

CYBER ASSET VULNERABILITY RANKING ALGORITHM FOR SECURITY RISK  
MANAGEMENT

A Thesis

Presented to

The Faculty of the Department of Computer Engineering  
San José State University

In Partial Fulfillment

of the Requirements for the Degree  
Master of Science in Software Engineering

by

Partha Sarathi Ghosh

December 2019

© 2019

Partha Sarathi Ghosh

ALL RIGHTS RESERVED

The Designated Thesis Committee Approves the Thesis Titled

CYBER ASSET VULNERABILITY RANKING ALGORITHM FOR SECURITY RISK  
MANAGEMENT

by

Partha Sarathi Ghosh

APPROVED FOR THE DEPARTMENT OF COMPUTER ENGINEERING

SAN JOSÉ STATE UNIVERSITY

December 2019

Younghee Park, Ph.D.

Department of Computer Engineering

Kaikai Liu, Ph.D.

Department of Computer Engineering

Vikrant Nanda, M.S

Adjunct Professor, School of Business

## ABSTRACT

### CYBER ASSET VULNERABILITY RANKING ALGORITHM FOR SECURITY RISK MANAGEMENT

by Partha Sarathi Ghosh

In cybersecurity, risk management is a dominant topic today. Comprehensive cyber regulations like, General Data Protection Regulation (GDPR) in European Union (EU) and California Consumer Protection Act (CCPA) in California are tightening the grip on every company with cyber infrastructure. Security experts and Chief Information Security Officers (CISO) are looking at cyber risk management to comply with cyber laws and to minimize impact of cyber attacks. Security engineers struggle to create inventory of cyber assets and evaluate their risk profiles and vulnerabilities. There is no software available to automate inventorying of cyber assets, classify their risk profiles, and perform risk management on them. In this thesis I intend to create an algorithm to rank the vulnerability of assets in an organization using stochastic methods and provide a risk management tool. I review technical papers to identify a set of cyber assets that are most common in an enterprise network, evaluate if Markov process is the right mathematical model to derive the ranking algorithm, and derive the foundational mathematical equations for the ranking algorithm.

## ACKNOWLEDGMENTS

I would like to thank my advisors Dr. Younghee Park, Prof. Kaikai Liu, Ph.D., and Prof. Vikrant Nanda, M.S., for their encouragement and support while working on this thesis. I would also like to thank Prof. David C. Anastasiu, SJSU, for creating this LaTeX template and ensuring that it accurately follows the graduate studies thesis style requirements. Last, but not least, I would like to thank my wife Gargi Ghosh Hazra; without her support, I would not have been able to pursue graduate studies. Finally, I would like to thank my parents for all that they have done to support me financially during my early education with all their hardships.

## TABLE OF CONTENTS

|  |      |
|--|------|
| List of Tables .....                                     | vii  |
| List of Figures .....                                    | viii |
| 1 Introduction.....                                      | 1    |
| 2 Literature Review .....                                | 3    |
| 2.1 Cyber Assets .....                                   | 4    |
| 2.1.1 Software Configurations .....                      | 4    |
| 2.1.2 Ubiquity of Networked Devices .....                | 5    |
| 2.1.3 Classification of Cyber Assets .....               | 5    |
| 2.1.4 End Points in IT Infrastructure Assets.....        | 6    |
| 2.1.5 Cryptocurrency Assets.....                         | 7    |
| 2.1.6 Cloud Software Assets.....                         | 7    |
| 2.1.7 Manufacturing Industry Assets.....                 | 8    |
| 2.1.8 Smart City Assets .....                            | 9    |
| 2.2 List of Cyber Assets .....                           | 9    |
| 2.2.1 Asset Vulnerability.....                           | 10   |
| 2.3 Ranking Algorithm .....                              | 10   |
| 2.3.1 Stochastic Methods .....                           | 11   |
| 2.3.2 Markov Process .....                               | 11   |
| 2.3.3 Nature of a Cyber System .....                     | 13   |
| 2.3.4 Asset Vulnerability Equation Derivation .....      | 13   |
| 2.4 Literature Review Conclusion .....                   | 14   |
| 3 Algorithms.....  | 15   |
| 3.1 Asset Tracking Crawler .....                         | 15   |
| 3.2 Vulnerability Ranking Algorithm.....                 | 15   |
| 3.2.1 Mathematics of the Ranking Algorithm .....         | 15   |
| 3.2.2 Complexity Analysis of the Ranking Algorithm ..... | 15   |
| 4 Future Work .....                                      | 16   |
| 5 Conclusions .....                                      | 17   |
| Literature Cited.....                                    | 18   |
| Appendix A: Markov Chain Analysis.....                   | 20   |
| Appendix B: Bayesian Statistics .....                    | 21   |

