5-1-2022

Vulnerabilities Posed by Information Technology Systems of Firms Fixed Effect Panel OLS

Aparna Narayanan

VULNERABILITIES POSED BY INFORMATION TECHNOLOGY SYSTEMS OF FIRMS FIXED EFFECT PANEL OLS

by

APARNA NARAYANAN

Presented to the Faculty of the Honors College of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

HONORS BACHELOR OF SCIENCE IN INFORMATION SYSTEMS

THE UNIVERSITY OF TEXAS AT ARLINGTON
May 2022

ACKNOWLEDGMENTS

I would like to thank my mentor, Dr. Ruochen Liao, for his continuous encouragement and guidance throughout the course of this research. He was consistent in helping me from understanding the significance of research to comprehending the results accurately. This research would not have been possible without his knowledge, experience, and willingness to teach.

I would also like to thank my Honors College Advisor, Ms. Bobbie Brown, for her support throughout this research.

Lastly, I would like to thank my Honors College peers who provided honest and constructive peer reviews, which helped me accomplish this research.

April 22, 2022

ABSTRACT

VULNERABILITIES POSED BY INFORMATION TECHNOLOGY SYSTEMS OF FIRMS FIXED EFFECT PANEL OLS

Aparna Narayanan, B.S. Information Systems

The University of Texas at Arlington, 2022

Faculty Mentor: Ruochen Liao

The information technology infrastructure of firms plays an essential role in their operation. With this, many cybersecurity risks arise at an organizational level. This research seeks to understand if the age and size of a firm are related to this risk factor. The research utilizes a dataset containing records of 13,075 companies across the period of 2014-2020, extracted from the directory of records of autonomous system numbers from CyberGreen. This study hypothesizes the following:

H1. Organizational age will have a positive relationship to the organizational risk score, making them more vulnerable to attacks.

H2. The organizational size will also have a positive relationship to the organizational risk score, making them more vulnerable to attacks.

The age and size variables were standardized, and we found that for a unit increase in age, the risk score increased by 0.0256, and for an increase in size, the risk score increased by 0.1427. This confirmed our hypotheses to be correct, concluding that older and/ or larger firms are more prone to risk.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Table Page

CHAPTER 1

INTRODUCTION

1.1 Background and Significance of Research

The IT infrastructures of firms have been playing an increasingly important role in the competitive advantage and operation of the firms. On the other hand, they are also becoming more embedded than ever, thus creating significant obstacles in replacing or upgrading them. Cybersecurity is necessary to protect data and services from theft: especially third-party data or client data such as Non-Public personal Data. However, there is always the inherent risk with cyberspace, and it is crucial to weigh the positives against the negatives of the services.

Hacked says, "According to Cloudflare, a company that provides DDoS mitigation services, the number of DDoS attacks have increased rapidly over the past year. From Q3 to Q4 of 2021, there was a 175% increase in the number of DDoS attacks." We learn that the importance of cybersecurity space and understanding the effects of characteristic features of firms with regards to the risk they possess is more critical than ever.

1.2 Objective

This research project explores where a firm's vulnerabilities lie, the reasons behind using certain services despite their vulnerabilities, and potential breaches. Organizational inertia literature has pointed out that firms tend to prioritize areas where they can reap the most return on investment and are unwilling to make changes in others. As a result, preventative measures against a future attack that may or may not happen after all is an investment that firms are reluctant to make. This project seeks to discover how firms of different sizes and ages react to these cybersecurity risks and the factors that motivate or hinder changes. The insight would help gain exposure to the current cybersecurity landscape in the business world and the challenges that firms face.

Several factors make IT systems of organizations vulnerable to risks. The objective of this research is to study and conclude if the two main factors, age, and size, are correlated to the risk score of an organization.

