SAND2016-6291C

Refining the Foundations for Cyber Zone Defense

Robert Mitchell
Sandia National Laboratories
Albuquerque, NM 87185
Email: rrmitch@sandia.gov

Paul Sery
Sandia National Laboratories
Albuquerque, NM 87185
Email: pgsery@sandia.gov

***Abstract*—Since our last paper, cyber attacks have shown no evidence of declining in frequency or sophistication. We claim that applying isolation zones is an effective way to defend cyber systems; our team proposes a simulation and mathematical model that provide numerical data that supports this claim. This paper extends our earlier cyber zone defense (CZD) framework in two critical ways. First, we relax our assumption that zones completely isolate nodes and consider interzone boundaries to be porous. Second, we investigate methods to estimate one of the legacy parameters inherited from our earlier work and the new porosity parameter. The extended simulation and model more closely approximate real world cyber systems and have lower residuals than our previous investigation.**

I. Introduction

Our last paper [9] asserted that the ever increasing complexity of the cyber world provides the attacker with a strategic advantage over the defender. This is illustrated by the increasing capabilities and size of operating systems and applications, the proliferation of new devices like the smartphone and the Internet of things (IoT). Each interconnection and piece of software is created by humans and vulnerable to error. We maintain that relying on patching all the vulnerabilities in increasingly complex cyber systems and/or anticipating all zero-day exploits is an insufficient strategy. The cyber environment makes traditional defensive techniques fundamentally reactive; they must be supplemented with proactive approaches. Thus, we focus on intrusion tolerance as opposed to intrusion prevention, detection or response.

Much of cyber defense today focuses on protecting the end points that attackers break into and use to their advantage. Figure 1 illustrates the traditional cyber defense sequence where the defender first minimizes vulnerabilities, then attempts to detect attacks against inevitable defects and finally responds to attacks when detected. Each step becomes more difficult as systems become more complex, and the defender is never assured of detecting all or even most attacks, successful or not. This is a fundamentally reactive strategy and more art than science.

Cyber zone defense (CZD) adds tolerance in parallel to the traditional cyber defense sequence as shown in Figure 2. It borrows successful defensive tactics and strategies from the physical world. For instance, naval architects partition ships into survivable cells; infantry select advantageous terrain and place obstacles to channelize the adversary and mitigate their capabilities. These strategies are proactive by nature.

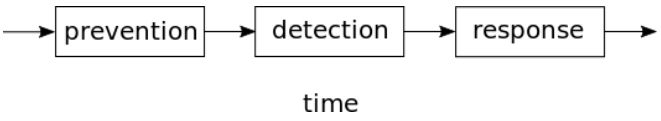

Fig. 1. Traditional cyber defense timeline.

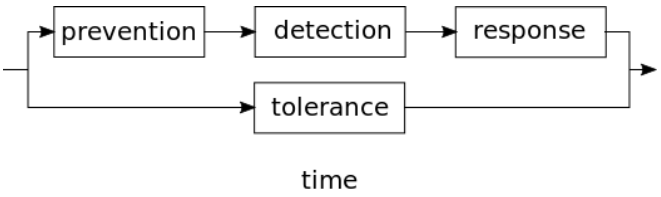

Fig. 2. Emerging cyber defense timeline.

Adapting successful physical strategies to the cyber world should increase cyber system resiliency and survivability by disrupting the reachback and lateral movement of malware. CZD organically partitions networks by device class or activity to isolate hosts from each other. We claim this effectively attenuates or, in some cases, eliminates the impact of a cyber attack without knowing how malware works or whether it even exists.

However, before applied researchers and practitioners proceed with technology transfer, prototyping and implementation, we are compelled to provide some prediction of the conditions under which this approach is viable. This article provides an approach to gain this insight.