## LIST OF TABLES

## LIST OF FIGURES



## 1 INTRODUCTION

In a business enterprise, Information Technology (IT) infrastructure has numerous physical and virtual computing assets. These assets could be physically co-located or be in a data center facility, which could be thousands of miles apart, beyond the physical reach of a system administrator. Security engineers have to account for each of these assets, using different methods like Application Programming Interface (API) calls to cloud service providers, running network scanners, or perform penetration testing. Virtualization and containerization are becoming the de facto mode to provide application services. Having only bare metal servers in computing infrastructure is becoming obsolete in enterprises. IT infrastructure will continue to use a hybrid computing environment, consisting of bare metal servers and cloud service infrastructure. Looking from a security engineers' perspective, the challenge to identify a cyber asset, risk profile it, and then mitigate the risk, becomes much more daunting with the unavailability of a cyber asset inventorying software.

Securing the computing infrastructure is achieved with a layered or perimeter security around the connected assets using Access Control Lists (ACLs) in routers, application firewalls, deep packet inspections, etc. Even with all these bells and whistles for securing the computing infrastructure being in place, attack on networks and data breaches are common. The magnitude of the number of assets in an enterprise multiplied with the vulnerability present in those assets, makes cyber asset management a complex task. Apart from securing the computing resources using security devices and algorithms, Chief Information Security Officers (CISOs) are looking for risk management of the cyber assets. Cyber risk management involves identifying assets, creating risk profiles for them, and classifying them into different categories of risk profiles. To make this task even more complex, new vulnerabilities are discovered thereby making the risk profiles of the assets change dynamically with time.

The subject of this work is to classify cyber assets, identify them, and then rank them in order of their vulnerability. The purpose of this thesis is to tackle the problem of dynamic changes in threat landscape for assets in enterprise and propose a solution that would perform asset inventorying, risk classification, and cyber risk management.

In this research paper (as part of ENGR-200W course work), I plan to focus on doing literature review in section 2, to classify assets, identify the assets algorithmically, assign a risk profile, and then create a vulnerability ranking of the cyber assets, provide direction on future work in section 4, and sum up the key achievements in conclusion in section 5. This work will create a solid foundation for my future work on this thesis in the summer of 2020.

The domain of this work is broad and in the time available, I would restrict the scope of this work to do a study of contemporary work in similar areas and then create a comprehensive list of cyber assets, and look into the use of Markov process for finding the ranking algorithm based on vulnerabilities. Risk management methodology is beyond the scope of this work.

This is a software product space in which there is hardly any comprehensive software suite available in the market. There is no software today which does asset inventorying, asset ranking, risk classification, and perform cyber risk management. That makes this research an interesting and promising subject.

## 2 LITERATURE REVIEW

In this section, I review a number of papers to dissect the subject of this thesis. The title has three key topics that need to be demystified by referring to contemporary research in similar areas of cyber asset detection, and use of stochastic methods in network security. These topics are: cyber assets in section 2.1, ranking algorithm in section 2.3, and security risk management (*this topic is not covered in this literature review*). Classifying cyber assets and deriving the ranking algorithm would need a perspective on the current and previous researches in these areas. Vulnerability could be a quantitative or a qualitative parameter, which would be used as the parameter in the predictive algorithm for the asset ranking. Numerous attributes determine the cyber asset vulnerabilities. It could be Operating System (OS), hardware, zero-day vulnerabilities, software used, etc. The focus of the literature review would be to look for the different types of assets that could be classified as cyber threat and to try to parameterize the attributes of the vulnerability matrix.