Cybersecurity deals with risk management, data protection, and network security. Hence, establishing the relationship between the various factors that affect a firm's risk, such as age and size, can benefit the cyber ecosystem. This research utilizes Python to analyze the effects of age and size of an organization on its risk score. Fixed Effect Panel OLS on Python accomplishes that.

1.3 Overview

We see that, as age increases, the risk score of a firm increases. Likewise, as the size gets large, the risk also gets bigger. Larger firms generally tend to be older, and much intellectual property is at stake. Hence, it is vital to understand the effects of characteristic features of firms such as age and size regarding the risk they possess.

CHAPTER 2

LITERATURE REVIEW

2.1 Information Technology Systems

The IT systems of firms are blessings and boons at the same time. There are several instances when these IT systems can be misused. For example, let us look at a straightforward IT concept of open source. When code is openly published, programmers review the code. They can find bugs in the code and report them so that the developers can fix them. There is also such a thing called a "patch," which is a set of changes that a person who finds a bug submits that the developers can directly use to resolve the issue. (Payne, 2002) An open-source software setup is used as a form of "peer-review" to evaluate the effectiveness of a system. (Payne, 2002) It especially comes in handy when some users try to evaluate the level of security the code offers before they choose to use it. Unfortunately, however, certain vulnerabilities come with open-source software. One such instance is called a Back Door. (Payne, 2002)

With the software code being completely open-source, attackers can easily insert malicious code that bypasses all the security measures. (Payne, 2002) For example, it can be done when they log in as a system administrator without needing a password if they connect from a specific IP address. An example of this would be security-sensitive software. For example, attackers hacked the FTP site containing Wietse Venema's TCP Wrapper software and tried inserting a "back door." However, the malicious code was discovered and fixed within a day. (Payne, 2002) If the code, in this case, were proprietary, it would have stayed vulnerable for much longer than a day. Open-source software also allows companies to implement their process controls. Such as installing in-house security auditing work, which would not be possible with proprietary (Payne, 2002).

2.2 DNS Cache Poisoning/Spoofing

A local copy of the DNS "phone book" is held to make the DNS query process faster. This phone book is called the DNS Cache (Open Web Application Security Project OWASP). It is prevalent in operating systems to store a temporary DNS cache database that contains a list of all recently accessed domain names and their addresses from the first time they were requested. When a user tries to browse a website, the computer looks through the cached DNS for the IP address (OWASP). Unfortunately, malicious actors corrupt the cached IP address from the local network, which can cause the query to return incorrect results (OWASP). In other words, these hackers can enter wrong information in place of the correct records, leading to redirection to a malicious website. Examples of DNS spoofing are below.

2.2.1 Malaysian Airlines

A hacker group redirected the visitors of Malaysia Airlines' official website to another site with malicious content. (Raghuvanshi, 2015) An example of DNS cache poisoning would be this hacking of the website of Malaysia Airlines in 2015. When users tried to access the website of Malaysia Airlines, they were redirected to a page with a picture of a lizard and a "404 - Plane Not Found" message. (Newcomb, 2015) The attackers used the Cached IP address in the DNS server to redirect to their site. Although Malaysian Airlines claimed that their systems were not hacked and the data was intact, the hacked website released a screenshot of a passenger, Amy Keh’s travel plans. (Chicago Tribune, 2015)

2.2.2 Sony Pictures

In 2014, Sony Pictures was hacked due to a combination of weak passwords and server hardening with Sony Pictures entertainment. (Deadline, 2015) As a result, personal information of employees and their dependents was released to the public. It included emails between employees and information about executive salaries. Additionally, copies of unreleased films were released. The entity's operations were affected due to the release of personal information as there was a class-action lawsuit against the company for failure to maintain reasonable security measures to protect their information. (BBC, 2015)

2.2.3 Organizational Impact

The organizational impact of the DNS Cache poisoning was significant in the case of Malaysian Airlines. The Malaysian Airlines website was down for 22 hours, preventing the users from getting Malaysian Airlines-related services online (Newcomb, 2015). In addition to that, the extent of data loss was unclear, and users’ personal data might have potentially been hacked. This was primarily a concern as Malaysian Airlines’ travelers had their credit card information, personal data, and travel plans stored on the website, which can cause a threat to their security and privacy, which eventually affected the organization’s bottom line (Newcomb, 2015).