In this study, we propose improvements to our prior work: Our previous simulation and model assumed a highly simplified configuration; we propose extensions that will provide a more realistic level of sophistication and fidelity. Also, space constraints prevented us from providing techniques to estimate all of the simulation and model parameters; here we provide multiple ways of calculating the parameters based on real data. The rest of this paper is organized as follows: First, Section II summarizes the state of the art of this topic. Next, Section III describes our extended simulation and model. Third, Section IV visualizes the numerical results from the

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. Approved for unlimited release: SAND2016-XXXX C.

simulation and model and interprets the trends the figures show. Following this, Section V proposes methods for applied researchers or practitioners to estimate the parameters in the simulation and mathematical model. Finally, Section VI discusses our conclusions and future research directions.

II. Literature Search

In our earlier work, we [9] introduced a new framework called CZD that treats malware like a black box and proposed a simulation technique and closed form mathematical model that predicted the performance of CZD.

Our earlier literature search [9] revealed that there was a gap in the literature with respect to predictive modeling applicable to CZD performance. Since that time, a few new contributions have appeared.

Many interesting papers studying the cyber security impact of zoning have been published in the interval between [9] and the present work. Some investigations [3], [2], [6], [5] and [11] have proposed cyber security metrics and modeling techniques. Two papers [12] and [7] researched supervisory control and data acquisition (SCADA) systems where zoning is already known to be the best practice. One paper [1] investigated detecting malware based on samples' reachback signature. Another paper [4] studied specific zoning techniques.

Holm, et al. [3] propose Predictive Probabilistic Cyber Security Modeling Language (P2CySeMoL). P2CySeMoL generates attack graphs based on models; these attack graphs can predict the security of the cyber systems under consideration. The authors' primary metric is probability of attack success. Holm, et al. also provide Pearson correlations between their approach and those of their contemporaries. Their threat model assumes the adversary has one week to complete their attack.

Gontarczyk, et al. [2] model the impact of zoning as applied to cloud computing. While our zones are collateral, the zones the authors propose are concentric. Specifically, they envision tiers comprising at least: unmanaged, externally managed and directly managed zones; the common thread within each tier is a shared level of trust. They propose a mapping of nineteen business application classifiers, technologies and vulnerabilities to zones. Gontarczyk, et al. do not provide any numerical data or propose any metrics for measuring the effectiveness of this allocation. The authors' threat model focuses on the insider threat; Gontarczyk's basic idea is to minimize the number of insiders for more sensitive subsystems.

Leuprecht, et al. [6] challenge what they call the Castle Model of cyber security in a thought experiment. The Castle Model includes antivirus applications, inward and outward facing firewalls, spam filters and authentication mechanisms. The authors argue that application of the Castle Model decreases the agility of an enterprise, and the sophistication required to overcome this obstacle creates even more vulnerabilities. They consider three alternatives to walls: attack graphs, real-time intrusion detection systems and privilege separation; however, Leuprecht, et al. say these will not do the job either. The authors endorse the computing in compromised environments metaphor which includes virtual castles, virtual software, behavioral modeling, secret sharing and decoys of sensitive data. This study did not include any numerical data or quantitative evidence. An important similarity between this work and ours is that we both advocate intrusion tolerance as the way forward as opposed to prevention, detection or response (cf. Figure 2).

Khan, et al. [5] discuss metrics related to cyber security by design. Specifically, the authors propose the Cyber Resilience Engineering Framework (CREF) and work through a case study where they begin with an unsecure system and shore up its weaknesses. The inputs to CREF are a network model, attacks and resilience metrics; its output is a resilience axis. Khan, et al. propose that network resilience metrics fall in to three categories: proactive, resistive and reactive. CREF seems able to accommodate an arbitrary attack model. Unfortunately, they don't provide a clear definition for a metric in any of the three categories and measure their results based on resilience in the abstract.

Sack and Ierache [11] study key cyber security controls and metrics. The authors intend to survey the literature, compile the existing security controls and metrics and unify them into a framework that can evaluate all systems on a level playing field. They organize controls by five phases of cyber defense and four iterations of cyber defense: Sack and Ierache's five cyber defense phases are: identify, protect, detect, respond and recover. The authors' four iterations are: quick wins, visibility and attribution measures, improved information security configuration and advanced subcontrols. This yields a five by four matrix of metric classes. They show how to calculate the score for one of these twenty classes. Our solution would fall into the protect phase and the improved information security configuration iteration of Sack and Ierache's framework. There are no numerical results in this article. The authors do not define their attack model, but their approach could be generic enough to accommodate anything reasonable.

