A system dynamics, epidemiological approach for cyber-resilience to zero-day vulnerabilities

Daniel A. Sepúlveda-Estay
Department of Technology, Management and Economics,
Technical University of Denmark DTU[*]

September 9, 2020

Abstract

Cyber-attacks are serious threats to operations in most industries, enabled by a growing dependence on Information Technology (IT). To minimize disruptive effects on operations, organizations with complex system derive value both from preventing cyber-attacks and from responding promptly and coherently when cyber-attacks happen, capacity is known as cyber-resilience. Frameworks have been presented in literature to promote cyber-resilient response, yet little is known about the structures that result in a cyber-resilient behavior. This paper explores an approach to modeling the structure of a system that is subject to an infection an eventual recovery from zero-day malware cyber-attacks, based on mechanisms derived from epidemiology. By analyzing the relationship between the system vulnerabilities and the incidence of malware infections in a population of systems, this paper derives structural recommendations for resilience response, and policy requirements based on the claim that cyber-threats are a public-cyber-health issue instead of merely a competitive factor.

Keywords— System Dynamics, Cyber-Epidemiology, Cyber-Resilience

1 Introduction

Increasingly complex industrial systems, largely enabled by their dependence on Information Technology (IT), are a fundamental feature of modern international operations. The continuous inter-connectivity of critical infrastructure provided by IT is essential for the operation of these complex, critical systems in areas such as healthcare, transportation systems, and electrical power systems, for example. Moreover, many of the support services required for these industries are online and thus dependent on connectivity, to simplify the identification and solution of operational issues.

Cyber-attacks are malicious and deliberate attempts by agents (individuals or organizations) to access the information system of another individual or organization, with the objective to obtain a benefit from accessing the victim's network. The capacity of a system to recovering from the consequences of a cyber-attack, called its cyber-resilience, has been thus identified as

[*]corresponding author email: dasep@dtu.dk, ORCID:0000-0001-5224-622X

a desirable system capability (Khan & Sepulveda Estay, 2015). The process of maintaining or returning to a level of operational performance after a disruption, is represented through a disruption curve (Sheffi & Rice Jr, 2005). This curve has served as a basis for the description and quantification of the resiliency of a system (Munoz & Dunbar, 2015) (Barroso, Machado, Carvalho, & Machado, 2015), and as a basis for the quantification of the effects of cyber resiliency (Sepulveda Estay & Khan, 2015).

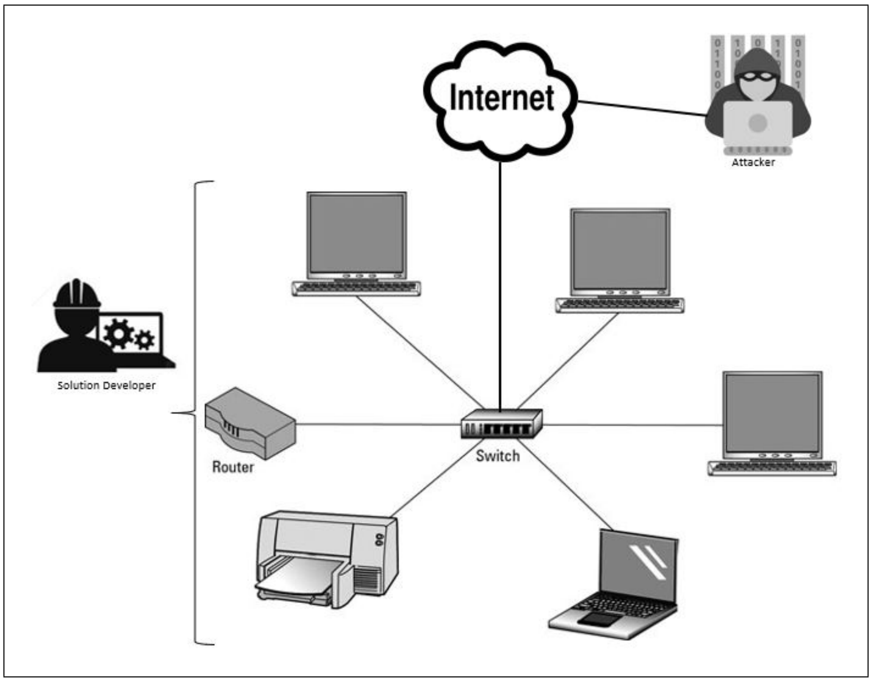


Figure 1: Network example with attacker and Solution developer

Multiple examples of successful cyber attacks with effects on operations (Walters, 2015) (Leslie, Harang, Knachel, & Kott, 2018) evidence that such enhanced connectivity also results in new challenges for maintaining the availability, integrity and confidentiality of these services and their associated information. The accidental creation of vulnerabilities which expose interconnected systems to attacks is an unintended negative consequence of building complex systems. As an example of complexity affecting response, a 2020 survey by Cisco found that after a cyber attack, 39% of companies with 1 security vendor on average had more than 4 hours downtime. in contrast, 73% of companies with more than 20 security vendors had more than 4 hours downtime. Organizations must prepare, identify, respond and adapt once these vulnerabilities are breached affecting operations.

The way in which networks are connected provides a contact between computers that results in their exposure to cyber-attacks. Baldwin et al.,(Baldwin, Gheyas, Ioannidis, Pym, & Williams, 2017) revealed that attacks on individual ports within networks are inter-related, and as a result the management of cyber attacks should include "*a degree of correlation as it represents the impact of contagion on the level of risk*".

Cyber-attacks can happen in a variety of ways such as through Malware (Grégio, Afonso, Filho, Geus, & Jino, 2015) (Soliman, Sobh, & Bahaa-Eldin, 2017), Phishing (Gupta, Arachchilage, & Psannis, 2018) (Pienta, Thatcher, & Johnston, 2018), Man-in-the-middle attack (Conti, Dragoni, & Lesyk, 2016), Denial-of-Service attack (Specht & Lee, 2003) (De Donno, Giaretta, Dragoni, & Spognardi, 2017), SQL injection (Som, Sinha, & Kataria, 2016) (McWhirter, Kifayat, Shi, & Askwith, 2018), Zero-day exploits (Fagioli, 2019), and DNS Tunneling (Soliman et al., 2017) (Alieyan et al., 2019). Table 1 compares cyber-attack types according to their activation, reproducibility, passive versus active strategy, and provides some examples for each

Table 1: Cyber-attack types and characteristics

Type	Description	Activation	Reproducible	Active/ Passive	Examples
Malware	Set of instructions that run on a system to make it do arbitrary activities on behalf of an attacker, or to act in a way (automated or not) that threatens security aspects of the compromised system, its users and associated data (Grégio et al., 2015).	Automatic /Human-activated	Perfect	Both	Virus, worms, ransomware, troyan horses.
Phishing	Process to steal user's credentials using fake emails or websites or both (Gupta et al., 2018)	Human-activated	Imperfect	Passive	Email scams, spear phishing.
Man in the middle	Malicious access to a communication channel between two endpoints, As a result, the attacker has convinced both victims that they use secure channel, but in reality it has access to and can manipulate all encrypted messages (Conti et al., 2016)	Human-activated	Imperfect	Active	Spoofing-based, actve network interception, IP Hijacking, False Base station.
Denial of Service	Coordinated attack on given target system or network by requiring an very high number of services, and thus effectively blocking the system to legitimate service requests (Specht & Lee, 2003)	Automatic /Human-activated	Perfect	Active	Agent-handler, Reflector, IRC-based, Web-based
SQL injection	Alteration of SQL (Standard-Query-Language) statements that are used within a web application through the use of attacker-supplied data, to corrupt or expose victim data (Som et al., 2016)	Automatic	Perfect	Passsive	Authentication bypass, imformation disclosure, remote command execution
Zero-day exploits	Security exploit based on a programming or hardware flaw that has not yet been disclosed to the vendor or developer (Fagioli, 2019)	Automatic	Perfect	Both	Zero-day code
DNS Tunneling	Exploit where the DNS protocol to tunnel malware and other data through a client-server model (Alieyan et al., 2019)	Automatic	Perfect	Both	Botnets

cyber attack type.

Malware (MALicious softWARE) in particular has been identified as one of the top threats in the cyber-security landscape (Sfakianakis, Douligeris, Marinos, Lourenço, & Raghimi, 2019). New malware attacks, such as ransomwares trojans, worms, viruses, rootkit, spyware, and adware, have affected more than 100 million devices in over 200 countries in 2017 in a trend that is on the rise (Malwarebytes, 2018). Negative effects from malware stretch beyond their financial costs to include stolen personal identifiable information, business secrets or intellectual property, even with the possibility of leading to deaths and injuries from CPS disruption (Redondo & Insua, 2019).

