An Integrated Cyber Security Risk Management Framework and Risk Predication for the Critical Infrastructure Protection

Halima Ibrahim Kure[1], Shareeful Islam[2], Haralambos Mouratidis[3]
[1]School of Architecture, Computing and Engineering, University of East London, UK
[2]School of Computing and Information Science, Anglia Ruskin University, UK
[3]Institute for Analytics and Data Science, University of Essex, UK
h.kure@uel.ac.uk, shareeful.islam@aru.ac.uk, h.mouratidis@essex.ac.uk
Corresponding author: Shareeful Islam (shareeful.islam@aru.ac.uk)

Abstract

Cyber security risk management plays an important role for today's businesses due to the rapidly changing threat landscape and the existence of evolving sophisticated cyber attacks. It is necessary for organizations, of any size, but in particular those that are associated to a critical infrastructure, to understand the risks, so that suitable controls can be taken for the overall business continuity and critical service delivery. There are a number of works that aim to develop systematic processes for risk assessment and management. However, the existing works have limited input from threat intelligence properties and evolving attack trends, resulting in limited contextual information related to cyber security risks. This creates a challenge, especially in the context of Critical Infrastructures, since attacks have evolved from technical to socio-technical and protecting against them requires such contextual information. This research proposes a novel Integrated Cyber Security Risk Management (i-CSRM) framework that responds to that challenge by supporting systematic identification of critical assets through the use of a decision support mechanism built on fuzzy set theory, by predicting risk types through machine learning techniques, and by assessing the effectiveness of existing controls. The framework is composed of a language, a process and it is supported by an automated tool. The paper also reports on the evaluation of our work to a real case study of a critical infrastructure. The results reveal that using the fuzzy set theory in assessing assets' criticality, our work supports stakeholders towards an effective risk management by assessing each asset's criticality. Furthermore, the results have demonstrated the machine learning classifiers' exemplary performance to predict different risk types including denial of service, cyber espionage, and Crimeware.

KEYWORDS: Cyber Security Risk Management, Threat Intelligence, Fuzzy Theory, Control Effectiveness, Risk Prediction, Machine Learning, Case Study

1. Introduction

Critical Infrastrucures (CIs), such as energy and healthcare, heavily rely on Information and Communication Technology (ICT) to support reliable service delivery. Such integration of ICT to CIs introduces a number of advantages, such as higher degree of flexibility, scalability and efficiency in the communication and coordination of advanced services and processes. On the other hand, the increase usage of ICT in CIs creates new opportunities for cyber attacks and increases the vulnerability of those systems. Due to the importance of Critical Infrastructures, there is recently an increased number of attacks that are evolving in terms of sophistication, persistence and the resources that attackers have available. Such attacks consider not just the technical limitations of the relevant technologies but also the contextual information related to the Critical Infrastructure.

Despite of several existing works on cybersecurity risk management, the literature fails to present works that consider such contextual information when performing risk management for Critical Infrastructures. Moreover, existing works focus more on the prediction of risks and do not consider – as part of the same process – necessary controls that mitigate those risks. Our work advances the state of the art through the integration of Cyber Threat Intelligence (CTI) to the risk management process, to understand contextual information related to the threat actor's behaviour, Tactics, Techniques and Procedures (TTP) and Indicators. Moreover, it provides a unified process that integrates both risk prediction and risk mitigation with the aid of machine learning.

This paper presents an Integrated Cyber Security Risk Management Framework(i-CSRM) for the effective security risk management of Critical Infrastructures, which brings together the above contributions of our work. The proposed i-CSRM includes three main components: a conceptual view, a process and a tool. It also makes use of the latest relevant threat, attack and vulnerability repositories and standards such as the Common Weakness Enumeration (CWE), the Common Attack Pattern Enumeration and Classification (CAPEC), the CIS Critical Security Control (CSC) and CTI methods for risk management activities.