Soltan, et al. [12] propose the Post Attack Recovery and Detection (PARD) algorithm for power grids which fall under the umbrella of SCADA systems. The authors demonstrate the ability to recover phase angles (important in the context of power grids), identify broken links, find graph classes for which these questions can be answered and partition a power grid into zones that can resist attacks. This works makes it clear that power grids are a unique cyber security application. In future work, we will pursue the identification of attack resistant zones in an IT enterprise.

Machii, et al. [7] claim zoning delivers two benefits: cyber attack isolation and increased attack detectability. The authors consider dynamic zoning for SCADA systems; in their application, zones change due to the state or mode of the system. Machii, et al. identify four system states and determine the connectivity required for each state. The authors use OpenFlow [8] to implement the zoning and developed OpenState to interface with this technology. Unfortunately, this

study does not include any numerical data. When we consider dynamic zones in future work, zones will change according to human collaboration patterns.

Celik, et al. [1] apply unsupervised machine learning to defeat malware reachback techniques. In particular, the authors use one class support vector machines, k-nearest neighbors and least squares anomaly detection and k-means clustering. They provide clear numerical results in the form of receiver operating characteristic plots. The threat model of Celik, et al. considers several malware families: Sality, Tbot, Spyeye, ZeusV1, ZeusGameOver, ZeusPonyLoader, ZeusV2, Kelihos Zeroaccess, Agobot, Donbot and Kaiten. The authors concede these families may not be representative of the population. While Celik, et al. propose an intrusion detection approach, our work focuses on intrusion tolerance.

One special case of a zone is a demilitarized zone (DMZ) which is a popular cyber security technique. Sunghyuck Hong [4] studies the viability of a DMZ that provides a shelter from domain name system (DNS) based distributed denial of service (DDoS) attacks. The author's metric is response time in seconds. Their threat model is limited to a DNS based DDoS attack. Whereas an outsider typically launches a DNS based DDoS attack, our attack model considers the insider threat and is generic to any attack scenario.

III. Simulation and Model

In our earlier work [9], we proposed two metrics: the probabilities of compromise (p_c) and reachback (p_r). p_c is the probability an adversary has implanted malware on a host. p_r is the probability an infected host can communicate with the adversary's command and control (C2) site.

In this work, we relax our assumption that zones are completely isolated. We add an additional parameter ϕ that measures the porosity of zone boundaries. We considered several physical phenomena to use as an analogy for malware crossing zone boundaries: these included electrical impedance, electrical resistance, permeability and porosity. Electrical impedance and resistance are closely related: The former measures a circuit's opposition to an alternating current (AC) while the latter measures the difficulty of passing a direct current (DC) through a conductor. The International System (SI) unit for both of these metrics is the ohm; unfortunately, ohms do not translate well into the cyber domain. Permeability measures a material's ability to pass fluids; its SI unit is m^2. While researchers frequently discuss the "attack surface" of a cyber system, equating a system vulnerability count with the square of physical distance is questionable. Porosity is related to permeability and measures the proportion of void in a material; it is unitless. This unitlessness of porosity is especially appealing.

We continue our approach from [9] where the math model forms the basis for and naturally lends itself to complex simulations.

A. Simulation

We revised the simulation from [9] to include ϕ. This simulation iteratively calculates the probability of compromise p_c^i for each host individually. The metric of interest is the overall probability of compromise p_c which is still the arithmetic mean of p_c^i. In addition to considering attackers moving laterally within a zone with probability $p_c \cdot p_e$, the updated simulation considers attackers moving laterally across zones with probability $p_c \cdot p_e \cdot \phi$.

p_r is still calculated as described in [9] and specified in Equation 1:

$$1 - \prod_i (1 - p_c^i) \tag{1}$$

where i covers the set of externally facing hosts.