Sony Pictures faced a loss of about $15 million to settle the case. This was in addition to a “max of $10,000 per individual plus around $1,000-$3,000 to the group of initial plaintiffs (Patten, 2016). With 435,000 class action members certified by Klausner last November, attorneys for the plaintiffs will are expected to receive $3.49 million”

(Patten, 2016). Lastly, Sony spent over $4 million to help employees protect themselves from identity theft. This also included the investigation and remediation costs. In addition, the personal information of employees and their dependents was released to the public.

Malaysian Airlines was 65 years old at the time of the hack, and Sony Pictures was 26 years old. According to these examples, we have reason to correlate the age of a firm to the level of cyber risk. This is because of two main reasons. First, older firms are likely to stick to their legacy system (Indeed, 2015). This also increases the cost of an upgrade, which would lead to the eventuality of not upgrading the systems hence, making them vulnerable to risks. Second, with older firms, employees are trained to fit needs that could be outdated. (Ruff, 2020)

2.3 Distributed Denial of Service Attacks

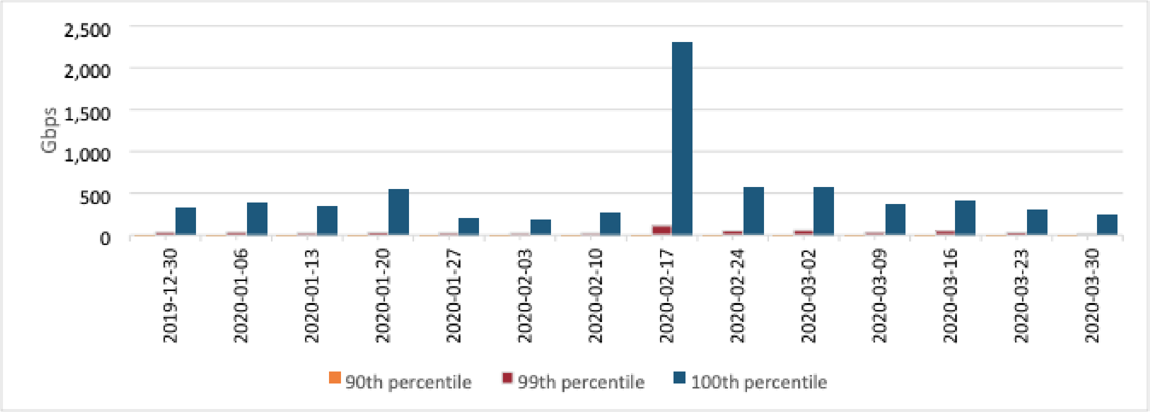
A Distributed Denial of Service (DDoS) attack is a type of Denial of Service (DoS) attack where the target server is flooded with network traffic until it becomes unresponsive (Cloudflare). This is caused by several infected devices on the network instead of the conventional single-origin DoS. A DDoS especially become hard to track down and mitigate as it does not have a single origin and instead has several hacked systems on the same network (Cloudflare).

2.3.1 Amazon

One reason for a DDoS attack is the Connectionless Lightweight directory access protocol (CLDAP). (Amazon, 2020) CLDAP is a connectionless alternative to Microsoft's Lightweight Directory Access Protocol (LDAP). Paul Nicholson explains that "CLDAP is designed to reduce the connection overheads at retrieving organizational resource information from a directory service database when using LDAP." While it is more efficient for larger organizations, CLDAP maximizes the scale of attacks. (Amazon, 2020)

After the first attack in Q1 2020, Amazon AWS Shield released the Threat Landscape Report, which included the Q1 2020 UDP reflection attack. (Amazon, 2020) This particular attack was directed at a single undisclosed AWS user. A CLDAP reflection caused the attack as they were hijacked to amplify DDoS traffic. With a magnitude of 2.3 Tbps, there was a three-day period of elevated threat during the attack. This was also about 44% larger than any other volumetric event detected by AWS. An attack of this magnitude towards just a single customer from a large organization such as Amazon AWS caused significant reputational damage.