The i-CSRM framework introduces three main novel elements: (a) At conceptual level, it combines concepts from the risk management and the cyber threat intelligence areas and through those defines a unique process that consists of a systematic collection of activities and steps for effective risk management of CIs; (b) It adopts Machine Learning (ML) models such as K-Nearest neighbours (KNN), Naïve Bayes (NB), and Neural Network (NN) to predicate possible risk types, empowering organisations with early warnings that allow them to plan ahead and prevent potential attacks; (c) It introduces a software tool (i-CSRMT) to automate the risk management activities. The tool provides a comprehensive workflow to guide users through individual activities, starting with defining the risk analysis context and applying risk controls. i-CSRMT serves as an additional component of the proposed framework that enables asset criticality, risk and control effectiveness calculation for a continuous risk assessment.

The paper also describes the application of i-CRSM and i-CSRMT to a real case study to demonstrate the overall applicability of the framework. The result shows that ML classifiers are able to predicate the risk types with high accuracy and i-CSRM is effective to management the risk on the studied context. To the best of our knowledge, this is the first time that such application and results are reported on the literature.

2. Related Works

This section provides an overview of existing works which are relevant to the proposed i-CSRM. We focus on four main areas: cyber threat intelligence, risk management, risk predictions using machine learning techniques and risk management standards.

2.1. Cyber Threat Intelligence (CTI)

(Conti, Dargahi and Dehghantanha, 2018) elucidates the increasing number of cyber-attacks that requires cybersecurity and forensic specialists to detect, analyse and defend against cyber threats in almost real-time. In practice, timely dealing with such a large number of attacks is impossible without intensely perusing the attack features and taking similar intelligent defensive actions; this, in essence, defines cyber threat intelligence notion. (Kure and Islam, 2019) targeted to progress the appreciation of the perception of CTI by awarding a much-needed definition of CTI and producing an idea of the intelligence creation method. In their paper, Abu et al., (2018) identify challenges related to CTI, such as threat data being overloaded, quality of threat data that is shared amongst community members, privacy and legal issues, which governs the lawful sharing of data and the interoperability issues faced by threat sharing platforms and standards used by the platforms. However, with all these challenges, adopting CTI by organisations to help them minimize future threats still outweighs its lack of adoption.

2.2. Risk Management

(Rød et al., 2020) presents an approach on how risk management standards can be extended to a critical infrastructure resilience management framework. Focusing in particular on the organizational and technological resilience domains, which are considered those that can most readily be controlled by critical infrastructure operators, the article presents one of the resilience assessment techniques in some detail to operationalize the overall management framework. (Izuakor and White, 2016) suggested a novel approach for assessing critical infrastructure asset identification using a multi-criteria decision theory to address the difficulties of identifying crucial assets. The current methodology stops short of providing a systematic framework for making important decisions.(Bialas, 2016)proposed a novel formal risk assessment approach for dealing with dangerous accidents' internal and external effects in sensitive infrastructure. This study did not consider interdependencies, and it also did not have a framework for determining risk level and control mitigations. (Cherdantseva et al., 2016) conducted a study of the current state of cybersecurity risk management utilising SCADA systems. The study looked at the plurality of risk management methods that evolve or contribute in the framework of the SCADA method. They were evaluated and tool-supported in terms of their goals, implementation domain, risk management principles, effect assessment, and sources of probabilistic evidence. Regardless of the various risk reduction methods for SCADA structures, the requirement for a holistic solution that

includes all risk management processes remains unmet.(Sapori E, Sciutto M and Sciutto G, 2014) proposes a risk-based methodology assess security management systems that were applied to railway infrastructure. The methodology analysed the system, integrates technological, human and procedural aspects by using flow charts. However, identifying critical assets were not the main focus of this paper.In (Islam et al., 2017), there is an illustration of a risk management framework that helps users with cloud migration decisions, following the necessary risk management principles. This framework is essential as it enables users in identifying risks based on the relative importance of migration objectives and risk analysis with the semi-quantitative approach.

2.3. Machine Learning for Risk prediction

The literature has also presented work on fundamental concepts and principles of machine learning and their application to critical infrastructure systems. (Gupta et al., 2020) explored machine learning and deep learning models to make intelligent decisions concerning attack identification and mitigation. They proposed ML-based secure data analytics architecture (SDA) to help classify attack input data. Their threat model addresses research challenges in SDA using different parameters such as reliability, accuracy and latency. (Husák et al., 2018) provides a survey of prediction and forecasting methods applied in cybersecurity. They discuss four main tasks: attack projection, recognition of intention, the prediction of next moves, and intrusion prediction. They further discussed the application of machine learning and data mining in threat detection. The results indicate that suitability for machine learning is needed to understand risk and intrusion predictions. Future research needs to focus more on improvements in attack prediction and its utilization in practice.

