University of Macedonia

School of Business Administration
Department of Business Administration

MASTER'S THESIS

A THESIS SUBMITTED FOR THE

INTERDEPARTMENTAL MASTER OF BUSINESS ADMINISTRATION

Analyzing and mitigating the financial impact of Cyberattacks on businesses through clustering

Dimitrios Vavatsioulas

Supervisor
Efstratios Livanis

Thessaloniki, Greece
September 2023

Acknowledgements

Many people helped and supported me while creating this Master's Thesis, which is why I would like to express my gratitude to them.

I want to begin by expressing my gratitude to my supervisor, Efstratios Livanis, for his invaluable guidance and support throughout the duration of this project. The knowledge, expertise, and encouragement that I received from my supervisor have been invaluable resources and have greatly contributed to the successful completion of this thesis.

I would also like to extend my gratitude to my Master's professors for their excellent lectures, insights, and guidance. Their passion for their subjects and the dedication to their students have had an important impact on my academic journey.

Finally, I want to express my gratitude, once again, to my family and parents for their unconditional support during my academic journey. Their belief in my abilities and constant encouragement have been a driving force behind my success for my Master's degree.

Lastly, I would like to thank my friends for their continuous support and companionship throughout my Master's program, as their encouragement was critical for the completion of this Thesis.

Abstract

This study aims to highlight the financial impact of cyberattacks on organizations and to identify patterns and trends in the data using clustering techniques. The growing trend and impact of cyberattacks has made it increasingly important for organizations to understand and manage the financial risks associated with such attacks.

To achieve the research objectives, the study will first conduct a literature review on this subject and will examine the financial consequences that cyberattacks have. Next, a framework is presented, the Cyberattack Financial Analysis framework, which has four distinct modules and each one performs a seperate process. This framework is able to provide a theoretical concept which could be used in order to analyze and categorize cyberattacks based on their financial impacts, by using clustering techniques, and then suggest adaptive response measures and cybersecurity strategies.

The results of the study will provide valuable insights about the general financial consequences of cyberattacks and will help organizations better understand the risks and costs associated with these attacks. The suggested framework offers a novel approach on how to analyze and protect against cyberattacks. The findings of the study will also be useful for policymakers and researchers interested in understanding the aforementioned phenomenon and in developing strategies to mitigate these threats.

Contents

List of Figures

Chapter 1

Introduction

1.1 Topic description

Businesses and governments around the world, are becoming increasingly worried about the consequences of Cyberattacks. Apart from the expenses of dealing with and bouncing back from an attack there can be costs that arise as a result, like loss of revenue or harm to a company reputation. The financial impact of a Cyberattack can differ greatly depending on factors such as the type of attack, the specific industry targeted and how well prepared and responsive an organization is.

The purpose of this thesis is to evaluate and comprehend the aspects that contribute to the financial impact of Cyberattacks using clustering methods. The current literature on the monetary impact of Cyberattacks and the application of clustering algorithms in cybersecurity will be reviewed. Then, we will propose the use of Cybercrime Financial Analysis Framework; a theoretical system which consists of four (4) discrete Modules, each one responsible for different jobs.

This study's findings will illuminate the financial harm of Cyberattacks and the elements that contribute to that cost and give a useful tool for analyzing and comprehending those threats. Researchers in the domains of cybersecurity and financial risk management, as well as professionals and policymakers in the sector will find the

study's conclusions valuable for mitigating the economical effect of Cyberattacks.

1.2 Purposes

The purpose of this thesis is to investigate the financial impact of Cyberattacks and to develop a methodology (the CFA Framework), which uses clustering in order to gain insights and to adapt defense strategies against Cyberattacks. Specifically, this research aims to:

- provide a comprehensive review of the existing literature on the financial impact of Cyberattacks and the use of clustering algorithms in cybersecurity.
- present a methodology for using clustering algorithms to group similar attacks based on their financial impact (CFA Framework).
- identify the factors that contribute to the financial impact of Cyberattacks and how these factors are different from industry to industry.
- provide a theoretical framework for analyzing and understanding the financial impact of Cyberattacks, which can be used by organizations and policymakers to better assess and manage their financial risks and also take measures and develop a cybersecurity strategy.
- contribute to the fields of cybersecurity and financial risk management by providing new insights and perspectives on the financial impact of Cyberattacks.
- provide businesses a useful framework for protecting against Cyberattacks and mitigate their financial losses.

1.3 Methodology

This study's methodology uses a structured approach that follows the "problem-theory-method" paradigm, which makes it easy to develop and explore the thesis issue.

A thorough examination of the financial impact of Cyberattacks and the creation of a framework for a security strategy are ensured by the combination of problem identification, theoretical underpinning, and methodological rigor.

The first issue is the difficulty in determining the financial consequences of Cyberattacks and developing strong defense measures. A comprehensive review of relevant research on Cyberattacks, clustering analysis, and financial effect assessment is then used to develop the suggested theoretical framework (CFA Framework).

This study's methodology is divided into stages. Regarding the Cyberattack Financial Analysis Framework, data collecting techniques are initially used in Module 1 to obtain information on cyber attack occurrences, while clustering analysis is used in Module 2 to group related attacks based on common traits. The financial impact evaluation is then being executed by Module 3 by calculating the direct and indirect costs related to Cyberattacks. Finally, Module 4 develops and suggests defense measures and adaptive response measures that reduce risks and improve cybersecurity resilience using the knowledge obtained from the other modules but also from the framework as a whole.

The research methodology used in this work allows us to examine the financial effects of Cyberattacks and the development of protection strategies. This technique ensures a thorough understanding of the subject and improves knowledge in the fields of cybersecurity and financial risk management.

Chapter 2

Background

2.1 Cyberspace and its derivatives

2.1.1 Cyberspace definition

The virtual environment that is produced as a result of the interconnectivity of computer networks, such as the internet and other communication networks, is referred to as Cyberspace[1]. All of the information, communication, and commercial transactions that take place online are included in this enormous and intricate field. The following are some of the fundamental components that together make up cyberspace:

- Hardware: computers, servers, routers, and other electronic devices are examples of hardware. Hardware refers to the physical infrastructure that underpins and sustains cyberspace.
- Software: the term ”software” refers to the programs and applications that are run on top of the hardware to enable communication, data processing, and other activities that take place in cyberspace.
- Data: it refers to any information that can be stored, processed, and communicated in cyberspace. This includes text, photos, audio, and video as well as other types of information.

- Protocols: the sets of rules and standards that govern communication and the exchange of data in the digital environment.
- Users: the individuals and organizations that access and use cyberspace for a variety of different reasons.

In today's world, Cyberspace is incredibly important as it enables socializing in new ways, gives us access to knowledge and resources that were once out of reach or difficult to find, and expands communication options. It has a major impact on our everyday life: it changes how we collect information, how we work together in e-society, and it creates new ways to communicate. Also, it helps businesses to grow, to improve their effieciency and therefore make more money.

Despite its many benefits, the online world also has a dark side to it. Committing illegal activities like hacking into systems, stealing someone else's identity, spreading fake news and engaging in cyberbullying are all factors that create a concerning problem. Offenders who operate on virtual platforms take advantage of anonymity and convenience offered by the internet to perform unlawful acts while exploiting vulnerabilities present within computer networks or personal devices. As technology continues to advance and become more integrated into our daily lives as humans, the cybercriminals also adapt to these changes, which makes it necessary to take constant protective measures.

In conclusion, people are able to interact with one another, communicate with one another, as well as access information and services all around the world, however it is crucial to be aware of the numerous threats that exist and take precautions to protect ourselves, our companies and our systems.

