Understanding Bifurcation of Slow Versus Fast Cyber-Attackers

Maarten van Wieren[1(✉)], Christian Doerr[2], Vivian Jacobs[1], and Wolter Pieters[2]

[1] Deloitte Nederland, Amsterdam, The Netherlands
mvanwieren@deloitte.nl
[2] Delft University of Technology, Delft, The Netherlands

Abstract. Anecdotally, the distinction between fast "Smash-and-Grab" cyber-attacks on the one hand and slow attacks or "Advanced Persistent Threats" on the other hand is well known. In this article, we provide an explanation for this phenomenon as the outcome of an optimization from the perspective of the attacker. To this end, we model attacks as an interaction between an attacker and a defender and infer the two types of behavior observed based on justifiable assumptions on key variables such as detection thresholds. On the basis of our analysis, it follows that bi-modal detection capabilities are optimal.

Keywords: Cyber-attack · Economic models · APT · Smash-and-Grab · Information security · Behavioral optimization · Bifurcation

1 Introduction

The exponential rise of connectivity thanks to ICT has made many ways of value creation more efficient. The associated web of connectivity, commonly referred to as "cyberspace", has different scaling properties than our physical world [3] leading to the reduction of typical timescales for interactions, eliminating the need for middlemen and ensuring far more efficiently operating markets (see for instance Van Ark, Inklaar, and McGuckin [13]). As an undesired, but natural side-effect, we have also seen a rise of more "parasitic" forms of value creation in this cyber space. These are agents that make use of its scaling benefits at the expense of other agents' value, e.g. through cyber-attacks. This concerns cyberspace activity linked to commonly known criminal activities such as acts of espionage, fraud, scams, vandalism and terrorism.

Although attribution of cyber-attack to threat actors is still a hard problem [12], it has become apparent that cyber-attacks can be broadly categorized into two groups. On the one hand, there are "Smash-and-Grab"(S&G) type attacks where the threat actor for instance employs malware linked to known vulnerabilities. On the other hand, there are the so-called "Advanced Persistent Threat" (APT) type of attacks, where the threat actor employs the tactic to avoid detection by the defender for as long as possible while slowly realizing their goals. Well known examples are Stuxnet, Duqu, Flame and Red October, which in some cases evaded detection for years [15].

G. Livraga et al. (Eds.): DPM and QASA 2016, LNCS 9963, pp. 19–33, 2016.
DOI: 10.1007/978-3-319-47072-6_2

However, the rationale for the existence of those two groups of fast versus slow attackers is still poorly understood. This paper describes a model for analyzing optimal attack strategies for cyber-attackers depending on detection capabilities of defenders. Attackers having an incentive not to be detected, adopt a type of behavior aimed at remaining unnoticed by the defender's detection capabilities, this means acting slowly. Since acting faster increases the probability of being detected, attackers cannot at the same time act fast and remain undetected, and therefore need to make a choice between these two approaches. In some cases, it is rational to act slow in order to avoid being detected, while in other cases a quick attack makes sense. However, this intuitive argument alone does not explain why attackers may want to choose either fast (S&G) or slow (APT) strategies, since intermediate attack speeds should then also appear. The model explains the observed bifurcation of attack behavior, distinguishing slow and fast cyber-attacks, by showing that intermediate attack speeds are associated to a smaller return on investment for the attacker.

Based on the model, a defense strategy is suggested that implements a stochastic optimization of the parameters under control of the defender. The formulation is kept abstract on purpose, in order to ensure a broad applicability of the model to organizations that differ in their cyber risk capabilities and management, while nonetheless giving insight in the relevant metrics to consider in the first place and the general, organization-independent behavior of attackers and defenders. In practice, to test the bifurcation hypothesis and optimize the defense capabilities for a specific organization, more work is required. Detection parameters should be determined based on the details of the defender's analytics, activity level has to be defined and the loss has to be measured for different activity levels. A relatively simple attacker model like the one presented here, could help to interpret the measurements and put the right capabilities into place.

The remainder of this article is structured as follows. Section 2 describes related work, and Sect. 3 defines the main concepts used throughout the remainder of the article. Section 4 introduces a basic defender detection model. Section 5 relates the defender model to attacker behavior and Sects. 6 and 7 provide analysis of the behavior with respect to this model. Section 8 describes the associated perspective of the defender and Sect. 9 provides conclusions and discussion.

2 Related Work

This paper fits in a tradition of economic modeling of behavior of attackers and defenders, in order to predict or explain real-world phenomena (e.g. Gordon and Loeb [6]). In this paper, we frame the optimization question in terms of the optimal choices for a defender, under the assumption that the population of attackers will also optimize their behavior. This is essentially a minimax optimization in a two-step game, in which the defender moves first [4]. For illustrating the explanation of the bifurcation phenomenon this is sufficient. When assuming that attackers know that defenders take the bifurcation into account in

their strategies, more advanced game-theoretic models are of use. We will come back to this in the discussion.

