Game Theoretic Analysis of Ransomware: A Preliminary Study

Rudra Prasad Baksi and Shambhu Upadhyaya
Department of Computer Science & Engineering, University at Buffalo, SUNY, Buffalo, NY, USA
{*rudrapra, shambhu*}*@buffalo.edu*

Keywords: Cryptography, Computer Security, Cyber-Security, Game Theory, Ransomware.

Abstract: Ransomware attacks have been frequent and wreaking havoc of the kind never seen before. This paper presents an analysis of a basic type of ransomware. When faced with a ransomware attack, the victim needs to address a question whether to pay or not to pay the ransom. In this regard, we develop a game-theoretic model to analyze the attack landscape and to determine under what conditions the defender is in a position of advantage to successfully neutralize the attack. In this preliminary analysis, we develop strategies which would help the victim to make an informed decision. We put forward two parameters that help the defender make an informed decision in the face of an attack. We perform a sensitivity analysis to show how the variation of the parameters affect the outcomes of the attacker and the defender and thereby affecting the equilibrium strategies. We then discuss how the outcome of the model can help defenders to come up with an effective defense mechanism against similar future attacks.

1 INTRODUCTION

Ransomware are a type of malware which encrypt critical data of a system and hold them for a ransom. If the ransom is paid, then the data is released, else it is made inaccessible to the victim. There are primarily three types of ransomware, *the locker*, *the crypto*, and *the hybrid* (Zakaria et al., 2017). The locker locks the entire system and denies the user any access to it. On the other hand, crypto encrypts only critical data found on the system. It targets specific files and/or folders. The hybrid variant of the malware possesses capabilities of the other two types of ransomware. More recent variants of ransomware come with additional features like campaign abort strategy or a contingency plan of attack upon being discovered prior to the launch of attack. They qualify as advanced persistent threats (APT) (Baksi and Upadhyaya, 2018). In this paper, we limit ourselves to non-APT type ransomware so as to obtain some preliminary results which can later be extended to cover more sophisticated malware and/or APT type ransomware.

Ransomware attacks have been wreaking havoc in the industry and/or government organizations. It is a nuisance which not only hampers daily working of the government and/or industry but also makes it difficult for common people to carry out their normal activities. Ransomware attacks may cause healthcare facilities, schools, public transport organizations, police stations, gas stations, IoT infrastructure of many organizations, and many more institutions to suspend their daily services (Milosevic et al., 2016). In April 2017, Erie County Medical Center (ECMC) came under ransomware attack by a ransomware named SamSam. ECMC refused to pay the ransom of $30,000 in crypto currencies and ended up spending around $10 Million in system restoration. But they were covered by cyber insurance (Davis, 2017) (Goud, 2017). In March 2018, the city of Atlanta came under the SamSam attack. The officials refused to pay the ransom of $51,000 in crypto currencies and ended up spending somewhere between $2.6 Million to $17 Million in restoration of the systems with years of police data lost (Deere, 2018) (Newman, 2018). Colonial Pipeline came under attack by a ransomware called DarkSide and they paid a ransom of $4.4 Million (75 bitcoins) (BBC, 2021). This ransomware attack caused fuel shortages in many areas including cities and airports. It caused loss of business and disruption of public life and fuel prices went up. These are some of the examples of ransomware attacks the country has faced recently. Many attacks go unreported for some reason or the other. But the threat is real and so is the damage suffered from the attacks. The affected agencies and institutions would like to make a quick decision in order to restore services so that public life can return to normalcy. In doing so, they need to make not only a quick decision but it should also

be strategic. Our paper dives into the area of this decision making. The research presented here would help a victim to make an informed decision when faced with such an attack.

We use game theory to analyze a form of ransomware which we call a basic ransomware in the rest of this paper. We design a two player sequential game. One of the players is the attacker whereas the other is the victim and/or defender. Through the analysis we determine optimal strategies for both the attacker and the defender. Then we present equilibrium solutions for different conditions. We also perform a sensitivity analysis to examine how the decisions of both players are being affected when the values of the decision parameters change. The major contribution of this paper is the presenting of a parameter to quantify the importance of the value of resources under attack and the introduction of a new parameter to better understand the reputation of the attacker. We, thus make the importance of the resources under siege and the reputation of the attacker as quantifiable metrics and a part of the decision making process. The paper is organized as follows. In Section 2, we discuss some background information and related work in this area. In Section 3, we present our game theory based analysis of the basic ransomware. We then analyze the results in Section 4 and investigate the sensitivity of the values of the parameters while making an informed decision and put forward a prescriptive solution of preparedness and mitigation of the ransomware attack. Finally, we conclude the paper in Section 5 and explore the possibilities of future research in tackling more sophisticated ransomware attacks.

2 Background and Related Work

Zakaria et al. (Zakaria et al., 2017) investigated the rise in spread of ransomware, and laid down the main areas for research on ransomware starting with the detection using indicators of compromise (IoC), signatures of the malware and analysis of network traffic. Then depending upon the type of attack mounted by the malware, the ransomware is classified into one of the three categories, locker, crypto, and hybrid. This is followed by two other areas of research identified by the authors, which are recovery from the attack and prevention from future attack. In our paper, we investigate the strategies of the attacker as well as the defender, and examine "recovery from attack" and "prevention of future attacks" mentioned in (Zakaria et al., 2017).