2.1.2 Cybercrime, Cyberattack, Cybersecurity, Cyberthreat, Cyberrisk

In order to mitigate Cyberspace's dangers, it is critical to examine and differentiate the theoretical entities that exist in Cyberspace.

Cybercrime:

Casey [2] defines Cybercrime as any criminal activity involving computers or networks. Cybercrime, in general, refers to any illegal conduct involving digital tools and technology, including computers, cellphones, networks, and ICT infrastructure, with the intention of harming other people or businesses. It comes in a variety of forms including hacking, cracking, device misuse, data modification, identity theft and more. In addition to these, there are other cybercrimes that include violence such as child pornography, cyberbullying and harassment.

According to the European Commission [3], three forms of criminal activity are included in the term cybercrime. The first involves conventional criminal behavior like forgery or fraud, but in the context of cybercrime, it explicitly refers to crimes committed across electronic communication networks and information systems. The second is the dissemination of illicit material through electronic media (i.a. child sexual abuse material or incitement to racial hatred). The third category consists of offenses specific to electronic networks, such as hacking, denial-of-service assaults, and attacks on information systems.

We can see that cybercrime has a variety of acts and has many aspects, making it challenging for authorities to prevent and address every cybercrime occurrence successfully.

Cyberattack:

The National Institute of Standards and Technology (NIST) defines Cyberattack as an attack via cyberspace that targets an enterprise or a person for the purpose of destroying or maliciously controlling their infrastructure, stealing sensitive data or destroying the integrity of it [4].

The difference from Cybercrime is that, as a general rule, Cyberattacks' main target is machinery, while Cybercrimes' main target is humans.

Cybersecurity:

In order to confront Cyberspace's dangers, experts attempted to create a broad, general framework that includes a variety of techniques and counter-measures.

Schatz [5] defines Cybersecurity as the strategy and measures used by businesses and governments to manage security risks and ensure the availability, confidentiality, and integrity of their digital assets. In order to provide the best protection for the current state of Cyberspace and its users, the concept includes guidelines, regulations, and a collection of precautions, technologies, tools, and training.

Cybersecurity experts must be constantly alerted to new threats that Cybercriminals are developing. Day to day, Cybercriminals are finding different ways to attack their target, and the people responsible for protecting that target must be able to respond, if not prevent, these attacks.

Cyberthreat:

According to the National Institute of Standards and Technology (NIST)[1], a cyberthreat is any circumstance or event that has the potential to have a negative impact on the operations of an organization (such as its mission, functions, image, or reputation), assets, people, other organizations, or the country through the use of an information system. This can occur through unauthorized access, the destruction of information, disclosure of information, modification of information, or denial of service.

Cyber-risk:

Cyber-risk is the possibility of damage or harm brought on by Cyberattacks or other online threats. Financial loss, reputational harm, and legal obligations are just a few of the significant consequences that might result. Strong security measures, staff training on how to recognize and prevent online dangers, and having a strategy in place to respond to cyber events are all crucial for enterprises and people to defend against cyber-risk.

Cyber-risk is an increasing worry for companies and individuals in today's digital world. As the quantity and sophistication of online threats continue to rise, it is more crucial than ever to be aware of the possible dangers and to take preventative measures. This can reduce the possibility for injury or damage and guarantee that businesses and individuals are better prepared to respond to cyber incidents.

2.2 Machine Learning

Machine learning is a subfield of artificial intelligence that focuses on the development of algorithms and models that allow computers to learn from and make predictions or decisions based on data, without explicitly being programmed to perform specific tasks. In other words, machine learning allows computers to learn without being explicitly programmed to perform specific tasks. There are two primary categories into which machine learning algorithms can be placed: supervised learning and unsupervised learning.

During the training phase of supervised learning, the algorithm is trained using a dataset that has already been labeled. This ensures that the correct output labels are attached to the appropriate input data. The algorithm is taught to map the input data to the output labels that correspond to those data, and it can then be used to predict the output labels for future input data that has not yet been seen. Classification, also known as the prediction of a categorical label, regression, often known as the prediction of a

continuous value, and sequence labeling are all examples of supervised learning tasks (predicting a sequence of labels).

Unsupervised learning is another form of machine learning in which the algorithm is not provided with any labeled training data. Instead, the program is tasked with determining the underlying structure of the data by employing various methods, such as clustering. Clustering is a method in which a dataset is broken up into groups, or clusters, on the basis of the similarity of the data points that are contained inside each cluster. Detecting anomalies, compressing data, and summarizing it are just some of the activities that may be accomplished with clustering algorithms. These algorithms are used to detect patterns and relationships hidden in the data.

2.2.1 Clustering

Clustering is a unsupervised technique in machine learning that involves dividing a dataset into groups, or clusters, of similar data points. The goal of clustering is to partition the data in such a way that points within a cluster are more similar to each other than to points in other clusters.

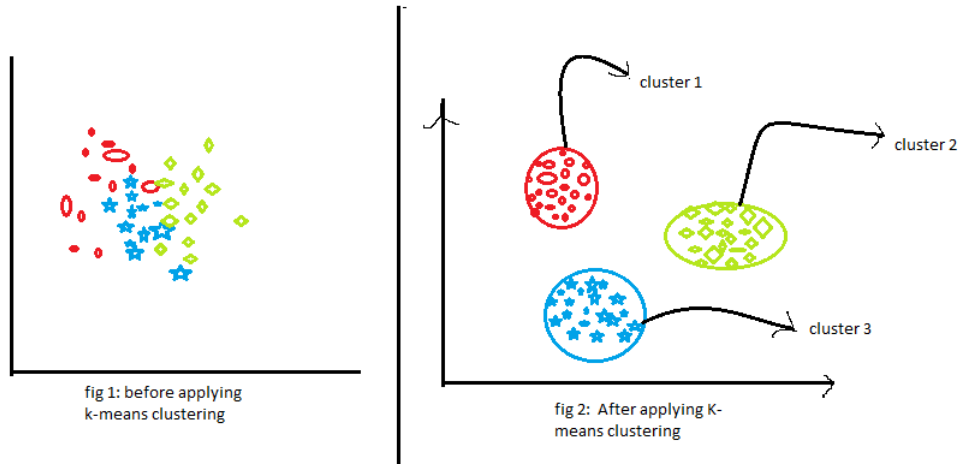

Figure 2.1: A simple overview of clustering process. First, the data points are unlabeled and scattered. After clustering, similar data points are grouped together and form a ”cluster”. Source: AnalyticsVidhya.com [6]

There are many different clustering techniques, but some common ones include:

- K-means clustering: This is a centroid-based algorithm, or a distance-based algorithm, in which the goal is to partition data into K distinct clusters. The algorithm works by first randomly selecting K centroids, and then iteratively assigning each data point to the nearest centroid and updating the centroids based on the mean of the points assigned to it.

- Hierarchical clustering: This algorithm creates clusters in a hierarchy where each cluster is layered inside of every other cluster. There are two main types of hierarchical clustering: Agglomerative (bottom-up) and Divisive (top-down). Agglomerative clustering joins together groups of individual points depending on their similarity. Divisive clustering separates the initial cluster of all points into smaller groups.

- Density-based clustering: This clustering technique works by locating areas with a high concentration of data points and expanding clusters around those areas. The technique can handle noise and outliers in the data and locate clusters of any shape.

- Spectral clustering: This graph-based clustering algorithm divides the data into clusters by constructing a similarity matrix of the data points and applying graph theory to the matrix. In non-linearly separable data, the approach is helpful for locating clusters.