Figure 2.1: Amazon Volumetric Events for Resources on AWS during Q1 2020



Prior to the Q1 attack in 2020, a different DDoS attack of a smaller magnitude took place in October 2019 (Spadafora, 2020). During this, the Amazon route 53 DNS web service was affected. AWS Shield could not fully mitigate the attack, and it affected the enormous customer base of AWS. Websites of thousands of their users were affected. AWS went offline due to the attack. The attack lasted for 8 hours. According to Whalebone (2021), this AWS attack was a Slow Drip Attack. "In this kind of attacks, a malicious actor continuously sends queries to the authoritative nameservers of the domain that they are targeting. The queries contain mainly non-existing pseudo-random subdomains. A direct result of this flood of queries is that the resources of the victim's nameservers are depleted, and eventually, they stop answering even to legitimate requests." This also caused some major reputational damage as that attack took place despite the presence of Amazon's very own DDoS mitigation service, AWS Shield (Amazon, 2020).

2.3.2 Organizational Impact

In 2019, Amazon AWS was under an 8-hour Denial-Of-Service attack that affected Amazon's web service. Due to the large size of the organization, Amazon could not entirely mitigate the attack, and it affected its enormous user base. This led to Amazon going offline for a few hours. (Spadafora, 2020)

In 2020, due to the usage of CLDAP within a large organization such as Amazon, AWS had to undergo a severe DDoS attack (Amazon, 2020). These two examples show that IT systems in place for larger organizations are risky; hence we will study the correlation between the size of an organization and its threat levels.

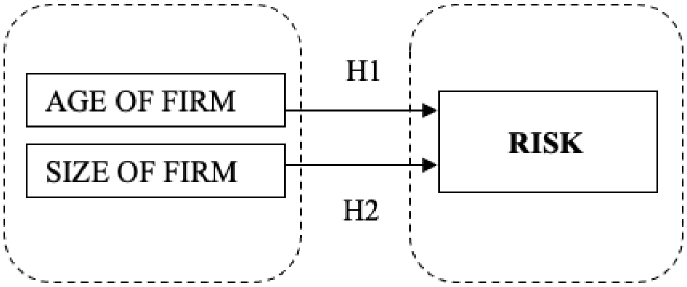
2.4 Hypothesis

Through the examples of Malaysian Airlines, Sony Pictures, and Amazon, it can be concluded that the age and size of an organization play an essential role in the cybersecurity landscape of the organization. So this research draws two main hypotheses based on these two characteristics of the organizations. First, as companies grow in age, they are more likely to have their legacy systems in place. This can be software or hardware. We see through Malaysian Airlines' example that the improper management of their system that caused the DNS Cache Poisoning attack might have been due to the age of the organization, which is about 65 years. Second, larger firms tend to use technology that puts the firm at a cybersecurity threat. We see that happen with Amazon's DDoS attacks due to the Connectionless Lightweight Directory Access Protocol. Thus, the following hypotheses are formed:

H1. Organizational age will have a positive relationship to the organizational risk score, making them more vulnerable to attacks.

H2. The organizational size will also have a positive relationship to the organizational risk score, making them more vulnerable to attacks.

Figure 2.2: Hypothesis



CHAPTER 3

METHODOLOGY

3.1 Cyber Green Initiative

We learn through the CyberGreen Metrics article about the importance of security metrics. We learn that the parties mainly affected by cybersecurity threats are the Defense Industrial Base and technology firms with global reach. Although these main targets might be able to protect themselves, oftentimes, the third parties or the secondarily involved parties receive little to no protection from Cybersecurity threats. Another key takeaway from this article would be that CyberGreen evaluates the risk factor that entities possess.