(Lilly et al., 2019) argued that despite significant advancements in identifying, deterring, and mitigating cyber incidents, NATO agencies are discontented, along with the intelligence agencies whose strategy against cyber incidents is primarily reactive and implemented rather than being executed before attacks. They proposed an indications and warning (I&W) framework for the cyber-domain. They have applied that framework and examining its effectiveness in the private sector and also deployed it on an actual case. The research finds that indications and warning frameworks effectively detect cyber threats and risks even before they occur in the private sector infrastructure networks. Future research should close the gap and increase understanding of how governments can apply this framework and integrate it within the existing processes.(Singh et al., 2020) describe that machine learning techniques are used to understand 5G network infrastructures with the emerging IoT and 5G infrastructures. The researchers find that it is possible to deploy power-optimized technology in a way that promotes the network's long-term sustainability. They propose a machine learning-based network sub-slicing framework in a sustainable 5G environment to optimise the network load balancing issues. Future research should focus on using machine learning to enhance the stability and sustainability of network infrastructures. (Tanwar et al., 2019) make use of blockchain technology and machine learning to improve a system's accuracy and provide precise network results and resilience against attacks. The researchers find that machine learning and blockchain technology can be used in intelligent applications such as Unmanned Aerial Vehicle (UAV) and smart cities. Future research should consider the issues and challenges in risk management and assessment in blockchain technology.

2.4. Risk Management Standards

There are a number of standards that provide a comprehensive guideline for performing risk management activities. ISO 31000 (ISO 31000) emphasises on understanding organisational internal and external context before performing any risk management activiteis. It includes a systematic process which can be applied to different risk type including project, financial and safety and used by any organisation type. Additionally, the standard provides a list of definitions and set of principles for risk management. ISO 27005 (ISO 27005) provides guidelines for a systematic and process-oriented information security risk management approach. A process is described for the systematic identification, assessment, and treatment of risks, the result of which is a prioritised list that is then to be continuously tracked. The assessment of risk is based on various influencing variables, such as criticality of company assets, extent of vulnerabilities, or impact of known security incidents. ISO 27001 (ISO 27001)provides a list of requirements for the information security management system. ISO 27005 satisfies the requirements related risk management defiend by ISO 27001. ISO 27001 considers risk management as a core component for the overall security management. The National Institute of Standards and Technology SP 800-39 (NIST 800 -39) provides guidelines for managing risk to organizational operations and assets. Risk management is considered as a holistically from every aspect of the organisation including organisation, mission, process and information system level. NIST cyber security improvement framework(NIST -CSF) aims to improve security for the the Critical Infrastrcuture (CI) using four implementation tiers (i.e., partial, risk informed, repeatable, adaptive) to demonstrate the organisation view about cyber-security risks and the processes in place to manage those risks. The tiers includes three main components. i.e., risk management process, program and external participation. It considers profiling to move from current state to target state based on

the achievement of cyber security risk management goals. The Centre for Internet Security Critical Security Controls (CIS) provides a prioritised set of actions that alleviate the most coordinated attacks against systems and systems and can be applied for critical infrastructure sectors. CIS provides an effective security defence based on 20 critical high level controls which are classified as basic, foudantional and organisational. Each control inclues a number of sub-controls , hence there are total 148 sub-controls that map with the other relevant standards.

The above mentioned works are important and contribute to the improvement of the cyber security risk management domain. However, little effort is taken relating to integrating the threat intelligence data and risk prediction for the overall risk management activities. The existing standards provide a generic guideline for risk management activities. For instance, ISO 31000 is generic and lack of guideline how to manage specific risk, whereas ISO 27005 does not include any specific information or details on the implementation of the risk management process. NIST CSF also provides limited detailed to measure a specific implementation tier for risk management. Our work contributes to address these limitations by proposing an Integrated Cyber Security Risk Management (i-CSRM) framework and relevant tool support. i-CSRM integrates threat intelligence, risk predication and effectiveness of controls within an automated risk assessment and management process for the Critical Infrastructure protection.