Baksi and Upadhyaya (Baksi and Upadhyaya, 2017) used a hardware-based defense architecture by leveraging the capabilities of the trusted platform module (TPM) as a defense against APTs. Cekar et al. (Çeker et al., 2016) used deception to counter denial of service (DoS) attacks. They also used a game-theoretic model based on the signaling game with perfect Bayesian equilibrium (PBE) to investigate the implications of deception to counter the attacks. Deception as a potential defense tool has been used to lure attackers to high interaction honeypots in (Pauna, 2012) and thereby designing an effective malware detection system. The author proposed an adaptive honeypot system based on game-theoretic concepts to entice the attackers, leading to the detection of the rootkit malware by the defender. Baksi and Upadhyaya (Baksi and Upadhyaya, 2020) used a Hidden Markov Model (HMM) based approach to identify intrusion by a ransomware. The HMM based IDS designed by them was aimed at tackling APT type ransomware but works well against basic ransomware as well. Yet, the question is, once a defender has detected an intrusion by a ransomware, what should be the next *plan-of-action*. The research reported in Baksi and Upadhyaya (Baksi and Upadhyaya, 2017), Cekar et al. (Çeker et al., 2016), and Baksi and Upadhyaya (Baksi and Upadhyaya, 2020) are aimed at an APT type malware threat and have deception as a defense strategy. In our paper, we use basic ransomware as the threat model, and put forward a scheme which would help a defender make an informed decision when faced with an attack, especially when the defenders do not have a deception based defense system against such attacks.

Game theory opens up new avenues for malware analysis. It has been used to analyze the strategies of the malware and the victims and study the attack landscape. Khouzani et al. (Khouzani et al., 2011) used a zero-sum dynamic game-theoretic model as a solution to malware attack. They analyzed the structural properties of saddle-point strategies, which are simple threshold-based policies, and showed the possibility of a robust dynamic defense system against malware attacks. They have investigated the network defense landscape of mobile wireless networks. The strategies investigated on the part of the defender were reception and patching rates. The strategy of the attacker that was investigated was the annihilation rate of infected nodes. Through the formulation of a dynamic game it was proved that threshold-based policies form an effective robust solution to malware attacks. Spyridopoulos et al. (Spyridopoulos et al., 2013) investigated a game-theoretic approach for the cost-benefit analysis of malware proliferation, and modeled it on the lines of epidemic spread models, namely, SIR and SIS models. They applied their models on the Code-

Red worm. The idea was to develop a cost-benefit game-theoretic model to apply malware proliferation strategies including "patching" of infected nodes in a network, "removal" of infected nodes in the network, and/or the combination of both. They used "FLIPIT" game as the basis for the development of their model. In our paper, we use game theory to analyze a basic form of ransomware attack through a sequential game. We put forward two new parameters to help the defender make an appropriate decision when faced with an attack.

Cartwright et al. (Cartwright et al., 2019) came up with a game-theoretic model to analyze generic ransomware attacks. They used the kidnapping game as the basis for the model (Gintis, 2009) (Selten, 1977) (Selten, 1988). The malware was modeled as the kidnapper whereas the database of the victim was modeled as the hostage. The goal of the paper is to help the defender to make an informed decision regarding the payment of the ransom, when attacked by a ransomware. But the limitation is that the game is applicable to ransomware attacks wherein the attackers are bound by the same *law of the land* as the defender. The authors assume that if the attackers are apprehended, then ransom payment as well as the encrypted resources could be extracted from them. In our research, we consider that the attackers could be anywhere in the world while staging the attack. In this scenario, it becomes really difficult to apprehend the attackers both in terms of legal and logistical fronts. But, this makes the ransomware more generic in nature and takes care of a vast majority of the attacks which are staged from foreign land. We use game theory to analyze the attack and help the defender in making an informed decision.

3 Basic Ransomware

3.1 The Threat

A malware is a software program which is designed with malicious intent to cause harm to the victim by the attacker. When the intent of a malware is monetary gain by *hijacking* victim's resources for a ransom, it is called a ransomware. Depending upon the nature and level of sophistication, a ransomware can be of an APT type or of a basic nature. An APT type ransomware is generally created by nation state actors. They are highly sophisticated attacks and are mounted through multiple clandestine stages (Baksi and Upadhyaya, 2018). For such attacks, even though monetary gain is generally the primary goal, they may have other concealed and/or disguised agenda. On the other hand, in a basic ransomware attack, the attacker encrypts the resources under risk and charges a ransom. If the ransom is paid, the attacker releases the encrypted resources, else the victim loses the resources forever. Such attacks generally have only one goal, i.e., to make the resources inaccessible to the victim until the ransom is paid. Both the APT type ransomware and the basic ransomware can be of three types, namely, crypto, locker and hybrid as discussed earlier. Attack and threat models for them exist in the literature (Kolodenker et al., 2017) (Zimba and Chishimba, 2019). In our paper, we restrict ourselves to the research concerning defense against basic ransomware.