More specifically, this paper focuses on the time dimension of the behavior of attackers. In this context, several related questions have been addressed, mostly focusing on the optimal timing of attacks. The FlipIt game [8,10,14] investigates timing decisions of attacker and defender moves in order to maximize control over a shared resource, with minimal cost. The basic game only considers a single attack type. Pieters and Davarynejad [11] present a model for deriving attack frequencies from optimal timing decisions of attackers, with different attack vectors and a fixed income for the attacker per unit of time. Axelrod and Iliev [2] discuss the optimal timing of the use of exploits in cyber conflict, taking into account that using an exploit now may make it unavailable for later use. In contrast to this related work, the present paper focuses on the speed of attacks, in order to explain the observed separation between fast and slow strategies. As far as we are aware, this aspect has not been investigated yet.

In our work, key defender parameters are related to detection thresholds. Similar considerations have been studied by others in game-theoretic settings involving multiple attacker types [5]. However, in our current work, the attacker types (fast and slow) are what is explained by the analysis, rather than a starting point. More generally, we are not aiming at developing attacker personas or profiles [1], nor on using those in a security analysis [9], but rather on explaining different styles of attacker behavior that follow from optimization.

3 Definitions

The optimal way to organize defense capabilities for the various assets in an organization depends on the precise incentives of the cyber-attackers targeting specific assets via various attack vectors. Before describing the model in more detail, in this section we discuss various types of assets, attacker motives, and attack vectors, that may assist defenders in their considerations. The rest of the paper should be seen as separately applicable to each of the concepts discussed here.

Cyber-attacks are defined here as an attempt by a threat actor to abuse Information Assets of some defending party. Information Assets are defined as the set of information that holds value to the defender, either direct (i.e. abuse directly reduces value of the defender) or indirect (i.e. where abuse leads to loss of value for third parties associated to the defender). Indirect losses may of course materialize in further direct losses through fines and/or claims. Value can take multiple forms, the most commonly ascribed values are: economic, financial, well-being, human lives, culture, nature, political, etc. For the purpose of this article and without prejudice to other forms of value, we have foremost financial value in mind.

Information Assets may be characterized through the well-known Confidentiality - Integrity - Availability triad. Confidentiality means an Information Asset may contain information asymmetry that leads to the potential to create value

and/or to the potential to destroy value. Integrity means an Information Asset may contain records of reference that if the reference is changed it destroys value for the defender. Availability means an Information Asset may be (partially) lost so that, even if it is only temporarily, it destroys value for the defender. Information Assets may fall into all three categories or combinations.

Abuse of Information Assets typically arrives in three forms: cyber-espionage, cyber-fraud and cyber-destruction. Cyber-espionage concerns breaking information asymmetry (confidentiality), where usually it has more value to the attacker if this remains unknown to the defender. Typical motives for the attacker include: market competition, geo-politics, national-defense and insider trading. Cyber-fraud concerns breaking the integrity of the Information Assets. Here it depends on the motive if the attacker is even able to keep the cyber-fraud hidden to the defender after the attack. Some motives for cyber-fraud include: payments and transactions fraud, cover-ups of criminal acts, terrorism, war, accreditation and smuggling. Finally, cyber-destruction means making an Information Asset (temporarily) unavailable. Some motives for cyber-destruction include: hacktivism, terrorism, war, extortion and competition.

Of course combinations of these three forms into a composite attack is also possible. This means there is an initial attack followed by another type of attack. One example of this is where intelligence gets stolen to assist in a follow-up attack. Another example is a DDoS attack to momentarily distract the defender from another attack. In particular, the most dangerous type of attacks concerns abuse of the integrity of (security) controls as a pre-cursor for any other type of attack. Clearly, such a composite attack would classify as a sophisticated attack given that it requires a wide range of capabilities from the attacker. On the other hand, "unsophisticated" attackers that employ a more limited set of (known) techniques to exploit (known) vulnerabilities, may still cause significant levels of abuse since they can operate more agile thus quickly and on a larger scale.

We define two layers of defense in the security architecture description. The first layer concerns the technical/physical boundary between the public and the private domains of any network. If an internet connection exists, then vulnerabilities are likely to be identifiable. The second layer concerns the boundary between the private domain and the Information Asset at risk. All (technical) protection measures that are in place to prevent any form of abuse of Information Assets is part of the second protection layer. With respect to these defense layers, three channels can be identified for cyber-attacks. The first attack channel makes use of critical vulnerabilities in the first defense layer to gain access. The second attack channel works through insiders (knowingly or unknowingly, effectively circumventing the first protection layer with the knowing or unknowing help of insiders within the firm, granting them instant access across the first layer of defense. The third attack channel is through third parties, effectively circumventing the first two protection layers. This is the case for instance with a DDoS attack or if data gets abused in the "cloud" (which is in effect a third

party's computer). Combinations of attack channels is of course again possible and associated to more sophisticated attackers.