B. Closed Form Mathematical Model

In addition to the simulation, we also revised the closed form mathematical model for p_c from [9] to include ϕ. In this way, in addition to considering the adversary's lateral movement within a zone with probability $\frac{z}{n} \cdot p_e$, the updated mathematical model considers adversarial lateral movement across zones with probability $\frac{n-z}{n} \cdot p_e \cdot \phi$. Equation 2 specifies the updated prediction for probability of compromise.

$$p_c = \frac{1 + p_e \cdot (\frac{z}{n} + \phi \cdot \frac{n-z}{n}) \cdot (n-1)}{n} \tag{2}$$

The calculation of p_r remains the same as in our prior work [9]; we include it here in Equation 3 for convenience.

$$p_r = 1 - (1 - p_c)^x \tag{3}$$

IV. Results

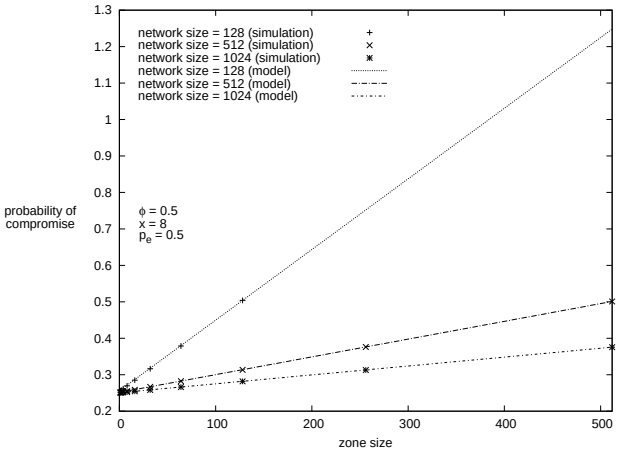

Fig. 3. Probability of compromise versus zone size and network size.

In Figures 3 through 10, the points come from the simulation data set, and the curves plot the mathematical model. We extended the domain of these figures, the zone size, from 256 (in our prior work [9]) to 512.

For convenience, we reprint the simulation and model parameters from our prior work [9].

Figures 3 through 6 show a linear relationship between zone size and probability of compromise, while Figures 7

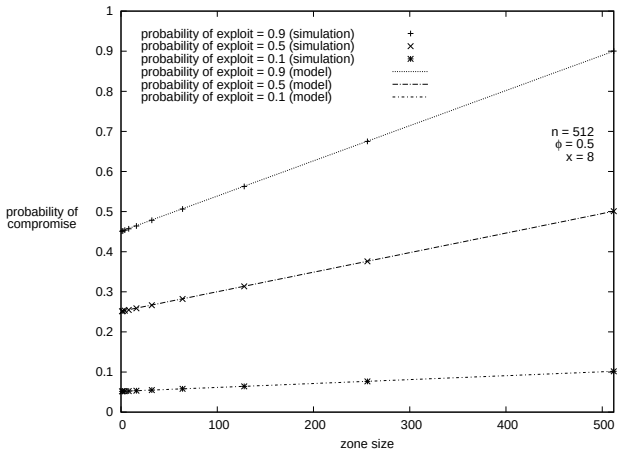

Fig. 4. Probability of compromise versus zone size and probability exploit exists.

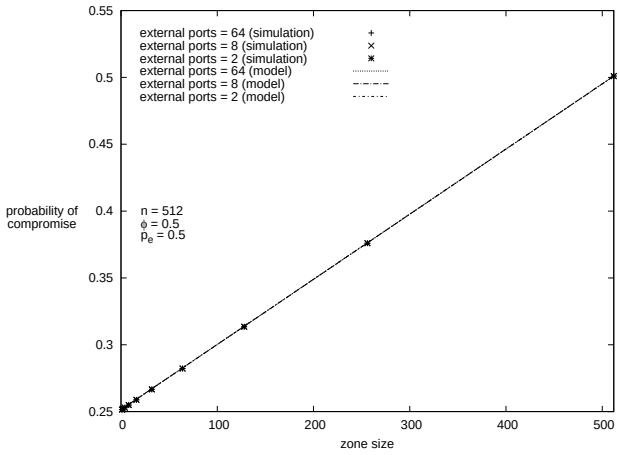

Fig. 5. Probability of compromise versus zone size and number of external ports.