Parameterized attack graphs have been proposed in the literature to model attacks that exploit vulnerabilities. The attack graphs capture attacker's preconditions, system and network vulnerabilities, attacker effects, and the impact of the attack on the network (Sheyner et al., 2002). The attacker precondition component include the attacker's capabilities and the knowledge needed to stage the attacks at an atomic level. However, attack graphs were found to be not very useful due to scalability concerns regarding both model specifications and eventual threat analysis (Chinchani et al., 2005). Even with automated tools for attack graph generation (Sheyner and Wing, 2003), such traditional approaches are not feasible in the context of ransomware where the attacker might use social engineering tactics and launch the attack in multiple stages. Game theory can effectively model this type of attacks and capture the interactions between the attacker and the defender. In order to facilitate the development of the game model, we introduce two parameters, that are specific to ransomware type attacks, as described in the next section.

3.2 The Game

We now present a game to depict the ransomware attack on a vulnerable and under-prepared system. We assume that the attacker exploited some form of vulnerability, thereby not giving the defender any time for preparedness. Once the attack has occurred, the defender is left with one of two choices. The first choice is to pay ransom and hope the decryption key is released by the attacker, while, the other option involves not paying ransom. The defender can make these choices based on certain conditions. In this section, we analyze two conditions which would help the defender make an informed decision on the payment of ransom and decryption of the encrypted resources held for ransom. In accordance with our assumptions, the *willingness* of the defender to pay ransom primar-

ily depends on two factors, the value of recovered resources under siege and the reputation of the attacker.

Figure 1 gives a pictorial representation of the game. The game begins with the attacker choosing either of the two strategies "Attack" or "No Attack." If the attacker chooses the strategy "No Attack", then the defender has nothing to do in order to respond to the attack. But if the attacker chooses to mount an "Attack", the vulnerable resources are encrypted and then the defender is left with one of the two choices, "Pay" or "No Pay" of the ransom. Once the defender has made its move, the attacker has two more strategies to choose from, "Release" or "No Release" of the decryption key. If the attacker is a rational player, it will only release the decryption key if a ransom payment is received. If it does not receive the ransom, it will not release the decryption key. But there can be situations where the attacker may choose to do otherwise. That way the attacker chooses not to be rational. The reasons for the attacker not being rational can be many but it is outside the scope of this paper. Since the attacker can be rational or irrational, the reputation of the attacker can play an important role in making an informed decision on the part of the defender when it is under attack.

In eq. (1), parameter r_{Rec} gives the ratio of recovered resources after payment of the ransom to the total value of assets of the defender. Variable R is the value of the resources under risk and/or siege. The ransom value charged by the attacker is denoted by β. The value of recovered resources, once the ransom is paid, is given by $R-\beta$. The value of total assets of the defender is denoted by $R_{TotalAssets}$. Therefore, r_{Rec} indicates the importance of the recovered resources for the defender, given the total value of assets, once the ransom is paid and assuming the encrypted resources have been released.

$$r_{Rec} = \frac{R-\beta}{R_{TotalAssets}} \tag{1}$$

In eq. (2) below, parameter r_{Rep} establishes the reputation of the attacker either as a rational or irrational player in the game. The higher the value of r_{Rep}, the more rational is the attacker and the higher is its trustworthiness. In an incident, if the attacker releases the decryption key on receiving the ransom payment, we assign 1 as the value of reputation for that incident. If the attacker chooses not to release the decryption key when the defender has not made the ransom payment, we assign 1 as the value of reputation for that incident. For other cases we assign 0 as the value of reputation for that particular incident. Then we take a mean of the reputation values of all the last known reported incidents to calculate the overall reputation of the attacker. If the attack is a first time attack, we assign a value of 0.5 to r_{Rep} for the purpose of decision making. The attackers when they act rationally, the r_{Rep} value for them for the next game goes up. If they act irrationally, then they incur penalty and the r_{Rep} value goes down which results in lower willingness to pay the ransom on the part of the defender.

Table 1: Attacker Notations

Notation	Description
x_1, x_2	Attacker's first and second strategies, respectively
$x^*, \widehat{x}$	Optimal strategy and best response, respectively
U_A, U_A^*	Expected Utility and Optimal Utility, respectively
Release	Decision to release the encryption keys (value 0 or 1)
(1-Release)	Decision to not release the encryption keys (No Release)

Table 2: Defender Notations

Notation	Description
y	Defender's Strategy
$y^*, \widehat{y}$	Optimal strategy and best response, respectively
U_D, U_D^*	Expected Utility and Optimal Utility, respectively
Pay	Decision to pay the ransom (value 0 or 1)
(1-Pay)	Decision to not pay the ransom (No Pay)

$$r_{Rep} = [Mean\ of\ all\ last\ known\ reported\ incidents] \tag{2}$$

Tables 1 and 2 list the notations used for describing the utility functions and strategies of the attacker and the defender, respectively.

With all the parameters under consideration, we look into the strategies of both the attacker and the defender. Variable x_1 represents the strategy for the attacker which can take up values "Attack" or "No Attack." We assign the value of the strategy "Attack" as 1 when the attacker decides to attack and 0 otherwise. Similarly for the "No Attack" strategy, the value is 1 when there is no attack, and 0 otherwise. The strategy variable x_2 for the attacker can take up values "Release" or "No Release." The value of "Release" is 1 when the attacker decides to release the decryption key, and 0 otherwise. Similarly, the value of "No Release" is 1 when the attacker decides against releasing the decryption key, and 0 otherwise. The decision variable for the defender is denoted by y. It takes up either of the two strategies as its value, viz. "Pay" and "No Pay." After an attack has taken place, if the defender decides to pay the ransom then the value of the strategy "Pay" is 1, and 0 otherwise. Similarly, if the defender decides against payment of the ransom, the value of "No Pay" strategy is 1, and 0 otherwise. Now, with $x_1 = Attack$, $y = Pay$, and $x_2 = Release$, we define the utility functions of the attacker and the defender. Equations (3) and (4) show the utility functions of the defender and the attacker, respectively.

