in the case of a defender in charge of one resource, he/she might want to limit the number of reboots to 3 times in 5 time units, in order to guarantee minimum access and usability. In Section 3, we assumed that there was no such constraint.

Definition. *Define budget-limit condition.*

5 Conclusion

References

[1] Sajjan Shiva, Sankardas Roy, and Dipankar Dasgupta. Game theory for cyber security. In *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research*, CSIIRW '10, pages 34:1–34:4, New York, NY, USA, 2010. ACM.

[2] Milind Tambe. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned.* Cambridge University Press, New York, NY, USA, 1st edition, 2011.

[3] Eric Cole. *Advanced Persistent Threat: Understanding the Danger and How to Protect Your Organization.* Syngress Publishing, 1st edition, 2013.

[4] Wil Allsopp. *Advanced Persistent Threat Modeling: Defending Against APTs.* O'Reilly Media, Inc., 1st edition, 2016.

[5] P. Hu, H. Li, H. Fu, D. Cansever, and P. Mohapatra. Dynamic defense strategy against advanced persistent threat with insiders. In *2015 IEEE Conference on Computer Communications (INFOCOM)*, pages 747–755, April 2015.

[6] Stefan Rass, Sandra König, and Stefan Schauer. Defending against advanced persistent threats using game-theory. *PLOS ONE*, 12(1):1–43, 01 2017.

[7] Stefan Rass and Quanyan Zhu. *GADAPT: A Sequential Game-Theoretic Framework for Designing Defense-in-Depth Strategies Against Advanced Persistent Threats*, pages 314–326. Springer International Publishing, Cham, 2016.

[8] Marten Dijk, Ari Juels, Alina Oprea, and Ronald L. Rivest. Flipit: The game of "stealthy takeover". *J. Cryptol.*, 26(4):655–713, October 2013.

[9] George Mailath and Larry Samuelson. *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, 2006.

[10] Eric A. Hansen, Daniel S. Bernstein, and Shlomo Zilberstein. Dynamic programming for partially observable stochastic games. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence*, pages 709–715, San Jose, California, 2004.

[11] Akshat Kumar and Shlomo Zilberstein. Dynamic programming approximations for partially observable stochastic games. In *Proceedings of the Twenty-Second International FLAIRS Conference*, pages 547–552, Sanibel Island, Florida, 2009.

[12] Aron Laszka, Gabor Horvath, Mark Felegyhazi, and Levente Buttyán. *FlipThem: Modeling Targeted Attacks with FlipIt for Multiple Resources*, pages 175–194. Springer International Publishing, Cham, 2014.