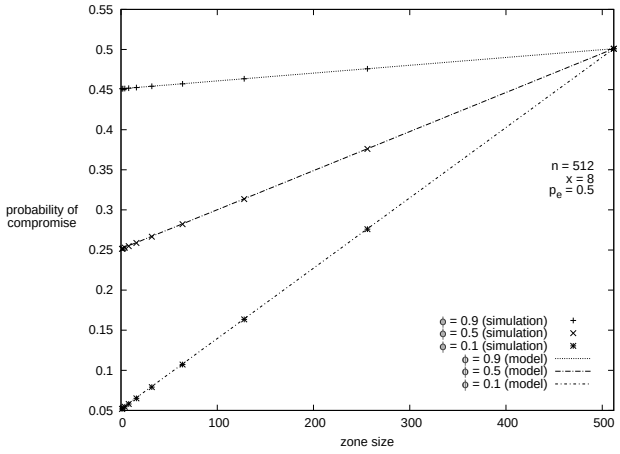

Fig. 6. Probability of compromise versus zone size and porosity.

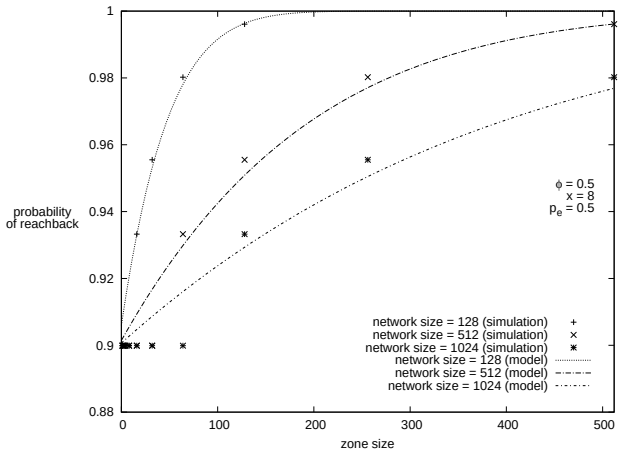

Fig. 7. Probability of reachback versus zone size and network size.

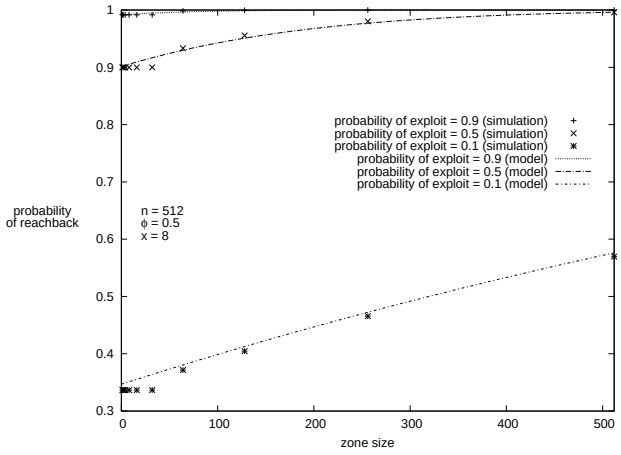

Fig. 8. Probability of reachback versus zone size and probability exploit exists.

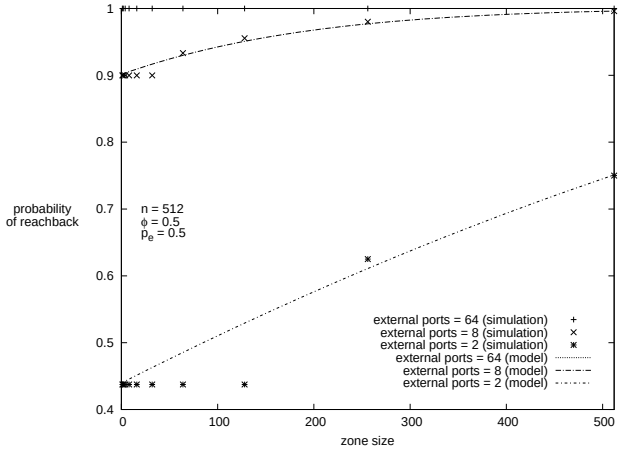

Fig. 9. Probability of reachback versus zone size and number of external ports.