In order to explore policies to manage zero-day malware attacks, this paper proposes a simulation model based on the System Dynamics method (SD) taking a Local-Area-Network (LAN) as the unit of analysis.

This paper presents and analyzes the proposed model sequentially as follows. Section 2 sets the context for this work with respect to previous efforts to model the spread of malware and their operational consequences by using epidemiological models. Section 3 develops and describes the model's structure. Section 4 analyses the model validity while Section 5 presents and analyses the results from running the model. Finally, Section 7 discusses the findings from the analysis section, and outlines areas of future work derived from these findings.

2 Theoretical background

A Zero-day vulnerability is technically defined as *a security flaw that has not yet been disclosed to the vendor or developers* (Sood & Enbody, 2012). Malware that successfully take advantage of a zero-day vulnerability are called zero-day malware.

As it has been argued that it is improbable that all security flaws will have been removed from complex systems before implementation, attackers spend substantial time and resources for the discovery and exploitation of these zero-day vulnerabilities through zero-day malware, making these even available in the worldwide market (Miller, 2007). This results in an "arms race" between attackers and solution developers.

As illustrated in Figure 1, the model that has been chosen for this paper is an adversarial model between attackers who activate zero day malware and solution developers who propose solutions to these zero-day malware. Solution developers include software developers but also consider any role and activity that actively works towards solving the zero-day malware attack, such as first-responder teams and IT engineers. The attackers do not only include external attackers, but also internal personnel for example that either plan or stumble upon the existence of a zero-day vulnerability and activates it.

When a zero-day malware is detected, a process of analysis is followed by developers within but also outside the affected companies, in a race to release a solution as soon as possible. External solution typically consist of software patches or upgrades, while internal solutions may include network reconfiguration, hardware update, and the training of users (Horowitz, 2019).

Models have been used to understand the cyber resilience of systems when exposed to malware cyber-attacks.

Lelarge (Lelarge, 2009) describes the modelling of interconnected agents subject to epidemic risks such as the ones caused by computer viruses and worms, through the use of graph theory. Lelarge argues that network externalities have both a public and a private part, relationship that results in the identification of counter-intuitive phenomena between the public protection policies and private protection initiative.

When there is a decrease in the public incentive to invest in self-protection, the fraction of the population investing in self-protection increases. On the other extreme, in a situation where the public protection ensures that the agent cannot be harmed by the decision of others, the results highlight the appearance of the free-rider problem. The middle ground is found where the public protection is weaker, situation in which the network can exhibits critical mass behavior. Lelarge also goes to show that when security against attacks is provided by a monopolist, the monopolist takes advantage of positive network externalities to providing a low quality protection.

In 2013, Schramm (Schramm & Gaver, 2013) used epidemiological models together with the Lancaster model of armed conflict to understand the propagation of cyber attacks through a network within the context of military operations. This research uses the `SIR` Model (Susceptible- Infected- Recovered)(Kermack & McKendrick, 1927), and a Kermack-McKendrick infection process, dependent and proportional to the population of `S` and `I`. The Lancaster model provided the dynamics through which the contact between `S` and `I` took place. This contact could either be through an *Aimed fire*, an *Area Fire* or a *Mixed effects* process.

In 2015, Xu et al., (Xu, Da, & Xu, 2015) explored the question of the correct application conditions for cyber-epidemic models. The model used is a unidirected finite network. This paper systematically explores the assumptions in models (dependencies), covering the effect on epidemiological models of aspects such as equilibrium thresholds, Equilibrium infection probabilities, network topology, and the effect of disregarding these assumptions.

Cisotto et al., (Cisotto & Badia, 2016) proposed the application of the `SIR` epidemic model to cellular automata, evaluating aspects such as agent mobility and grid connectivity in the spread of infections.

In 2016, Tran et al., (Tran, Campos-Nanez, Fomin, & Wasek, 2016) used an epidemiological approach to model the outbreak of zero-day attacks in a closed network, proposing what they named a cyber-resilience recovery model (CRRM). This research is to the best of our knowledge, the first to use a System Dynamics (SD) model to represent and simulate the *"Code Red"* worm infection i 2001, when 359.000 computers connected to the internet were infected in less than 14 hours. The model used the NIST-800-61 incident response framework. The SD model was based on the SIQR (Susceptible- Infected- Quarantined- Recovered)

-Yan et al., (Yan, Liu, Zhang, & Jia, 2019) presented in 2019 a simulation model that combined an epidemiological model with an individual defense model in what was described by them as a two-way dynamical framework. For the epidemiological model Yin used the `SIS` model (Susceptible- Infected- Susceptible) and a MArkov Decision Process for the individual defense model.

3 Dynamic model

The mathematical modeling of an epidemic is based on the epidemiological contagion model first proposed by Kermac and McKendrick (Kermack & McKendrick, 1927), and later developed by Sterman (Sterman, 2010). System Dynamics models with an epidemiological approach have been used extensively in medical and public health research to study decision control (Flessa, 1999) (van Ackere & Schulz, 2019), trend projection (Huang, Lin, Chen, Huang, & Wu, 2013) and the identification of opportunities and new directions of research (Homer & Hirsch, 2006). However, we have found no applications of system dynamics to the research of the contagion of information systems through cyber attacks. Models for cyber attacks that have been proposed focus on meta-features of an attack and a chain-of-action that lead to an attack, known as *kill-chains* and attack graphs (Al-Mohannadi et al., 2016), Bayesian networks (Comert, Pollard, Nicol, Palani, & Vignesh, 2018), big data analysis (Ju, Guo, Ye, Li, & Ma, 2019), Markov processes (Lalropuia & Gupta, 2019) and Stochastic Networks (Kotenko, Saenko, & Lauta, 2019).

The epidemiological model approach proposes the detail analysis of the role of infected individuals in order to understand the spread of an infection in a population. This article extends the use of the epidemiological approach to the case of information systems (further called *Systems*). The structure of this model is constructed from the following sub-structures:

1. An *epidemiological contagion sub-structure*, to represent the process where a system that has been compromised by a cyber-attack, comes into "contact" with an uncompromised system, thus exposing it to the probability of contagion,

2. A *feature-building-and-securing sub-structure*, to represent the system features that are transformed into vulnerabilities by the work of attackers, and then turned back into safe features by solution developers,

3. A *systems recovery sub-structure*, to represent the process through which compromised systems recover back to an uncompromised state, and

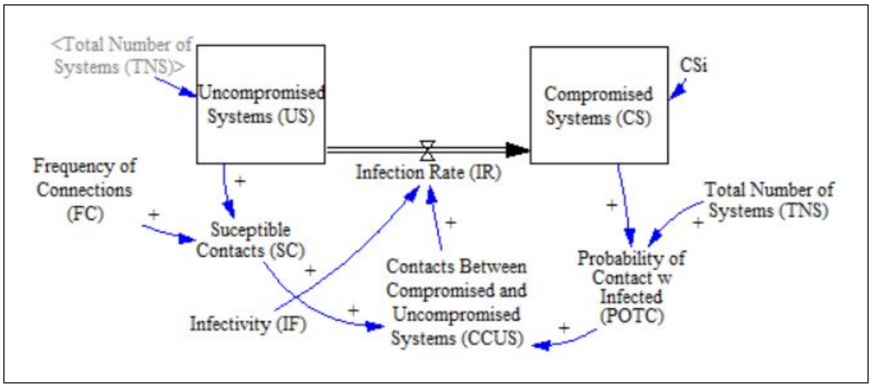
4. A *variable resource-building sub-structure*, to represent the resources that increase or decrease according to the number of compromised and uncompromised systems, in the form of the attackers that work to compromise the system and the solution developers that work to solve the cyber attacks and restore compromised systems.

5. *Disruption substructures*, to represent the way in which the system is changed in a controlled way in order to analyze its response.

These sub-structures have a causal connections between them. For example, the contagion substructure depends on the infectivity level of the system features, and the level of the resources (attackers and solution developers) are directly dependent on the proportion of compromised systems. The sub-structures in the model and their connections are described next.

3.1 Epidemiological contagion sub-structure

The epidemiological sub-structure reflects the contagion process that happens when a compromised system (CS) comes into contact with an uncompromised system (US). An epidemiological model considers that some of those contacts will result in an infection of an US through an infection rate (IR).