When considering a defense strategy, defenders must know what Information Assets to protect, which attacker types these Information Assets attract, and which potential attack channels may be used. Some attacks like cyber-espionage and cyber-fraud on intangibles are only likely to occur as slow attacks. In contrast, cyber-fraud on tangibles as well as cyber-destruction attacks will likely end in a fast phase. Defenders thus will want to be capable to deal with slow as well as fast attacks, which must be dealt with by developing bi-modal detection capabilities: one slower regime with as low granularity as possible against slow attacks and one fast regime with higher granularity against fast attacks. We will come back to this in Sect. 8.

4 Modeling Detection and Response

In this section we set up a model for the defender's detection capability. Attackers will adapt their behavior in line with their goal(s). We assume that they can adapt their activity level, i.e., the number of attack-related moves against or in the defender's system per unit time. We assume that attackers get closer to their goal(s) by abusing Information Assets, and define the (average) rate of abuse of an Information Asset by the attacker as proportional to its activity level. This means that attackers in absence of defense simply have an incentive to act as quickly as possible to realize their goal(s).

The defender has the capability to detect and respond to an (attempted) attack. With the typical detection setup described below there is a certain monotonically increasing probability per activity level of the attacker that an attack will be detected. This also means that there is a typical time it takes the defender to neutralize the attack. Initially, we set the detection capability to be fixed, later we will consider that it may be varied by the defender.

Detection depends on identifying suspect activity with respect to normal activity. For this purpose, the defender will continuously sample a given scope containing a number of continuously changing elements $0 \ll S \in \mathbb{N}$ to test for suspect behavior.[1] For this test, a selection threshold $\theta_0 > 0$ is set that is defined through the expected number of elements S_0 that will be considered suspicious based on detection granularity $a_0 > 0$ without being associated to a specific attack (false positives):

$$S_0 = Se^{-\frac{\theta_0}{a_0}}. \tag{1}$$

This indicates that increasing the threshold θ_0 will reduce the number of suspicious elements, while lower detection granularity a_0 reflects an improving capacity of the defender to pick out suspect behavior. Lower detection granularity a_0 would thus reduce the number of false positives.

[1] A typical example is an analytics capability scanning through a large number of log files generated periodically by the system, checking them against predefined (mis)use cases or rules.

Depending on the activity level a of the attacker,[2] some small fraction of the scope is associated to a specific attack. The number of elements $S_a > 0$ that are actually detected as suspect elements depends on the threshold:

$$S_a = Se^{-\frac{\theta_0}{a}}. \tag{2}$$

Thus, the total number of suspicious elements is the sum of the false and true positives, $S_0 + S_a$. Suppose the defender randomly picks a suspect element, then the conditional probability that investigation of this element will lead to detection of the attack is defined as:

$$P_D = \frac{S_a}{S_0 + S_a} = \frac{1}{1 + e^{\theta_0(\frac{1}{a} - \frac{1}{a_0})}}. \tag{3}$$

Now suppose the scope is refreshed on regular intervals of duration T_r and let $p \in \mathbb{N}$ denote the investigative power of the defender, determining how many of such suspect elements can be investigated in time $T_r > 0$. Then, in case $p \ll S_a + S_0$, the rate (probability per unit time) of detection may be approximated by:

$$r_D = \frac{1 - (1 - P_D)^p}{T_r}. \tag{4}$$

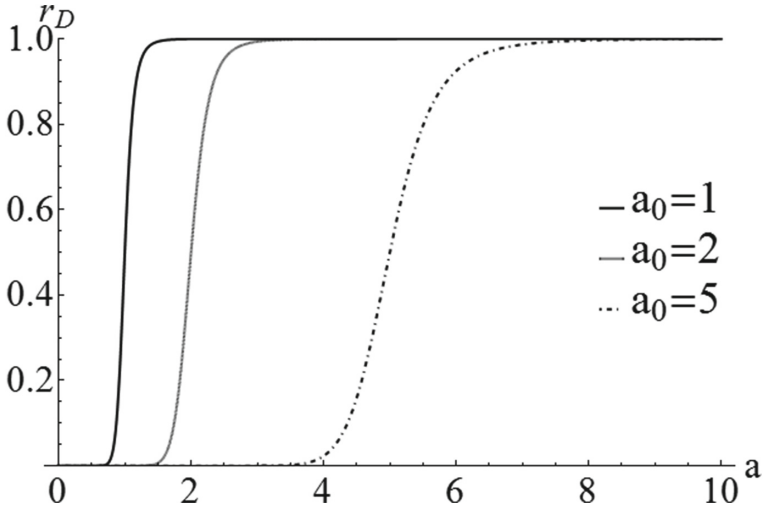

Fig. 1. Detection rate r_D as a function of attacker activity a for $T_r = p = 1$, $\theta_0/a_0 = 15$ and for various values of a_0, where a_0 determines the a-value for which detection has a crossover from a low to a high rate.