through 10 show a sublinear relationship between the same variables. These relationships and the effects of network size and probability of exploitability on probability of compromise and probability of reachback are also consistent with our

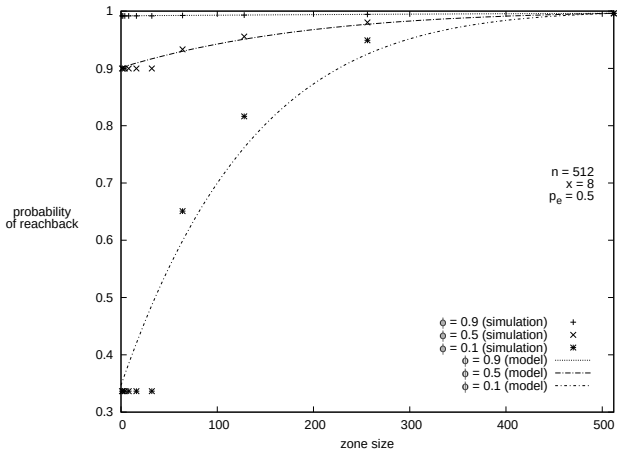

Fig. 10. Probability of reachback versus zone size and porosity.

TABLE I
SIMULATION PARAMETERS.

Parameter Name	Description	Source
p_c	probability of compromise	Simulation
p_r	probability of reachback	Simulation
external ports	Internet facing ports	Defender
neighbors	network adjacencies	Defender
insiders	initial insiders	Attacker
p_e	probability an exploit exists	Attacker

earlier work [9].

Figures 6 and 10 show porosity has a direct relationship probability of compromise and probability of reachback, respectively. This is intuitive: as the zones become more porous, malware is more likely to infect hosts outside of the zone. Figure 6 shows the simulation results match the mathematical model exactly when predicting probability of compromise. On the other hand, Figure 10 shows a small residual between the simulation results and the mathematical model when predicting probability of reachback; this error decreases as porosity and zone size increase. The mean squared error (MSE) between the simulation and mathematical model data sets is 0.002. The slopes of the lines in Figure 6 and the range of the curves in Figure 10 indicate the effect of zone size increases as the porosity decreases. This is because as the zones become more porous, the protection they provide becomes less impactful.

TABLE II
CLOSED FORM MATHEMATICAL MODEL PARAMETERS.

Parameter Name	Description	Source
p_c	probability of compromise	Model
p_r	probability of reachback	Model
p_e	probability an exploit exists	Attacker
z	zone size	Defender
n	network size	Defender
x	number of external ports	Defender

We observe a couple of trends among the original parameters that did not make it into the prior paper [9]: First, the slopes of the lines in Figure 3 and the range of the curves in Figure 7 indicate the effect of zone size on the probabilities of compromise and reachback increases as the network size decreases. Also, the slopes of the lines in Figure 4 indicate the effect of zone size on probability of compromise increases as the probability of exploitability increases; however, the range of the curves in Figure 8 indicate the effect of zone size on probability of reachback decreases as the probability of exploitability increases.

V. PARAMETERS

In this section, we propose methods to estimate p_e from our earlier work [9] and ϕ which we introduce in this work. We provide estimation techniques for these parameters because the remainders of Tables I and II are more intuitive and ϕ is a newcomer.