<Total Number of Systems (TNS)>
Uncompromised Systems (US)
Compromised Systems (CS)
CSi
Infection Rate (IR)
Frequency of Connections (FC)
Suceptible Contacts (SC)
Total Number of Systems (TNS)
Probability of Contact w Infected (POTC)
Contacts Between Compromised and Uncompromised Systems (CCUS)
Infectivity (IF)
+
+
+
+
+
+
+
+

Figure 2: Epidemiological contagion sub-structure

The contacts between a compromised and an uncompromised system (CCUS) will depend on the number of susceptible contacts (SC) and the probability of a contact with a infected system (POTC). SC is defined as the frequency of connections (FC) undertaken by US and the POTC is defined as the probability of meeting a CS from the total number of systems (TNS) in the population. The IR is determined both by the CCUS and the infectivity (IF) of the compromised systems. Figure 2 represents the epidemiological contagion substructure for this model, following equations (1) through (6).

$$US = \int_0^t (RR - IR)dx \tag{1}$$

$$CS = \int_0^t (IR - RR)dx \tag{2}$$

$$IR = CCUS * IF \tag{3}$$

$$POTC = CS/TNS \tag{4}$$

$$SC = US * FC \tag{5}$$

$$CCUS = POTC * SC \tag{6}$$

In this model, the initial value for US (US$_{\mathtt{i}}$) is TNS minus the initial number of compromised systems (CS$_{\mathtt{i}}$). The value of CS$_{\mathtt{i}}$ has to be at least 1 for an epidemiological model to have sense.

$$US_{\mathrm{i}} = TNS - CS_{\mathrm{i}} \tag{7}$$

Both FC and TNS are assumed to be an exogenous variables. FC is a characteristic of the systems, reflecting the number of connections per day that are performed on average by a system, either CS or US. On the other hand, TNS are the total number of systems in the population, whose growth is not relevant for the purposes of the model, and thus considered as a constant.

Finally the IF is not exogenous or constant, but is a variable that reflects the capacity of a cyber-attack to compromise systems. The value of IF is a result of the number of features in existing systems and the capacity of systems of being secured against cyber-attacks, features of the model discussed in the next sub-structure, concerned with feature-building and system securing.

3.2 Feature building and securing sub-structure

Systems that are susceptible to cyber-attacks, have a finite number of features that are either safe or unsafe, and at the same time known or unknown. The features relevant to this model are three, the safe features (SF), the known vulnerable features (KVF) and the unknown vulnerable features (UVF). These features follow a progression from being unsafe and unknown to being unsafe and known, to finally being safe and known. This is modeled as a stock and flow structure with conversion rates from UVF to KVF to finally SF.

The evolution of a hazardous system feature starts by a vulnerability activation (VC), process through which a previously safe set of features acquires a vulnerability through adversarial activation by attackers. These vulnerable features are at some point identified through a process of vulnerable feature discovery (VFD), and eventually a vulnerable feature solution (VFS) is generated. The average time for each of these processes are represented by the vulnerability creation time (VCT), the discovery time (DT), and solution time (ST).

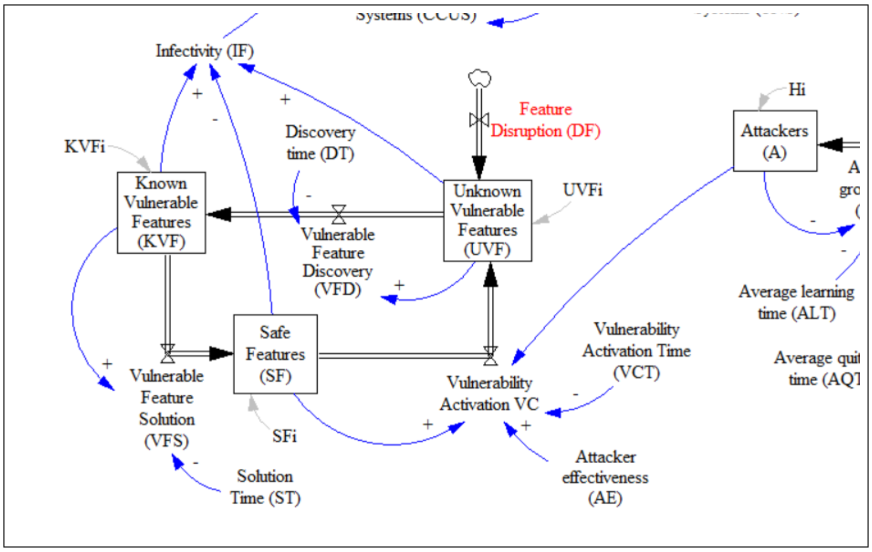


Figure 3: Feature building and securing sub-structure

Figure 3 represents the feature-building and securing sub-structure as proposed by this model, following equations (8) through (12).

$$UVF = \int_0^t (VC - VDF)dx \tag{8}$$

$$KVF = \int_0^t (VDF - VFS)dx \tag{9}$$

$$SF = \int_0^t (VFS - VC)dx \tag{10}$$

$$VDF = UVF/DT \tag{11}$$

$$VFS = KVF/ST \tag{12}$$

The feature building and securing sub-structure is connected to the epidemiological contagion sub-structure through the infectivity variable. In our model, the infectivity is represented as the proportion of the features that are unsafe from the total number of features, as is shown in equation (13).

$$IF = (UVF + KVF)/(SF + UVF + KVF) \tag{13}$$

Also, the `VC` is dependent on the existing features that are considered safe (`SF`), the effectiveness of attackers in the creation of vulnerabilities (`AE`), the vulnerability creation time (`VCT`), and the number of Attackers (`A`), as is shown in equation (14).

$$VC = (AE * SF * A)/VCT \tag{14}$$

The initial values for `SF`, `UVF`, `KVF`, are considered according the conditions of the systems that are being analyzed, and for this model, the total number of features in constant over time, as is shown in equation (15).

$$SF_{\mathrm{i}} + SF_{\mathrm{i}} + SF_{\mathrm{i}} = SF_{\mathrm{j}} + SF_{\mathrm{j}} + SF_{\mathrm{i}}, \forall i, j \in [0, t] \tag{15}$$

3.3 Systems recovery sub-structure

The systems recovery sub-structure represents the way in which `CS` recover to `US`. This process, represented through the Recovery Rate flow (`RR`) from `CS` to `US` is mediated by the number of existing Solution Developers (`D`), by the Developer Effectiveness (`DE`), the Solution Adoption Time (`SAT`), and the number of compromised systems `CS` as indicated by Equation 16.

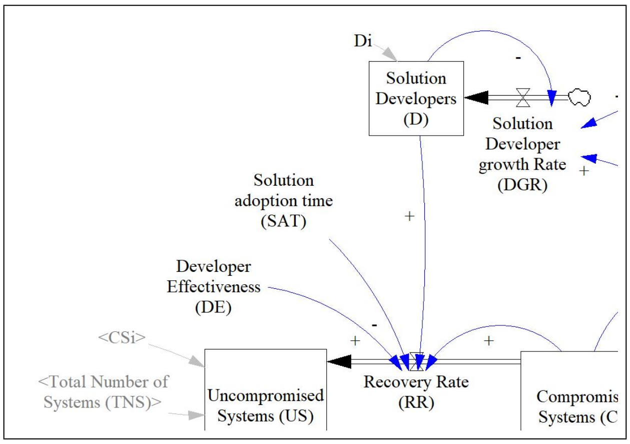

Di
Solution Developers (D)
-
Solution Developer growth Rate (DGR)
+
Solution adoption time (SAT)
+
Developer Effectiveness (DE)
-
<CSi>
+
+
<Total Number of Systems (TNS)>
Uncompromised Systems (US)
Recovery Rate (RR)
Compromis Systems (C

Figure 4: System recovery sub-structure

$$RR = (CS - SAT) * DE * D \tag{16}$$

Figure 4 represents the systems recovery substructure as proposed by this model, following equation 16. `SAT` and `DE` are considered as exogenous variables to the model, and their values are indicated in Table 3.

3.4 Variable resource building sub-structure

The variable resource building for this model is both the number of Developers (`D`) and number of Attackers (`A`) that exist in the system. By contrast the other resources in the system, this is the Systems (`CS` and `US`) and the system features (`UVF`, `KVF` and `SF`) are considered constant for this model, and only vary in a cycle between different states.

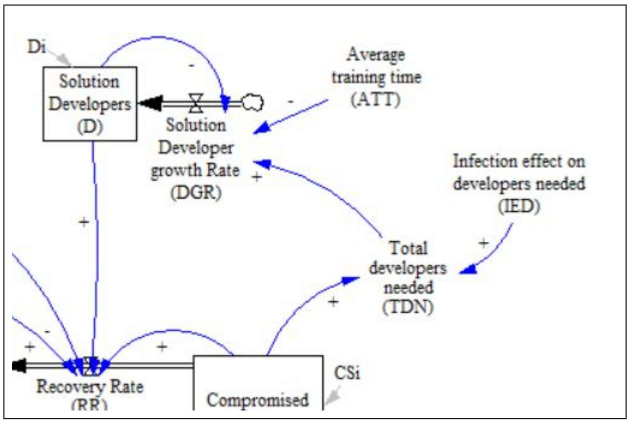

Di
Solution Developers (D)
-
Solution Developer growth Rate (DGR)
Average training time (ATT)
-
+
Infection effect on developers needed (IED)
+
Total developers needed (TDN)
+
+
-
+
+
+
Recovery Rate (RR)
Compromised
CSi

Figure 5: Developer resource building sub-structure

Additionally, D and H do not vary instantaneously as they are dependent on adjustment times, giving as a result inertia to the system. H and D are brought into the system or taken out of it according to the structural incentives that exist in the system.