The ranking algorithm will be a mathematical methodology that would be used to derive a rank for each cyber asset in an enterprise. The threat landscape for the cyber assets change with time. So, the risk profile associated with the cyber asset is a function of time. Each cyber asset needs to be treated independently for vulnerability assessment, based on their attributes and parameters before deducing a rank with respect to all the cyber assets. Stochastic methods are well suited for dynamic ranking, where the state of the asset changes and so do the attributes that constitute its vulnerability. So, a study of the stochastic methods is very relevant and appropriate for this subject.

The objective of this literature review is to create a pedestal for the individual component of this thesis so that a complete solution could be hypothesized. The algorithms in section 3 can take novel approaches by choosing to learn from earlier works and eliminating improbable choices.

## 2.1 Cyber Assets

Waedt *et al.* [1] define an asset as “... *something that has a potential or actual value for an organization*”. In this thesis, the word “asset” always means a cyber asset. In this section I write about classification, identification, and visibility of cyber assets in an enterprise. In this thesis, I have used the vocabulary of systems and software engineering [2]. Evaluation of the different cyber assets could possibly be, but not limited to, entities in cyber risk management. I review the dependencies among cyber assets in [3]. Cloud computing system assets are looked into, in detail, in [4].

### 2.1.1 Software Configurations

Software configurations are one of the most vulnerable components in the IT infrastructure. Software configurations build the network segment and create the compute and storage blocks in the IT infrastructure. Software configurations are used in a computing device, a network device, a virtual device, an application program executing on a multitude of computing devices, security configurations, etc. This poses a high risk because these configurations are often hand crafted and visually validated. Humans are the weak links in cybersecurity. Human gullibility and fallibility are the primary reasons for majority of the cyber incidents. Hagen, in [5], questions if human relationship with security education is an ongoing challenge. Note here that a person who creates a configuration, also knows about the drawback or weak links in a configuration. Trust is the most important credential for the people who create these configurations. Thus, human errors are bound to happen. Automated validation of configurations can mitigate risks caused by human errors. Errors happen even with checks and balances, due to oversights or errors in validation software. So, the human errors pose a huge challenge in cybersecurity. This risk needs to be addressed in cybersecurity risk management. Software configurations are cyber assets that need to be protected and classified, and a risk profile needs to be associated with them. For this entity, I study [6], [7].

### *2.1.2 Ubiquity of Networked Devices*

Devices with network connectivity are ubiquitous. The importance of connected devices in any enterprise is tied to their profitability because any activity in an enterprise (inventory assessment, procurement, billing, delivery, delivery tracking, etc.) becomes highly automated with connected devices. That increases productivity and profitability. As a large scale example of connected devices, a smart city is a complex grid of connected devices. Network connectivity introduces an element in cyber risk that can have an impact on life and property, unless the cyber assets are inventoried and risk profiled. I review the cyber assets in a smart city in [1]. Connected devices in the manufacturing sector bring in new challenges in cybersecurity. The devices used in a manufacturing pipelines are networked and remotely operated. They often use the same set of base software akin to any desktop or laptop. Stuxnet, an industrial sabotage causing computer worm, as described by Langner [8], was the cause of a cyber breach in Iran's nuclear reactors. The devices in these nuclear reactors were using Windows OS, an OS used in any home and office computing device. Cyber perpetrators used the vulnerability in Windows OS to cause malfunction in the nuclear reactors, causing the nuclear centrifuge to go out of control and eventually damaging the nuclear reactor beyond repair. The point here is that, large or small scale industrial process controls use the same set of software and hardware, exactly similar to consumer grade software and hardware. The context of deployments and functionality could be different in the industrial space. However, though the vulnerabilities posed are similar, the financial impact in industrial context is significant. So, the identification of the assets in an industrial process control is important and is looked into in [9].