3. Integrated Cybersecurity Risk Management (i-CSRM)

The proposed integrated Cybersecurity Risk Management (i-CSRM) framework makes use of a number of open security, vulnerability and control repositories (such as the Common Weakness Enumeration (CWE), the Common Attack Pattern Enumeration and Classification (CAPEC), the CIS Critical Security Control (CSC)) and widely used techniques such as CTI and Machine Learning models for the risk assessment, predication, and management activities. In particular, CTI provides a detailed understading of existing threats in terms of threat actor properties, indicator of compromise and TTP. CTI reviews the context of threats that impact on the organisation and supports the risk management activities to determine the risk level and relevant controls. Therefore, CTI provides a number of benefits interms of detecting the relevant threats, possible IoCs and TTPs to guide which vulnerabilities are more exploitable within a specific context. i-CSRM also integrates the CWE and CAPEC to identify the weakness and relevant threats within the systems. It also includes the CIS controls to determine the relevant controls for the risk mitigation. These standards and practive provide identification of vulnerabilities and controls for any specific context. Finally, a number of machine learning modles are considerd for the risk predication including K-Nearest neighbours (KNN), Naïve Bayes (NB) , and Neural Network (NN). The adoption of standards and techniques provide a wider applicability of the i-CSRM and improve an efficient risk management practice. i-CSRM consists of three components, i.e., conceptual view, process and tool. The conceptual view includes a necessary concepts for the risk management activities. The process provides a systematic list of activities that helps organisations to understand the associated risks and the necessary control measures to align with the business goal This section provides an overview of these components.

3.1. i-CSRM Conceptual Language

An important part of understanding the context of the risk analysis depends on clear conceptual elements to represent and model that context. It basically requires a straightforward understanding and precise reinterpretation of abstract ideas or principles to understand what a the system, service etc are, what they do, how they achieve clear objectives, and how they can be implemented (Chen, 1976). To support this, i-CSRM includes a conceptual language with concepts that supports risk assessment and management activities as well as contextual information. The concepts of the language are given below:

- **Actor:** An actor represents an individual, such as an organisation or a human user, that has a strategic goal within its organisational context and performs specific activities (Castro, Kolp and Mylopoulos, 2002). In other words, actors could be an organisation, functional department or set of people involved in providing, requesting or receiving critical services through many forms of information exchange. Actor can be internal or external. The internal Actor is the critical infrastructure organisation that supplies infrastructure and other services needed to run its operations and has skilled personnel who play different roles such as risk manager, information technology security analyst, senior engineer. External actors are mainly users outside the organisation who make use of the services provided by the organisation.

