
Threat modelling with UML for cybersecurity risk management in OT-IT integrated infrastructures

Master of science in Technology Thesis
University of Turku
Faculty of Technology
Security of Networked Systems
May 2022
Ali Sharif

Supervisors:
Professor Jouni Isoaho
Dr. Ali Farooq

Acknowledgements

I am proud to thank my supervisors Professor Jouni Isoaho and Dr. Ali Farooq both from the University of Turku for their guidance in accomplishing this thesis and all my friends for their moral support throughout this work. I would like also to send my best regards to the EIT Digital master school organization to provide such an incredible atmosphere for the master students who are willing to study the masters in two different European countries as entry and exit year studies. I want to thank ResilTech company based in Italy to offer me such a precious internship position. I want to thank Dr. Francesco Brancati to be my industrial supervisor at ResilTech during my internship. I want to thank Professor Adam Nagy from the Eötvös Loránd University (ELTE) in Hungary also to be my entry year thesis supervisor and reviewer. Professor Seppo Virtanen from the university of Turku was a great motivator with his encouragement and moral support.

UNIVERSITY OF TURKU

Faculty of Technology, Department of Computing

Ali sharif: Threat modelling with UML for cybersecurity risk management in OT-IT integrated infrastructures

Master of Science in Technology Thesis, 76 p. Security of Networked Systems
May 2022

A strong cybersecurity threat management can provide a good security situation against malicious attacks designed to access, modify, delete, destroy or capture user or organization systems and sensitive data. In this work, first the issue of cybersecurity is described, then the common attacks of OT-IT integrated systems as target systems are examined. The concentration area of this thesis is about the security of OT-IT systems. The purpose of this thesis is to provide a Cybersecurity risk management solution fundamentally focused on detecting common cybersecurity intrusions which are widely being used by the malicious attacks to forcefully abuse or take advantage of precisely a computer network. The main idea of this project is to providing a solution which can help the cybersecurity experts of OT-IT companies to catch the abnormalities of the network practically by the time a pre-defined intrusion is being executed by an attacker, in order to give more defensive power against the possible threats. In chapter 3 There will be proposed model is designed with UML and SysML in Eclipse Papyrus software which is a great tool to model a system. Here, I presented a threat modeling detection system which is practically an IDS. Finally, the model will be implemented using the PCA methods and the SVM, which are part of machine learning techniques. The Intrusion Detection System is implemented and the results show the high efficiency of the proposed method.

Key Words: Cyber security, Intrusion detection system, UML, SysML, Support vector machine, principal component analysis, OT-IT

Contents

1	Introduction	1
1.1	Problem statement	2
1.2	Background... ..	3
2	Cybersecurity risk management	6
2.1	Cybersecurity in OT-IT SMEs... ..	8
2.1.1	Assets of OT-IT in SMEs... ..	10
2.1.2	Cybersecurity attacks in OT-IT systems	14
2.2	Risk Management Concepts... ..	16
2.3	Risk Management Frameworks... ..	22
2.4	System architecture	28

3	Threat modelling approaches	
3.1	System modelling languages (UML and SysML)	32
3.2	EMF as a modelling environment	39
3.3	Case study based on a proposed model.... ..	41
3.4	Cybersecurity Threat modelling	43
4	A model-based threat implementation	46
4.1	Framework architecture	52
4.2	Implementation	56
4.3	Testing and verification	67
5	Conclusion	71
	References	73

List of Tables

Table 1: Common Threats in the System of OT-IT.....	14
Table 2: Table 2. Risk management steps in the System of OT- IT.....	15
Table 3. Evaluation of proposed method.....	70

List of Figures

Figure 1: Integrate XML, UML and Java with the EMF framework.....	39
Figure 2: A simple view of the Ecore meta-model.....	40
Figure 3. Proposed model of our OT-IT case study.....	41
Figure 4. An Integrated Cyber Security Risk Management Approach for OT-IT.....	42
Figure 5. Use Case Diagram.....	43
Figure 6. Sequence Diagram.....	44
Figure 7. Activity Diagram.....	45
Figure 8. Proposed method.....	51
Figure 9. Dispersion of attacks in the database.....	52
Figure 10. Rate and percentage of attacks in the database.....	53
Figure 11. The depiction of primary data from two-dimensional space.....	57
Figure 12. Part a) Before PCA is applied Part b) After PCA application.....	58
Figure 13. Data in two-dimensional space.....	59
Figure 14. Linear classification in two-dimensional space.....	60
Figure 15. Classification in space a) one-dimensional, b) three-dimensional.....	60
Figure 16. Lines distinguishing two classes.....	61
Figure 17. Support vectors of two classes	61
Figure 18. Datasets with errors.....	62
Figure 19. Classification with a soft margin.....	63
Figure 20. Margin calculation process.....	64
Figure 21. View of converting a nonlinear to linear separator.....	65

List of acronyms

ACO	Ant Colony Optimization
ANN	Artificial Neural Network
CI	Critical Infrastructure
CPU	Central Processing Unit
DARPA	Defense Advanced research projects Agency
DBN	Deep Belief Networks
DL	Deep Learning
DMZ	Demilitarized Zone
DT	Decision Trees
SysML	Systems Modeling Language
ES	Evolution Strategies
GA	Genetic Algorithm
ICA	Integrity, Confidentiality and Availability
ICT	Information and Communications Technology
IDS	Intrusion Detection System
INCOSE	The International Council on Systems Engineering
IoT	Internet of things
IPS	Intrusion Prevention System
KNN	Bayesian K-Nearest Neighbor
MBSE	Model Based System Engineering
ML	Machine learning
NIST	The National Institute of Standards and Technology
OMG	Object Management Group
OT-IT	Operational Technology - Information Technology
OWASP	The Open Web Application Security
PCA	Principal Component Analysis
RF	Random Forest
RNN	Recurrent Neural Networks
SAE	Self Assembled Encoders
SME	Small and Medium-sized Enterprises
SVM	Support Vector Machine
UML	Unified Modeling Language
U2R	U2R User to Root
VM	Virtual Machine
XML	Extensible Markup Language
WWW	World Wide Web

1 Introduction

Today, we are witnessing the expansion of the presence of computers in all aspects of our lives. Increasing the number of Internet users considering people's familiarity with programs to have access to computer networks, and increasing the information available on the servers of organizations, the need to increase the security of computer networks has become crucial. Computer devices contain valuable information and resources, which must be protected against attackers. In recent years, we have seen many attacks on computer devices, as an example In May 2021, a cyberattack on the Colonial Pipeline, the largest fuel pipeline in the US disrupted fuel deliveries in twelve states for several days [1] and as other example we could mention other SolarWinds, a major US information technology firm, was the subject of a cyberattacks in early 2020, hackers secretly broke into Texas-based SolarWind's systems and added malicious code into the company's software system. The system, called "Orion," is widely used by companies to manage IT resources [2].

Maintaining the security of devices and their information is one of the biggest challenges for technology managers and even end users. Using of network security technologies can reduce this risk to some extent [3]. Various security devices such as firewalls, IDS and IPS have been built to detect cyber-attacks. The mentioned two examples of cyber-attacks are in a close relation with OT-IT integrated environments cyber threats and in this thesis, I took the OT-IT combined infrastructures as the target system to implement an intrusion detection system which is capable of intercepting and detecting some of the major threats which is being used by the malicious attackers as an objective of this work. The proposed system which is built on a cybersecurity threat detecting model based on a common threats database will offer a tool for the information security staff of OT-IT companies to stay one step ahead of possible attacks.

1.1 Problem Statement

The issue of network security is one of the topics that have been more or less discussed for a long time and its importance has been proven to everyone today. Therefore, every day, new security mechanisms and solutions are presented to the security community for use in the world of the Internet. It is interesting that none of these mechanisms can claim that it can completely prevent any sabotage and intrusion. Meanwhile, due to the influence of the Internet in all jobs and organizations, and the Internet of Things, which is growing rapidly, the importance of security in it has become clearer. Therefore, the issue of Internet security and IoT security especially in OT-IT integrated environments is one of the hot topics in the comprehensive security of the world, and providing new security solutions in the country can be very useful.

Risk management is always on the side of planning to deal with possible future events. Securing organizations and investments against risks and losses requires the formation of an intellectual and practical system, through which policy-making against risks is integrated [4]. Based on Cybersecurity risk management, the system supervisor should be able to identify existing problems, define and obtain structural analysis of them, and by collecting relevant and classified information, provide the most appropriate methods of risk reduction. The term maturity in risk management approaches the state of perfection or evolution or readiness, as well as the path to perfection, growth and development [5]. In this thesis, I will discuss the roles and position of risk management and the steps that will be presented to manage these Cybersecurity risks, and at the end, and at the final chapters I will be proposing a model-based Cybersecurity risk management solution and implement a portion of it with focusing on Intrusion Detection Systems over OT-IT solutions.

1.2 Background

Cyber security is at the root of technologies, processes and practices designed to protect networks. Computers, programs and data against attacks, damage and unauthorized access [6] these types of incidents make your performance more vulnerable to external attacks and hackers [4]. Without the right security strategy, irreparable things can happen. Attackers know how to find and exploit vulnerabilities, opening up gaps that cause huge systems to collapse [7]. Any expert hacker can bypass any simple defense [9]. As a company expands, cybersecurity becomes more difficult. For example, the attack level and the consequences of a large company is much severe than Small and medium-sized enterprises (SMEs) [10]. In a connected world, everyone benefits from advanced cyber applications. At the individual level, cyberattacks can take everything from identity theft, extortion efforts, and the loss of important information such as family photos. Everyone depends on critical infrastructure such as power plants, hospitals and financial services companies. Providing this and other organizations is essential to maintaining the functioning of society [11].

Risk management can be considered a process to maintain assets and revenue streams, and value engineering is a value management integral to risk management. An integration of value engineering and risk management creates synergy. The issue of information security in organizations has made the use of information security systems risky. If the risk management process in these systems is done properly, it can be successfully implemented. In general, risk management consists of three basic steps: risk identification, risk assessment, and risk reduction planning. In identifying and determining the amount of information system risk, problems such as the lack of statistical data cause incorrect values for information system risk to be calculated.

The purpose of cybersecurity management of any organization is to protect the tangible and intangible assets of the organization (software, hardware, information and communication and manpower against any threat (unauthorized access to information, risks from the environment and the system and the risks posed by users, to achieve these goals requires a coherent plan managing the information security system a solution to achieve these goals [10].

The optimized value index is obtained by multiplying the risk factor by the usual value index. This index facilitates effective senior management decisions by clarifying the advantages and disadvantages of each idea [14]. In fact, risk is an unknown circumstance or event that, if it occurs, has a positive or negative effect on at least one of the project objectives. Value is the ratio of work to cost. Value can be increased by improving performance or reducing costs. Value studies provide good opportunities to reduce the cost of longevity, improve quality, reduce manufacturing time, extend longevity, and sometimes a combination of these.

Managing security risks in an integrated OT-IT system is difficult and costly. In fact, if a new vulnerability or a new virus is identified, these results can be very costly. In addition [12], organizations need a systematic security risk management approach to provide a rapid and appropriate response to security incidents and to protect their assets. In addition, enterprise or individual users expect information systems to be secure, able to anticipate their risks, and their strategies to reduce those risks. Secure organizational information management has led to the development of better criteria for understanding the status of an organization's security attitude. On the other hand, risk management is one of the basic components of an organizational risk management process. It is based on security criteria for managing security risks [15]. The Information Technology (IT) merged with Operational Technology (OT) industry has developed greatly during the

second half of the last century. The core of IT industry has also become much closer and more integrated with the OT industry. This technology is available and largely integrated with modern society. A software solution can reduce costs, speed up and facilitate the security management process. The output of this thesis is to provide a model-based system for cyber security risk management in OT-IT infrastructure. I intend to provide an intrusion detection system to detect attacks and prevent information from being compromised. In the first chapter, the problem, goals and outputs of the research are stated. The second chapter deals with the basic concepts of cyber security, cyber security in small and medium organizations and the cyber security risk management framework. In the third chapter, threat modeling methods, system modeling languages (UML, SysML) are discussed. Chapter 4 also includes the proposed framework, implementation and evaluation.

2 Cybersecurity risk management

A Cyberspace is an interconnected network of IT infrastructures that includes the Internet, telecommunications networks, computer systems, and processors and internal controllers in major industries [16].

Firewall: A firewall can be a hardware device or a software program, or a combination of both. A good firewall can prevent intruders from accessing your system and prevent any information from leaving your computer without your permission. A firewall cannot directly prevent viruses from attacking, but sometimes it does prevent viruses from sending emails from an infected computer. In general, this system, by defining a series of rules in it, can prevent a series of known attacks and restrict access to the computer ports on the network.

Intrusion Detection Systems (IDS): software or hardware devices are that automatically monitor the flow of network traffic or a single host and accurately analyze existing traffic based on security issues and symptoms. Finally, they can send the necessary warnings to the system security administrators.

Elements of cyber security

Ensuring cybersecurity requires coordinating efforts in an information system that includes:

- ✓ Application security
- ✓ Information security
- ✓ Network Security
- ✓ Accident recovery / business continuity planning
- ✓ Operational security
- ✓ End user training [14]

IoT creates new avenues for technology, media and telecommunications businesses, creating a whole new business and revenue stream or providing an efficient consumer experience. But this creates new opportunities for all information that is compromised. Not only is more data being shared through IoT, but the more users use this data, the more sensitive information is being shared. As a result, the risks are symbolically greater [18]. Establishing cybersecurity is very difficult due to the nature of cyberspace. Cyber technology can certainly be used as a tool of conventional warfare to attack government agencies, financial institutions, energy and national transport infrastructure, and public morale, so insecurity in cyberspace is not just about insecurity in information systems. It includes all the infrastructures that are somehow related to information technology. The risks associated with any attack depend on three factors: the threats of who attacks), the vulnerability of how they attack, and the effects (what effects this attack has on risk management of information systems [19].

People who commit cyber-attacks generally fall into one or more of the following categories: criminals who commit crimes such as extortion for the purpose of making money, spies who seek to obtain information used by the government, or identity [20]. Private individuals, national fighters who have increased their ability to carry out cyber-attacks in support of their country, hackers who carry out cyber-attacks for non-financial and non-monetary reasons, and terrorists who act as a governmental or non-governmental welfare agent. Government-sponsored attacks. ICT systems are very complex and attackers are constantly proving vulnerabilities that occur in many places [9].

2.1 Cybersecurity in OT-IT SMEs

In fact, in the 21st century, some of the biggest cyberattacks have happened to small businesses. Security experts have recently found that small and medium-sized companies are much more vulnerable to cyber-attacks and hacker intrusion. Hackers are aware that large companies use more layers of security to store their information, and usually seek the help of skilled cybersecurity professionals, so they turn to small companies.