Depending on the goals and purposes of the analysis, clustering can be used to draw conclusions from a dataset. Implementing clustering techniques, on a dataset could potentially lead to the below findings:

- Understanding the underlying structure of the data: By grouping the data into clusters, it is able to learn more about its structure and spot patterns and connections that might not be immediately obvious.

- Identifying groups or segments of similar data points: Clustering can be used to identify groups or segments of data points that are similar to each other and potentially belong to the same class or category. This can be useful for tasks such as customer segmentation, market analysis, and recommendation systems.
- Simplifying data: Clustering is a technique that can be employed to group points thereby simplifying and reducing the complexity of the data. This grouping makes it easier to comprehend and analyze the information.
- Improving the performance of machine learning models: The effectiveness of machine learning models can be enhanced by using clustering as a preprocessing step in their construction. For instance, clustering can be used to collect data points that are very similar to one another. After that, distinct models can be constructed for each cluster, which, in comparison to the construction of a single model for the entire dataset, can lead to superior results.

It's important to note that the conclusions drawn from clustering depend on the specific clustering algorithm and parameters used, as well as the quality and characteristics of the data.

2.2.2 Classification

A machine learning approach called classification involves putting data points into predetermined groups or classes based on specific traits or properties. It is a frequent activity in data analysis and has a variety of uses, including fraud detection, text classification, picture classification, and consumer segmentation.

Healthcare, banking, and cybersecurity are just a few of the industries that could benefit from the employment of classification algorithms. Based on certain characteristics, such as age, gender, and medical history, classification algorithms can be used in the healthcare industry to forecast the risk that a patient would contract a specific

disease. Algorithms for classifying data can be used in finance to spot fraudulent activity or anticipate credit risk. Classification may be used in cybersecurity to categorize occurrences and then carefully examine the data generated, allowing businesses and people to defend themselves.

Overall, classification is a potent tool for categorizing and comprehending data, and it has a wide range of real-world uses in several industries. It is a significant field of data science and machine learning research and is still undergoing active growth and innovation.

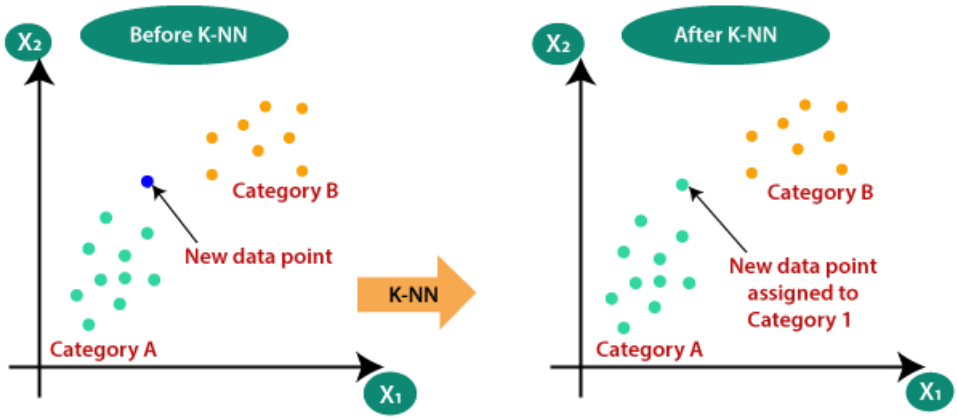

Figure 2.2: A simplified overview of classification process. First,the data point is unlabeled. After executing the classification process, the data point has a class label. Source: AnalyticsVidhya.com [7]

A summary of the three most popular classification techniques is presented:

1. Decision trees: A decision tree is a form of machine learning method that creates a decision tree by assigning weights to features. The algorithm begins at the root and proceeds to the branches by first asking a series of questions about the values of the characteristics, and then splitting the data points into distinct branches depending on the answers. Once a data point has been allocated to a class, the procedure is repeated recursively until all data points have been assigned to a leaf node. Easy to learn and comprehend, decision trees can process numerical and categorical information. In spite of this, they are not immune to overfitting

and may not always provide the best forecasts.

2. Using the probability of specific traits given a certain class, Naive Bayes classifies data points into categories. It relies on the ”naive” assumption that the characteristics may be treated separately. Under this assumption, the algorithm determines the likelihood of each class given the feature values and assigns the data point to the class with the greatest probability. Naive Bayes performs well on big datasets with numerous characteristics and is easy to build up and execute. It may not always produce the best accurate forecasts and it might be too sensitive if the data includes irrelative information.

3. Classifying data points according to the class of the K-nearest data points in the feature set is the goal of the K-nearest neighbors (KNN) instance-based learning method. KNN uses a distance measure between a data point and all other data points in the feature space to determine which K data points are most similar to it. After that, we look at the K nearest data points and use the data point’s majority class to decide its classification. Due to its intuitive nature and lack of training or data presumptions, KNN may be quickly and easily implemented. However, it may not function well on big or complicated datasets and may be computationally costly to operate.

2.2.3 Association rules

Machine learning algorithms like association rules can be used to explore big datasets in search of patterns of correlation between previously unnoticed variables. They have several uses, such as in recommendation systems, fraud detection, and text mining and are commonly employed in market basket analysis to determine which products are frequently bought together.

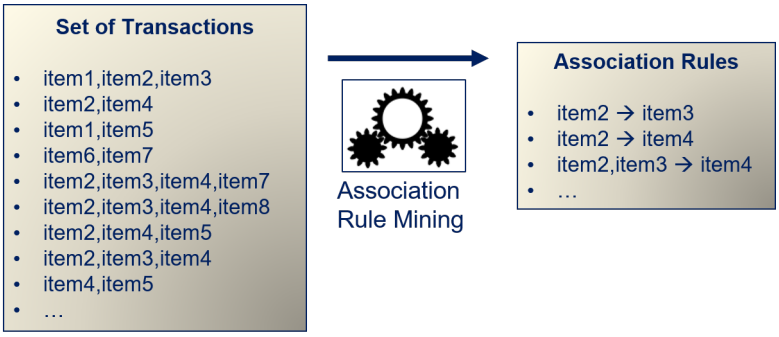

Figure 2.3: Simple example of Association Rules. Source: sherbold.github.io [8]

Association rules can be discovered using a number of different algorithms, including:

Apriori: Apriori is a classical algorithm for discovering association rules, and it takes a bottom-up approach by first generating collections of frequently occurring items and then deriving association rules from these sets. Statistically significant association rules are chosen based on a support threshold and a confidence level, respectively. Apriori works well and is simple to implement, however it may not be suitable for huge, high-dimensional datasets.

One such more recent technique is called FP-growth, and it employs a top-down strategy to find rules of association by way of constructing a compact data structure known as an FP-tree. It has a support threshold to find clusters of often occurring items and a confidence level to pick out rules of association with high statistical reliability. While FP-growth outperforms Apriori in efficiency and scalability, the FP-tree itself may be memory-intensive.

In contrast to Apriori and FP-growth, ECLAT employs a depth-first search strategy to create frequent item sets and association rules. It employs a support threshold to zero in on sets of frequently occurring items and a confidence threshold to zero in on statistically significant rules of association. While efficient and straightforward to deploy, ECLAT may struggle when applied to big, high-dimensional datasets.

As a whole, association rules are a powerful method for mining enormous datasets for hidden patterns of correlation between previously unnoticed variables. Distinct al-

gorithms may have different advantages and disadvantages, but they all use support and confidence criteria to find statistically significant groupings of often occurring items and association rules. Association rules are useful in many fields, from business and medicine to economics and public health, because of the insights they can provide and the way they can guide decision-making.