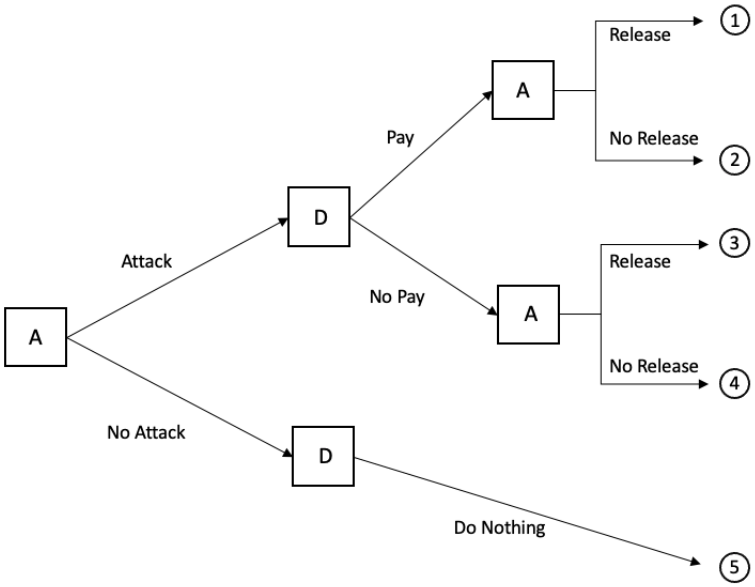


Figure 1: Basic Ransomware Attack

$$
\begin{aligned}
U_D \equiv\ & (x_1) * [(y) * \{(x_2) * (R-\beta) \\
& + (1-x_2) * (-R-\beta)\} \\
& + (1-y) * \{(x_2) * (R) + (1-x_2) * (-R)\}] \\
& + (1-x_1) * (0)
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
U_A \equiv\ & (x_1) * [(y) * \{(x_2) * (\beta) \\
& + (1-x_2) * (\beta)\} + (1-y) * \{(x_2) * (0) \\
& + (1-x_2) * (0)\}] \\
& + (1-x_1) * (\lambda)
\end{aligned} \tag{4}
$$

The above equations show all strategies and all scenarios including the ones which generated a payoff of 0 for the player. The simplified equations are:

$$
\begin{aligned}
U_D \equiv\ & (x_1) * [(y) * \{(x_2) * (R-\beta) + (1-x_2) * (-R-\beta)\} \\
& + (1-y) * \{(x_2) * (R) + (1-x_2) * (-R)\}]
\end{aligned} \tag{5}
$$

$$
\begin{aligned}
U_A \equiv\ & (x_1) * [(y) * \{(x_2) * (\beta) + (1-x_2) * (\beta)\}] \\
& + (1-x_1) * (\lambda)
\end{aligned} \tag{6}
$$

The defender makes a decision of paying the ransom based on the value of resource which it might get back on payment of the ransom, and the reputation of the malware. If the values of r_{Rec} and r_{Rep} are "significantly high" then the defender should pay the ransom. If the values of r_{Rec} and r_{Rep} are "significantly low", then the defender should decide not to pay the ransom. Table 3 shows the four main scenarios for

Table 3: Recovered Resources and Reputation Value for Defender

	Value of Affected Resources (r_{Rec})		
		High (H)	Low (L)
Reputation (r_{Rep})	High (H)	H, H	H, L
	Low (L)	L, H	L, L

different values of r_{Rec} and r_{Rep}. The high and low values of r_{Rec} and r_{Rep} are set by a threshold defined by the defender. The threshold for r_{Rec} is t_{Rec}, value of which is decided by the defender based on the total value of assets it owns. If $r_{Rec} \geq t_{Rec}$ it is said to have a "High (H)" value. Otherwise r_{Rec} is said to have a "Low (L)" value. Similarly, If $r_{Rep} \geq t_{Rep}$ it is said to have a "High (H)" value. Otherwise r_{Rep} is said to have a "Low (L)" value. Predetermining the threshold values helps the defender in attack preparedness. This also helps in making economic decisions with contingency plans in place.

With the strategies and utility functions in place, we now describe how the players make strategic decisions and how the game proceeds. Then we present equilibrium solutions depending upon the various conditions. Thereafter we conduct a sensitivity analysis so that the defender can visualize the expected change in the decision from the change in the parameters. The change in the value of parameter r_{Rec} signifies a change in the importance of the value of encrypted resources to the defender. Change in r_{Rec} is caused by a change in any of the three parameters, viz. R, β and $R_{TotalAssets}$. This would help defender in attack preparedness, and as may be seen through the sensitivity analysis, the effects of change in the

Table 4: Pay-off table for the ransomware attack game

Outcome	Attacker	Defender
1	β	$R-\beta$
2	β	$-R-\beta$
3	0	R
4	0	$-R$
5	λ	0

value of the parameters on the decision making process would be vivid.