### *2.1.3 Classification of Cyber Assets*

A cyber asset could be a hardware, a Virtual Machine (VM), or a software service running in a cloud server, which is not owned by the enterprise, a database server, or a

process running in any server. A cyber asset's location could be a home, a business facility, a sports arena, a theater, a multi storied building, a smart city, etc. The presence of networked devices in any proximity to users makes it challenging to classify a device that can be a cyber threat. A non-connected device like a USB storage (commonly known as a pen drive) could be the unlikeliest source of cyber threat when connected to a network infrastructure through a computing device like laptop. Cyber threats through non-connected devices are also cyber risks. This topic is beyond the scope of this thesis because it's a specialized case of cyber policy and cyber hygiene. The variety of devices that connect to the network makes the classification of the cyber assets a difficult task. For this thesis to be developed into a Minimal Viable Product (MVP), the scope of the cyber assets is restricted to more common asset types, like, cloud infrastructure [9], virtual cyber assets [10], and a large scale infrastructure like smart city [1]. Assets in an enterprise network that consist of wired and wireless networks [11] [12] need to be considered in the classification of cyber assets.

#### *2.1.4 End Points in IT Infrastructure Assets*

In a study on enterprise network security by Chen *et al.* [11], they focus on the end point security. End points are applications, network devices, and policy servers where security mechanisms would be implemented. Gustavo *et al.* [12] while proposing a new security architecture for mobile enterprise, discuss an enterprise network segment as illustrated in [12, Figure 1, 2]. The cyber assets [12] are predominantly networked devices. Virtual Private Network (VPN) is determined as an asset in this study. With the focus on using VPN to secure the enterprise infrastructure as described in [12], the classification of *network connectivity* as a cyber asset becomes new concept. When hybrid infrastructure is used, like, on-prem (compute and storage resources in a facility owned by the enterprise) and cloud infrastructure, the connecting pipelines become key targets for attacks. Often these these data pipelines (physically optical fiber) are shared between multiple tenants.

These data pipelines are not point-to-point, but rather multi-hop and span the infrastructure of multiple Internet Service Providers (ISPs). Protecting data pipelines using encryption technologies is key to security.

#### 2.1.5 *Cryptocurrency Assets*

Peng *et al.* [10] talk about cryptocurrency, one of the most interesting virtual assets in any enterprise. They say “In recent years, to facilitate the user to purchase online paid services, the game operators and Portal network enterprise have launched the network virtual currency of a variety of names . . .”. Virtual currency is gaining prevalence. This form of money does not have a physical form nor is supported by major financial organizations and governments, yet businesses have started accepting it. Securing this virtual asset is of great importance. The solution seems to be like securing data that is through encryption. There would be a lot of work that would be undertaken in the area of securing virtual money as it becomes more acceptable by governments and financial institutions.

#### 2.1.6 *Cloud Software Assets*

Ease of use and cost benefits have caused a paradigm shift for the enterprises to use cloud infrastructure more than often. For productivity needs of any enterprise, the choice of office software suites is provided by two major providers — Microsoft and Google. These are browser based softwares, which have made collaboration and document versioning easy for the user. On the other hand, this has become a pain for the security professionals. Most of these cloud based softwares are encrypted end-to-end even in the corporate networks. Often the security professionals are concerned about losing corporate documents through this browser based productivity software. Kushwah *et al.* [4] detail some key cyber assets in a cloud infrastructure. This service is known as Infrastructure as a Service (IAAS). These are VMs, which run compute resources provided by the infrastructure provider and are often shared with multiple tenants. This

means that a physical server could be shared by multiple enterprises and a different instance of the OS running on a physical server would be the security boundary. Service providers ensure security of the VMs for each client. The most interesting part of the VMs is their mobility. This means that if a server is running out of capacity in terms of compute performance, then the VMs could be shifted to another physical server. In this circumstance, the key requirement is to ensure that there is no downtime for the services provided by the VM that has been moved. The challenge in VM mobility is that the network security policy associated with the VM also needs to be shifted. This is a complex operation, often manual and error prone. With no real-time audit of the security of the VM mobility, this is a potential security threat. Virtual assets are more prevalent in computing environment and keeping a track of all the VM resources is a challenge for the IT security staff.