D is influenced by CS, and the degree of this influence is represented by the variable Infection Effect on the Developers needed (IED). This puts a maximum on D, the Total number of Developers Needed (TDN). As this is not an instantaneous process, the time that takes D to adjust to TDN is represented by a first order delay, with time constant Average Training Time (ATT), and through the Solution Developer growth rate variable (DGR).

Figures 5 represents the Solution Developer system resource building sub-structure as proposed by this model, following equations (17) through (19).

$$D = \int_0^t (DGR)dx \tag{17}$$

$$DGR = (TDN - D)/ATT \tag{18}$$

$$TDN = MAX(0, CS * IED) \tag{19}$$

The initial value for D, D_i, is assumed to be zero as shown in Table 2, and for this model, ATT and IED are considered as exogenous variables.

A similar structure is considered for the number of Attackers in the system (A). The Total number of Potential Attackers (TPA) is determined by the variable Infection Effect on Attacker numbers (TPA), and this maximum number is realized over time through a first-order delay.

A difference with the structure for Developer resource building is that the Average Learning Time (ALT), this is, the adjustment time considered when A is lower than TPA as per equation 20, is different from the adjustment time considered when A is higher than TPA, the Average Quitting Time (AQT) as shown in equation 21. Both of these equations determine the Attacker Growth Rate (AGR).

If $(TPA - A) > 0$ then

$$AGR = (TPA - A)/ALT \tag{20}$$

otherwise if $(TPA - A) <= 0$

$$AGR = (TPA - A)/AQT \tag{21}$$

Figures 6 represents the Attacker system resource building sub-structure as proposed by this model, following equations (22) through (23).

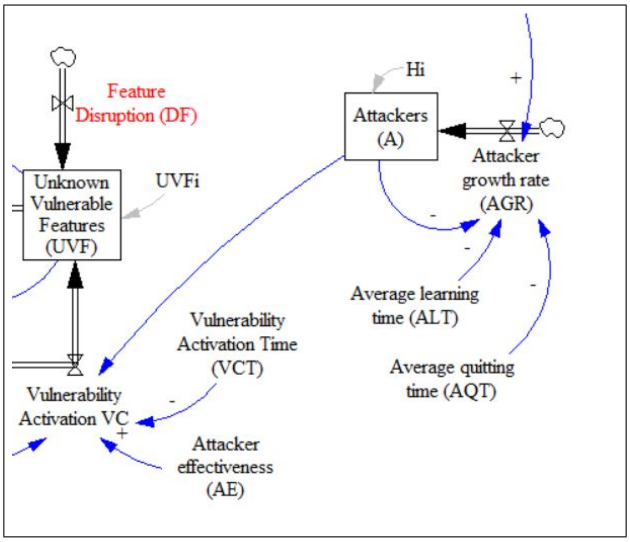
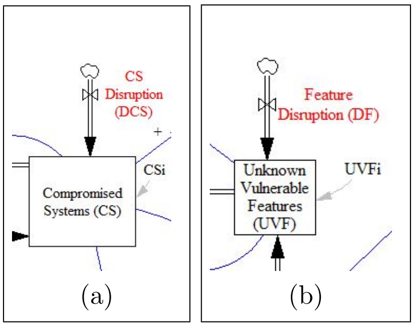


Figure 6: Model structure for resource building of Attackers



(a)

(b)

Figure 7: Model disruption sub-structures

$$A = \int_0^t (AGR)dx \tag{22}$$

$$TPA = MAX(0, CS * IEA) \tag{23}$$

$$\tag{24}$$

The initial value for A, H_i, is assumed to be zero as shown in Table 2, and for this model, ALT, AQT and IEH are considered as exogenous variables.

3.5 Disruption sub-structures

The last sub-structures in the model is the disruption sub-structure. This substructure is required in order to stimulate the system, understanding stimulation as a controlled change in the operating conditions of the system in order to analyze its behavior. The stimulation is created in the model in two ways. First, the model is stimulated through the disruption of the compromised systems (DCS), by adding a number CS_{stim} of compromised systems at the time time t_{CSstim}, as per equation 25. This is represented in the model by a flow into the CS Stock, as represented in Figure 7(a).

Table 2: Initial stock values

Stock Name	Init.Value	Unit
Compromised systems ($\mathrm{CS_i}$)	0	[System]
Attackers ($\mathrm{H_i}$)	0	[Person]
Known vulnerable features ($\mathrm{KVF_i}$)	0	[Feature]
Safe Features ($\mathrm{SF_i}$)	2000	[Feature]
Solution Developers ($\mathrm{D_i}$)	0	[Person]
Unknown Vulnerable Features ($\mathrm{UVF_i}$)	0	[Feature]

Table 3: Values for exogenous variables

Exogenous Variable	Value	Unit
Average learning time (ALT)	30	[day]
Average Training Time (ATT)	120	[day]
Average quitting time (AQT)	360	[day]
Developer effectiveness (DE)	0.8	[1/Person]
Discovery time (DT)	50	[day]
Frequency of connections (FC)	1500	[1/day]
Attacker effectiveness (AE)	0.0005	[1/Person]
Infection effect on developers needed (IED)	1.5	[Person/System]
Infection effect on Attacker numbers (IEA)	0.45	[Person/System]
Solution adoption time (SAT)	60	[day]
Solution time (ST)	15	[day]
Total Number of systems (TNS)	5000	[System]
Vulnerability Activation time (VCT)	360	[day]

$$DCS(t) = \begin{cases} 0 & t < t_{\mathrm{CSstim}} \\ CS_{\mathrm{stim}} & t = t_{\mathrm{CSstim}} \\ 0 & t > t_{\mathrm{CSstim}} \end{cases} \tag{25}$$

Second, the model is stimulated through the disruption of the vulnerable features (DF), by adding a number $\mathrm{UVF_{stim}}$ of unidentified vulnerable features at the time time $\mathrm{t_{UVFstim}}$, as per equation 26. This is represented in the model by a flow into the UVF Stock, as represented in Figure 7(b).

$$DF(t) = \begin{cases} 0 & t < t_{\mathrm{UVFstim}} \\ UVF_{\mathrm{stim}} & t = t_{\mathrm{UVFstim}} \\ 0 & t > t_{\mathrm{UVFstim}} \end{cases} \tag{26}$$

The complete model is shown in Figure 8, Table 2 lists the initial value for all stock variables and Table 3 list the values for all exogenous variables considered in the model.

4 Model Validity

The validation of system dynamics models is defined as *their comparison to empirical reality for the purpose of corroborating or refuting the model* (Senge & Forrester, 1980). The validity of a model can be in general be divided into validity of structure and validity of behaviour.

Additionally, structural validity has a temporal precedence over behavioural validity (Qudrat-Ullah & Seong, 2010). For more information about the validation of system dynamics models, we recommend the work by Barlas about the philosophy of model validation (Barlas & Carpenter, 1990), and about the formal aspects of system dynamics model validation (Barlas, 1996).