Yurie Ito, Executive director of CyberGreen, defines Cyber Health as “a condition of cyber systems and networks that are not only free from infection from malware and botnets but also contribute more broadly to the overall trust and usability of the cyberspace for the well-being of all.” This definition can be related to CGI’s goal of providing statistics on the risks that affect cyberspace.

CyberGreen Institute provides various services related to Cyber health. Some of these include developing statistical models to measure key risk indicators in organizations, detecting the vulnerabilities in a Cyber Ecosystem, and providing risk mitigation plans.

3.1.1 Risk Score

A significant active attack that affects various organizations is the Distributed Denial of Service (DDoS) attack. A Denial of Service (DoS) attack occurs when authorized users are hindered from accessing their systems. This is generally accomplished when hackers target the system by sending network traffic until the system becomes unresponsive. On the other hand, a DDoS is a much more severe case of Denial of Service. With DoS, only one origin floods the target system with network traffic whereas, with DDoS, several distributed systems flood the target system with traffic, making it unresponsive. DDoS also takes longer to be mitigated as the origin of the hack is not just one system so tracking it down becomes complex. Kindly note that there is such thing called an amplification factor which determines the multiplicity of a DoS or DDoS attack.

Various protocols can cause Distributed Denial of Service (DDoS) Attacks. The four main ones are Domain Name Service (DNS), Network Time Protocol (NTP), Simple Network Management Protocol (SNMP), and Simple Service Discovery Protocol (SSDP). It is important to note these protocols as CyberGreen further has data ranking the DDoS risk of entities/ enterprises. CyberGreen utilizes these four protocols as risk indicators to calculate the Index Score with the help of internet scan data.

The primary research method for this project was studying reports on existing technology that companies utilize. This was to see if the firm's size and age impacted the risk scoring. The research began with a foundational hypothesis which included initial assumptions from preliminary research findings.

A dataset containing records of 13,075 companies across the period of 2014-2020, extracted from the directory of records of autonomous system numbers from CyberGreen, was analyzed using Fixed Effect Panel Ordinary Least Squares.

Table 3.1: CyberGreen Dataset Variables

State	The state in which the firm is located in
ASN	The ASN ID of the firm
Year	The year in which the report is generated
Transit_addr	The number of downstream services that rely on this firm's ASN service
Originate_addr	The number of services that this firm operates.
sumscr_avg	The risk score evaluation of this firm, a higher score means higher risk.
Age	The age of the firm in years.

3.2 Python Data Analysis

The data analysis was conducted using Python 3.9.7 in the Anaconda distribution; the data was loaded and cleaned by imputing missing variables using average values and normalizing variables. The predictor variables, Age (Age of firm) and Originate_Addr (size or number of services the firm operates), were also standardized. The analysis used a Fixed Effect panel OLS in the scikit-learn module. The normalized variable risk score (std_risk) was used as the outcome (dependent) variable, with the main predictors being std_age (H1) and std_size (H2). Predictors are std_age and std_size, standardized variables from Age and Originate_Addr, respectively. The target variable is std_risk, derived from sumscr_avg.

3.2.1 Fixed Effect Panel OLS

In order to run a Fixed Effect panel OLS on Python, the dataset was treated as panel data. The panel data observed the various variables at several points, including the fixed effects. An example of a fixed effect would be periods.

CHAPTER 4

DISCUSSION

By running a fixed effect panel of ordinary least squares on Python, it was found that for a unit increase in standardized age, the risk score increased by 0.0256, and for an increase in size, the risk score increased by 0.1427.