#### *2.1.7 Manufacturing Industry Assets*

Process Control Systems (PCSs), also known as Industrial Control Systems (ICSs) are built with pieces of equipments that are used in production lines in manufacturing process. Over the course of the past few decades, off-the-shelf components have been making inroads into the PCSs. These network-connected and automated components are known as Internet of Things (IOT) in cybersecurity parlance. Cyber attacks affect the PCSs. The impact of cyber attacks on assets involving PCSs has been described in great detail by Kiss *et al.* [9]. In their study they not only evaluate but try to quantify the impact of the cyber attacks on the elements that consist of the PCSs. Note here that the test bed used in [9], is a chemical engineering plant, Tennessee Eastman Chemical Process, Tennessee. These devices are numerous. Each of them needs to be identified, and can be identified, using OS scanning techniques. Since most of these devices operate in a production system, it is not only important to monitor the network data that is originating from these devices but also the processes running in them. If a malicious process runs in these



devices then the production system is compromised. In this thesis, I aim to introduce a technique that would identify all the devices in a PCS.

#### *2.1.8 Smart City Assets*

An example of a large scale asset inventory is a smart city [1]. A smart city has a network of devices that automates the operations of a city. The boundary of automation is not well-defined but left to human ingenuity. Traffic signals, trains, light rails, air conditioning, street lights, parking ticket kiosks, traffic lanes that are tolled, access to buildings, and so on, are some examples of the assets in a smart city. Waedt *et al.* in [1] broadly classify the cyber assets in a city by grouping them as “physical, infrastructure and movable assets”, “Information & Communication Technology (ICT) assets”, “intangible assets”, and “critical assets”. In [1, Figure 3], the main concept of asset management has been laid out. Each element in a city has to be identified and appropriately tagged before that asset can be part of any automation operation in the city. The problem to solve this is enormous and this is improbable to be implemented in any existing city. This problem is not only a computing problem but is also a public policy issue. Securing such an infrastructure using asset management is only possible when each component is risk-analyzed independently.

### **2.2 List of Cyber Assets**

The classification of cyber assets through automation is the key to simplify the challenge of cyber asset identification. Cyber assets in organizations may vary, but they could be classified in one of the following categories:

- 1) Network connectivity devices, e.g., switches, routers, wireless access points, etc.
- 2) Compute devices, e.g., servers, laptops, etc.
- 3) Virtual compute resources, e.g., VMs, virtual routers, etc.
- 4) Network data pipelines, e.g., VPN, fiber, Local Area Network (LAN) cables, etc.

- 5) Data assets, e.g., documents, code, logs that are electronically stored, etc.
- 6) Identity management, e.g., user login credentials, etc.
- 7) System configurations, e.g., switch, router configuration, firewall configurations, etc.

In this thesis, I will retain the focus on these cyber asset classifications, and identify and rank them based on their vulnerabilities.

### *2.2.1 Asset Vulnerability*

Vulnerability is defined as the weakness in a system and covers the various cyber assets and their weak links that make the software or hardware vulnerable to hackers. There are two common methods used by organizations to ascertain vulnerability of cyber assets. They are Open Web Application Security Project (OWASP) and Common Vulnerability Scoring System (CVSS). In both these methods, the vulnerability of a software program is metered on a scale of 1 to 10, 10 being the highest in terms of severity. This method of assessing software vulnerability works because only a part of a device gets impacted by a bug. The impact of a bug depends on how that software is being used. If the bug is in a database then the system's vulnerability is high because usually it is the database that maintains the application or systems data. Though these methods do not take into account the vulnerability of a device, they are the apt metrics in calculating ranking of asset vulnerability for this thesis.