4 Preliminary Results

4.1 Decision Making Conditions

Table 4 shows the pay-off for the defender and the attacker for each outcome. When the attacker decides to attack, the maximum pay-off for the attacker is the ransom amount it receives, as represented by $U_A(x_1 = Attack) = \beta$. For this ransomware, the main goal is monetary gain from the ransom received from the victims. If the attacker decides not to attack, then its pay-off is the savings by avoiding the cost of attack as represented by $U_A(x = No\ Attack) = \lambda$.

For $x_1 = Attack$ the following condition must hold,

$$F_1 \equiv U_A(x = Attack) \geq U_A(x = No\ Attack) \equiv \beta \geq \lambda$$

For $x_1 = No\ Attack$ the following condition must hold,

$$F_2 \equiv U_A(x = Attack) < U_A(x = No\ Attack) \equiv \beta < \lambda$$

From the conditions F_1 and F_2 we get,

$$x_1^* = \begin{cases} \text{“Attack”} & \beta \geq \lambda \\ \text{“No Attack”} & \text{Otherwise} \end{cases} \quad (7)$$

For $x_2 = Release$ or $x_2 = No\ Release$ the following condition should hold so that it is in the best interest of the attacker,

$$x_2^* = \begin{cases} \text{“Release”} & \text{y=“Pay”} \\ \text{“No Release”} & \text{Otherwise} \end{cases} \quad (8)$$

The attacker can make decisions based on the pay-off table. The attacker starts the game by making the first move. The first mover's advantage goes to them. When the attacker attacks, the defender is left with the choice of paying or not paying the ransom. Once the defender has made the decision, the attacker decides to release or not release the decryption key. With this decision, the attacker ends the game. The decision for the defender cannot be made easily in a similar fashion. The pay-off table does not quantify the importance of the value of resources for the defender. Moreover, the pay-off table does not guarantee or give insight into the rationality and reputation of the attacker. Consequently, the defender needs to depend on other parameters. Considering this aspect, in this paper we introduced two parameters to help make the defender an informed decision, viz. r_{Rec} and r_{Rep}.

The defender decides the threshold for both the parameters. If the value of the parameter is above the threshold, then it is quantified to have a “High (H)” value, else “Low (L)” value. Once the defender has the values for both parameters, they need to refer to Table 3 in order to make the decision. Therefore, the optimal strategy is

$$y^* = \begin{cases} \text{“Pay”} & r_{Rec} \geq t_{Rec} \text{ AND } r_{Rep} \geq t_{Rep} \\ \text{“No Pay”} & r_{Rec} < t_{Rec} \text{ OR } r_{Rep} < t_{Rep} \end{cases} \quad (9)$$

4.2 Equilibrium Solutions

The game presents the conditions, the strategies, and the pay-offs for each strategy. Considering these factors, Table 5 presents the best responses for both the attacker and the defender. Given the conditions, these best responses translate to equilibrium solutions for the game.

Parameter λ denotes the cost of attack on the part of the attacker. When they decide not to mount an attack, i.e., the strategy is “No Attack”, the pay-off is λ. This is the financial saving they make by avoiding the cost of attack. If the system is harder to infiltrate, then the value of λ is higher. For the defender it means, if the system they build is more secure against infiltration, the value of λ increases which discourages the attacker from mounting the attack. This information is important because this encourages to use a stronger encryption system to secure the database and use security best practices to lower the number of vulnerabilities that might exist in the system.

In the event of an attack, the defender is left with either of the two choices, viz. “Pay” and “No Pay.” This is where the threshold values for r_{Rec} and r_{Rep} comes handy in the decision making process. The defender needs to decide a value for t_{Rec} based on the value of resources under siege and the total value of resources owned by them. This helps to decide the limit at which the defender is comfortable in paying the ransom. For different values of the resources, ransom, and total assets, as r_{Rec} goes below t_{Rec}, the willingness to pay the ransom decreases. The reason being the value of recovered resources becomes less important to the defender.

Table 5: Best Response of the Attacker and the Defender given the conditions (Equilibrium Strategies)

Conditions		**Strategies**	
Attacker	**Defender**	**Attacker**($\widehat{x_1}, \widehat{x_2}$)	**Defender**($\widehat{y}$)
$\beta < \lambda$	N/A	No Attack, Nothing	Do Nothing
$\beta \geq \lambda$	$r_{Rep} \geq t_{Rep}$ AND $r_{Rec} \geq t_{Rec}$	Attack, Release	Pay
$\beta \geq \lambda$	$r_{Rep} < t_{Rep}$ OR $r_{Rec} < t_{Rec}$	Attack, No Release	No Pay

The defender shouldn't rely on the r_{Rec} parameter alone. The reputation of the malware is also important. If the value of r_{Rec} is low, i.e., less than t_{Rec}, the resources are less important to the defender and it can decide not to pay the ransom. But if it is high, i.e., $r_{Rec} \geq t_{Rec}$, the next action the defender should take is to check the value of r_{Rep}. If $r_{Rep} < t_{Rep}$, the reputation of the malware is low. This signifies that if the defender pays the ransom, there is a very high possibility that the attacker wouldn't release the decryption key either. But if $r_{Rep} \geq t_{Rep}$, the reputation of the malware is high and it can be trusted with the payment of the ransom. The best strategy for the defender therefore would be to "Pay" the ransom when $r_{Rec} \geq t_{Rec}$ and $r_{Rep} \geq t_{Rep}$ and "No Pay" otherwise.