[2] The activity level parameterizes in an abstract and general way the number of actions performed during the attack per unit time. A concrete value depends on the details of the attack and the system. E.g., it may be the rate of data exfiltration from the defender's network.

Table 1. Model parameters and their meaning.

a_0	Detection granularity
θ_0	Detection threshold
T_r	Time interval between refreshes
p	Investigative power

In Fig. 1 some sample graphs are displayed for the detection rate as a function of the activity, with $T_r = 1$ and $p = 1$.

An interpretation guide to the various degrees of freedom of our detection model is given in Table 1.

The expected time before the attack will be detected is then given by

$$T_D = \frac{1}{r_D} = \frac{T_r}{1 - (1 - P_D)^p}. \tag{5}$$

After detection, it will take some additional time $T_N > 0$ before the attack will actually be neutralized through the response function so that the total expected maximum duration of an attack is

$$T_{A,\max} = T_D + T_N. \tag{6}$$

We can be more concrete by estimating the typical values and ranges for the parameters in the model based on a realistic situation. A typical refresh time is of the order of hours to days, so $T_r = 1$ (day). Based on our experience in the field, the fraction of false positives may vary between $10^{-7} - 10^{-3}$, depending on the maturity of the defender's analytics. This implies a range for θ_0/a_0 of 7 to 16 for a typical and mature defender, respectively. Furthermore, the investigative power p for a typical and mature defense system will lie between 1 and 1000 respectively. For instance, the number of employees judging suspicious elements can be a proxy for the investigative power p of the organization at hand.

5 Optimizing Attacker Strategy

Consider the return on investment for a collective of attackers with varying activity levels. For a given pair of fixed defender-attacker, we assume that the rate at which the attacker accumulates benefits is equal to the rate at which the defender accumulates losses[3]

$$r_{\text{abuse}} = C \cdot a, \tag{7}$$

[3] This assumption takes into account loss occurring within any time interval after an attack. Not only incidents with a direct financial loss result in value loss for an organization. Also indirect impact in the form of lost investments and future income, as well as the consequences of (so far) unnoticed attacks usually lead to value loss for the defender in the long term.

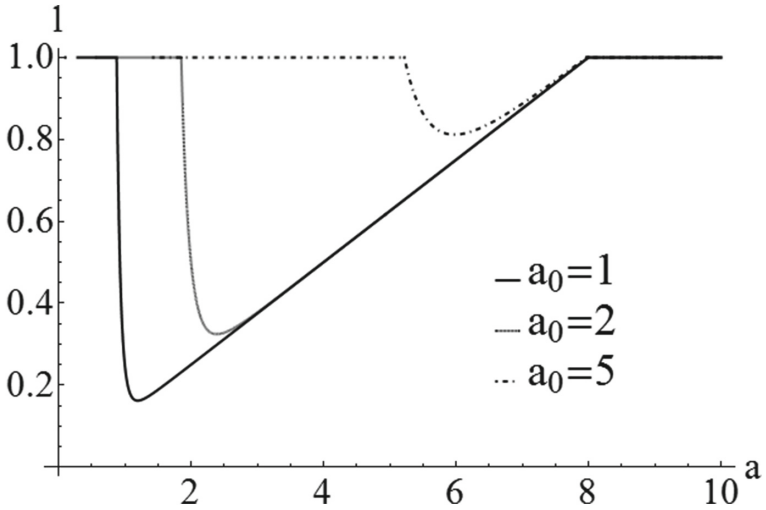

Fig. 2. Loss fraction l as a function of a for various values of the parameter a_0, and for $C = p = T_r = 1$, $T_N = 0$, $L_{\max} = 8$, and $\theta_0/a_0 = 15$. The minimum $a_{\min}$ is apparent. Clearly, lowering a_0 results in a decrease of the minimum $a_{\min}$, while at the same time lowering the loss fraction $l(a_{\min})$.

where $C > 0$ is some constant that does not depend on the defender's detection and response function. The expected total loss cannot exceed a certain limit given a finite size of the Information Asset being abused. It is therefore bounded from above by the total exposed value $L_{\max}$ and is given by

$$L(a) = \min(r_{\text{abuse}} \cdot T_{A,\max}\,,\, L_{\max}). \tag{8}$$

Assuming an undetected attacker leaves the system when the maximal value $L_{\max}$ is extracted, the maximal attack time is $L_{\max}/C\,a$. Figure 2 displays the expected fraction of value loss $l = L/L_{\max}$ for the same parameter values as used in the graph for the detection rate in Sect. 4.