Chapter 3

Literature review

A thorough enough literature review will be offered in this chapter. This in-depth look at current research will cover the financial effects of Cyberattacks, cyber insurance, ways to build resilience against cyberthreats. Also, it will cover the use of ML algorithms and analytics, like clustering algorithms, to better understand and reduce the financial damage of Cyberattacks.

3.1 Cyberattacks and their financial impact

During the past few decades, Cyberattacks have become increasingly common and sophisticated, in part due to our growing reliance on digital systems and networks [9]. Cyber threats have progressed from relatively straightforward, isolated occurrences to massive, coordinated operations against essential infrastructure, businesses, and governments [10]. Attacks using ransomware, distributed denial-of-service (DDoS), and advanced persistent threats (APTs) have increased in frequency and have the ability to seriously harm the finances of targeted businesses [11].

Cyberattacks can have both direct and indirect financial consequences; direct having short-term financial losses and indirect including more difficult to estimate long-term expenses [12]. Theft of money, sensitive information, or intellectual property is one type of direct financial cost. Damage to one's reputation, a decline in customer con-

fidence, and the subsequent effects on future revenue sources are examples of indirect costs [13].

Also, firms may have opportunity costs as a result of downtime or the requirement to redirect resources from other initiatives in order to solve cybersecurity risks [14]. These indirect costs can be particularly significant for businesses because they may have broad-reaching and long-lasting effects, which may result in a loss of market share and competitive advantage.

3.1.1 The cascading effects of Cyberattacks on the economy

Cyberattacks can have wider economic consequences than only the organizations they target in terms of financial loss. For instance, a significant Cyberattack on a sector of key infrastructure, like the energy or financial sectors, might cause extensive disruption and significant financial losses, influencing other businesses and ultimately having an effect on national economies [15]. Cyber espionage can also damage a country's economic competitiveness and prospects for future growth by stealing intellectual property or trade secrets [16].

However, because of the expenses associated with Cyberattacks, funds that could be used for more beneficial investments could instead be diverted to cybersecurity measures, thus impeding innovation and economic growth [17]. This emphasizes the necessity for a comprehensive understanding of the financial effects of Cyberattacks and the creation of practical methods to lessen their effects on companies and the economy as a whole.

Accurately estimating the costs of Cyberattacks remains difficult due to the complicated and multidimensional nature of their financial impact. Several approaches, including cost-benefit analysis, event study analysis, and econometric modeling, have been put forth to calculate the financial effects of cyber attacks [18]. It is challenging to fully understand the financial impact of these methods since they frequently rely on insufficient or inaccurate data [19].

3.1.2 The role of cyber insurance in mitigating financial risks

Cyber insurance is a specialized insurance policy created to assist businesses in mitigating the financial risks connected with Cyberattacks and data breaches. Typically, these plans cover a variety of expenditures incurred during and after a cyber attack, including incident response costs, legal bills, public relations efforts, regulatory fines, and business interruption [20]. Since cyber risks continue to change and become more sophisticated, cyber insurance has become a vital component of firms' complete risk management strategies. Due to the continually evolving threat landscape and the difficulties in projecting the financial impact of cyber accidents, however, precisely estimating cyber risks and underwriting cyber insurance policies remains a difficult task [21].

3.1.3 Building resilience to minimize financial impact

Given the growing threat of Cyberattacks and their potential financial consequences, it is critical for organizations to build resilience against these incidents [22]R. esilience means spending money on strong cybersecurity measures and taking a proactive approach to risk management, which includes continuous monitoring, threat intelligence, and planning for how to respond to an incident [23]. By finding and fixing weaknesses, organizations can reduce the chances of successful attacks and the financial risks that come with them.

Training and awareness programs for employees are very important for building resilience, since human error is often a major cause of cybersecurity incidents [24]. Educating employees about the importance of cybersecurity, common attack vectors, and the best ways to keep sensitive information safe can help an organization develop a culture of security.

Collaboration between businesses, governments, and industry groups is also important for sharing information, resources, and best practices, which can improve cyber-

security and make the digital ecosystem more stable. Cyberattacks can cost businesses and the economy as a whole a lot of money. Public-private partnerships, information-sharing programs, and coordinated responses to cyber threats can help reduce this cost [25].

3.1.4 Leveraging advanced analytics to understand financial implications

Because it's hard to figure out how much money Cyberattacks cost, more and more people are interested in using advanced data analysis techniques, like clustering algorithms, to find patterns and insights that can help organizations and policymakers better understand and reduce the financial risks of Cyberattacks [26]. By finding groups of Cyberattacks with similar features and analyzing their financial effects, researchers can get a better idea of the factors that drive the costs of Cyberattacks and come up with targeted plans to reduce their effects.

By using clustering techniques, organizations can better categorize and understand the types of Cyberattacks they face. This lets them better use their resources and put in place security measures that are specific to their needs [27]. This approach is based on data, which can help make stronger cybersecurity plans that take into account the different financial risks that come with different types of cyber threats.

In conclusion, the financial effects of Cyberattacks are complicated and involve more than just the immediate costs that the organizations that are attacked have to pay. By understanding how these events affect the economy as a whole and using advanced analytics to learn more about what causes their financial effects, stakeholders can come up with better strategies to build protection and lower the risks that come with cyberthreats.

3.2 Overview of Clustering algorithms and their applications

Clustering algorithms are useful for analyzing data because they group things that are similar together so that similarities and relationships can be easily identified. One advantage is that they also work well with large datasets, so we can utilize them in order to extract useful information from raw big data. This section gives an overview of a few popular clustering algorithms and then we will see how they are used in different fields.

K-means is a well-known algorithm for dividing and grouping data points into a certain number of groups. It works by finding the center (mean) of each cluster and reducing the distance between data points and their respective cluster centers. This method is known for being easy to use and quick and this it is a good choice for big datasets. But the performance can be affected by where the cluster centers are put at first and how the data is spread out.

Based on how similar the data points are, hierarchical clustering algorithms group the data points into a tree-like structure. As a result, it is simpler to understand how the data points are related. There are two main types of hierarchical clustering: agglomerative and divisive. Agglomerative approaches begin with a single data point and gradually add more to it to create larger clusters. On the other hand, divisive approaches begin with a single cluster and divide it into more smaller clusters. Dendrograms generated by these algorithms tend to be effective at showing hierarchical relationships in the data and making them simple to understand.

DBSCAN is a density-based clustering algorithm that looks at how many data points are in a dataset to find groups. It groups points that are close to each other, making it possible to find clusters with various characteristics. DBSCAN can also deal good with noise and find outliers (extreme data points in clusters).

Clustering techniques can be used in a wide range of different fields. Clustering algorithms can be used in image processing to divide images into parts, sort objects into

groups, and compress images by reducing the number of colors used. Also, clustering methods are used in bioinformatics to look at gene expression data, sort proteins based on how similar their sequences are, and group organisms based on their evolutionary relationships [28].

In social network analysis, clustering algorithms are used to find communities within networks, find users who have a lot of influence, and look at how people are connected to each other. For example, clustering methods can be used to put people in a social network into groups based on their interests. This shows sub-communities and makes it easier to show targeted ads or recommend content [29].

Clustering techniques have been used in finance to look at stock market data, divide customers into groups, and spot possible fraud. Clustering can help investors see patterns in how stock prices move, which helps them make better decisions. Using clustering algorithms to divide customers into groups can help financial institutions make products and services that fit the needs of each group. This improves customer satisfaction and retention [30].