Small and medium-sized companies do not have as much information as large organizations. So why are such attacks targeted? There are several main reasons:

Valuable data: Hackers know that even small companies have useful data, such as personal medical information, credit card information, bank account information, or proprietary business information. Cybercriminals will certainly benefit from using or selling this information.

Power of computer systems: Sometimes hackers are only interested in using corporate computers. Using a DOS attack (DDoS), a large amount of false demand is deliberately sent to the target server to cause the server and systems to crash. Hackers can use systems for their own benefit.

Business Partners: Today, businesses conduct many transactions, supply chain management, and information sharing online. Because larger corporate networks are (though not necessarily) more difficult to penetrate, hackers can target smaller partners as a way to gain access to large corporate systems.

Liquidity: In the first place, hackers carry out cyber-attacks for personal gain. Some attacks are also intended to cause damage, but usually the motive is money.

Small businesses usually have more liquidity than large organizations and companies. Large organizations and companies have a dedicated team to deal with cyber-attacks. In many small businesses, the task of protecting the network against cyber-attacks is performed by an employee, which is likely to be one of the tasks of protecting the network against cyber-attacks. This makes small businesses vulnerable to hacker attack. This is a very important position that 100% of the time should be devoted to repelling an attack.

To have a security strategy, let's first have a clear vision of the current threats:

Phishing: Phishing usually works by prompting users to click on an email or URL that contains a virus. The message is sent to the victim, and the victim enters their sensitive and confidential information directly on fake websites that look exactly like healthy and legal websites. Phishing is becoming more and more complicated and it is very difficult to distinguish a fake message from the original. Especially when hackers target certain people and with prior knowledge send them messages that the victim cannot resist.

Ransomware: In Ransomware, hackers use a wide range of methods to target businesses, enter into users' computers, and start encrypting files after execution. These files are encrypted with complex patterns and are practically unusable for users. Victims must pay a ransom to hackers to regain access to their data.

Malvertising: Malvertising, or malicious advertising, is used to spread malware and is spread through reputable sites using ad networks. A network crashes after a user clicks on a seemingly legitimate ad.

Software vulnerabilities: Hackers exploit vulnerabilities in popular platforms such as WordPress, tools such as Java, or files such as HTML, PDF, and CSV to disable systems and networks.

2.1.1 Assets of OT-IT in SMEs

Small and medium-sized enterprises (SMEs) have a large share in the economy as a whole. The share of these businesses in the macro-economies is generally considered separately, and of course they have a great impact on the economy on a macro-scale. In general, small and medium-sized businesses may not have received as much attention from the security industry as large companies and organizations, but there are many needs that need to be addressed in order to maintain security in these businesses.

Small and medium-sized businesses generally avoid large capital expenditures, as these businesses generally lack the financial resources to easily cover security costs. Most of these businesses are young and looking to grow and prosper, so a significant reduction in their cash reserves could be potentially catastrophic. At the same time, these jobs need to protect their physical and cyber security as much as others. Here security appears as a monthly service at a low cost. The security department provides services to small and medium-sized businesses in the same way that you can buy a new cell phone, house or car in installments, by distributing costs over a period of time, helping more small businesses gain benefits from the latest technologies in the security industry. Another way to help these businesses is to contract with reasonable incremental costs. In this way, the supplier remains the owner of the installed systems after installing security systems in small businesses, but also earns a regular income by providing services over time. This method helps small businesses to benefit from the latest security systems without fear of unexpected capital costs. The biggest challenge is for small and medium-sized businesses to simplify the systems that protect their business by choosing a platform for security products. Business systems that are different in design may not only improve job security if used improperly on a small scale, but may also create more security problems for them.

Therefore, our experts believe that small and medium-sized businesses will benefit the most by using a single solution that meets all the security needs of small businesses. Smart business platforms, for example, that also have insight and intelligence, help small business owners better protect their assets and manage their business and employees using modern science. These systems provide you with very easy and convenient facilities that in case of cancellation, the trade will be provided automatically and the delivery of goods and services will be done safely through video verification. The systems also provide customer experience data on details such as customer activity, busy areas and sales queues.

Small and medium-sized businesses usually have to achieve great success with limited resources. When it comes to physical security, the choice of technology, such as CCTV, is often made with limited financial resources. In the end, many instead of preparing a strategic plan to upgrade their security system as a coherent solution, choose different solutions for access control and video surveillance, which may eventually be due to their basic security needs. Do not cover savings. This can put small businesses at greater risk, such as cyber-attacks. Because just one unprotected device is enough to endanger your entire system and businesses to put your entire system at risk.

Another important point is that small business owners think that because they do not have a billion-dollar deal, cybercriminals will not bother to attack their networks. However, statistics show that small and medium-sized business owners are actually more at risk of hacking, cyber-attacks, and even building theft. This can put their businesses at serious risk. Small and medium-sized businesses tend to choose products over the system, and as a result often face problems integrating security systems. They are more likely to buy security systems piecemeal, sometimes without looking at a coherent whole and with a much more detailed look.

Small and medium-sized businesses also face challenges such as limited budgets, shortages of in-house human resources and IT and administrative resources. That's why these businesses are far more vulnerable to cybersecurity and less protected. Also, the small business market has been severely disrupted during the Corona epidemic, making it even more challenging to take advantage of new security industry technologies.

Cybersecurity is a matter that should be taken seriously by all organizations. Because of the huge statistics of cyber-attacks on big and small organizations, organizations must be aware of the risks and be prepared to allocate resources to protect their security. In the following, I will mention 5 important tools and services that every organization should consider investing in to strengthen their cybersecurity.

1. Firewall

Firewalls observe network traffic and connection attempts and decide to allow them or not within the computer or networks. Although firewalls are useful, they have their limitations. Hackers have found out ways to deceive firewalls to allow faulty data and software. This means a program can bypass a firewall without issues. Regardless of these problems, firewalls are useful in detecting and preventing harmful attacks on business entities.

2. Antiviruses

If you do not have much experience in cybersecurity, you might think firewalls and antiviruses are the same, but this is not true. Having both of these elements is crucially vital to having a safe system.

Antiviruses usually warn about viruses and malware and can have extra services like scanning emails for bad attachments or links. Modern antiviruses do things like quarantine and eliminate prospective threats. There is a wide array of antiviruses available and determining an appropriate antivirus for your organization is easy.

3. PKI services

Many people associate PKI with SSL and TLS, password protection programs, server communications, HTTPS, and the lock sign in the browsers. While SSL is important in general websites and internal networks, PKI can solve some issues with cybersecurity and play important roles in the safety of an organization.

4. Managed identification services

As cyber-attacks and hacks have become more complicated their techniques and software have also become stronger. Organizations need to invest in stronger methods for protection. Today, just having a system that reacts to threats is not enough and attacks need to be identified before they cause issues. Cybersecurity has changed course from investing in technologies that try to prevent attacks to systems that detect security weaknesses and quickly respond to them. Addressing an ongoing attack in an IT network is much more destructive than preventing it from spreading.

2.1.2 Cybersecurity attacks in OT-IT systems

Table 1 represents the most common threats and the concerns they may arise:

Threat	concern
Improper validation of input to systems	<ul style="list-style-type: none"> • Buffer overflow • Code injection • XSS attacks • Redirect data
Permissions, access levels and controls	<ul style="list-style-type: none"> • Weakness or lack of access level control • Perform various operations without having access level • Activation of communication networks of control systems • Inadequate configuration security
Improper authentication system	<ul style="list-style-type: none"> • Access to control systems by bypassing authentication systems • Lack of authentication systems to perform some critical and sensitive operations • Use of user-side authentication systems
Cryptography	<ul style="list-style-type: none"> • Lack of encryption in sensitive information • Use of unreliable and vulnerable cryptographic algorithms
Manage access to systems	<ul style="list-style-type: none"> • Use a simple or default password • Failure to properly maintain a password to access systems
Weakness in the rules defined in firewalls	<ul style="list-style-type: none"> • Lack of complete knowledge of the configuration of industrial control systems networks • Lack of security of special and sensitive ports in network equipment

Table 1. Common Threats in the System of OT-IT

The following table 2 summarizes each step of the system creation cycle, and the steps required to implement risk management at each step:

Risk management	Stage features	Risk management activities
Step 1 Start with system planning	The need for an IT-OT system in the relevant field is identified	In the stage of determining the needs of the identified system risks, security needs and a security strategy for the system are considered
Step 2 Create a system	systems are designed, planned and created	The risks identified in this step help to analyze the system security issues and thus complete the architecture and design of the system
Step 3 Implement the system	IT is implemented taking into account the security features of the system	Assist in identifying system implementation needs by considering the hardware environment and deciding to address identified risks before the system becomes operational.
Step 4 Operation	The system starts its operation and is continuously adjusted by adding or reducing the hardware and software of the battery in the process.	Perform risk management activities to periodically upgrade the system whenever significant changes in the operating and production environment of the IT system join him.
Step 5 Protection and Assignment	Changes in hardware and software information or processes such as information transfer or hardware placement	Risk management activities for hardware or software components and ensuring that residual data is properly controlled and handled safely

Table 2. Risk management steps in the System of OT- IT

2.2 Risk Management Concepts

Risk management is always on the side of planning to deal with possible future events. Securing organizations and investments against risks and losses requires the formation of an intellectual and practical system by which risk-based policies are integrated. On this basis, the system should identify existing problems, define and obtain structural analysis of them, and by collecting relevant and classified information, provide the most appropriate methods of risk prevention, control and financing. According to the cases, the term maturity in risk management approaches means the state of perfection or evolution or readiness, as well as the path to perfection, growth and development. In this research, we discuss the roles and position of risk management and the steps that will be taken to manage these risks, and at the end, we take a look at the estimated costs and risk analysis.

Entering the information age and the wide and deep entanglement of various aspects of human societies from economics and business to politics and social relations with information technology and also providing services and mass distribution of knowledge and information in the context of the World Wide Web and its transformation The great bridge to the most important bridge of individuals, organizations and governments, has been the beginning of the creation of a new chapter in human civilization, a chapter in which information is the most fundamental component and the most vital element. Hence, new threats with a very different nature from the past, with increasing intensity and increasing complexity, attack the shaky foundations of the technical, human and organizational infrastructures of information and communication technology, with the aim of bringing the basic components to their knees. Considers information security, which includes confidentiality, accessibility and integrity. From this perspective, it is not far from the truth if we consider information security as one of the biggest challenges facing humans.

Thus, the art of risk analysis and its management knowledge are very important for the dynamic and leading organizations of the present era that operate in a competitive and challenging environment. Information technology is one of the most important and vital sources of maintaining and creating a competitive advantage for companies and organizations. This is crucial in terms of information security and thus maintaining a competitive advantage, and is the only way that either leads to security and continued presence in the market or leads to destruction and elimination by competitors. Some of the most important reasons for the importance and position of risk analysis and management in information security are:

- Information security is complex and costly, and the resources available to the organization are limited, and therefore, by conducting a comprehensive and accurate risk analysis of the organization's limited and scarce resources in the way of proper protection of valuable information assets and capital is spent and wasted. Resources will be blocked.
- The beginning of the failure of many information security schemes is the wrong protection of the right things or the right protection of the wrong things. Simply put, the plan fails because its designers did not bother to identify the organization's valuable assets and strategic information assets and acted on the organization's missions and goals without any knowledge or awareness of the threats and consequences of their occurrence.
- To design and implement cyber security solutions.

Therefore, proper understanding and identification of existing risks and the ability to use the data obtained from risk assessment and analysis can lead to the following:

- Effective and efficient design based on facts and deep knowledge of valuable capital and assets

- Intensity of the effect and consequence of threats on the goals and policies of the organization
- Select appropriate and cost-effective tools and procedures to contain and reduce threats
- Select appropriate risk management strategies such as accepting, avoiding, reducing or transferring risk

Risk management cycle

The perception of risk management presents the need for risk awareness in society and economic institutions. Only in the light of such awareness can the risk management system be organized and used in the community and all related institutions. Risk management tasks can be performed and managed according to the size and scope of each organization, by an organizational unit and under the supervision of the relevant manager called the risk manager.

To identify the types of risks that each organization is exposed to, we can search in specific ways among the resources within the organization and outside it. From the perspective of Resource risk management within the organization is everything that exists in any organization and can be exploited. After recognizing the risks of each organization, it is necessary to examine the impact of each loss on the entire organization, and this requires determining what is the probability of each loss occurring, and secondly, what amounts if it occurs. And how these amounts will affect the financial structure of the organization. In other words, the actions of the first stage (identification of risks) provide a set of data (raw information) about the risks that threaten the organization, and the next step provides the obtained weaves (raw data), Based on the foundation of the risk management system in the organization, they are classified and processed, which leads to the processing of measured, valid and prioritized information.

In summary, the turnover of risk operations can be considered as consisting of the following:

- Risk identification
- Identify and define needs
- Risk structure analysis
- Collection and classification of information
- Develop a strategy for dealing with risks
- Monitoring and follow-up

Basic risk management steps

Risk management is an integral part of an ideal management. Therefore, the implementation of risk management is important to provide appropriate and continuous services and organizational responsibilities. Creating a risk management framework can be effective in achieving organizational goals and integrated, uniform and standard risk management. In this section, we will cover the basic steps in the field of risk management and provide explanations about each section.

Step 1: Risk assessment

All activities related to the information security risk management process start from this step, which includes two steps:

- **Risk analysis**

Mapping vulnerabilities and threats can be included in this step. However, in addition to identifying the above, appropriate values should be assigned to various factors in order to calculate the probability of occurrence of risk and the effects and consequences of risk; In other words, using the above, the risk should be able to estimate.

- **Risk assessment**

Risk assessment is a process that helps to determine the importance of this risk for the organization, which is calculated by comparing the estimated risk, the output of the previous stage and the organization's risk criteria. In fact, in order to understand how an organization views the results of risk calculation and analysis, we must first determine what criteria or criteria are available for accepting or rejecting risk.

The organization can determine these criteria based on risk assessment methodology, numerically, linearly or matrix. Based on this, it can be determined from the organization's point of view, which risks are accepted and which are not.

Step 2: Deal with risk

In this step, the organization must determine what strategy and program it has to deal with unacceptable risks. The most important output of this section is the Risk Management Plan (RTP). In this plan, the organization, while prioritizing the risks, describes each of its measures to deal with the risks.

Step 3: Risk communication

At all stages of the risk management process, communication must always be considered as an important factor. That is, at each stage of the process, effective communication must be established between stakeholders, consultants, senior managers, asset owners, risk owners, or other stakeholders. For example, in the risk assessment step, all the named factors that have a significant role in the organization's decisions should be involved. On the other hand, senior managers have a pivotal role in deciding on the range and criteria of accepted risk, as well as approving and allocating sufficient resources to risk management plans.