Disregarding the trivial case when the loss fraction is equal to 1 for all a (not sufficiently low a_0), it has a non-trivial minimum as a function of attacker activity at some $a_{\min}$ so that

$$l(a) \geq l(a_{\min}) \qquad \forall\, a > 0. \tag{9}$$

The minimum $a_{\min}$ can be determined analytically for $p = 1$, in which case it is given by

$$a_{\min} = \frac{\theta_0}{1 + X}, \tag{10}$$

where X is the principal solution of $Xe^X = (T_r + T_N)e^{\theta_0/a_0 - 1}/T_r$, also known as the Lambert-W function. For $p \neq 1$, $a_{\min}$ can be proven to exist given continuity and boundedness of $l(a)$. In this case, $a_{\min}$ can be computed numerically.

For a given type of Information Asset, the loss for the defender will in first approximation be linearly proportional to the (gross) gains for the attacker $\tilde{G}(a)$, which we define by

$$\tilde{G}(a) = G_0 \cdot L(a). \tag{11}$$

Here, the proportionality factor G_0 is usually of the order $10^{-2} - 10^{-1}$. The existence of a minimum in the loss fraction combined with the fact that defender loss is proportional to attacker gain, implies that for attackers to optimize gains, they have an incentive to act either more slowly or more quickly (the derivative is either positive or negative, depending on which side of the minimum the attackers find themselves). In fact, given that a singular minimum $a_{\min}$ exists (a_0 low enough), it also follows that for every activity level $a_{\text{slow}} < a_{\min}$ there is at least one value $a_{\text{fast}} > a_{\min}$ such that

$$l(a_{\text{slow}}) = l(a_{\text{fast}}). \tag{12}$$

From this it already follows that attackers (also depending on their exact properties) will tend to split into two categories: slow and fast attackers. This is a bifurcation phenomenon and heuristically represents the two main strategies that attackers may follow with respect to detection by the defender: stealth (not get detected) or speed (act quicker than defenses). This may be a rational strategic choice by the attacker related to the details of their objectives and capabilities as portrayed in Sect. 2. However, even without such rational decision making, this will be the result of a selection mechanism where successful attackers amplify strategies that have worked best in the past.

6 Economic Considerations Attacker

For better understanding of attacker behavior, we need to consider the net gains for the attacker by including the costs and limitations associated to an attack. For this purpose we observe the two extremes of very slow and very fast attacks.

For very slow attacks, the time required to accumulate gain becomes prohibitively long given that attackers need to invest an increasing amount of time. During this time they will have a fixed level of expenses (living cost as well as the cost of invested capital at risk due to uncertain returns). This implies that costs are proportional to the time required for the attack so we include $c_0\, T_A(a)$ as a cost term (we assume zero interest returns on the invested capital).

For very fast attacks, it will be increasingly costly for the attacker to arrange the required infrastructure and capabilities. This effect can be summed up as the law of diminishing returns. A given increase in activity level will cost an exponentially increasing amount of investment for the required capabilities

$$J(a) = J_0 e^{\frac{a}{\theta_J}}. \tag{13}$$

Here, J_0 is a (small) fixed investment cost and θ_J is the capability investment threshold.

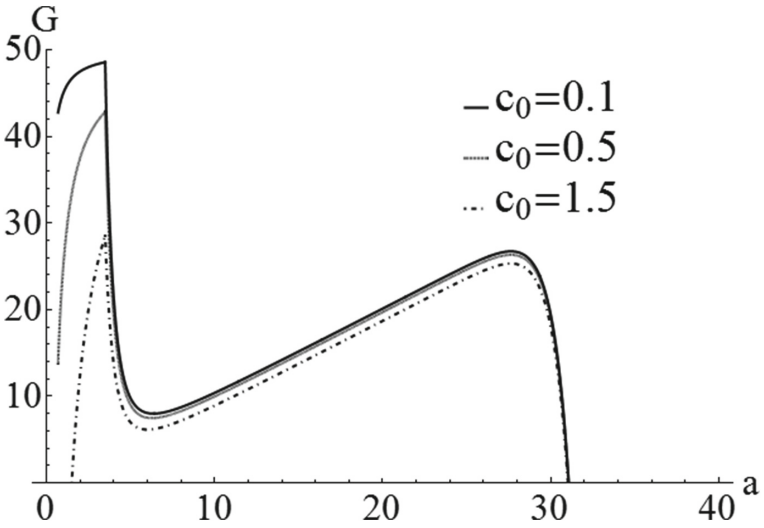

Fig. 3. Net gain function G as a function of a for various values of the operation cost c_0, and for $C = p = T_r = G_0 = \theta_J = 1$, $T_N = 0$, $L_{\max} = 50$, $J_0 = 10^{-12}$, and $\theta_0/a_0 = 6$.