## **2.3 Ranking Algorithm**

Cyber events are unpredictable in nature but they are happening all the times. It's unpredictable because the modus operandi of the perpetrators are unknown. Visibility of the assets would reduce the response time to react on a cyber event. There could be thousands of cyber assets that need to be monitored. If the assets are not ranked then monitoring the assets, the mitigation techniques to be applied, or budgeting of dollars for fixing the vulnerabilities become a humongous task. Hence, there is a need to rank the

assets for different processes involved in managing cybersecurity. New bugs get detected every day. Zero-day vulnerabilities are hidden bugs and are present in systems. The security issues are dynamic and ever-changing, to be more specific “random” in nature. So an algorithm to rank these assets should be capable to take into account the continuously changing vulnerability of the assets.

There are multiple mathematical models to explore and derive a ranking algorithm. In this thesis, I take a cue from the work of [13], and restrict the scope to exploring the Markov chain for the derivation of the asset ranking algorithm.

### 2.3.1 Stochastic Methods

Processes in nature are random and can most of the times be explained by the laws of probability. Pukite *et al.* [14] introduce Markov chain analysis as “*These processes that include growth and decay of living organisms, spread of epidemics, decay of radioactive material, traffic on freeway, and failure and repair of electronic systems. The study of stochastic process can be defined as the 'dynamic' part of probability theory in which we study a collection of random variables, their interdependence, their change in time, and limiting behavior*” [14]. There are two types of stochastic processes — stationary and nonstationary or evolutionary. In a stationary stochastic process, the probability of an event happening remains constant over a period of time, while in an evolutionary process, changes can be found over a period of time.

### 2.3.2 Markov Process

“*Markov process is a random process, whose future probability is determined by the probabilities of the current process*” [15]. To explain it lucidly, we can estimate the probability of a rain tomorrow, from the probability of rain happening today. Similarly, the rain probability of tomorrow would determine the rain probability for the day after. Note that this is not deterministic. For the fundamentals of Markov process, please refer to [14]. In deriving the algorithm to find out the vulnerability of the asset, we will be

using this statistical principle. I will review a few contemporary research, in the area of network security, where Markov chain analysis or Markov process is being used.

Karras *et al.* proposed an efficient security model based on Markov process [16]. Networks are dynamic in nature with compute end points being attached and removed simultaneously. Concurrently, the network security policy for the network gets changed. The state of the network device and the rate of changes that the network device undergo with changes of the security configuration are hardly static. This work is of great relevance to this thesis and is an encouragement for choosing Markov process. Markov process has been used by Shing *et al.* in processing information security risks. The researchers use Markov chain to monitor state changes dynamically in information security, which is helpful in making better decisions to protect information in organizations. Cheng *et al.* use Markov model in determining the optimal strategy in network security for moving targets (continuous attacks in networks). Wang *et al.* use Markov chain along with game model for analyzing network security models. One of the key learnings from this paper is the use of nonlinear programming to achieve equilibrium so that a defending policy can be effectively applied against the attacker's methods. Abdulmunem *et al.* combine Markov model with attack tree analysis to describe Building Automation System's (BAS) availability and security models. They base their research on finding faults causing disruption in building operations. They use Markov model to observe fault behavior during operation time and calculate system availability in the BAS. The domain of this work is not based on the TCP/IP networking stack, though it serves as a good model for BAS security. Note here that BAS has been transitioning to use TCP/IP based networks. Thus, if only the source of the data point is changed in the BAS systems, the work is still relevant.