- **Assets:**Assets are necessary and have values to the organisation, such as an organisation's application or software. The asset concept consists of sub-classes such as asset types, criticality and asset goal.Asset profiledescribes the necessary descriptive information about the many components of all the organisation's asset types.
- **Goals:**The goal of any critical infrastructure includes; the concealment of sensitive data against unauthorised users, ensuring the organisation's assets are made available and accessible to the end-users, and the assets' ability to perform their required functions effectively and efficiently without any disruption or loss of service. The asset goals include;Availability (A), Integrity (I), Confidentiality (C), Accountability (ACC) and Conformance (CON).
- **Threat Actor:**Threat actors are individual, groups or organisations with malicious intents to execute a cyber-attack. It is necessary to identify and characterize possible threat actors for the organisation. It includes a number of properties to under the theat actors such as Skill, Motivation, Location, Resources, Size, and Opportunity for intelligence analysis of the threat.
- **TTP:**This concept describes the specific adversary behaviour that a threat actor exploit for an attack. TTP needs a number of resources such as tools, infrastructures, capabilities and right skill for a theat actor. TTP is one of the core properties for the cyber threat intelligence analysis. A threat actor uses TTP to plan and manage an attack by following a specific technique and procedure. They involve the pattern of activities or methods associated with a particular threat actor and consist of the threat actor's typical behaviour (attack pattern) and specific software tools that can be used to perform an attack.
- **Indicator of Compromise:** This concept contains a pattern that can be used to detect suspicious or malicious cyber activities. IOC is detective in nature and are for specifying conditions that may exist to indicate the presence of a threat along with relevant contextual information. Organisations should be aware of the data associated with cyber-attacks, known as indicators of compromise (IOC) as a prt of CTI analysis. The sub-classes includes network indicator, host-based indicator and email indicator to detect the pattern.
- **Vulnerability:**Vulnerability is the weakness or mistake in an organisation's security program, software, systems, networks, or configurations targeted and exploited by a threat actor to gain unauthorised access to an asset (system or network) using TTP. There are several ways an attacker can exploit vulnerabilities in critical infrastructures, thereby causing severe damage. This could be from a threat actor only being able to view information and to a worst-case scenario.
- **Threat:**The threat is the possibility of a malicious attempt to damage or disrupt an organisations asset (systems or networks), access files and infiltrate or steal data. The threat is identified as an individual or group of people attempting to gain access or exploit a vulnerability of an organisation's asset or the damage caused to hinder the organization’s ability to provide its services. Threats such as denial of service or malware attacks are famous threats to critical infrastructures, causing security challenges to the interconnected devices (Baldoni, 2014).
- **Risk:** Risk is defined as the probable exposure of a threat due to the exploitation of the relevant vulnerabilities which impact on the confidentality, integrity and availability of the assets. Organisations cannot wholly avoid the Risk; however, it is the actors' role to ensure that risks are kept to a minimum level to achieve their goals. The risk can pose any potential conseqncnes relating to financial loss, reputational damage, privacy violation non-compliance consequences, disruption of any service delivery. To understand a cyber-attack, we have to study the nature of the attack and its motivation (Gandhi et al., 2011). The severity of risk is estimated based on the information about the threat actor, vulnerability factors and the impact of a successful exploit affecting the security goals of the assets to be gathered.
- **Controls:** These are the security mechanism to tackle the identified risks for the overall business continuity. Generally, the controls are modelled based on its functions such as corrective, detective and preventive. Preventative controls aims to stop unwanted or unauthorized activity from occurring and are designed to be implemented prior to a threat can materialised; Detective controls detect errors and irregularities, which have already occurred and ensured their immediate correction. Corrective controls help to mitigate damage once a risk has materialised. This means that the level of attack determines the type of control used, and the effectiveness of the existing controls is evaluated. The CIS_CSC recommended a list of controls that we adopt for the proposed framework. This means that the level of attack determines the type of control to be used and the effectiveness of the existing controls. To evaluate the effectiveness of the existing controls, an assessment of each control objective is carried out. We apply a set of criteria: Relevance- The level to which the control addresses the relevant control objectives under analysis. Strength- The strength of the control is determined by a series of factors. Coverage means the levels at which all significant risks are addressed. Integration- The degree and manner in which the control reinforces other control processes for the same objective—traceability- How traceable the control is, which allows it to be verified subsequently in all respects. The sub-class is control type and control effectiveness.

The relationships between those concepts are shown in Figure 1. An actor represents an entity, an organisation or a human user that generates strategic, operational and tactical plans within its organisational setting. An actor owns a wide range of assets that require several security goals for supporting the business process. In the context of our framework, an Actor is represented as having an interest in the organisation's assets. These assets have security goals such as Confidentiality, Integrity and Availability for the business's continuation and reputation, and the attainment of one or more of the goals is always their focus. Vulnerability is a weakness in an organisation's security program, software, systems, networks, or configurations targeted and exploited by a threat actor to gain unauthorised access to an asset (system or network) using TTP. Risk is the failure of an organisation or individual to achieve its goals due to the malicious attempt to disrupt its critical services by a threat. The threat actor is a type of Actor with malicious intent characterised by their identity, suspected motivation, goals, skills, resources available for them to carry out a successful attack, past activities, TTP used to generate a cyber-attack and their location within the organisation's network. A threat actor uses TTP to plan and manage an attackby following a specific technique and procedure. The CTI information such as TTP and threat actor properties are used for the risk assessment activities. They involve the pattern of activities or methods associated with a specific threat actor and consist of the threat actor's specific behaviour (attack pattern) and specific software tools that threat actors can use to perform an attack leaving behind the attack's incident. The incident is the type of event that represents information about an attack on the organisation. Some specific components determine the type of incident, such as threat types, threat actor's skill, capability and location, assets affected, parties involved, and time. With a specific attack pattern, the organisation tends to think broadly by developing a range of possible outcomes to increase their readiness for a range of possibilities in the future. With Indicators, a pattern that can be used to detect suspicious or malicious cyber activity is gathered.