A. Probability of Exploitability

p_e is the probability of exploitability. It can be estimated based on a vulnerability scan of the subject network; tools like Security Center and Nessus [10] can be applied to an enterprise or a representative sample of an enterprise to collect raw vulnerability data. One way to transform this raw data into p_e is to divide the number of hosts with a vulnerability by the total number of assessed hosts.

$$p_e = \frac{\text{vulnerable hosts}}{\text{hosts scanned}} \tag{4}$$

A more sophisticated way to estimate p_e is to weight the host vulnerabilities.

$$p_e = \sum_i \frac{s_i}{s_{\max}} \tag{5}$$

s_i could be the number of vulnerabilities on the host or a numerical score associated with the most severe vulnerability present. In the former case, $s_{\max}$ is the maximum vulnerability count; this can be lowered and the actual count clipped to avoid a few hosts biasing p_e. In the latter case, $s_{\max}$ is the maximum severity; one mapping scheme assigns 4 to critical severity vulnerabilities, 3 to high and so on for medium, low and informational. As a proof of concept, we studied a Security Center report from an enterprise comprising 51,199 hosts. We wrote a Python script to extract the number and severity of vulnerabilities associated with each host and estimated p_e per Equation 4 and Equation 5 using count and severity weighting. Consider the following microexample with hosts A, B, C and D where A and B have no vulnerabilities, C has

five low severity vulnerabilities and host D has one critical vulnerability: p_e per Equation 4 is

$$p_e = \frac{2}{4} = 0.5 \tag{6}$$

while p_e per Equation 5 weighted by vulnerability count is

$$p_e = \frac{0+0+5+1}{4 \cdot 5} = 0.3 \tag{7}$$

and p_e per Equation 5 weighted by vulnerability severity is

$$p_e = \frac{0+0+1+4}{4 \cdot 4} = 0.3125 \tag{8}$$

B. Porosity

ϕ is the porosity of the zone boundaries. It can be estimated based on the ports permitted to communicate. One way to do this is to divide the number of well known (1024) or Internet Assigned Number Authority (IANA) registerable (49,152) ports for which traffic is allowed to cross a zone boundary by the total number of ports:

$$\phi = \frac{\text{ports open}}{\text{total ports}} \tag{9}$$

A more sophisticated way to estimate ϕ is to weight the ports because some ports may create more vulnerability than others:

$$\phi = \frac{\sum_i w_i o_i}{\sum_i w_i} \tag{10}$$

where w_i is the weight of port i and o_i is a Boolean variable which is 1 if the port is open and 0 otherwise. Weights could be collected by extracting data from a vulnerability database: ports that appear more often and/or more recently could have larger w_i than those appearing less often and/or less recently. As a proof of concept, we downloaded MITRE'S Common Vulnerabilities and Exposures (CVE) database, which has 90,880 weaknesses as of the time of this writing. We wrote a Python script to count the number of vulnerabilities associated with each port number and calculate a weight for each port in the following fashion:

$$w_i = K + v_i \tag{11}$$

v_i is the number of vulnerabilities associated with port i, and the K term allows a port with no vulnerabilities to impact ϕ. K is a constant; it is the relative weight of a port with no known vulnerabilities. For example, if we set $K = 1$, $w_{113} = 2$ because there is one vulnerability (CVE-2007-2711) associated with this TCP port (an authentication service used by IRC servers). On the other hand, $w_{80} = 15$ because there are fourteen vulnerabilities associated with this TCP port (HTTP). For reference, the five highest weighted TCP ports are:

1) 80 (HTTP)
2) 443 (HTTPS)
3) 3050 (gds_db)
4) 25 (SMTP)
5) 23 (telnet)

Consider the following microexample: only considering TCP ports 80 and 113, if connections on port 80 are blocked while those port 113 are allowed,

$$\phi = \frac{0 \cdot 15 + 1 \cdot 2}{15 + 2} = 0.12 \tag{12}$$

However, we do not advocate blind reliance on numerical data: although TCP and UDP ports 53 don't appear in CVE entries, they are particularly important to shut down if the operating environment allows this. Disabling these ports will prevent the DNS lookups that are prevalent in malware C2. Estimating K and performing a sensitivity analysis are possible future research topics.

VI. Conclusions

In this paper we proposed revisions to our earlier simulation and model that allow for porous zone boundaries and methods to estimate these artifacts' input parameters. Our prior work [9] accommodated this extension neatly, and the fidelity of the earlier work benefits from this addition. The parameter estimation procedures address an open thread in our previous paper and makes the investigation more germane to practitioners.