Step 4: Monitor and control the risk

In this step, the most focus is on controlling and following up on risk management plans. So, at first glance, you need to make sure that the planned actions are in line with the set schedule. Also, the degree of risk reduction should be monitored based on initial expectations. Thus, the status of RTP projects should be frequently reviewed and monitored for timing, impact, quality, and optimality, and corrective action taken if necessary. These actions can include changes in how plans are implemented, review of risk scenarios, allocation of new resources, and the like.

Step 5: Review the risk

A project view of risk management is not correct as a process that has a starting point and an end point; the cyclical nature of this process indicates that it is not a one-step process and its characteristics are constantly improving. In other words, the necessary steps in the risk management process should be repeated at regular intervals and the results and experiences gained in each stage, in the role of input and feedback, the new stage should be used.

Planning to deal with risk

Information security risk management planning involves the process of developing programs to mitigate threats and deal with risks that have a higher priority than other risks in the risk assessment phase.

At this stage of risk management planning, maximum attention should be paid to those risks that are more important and effective. The following methods are commonly used to deal with project risk:

- **Prevention**

Plan for preventive measures to reduce the likelihood and prevent the occurrence of risks.

- **Reduce the destructive effects of risk**

If preventive measures are not effective, they should be replaced by planning corrective actions in order to deal with the negative effects of risk occurrence and also to control its effects. This program also includes the use of designs.

- **Attract negative effects**

For risks that have a lower priority and also, in cases where corrective actions do not have the desired effects, the absorption of negative effects and dealing with them should be foreseen in the plan. The risk management plan and plan should be set in such a way that the priority and urgency of the risks are considered. Also, in order to allocate resources and implement the necessary measures to deal with risks, the budget, schedule and planning of other projects for risk management should be considered.

2.3 Risk Management Frameworks

Implement a risk management plan

The risk management plan must be very precise and planned, and the following important factors must be considered in its implementation.

- **Proportionality to risks**

It does not make sense to have a lot of resources and time to deal with | Assign risks that have a lower priority or are less likely to occur. Because this issue leads to neglect of more important risks and not paying enough attention to important and effective risks can be very dangerous. As a result, planning and allocating resources to risks should be commensurate with prioritizing project risks and project size and importance.

- Cost-oriented

Excessive spending to deal with risks can be challenging for the whole project. Therefore, in the project risk management plan, special attention should be paid to the issue of cost Be realistic management.

- Be realistic

Developing theoretical and non-practical plans not only wastes time and resources, but may also divert the plan from its original path and create more problems even compared to lack of planning. Achieving a realistic plan requires paying close attention to limitations and assumptions.

- Scheduling

It is clear that in any plan, there must be a specific schedule and, in its implementation, one must be committed to the desired schedule.

- Project Manager

In large projects, the risk management plan must have a specific manager to be able to follow the plan process and ensure its proper implementation. Also, important risks must have a specific manager to take overall responsibility for risk management, project risk identification and response.

The inputs of the project risk management plan are:

- General risk management plan; In particular, determining who is to be responsible for developing the project.
- Risks that require a risk management program; As mentioned earlier, you should only plan for higher priority risks.
- Dependence of risks, which means the same characteristics of risks, such as the root and source of risk, signs and symptoms of risk, effects and threat, probability of occurrence, time of occurrence and severity of impact.

- Prioritize the risk in order to determine the amount of work that needs to be done for each risk and the resources required.
- Determining constraints is very important in preparing a risk management plan and the effects of these constraints on others. The most important limitations in the project risk management plan are:
 - the budget
 - Resources
 - Time limit in implementing the risk management plan)
 - Changes (changes that are not possible due to various reasons).

At present, information is the most important treasure of organizations and individuals, and the loss and even the slightest damage to it requires time, money and unimaginable labor to compensate, and sometimes threatens the working principles and existence of an organization. Slowly In this regard, information security management to improve security in the emergence and exchange of information, with the help of a management system based on standards and technical guidelines and correct management decisions [21], can improve the performance of the information and communication system. Therefore, the risks associated with this process must be controlled. Risk management is a good tool for risk control. Applying risk management and assessment methods based on value management has a tremendous impact on how organizations organize their activities in the field of information security. In risk management, the first and most basic step is to identify the risk. With the advent of the first information security management standard in 1995, a systematic approach to securing the information exchange space was formed. Appropriate early decision-making in information security risk management can reduce costs and facilitate risk control. There is a high level of uncertainty in the data set in security decisions.

Due to various limitations such as the occurrence of certain accidents, human mentality and economic considerations, access to quantitative data is difficult, and even if data are available, they are often inaccurate or unreliable. Cyber security management is implemented through standards and information security management systems in organizations. Also, project risk management is one of the major topics in project management that includes planning, organizing, monitoring and controlling all aspects of the project and includes risk identification, measurement, risk response development and risk response control. Therefore, risk identification and assessment play an essential role in prioritizing and providing the right solution for corrective and preventive measures. In this study, the introduction of value engineering in the study and interpretation of the difficulty of each risk for each organization, has strengthened the risk identification stage. This provided criteria that can be expanded as effective risks in different organizations.

The results of the study identified and evaluated the risks of organizational information security. Therefore, in order to reduce and control them, it seems necessary to provide management solutions based on risk and value management. On the other hand, risks that are less important than other risks for the organization in question, require less attention and time and money. Considering that the combination of risk and value engineering, in addition to reducing the time required for pre-study and risk and value analysis separately, will save more and money, organizations can use the risks introduced in this thesis work to strengthen information security systems for managers in information security risk management. In this way, the parts of the organization that need more attention, time and money are identified. The main emphasis of the risk management group is on the initial costs of information security, i.e., the costs of equipment and manpower.

While value engineering always examines the cost to the organization. Therefore, the combined use of these two methods in the project expands the scope of the employer in choosing the appropriate option. Given that the most important part of an organization's security is strengthening and preventing damage, because often any correction after the deterioration of information security in the organization cannot be very fruitful and effective, pay attention to risk management, risks the high and costs they impose on the organization, which in many cases can significantly reduce vulnerabilities or minimize the consequences of a threat.

Entering the information age and the wide and deep entanglement of various aspects of human societies from economics and business to politics and social relations with information technology and also providing services and mass distribution of knowledge and information in the context of the World Wide Web and its transformation The great bridge to the most important bridge between individuals, organizations and governments, has been the beginning of the creation of a new chapter in human civilization, a chapter in which information is the most fundamental component and the most vital element.

Hence, new threats with a very different nature from the past, with increasing intensity and increasing complexity, attack the shaky foundations of the technical, human and organizational infrastructures of information and communication technology.

With the aim of bringing only the fundamental components to their knees. Considers information security, which includes confidentiality, accessibility and integrity. From this perspective, it is not far from the truth if we consider information security as one of the biggest challenges facing humans. Thus, the art of risk analysis and its management knowledge are very important for the dynamic and leading organizations of the present era that operate in a competitive and challenging environment. Information technology is one of the most important and vital sources of maintaining and creating a competitive advantage for companies and organizations. This is crucial in terms of information security and thus maintaining a competitive advantage, and is the only way that either leads to security and continued presence in the market or leads to the destruction and elimination by competitors. Information security is complex and costly, and the resources available to the organization are limited, and therefore, by conducting a comprehensive and accurate risk analysis of the organization's limited and scarce resources in the direction of proper protection of valuable assets and information is spent and wasted Resources will be blocked.

Therefore, proper understanding and identification of existing risks and the ability to use the data obtained from risk assessment and analysis can lead to the following:

- Effective and efficient design based on facts and deep knowledge of valuable capital and assets
- Intensity of the effect and consequence of threats on the goals and policies of the organization
- Choose the right and cost-effective tools and procedures to contain and reduce threats
- Select appropriate risk management strategies such as accepting, avoiding, reducing or transferring risk

2.4 System architecture

Principles of risk management (RM) work has several different elements that are primarily noteworthy:

- Perform risk analysis including cost and benefit analysis.
- Implement, review and maintain support.

In fact, the main part of the risk management process goes back to these threats and how they are likely to occur, to do so formulas and terms have been developed under the heading "RM". The effect of potential threats is to clarify some of the principles of risk management to clarify this:

- Property

Includes product information process information resources and computing infrastructure, which can be identified under examples such as development information, support, product replacement, public credit, and estimated costs, which are directly related to the life of the organization. The loss of assets can jeopardize the key concepts of information security, confidentiality, integrity and availability.

- Threats and protection

Threat In simple terms, any potential presence that causes adverse effects and compromises the security of our organizational information resources, whether human or machine, is protection to control and reduce the risk associated with a particular threat or group of threats.

- Vulnerability

Vulnerability results from a lack of weakness of protection, and it should be borne in mind that a minor threat has the potential to become a larger threat.

- Percentage of asset value

Represents the percentage value of the loss of a particular asset due to a threat, such as the effect of losing some hardware or catastrophically losing all computing resources.

- Estimated frequency

Represents the estimate that a threat is expected to occur and the range of this value never exceeds (100) and is usually considered based on the probability of the event and the number of people in creating the error that occurs and the damage is very small.

- Risk analysis

This key step has the following three main steps:

- ✓ Risk quantitative analysis
- ✓ Risk qualitative analysis
- ✓ Asset valuation and protection process

Cybersecurity architecture (cybersecurity architecture, network security architecture or cyber architecture for short) defines the organizational structure, functional behavior, standards and policies of the computer network, which includes network features and security. Cybersecurity architecture is a tool to reduce the risk of cyber breaches and protect your assets against digital damage [22]. New infrastructure is facing new challenges which lead to cybersecurity, network security, and the formation of infrastructure operators. Security of the systems was the most important parameter in the past, but nowadays, cybersecurity is a challenging topic which aims to protect the systems from growing threats. The current level of science, knowledge, and budgets are not sufficient for providing the required satisfactory security in this field.

Collective features of cybersecurity architecture include the following:

Network elements

- Network nodes (computers, repeaters, hubs, bridges, switches, routers, modems, gateways, etc.)
- Network communication protocols (IMAP, HTTPS, HTTP, FTP, DNS, DHCP, TCP / IP, etc.)
- Network connections between nodes using special protocols
- Network topology between nodes (point-to-point, bus, star, ring or circular, mesh, tree, daisy chain, hybrid)

Security elements:

- Cyber security devices (firewalls, IDS / IPS intrusion detection / protection systems, encryption / decryption devices, etc.)
- Cyber security software (anti-virus software, spam software, anti-malware software, etc.)
- Secure network communication protocols (IMAP, HTTPS, HTTP, FTP, DNS DHCP, TCP / IP, etc.)
- Powerful encryption techniques (end-to-end encryption, zero-knowledge privacy, blockchain, etc.) [23]
- Application security threats

Regardless of whether the applications are created by a specific team or are prepared in the form of software packages, there are threats to them. These threats are mainly for the purpose of obtaining confidential information or making changes to the system, in order to impersonate, manipulate its data, or even change the homepage in order to destroy the credibility of the collection.

Security in applications includes the following:

1. Application security against known attacks on the web
2. Vulnerabilities and platform security vulnerabilities and application framework
3. Security problems in the mechanisms used in the application
4. Application security flaws due to problems with its operating logic
5. Vulnerabilities and security issues due to incorrect configuration of the web server to host the application.

3 Threat modelling approaches

3.1 System modelling languages (UML and SysML)

UML is a modeling language used to analyze and design object-oriented systems. UML was first introduced by Rational and has since been endorsed by many computer companies and industrial and software communities around the world, so that only a year later it was accepted as the standard modeling language by the Object Management Association. UML is a language for defining, depicting, building, and documenting software system products, commercial systems, and other non-software systems. In fact, the main purpose of using UML is to use its high descriptive power to model the software architecture. UML supports the production of object-oriented software and offers a number of diagrams for modeling various aspects of the system. Some of them represent static aspects of the system and others describe behavioral aspects.

Diagrams of UML include:

- User diagram
- Class diagram
- Sequential diagram
- Collaboration diagram
- Mode diagram
- Activity diagram

- Deployment diagram
- Object diagram
- Component diagram

In software engineering, UML is a modeling language with the possibility of designing the system in the form of graphs and graphic templates. And the programmer can do the coding with a broader view of the project with the help of the UML language. Software has also been designed and developed to facilitate the UML language process.

Features of Unified modeling language

So far, we have said that UML is a modeling language for describing features and documenting the implementation of a software system being developed graphically. Features of UML language help to understand, design and maintain software information of the application developer.

Some features of the UML language include the following:

- Facilitate communication between project members
- Ability to convert UML language models to other programming languages such as Java and C ++
- Reduce training costs and plan changes
-
- Support for high-level object-oriented concepts such as Collaboration, Pattern

UML chart

The UML language process is such that the behavior and structure of the system are represented by UML diagrams. Therefore, it can be concluded that when a programmer or developer wants to work on a software project, he can view a set of objects in the UML language using the types of UML diagrams, the types of

which we will refer to below. Slow analysis. In general, UML diagrams have the following two views on the model and description of the system

- 1- A structural or static view that refers to objects, operations, and properties.
- 2- Behavioral or expressive view that refers to the relationship and coordination between objects and showing changes.

Types of UML diagrams

1- Structural diagram

Class diagram: Class diagram in UML as one of the most widely used UML diagrams when building a system, it is responsible for depicting the object-oriented nature of the program under production or development.

Object diagram: To build a prototype project, the object diagram depicts the relationship between objects and shows a static view.

Component diagram: Components are divided into different groups according to their relationship with each other. The component diagram is responsible for depicting the relationship between the components of the system, which consists of classes.

Deployment and Distribution Chart: A description of all the hardware, other components, and system implementation environment can be used by the deployment team in the deployment and explanation diagram, and shows what the components are based on.

2- Behavioral diagram

Use case diagram: In fact, shows an overview of the system and specific goals, dependencies and capabilities are modeled.

Sequence diagram: As a subset of the interaction diagram, it shows the sequence and sequence of the flow of messages from one object to another, and because the interaction between the components is important in execution and implementation, the sequence diagram is very useful.

Collaboration diagrams: There are many similarities between sequence diagrams and collaboration diagrams, and the most important difference is the appearance of the diagrams. Unlike sequence diagrams, which show the actions of objects in chronological order, collaboration diagrams emphasize the relationship between objects.

State diagram: Different modes of an object can be viewed by mode diagrams and depicting the object life cycle for software developers.

Activity diagram: Description of the control cycle and step-by-step description of the workflow along with the links and topics that make up the system is done by the activity diagram.

SysML, a system modeling language, is used to support model-based system engineering (MBSE). The model-based system engineering approach refers to a set of standard system development methods that, through modeling software, Relevant methods and tools are performed for the entire life cycle of a system.