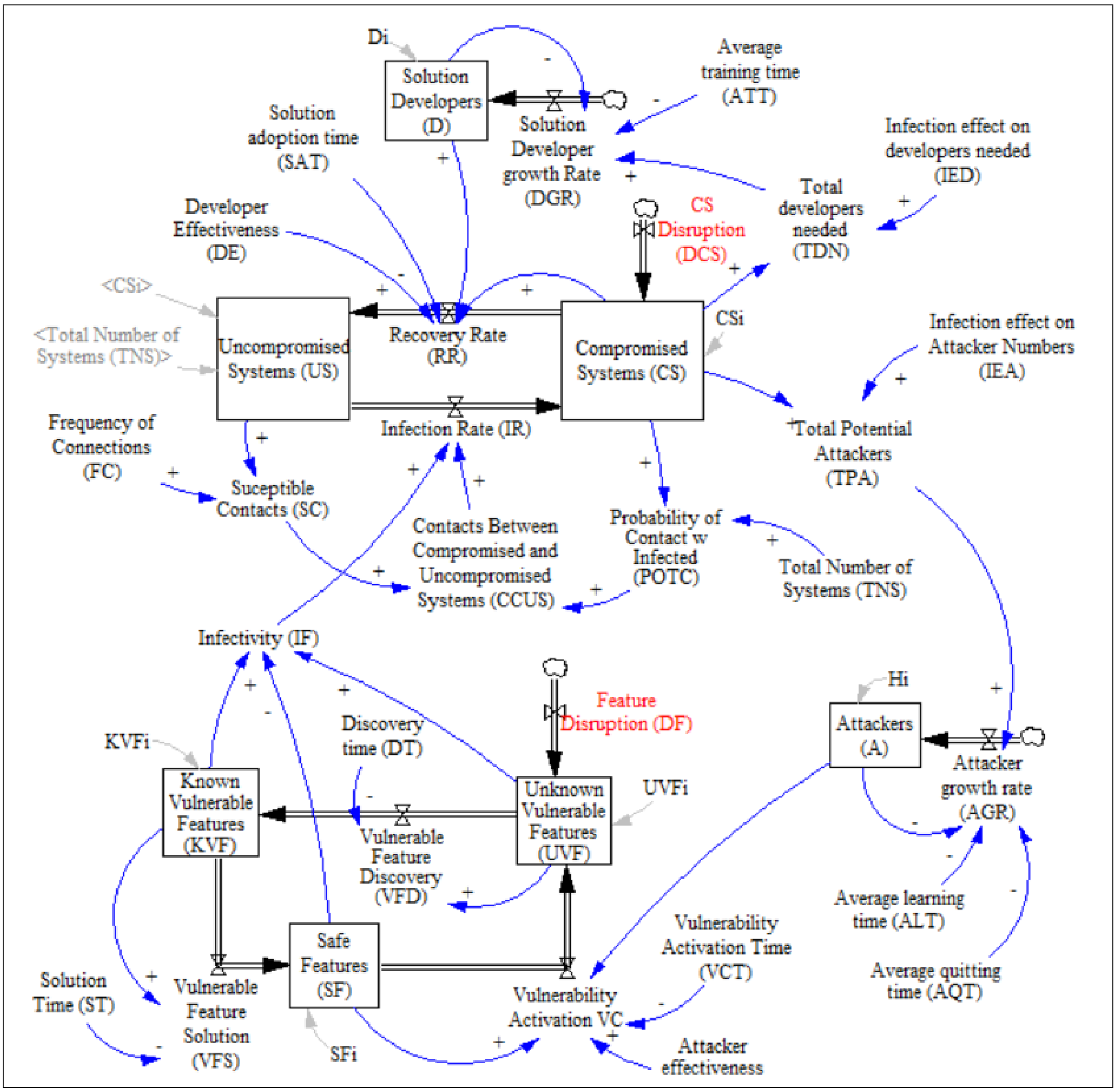

Di
Solution Developers (D)
Average training time (ATT)
Solution Developer growth Rate (DGR)
Solution adoption time (SAT)
Infection effect on developers needed (IED)
Total developers needed (TDN)
CS Disruption (DCS)
Developer Effectiveness (DE)
<CSi>
<Total Number of Systems (TNS)>
Uncompromised Systems (US)
Recovery Rate (RR)
Compromised Systems (CS)
CSi
Infection effect on Attacker Numbers (IEA)
Infection Rate (IR)
Total Potential Attackers (TPA)
Frequency of Connections (FC)
Suceptible Contacts (SC)
Contacts Between Compromised and Uncompromised Systems (CCUS)
Probability of Contact w Infected (POTC)
Total Number of Systems (TNS)
Infectivity (IF)
Hi
Feature Disruption (DF)
Discovery time (DT)
KVFi
Attackers (A)
Attacker growth rate (AGR)
Known Vulnerable Features (KVF)
Vulnerable Feature Discovery (VFD)
Unknown Vulnerable Features (UVF)
UVFi
Average learning time (ALT)
Safe Features (SF)
Vulnerability Activation Time (VCT)
Average quitting time (AQT)
Solution Time (ST)
Vulnerable Feature Solution (VFS)
Vulnerability Activation VC
SFi
Attacker effectiveness

Figure 8: The complete epidemiological model of Cyber resilience

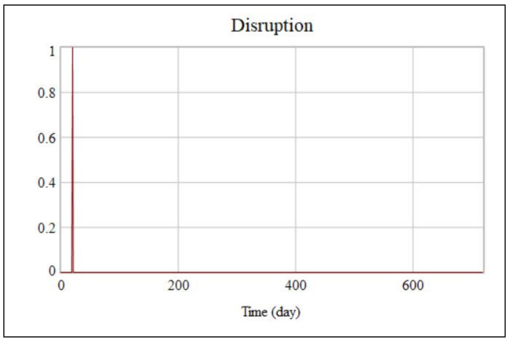


Figure 9: Disruption input for the model

The validation tests that have been carried out on the model presented in this paper address both the structural validity and the behavioural validity of the model.

–Part of the appeal of SD due to its ability to link the observable patterns of behaviour with an underlying structure. As a result SD creates causal models. Methods for structural validity of SD models have been proposed in literature. (Qudrat-Ullah & Seong, 2010).

The structural validity of the model was tested by performing Dimensional consistency, structure verification, boundary adequacy and the verification of some of the parameters with approximation of the other parameters

For example, the estimation of the persistence of the zero-day attack exploit, the Discovery Time DT, and the solution development once the exploit is found, the Vulnerable Feature Solution VFS, are obtained from industrial cases and reports (Ablon & Bogart, 2017)

5 Results

The model has been tested by changing the exogenous variables in the model, which can be either inputs to the model state, as described in subsection 3.5, or some of the model parameters as described in subsections 3.1, 3.2, 3.3 and 3.4. These are changed one at a time, to understand the influence of each of these changes have in the system behavior. The results for changing the model input is shown in Subsection 5.1, and the changes to different exogenous parameters are presented in Subsection 5.2.

5.1 Changes in model inputs - base and comparison simulations

The model is set up according to the structures described in Section 3, and following the exogenous variable values listed in Table 3 and the initial stock values listed in Table 2. The system is simulated over 1000 days, starting from a balanced to be disrupted by a single compromised system ($CS_{\text{stim}} = 1$) being introduced into the CS stock at $t_{\text{CSstim}} = 20$, as illustrated in Figure 9. This simulation does not disrupt UVF, and therefore considers $UVF_{\text{stim}} = 0, \forall t$. A number of relevant system states and variables that result from changing CS are shown in Figure 10. All of the graphs in Figure 10 reach equilibrium after over 1000 days.

The model is subsequently simulated under similar conditions, by using the second disruption structure, with $UVF_{\text{stim}} = 1$ and $t_{\text{UVFstim}} = 20$ also following the disruption pattern shown in Figure 9. The results are compared with the first disruption results and these are reflected in Figure 11.

5.2 Changes in model exogenous parameters

Some of the exogenous variables shown in Table 3 were also changed to understand the influence of these variables in the behaviour of the model. Three types of exogenous variables were