The attacker, if feels that the pay-off is higher when $x_1 = Attack$ as compared to $x_1 = No\ Attack$, then mounts the attack. For the basic ransomware considered in this paper, with attacker being rational, it is in its best interest to release the decryption key on receiving the ransom payment and not releasing decryption key, otherwise. Therefore, the attacker's best strategy would be $\widehat{x_1}$, $\widehat{x_2} = Attack$, $Release$ on receiving the ransom payment and $\widehat{x_1}$, $\widehat{x_2} = Attack$, $No\ Release$ if the ransom payment is not made.

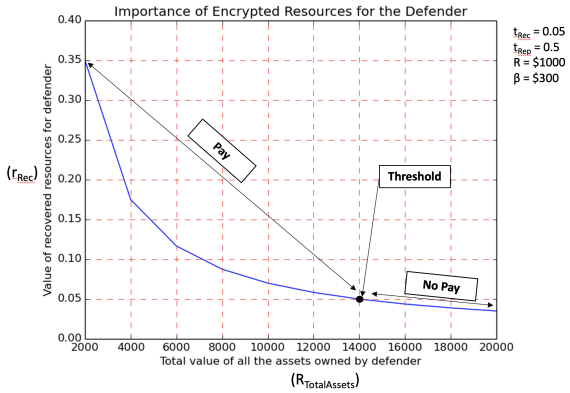
4.3 Sensitivity Analysis

In this section, we perform a sensitivity analysis to determine how the values of ransom and total assets affect the decision making of the defender.

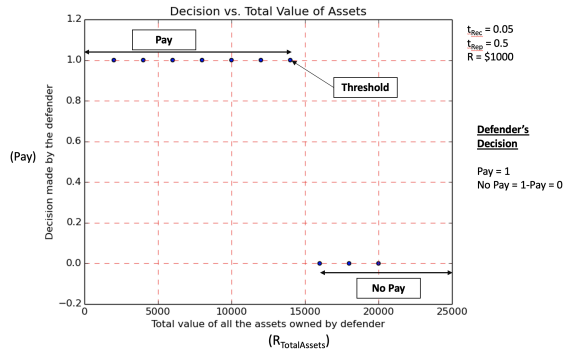
To begin with, we assume an organization with a value of total assets ($R_{TotalAssets}$) of \$10,000. The value of resources (R) under siege is \$1,000. The threshold values for the recovered resources parameter (t_{Rec}) and the reputation parameter (t_{Rep}) are assigned 0.05 and 0.5, respectively. For the purpose of analysis, we use the value of r_{Rep} to be 0.618 (this is obtained by randomly generating a few reputations for past incidents and taking the mean). The ransom value (β) was set at \$300. Now in order to understand how much the total value of all assets affect the decision making process, we vary $R_{TotalAssets}$ while keeping the values of other parameters unchanged. When the value of $R_{TotalAssets}$ is \$14,000, the r_{Rec} value is at the threshold. Figure 2a shows how the value of r_{Rec} changes as we increase the value of total assets owned by the defender. The idea is to visualize the change in the importance of the value of resources under siege for the defender when the value of total assets owned by them changes. The higher the value of r_{Rec}, the more important is the encrypted resources to the defender. Figure 2b shows how the decision of the defender changes as the importance of the value of encrypted resources changes. The decision "Pay" is denoted by the value 1. The decision "No Pay" or not to pay the ransom is denoted by the value 0, i.e., $No\ Pay = 1 - Pay = 0$. The plot shows that when the importance of the encrypted resources diminishes, the willingness to pay the ransom decreases. From equation (1) and Figure 2a, it is apparent that r_{Rec} is inversely related to $R_{TotalAssets}$.

Now, we vary the ransom value from \$100 to \$1,000. We keep the value of R at \$1,000, t_{Rec} 0.05, t_{Rep} 0.5, and r_{Rep} 0.618. Figure 3a shows how increasing the ransom value affects the r_{Rec} value. When the ransom value is \$500, it is the threshold value for r_{Rec}. The figure shows that as the value of ransom increases, the effective value of the recovered resources decreases. Therefore, the importance of the same to the defender decreases and so does the willingness to pay the ransom. Figure 3b shows how increasing the ransom value affects the decision of the defender. The decision "Pay" is denoted by value 1 for a ransom value and the decision "No Pay" is denoted by 0. With an increase in the value of the ransom, the willingness to pay decreases owing to the fact that the effective value of the recovered resources diminishes. From equation (1) and Figure 3a, it is evident that the relationship between r_{Rec} and β is linear with a negative slope.

The sensitivity analysis shows how one can visualize the importance of each parameter in the decision making process. In this paper we presented an example with synthetic data. But, this is applicable in the real world if the defender wants to plug-in real values. This decision making process helps to make an informed decision when faced with an attack. The sensitivity analysis helps to visualize the effect of the attack and helps in attack preparedness on the part of the defender.