Combining the gross gain with the time-dependent cost and the up-front investment for the capabilities, we define the net gain as

$$G(a) = G_0 \cdot L(a) - C_0 \cdot T_A(a) - J(a). \tag{14}$$

This net gain function is plotted for some parameter values in Fig. 3.

The net gain has two local maxima, corresponding to a slow and fast optimal activity level with respect to net gain. Which one of these is the global maximum depends on the choice of parameter values. In Fig. 3 we see that the highest gain is obtained by slower attackers. In between these two local maxima there is a minimum in the gain function, meaning the bifurcation mechanism is still intact.

7 Attacker Behavior Analysis

We now revisit the types of cyber-attack we defined earlier (cyber-espionage, cyber-fraud and cyber-destruction) and will relate these to the analysis made in the previous section and will see what considerations attackers will have to move either fast or slow. Here, we define fast attackers as having an activity $a > a_{\min}$, while slow attackers have an activity $a < a_{\min}$.

In case of cyber-espionage, there is typically a clear incentive that the attack does not get uncovered after the fact. This means there is a penalty involved for detection that changes the gain function to have only a single optimum. This means that there is an incentive to move slowly, consistent to what is for example being observed with APT's. (This obviously does not mean that all

such attacks go undetected.) In case that it doesn't matter to the attacker if the defender knows about the attack however, then fast attacks for the purpose of cyber-espionage are still perfectly rational and this has indeed recently been observed.

In case of cyber-fraud, it will depend on the details whether slow or fast attacks are attractive. If the fraud concerns tangibles (for instance money transactions), then it can safely be assumed by the attacker that it will be uncovered after the abuse has succeeded given the many controls in place on the defender side. In this case it becomes attractive to move fast as is commonly observed. If the fraud on the other hand concerns intangibles, e.g. hiding criminal acts, then the goal is to remain undetected after the fact. Other types of attack may also benefit from cyber-fraud attacks on intangibles, e.g. through placement of back doors or covering tracks.

In case of cyber-destruction, it is clear that the abuse will get detected after the fact, so in principle there is no real incentive to move slowly. Even more strongly, these attackers have the incentive to move quickly to facilitate a broad reach before the unavoidable detection will lead the defender to block further attempts.

Of course, attackers do not need to choose a fixed strategy i.e. activity level. For all fast attacks, a preparatory attack (for instance to weaken controls of the defender) is an option. These are the composite attacks referred to in Sect. 3. The incentive for such preparatory attacks is to remain undetected at least until the fast attack ensues. This means that the preparatory attack will benefit from slow movement. Conversely, fast attacks like DDoS attacks have also been used in composite attacks to serve as distraction for the defender. Such composite attacks however require a significant level of sophistication that is fortunately still relatively rare. For less sophisticated attackers, the smash-and-grab tactics still make perfect sense.

Finally, another consideration that may favor fast attacks as compared to slow attacks is that fast attacks allow for a larger number of targets in a given time-span than is possible for slow attacks. This larger sample of defenders means the fast attacker likely encounters multiple distinct realizations of detection capabilities. This implies that diversification effects will tend to dampen the volatility in results that fast attackers will experience, thus leading to a more stable (criminal) return on investment. In terms of natural selection principles acting to favor one attacker over the other, this helps the persistence of fast attackers.

8 The Defender's Dilemma

Typical defenders have to make choices on what capability to invest in to obtain sufficient and optimal security. Given the gain function for attackers described above, which has two local maxima, the defender's hypothesis is that the attacker population is split into slow and fast attackers. With respect to each population, two optimal defense configurations exist for each population that minimize their impact. Here we observe the considerations from the perspective of the defender related to this optimization.

Consider the capabilities associated to the model introduced in Sects. 4 and 5. There are four degrees of freedom that can be controlled by the defender: the detection threshold θ_0, the detection granularity a_0, the power p and the time to refresh the scope T_r. From the model as well as logic, it follows that the defender should typically aim to reduce the detection threshold, limit and refresh time, while increasing the detection power.

Optimizing the value of these parameters will also lower $a_{\min}$ as can for example be seen in Fig. 2, reflecting that it requires ever lower attacker activity levels to remain undetected sufficiently long to benefit as attacker. Although this effect may be small on short time-scales, attackers will tend to also adapt to changing properties in the environment created by the defender (and making the attacker move slower is of course still a good thing since it will take longer for the same loss to accumulate).