Model-based system engineering can greatly help manage complexity, while improving design quality and cycle time, by increasing communication between different development teams and facilitating the absorption of design knowledge and evolution. A standard and robust modeling language that is considered to be a vital enabler for model-based system engineering (MBSE). System Modeling Language (SysML) is a graphical and standard modeling language designed for process modeling in system engineering.

In particular, this language has been developed to describe, analyze, and validate complex systems in which a system may consist of different components such as hardware, software, or procedures. Since designing a complex system accounts for about eighty percent of the cost and time of a balanced life cycle, modeling based on the rules of a coherent language can be a good way to keep design principles in place while making changes.

SysML is an integrated language in the field of systems engineering and is in fact the counterpart of UML in software engineering. This language is developed based on standard graphical symbols and its design structure is object-oriented. System Modeling Language (OMG SysML TM) is a general-purpose modeling language that features, designs, analyzes and validates the system as possible. Includes hardware and equipment, software, information, personnel, methods and facilities, supports, SysML is a graphical modeling language with a semantic basis for presenting the requirements, behavior, structure and features of the system and its components. This is the case for systems modeling in a wide range of industries such as aerospace, automotive, medical care, etc.

SySML is based on UML (Integrated Modeling Language), which is a common modeling language in software engineering and is designed based on deep concepts of object-orientation. For several years, UML was also used to model physical systems, but there were problems with modeling real systems with UML, which is specifically software. The solution to these problems is SySML, which is not really a new language but an expanded form of UML to describe the model of physical systems. INCOSE set its goal in 2001 to create UML as a standard language for system engineering. Because SysML and UML are so closely related, their structures and concepts are very similar.

While SysML is a format and part of UML, it also removes some UML elements. The extension introduced for SysML is limited to stereotypes and several new diagrams, stereotypes can be defined in any UML tool, no special tools are required for SysML. Only new charts need special support; But these concerns are only at the chart level and not at the model. Nevertheless, SysML modeling tools can naturally support better language and performance. The figure shows the relationship between UML and SySML. It is clear from the figure that SySML has borrowed a great deal of UML structure.

Four thematic contexts are defined for SysML modeling. Structure, Behavior, Parametric, and Requirements, this classification is provided by the Object Management Group (OMG), which has produced specifications for SysML. The SysML modeling language includes different perspectives on a system, which are known as the four pillars of this modeling language. The first pillar refers to the requirements of a system and includes all its functional and non-functional requirements.

The second pillar provides the structure of the system taking into account all the different subsystems and connections. The third pillar considers the behavior of the system and includes the creation of functional activities, scenarios and different modes of the system. The last pillar includes detailed features, rules and physical constraints on the system.

A) Structural diagrams

Any basic unit used in the system. It is called a block and is the central element in SysML, which can be hardware or software, the block represents any of the higher levels of the system, subsystem or component of the system or even the environment. Blocks not only have structural features between system blocks and are:

- **Block Definition Diagram:** The core of SysML diagrams is that it shows the structural hierarchy of the system and system components and provides a static view of the system.
- **Internal block diagram:** Describes the internal structure of each block in the system and provides a design view.
- **Closed diagram:** Used to organize the model and express its structure, and provides a model management perspective.

B) Behavioral diagrams

Demonstrate dynamic behavior between system blocks and are:

- **Application diagram:** shows a high-level expression of functions resulting from the interaction between systems or their components and contains a practical view
- **Sequence diagram:** Shows the interaction between the collaborating parts in the system and includes the view of the interaction.
- **Activity diagram:** Shows data flow and control between activities and includes activity vision.
- **State machine diagram:** shows the transition between the states of the whole system or its parts in response to events and contains a status view.

C) Requirement's diagram

Indicates system requirements, their hierarchy, and the relationships that meet or review those requirements. These relationships allow requirements to be related to each other, as well as to system design models and test cases. Requirements diagrams act as a bridge between system models and common requirements management tools, providing a good view of requirements.

3.2 EMF as a modelling environment

EMF is a powerful framework and tool to build Java applications based on simple definitions of models. This tool actually integrates Java, XML and UML technologies. Figure 1 illustrates this alliance. In other words, if you have one of the three models, you can automatically generate the other models using the tools and features of the EMF framework. The model can be defined using a UML modeling tool or an XML schema or even by describing notes on Java interfaces. The programmer simply writes a subset of the model-related abstract relationships, and the rest of the code is generated automatically. Figure 1 simply shows the type of integration of XML, UML and Java with the EMF framework

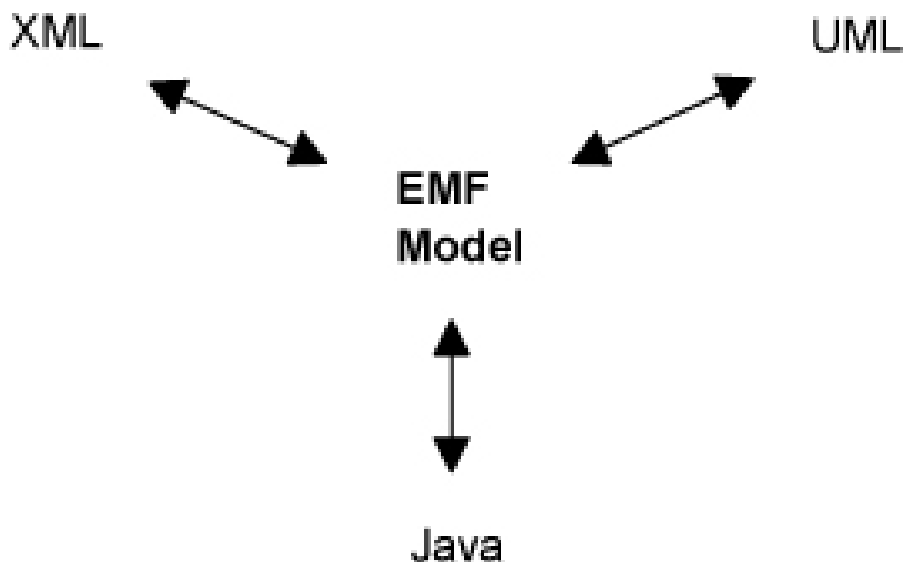


Figure 1: Integrate XML, UML and Java with the EMF framework

EMF is simply a set of plugins that can be used to model a data model and generate another code or output based on that model. EMF distinguishes between meta-model and model. The meta-model expresses the structure, while the model is an example of the meta-model.

The model used in EMF to display models is called Ecore. Ecore itself is an EMF model and therefore its own meta-model. A simple model view of Ecore can be seen in Figure (2). Figure 2 shows the elements required for model definition and meta-modeling with Ecore. As you can see, four Ecore classes are required to display the model:

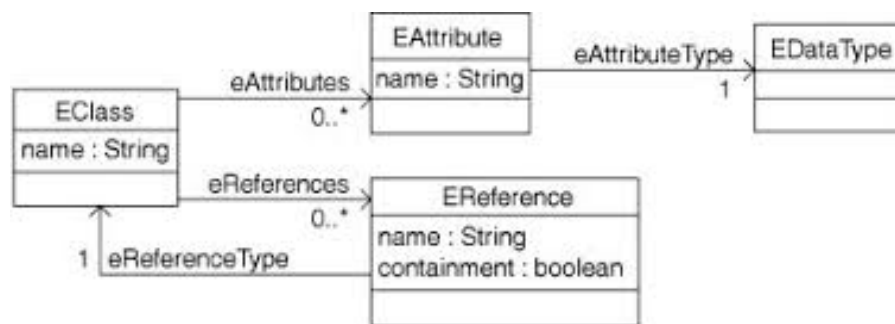


Figure 2: A simple view of the Ecore meta-model

- EClass is used to display a modeled class that has a name, zero or more attributes, and zero or more references.
- EAttribute is used to display a modeled attribute. Attributes have a name and a type.
- EReference is used to display one side of a 14-class association that has a name, a Boolean flag to indicate its inclusion, and a reference type (destination) that is another class.
- EDataType is used to display an attribute type. The data type can be a primitive type like int or float or an object type like java.util.Date[24]. It should be noted that Ecore is a small and simplified subset of complete UML, in other words, its class diagram. For example, full UML also supports modeling the behavior of an application.

3.3 Case study based on a proposed model

One of the most important challenges in networks is network security. In this study, we were looking for a solution to maintain network security. One way to maintain network security is through intrusion detection systems. In this thesis, I provide network security by providing a penetration detection system using support vector machine (one of the machine learning methods). The aim of this study is to thesis security by providing a high-performance intrusion detection system. Here, the Figure 3 depicts the proposed model based on an OT-IT integrated system joint with the case study and the final results.

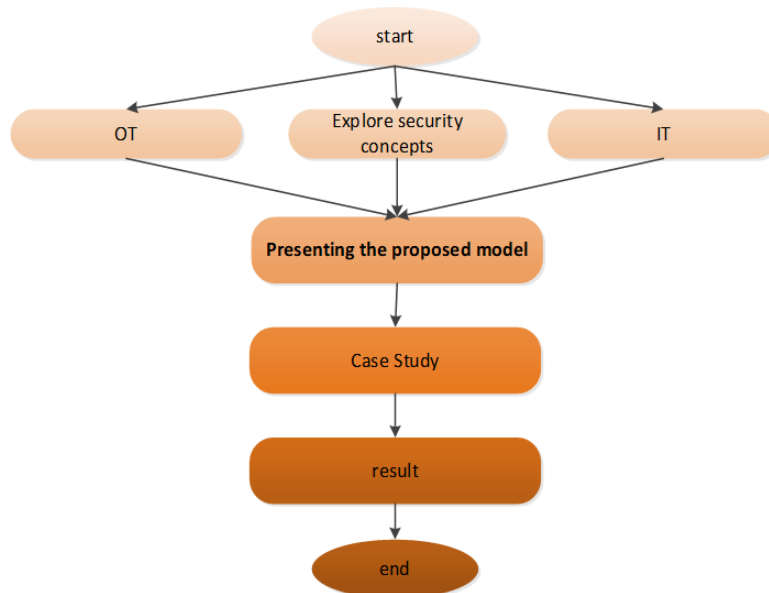


Figure 3: Proposed model of our OT-IT case study

Our path:

- Gather information by studying the available documents.
- Study and knowledge of the concepts of risk process management and search in the literature on cyber process management systems and extract all requirements and demands.
- Study and knowledge of architectural concepts and search in the literature on risk management architecture and its study.
- Study and study security concepts.
- Review of work done in the field of process management systems and architecture
- Provide a secure model for cyber process management systems in industrial and information environments.

Figure 4 shows the relations and the connection path of the elements of an accepted cybersecurity risk management approach based on OT-IT systems.

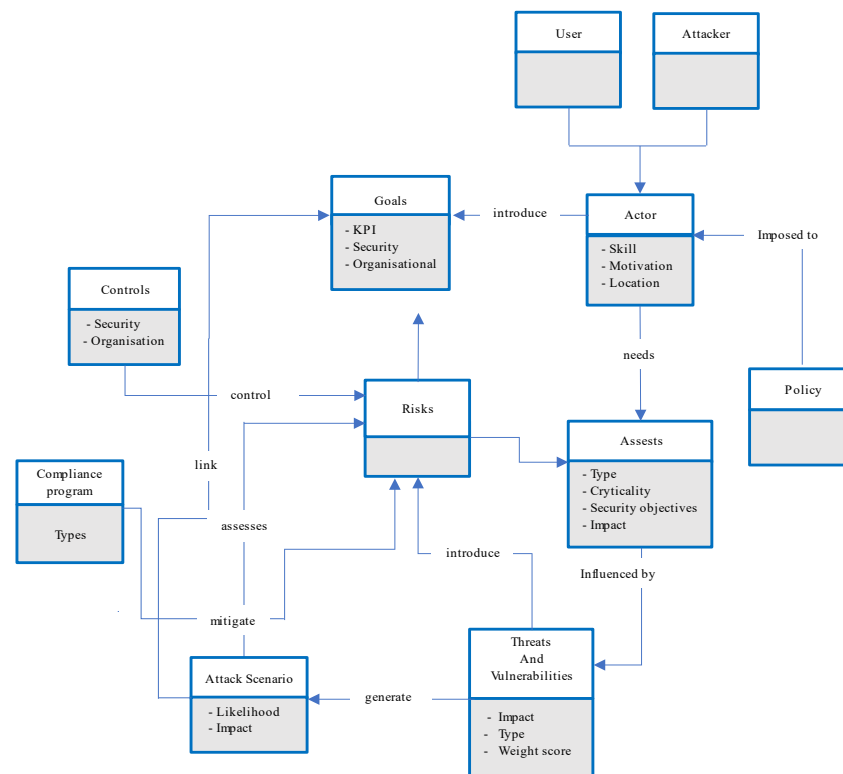


Figure 4. An Integrated Cyber Security Risk Management Approach for OT-IT integrated systems

3.4 Cybersecurity Threat Modelling

In this work, using Eclipse Papyrus, the diagrams required to design the intrusion detection system have been drawn. Figure 5 shows Use Case Diagram, Figure 6 shows Sequence Diagram and Figure 7 shows Activity Diagram.