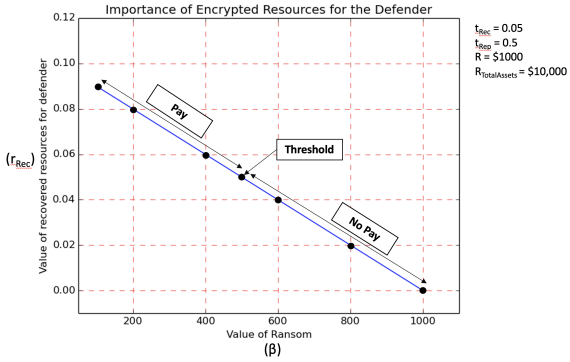


(a) Importance of the resources for the defender

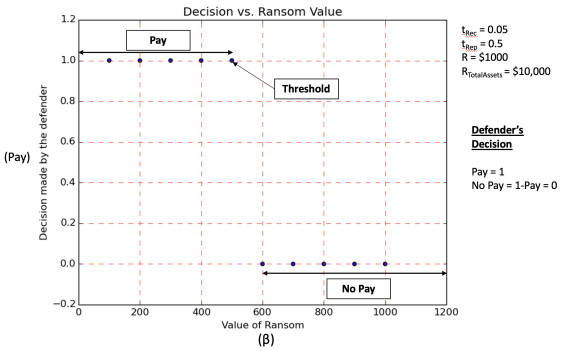


(b) Decision made by the defender

Figure 2: Sensitivity Analysis by varying $R_{TotalAssets}$



(a) Importance of the resources for the defender



(b) Decision made by the defender

Figure 3: Sensitivity Analysis by varying Ransom Value

4.4 Prescriptive Solution

The sensitivity analysis in Section 4.3 shows how the change in the value of ransom or change in the total value of assets owned by the defender affects the decision made by them. The equilibrium strategies for the attacker and the defender are shown in Table 5. The game is designed without any data and the equilibrium strategies were obtained through backward induction. Even though the sensitivity analyses were performed on an example set of data, the game would work fine for any range of data for a basic ransomware attack. Data on ransomware attacks are hard to come by as institutions and organizations often do not report the details fearing the leakage of sensitive information in the public domain and/or adverse effects on their reputation. Therefore, there can be an argument here about the incomplete information game (Harsanyi, 1994). To begin with, one can argue that λ value is an unknown entity to the defender. But through penetration testing and/or employment of ethical hackers that value can be known with quite precision (Krishnan and Wei, 2019) (Zantua et al., 2018). Another value r_{Rep} may seem to be unknown. But through media reports and browsing through historical attacks by the same malware or the same attackers, it can be calculated. If no information is available whatsoever, then the value of r_{Rep} is assumed to be 0.5. This way the game no longer becomes an incomplete information game. With the values under consideration and assumption, Table 5 and Algorithm 1 present a prescriptive solution to a basic ransomware attack. The equilibrium strategies and the algorithm present an opportunity for the defender to prepare in advance and/or make an informed decision when under attack from ransomware.

The basic ransomware may come with few more features. An important feature being an early deadline for ransom payment. After this early deadline, often the value of ransom demanded is doubled. If the defender wishes to pay the ransom, then they will have to pay double the amount after the early deadline. In this scenario, the defender can simply update the values of the parameters wherever applicable, including the value of the ransom. Another feature in the game can be the existence of a bargaining stage between the attacker and the defender. After the bar-

Algorithm 1 Choosing Defender's strategy based on the Optimized Strategy of Attacker from Table 5

t_{Rec} = # Set by Defender based on system config.
t_{Rep} = # Set by Defender based on info collected
if $\beta \geq \lambda$ **then**
  if $r_{Rec} \geq t_{Rec}$ AND $r_{Rep} \geq t_{Rep}$ **then**
    return Pay
  else
    return No Pay
  end if
else if $\beta < \lambda$ AND $\widehat{x_1} = No\ Attack$ **then**
  return Do Nothing
end if

gaining process, if the ransom value changes and/or value of resources under siege changes, then the defender can update the value of the parameters in the game. The value of r_{Rep} is calculated by observing the ransomware attacks which have been known to the defender and/or reported publicly as shown in the eq. (2). If the attack is happening for the first time and/or there exists no reports of historical occurrence of the same, then the defender can proceed with the value of 0.5 for r_{Rep}. But through proper investigation, and if any further clue can be found that links the ongoing attack to some other attack and/or attacker, then the defender can update eq. (2) using the values from those attacks. Thereafter, the defender can update the values of the parameters and tables in the game.

The defender, with the updated game can refer to Table 5 and Algorithm 1 to make an informed decision. Having a strong security system, effective intrusion detection system, strong encryption system and following proper security practices, the defender effectively increases the value of λ. This acts as a deterrent against probable attacks on the system.

5 Conclusion and Future Work

In this paper we used game theory to analyze the attack defense scenario involving a basic ransomware and a victim. We introduced two parameters which would help the defender in making an informed decision when under attack. The r_{Rec} parameter is the ratio of the value of recovered resources after payment of ransom to the total value of all the assets owned by the defender. This helps the defender to quantify the importance of the resources under siege. The higher the value of r_{Rec}, the more willing the defender is to pay ransom. This paper presents a second parameter for decision making so that the defender doesn't lose out much in the game. This parameter r_{Rep} helps the defender to guess how "trustworthy" the malware is. The higher the reputation, the more willing is the defender to pay the ransom. The algorithm presented in this paper summarizes the process to obtain the equilibrium solutions when the defender is under attack from a basic ransomware. The parameters and the algorithm presented in this paper would not only help the defender in preparedness but also help in making an informed decision when under attack.