However, the impact for each parameter is not the same for slow or fast attackers as can be seen in Fig. 4. The impact of lowering the detection granularity a_0 is foremost lowering the optimal activity rate for slow attackers, while there is hardly any impact for fast attackers, thus forcing attackers to act fast. The same holds true for improving the detection threshold θ_0 and power p. In contrast, lowering the time to refresh the detection scope T_r and time to neu-

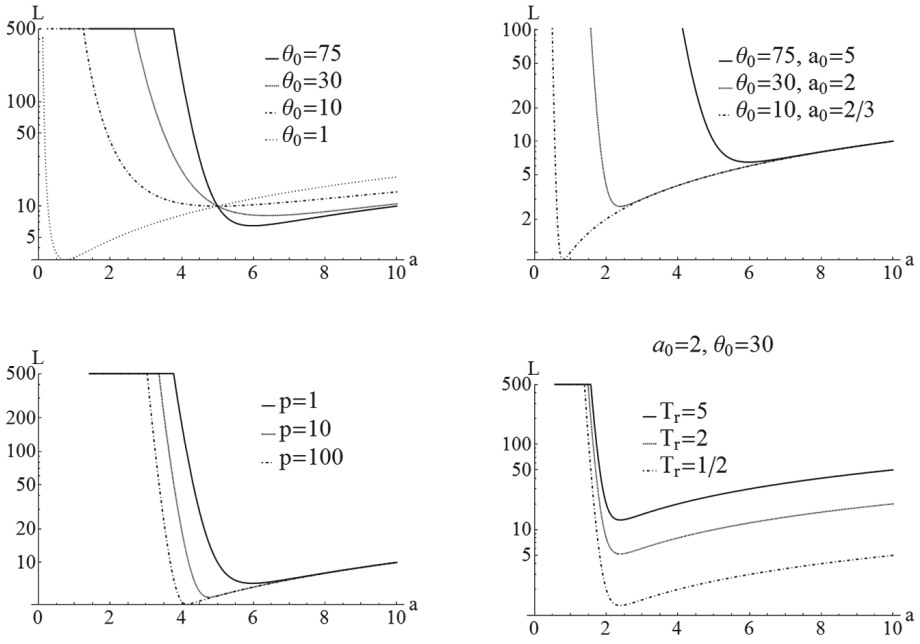

Fig. 4. Effects of parameter variation on the loss function. When parameter values are not explicitly mentioned we have taken the values $p = T_r = C = \theta_J = 1$, $T_N = 0$, $A_0 = 5$, $L_{\max} = 500$ and $\theta_0 = 15$. The panels in this figure illustrate the effects discussed in Sect. 8 and represented in Table 2.

Table 2. Effects of parameter optimization on fast and slow attacks. In each row all unmentioned parameters are kept fixed. The term "reduces" indicates that the defender's loss $L(a)$ decreases for fast or slow attacks, i.e., high or low values of a, respectively.

Effect on loss	Slow attack	Fast attack
Decrease a_0	Reduces	No
Decrease θ_0	Reduces	Increases
Decrease T_r	Reduces very little	Reduces
Decrease T_N	Reduces very little	Reduces
Increase p	Depends	Reduces slightly

tralize the attack T_N are only effective to reduce the impact from fast attacks. However, given that a decrease of the time to refresh T_r will likely reduce the detection power p, this may in effect work counterproductively with respect to countering slow attacks. The abovementioned effects of optimizing the four parameters on fast and slow attackers are summarized in Table 2.

From this follows a defender's dilemma when investing in capabilities. As follows from the analysis in Sect. 7, most defenders will need to defend against slow and fast attacks alike and will thus have to make tough choices on how to optimize with respect to both types of attack. The dilemma here is: do we make sure we detect even very slow attackers at the expense of reacting to fast attackers, or do we make sure that fast attackers are dealt with quickly and hope that no attacker arrives that is too slow? The way to deal with this dilemma, is to have the defense capabilities act in two regimes simultaneously. We refer to this as bi-modal detection where part of detection resources should be spent on acting quickly with higher granularity (i.e. lower resolution), while another part of the resources should be spent on carefully checking all elements derived with low threshold θ_0 and feeding the results back into reducing the detection granularity a_0 further and further (i.e. increasing resolution). By creating a linear combination of these two regimes in this way, the resulting optimum will be better for the same number of resources as when only optimizing a single configuration (i.e. set of parameters).

9 Conclusion and Discussion

This paper describes a model aimed at analysis of the interaction between attackers avoiding detection and defenders attempting to detect and neutralize attacks. We have seen that there are two natural optima for attackers: moving fast and moving slow. This coincides with observed properties of real world attackers. As far as we are aware, this is the first analytical model showing this bifurcation in time of attacker behavior under minimal and logical assumptions.