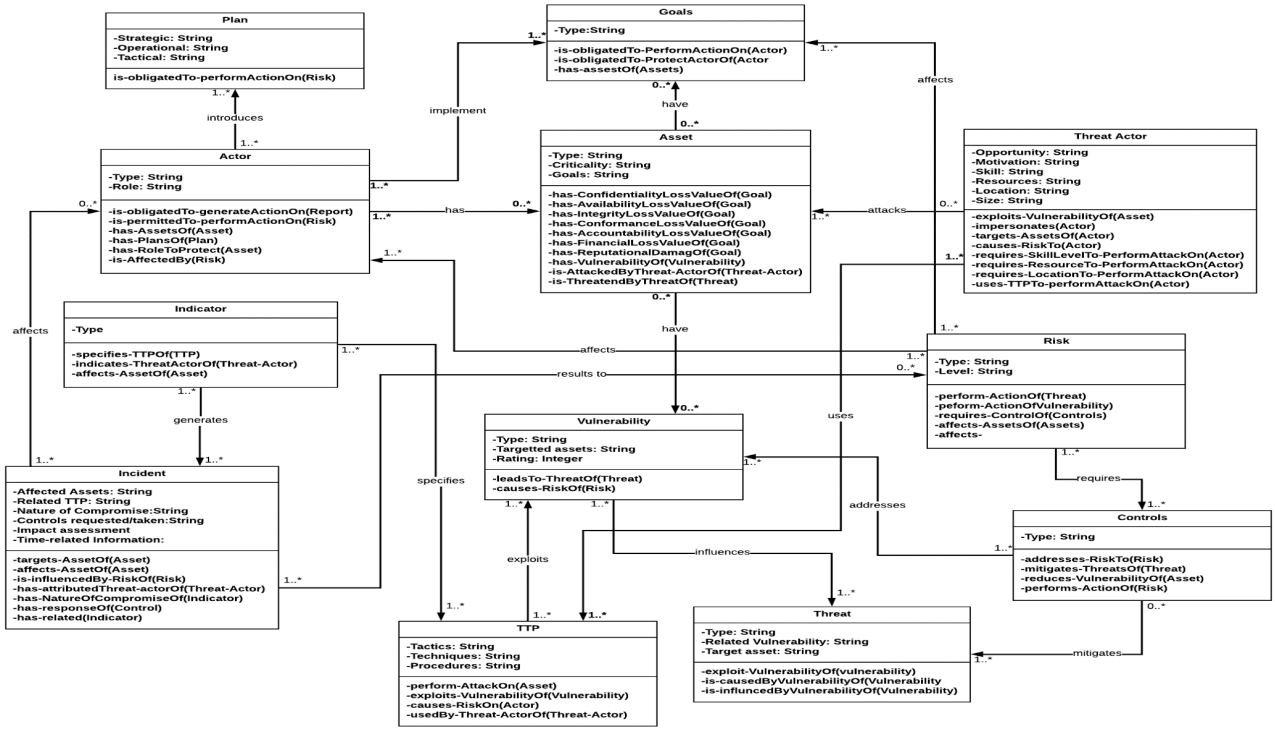

Figure 1: A meta-model for i-CSRM at an organisational level

3.2. Process

The i-CSRM considers a systematic process for risk assessment, predication, and control as presented in Figure 2. The process establishes a solid relationship between multiple steps using the concepts for the effective delivery of an expected outcome. An activity deals with linked tasks that are interdependent that receive and convert one or more input into an output artefact (Knight and Burn, 2005). The i-CSRM process is decomposed into activities and steps that provide a lower level of detail. Activities 1 and 2 focus on the context of the risk assessment and in particular an organisation's scope. This helps to gain a comprehensive understanding of supported assets, functions, goals and essential security requirements. Activity 3 gathers vulnerability and threat information from multiple sources through various means, to address vulnerabilities protect assets and respond to threats. Activity 4 determines the risk level and provides a risk register with the previous activities' data. Activity 5 implements control measures and evaluate the effectiveness of the existing control. The effectiveness of one activity determines the essential elements of information needed for the next activity. Therefore, activity 5 evaluates the effectiveness of the existing controls. Each activity specifies the steps that need to be followed, and each step identifies the needful inputs, participating actors and final output. Primarily, the