### 2.3.3 Nature of a Cyber System

A cyber system (enterprise network with assets) state is dynamic. The rate of change of the elements in this system that impacts the current state are also dynamic. Prediction of future events could be achieved based on the probability of the current events. The authors of the research papers on Markov chain overwhelmingly support the use of stochastic methods and specifically Markov chain. So these algorithms would prove to be very crucial for validation of the asset vulnerability rank algorithm that I intend to develop.

### 2.3.4 Asset Vulnerability Equation Derivation

Asset vulnerabilities change with time based on a multitude of parameters that are constantly changing. In this scenario, a non-deterministic stochastic method would reflect the correct vulnerability of an asset and hence, is the most appropriate method to use.

Let's try to explain this mathematically.

Let's say that Asset **A** (4) has a vulnerability **V** (1). So, the vulnerability of an asset can be expressed as:

$$V_A = \sum_{i=1}^n v_i \quad (1)$$

Vulnerability of each component in the asset is:

$$v_i = \sum_{i=1}^n f(P_i) \quad (2)$$

Each component is made up on multiple entities. So, the probability of the entity being vulnerable is:

$$f(P) = \prod_{i=1}^n p_i \quad (3)$$

The ranking algorithm would derive the set of vulnerability of assets (**V** (1)), as defined by:

$$A = \{V_i \mid 0 < i < n\} \quad (4)$$

A ranking algorithm would use the stochastic methods to create a rank of an asset based on the vulnerability of the assets. The ranking algorithm is similar to web page ranking used by search engines in [13]. In case of search engines, the relevance and weights are based on the web page. In the scenario of cyber assets, the ranking algorithm would be based on the vulnerability assessment of the asset. The rank of the asset would not be the same always because the threat landscape changes.

## **2.4 Literature Review Conclusion**

A cyber asset's vulnerability is changing constantly. So, any method to determine a ranking algorithm based on the static stochastic process would be inaccurate, less reliable, and its effectiveness would be questionable. Thus, the obvious choice for the ranking algorithm is an evolutionary stochastic process. The key elements that need to be considered in the evolutionary stochastic process are the state of the asset and the rate at which the vulnerabilities change for the asset. Here again, the rate of vulnerability change is a real-time function of the changes in the vulnerability of the individual component of the assets. This approach is different from the current practiced method in the industry. The general method of tracking a vulnerability is based on the application. In this approach, the totality of a system is not encompassed. It's possible that a vulnerability may exist in a system, however, the impact of it may not be relevant at all because the system might not be using the vulnerable part at all.

### 3 ALGORITHMS

In this section, I will discuss the different algorithms needed for this thesis. Since literature review is the only ask for this assignment, I have left this and the subsequent section in ([**To be completed in 2020**]) status. I will write this section once I start working on this thesis again in the summer of 2020.

#### 3.1 Asset Tracking Crawler

#### 3.2 Vulnerability Ranking Algorithm

[**To be completed in 2020**]

##### 3.2.1 *Mathematics of the Ranking Algorithm*

[**To be completed in 2020**]

##### 3.2.2 *Complexity Analysis of the Ranking Algorithm*

[**To be completed in 2020**]

#### **4 FUTURE WORK**

In this thesis I aim to develop a coherent solution to identify ranking of assets based on their vulnerabilities. The design of a crawler to look for assets is a never-ending process. An optimization calculation needs to be done to ascertain software architecture. The asset tracking is a highly compute intensive process and so a distributed algorithm has to be devised for high performance. A huge volume of data would be generated on a per minute basis. For storing this huge volume of data, one needs an efficient algorithm. The ranking algorithm is a NP-Hard problem, which means that it will not be complete in polynomial time. So another area of work would be to optimize this NP-Hard problem. With advances in distributed computing, future work would also be to fine-tune the ranking algorithm using Machine Learning algorithms, so that prediction can also leverage previous cyber events.