In Sect. 8 we recommend creating a bi-modal detection capability. Such a bi-modal detection would benefit from a quantitative analysis of the performance

of detection modes against the framework set out in Sects. 4, 5, 6 and 7. This is left for future research.

In the present paper, we have assumed that attackers will naturally tend to optimize their behavior for a given defense configuration. In reality however, the defenders as well as the attackers may choose to adapt their strategies. For instance, we have not considered the possibility of multiple attacks by the same attacker. This could also be of interest when to each attack an initial cost is associated [7]. In our case, this cost depends on the quality of the protection capability. An example of such a strategy could be attackers attempting to exhaust the defensive detection capability with many fake fast attacks for the purpose of hiding the actual slow attack. Defenders could adapt again to such strategies, and investigating this interaction in a game-theoretic model would be interesting as well for future research.

Acknowledgements. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007–2013) under grant agreement ICT-318003 (TRESPASS). This publication reflects only the authors' views and the Union is not liable for any use that may be made of the information contained herein.

References

1. Atzeni, A., Cameroni, C., Faily, S., Lyle, J., Fléchais, I.: Here's Johnny: A methodology for developing attacker personas. In: Sixth International Conference on Availability, Reliability and Security (ARES), pp. 722–727. IEEE (2011)
2. Axelrod, R., Iliev, R.: Timing of cyber conflict. Proc. Nat. Acad. Sci. **111**(4), 1298–1303 (2014)
3. Barabási, A.L., Albert, R., Jeong, H.: Scale-free characteristics of random networks: the topology of the world-wide web. Physica A Stat. Mech. Appl. **281**(1), 69–77 (2000)
4. Cox Jr, L.A.T.: Game theory and risk analysis. Risk Anal. **29**(8), 1062–1068 (2009)
5. Dritsoula, L., Loiseau, P., Musacchio, J.: Computing the nash equilibria of intruder classification games. In: Grossklags, J., Walrand, J. (eds.) GameSec 2012. LNCS, vol. 7638, pp. 78–97. Springer, Heidelberg (2012)
6. Gordon, L.A., Loeb, M.P.: The economics of information security investment. ACM Trans. Inf. Syst. Secur. (TISSEC) **5**(4), 438–457 (2002)
7. Herley, C.: The plight of the targeted attacker in a world of scale. In: WEIS (2010)
8. Laszka, A., Horvath, G., Felegyhazi, M., Buttyán, L.: FlipThem: Modeling targeted attacks with FlipIt for multiple resources. In: Poovendran, R., Saad, W. (eds.) GameSec 2014. LNCS, vol. 8840, pp. 175–194. Springer, Heidelberg (2014). doi:10.1007/978-3-319-12601-2_10
9. Lenin, A., Willemson, J., Sari, D.P.: Attacker profiling in quantitative security assessment based on attack trees. In: Bernsmed, K., Fischer-Hübner, S. (eds.) NordSec 2014. LNCS, vol. 8788, pp. 199–212. Springer, Heidelberg (2014)
10. Nochenson, A., Grossklags, J., et al.: A behavioral investigation of the FlipIt game. In: Proceedings of the 12th Workshop on the Economics of Information Security (WEIS) (2013)

11. Pieters, W., Davarynejad, M.: Calculating adversarial risk from attack trees: control strength and probabilistic attackers. In: Garcia-Alfaro, J., Herrera-Joancomartí, J., Lupu, E., Posegga, J., Aldini, A., Martinelli, F., Suri, N. (eds.) DPM/SETOP/QASA 2014. LNCS, vol. 8872, pp. 201–215. Springer, Heidelberg (2015)
12. Rid, T., Buchanan, B.: Attributing cyber attacks. J. Strateg. Stud. **38**(1–2), 4–37 (2015)
13. Van Ark, B., Inklaar, R., McGuckin, R.H.: Changing gear: productivity, ICT and-service industries in Europe and the United States. The Industrial Dynamics ofthe New Digital Economy, Edward Elgar, pp. 56–99 (2003)
14. Van Dijk, M., Juels, A., Oprea, A., Rivest, R.L.: FlipIt: the game of "stealthy takeover". J. Cryptology **26**(4), 655–713 (2013)
15. Virvilis, N., Gritzalis, D.: The big four - what we did wrong in advanced persistent threat detection? In: Eighth International Conference on Availability, Reliability and Security (ARES), pp. 248–254. IEEE (2013)

Springer

http://www.springer.com/978-3-319-47071-9

Data Privacy Management and Security Assurance
11th International Workshop, DPM 2016 and 5th International Workshop, QASA 2016, Heraklion, Crete, Greece, September 26-27, 2016, Proceedings

Livraga, G.; Torra, V.; Aldini, A.; Martinelli, F.; Suri, N. (Eds.)

2016, XIV, 247 p. 74 illus., Softcover
ISBN: 978-3-319-47071-9