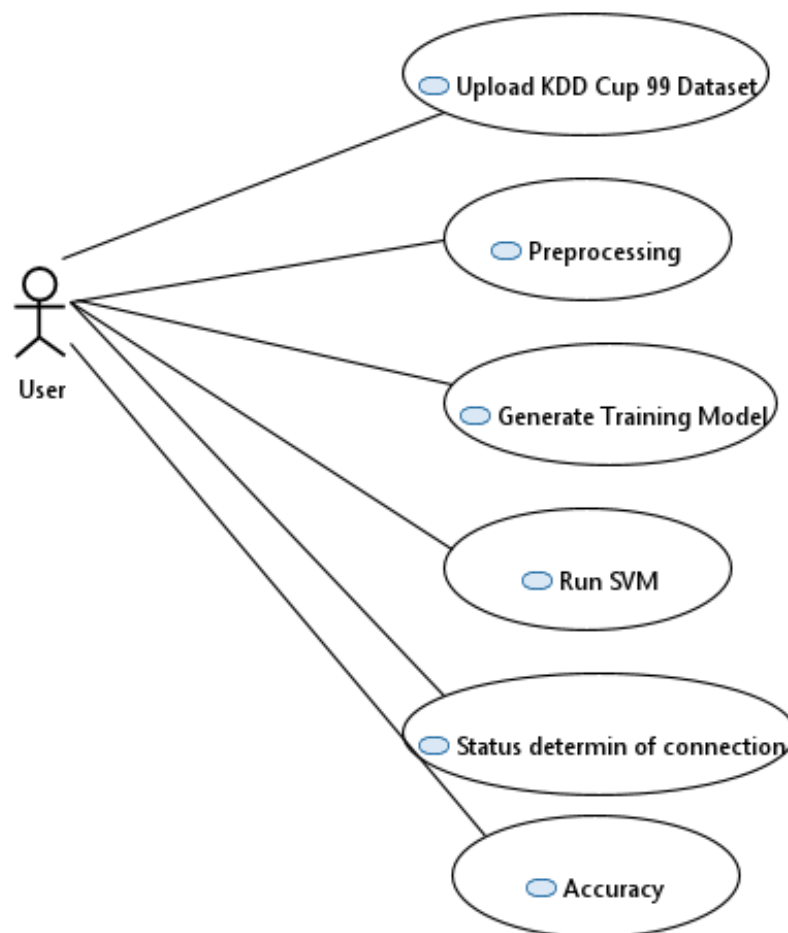


Figure 5. UseCase Diagram

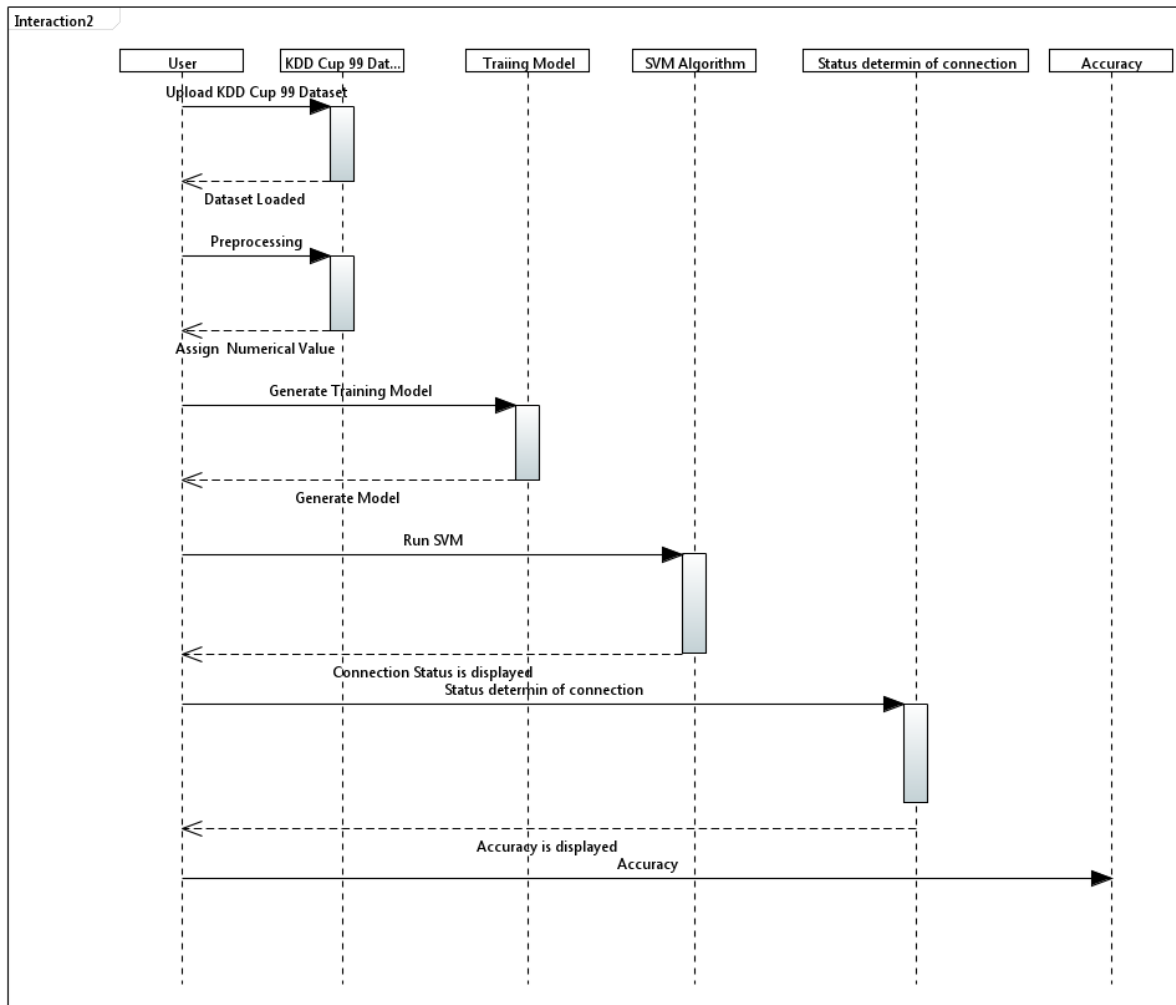


Figure 6. Sequence Diagram

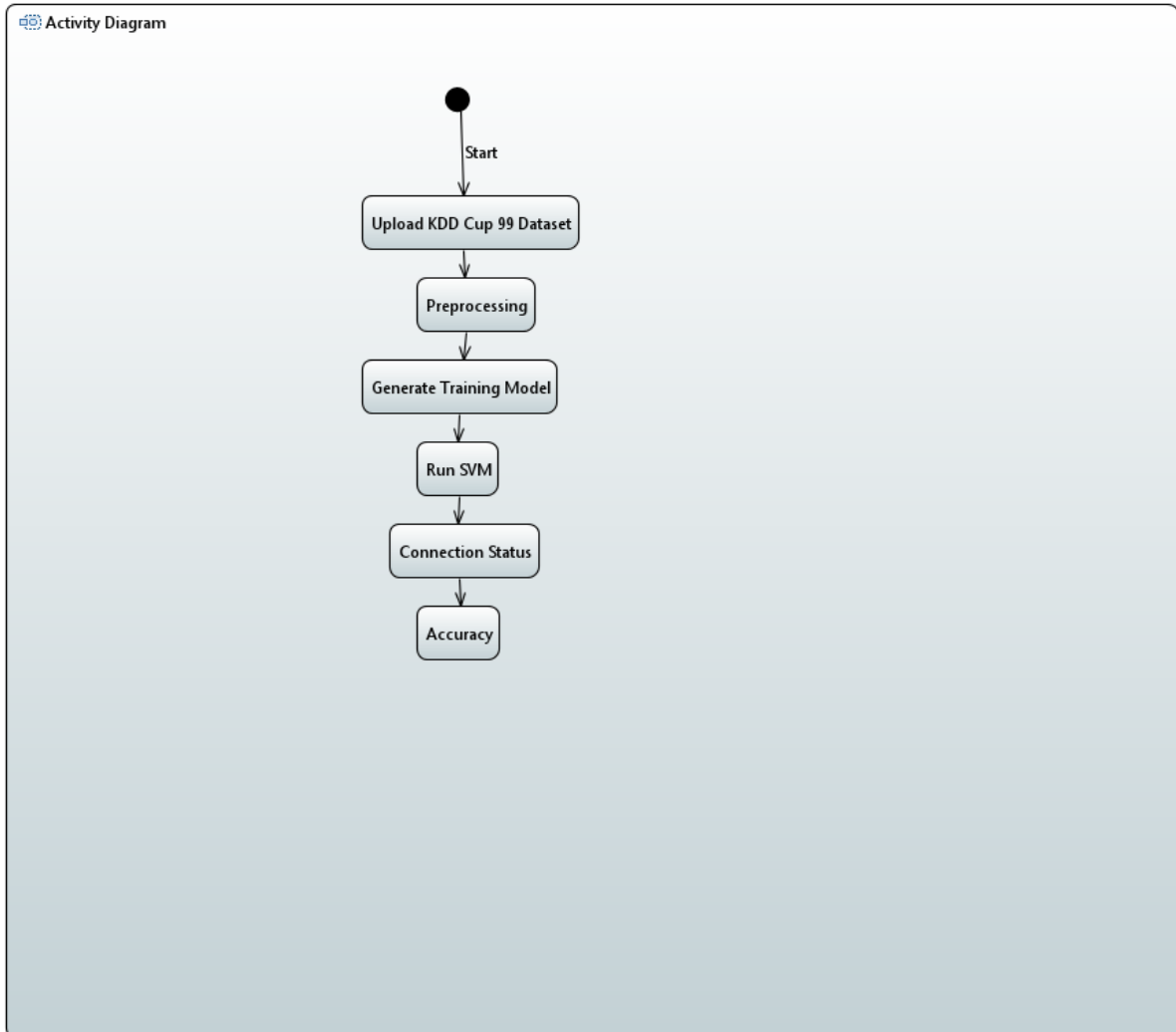


Figure 7. Activity Diagram

4 A model-based framework implementation

Industrial revolutions have created a platform in which industrial activities can be carried out automatically by machines, which we have witnessed four industrial revolutions to date. The first industrial revolution with the advent of steam engines and automation took place in the late eighteenth century. Mass production was the second industrial revolution that took shape in the early twentieth century, using electric power. Then, in the mid-1970s, the Third Industrial Revolution was formed with the aim of further supporting automated production based on electronics and information technology. In recent years, with the increase of research in the field of IoT and physical cyber systems, industries, governments and societies have noticed the tendency towards the fourth industrial revolution.

Accordingly, with an overview of the goals of industrial revolutions, we can understand the main role of production in the occurrence, establishment and spread of industrial revolutions among societies. In the fourth industrial revolution, this trend has continued, as in previous industrial revolutions, and in this regard, the level of investment in intelligent production is increasing rapidly. Of course, what is worrying is that in the process of these investments, the security of development and establishment has always been a secondary concern and has not been of major importance to the productive institutions. Exploring technologies such as the Internet of Things, cloud computing, bulk data, digital twins, augmented reality, 3D printing, artificial intelligence and the new generation of physical cyber systems that play a key role in establishing intelligent production in the Fourth Industrial Revolution. The security gap in intelligent manufacturing systems can be further understood.

Research also shows that only 16% of companies are prepared to meet the challenges of cybersecurity. But when it comes to cyber security, there is also the added concern that it is possible to attack without geographical restrictions and from anywhere in the world. Therefore, organizations will face a wide range of cyber-attacks. Summarizing the above, we will be aware of the need for security checks in intelligent manufacturing systems. However, due to the wide range of security issues in the mentioned technologies, in this article, we have tried to review the research related to the use of artificial intelligence in cyber security and in particular in physical cyber systems related to intelligent production.

Accordingly, we will first describe our research method in the second section. Then, in the third section, we will describe the role of cyber security in the two technical and managerial areas related to the Fourth Industrial Revolution. After that, in the fourth section, we will briefly introduce some artificial intelligence methods used in cyber security. Then, in the fifth section, we will examine the effects of artificial intelligence on attacks and countermeasures in cyber security. In the sixth section, by summarizing the issues raised, we introduce the benefits and challenges that artificial intelligence creates in cybersecurity and can affect the performance of intelligent manufacturing systems. Finally, in this chapter, we will provide suggestions for further research into the development of artificial intelligence in cybersecurity related to intelligent manufacturing systems.

Artificial intelligence methods in cyber security

- Learning algorithms

Artificial intelligence seeks to produce intelligent machines similar to humans, and to achieve this goal, these machines need to learn, and to achieve more accurate results, they need to practice using learning algorithms.

There are three types of learning algorithms for practicing machines:

- Supervised learning

This type of learning requires a learning process with big, precise data that is pre-determined. These types of algorithms are usually used as classification or regression mechanisms.

- Unsupervised learning

In contrast to regulatory learning algorithms, there is non-regulatory learning that uses unspecified data sets. These methods are often used in data clustering, dimensionality, or density estimation.

- Reinforcement learning

Reinforcement learning is a type of learning algorithm that is used in situations where the amount of data is limited by not provided.

Machine learning methods

Machine learning (ML) is one of the branches of artificial intelligence in which they try to empower systems by learning and improving the machine without using explicit programs. Machine learning involves the mathematical sciences that carry out the process of extracting information, discovering patterns, and inferring data. Machine learning algorithms include various types that can be classified into three general categories mentioned in learning algorithms. Standard machine learning algorithms in the field of cyber security, Decision Trees (DT), Support vector machines (SVM), Bayesian K-Nearest Neighbor (KNN), Random Forest (RF), and Principal Component Analysis (PCA).

Deep learning methods

Deep Learning (DL) is a subset of machine learning used to teach computers to do things that only humans can do in the amount of time expected. The most important advantage of deep learning over conventional machine learning is its better performance in large data sets. Common deep learning algorithms used in the field of cyber security include: Artificial Neural Network (ANN), Recurrent Neural Networks (RNNs), Deep Belief Networks (DBNs), Self-Assembled Encoders (SAEs).

Biological-based computing methods

Bio-based computing is an area of artificial intelligence that has undergone extensive research in recent years and includes a set of intelligent algorithms and methods that are compatible with bio-based features and behaviors and to solve Problems are used in real areas. Genetic algorithm (GA) techniques, Evolution Strategies (ES), Ant Colony Optimization (ACO), and artificial security systems have been most widely used in cybersecurity.

The effects of artificial intelligence on cyber security

What is clear today is that many cyber attackers use artificial intelligence and machine learning to improve and reinforce cyber-attacks. Of course, with the development of the use of artificial intelligence in the field of cyber and attention to this field in establishing security, tools based on artificial intelligence that can play a role in this field, have been considered. In this section, we try to introduce some of these cases.

Artificial intelligence enables the processes and abilities of humans in attacks to be performed automatically and intelligently, and in some cases leads to exceeding the ability of attacks. Attackers can also use artificial intelligence to make attacks with defense systems difficult. In threat intelligence, two perspectives provide the possibility that there is a general diagnosis:

Malware boost: Malware is a term for malicious software such as virus, worm, trojan horse, botnet, etc. The use of artificial intelligence has made the new generation of malware more sophisticated, smarter, and faster, as well as more difficult to detect.

Social Engineering: Artificial intelligence can be used to extract large amounts of social media data sets to extract users' personally identifiable information. This information can lead to malicious communication or the automatic creation of phishing messages.

Tools to attack artificial intelligence models

Since artificial intelligence is also used in defense strategies, which will be explained below) Attackers are always looking for tools that allow them to infiltrate artificial intelligence systems. Here, too, attackers act in three general ways:

Conflicting inputs: In this method, the attacker designs inputs to prevent the artificial intelligence defense system from detecting the attack.

Poisoning training data: In this type of attack, the attacker tries to modify the machine learning training data based on the algorithm used, in a way that significantly reduces the ability of the system to detect the attack.

Model extraction: In this method, the attacker uses reverse engineering techniques to try to identify the machine learning algorithm used to detect detected behaviors that deployed systems seek to avoid those behaviors to avoid those behaviors. This chapter describes the proposed method based on support vector machine method. First, preprocessing is applied to the data set and the data that has no effect on the result is deleted. Then, the backup vector machine algorithm is applied to the data set and the intrusion or non-intrusion status is determined.