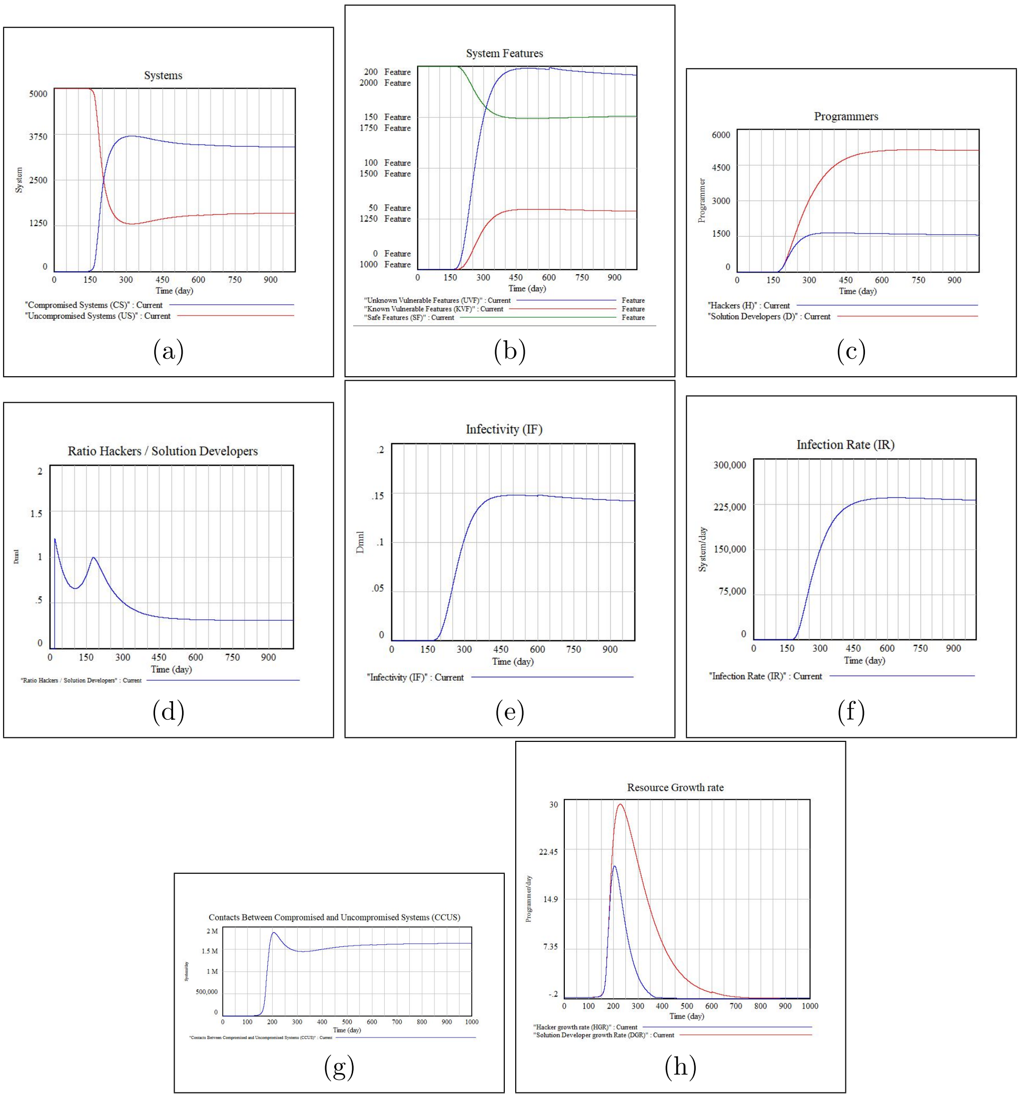


Figure 10: Outputs after changes to CS

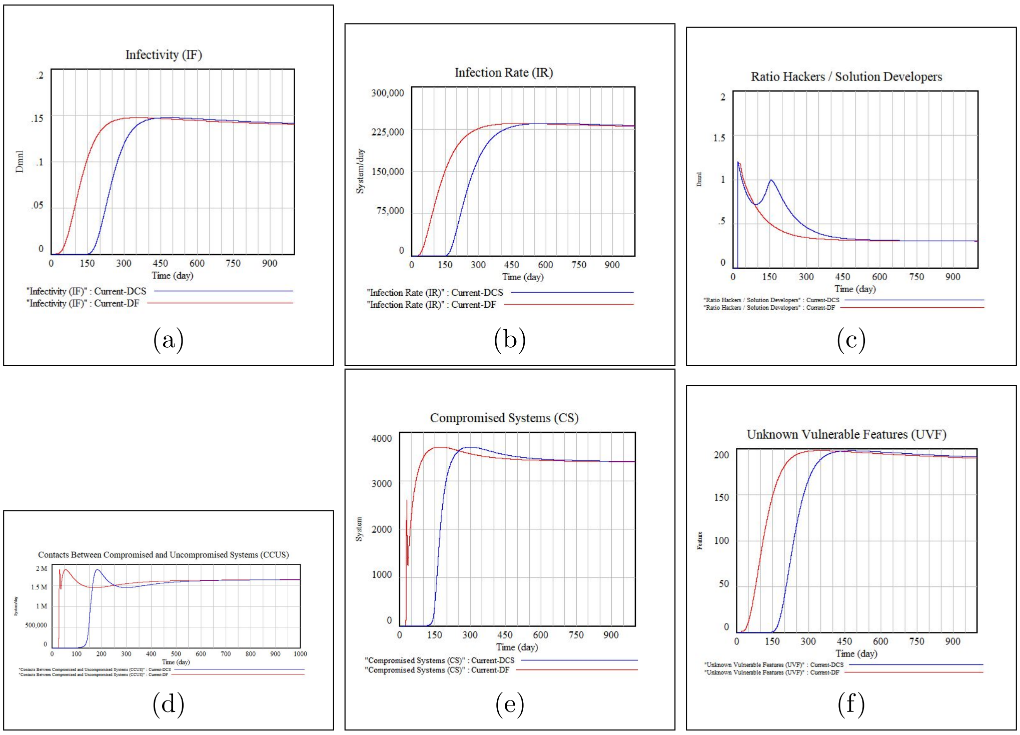


Figure 11: Outputs Comparison between DF and DCS Disruptions

Table 4: Values for exogenous variables for simulation

Variable name - `Acronym` [Unit]	**Min**	**Nominal**	**Max**
Average Training Time - `ATT` [day]	60	120	180
Solution Adoption Time - `SAT` [day]	30	60	90
Attacker Effectiveness - `AE` [1/Person]	0.00025	0.0005	0.00075

simulated, a training time represented by `ATT`, an adoption time represented by `SAT`, and an effectiveness index represented by `HE`. The values for which these exogenous variables were tested are shown in Table 4, and the ranges have been chosen around the nominal value, increasing by 50% as the maximum value, and 50% of the nominal value as the minimum value for the simulation. For comparison, all these simulations were carried out considering the same disruption, with $CS_{\text{stim}} = 1$ and $t_{\text{CSstim}} = 20$, and only one of the exogenous variables has been simulated at a time.

The results of the simulations are shown in Figures 12, 13 and 14.

6 Analysis

The dynamic behavior of this system shows that the model reaches an equilibrium state. In case of variations in the model inputs, the equilibrium levels that are obtained are the same, as seen in Figure 11. However, the time at which the dynamics develop are different, as seen for example in Figure 11b for `IR` or Figure 11d for the `CCUS`.

The growth rates of `D` goes on longer than `A`, as shown in Figure 10h, which ends up in a larger population of `D`, as shown in Figure 10c. As per the model parameters in Table 3, this occurs even though a) the training time for Attackers `ALT` is much smaller than the learning time for the solution developers `ATT`, and b) the reduction time for the Attacker population `AQT` is much larger than the reduction time for the solution developer population, also `ATT`.

The model reflects the pervasiveness and endurance of malware in the systems. This is a