Our team is pursuing five future research leads: First, we will model host remediation. The current model does not allow for compromised hosts to be restored to a legitimate state and returned to the system; this will close a significant gap between this research and cyber security practice. We will use a stochastic model instead of a mathematical model, but we will validate the result using an extended version of the present simulation. Second, we will consider the impact of time on a cyber attack. The passage of time will allow the defender to patch the vulnerabilities that the attacker exploits. Considering time will also create a game of attrition; at some point the attacker will run out of treasure to spend on the operational costs of the cyber attack. Third, we will study zones that are formed dynamically based on inter-host collaboration rather than statically by function. This will increase the relevance of our modeling and simulation to the computing enterprises largely comprising human-operated workstations rather than purpose built devices. Fourth, while we consider the interzone porosity using ϕ, intrazone porosity is a factor as well. We will include intrazone porosity in our future work. Fifth, ϕ parameterizes a global interzone porosity for the system while in reality, each zone boundary should have an independent parameter. The challenge here will be managing the number of these porosities, which will grow as the square of the number of zones (z). Following these five lines of investigation will yield significant progress towards our ultimate goal of a complete and orthogonal model for cyber zone defense.

References

[1] Z Berkay Celik, Robert J Walls, Patrick McDaniel, and Ananthram Swami. Malware traffic detection using tamper resistant features. In *Military Communications Conference*, pages 330–335, Tampa, Florida, USA, October 2015. IEEE.

[2] Andrew Gontarczyk, Phil McMillan, and Chris Pavlovski. Cyber Security Zone Modeling in Practice. In *10th International Conference on Information Technology and Applications*, Sydney, Australia, July 2015.

[3] Hannes Holm, Khurram Shahzad, Markus Buschle, and Mathias Ekstedt. P CySeMoL: Predictive, Probabilistic Cyber Security Modeling Language. *IEEE Transactions on Dependable and Secure Computing*, 12(6):626–639, 2015.

[4] Sunghyuck Hong. Efficient and secure DNS cyber shelter on DDoS attacks. *Journal of Computer Virology and Hacking Techniques*, 11(3):129–136, 2015.

[5] Yasir Imtiaz Khan, Ehab Al-Shaer, and Usman Rauf. Cyber Resilience-by-Construction: Modeling, Measuring & Verifying. In *Workshop on Automated Decision Making for Active Cyber Defense*, pages 9–14, Denver, Colorado, USA, October 2015. ACM.

[6] Christian Leuprecht, David B Skillicorn, and Victoria E Tait. Beyond the Castle Model of cyber-risk and cyber-security. *Government Information Quarterly*, 2016.

[7] Wataru Machii, Isao Kato, Masahito Koike, Masafumi Matta, Tomomi Aoyama, Hidemasa Naruoka, Ichiro Koshijima, and Yoshihiro Hashimoto. Dynamic zoning based on situational activitie for ICS security. In *10th Asian Control Conference*, pages 1–5, Kota Kinabalu, Sabah, Malaysia, June 2015. IEEE.

[8] Nick McKeown, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner. OpenFlow: Enabling Innovation in Campus Networks. *SIGCOMM Computer Communication Review*, 38(2):69–74, March 2008.

[9] Robert Mitchell, Paul Sery, and Tom Klitsner. Foundations for Cyber Zone Defense. In *International Conference on Computer Communication and Networks*, Waikoloa, Hawaii, USA, August 2016.

[10] Russ Rogers. *Nessus Network Auditing*. Syngress Publishing, 2 edition, 2008.

[11] Pablo G Sack and Jorge S Ierache. Initial proposal of a framework in the context of cyberdefense to articulate controls and metrics associated. In *International Conference on Computing, Communication and Security*, pages 1–6, Mauritius, December 2015. IEEE.

[12] Saleh Soltan, Mihalis Yannakakis, and Gil Zussman. Joint Cyber and Physical Attacks on Power Grids: Graph Theoretical Approaches for Information Recovery. In *SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, pages 361–374, Portland, Oregon, USA, June 2015. ACM.