Figure 8 shows the proposed method:

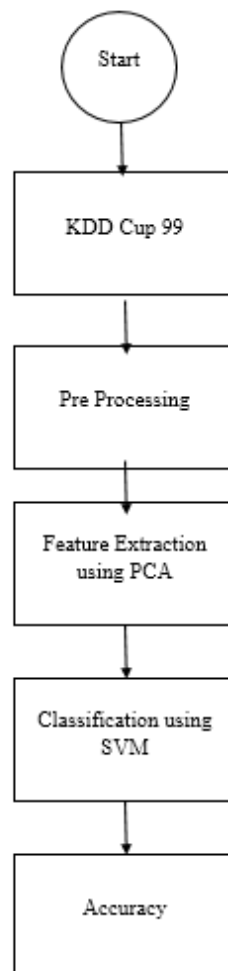


Figure 8. Proposed method

4.1 Framework architecture

Dataset: Due to the high volume of data, the training database, which contains 10% of the records, is used. This database contains 494021 communication records. Which includes all four types of attacks listed.

These four categories of attacks include 22 types of attacks and one normal mode. In addition to this database, there is also a test database that contains 311029 records that represent 37 types of attacks and a normal state. As a result, 15 attacks have more training than a database. Figure 9 is the dispersion of attacks in our database and as we can observe there is approximately around 300.000 Smurf attacks occurred:

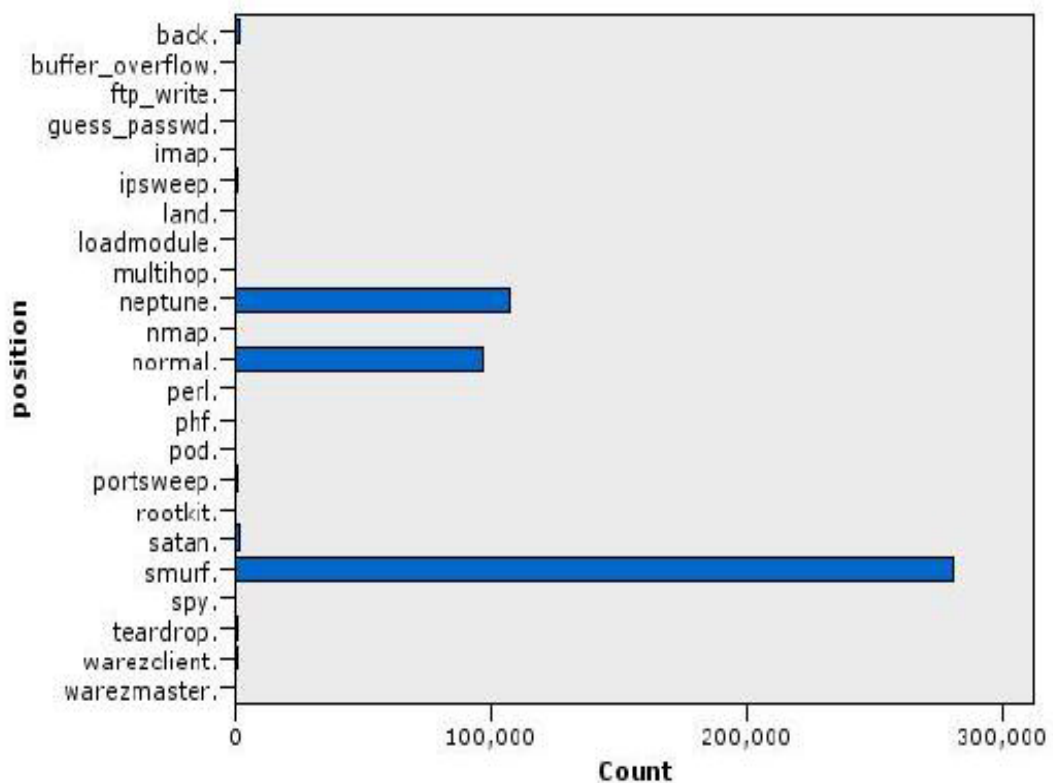


Figure 9. Dispersion of attacks in the database

Figure 10 represents the percentage of the discussed attacked which are existing in our database, for instance, the famous Smurf attack which is exactly equal to 280.790 count is about to fill the 56.84 percentage of our proportion.




Value	Proportion	%	Count
smurf.		56.84	280790
neptune.		21.7	107201
normal.		19.69	97278
back.		0.45	2203
satan.		0.32	1589
ipsweep.		0.25	1247
portsweep.		0.21	1040
warezclient.		0.21	1020
teardrop.		0.2	979
pod.		0.05	264
nmap.		0.05	231
guess_passwd.		0.01	53
buffer_overflow.		0.01	30
land.		0.0	21
warezmaster.		0.0	20
imap.		0.0	12
rootkit.		0.0	10
loadmodule.		0.0	9
ftp_write.		0.0	8
multihop.		0.0	7
phf.		0.0	4
perl.		0.0	3
spy.		0.0	2

Figure 10. Rate and percentage of attacks in the database

Figures 9 and 10 show the distribution of attacks in the database. As can be clearly seen, the highest number of records related to smurf attack is from the category of denial of service attacks and the lowest rate of attack related to spy attack is from the category of access attacks. Due to the low rate of some attacks compared to other attacks in the database, their scatter percentage is zero. A full description of the features used in the database can be seen in the table below.

Among the proposed attributes, 2 attributes (num_outbound_cmds, is_host_login) always have a constant value of zero in all records.

As a result, since they have no effect on determining the communication status, they can be eliminated and reduce the complexity of the issue and the time consumed. So, the input vector currently has 37 properties. When eliminating unnecessary features, focusing on the important ones can increase speed and efficiency without compromising statistical accuracy. Eliminating additional information also saves CPU and memory consumption in the training and testing process and storage space for the complex model, and as mentioned, memory and CPU constraints are highly regarded in sensor wireless networks.

The steps for selecting attributes are as follows:

- 1- Delete one of the input attributes
- 2- Test the new database with the selected model
- 3- Analyzed the obtained results according to the evaluation criteria
- 4- Calculate the importance of the deleted attribute according to the defined rules
- 5- Steps 1 to 4 are repeated for all available attributes.

To select the necessary attributes, the selection rules are set according to the evaluation criteria as follows:

- 1- If the detection rate increases, the accuracy rate increases, it is an unnecessary feature.
- 2- If the detection rate, the accuracy rate remains constant, the attribute is unnecessary.
- 3- Increasing the accuracy rate, decreasing the positive error rate is an unnecessary feature.
- 4- If the detection rate decreases, the positive error rate increases, the attribute is essential.
- 5- If the detection rate decreases, the accuracy rate decreases.
- 6- If the detection rate increases, the accuracy rate decreases.

7- The degree of accuracy decreases, the positive error rate increases the attribute is essential.

Feature extraction using PCA: The method of principal component analysis was first proposed in 1901 by Carl Pearson, which is one of the most important and basic topics in chemometrics and is one of the non-regulatory classification methods. Principal Component Analysis (PCA) is one of the simplest multivariate analysis methods used. The goals are:

- A) Extract the most important information from the data
- B) Reduce the size of the data
- C) Ease of describing the data set
- D) Analyze the structure of samples and variables and classify and identify variables

Due to the fact that multidimensional space is often difficult to understand, and also due to technical problems when it comes to high-dimensional problems, it is possible to create a matrix in order to organize such data, in which samples are rows and variables are columns. Be it. The data of such a matrix is called multivariate data. The analysis of the main components, through the linear combination of the main variables, leads to a reduction in the number of variables.

In order to better understand the data, chemists display it graphically, which can display data from one dimension to three dimensions, but if the dimensions (the number of variables measured for each sample exceeds three dimensions), It is not possible to graphically display them, in which case the variables must be reduced to three main components or less.

4.2 IMPLEMENTATION

In order to better understand the issue, we first consider the reduction of two variables to one, assuming that only one dimension can be observed. One solution is to image points from two-dimensional space to one-dimensional space, which under such conditions is very important for the line on which the points are depicted. Figure 10 shows two sets of data in two-dimensional space, group one marked with blue rhombuses, and group two with red squares. The data images on the red line have destroyed the information in the original data. For example, the images do not show that the main data are two groups, but the images of the data points on the green line separate the examples of groups one and two well. According to this figure, a good direction to visualize the main data is the axis on which the data is most scattered. This line is called the first major component, or PC1 - which represents the largest variance in the data and contains more information from the data. The point images of the original data from the two-dimensional space on the PC1 axis are called point points on the PC1.

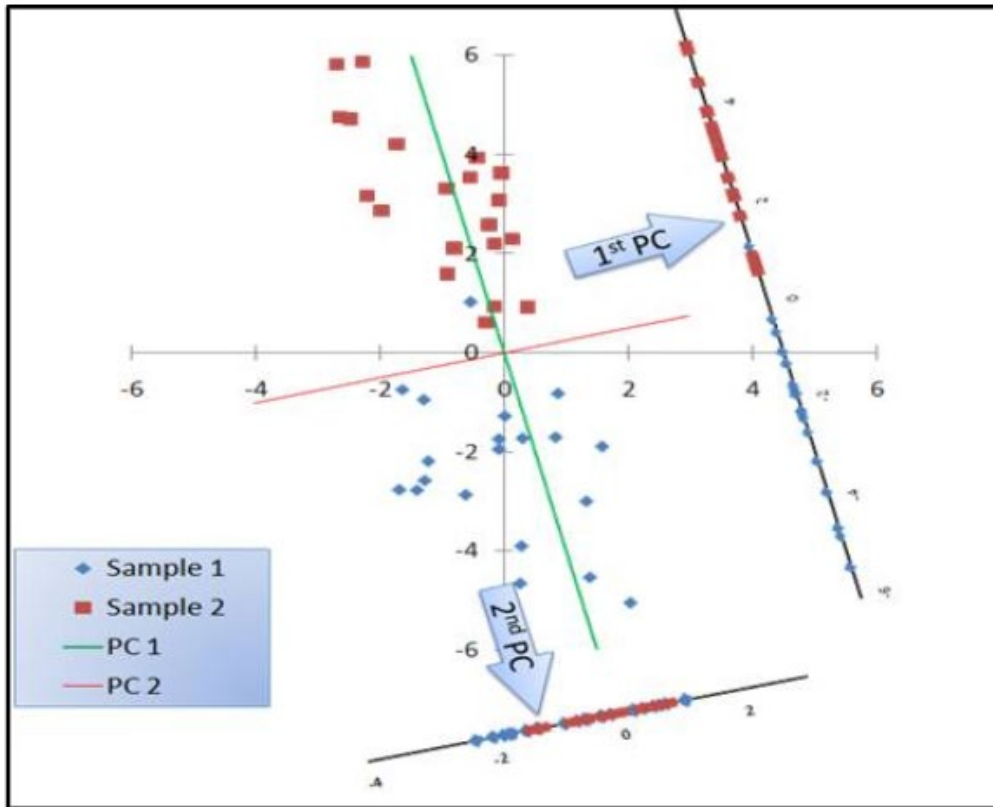


Figure 11. The depiction of primary data from two-dimensional space to one-dimensional space remains in the green line of information with imagery, but in the red line of information with imagery is lost.

PCA is a method of multivariate analysis that creates and selects a smaller number of new variables called principal components from the linear composition of the main variables, so that some less important information is removed. The first basic component extracted contains the largest amount of data scatter in the entire data set. The second extracted component also has two important features, first, that this component has the highest variance of data, which is not described by the first component, and second, that it is perpendicular to the first component.

Figure 12 shows this. According to the figure, part (a) is before the application of the PCA technique, which shows that the data information is evenly distributed on the X and Y axes, but part (b) shows that after the application of PC1 PCA accounts for the largest share of data variance

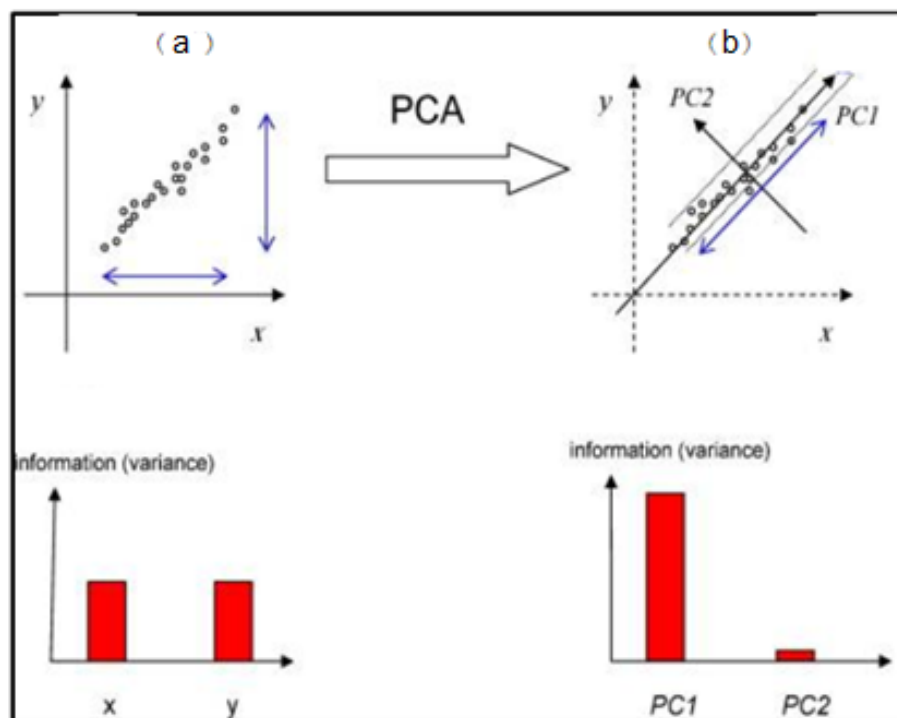


Figure 12. Part a) Before PCA is applied, the data are evenly distributed on the X and Y axes in two-dimensional space. Part b) After PCA application, most of the data are distributed on PC1 line

Classification using Support Vector Machine

A support vector machine, abbreviated to SVM, is a computer algorithm that learns by example how to assign associated tags to different objects. The purpose of this algorithm is to identify and differentiate complex patterns in data that are used depending on the application in clustering, classification, ranking, clearing, and so on. The backup vector machine generally consists of four basic concepts: separator plate super, maximum margin plate, soft margin, and RAR core function.

- Separator hyperplane

Hyperplane is basically a term that covers space above three dimensions. In one-dimensional hyperplane space there is only one point, in two-dimensional space there is one line, and in three-dimensional space there is one page, and in space more than three dimensions we have hyperplane. But for convenience, we call all these levels a hyperplane.