Clustering algorithms can be used to study and understand Cyberattacks by grouping similar events together. This can help security experts spot trends, find out how attacks are usually done, and come up with good ways to stop them. Clustering techniques can also be used to find strange things in network traffic or system logs, which could be signs of bad behavior [31].

In conclusion, clustering algorithms are a powerful way to find patterns and relationships that aren't obvious in a dataset. Their use in many different fields shows how flexible and useful they are as tools for analyzing data and gaining insight.

3.3 Using Clustering on Cyberattacks for financial assesments

Many sectors, including banking and technology, have utilized clustering techniques extensively to evaluate and group data sets with similar characteristics. In recent years, researchers have investigated the use of clustering techniques to evaluate the monetary impact of Cyberattacks.

A clustering-based method was proposed [32] to identify and classify Cyberattacks based on their financial impact. The authors utilized the K-means clustering technique to classify Cyberattacks according to their similarities in terms of financial loss, influence on business operations, and reputational harm. The results demonstrated that this method efficiently classified Cyberattacks into high, medium, and low financial impact categories.

Additionally, Dagoumas [33] demonstrates the use of clustering methods to assess the impact of cybersecurity attacks on power systems. The study applies clustering methods to evaluate the impact of different cybersecurity threats on the total operating cost and power grid adequacy. By clustering similar entities and analyzing their financial performance and reactions to Cyberattacks, clustering methods provide valuable insights into the financial impact of Cyberattacks and aid in developing effective strategies for mitigating their consequences.

We can see that the research indicates that clustering approaches can be utilized to evaluate the financial impact of Cyberattacks. By categorizing Cyberattacks according to their similarities in terms of financial impact, companies can have a better understanding of the potential dangers and deploy resources more efficiently. It is useful to highlight that the effiecincy of clustering algorithms depends on the quality and completeness of the dataset. Hence, future study should concentrate on enhancing the precision of the data and on researching novel clustering approaches that can more accurately ”understand” the complexity of Cyberattacks.

Chapter 4

Methodology

4.1 The Cyberattack Financial Analysis Framework (CFA Framework)

With the number and the complexity of Cyberattacks rising, companies not only have to worry about protecting their digital assets, but also about the possible financial consequences of these attacks. This chapter gives an overview of a comprehensive framework (CFA Framework) that uses advanced clustering methods to study the financial effects of cyber attacks and later develop a mitigating/defending mechanism. This framework aims to help cybersecurity experts and other parties make better decisions by combining the strengths of machine learning algorithms, data analysis and visualization, and statistical testing.

The main goal of the framework is to go beyond the usual focus on the technical parts of cyber threats and look at the financial costs that organizations have to deal with. The framework makes it simpler to understand how various forms of Cyberattacks impact finances by using clustering methods which group similar attacks based on their characteristics and patterns. This analytical approach provides decision-makers and cybersecurity experts with the tools they need in order to a) develop preventive strategies, b) make logical use of resources and c) develop defenses to reduce the after-attack cost

of Cyberattacks.

The system use machine learning algorithms to analyze big datasetes of Cyberattacks incidents, identify undetected patterns and incidents with similar financial impacts. The system provides a simple visual representation of attack clusters by using data analysis and visualization techniques. This makes it easier to make decisions based on accurate information. Statistical testing is used to figure out how important the financial effects of different attack groups are. This helps organizations decide how to respond and where to put their resources.

This framework gives insights that are useful not only for cybersecurity experts but also for other people in an organization, such as risk management teams, finance departments, and senior leadership. By putting a number on the financial effect of cyber attacks, the framework helps these stakeholders figure out how much they could lose, how best to use their resources and how to improve their organization's overall cybersecurity strategy.

The suggested framework gives organizations all the tools they need to analyze the financial effects of cyber attacks. It does this by combining clustering techniques, machine learning algorithms, data analysis and visualization, and statistical testing. By using this framework to help them make decisions, companies can protect their assets, limit their financial losses and become more resilient against Cyberthreats.

The CFA Framework is a theoretical system that can be applied to every company and organization that would like to protect its assets and minimize (or even prevent) the financial damage of a possible Cyberattack.

4.2 Modules

The CFA Framework consists of four (4) discrete components/modules. Each module is responsible for performing a specific job, through subprocesses, that will contribute to the main objective: mitigating the financial harm of Cyberattacks.

4.2.1 Module 1: Attack Pattern Identification and Analysis

Module 1 of the framework, ”Attack Pattern Identification and Analysis,” is all about finding and analyzing attack patterns, which is a very important task. This module is a key part of understanding the different types of cyber attacks and how they work. This helps organizations come up with good ways to defend themselves.

Overview

The first step of the module is to gather different kinds of data, such as security logs, incident reports, and threat intelligence feeds. These sources have a lot to say about past cyber attacks and the patterns they followed. The collected data is then put through steps to make sure it is ready to be analyzed.

Next, advanced data analysis techniques are used to get useful information from the data that has been collected. Methods like machine learning algorithms, statistical analysis, and pattern recognition are used to find patterns of attacks that keep happening. This analysis gives a better understanding of how different types of cyber attacks work by finding similarities and common methods used by threat actors.

In addition to the technical analysis, visualization techniques are used to show the attack patterns in ways that are easy to understand and give useful information. Visuals like charts, graphs, and diagrams help cybersecurity experts and other stakeholders understand complicated information and get insights that can be used.

Also, statistical testing methods are used to figure out how important and reliable the attack patterns found are. This step makes sure that the patterns found aren’t just a result of luck, but are instead signs of different attack behaviors.

The results of Module 1 give us important information about the traits, strategies, and techniques used by different kinds of cyber attacks. It makes it easier to figure out what attacks have in common and to put them into meaningful groups or clusters. These clusters are the basis for the rest of the framework’s modules. This lets defense

strategies be more targeted and resources be used more effectively.

By looking at attack patterns, Module 1 helps organizations better understand the threat landscape and make decisions about their defenses that are based on good information. With a better understanding of attack patterns, you can find potential weaknesses, take preventative steps, and make your cybersecurity stronger.

Vavatsioulas' thesis [34] presents a similar approach for creating a Classification System for Cybercrime. In short, the system receives feature vectors of Cybercrime incidents and then classifies them automatically, using Classification and Machnine Learning algorthims, into specific Cyberattack categories/classes. This helps the authorities to speed-up the process of collecting information about an incident and then responding to the incident.

In conclusion, Module 1 of the framework is all about finding attack patterns and figuring out how they work. Through data collection, preprocessing, advanced analysis techniques, visualization, and statistical testing, this module gives organizations valuable information about the characteristics and behaviors of different types of cyber attacks. The results of this module will be used to create targeted defense strategies in later modules, which will ultimately make organizations more resistant to cyber threats.

Processes

In this section, we'll look at the different steps that make up Module 1, which is about finding and analyzing attack patterns. This module uses advanced data analysis techniques to find patterns and behaviors that different kinds of cyber attacks do over and over again. Let's get into the details of how each process works.

1. Data Collection and Preprocessing: During the data collection and preprocessing process, different data sources, such as security logs, event reports, and threat intelligence feeds, are gathered. Its goal is to make a big list of all the cyber attacks that the group has been hit with. The collected data is then put through steps called preparation,

which include cleaning, transforming, and extracting features from the data. Cleaning up data means getting rid of noise, dealing with missing numbers, and fixing inconsistencies. Transformation methods can be used to ”normalize” or ”standardize” the data, making sure that the information from different sources is the same. Feature extraction is the process of finding useful information in the raw data that can be used for further research, such as the type of attack, the time it happened, or the systems that were affected. By doing these things, the data is set up for further research, which makes sure it is good and usable.