Table 4.1: Python Panel OLS Estimation Summary

```
                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
const      -1.821e-17     0.0038  -4.82e-15     1.0000     -0.0074      0.0074
std_size       0.1427     0.0038     37.768     0.0000      0.1353      0.1501
std_age        0.0256     0.0038     6.7743     0.0000      0.0182      0.0330
==============================================================================
```

CHAPTER 5

CONCLUSION

In this study, we explored the effects of characteristic features of firms such as age and size with regards to the risk they possess. First, Malaysian Airlines' website was hacked in 2015 due to DNS spoofing. This is a phenomenon where hackers manipulate a cached copy of the Domain name system to redirect users to a malicious website. The hacker group also released travel plans for passengers. Secondly, Sony Pictures was hacked, which led to employees' personal information and dependents being released to the public. This included emails between employees and information about executive salaries.

Next, we saw that Amazon was under several DDoS attacks. One of them took place in Q1 2020, which was of magnitude 2.3 Tbps. Although this particular attack was mitigated, it caused tremendous reputational damage. This attack took place due to the size of the organization. This is because they used Connectionless Lightweight Directory Access Protocol is generally used to retrieve information more efficiently in large organizations. Based on the above examples, it was hypothesized that a firm's size and age negatively impact it. The larger the firm, the higher the risk evaluation, and the older the firm, the higher the risk. This is because larger firms are more prone to attacks; hence the recovery time can be much longer, and older firms are more likely to stick to their traditional legacy systems, which might not offer proper up-to-date risk mitigation. To test the hypothesis, a python data analysis was run with a dataset containing 13,075 companies from 2014 to 2020 retrieved from CyberGreen. A Fixed Effect Panel data OLS was used. Age and size were used as predictors, and the risk was the target variable.

The age and size variables were standardized, and we found that for a unit increase in age, the risk score increased by 0.0256, and for an increase in size, the risk score increased by 0.1427. This confirmed our hypothesis to be correct, concluding that older and/ or larger firms are more prone to risk. Several recommendations can be made to firms to protect themselves from Cybersecurity breaches, including investing in sound risk mitigation systems and investing in proper employee training as it all comes down to people, process, and technology. Cloudflare's DDoS mitigation system and Amazon AWS Shield are a few examples of such protection services. In addition, employee training can include cloud security certifications such as AWS certifications and seminars on enterprise risk management.

REFERENCES

AWS Shield Threat Landscape Report – Q1 2020. AWS Shield. (n.d.). Retrieved April 22, 2022, from https://aws-shield-tlr.s3.amazonaws.com/2020-Q1_AWS_Shield_TLR.pdf.

Baker, M. (2016, May). *Striving for Effective Cyber Workforce Development*. Software Engineering Institute. Retrieved April 22, 2022, from https://resources.sei.cmu.edu/asset_files/WhitePaper/2016_019_001_473577.pdf.

Geer, D. (2016, October 21). *CyberGreen Metrics*. CyberGreen. Retrieved April 22, 2022, from https://www.cybergreen.net/img/medialibrary/CyberGreen%20Metrics%20v.2.pdf.

Hornyak, T. (2015, February 4). *2014 cyberattack to cost Sony $35m in IT repairs*. Computerworld. Retrieved April 22, 2022, from https://www.computerworld.com/article/2879480/2014-cyberattack-to-cost-sony-35m-in-it-repairs.html.

Indeed. (2021, July 15). *Why do companies use legacy software? (with definition)*. Indeed Career Guide. Retrieved April 22, 2022, from https://www.indeed.com/career-advice/career-development/legacy-software.

Ito, Y. (2014, November 17). *Improving Cyber Health through Measurement and Mitigation*. Cyber Green. Retrieved April 22, 2022, from https://www.jpcert.or.jp/research/GreenPresentation-20141117_en.pdf.

Lazarevski, B. (n.d.). *DNS cache poisoning attack*. Open Wed Application Security Project. Retrieved April 22, 2022, from https://owasp.org/www-pdf-archive/DNS_Cache_Poisoning(OWASP_GHANA).pdf.