Consider Figure 4, in this figure in the ALL and AML classes we have two-dimensional space with the all-related area at the bottom of the figure and the AML at the top of the figure. There are also dots of color near the ALL class, the class of which is not specified and needs to be classified. These two classes can be separated by a line, which we see in Figure 13 are based on this, the task of the unknown point will also be determined. This point is part of the ALL class.

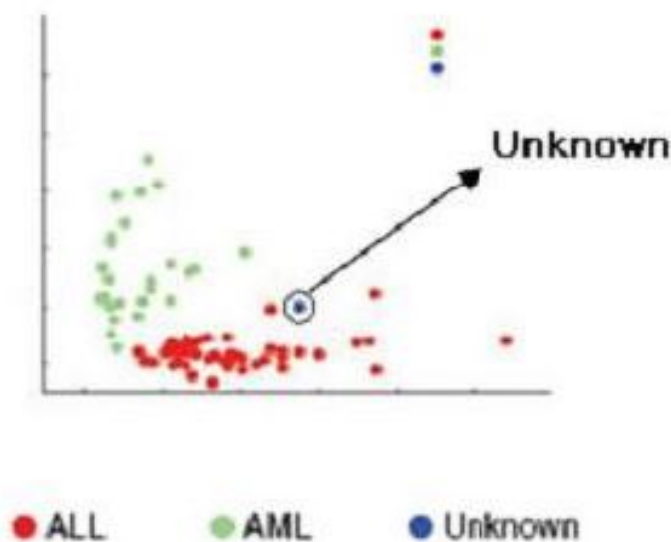


Figure 13. Data in two-dimensional space

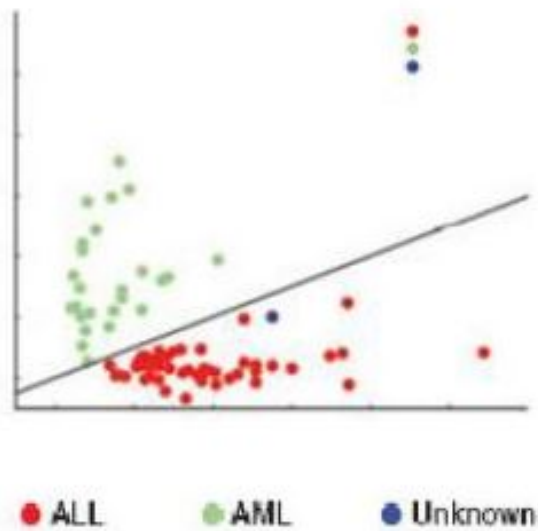


Figure 14. Linear classification in two-dimensional space

Now we see the same separation of two classes in one and three-dimensional space in the following figures. In Figure 15a, since the data has only one dimension, two classes can be separated by a point, and in Figure 15b, where the data has three dimensions, two classes can be separated by a plane.

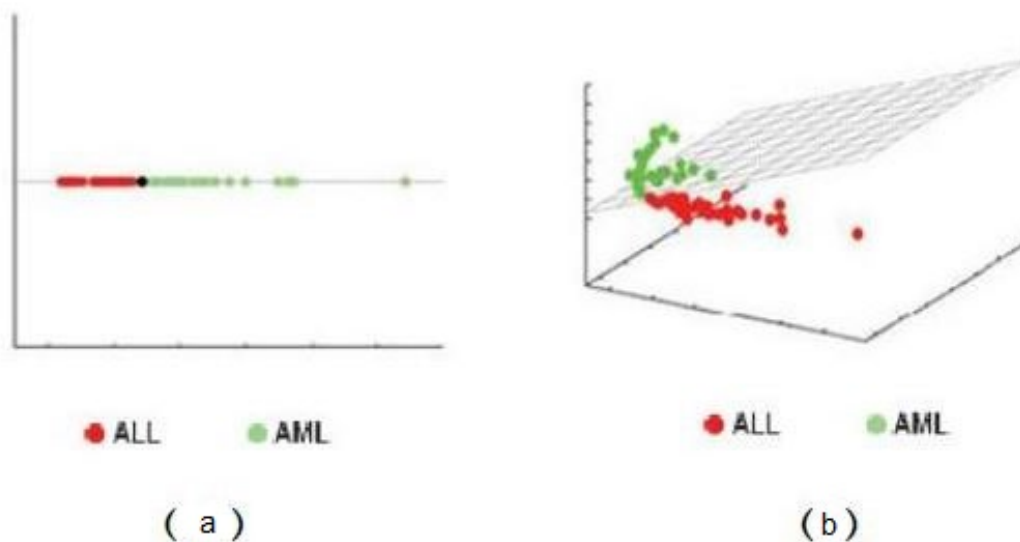


Figure 15. Classification in space a) one-dimensional, b) three-dimensional

- **Hyperplane with maximum margin**

So far, we know that the purpose of the backup vector machine is to separate the two classes ALL and AML in two-dimensional space with one line. There are many lines that do this. Figure 16 shows a view of these lines and Figure 17 represents the support vectors selective lines which is used to make an observable difference of the classes.

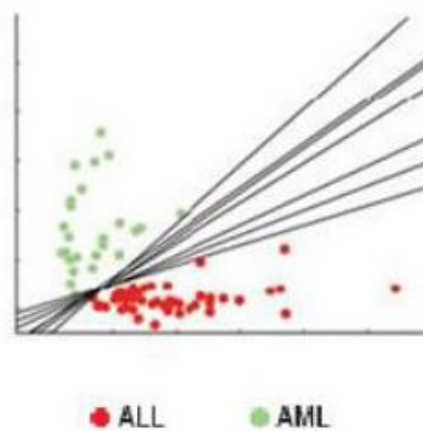


Figure 16. Lines distinguishing two classes

The question is, which of these lines is better than the other separating lines? The SVM algorithm selects the line in the middle as the separator cloud. In other words, it selects the line that has the maximum distance from each of the classes.

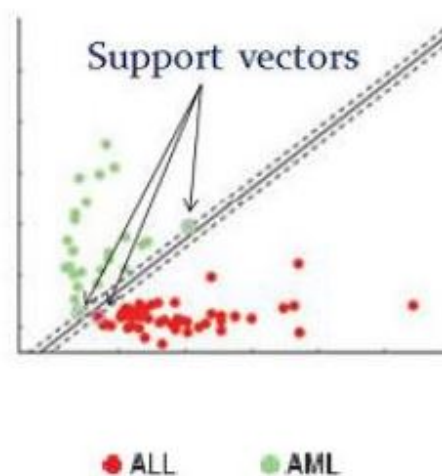


Figure 17. Support vectors of two classes

If we consider the distance between hyper plane and the nearest instructional vectors of the classes as the margin, the SVM selects the midline hyperplane with the maximum margin. Selecting this particular hyperplane maximizes the ability to accurately predict the class for unclassified instances.

In the figure, in addition to the hyperplane with the maximum margin, you see points called support vector. The training data closest to the separator hyperplane is called the support vector. In fact, the SVM detects the pattern between two classes by finding a decision level that maximizes the distance to the nearest points in the training set called the support vectors.

- Soft margins

Until now, it has been assumed that data sets with a straight line could be split into two completely separate parts, which in fact may not be the case. In fact, many real datasets cannot be easily separated, but rather have a distribution such as Figure 18, where the dataset contains an error.

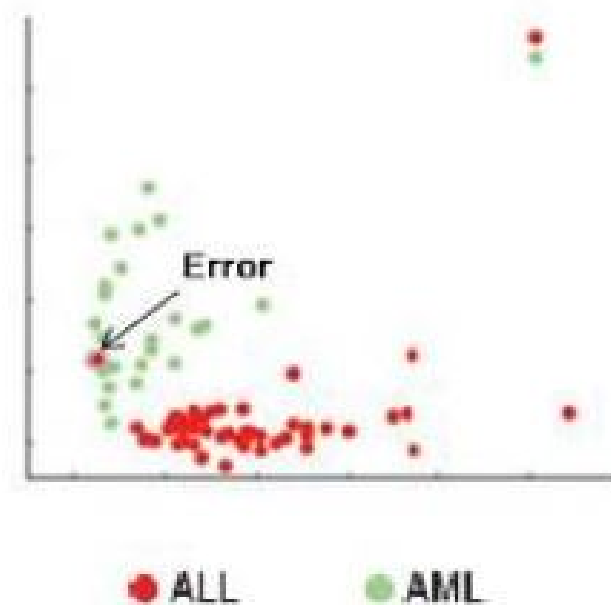


Figure 18. Datasets with errors

In this case, we want the support vector machine to be able to deal with data errors and allow the training data to be, to a certain extent, allowed to be in the wrong class, on the wrong side of the separator hyperplane. To manage such cases, it is necessary to make changes to the support vector machine by adding a soft border feature. This feature allows multiple data points to pass through the separator hyperplane and be on the wrong side, without affecting the final result. Figure 19 shows the soft margin solution.

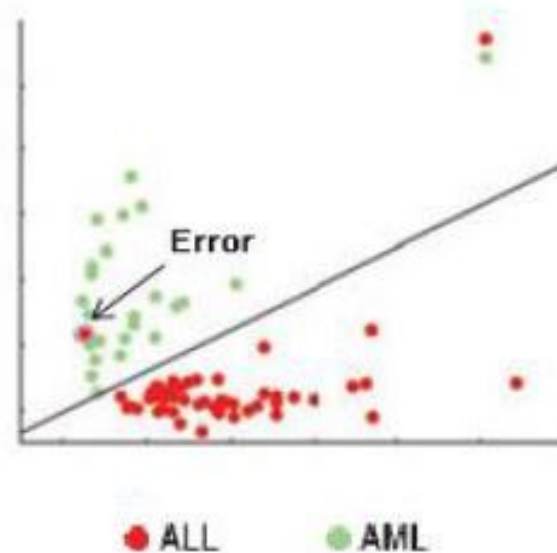


Figure 19. Classification with a soft margin

The sample with the error is now incorrectly classified. Of course, the point here is that we do not want SVM to allow too many of these misclassifications to occur. It is therefore necessary to specify a parameter to control the maximum number of permissible errors and the maximum allowable distance for accepting classification errors. The value of this parameter is determined by the user according to the application.

- **How to calculate Margin**

According to the above equations, the value of M can be obtained by b and W . Consider the two hypothetical points X_- and X_+ belonging to the positive and negative planes, respectively, in Figure 20.

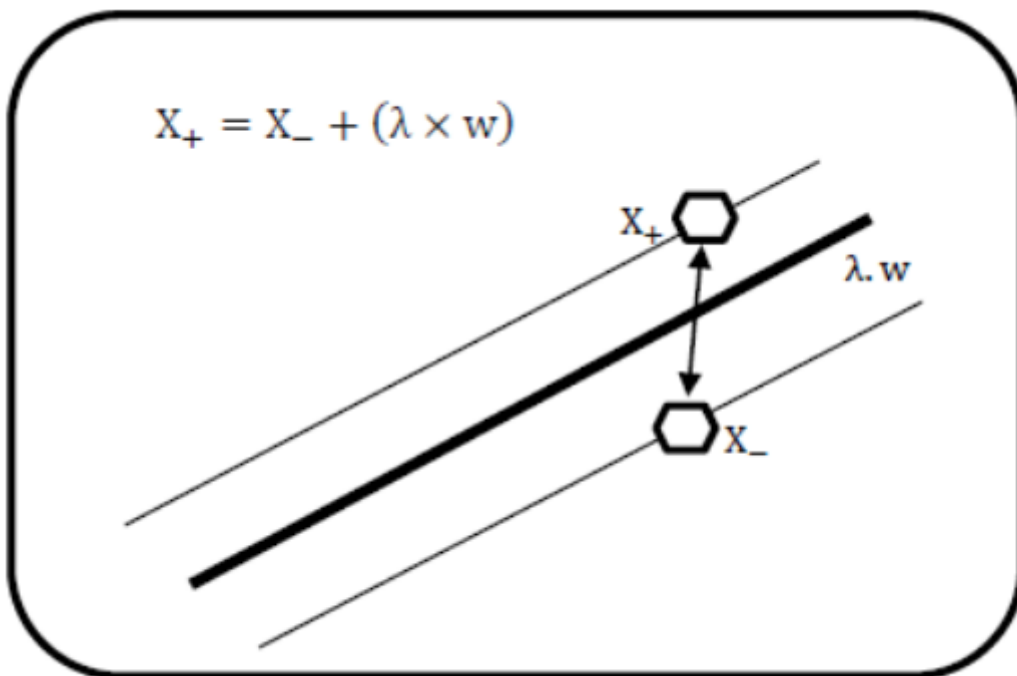


Figure 20. Margin calculation process

As we have the following relations:

$$(W * X_+) + b = +1$$

$$(W * X_-) + b = -1$$

$$X_+ = X_- + (\lambda * W)$$

$$X_+ - X_- = |\lambda * W|$$

Given the equation of the last relation, obtaining M with the help of b and W is a simple task.

Here is the process of obtaining M in relation:

$$(W * (X_+ + (\lambda * w))) + b = 1 \text{ then } (w * x_-) + b + ((\lambda * w) * W) = +1$$

$$-1 + b + b + ((\lambda * w) * W) = +1 \text{ Then } \lambda = \frac{2}{w * W}$$

$$M = |X_+ + X_-| = |\lambda * w| = \lambda * |w| = (\lambda * \sqrt{w * W}) = \frac{2 * \sqrt{w * W}}{(w * W)} = \frac{2}{(w * W)} = \frac{2}{|w|}$$

Now to maximize the value of M , $2 / (|w|)$ means the separation distance, $1/2 |w * W|$ Minimized. So, the main problem with SVM is minimizing $1/2 |w * W|$ Will be due to inequality.

$$y_i * ((W * X) + b) - 1 \geq 0$$

The method obtains the Lagrange coefficients w , b and a in such a way that $L(w, b, a)$ is minimized.

$$L(w, b, a) = \frac{1}{2} |w * W| - \left(\sum \alpha * [Y * ((W * X) + b) - 1] \right)$$

- Kernel functions

All of these topics were related to linear classification, but if the input data is nonlinear, the input vectors must be moved to a higher attribute space. Figure 21 shows a conversion of a nonlinear to linear separator.

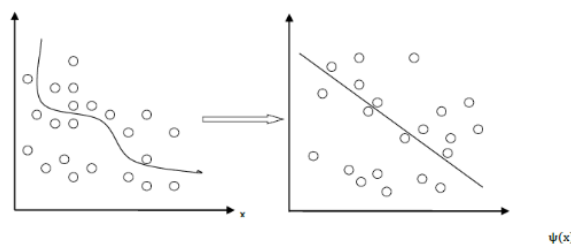


Figure 21. View of converting a nonlinear to linear separator

Suppose X is an example of an input property space with dimensions n ; The following equation is a nonlinear transformation that maps samples from N -dimensional space to a new N' -dimensional space. What this mapping does in an SVM algorithm is a function called a kernel.

$$\psi_j(x), \quad j = 1, 2, \dots, n', n' > n$$

So, one of the important factors in an SVM machine is the kernel function. Although several kernel functions have been proposed by researchers, the following basic kernel functions are used in most SVM machines.