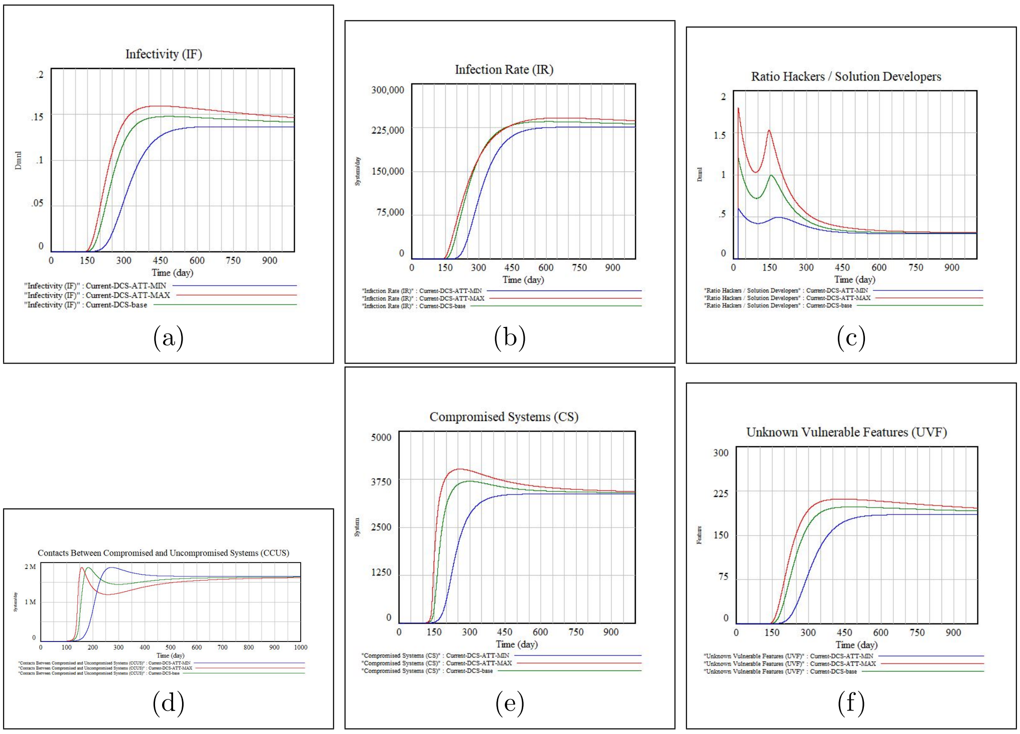


Figure 12: Simulation of ATT as per Table 4

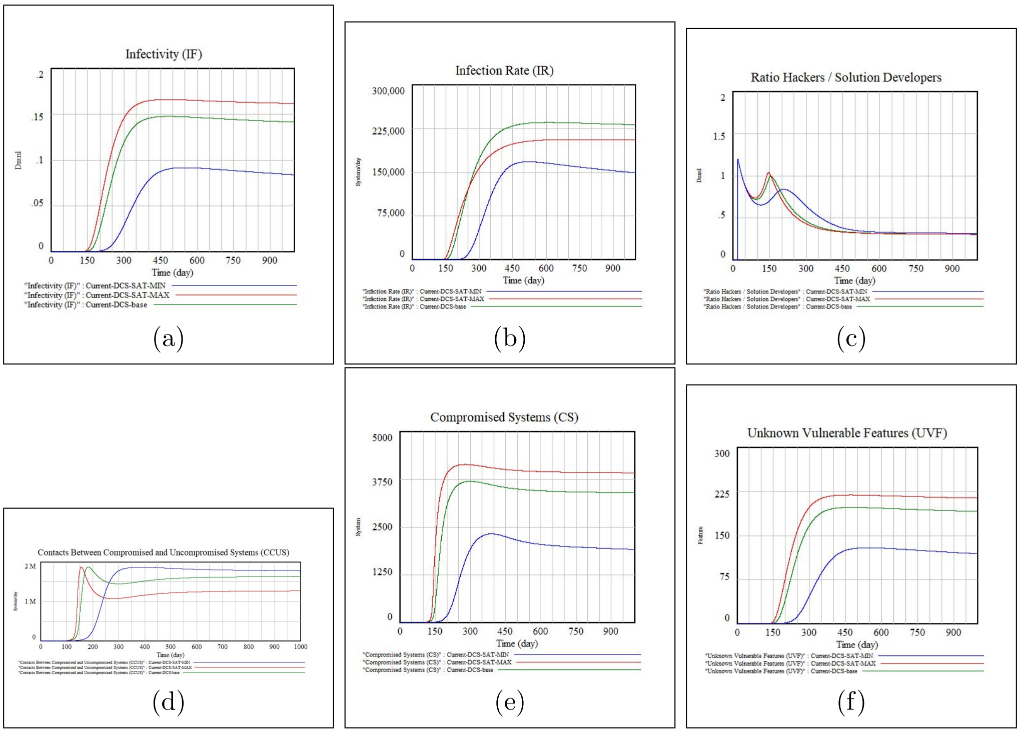
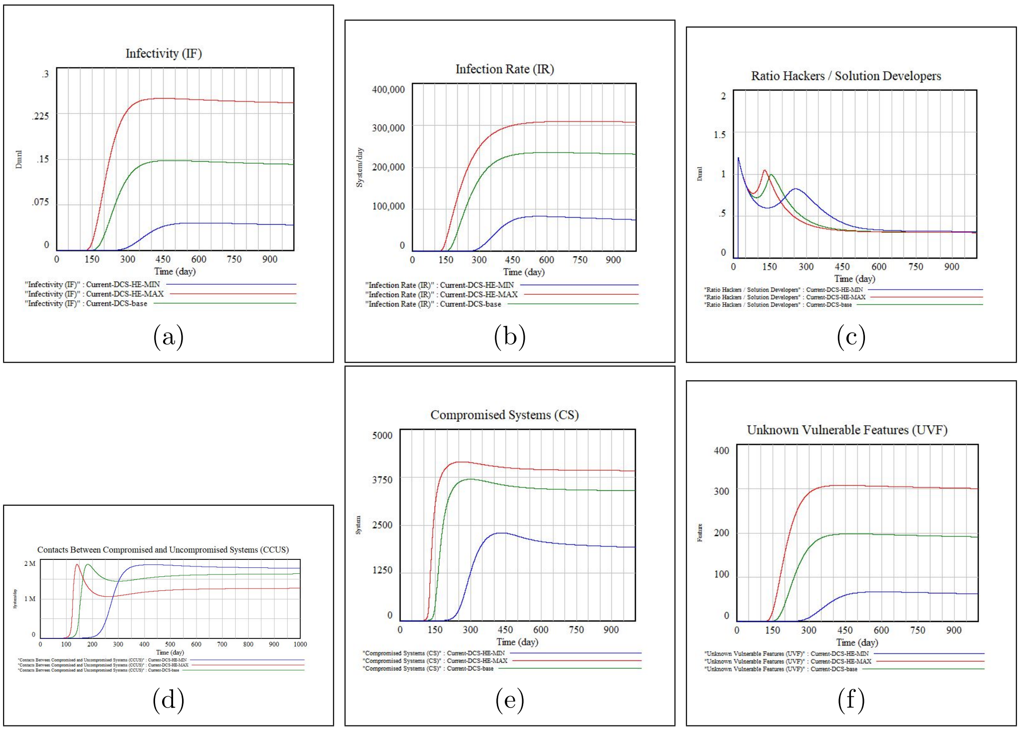


Figure 13: Simulation of SAT as per Table 4



(a)

(b)

(c)

(d)

(e)

(f)

Figure 14: Simulation of HE as per Table 4

result of the infections being a massive phenomenon and the solution implementation being a process which is necessarily localized. This is reflected in the model through the separation between the interaction between information system populations `US` and `CS`, from the population of system features `UVF`, `KVF` and `SF`. The result is that returning a system from `CS` to `US` does not necessarily mean returning a system feature from `KVF` to `SF`.

The model also reflects the latency state of cyber attacks through contagious vulnerable features. This is reflected in the model from an increase in `CS` does not start until some time later after the first infection. In the base simulation, the model experiences important changes after day 100 when as seen in Figure 10g `CCUS` first starts to increase, and is then followed by rapid increases in `IF` (Figure 10e), `IR` (Figure 10f), and `H` and `D` (Figure 10c).

The proportion of `CS` in equilibrium is over two times the number of `US`, as can be seen through Figure 10a, even though the number of `D` is more than 3 times the number of `A`, as can be seen in Figure 10c and 10d. This is reflects the pervasiveness of infectious computer attacks that can linger in the form of an ever present stock of `UVF`s in the system.

The Infection Rate `IR` balances at around 225,000 infections per day Figure 10f for a `TNS` of 5000, meaning that each system on average is exposed to 45 infections per day, out of the 1500 connections each of these systems make with other systems daily.

The values chosen for the model are subject to review and adjustment. Not all the variables have the same level of certainty in their values. However, the best approximation possible has been provided, taking into consideration that once the variable has been considered relevant for the causality of the model, then the only value we know to be wrong would be zero. This shifts the question towards finding scientific ways to approximate these values.

7 Conclusions and further work

The modeling and simulation described in this paper, explore the dynamics of cyber-attacks based on epidemiology dynamics as represented by one-to-one contagion between IT systems. System Dynamics is used as a way of modeling a simplified structure where these cyber-attacks occur, and to simulate the evolution over time of the system's behaviour, by representing the system's structure stocks (accumulations) and flows to identify and represent the causal connections between variables in the system.

The analysis process proposed in this paper makes a scientific contribution through the use of epidemiological mechanisms for the identification of important strategic-level variables and their causal connections, to represent and simulate IT systems, their features, their vulnerabilities, and the programmers that work either to create more vulnerabilities or to make these systems more secure.

Moreover, the process presented in this paper explicitly differentiates between variables that can be modified in the short term to influence the system's behavior, from those variables that require structural change for their modification and subsequent system influence. For example, the `SAT` can be quickly improved by choosing a different training program or screening the programmers that start their training, yet, in order to change the `POTC` for example, there a structure change that will require in a different probability. In the case presented in this paper, all the IT systems have the same probability of meeting each other, since as per Equation 4 the probability function is a constant. A different structure, where the connections between IT systems is organized around clusters, for example, would result in a different probability function, yet this is not actionable in the short term.

The controlled simulation of the model proposed in this paper provides a foundation based on structure for the relative influence of different variables in the system, either in amplitude, in phase or both. For example, similar changes in the number of unknown vulnerable features (`UVF`) or in the number of compromised systems (`CS`) create different outcomes as per Figure 11, with the former having a quicker effect in the model response pattern. However, both changes converge to the same equilibrium state.

All the assumptions presented in this paper can be analyzed and challenged in future work. Other types of cyber-attacks can be explored through dynamic models, or such a model could consider different network levels and hierarchies (CAN, IAN, WAN, etc).

The general model presented in this paper can be developed in greater detail to address more specific questions that will aid in the effective design of resilient response. Research about the infectivity of system features, about the shorter lead time for implementation of solutions or greater control about the cumber of susceptible contacts between compromised and uncompromised systems are all relevant questions that could be addressed through the use of simulation to identify investment effectiveness.

Additionally, the system dynamics model can be further improved by obtaining better approximations to the variables of the model for which there are multiple layers of causality that contribute to its value. An example of this is the solution adoption time `SAT`. Multiple levels of adoption are found in practice, such as machine isolation, software patches, retraining o infrastructure improvement, for example.

References

Ablon, L., & Bogart, A. (2017). *Zero days, thousands of nights: The life and times of zero-day vulnerabilities and their exploits.* Rand Corporation.