output of each activity serves as the input to the next activity that follows it. The effectiveness of the whole process is mainly achieved when conducted with the support of security experts delegated by an organisation to oversee the i-CSRM analysis. Hence, an organisation must delegate suitable actors to participate and supervise the implementation of the process. In the next few sections, we provide more information for each of the activities.

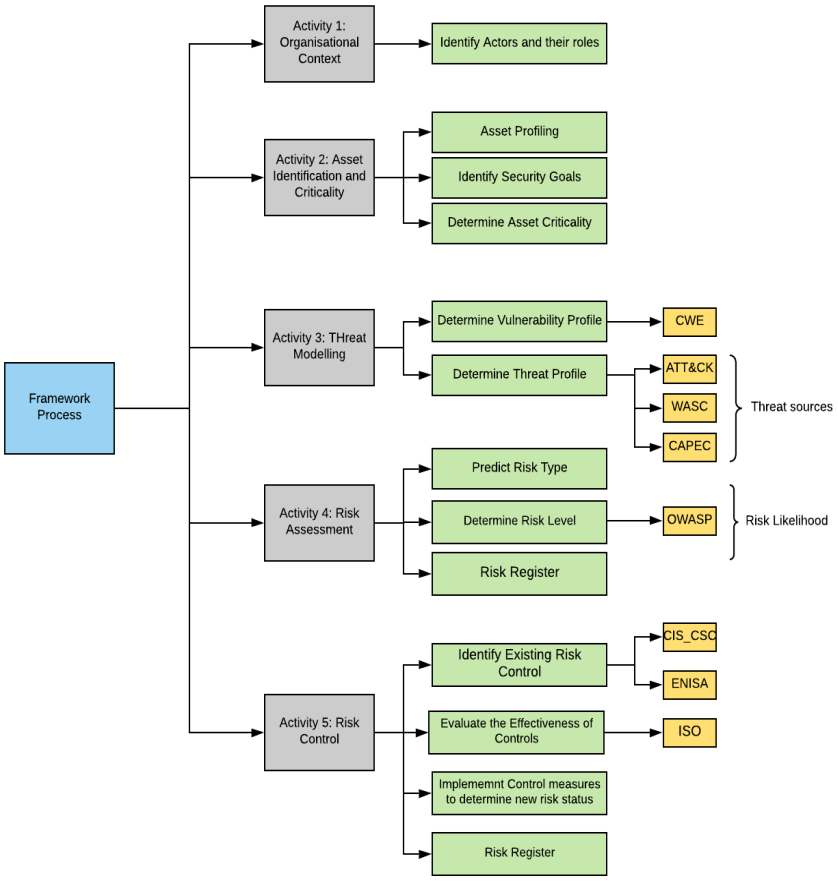

Figure 2: i-CSRM process

3.2.1. Activity 1: Organisational context: Every organisation exclusively operates within a defined scope and available resources. The Organisational context activity aims to better define the risk assessment organisational context by identifying relevant stakeholders, critical assets, security goals and how they impact risk management and viability. A stakeholder is any entity with a conceivable interest or stake in an activity (Goodpaster, 1991). A stakeholder can be an individual, group of individuals, or an institution affected by or influences an activity's impact, for example manager or adminstrator. In i-CRSM a stakeholder is modelled as an actor. To successfully execute the process and achieve this activity, it is essential to obtain a comprehensive picture of actors and their roles in meeting requirements. This becomes important in identifying and avoiding potential conflict of interests and other issues such as the actors responsible for the security and maintenance of organisational assets.