Through a formal analysis of a basic ransomware, our research provides a preliminary treatment of the mitigation strategies to counter advanced threats. This paper also provides a metric for the payment strategy. The results, approach and the methodology can be used to analyze a more sophisticated ransomware attack, viz. the APT type ransomware. This is part of our future work. The game theoretic analysis for such scenarios would include elaborate games. Another future research could be done on deciding proper threshold values through various forms of investigations including psychological investigations, if applicable. The analysis presented in this paper paves way for future research on APT type ransomware.

ACKNOWLEDGEMENTS

This research is supported in part by the National Science Foundation under Grant No. DGE –1754085. Usual disclaimers apply.

REFERENCES

Baksi, R. P. and Upadhyaya, S. J. (2017). Kidemonas: The silent guardian. *Secure Knowledge Management (SKM '17), Tampa, FL.*

Baksi, R. P. and Upadhyaya, S. J. (2018). A comprehensive model for elucidating advanced persistent threats (APT). In *Proceedings of the International Conference on Security and Management (SAM)*, pages 245–251. The Steering Committee of The World Congress in Computer Science, Computer Enigineering.

Baksi, R. P. and Upadhyaya, S. J. (2020). Decepticon: a theoretical framework to counter advanced persistent threats. *Information Systems Frontiers*, pages 1–17.

BBC (2021). Colonial pipeline boss confirms $4.4M ransom payment. *The British Broadcasting Corporation*.

Cartwright, E., Hernandez Castro, J., and Cartwright, A. (2019). To pay or not: game theoretic models of ransomware. *Journal of Cybersecurity*, 5(1):tyz009.

Çeker, H., Zhuang, J., Upadhyaya, S., La, Q. D., and Soong, B.-H. (2016). Deception-based game theoretical approach to mitigate DoS attacks. In *International conference on decision and game theory for security*, pages 18–38. Springer.

Chinchani, R., Iyer, A., Ngo, H., and Upadhyaya, S. (2005). Towards a theory of insider threat assessment. In *2005 International Conference on Dependable Systems and Networks (DSN'05)*, pages 108–117.

Davis, H. L. (2017). How ECMC got hacked by cyber extortionists – and how it's recovering. *The Buffalo News*.

Deere, S. (2018). Confidential report: Atlanta's cyber attack could cost taxpayers $17 million. *The Atlanta Journal Constitution*.

Gintis, H. (2009). *Game theory evolving*. Princeton university press.

Goud, N. (2017). ECMC spends $10 million to recover from a cyber attack! Cyber Security Insider.

Harsanyi, J. C. (1994). Games with incomplete information. In *Evolution and Progress in Democracies*, pages 43–55. Springer.

Khouzani, M., Sarkar, S., and Altman, E. (2011). A dynamic game solution to malware attack. In *2011 Proceedings IEEE INFOCOM*, pages 2138–2146. IEEE.

Kolodenker, E., Koch, W., Stringhini, G., and Egele, M. (2017). Paybreak: Defense against cryptographic ransomware. In *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, pages 599–611.

Krishnan, S. and Wei, M. (2019). Scada testbed for vulnerability assessments, penetration testing and incident forensics. In *2019 7th International Symposium on Digital Forensics and Security (ISDFS)*, pages 1–6. IEEE.

Milosevic, J., Sklavos, N., and Koutsikou, K. (2016). Malware in IoT software and hardware. *Workshop on Trustworthy Manufacturing and Utilization of Secure Devices (TRUDEVICE'16), Barcelona, Spain*.

Newman, L. H. (2018). Atlanta spent $2.6M to recover from a $52,000 ransomware scare. *Wired*.

Pauna, A. (2012). Improved self adaptive honeypots capable of detecting rootkit malware. In *2012 9th International Conference on Communications (COMM)*, pages 281–284. IEEE.

Selten, R. (1977). A simple game model of kidnapping. In *Mathematical Economics and Game Theory*, pages 139–155. Springer.

Selten, R. (1988). A simple game model of kidnapping. In *Models of strategic rationality*, pages 77–93. Springer.

Sheyner, O., Haines, J., Jha, S., Lippmann, R., and Wing, J. (2002). Automated generation and analysis of attack graphs. In *Proceedings 2002 IEEE Symposium on Security and Privacy*, pages 273–284.

Sheyner, O. and Wing, J. (2003). Tools for generating and analyzing attack graphs. volume 3188, pages 344–372.

Spyridopoulos, T., Oikonomou, G., Tryfonas, T., and Ge, M. (2013). Game theoretic approach for cost-benefit analysis of malware proliferation prevention. In *IFIP International Information Security Conference*, pages 28–41. Springer.

Zakaria, W. Z. A., Abdollah, M. F., Mohd, O., and Ariffin, A. F. M. (2017). The rise of ransomware. In *Proceedings of the 2017 International Conference on Software and e-Business*, pages 66–70.

Zantua, M. A., Popovsky, V., Endicott-Popovsky, B., and Holt, F. B. (2018). Discovering a profile for protect and defend: Penetration testing. In *International Conference on Learning and Collaboration Technologies*, pages 530–540. Springer.

Zimba, A. and Chishimba, M. (2019). Understanding the evolution of ransomware: paradigm shifts in attack structures. *International Journal of computer network and information security*, 11(1):26.