Alieyan, K., Almomani, A., Anbar, M., Alauthman, M., Abdullah, R., & Gupta, B. (2019). Dns rule-based schema to botnet detection. *Enterprise Information Systems*, 1–20.

Al-Mohannadi, H., Mirza, Q., Namanya, A., Awan, I., Cullen, A., & Disso, J. (2016). Cyber-attack modeling analysis techniques: An overview. In *2016 ieee 4th international conference on future internet of things and cloud workshops (ficloudw)* (pp. 69–76).

Baldwin, A., Gheyas, I., Ioannidis, C., Pym, D., & Williams, J. (2017). Contagion in cyber security attacks. *Journal of the Operational Research Society*, *68*(7), 780–791.

Barlas, Y. (1996). Formal aspects of model validity and validation in system dynamics. *System Dynamics Review: The Journal of the System Dynamics Society*, *12*(3), 183–210.

Barlas, Y., & Carpenter, S. (1990). Philosophical roots of model validation: two paradigms. *System Dynamics Review*, *6*(2), 148–166.

Barroso, A., Machado, V., Carvalho, H., & Machado, V. C. (2015). Quantifying the supply chain resilience. *Applications of contemporary management approaches in supply chains*, 13–32.

Cisotto, G., & Badia, L. (2016). Cyber security of smart grids modeled through epidemic models in cellular automata. In *2016 ieee 17th international symposium on a world of wireless, mobile and multimedia networks (wowmom)* (pp. 1–6).

Comert, G., Pollard, J., Nicol, D. M., Palani, K., & Vignesh, B. (2018). Modeling cyber attacks at intelligent traffic signals. *Transportation research record*, *2672*(1), 76–89.

Conti, M., Dragoni, N., & Lesyk, V. (2016). A survey of man in the middle attacks. *IEEE Communications Surveys & Tutorials*, *18*(3), 2027–2051.

De Donno, M., Giaretta, A., Dragoni, N., & Spognardi, A. (2017). A taxonomy of distributed denial of service attacks. In *2017 international conference on information society (i-society)* (pp. 100–107).

Fagioli, A. (2019). Zero-day recovery: the key to mitigating the ransomware threat. *Computer Fraud & Security*, *2019*(1), 6–9.

Flessa, S. (1999). Decision support for malaria-control programmes–a system dynamics model. *Health Care Management Science*, *2*(3), 181–191.

Grégio, A. R. A., Afonso, V. M., Filho, D. S. F., Geus, P. L. d., & Jino, M. (2015). Toward a taxonomy of malware behaviors. *The Computer Journal*, *58*(10), 2758–2777.

Gupta, B. B., Arachchilage, N. A., & Psannis, K. E. (2018). Defending against phishing attacks: taxonomy of methods, current issues and future directions. *Telecommunication Systems*, *67*(2), 247–267.

Homer, J. B., & Hirsch, G. B. (2006). System dynamics modeling for public health: background and opportunities. *American journal of public health*, *96*(3), 452–458.

Horowitz, B. M. (2019). Policy issues regarding implementations of cyber attack: Resilience solutions for cyber physical systems. In *Artificial intelligence for the internet of everything* (pp. 87–100). Elsevier.

Huang, S.-K., Lin, M.-T., Chen, H.-C., Huang, S.-C., & Wu, M.-H. (2013). Epidemiology of kawasaki disease: prevalence from national database and future trends projection by system dynamics modeling. *The Journal of pediatrics*, *163*(1), 126–131.

Ju, A., Guo, Y., Ye, Z., Li, T., & Ma, J. (2019). Hetemsd: A big data analytics framework for targeted cyber-attacks detection using heterogeneous multisource data. *Security and Communication Networks*, *2019*.

Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of London. Series A, Containing papers of a mathematical and physical character*, *115*(772), 700–721.

Khan, O., & Sepulveda Estay, D. A. (2015). Supply chain cyber-resilience: Creating an agenda for future research. *Technology Innovation Management Review*(April), 6–12.

Kotenko, I., Saenko, I., & Lauta, O. (2019). Modeling the impact of cyber attacks. In *Cyber resilience of systems and networks* (pp. 135–169). Springer.

Lalropuia, K., & Gupta, V. (2019). Modeling cyber-physical attacks based on stochastic game and markov processes. *Reliability Engineering & System Safety*, *181*, 28–37.

Lelarge, M. (2009). Economics of malware: Epidemic risks model, network externalities and incentives. In *2009 47th annual allerton conference on communication, control, and computing (allerton)* (pp. 1353–1360).

Leslie, N. O., Harang, R. E., Knachel, L. P., & Kott, A. (2018). Statistical models for the number of successful cyber intrusions. *The Journal of Defense Modeling and Simulation*, *15*(1), 49–63.

Malwarebytes. (2018). *State of malware 2017.* Retrieved 2010-08-25, from `https://www.malwarebytes.com/pdf/infographics/stateofmalwareinfographic.pdf`

McWhirter, P. R., Kifayat, K., Shi, Q., & Askwith, B. (2018). Sql injection attack classification through the feature extraction of sql query strings using a gap-weighted string subsequence kernel. *Journal of information security and applications*, *40*, 199–216.

Miller, C. (2007). The legitimate vulnerability market: the secretive world of 0-day exploit sales. In *Weis.*

Munoz, A., & Dunbar, M. (2015). On the quantification of operational supply chain resilience. *International journal of production research*, *53*(22), 6736–6751.

Pienta, D., Thatcher, J. B., & Johnston, A. C. (2018). A taxonomy of phishing: Attack types spanning economic, temporal, breadth, and target boundaries. In *Proceedings of the 13th pre-icis workshop on information security and privacy, san francisco, ca, usa* (Vol. 1).

Qudrat-Ullah, H., & Seong, B. S. (2010). How to do structural validity of a system dynamics type simulation model: The case of an energy policy model. *Energy policy*, *38*(5), 2216–2224.

Redondo, A., & Insua, D. R. (2019). Protecting from malware obfuscation attacks through adversarial risk analysis. *arXiv preprint arXiv:1911.03653*.

Schramm, H. C., & Gaver, D. P. (2013). Lanchester for cyber: The mixed epidemic-combat model. *Naval Research Logistics (NRL)*, *60*(7), 599–605.

Senge, P. M., & Forrester, J. W. (1980). Tests for building confidence in system dynamics models. *System dynamics, TIMS studies in management sciences*, *14*, 209–228.

Sepulveda Estay, D. A., & Khan, O. (2015). Extending supply chain risk and resilience frameworks to manage cyber risk.. (22nd EurOMA Conference : Operations Management for Sustainable Competitiveness ; Conference date: 26-06-2015 Through 01-07-2015)

Sfakianakis, A., Douligeris, C., Marinos, L., Lourenço, M., & Raghimi, O. (2019). Enisa threat landscape report 2018: 15 top cyberthreats and trends. *DOI*, *10*, 622757.

Sheffi, Y., & Rice Jr, J. B. (2005). A supply chain view of the resilient enterprise. *MIT Sloan management review*, *47*(1), 41.

Soliman, S. W., Sobh, M. A., & Bahaa-Eldin, A. M. (2017). Taxonomy of malware analysis in the iot. In *2017 12th international conference on computer engineering and systems (icces)* (pp. 519–529).

Som, S., Sinha, S., & Kataria, R. (2016). Study on sql injection attacks: Mode detection and prevention. *International Journal of Engineering Applied Sciences and Technology, Indexed in Google Scholar, ISI etc., Impact Factor: 1.494*, *1*(8), 23–29.

Sood, A. K., & Enbody, R. J. (2012). Targeted cyberattacks: a superset of advanced persistent threats. *IEEE security & privacy*, *11*(1), 54–61.

Specht, S., & Lee, R. (2003). Taxonomies of distributed denial of service networks, attacks, tools and countermeasures. *CEL2003-03, Princeton University, Princeton, NJ, USA*.

Sterman, J. (2010). *Business dynamics*. Irwin/McGraw-Hill c2000..

Tran, H., Campos-Nanez, E., Fomin, P., & Wasek, J. (2016). Cyber resilience recovery model to combat zero-day malware attacks. *computers & security*, *61*, 19–31.

van Ackere, A., & Schulz, P. J. (2019). Explaining vaccination decisions: A system dynamics model of the interaction between epidemiological and behavioural factors. *Socio-Economic Planning Sciences*, 100750.

Walters, R. (2015). Cyber attacks on us companies since november 2014. *The Heritage Foundation*(4487).

Xu, M., Da, G., & Xu, S. (2015). Cyber epidemic models with dependences. *Internet Mathematics*, *11*(1), 62–92.

Yan, D., Liu, F., Zhang, Y., & Jia, K. (2019). Dynamical model for individual defence against cyber epidemic attacks. *IET Information Security*, *13*(6), 541–551.