List of Autonomous System Numbers. BGP Looking Glass Database. (n.d.). Retrieved April 22, 2022, from https://www.bgplookingglass.com/list-of-autonomous-system-numbers.

National Archives and Records Administration. (2015, January 2). *Statement by the Press secretary on the Executive Order entitled "imposing additional sanctions with respect to North Korea"*. The White House President Barack Obama. Retrieved April 22, 2022, from https://obamawhitehouse.archives.gov/the-press-office/2015/01/02/statement-press-secretary-executive-order-entitled-imposing-additional-s.

Newcomb, A. (2015, January 26). *Malaysia Airlines Hit by Lizard Squad Hack Attack*. ABC News. Retrieved April 22, 2022, from https://abcnews.go.com/Technology/malaysia-airlines-hit-lizard-squad-hack-attack/story?id=28489244.

Patten, D. (2016, April 6). *Sony Hack Class Action Settlement Gets Final Approval*. Deadline. Retrieved April 22, 2022, from https://deadline.com/2016/04/sony-hack-lawsuit-settlement-approved-class-action-1201732882/.

Payne, C. (2002). On the security of open source software. *Information Systems Journal*, 61–78.

Raghuvanshi, G., Purnell, N., & Ng, J. (2015, January 26). *Malaysia Airlines website hacked by group calling itself 'Cyber caliphate'*. The Wall Street Journal. Retrieved April 22, 2022, from https://www.wsj.com/articles/malaysia-airlines-website-hacked-by-group-calling-itself-cyber-caliphate-1422238358.

Ruff, T. (2021, March 25). *When legacy becomes loss: The true costs of legacy systems*. Userlane. Retrieved April 22, 2022, from https://www.userlane.com/enterprise-legacy-systems/.

Sony attack yields another employee lawsuit. Deadline. (2015, January 7). Retrieved April 22, 2022, from https://deadline.com/2015/01/sony-hack-another-class-action-lawsuit-1201342693/.

Sony pays up to $8m over employees' hacked data. BBC News. (2015, October 21). Retrieved April 22, 2022, from https://www.bbc.com/news/business-34589710.

Spadafora, A. (2020, July 17). *AWS hit by major DDoS attack*. TechRadar. Retrieved April 22, 2022, from https://www.techradar.com/news/aws-hit-by-major-ddos-attack.

The importance of cybersecurity in business. BBC StoryWorks. (n.d.). Retrieved April 22, 2022, from https://www.bbc.com/storyworks/chubb-future-proof/the-importance-of-cybersecurity-in-business.

Tribune wire reports. (2015, January 26). *Malaysia air site hacked, some customer data appears online*. Chicago Tribune. Retrieved April 22, 2022, from https://www.chicagotribune.com/nation-world/chi-malaysia-airlines-site-hacked-20150125-story.html.

Whalebone. (2021, September 8). *Route 53 under attack!* Whalebone. Retrieved April 22, 2022, from https://www.whalebone.io/post/route-53-under-attack#:~:text=On%20October%2024%2C%202019%2C%20Amazon,system%20of%20the%20S3%20buckets.

What is a DDoS attack? Cloudflare. (n.d.). Retrieved April 22, 2022, from https://www.cloudflare.com/learning/ddos/what-is-a-ddos-attack/.

Worrall, W. (2022, February 25). *Will 2022 be the year of the ddos attack?* Hacked.com. Retrieved April 22, 2022, from https://hacked.com/will-2022-be-the-year-of-the-ddos-attack/.

BIOGRAPHICAL INFORMATION

Aparna Narayanan majored in Information Systems during her undergraduate degree at the University of Texas at Arlington. She is interested in cybersecurity, technology risk, and is passionate about acting as a liaison between the business and the technical world. Aparna plans on pursuing a career in technology risk consulting.