## 5 CONCLUSIONS

In this thesis, I intend to develop a cyber asset vulnerability ranking algorithm for security risk management. I conclude this thesis by listing down the key achievements so far:

- 1) I did a literature review of the cyber assets and created a list of preliminary cyber assets 2.2 that the crawler program 3.1 could work on.
- 2) I did a literature review of the stochastic methods with focus on Markov process 2.3.2 and establish that Markov process is the right mathematical model for the ranking algorithm.
- 3) I derived the base mathematical equations 2.3.4, which can be used as a starting point to develop the asset ranking algorithm.

## Literature Cited

- [1] K. Waedt, A. Ciriello, M. Parekh, and E. Bajramovic, "Automatic assets identification for smart cities: Prerequisites for cybersecurity risk assessments," in *2016 IEEE International Smart Cities Conference (ISC2)*, pp. 1–6, Sep. 2016.
- [2] "Iso/iec/ieee international standard - systems and software engineering – vocabulary," *ISO/IEC/IEEE 24765:2010(E)*, pp. 1–418, Dec 2010.
- [3] L. Buchanan, M. Larkin, and A. D'Amico, "Mission assurance proof-of-concept: Mapping dependencies among cyber assets, missions, and users," in *2012 IEEE Conference on Technologies for Homeland Security (HST)*, pp. 298–304, Nov 2012.
- [4] S. S. Kushwah and R. Tyagi, "Infrastructural assets provisioning in cloud computing systems," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, pp. 886–889, Dec 2015.
- [5] J. Hagen, "Human relationships: A never-ending security education challenge?," *IEEE Security Privacy*, vol. 7, pp. 65–67, July 2009.
- [6] X. Li, P. Avellino, J. Janies, and M. P. Collins, "Software asset analyzer: A system for detecting configuration anomalies," in *MILCOM 2016 - 2016 IEEE Military Communications Conference*, pp. 998–1003, Nov 2016.
- [7] M. Nodine, R. Grimshaw, P. Haglich, S. Wilder, and J. B. Lyles, "Computational asset description for cyber experiment support using owl," in *2011 IEEE Fifth International Conference on Semantic Computing*, pp. 110–117, Sep. 2011.
- [8] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security Privacy*, vol. 9, pp. 49–51, May 2011.
- [9] I. Kiss, B. Genge, and P. Haller, "Behavior-based critical cyber asset identification in process control systems under cyber attacks," in *Proceedings of the 2015 16th International Carpathian Control Conference (ICCC)*, pp. 196–201, May 2015.
- [10] H. Peng and L. Niu, "The risks of network virtual assets and its measurements," in *2009 International Conference on Networks Security, Wireless Communications and Trusted Computing*, vol. 2, pp. 434–437, April 2009.

- [11] C. Chen, K. Wang, and Y. Dai, “A novel architecture for enterprise network security,” in *2009 International Conference on Computational Intelligence and Security*, vol. 1, pp. 537–541, Dec 2009.
- [12] G. de los Reyes, S. Macwan, D. Chawla, and C. Serban, “Securing the mobile enterprise with network-based security and cloud computing,” in *2012 35th IEEE Sarnoff Symposium*, pp. 1–5, May 2012.
- [13] L. Page, S. Brin, R. Motwani, and T. Winograd, “The pagerank citation ranking: Bringing order to the web,” in *Proceedings of the 7th International World Wide Web Conference*, (Brisbane, Australia), pp. 161–172, 1998.
- [14] J. Pukite and P. Pukite, *Markov Process Fundamentals*, pp. 49–65. IEEE, 1998.
- [15] “Markov process – from wolfram mathworld.”  
<http://mathworld.wolfram.com/MarkovProcess.html>. (Accessed on 12/08/2019).
- [16] D. A. Karras and V. C. Zorkadis, “On efficient security modelling of complex interconnected communication systems based on markov processes,” in *2008 New Technologies, Mobility and Security*, pp. 1–7, Nov 2008.

## **Appendix A**

### **MARKOV CHAIN ANALYSIS**

**To be completed in 2020**

## **Appendix B**

### **BAYESIAN STATISTICS**

**To be completed in 2020**