Linear functions: $K(X_i, X_j) = (X_i, X_j)$

Polynomial functions: $K(X_i, X_j) = (\lambda * (X_i, X_j) + r)^d, \lambda > 0$

(RBF) Radial Basis Functions: $K(X_i, X_j) = \exp(-\lambda \|X_i - X_j\|^2), \lambda > 0$

Gaussian Radial Basis functions: $K(X_i, X_j) = \exp\left(-\frac{\|X_i - X_j\|^2}{2\sigma^2}\right)$

Sigmoid functions: $K(X_i, X_j) = \tanh(\lambda * (X_i, X_j) + r)$

Where λ , r and d are the kernel parameters and λ is a variable that plays an important role in the SVM learning machine. After describing the proposed method, which was based on the support vector machine, the proposed method is tested and verified in this chapter. First, the specifications of the computer system with which the simulation has been done are discussed, then the data set is described, then the implementation is described, and finally the evaluation is performed.

4.3 TESTING AND VERIFICATION

System specifications

This project was performed on a system with the following specifications:

CPU: Core i7

Ram: 8G

Hard disk: 1TB

GPU: 2 GB

Dataset presentation

A standard set of audit data is developed to compare and evaluate intrusion detection systems, including a variety of intrusions and this database is represented as (kdd (data_10_percent)).

This database has 5 different types of data, including 4 types of attacks and a series of normal data. Of course, the 4 types of attacks themselves include a subset of different types of attacks. Ten percent of the database itself is about five hundred thousand data, which is a large number that is not usually used in detection systems of all this data. Because each of these data records has 41 different properties. If we want to use this whole ten percent database, then intrusion detection system training takes a long time. One of the tasks done in this project is to select a suitable set of this database, so that it is not too large and includes different types of attacks and in appropriate numbers.

This thesis is implemented using MATLAB programming language and evaluations were performed on KDD 99 CUP database. The database used in this research is a database provided by DARPA in 1999, which is known as one of the most reliable databases for testing intrusion detection systems. This database, called KDDCup99 as following:

- 1- The first category includes the main attributes of a connection, which include the basic attributes of the TCP connection. Connection time, type of connection, type of network service is in this category.
- 2- The second category is the content attributes in a connection, such as unsuccessful communication attempts.
- 3- The third category is the host attributes that test the communications established in the last 2 seconds that have the same destination as the current connection and calculate statistics related to the behavior of the protocol, service, etc.
- 4- The fourth category, similar features of the same service, inspects last-second communications that have the same service as the current connection.

Attacks also fall into four categories:

- 1- Reject of service: Creating calculations or occupying memory resources to deny legal acceptance to use these resources
- 2- Exploration: Scan hosts and ports for information by finding known vulnerabilities
- 3- Access to the location: Unauthorized access from the vehicle outside the area in order to reveal vulnerabilities
- 4- Root Access: Unauthorized access to the root

In the first part, the codes related to reading and uploading the database are done in MATLAB environment and the information is read from the database and converted into an understandable format of MATLAB environment and finally stored in a file. Inside this file, the code related to the attacks is defined as follows:

The code for DOS attack is number 1 and U2R (user to root) attack is number 2, R2L attack is number 3, PROB attack is number 4, and finally the non-attack status or NORMAL is specified with number 5.

So far, the output shows the number of attacks. The inputs are the same number of features multiplied by the total number of samples, which is equal to the number $41 * 494021$.

Then the following results were obtained:

- 1) Total number of samples: 494021
- 2) Number of features: 41

70% of the data is selected for training and 30% for testing. After execution, the number of training data is 345814 and the number of test data is 148207. Attacks are also categorized from 1 to 5. Categorized by SVM. Finally, the Confusion function or the function of emptiness and confusion is used to correctly diagnose the results.

For this section, I have first divided the data into two categories of training and testing, of which 70% of the data is intended for training. Then trained the desired model using the dataset parameters. The comparison of the proposed algorithm in terms of detection rate, accuracy and false alarm rate parameters with existing algorithms is presented.

The various criteria used for this purpose, performance evaluation and experimental results are observed, which are obtained as follows:

- 1- Detection rate (DR)
- 2- False Alarm rate (FAR)
- 3- False Positive (FP)
- 4- False negative (FN)
- 5- True Positive (TP)
- 6- True negative (TN)

Definition of evaluation criteria:

- 1- False Positive (FP): Equivalent to each IDS alert time when an attack did not occur
- 2- False negative (FN): means the IDS failed to detect a real attack
- 3- True Positive (TP): Equivalent to a real attack that triggers IDS to generate an alert
- 4- True Negative (TN): Equivalent to a situation in which no attack occurs and no warning is generated [25]

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\ Score = \frac{2 * (Recall * Precision)}{(Recall + Precision)}$$

Evaluation parameter	Value
Accuracy	98.98 %
Precision	98%
Recall	99 %
F1 Score	99%

Table 3. Evaluation of proposed method

Conclusion

There are strong incentives to address cybersecurity risk management in OT-IT combined environments. The use of new technologies in connected networks, the pressure of innovation and desire to reduce costs, forces companies to consider the security aspects. Security is one of the major challenges in networks, both wired and wireless networks. This work is to provide a system that can enhance network security. Cyber-attackers are increasingly designing more sophisticated methods to attack industrial OT-IT combined systems which are vastly used in industrial environments. In this thesis, I tried to present a solution to identify and distinguish a wide range of methods and algorithms in such a way that if the attackers try to the security mechanisms, the system will identify and minimize the damage caused by the attack.

Intrusion detection systems nowadays are one of the critical and vital solutions to provide a secure network of computers. The goal of this thesis as demonstrated especially in chapters 3 and 4 is a network security software solution based on a proposed model over OT-IT to observe and detect common intrusions using support vector machine based on machine learning technique. The aim of this study is to provide security by providing a UML and SysML model-based intrusion detection solution. The model presented in this thesis is meant to eliminate the attack or reduce its effect for OT-IT systems. OT-IT combined environment like industrial systems are often developed based on operational and business requirements and little attention is paid to the potential security effects in them as described in this work.

To reduce existing vulnerabilities by detecting the network common cyber threats, a solution as an intrusion detection system introduced based on SysML and I used KDD cup 99 dataset interties to evaluate and assess my system. In this work, first, using my proposed model and the dataset, to providing an intrusion detection system that has high efficiency. This study presents a high-performance intrusion detection system using support vector machine. Support vector machine is one of the supervised machine learning algorithms. An example of implementation of this algorithm can be well understood by reading the chapter 4 of this thesis. The support vector machine is used to be able to deal with data errors and allow the training data to be, to a certain extent. Another technique which is used in this work is the Principal Component Analysis (PCA to identify key components and help us analyze a set of features that are more valuable than just examining them all, and was used to determine the desired space of the results. As explained earlier, we have data being extracted from our dataset, major portion the data is selected for training and the rest of it is for testing. Metrics of the evaluation parameters calculated using the Accuracy, Precision, Recall and F1-score formula.

References

- [1] B. Admin. "10 High Profile Cyber Attacks in 2021." cybermagazine.com. <https://cybermagazine.com/top10/10-high-profile-cyber-attacks-2021> (accessed: Apr. 22, 2021).
- [2] I. Jibilian and K. Canales. "Solarwinds-hack-explained-government-agencies-cyber-security." BUSINESSINSIDER.com. [Online]. Available: <https://www.businessinsider.com/solarwinds-hack-explained-government-agencies-cyber-security-2020-12>.(Accessed: April 19, 2022).
- [3] J. James, "Sharing Mechanisms for Information Technology in Developing Countries, Social Capital and Quality of Life." *Social Indicators Research*, vol. 94, no. 1, pp. 43-59, 2008, doi: 10.1007/s11205-008-9335-3.
- [4] "Policy Framework for Investment. Organization for economic co-operation (OECD)." OECD.org. 2006. <https://www.oecd.org/daf/inv/investment-policy/36671400.pdf> (accessed: Apr. 22, 2021).
- [5] M. Parchamijalal, S. Moradi, and M. Z. Shirazi, "Engineering Claim management office maturity model (CMOMM) in project-oriented organizations in the construction industry, *Construction and Architectural Management*", ISSN: 0969-9988 Article publication date: 22 December 2021.
- [6] I. Ajie, "A Review of Trends and Issues of Cybersecurity in Academic Libraries", *Libraries at University of Nebraska-Lincoln*, 2019. [Online]. Available: <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=5803&context=libphilprac>. Accessed: April 10, 2022.
- [7] P. Theron and A. Lazari, "IACS Cybersecurity Certification Framework (ICCF). Lessons from the 2017 study of the state of the art." Publications.jrc.ec.europa.eu.[Online].Available: https://publications.jrc.ec.europa.eu/repository/bitstream/JRC111611/the_iacs_cybersecurity_certification_framework.pdf. Accessed: April 10, 2022.

- [8] CRC, E.A. Fischer, "Cybersecurity Issues and Challenges: In Brief." Derechos.org.[Online]. Available: <http://www.derechos.org/nizkor/espana/doc/cybersecu2.html> (Accessed: April 10, 2022).
- [9] Y. Li and Q. Liu, "A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments," Energy Reports, Volume 7, 2021, Pages 8176-8186, ISSN 2352-4847, <https://doi.org/10.1016/j.egy.2021.08.126>.
(<https://www.sciencedirect.com/science/article/pii/S235248472100729>)
- [10] B. Dobran. "What is Cyber Security? Challenges and Threats Organizations Face." Phoenixnap.com. [Online]. Available: <https://phoenixnap.com/blog/what-is-cyber-security> (Accessed: April 10, 2022).
- [11] B. Harrell, "A Guide to Critical Infrastructure Security and Resilience", Ambassador Nathan Sales. Cisa.gov., 2019. Available: <https://www.cisa.gov/sites/default/files/publications/Guide-Critical-Infrastructure-Security-Resilience-110819-508v2.pdf> (Accessed: April 10, 2022).
- [12] M. Jouini and L.B. Rabai, "Comparative Study of Information Security Risk Assessment Models for Cloud Computing systems." Procedia Computer Science, Volume 83, 2016, Pages 1084-1089, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2016.04.227>.
(<https://www.sciencedirect.com/science/article/pii/S1877050916302605>)
- [13] M. Nieves, K. Dempsey, and V. Y. Pillitteri. "An Introduction to Information Security." NIST Special Publication 800-12 Revision 1. Available: <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-12r1.pdf>.
(Accessed: April 10, 2022).
- [14] B. A. Sabzkohi and T. Pourrostan. "Integrated Model of Value Engineering and Risk Management Approaches in Empowerment Projects (The Exterior Design)." International Journal of Advances in Mechanical and Civil Engineering (IJAMCE) (2016), pp. 52-56, Volume-3, Issue-3. DOI online no - [ijamce-iraj-doionline-4877](https://doi.org/10.1016/j.ijamce-iraj-doionline-4877).

- [15] D. Watson and A. Jones. "Digital Forensics Processing and Procedures." Chapter 5 - Risk Management, Syngress, 2013, Pages 109-176, ISBN 9781597497428, <https://doi.org/10.1016/B978-1-59749-742-8.00005-4>.
(<https://www.sciencedirect.com/science/article/pii/B9781597497428000054>).
- [16] K. Clarke. "Novel Alert Visualization: The Development of a Visual Analytics Prototype for Mitigation of Malicious Insider Cyber Threats." 2018. Doctoral dissertation. Nova Southeastern University. Retrieved from NSUWorks, College of Engineering and Computing. (1049) (https://nsuworks.nova.edu/gscis_etd/1049).
- [17] E. Conrad , S. Misener, and J. Feldman. "Operations Security, Eleventh Hour CISSP.". (Second Edition), Chapter 7 - Domain 7: Syngress, 2014, Pages 117-133, ISBN 9780124171428, <https://doi.org/10.1016/B978-0-12-417142-8.00007-8>.
(<https://www.sciencedirect.com/science/article/pii/B9780124171428000078>).
- [18] I. Saif. "Cyber risk in an Internet of Things world Flashpoint edition 4: More data, more opportunity, more risk." DELOITTE.COM. Available: <https://www2.deloitte.com/us/en/pages/technology-media-and-telecommunications/articles/cyber-risk-in-an-internet-of-things-world-emerging-trends.html>.(Accessed: April 09, 2022).
- [19] D. Watson and A. Jones. "IT Infrastructure, Digital Forensics Processing and Procedures." Chapter 7 -, Syngress, 2013, Pages 233-312, ISBN 9781597497428, <https://doi.org/10.1016/B978-1-59749-742-8.00007-8>.
(<https://www.sciencedirect.com/science/article/pii/B9781597497428000078>).
- [20] G. Singh, K. P. Yadav, and Y. Sharma, "A STUDY ON CYBER SECURITY: ISSUES AND CHALLENGES", International Journal of Multidisciplinary Consortium Volume - 2 | Issue - 1 | March 2015. ijmc.editor@rtmonline.in | <http://ijmc.rtmonline.in> | ISSN 2349-073X.

- [21] Z. A. Soomro, M. H. Shah, and J. Ahmed, "Information security management needs more holistic approach: A literature review". April 2016 *International Journal of Information Management* 36(2):215-225. DOI: 10.1016/j.ijinfomgt.2015.11.009.
- [22] "What is cybersecurity architecture?" cybersecurityforum.com. <https://cybersecurityforum.com/cybersecurity-faq/what-is-cybersecurity-architecture.html> (accessed: Apr. 22, 2022).
- [23] J. A. Bullock, G. D. Haddow, and D.P. Coppola. "Introduction to Homeland Security." (Fourth Edition) 8 - Cybersecurity and Critical Infrastructure Protection," Butterworth-Heinemann, 2013, Pages 283-321, ISBN 9780124158023, <https://doi.org/10.1016/B978-0-12-415802-3.00008-7>. (<https://www.sciencedirect.com/science/article/pii/B9780124158023000087>).
- [24] E. Gamma, L. Nackman, and J. Wiegand, "EMF: Eclipse Modeling Framework." SILO.PUB. Available: <https://silo.pub/eclipse-modeling-framework.html>. (Accessed: April 09, 2022).
- [25] B. A. Hosseini and E. Ashal. "Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity." *Emerg (Tehran)*. 2015 Spring;3(2):48-9. PMID: 26495380; PMCID: PMC4614595. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4614595>).