2.Statistical Analysis: Methods of statistical analysis are used on the attack data to find patterns and connections that are important. This means looking at the descriptive statistics of the dataset, like the means, medians, and standard deviations, to find out how the different variables are spread out and what their central trends are. Techniques for testing hypotheses like t-tests and chi-square tests can be used to figure out how important trends or differences between attack types are. With tools like the Pearson correlation coefficient, correlation analysis looks at the relationships between different factors, such as the type of attack and the business that was targeted. Time-series analysis looks at how attacks happen over time and looks for trends that change over time. These statistical analyses help us learn more about the attacks by giving us a better idea of their characteristics, trends, and connections.

3. Pattern Recognition Methods: Methods for recognizing patterns are used to find different attack patterns in the information. The goal of these methods is to find similar patterns, behaviors, or steps that different types of cyber attacks follow. Hidden Markov Models (HMMs) are a popular way to model the underlying states and changes in attack sequences. By training an HMM on attack data, it becomes possible to find common attack sequences and predict future states based on trends. This helps find attack sequences or mixtures of attack types that happen often. Pattern recognition methods make it easier to find attack trends that might not be easy to spot with manual analysis.

4. Visualization Techniques: Visualization is a key part of learning and making sense of attack patterns. Using data visualization methods, the attack data is turned into meaningful visual representations that help find patterns and figure out what they mean. Graphs, charts, and diagrams that show the attack patterns and connections can be made with tools like Tableau. Visualizations can include bar charts that show how often different types of attacks happen, scatter plots that show relationships between attack traits, or line charts that show how often attacks happen over time. Interactive visualizations let cybersecurity experts and other stakeholders explore the data, connect with visual representations, and learn more about patterns and trends. By using visualization methods, complex attack patterns can be explained in a way that is easy to understand. This helps with decision-making and makes it possible to come up with targeted defense strategies.

By going through these separate steps in Module 1, organizations can learn a lot about how attacks are planned and carried out. Experts in cybersecurity can use this technical analysis to find similar attack sequences, look at statistical trends, and see patterns that may not be clear from raw data. This information is the basis for making effective defense strategies, improving incident response plans, and allocating resources in a way that reduces the cost of cyber attacks.

4.2.2 Module 2: Clustering Analysis for Attack Grouping

The 2nd module of the CFA Framework focuses on clustering analysis to group cyber attacks that are similar based on their patterns and traits. Clustering is a powerful unsupervised machine learning method that lets us find groups or patterns in the attack data. By putting attacks that are similar together, this module hopes to give a better understanding of attack profiles, find common attack vectors, and make it easier to make more focused defense plans.

Overview

In Module 2, the process starts with choosing and pulling out important features from the attack data. These traits pick up on important things about the attacks that make them good for clustering. Feature selection makes sure that the clustering algorithm focuses on the most useful variables and feature extraction techniques help turn the high-dimensional attack data into a more easy-to-understand representation.

Once the features have been chosen and taken out, clustering algorithms are used to put threats that are similar together. These algorithms use different ways to figure out how similar threats are and put them into groups. K-means, hierarchical clustering, DBSCAN, and Gaussian mixture models are all examples of common clustering methods that could be used for this process.

After putting the attacks into groups, the groups are reviewed and validated to check for quality and consistency. Evaluation metrics and visual inspection methods help figure out how well the clustering algorithm works and give clues about how the clusters are put together and what their characteristics are. This process of evaluating makes sure that the groups make sense and can be understood.

To figure out what the clusters mean, you have to look at the characteristics, patterns and actions of each cluster to see what they have in common and what makes them different. By looking at the characteristics and features of the attacks in each cluster, cybersecurity experts can find the preferred attack vectors, businesses that are being targeted, or attack signatures for each group. This way of looking at the clusters gives a lot of useful information for making defense plans, finding weaknesses, and allocating resources to stop future Cyberattacks from happening.

Module 2 is a very important part of the system because it uses clustering analysis to group cyber threats that are similar to each other. This process helps us understand attack profiles better, makes it easier to find common attack routes and lets us come up with better ways to defend ourselves. By using clustering analysis in Module 2, businesses can learn important things about Cyberattack trends and become better at

protecting themselves and minimizing the potential harm.

Processes

1. Choosing and removing features: The first step of clustering analysis is to choose and remove important features from the attack data. These features should record important things about the attacks that can be used to group them together. Feature selection is the process of finding the most informative variables that contribute to the clustering process. Feature extraction techniques, such as Principal Component Analysis (PCA) or t-SNE (t-Distributed Stochastic Neighbor Embedding), turn high-dimensional attack data into a lower-dimensional representation that keeps meaningful information. We make sure that the clustering method focuses on the most important parts of the attacks by carefully choosing and extracting features.

2.Clustering Algorithms: Attacks that are similar can be grouped together using a number of clustering algorithms. K-means, hierarchical clustering, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), and Gaussian mixture models are all methods that are often used. These algorithms use different ways to figure out how similar threats are and put them into groups. K-means, for example, divides the data into a set number of groups based on how small the sum of squares is within each cluster. Hierarchical clustering creates a tree-like structure of groups with levels, while DBSCAN finds areas with a lot of data points. Gaussian mixture models are based on the idea that the data comes from a mix of Gaussian distributions. Which clustering method to use depends on the type of attack data and the structure you want for the clusters.

3. Evaluation and Validation: After the clustering algorithm has been used, it is important to evaluate and confirm the quality of the clusters that have been made. Metrics like the silhouette coefficient [35], the Davies-Bouldin index, and the Calinski-Harabasz index can be used to measure how close or far apart the groups are. These

metrics are numerical measurements of how well the clustering works. They help figure out the best number of groups and how well the clustering algorithm works overall. Validation also includes looking at the groups visually using tools like scatter plots, heatmaps, or dendrograms to learn more about their structure and how well they fit together. Validation makes sure that the clustering process gives data that can be understood.

4. Interpretation and Insights: Once the groups are set up, the next step is to figure out what the grouped attacks mean and what they can teach us. This means looking at the traits, patterns, and habits of each cluster to see what they have in common and what makes them different. Cybersecurity experts can learn more about the chosen attack vectors, targeted industries, or attack signatures for each group by looking at the features and characteristics of the attacks in each cluster. The interpretation of the clusters gives important information for making security plans, finding weaknesses, and allocating resources to make attacks less likely in the future.

4.2.3 Module 3: Financial Impact Assessment

Module 3 looks at how Cyberattacks affect a business's finances from a general point of view. Organizations can get a full picture of the financial risks that Cyberattacks pose by finding and categorizing cost components, analyzing financial data and qualitative factors, and finally coming up with ways to reduce those risks. This module gives companies the information they need in order to make decisions that protect their financial stability and resilience against cyberthreats.

Overview

In Module 3, we look at how Cyberattacks affect a business's finances. This lesson will help organizations understand the direct and indirect costs of Cyberattacks and figure out how much money they could cost them. By figuring out how much it will

cost, organizations can make smart choices, use their resources well and come up with plans to reduce the financial risks that Cyberattacks have.

The process in Module 3 starts with figuring out and putting together how much different parts of a computer attack cost. These cost components may include direct costs such as incident response and recovery costs, legal and regulatory penalties, and financial losses resulting from theft or fraud. Some examples of indirect costs are damage to a company's image, loss of customer trust, and less work getting done. By keeping track of and grouping these prices, organizations can get a full picture of how cyber attacks affect their finances.

Next, financial data and metrics are gathered and analyzed to figure out how much money cyber attacks have cost and could cost in the future. This study could include looking at financial statements, records of transactions, insurance claims, and other sources of relevant financial data. By looking at the financial effect, businesses can find out where cyber attacks cause the most money to be lost and how bad the damage is to their business operations.

In addition to the numbers, qualitative factors are also taken into account when figuring out what the bigger financial effects of computer attacks are. Some of these qualitative factors are market perception, customer churn, the name of the brand, and the long-term financial effects. Understanding the qualitative aspects helps organizations evaluate the intangible effects that may have long-term effects on their financial success.

Once the financial effects of cyber attacks have been studied, organizations can use the results to come up with plans to reduce and manage the financial risks. This could mean putting in place preventative measures, engaging in cybersecurity defenses, improving the ability to respond to incidents, and allocating resources to areas that are more likely to be affected financially. By handling financial risks ahead of time, organizations can limit the financial damage caused by cyber attacks and make sure their business operations can keep running.

Overall, Module 3 is an important part of the framework because it looks at how cyber threats affect money from a theoretical point of view. By putting a number on the financial effects and evaluating them, organizations can come up with plans to reduce the risks, use their resources well, and make decisions that are good for their finances.

Processes

1. Cost Identification and Categorization: The first step of the financial effect analysis is to find and classify the different parts of the cost of cyber attacks. This means taking into account both the direct costs, like incident response and recovery costs, legal and regulatory fines, and financial losses from theft or fraud, and the indirect costs, like damage to the company's image, loss of customer trust, and lower productivity. Organizations are able to gain a complete picture of how the impact of Cyberattacks on their finances by keeping track of these costs and grouping them together.

2. Financial Data Collection and Analysis: The process of collecting and analyzing financial data includes getting relevant financial information and metrics to figure out how much Cyberattacks have cost in the past and how much they could cost in the future. This study could look at financial statements, records of transactions, insurance claims and other sources of importat financial data. By looking at the financial effect, businesses can find out where Cyberattacks cause the biggest financial damage and it can influence their business operations.

3. Quantitative Impact Assessment: This part of the study is all about figuring out how much money cyber attacks cost. It means giving monetary values to the cost components that have been found and estimating the direct and indirect financial effects. Organizations can figure out what the real and possible financial effects of cyber attacks are by using financial data and metrics. This helps people understand how big the financial risks are and helps them make decisions.

4. Qualitative Implications Evaluation: Along with the quantitative research, qualitative factors are also taken into account when figuring out the bigger financial effects of cyber attacks. Some of these qualitative factors are market perception, the name of the brand and the long-term financial effects. Understanding the qualitative aspects helps organizations evaluate the effects that may have long-term effects on their financial success and can get a better idea of how Cyberattacks affect their finances as a whole.

4.2.4 Module 4: Defense Strategy Development

Module 4 applies theoretical knowledge from previous modules to cyber attack defense tactics. This module helps create targeted cybersecurity resilience measures by analyzing attack patterns, clustering results, and financial effect evaluation. Module 4 helps firms protect their digital assets and financial well-being from increasing cyber threats by understanding attack tactics and financial consequences.

Overview

Module 4 focuses on the development of efficient defense tactics to counter cyber threats. The main objective is to utilize knowledge acquired from previous modules in order to create specific strategies that improve an organization's cybersecurity stance. This module facilitates the development of defense plans that enhance resilience and decrease risks by integrating insights from attack pattern analysis, clustering, and financial impact assessment.

The procedure in Module 4 starts by using the insights obtained from Module 1 in order to identify patterns and attributes of cyber attacks. By comprehending the evolution of attacks and the employed methodologies, businesses have the ability to customize their security measures in order to proactively defend against these techniques. The findings of the clustering analysis in Module 2 provide additional insights for the refinement of these techniques, since they reveal similar attack groups and associated

behaviors.

Based on the knowledge gained from Module 3, businesses evaluate the financial consequences associated with different defense methods. This requires understanding of the manner in which the suggested measures influence potential financial liabilities and related costs. Organizations can achieve an effective balance between security and financial stability by strategically matching their defense plans with the corresponding financial implications.

Module 4 provides guidance on the development of proactive defense plans by equipping individuals with an in-depth understanding of attack patterns, clusters, and the associated financial consequences. These tactics may include a variety of actions, including as the enhancement of security protocols, the improvement of employee training, the strengthening of incident response capabilities, and the allocation of resources towards sophisticated cybersecurity technologies. Module 4 enables a comprehensive approach to cybersecurity defense by customizing these methods to the organization's unique risk profile and cost considerations.

Module 4 assumes a crucial role in the process of transforming theoretical knowledge into practical defense methods. By incorporating the findings from the past modules, businesses have the ability to strategically distribute resources, enhance areas of weakness, and develop a robust cybersecurity framework. The process of developing a defensive strategy helps businesses in taking proactive measures to protect their digital assets and financial stability from the ever-changing realm of cyber threats.

Processes

Insights Integration: The process of integrating insights includes putting together the findings from Modules 1, 2, and 3 so that they fit with the organization's cybersecurity goals. This requires a deep knowledge of attack patterns, common behaviors found through clustering, and the financial consequences of different attack scenarios. In practice, this process connects theory analyses with real-world applications. By

putting these insights together, organizations can find critical attack vectors, vulnerabilities, and possible financial risks that need to be addressed in their defense strategies.

Strategy Formulation: Once all the information has been gathered, the next step is to create and rank defense strategies. This means turning theoretical findings into actionable steps that improve the organization's cybersecurity posture. It needs careful thought about available resources, technological skills, and possible financial effects. In practice, this process includes making a strategic road map that outlines precise steps to prevent, detect, and react to cyber threats. Strategies could include improving network monitoring, implementing multi-factor authentication, teaching employees on security best practices, and setting up processes for responding to incidents.

Risk-Benefit Analysis: As defense plans are made, a risk-benefit analysis is done to see if there are any possible trade-offs between security improvements and their costs. This means figuring out how well each strategy works in terms of lowering the risk of an attack and minimizing any possible financial effects. It also looks at how easy it will be to put the plan into action, how resources will be used, and whether or not business operations will be interrupted. In practice, this analysis helps organizations make smart choices about which strategies offer the best balance of risk and reward and fit with the organization's risk tolerance and financial needs.

Allocating Resources and Planning for Strategy Implementation: Once strategies have been ranked and risks have been evaluated, the next step is to assign resources and plan for strategy implementation. This process includes figuring out the people, technologies, and financial investments needed to carry out the chosen strategies effectively. Practical considerations come into play when groups decide how much money to spend, who will do what, and when. By making sure that the resources are allocated in line with the chosen defense strategies, organizations can make sure that their cybersecurity efforts are well-done and will last.

Monitoring and Adaptation Framework: The last step is to design a framework for constantly monitoring and adapting defense tactics. This means setting up ways to track how well measures are working, find new threats, and adjust tactics as the threats are evolving. In practice, this process includes regular assessments, data-driven analysis, and continuous feedback loops that help change defense strategies as needed. By keeping a dynamic approach, organizations can deal with new vulnerabilities and keep up a reliable and effective defense posture.

Chapter 5

Conclusions and future work

5.1 Conclusions

In this Thesis, a theoretical framework, the Cybercrime Financial Analysis framework, is presented, which assesses the financial consequences of Cyberattacks on businesses and also suggests an adaptable framework for analyzing and protecting against the financial impacts of such attacks. The CFA framework consists of four distinct components and each one serves a specific purpose. Clustering techniques are used by the CFA Framework, making it clear that Machine Learning and Data Mining could be used in order to mitigate Cybercrime and Cyberattacks.

This work also demonstrates the potential efficiency of a practical application of the CFA Framework in evaluating and mitigating the financial consequences of Cyberattacks. This framework provides a comprehensive analysis of the financial impacts of Cyberattacks and aids in the development of defense strategies. The study's findings, even on a theoretical basis, have significant implications for policy-making and business practices. To mitigate the financial repercussions of future attacks, businesses that are more vulnerable to Cyberattacks should allocate more resources to cybersecurity measures, defense strategies and employee education on protecting themselves against common attacks.

In summary, the CFA Framework, when implemented in practice, could serve as a valuable tool for assessing and comprehending the financial effects of Cyberattacks. The study highlights the significance of the integration of cybersecurity and financial risk management by offering new perspectives into the Cyberattack field.

5.2 Future work

While this study provides valuable insights on the financial impact of Cyberattacks and the theoretical development of the CFA Framework, there are several things that could be examined for future research. Some potential improvements or extensibility for future work are:

- Developing a real-world, practical, software tool based on the theoretical CFA Framework presented in this Thesis, which provides real-time financial impact assessments and suggests response measures to cyberattacks.
- Conducting further research on the effectiveness of the CFA Framework in different industries and company sizes.
- Exploring the use of various machine learning algorithms to improve the accuracy of the CFA Framework metrics (e.g. financial impact).
- Investigating the impact of Cyberattacks on a company's reputation.
- Examining the role of insurance in mitigating the financial impact of Cyberattacks and developing insurance policies that align with the presented framework.

References

[1] Rebecca M Blank. “Guide for conducting risk assessments”. In: (2011).

[2] Eoghan Casey. *Digital evidence and computer crime: Forensic science, computers, and the internet*. Academic press, 2011.

[3] Council of Europe. *The commission communication ”towards a general policy on the fight against cyber crime”*. 2007.

[4] Michael Powell et al. *Protecting Information and System Integrity in Industrial Control System Environments: Cybersecurity for the Manufacturing Sector*. Tech. rep. National Institute of Standards and Technology, 2021.

[5] Daniel Schatz, Rabih Bashroush, and Julie Wall. “Towards a more representative definition of cyber security”. In: *Journal of Digital Forensics, Security and Law* 12.2 (2017), p. 8.

[6] Analytics Vidhya. *A Simple Explanation of K-Means Clustering*. 2020. URL: https://www.analyticsvidhya.com/blog/2020/10/a-simple-explanation-of-k-means-clustering/.

[7] Analytics Vidhya. *5 Classification Algorithms You Should Know: An Introductory Guide*. 2021. URL: https://www.analyticsvidhya.com/blog/2021/05/5-classification-algorithms-you-should-know-introductory-guide/.

[8] Sherbold. *Association Rule Mining*. URL: https://sherbold.github.io/intro-to-data-science/05_Association-Rule-Mining.html.

[9] Kim-Kwang Raymond Choo. “The cyber threat landscape: Challenges and future research directions”. In: *Computers & security* 30.8 (2011), pp. 719–731.

[10] Scott J. Shackelford. “From Nuclear War to Net War: Analogizing Cyber Attacks in International Law”. In: *Berkeley Journal of International Law* (2013).

[11] Ponemon Institute. *Cost of a Data Breach Report 2021*. 2021.

[12] Sasha Romanosky. “Examining the costs and causes of cyber incidents”. In: *Journal of Cybersecurity* (2016).

[13] Alessandro Acquisti et al. “The Economics of Privacy”. In: *Digital Privacy: Theory, Technologies, and Practices*. Auerbach Publications, 2006.

[14] Ross Anderson et al. “Measuring the cost of cybercrime”. In: *The Economics of Information Security and Privacy* (2013).

[15] Karthik Kannan, Rahul Telang, and Yan Xu. “Market reactions to information security breach announcements: An empirical analysis”. In: *International Journal of Electronic Commerce* (2007).

[16] S.O. Alashi and Dhuha H. Badi. “The Role of Governance in Achieving Sustainable Cybersecurity for Business Corporations”. In: *Journal of Information Security and Cybercrimes Research* (2020).

[17] Kristin Finklea and Catherine A. Theohary. *Cybersecurity: Authoritative Reports and Resources*. Congressional Research Service, 2013.

[18] Lawrence A Gordon and Martin P Loeb. “Managing cybersecurity resources: A cost-benefit analysis”. In: *McGraw-Hill* (2006).

[19] Sasha Romanosky, Rahul Telang, and Alessandro Acquisti. “Do data breach disclosure laws reduce identity theft?” In: *Journal of Policy Analysis and Management* (2011).

[20] Richard S Betterley. “The market for cyber insurance”. In: *Journal of Applied Corporate Finance* (2013).

[21] Christian Biener, Martin Eling, and Jan Hendrik Wirfs. “Insurability of Cyber Risk: An Empirical Analysis”. In: *The Geneva Papers on Risk and Insurance-Issues and Practice* (2015).

[22] Paul Cichonski et al. *Computer Security Incident Handling Guide*. Tech. rep. National Institute of Standards and Technology, 2012.

[23] PwC. *The Global State of Information Security® Survey 2020*. 2020.

[24] Einar Albrechtsen and Jan Hovden. “Improving information security awareness and behaviour through dialogue, participation and collective reflection. An intervention study”. In: *Computers & Security* (2010).

[25] Stein Schjolberg and Solange Ghernaouti-Helie. “A global protocol on cybersecurity and cybercrime? An initiative for peace and security in cyberspace”. In: *Cybercrime and cybersecurity* (2010).

[26] Tie Li et al. “An Integrated Cluster Detection, Optimization, and Interpretation Approach for Financial Data”. In: *Ieee Transactions on Cybernetics* (2022).

[27] Md Rayhanur Rahman, Mahdavi-Hezaveh Rezvan, and Williams Laurie. “What Are the Attackers Doing Now? Automating Cyber Threat Intelligence Extraction From Text on Pace With the Changing Threat Landscape: A Survey”. In: (2021). DOI: 10.48550/arxiv.2109.06808.

[28] Rui Xu and Donald Wunsch. *Clustering*. John Wiley & Sons, 2009.

[29] Santo Fortunato. “Community detection in graphs”. In: *Physics Reports* (2010).

[30] Amir Mahmud Husein et al. “Combination Grouping Techniques and Association Rules for Marketing Analysis Based Customer Segmentation”. In: *Sinkron* (2022). DOI: 10.33395/sinkron.v7i3.11571.

[31] Iqbal H. Sarker. “Machine Learning for Intelligent Data Analysis and Automation in Cybersecurity: Current and Future Prospects”. In: *Annals of Data Science* (2022). DOI: 10.1007/s40745-022-00444-2.

[32] Zhangyao Zhu and Na Liu. “Early Warning of Financial Risk Based on K-Means Clustering Algorithm”. In: *Complexity* (2021). DOI: 10.1155/2021/5571683.

[33] Athanasios Dagoumas. “Assessing the Impact of Cybersecurity Attacks on Power Systems”. In: *Energies* (2019). DOI: 10.3390/en12040725.

[34] Dimitrios Vavatsioulas. “Extending the Cybercrime Incident Architecture with a feature-based Cybercrime Classification System (CCS)”. BSc Thesis. University of Macedonia, 2021.

[35] Helio Jorge Regis Almeida et al. “Is There a Best Quality Metric for Graph Clusters?” In: (2011). DOI: 10.1007/978-3-642-23780-5_13.