Step 1: Identification of Actors and their roles: this step aims to identify and list the relevant actors. As described in the previous section, an actor represents an entity such as an organisation or human user with a strategic goal within its organisational setting, who carries out specific activities and makes informed decisions. Actors interact with the organisation's systems or relationships by providing technical and nontechnical support or services to the organisation. The nature of communications between actors needs to be clearly balanced, reconciled, interpreted and managed accordingly. The organisation's activities require an active set of actors to carry out various tasks to guide and lead the organisation in achieving its goals and ensuring its successful operations. In this case, actors can be identified as internal and external actors. The internal actor is the organisation itself that supply infrastructure, network facilities and other services needed to run its operations and has skilled personnel who play different roles such as information technology security analyst, risk manager and senior engineer. External actors mainly include users who use the organisation's services and third-party vendors who provide other services such as internet services.

3.2.2. Activity 2: Asset Identification and Criticality: This activity aims to identify and prioritise assets in terms of their boundary, components and assigning weights to the assets based on the importance they hold for the organisation. Assets are specific units such as hardware, a database, application, or program that support the delivery and usage of an organisation's services. Furthermore, to support organisations in assessing each asset's criticality, a decision support system using fuzzy set theory is provided. A fuzzy set theory provides a way of absorbing the uncertainty inherent to phenomena whose information is unclear and uses a strict mathematical framework to ensure precision and accuracy and the flexibility to deal with both quantitative and qualitative variables (Zimmermann, 2011). It can be used for approximate reasoning, easy to implement and adopt individual perception without incurring complexity within the risk management process. This activity includes three steps; identify assets and their goals, determining asset criticality, and identifying the business process. The resulting critical asset list is then used to assess vulnerability and threat identification in Activity 3.

- **Step 1: Asset Profile:** This step's basis is to profile assets in terms of their components, boundaries and assigning weight to the assets based on assets vital to the organisation. Assets are specific units such as a database, application, or program that support the delivery and usage of an organisation's services. To create asset profiles, a Security Analyst is involved in identifying assets by considering the core functions of the assets, alongside other subcomponents essential to achieving and maintaining crucial functions. Important asset information can be gathered by reviewing background materials, including independent audit/analytical reports, interviewing the critical infrastructure users, and physical observation of organisational assets. Besides, asset specification and management documentation provide essential details about the organisational asset.

- **Step 2: Identify Asset Security Goals:** Identifying assets security goals is vital for an organisation to determine what critical views of security must be ensured by each asset during processing, storage, or transmission by authorized systems, applications, or individuals. It also supports determining the impact that may result from accessing assets in an unauthorized manner for use, interruption, change, disclosure. Therefore, a Security Analyst considers a set of security goals that each asset aims to achieve. The consequential impact that may ensure the compromise of the security goals and the level of protection needed can be easily determined. There are different asset categories we consider for asset criticality. They include software, data, hardware, information communications and network and people. To better support this step, we have also defined a set of asset security goals every asset must aim to achieve. These are:
    - **Asset Availability (A):** Availability refers to ensuring that an asset is made available and accessible to authorised users when and where they need it.
    - **Asset Integrity (I):** Asset integrity refers to an asset's ability to perform its required functions effectively and efficiently without disrupting or losing its services.
    - **Asset Confidentiality (C):** Asset confidentiality refers to assets staying secured and trusted and preventing unauthorised disclosure of sensitive data.

- **Accountability (ACC):** This asset goal requires that attack or incident actions that occur on an asset are tractable to the responsible system or actor.
- **Conformance (CON):** This asset goal ensures that the assets such as services meet the specified standard.

- **Step 3: Determine Asset Criticality:** This step aims to identify and prioritise an organisation's critical asset by assessing those assets' primary security goals. In other words, the criticality of each asset is based on its relative importance. Asset criticality is imperative for prioritizing and developing actions that will reduce risks to the asset, improve asset reliability, and define strategies for implementing the appropriate controls. To ensure validity, consistency, and support stakeholders in assessing each asset's criticality, a decision support system using fuzzy set theory is provided. Fuzzy set theory plays a vital role in the decision process enhancement. It helps to deal with or represent the meaning of vague concepts, usually in situation characterization such as linguistic expressions like "very critical". Fuzzy logic, introduced by (Zadeh, 1988), is one of the best ways to deal with all types of uncertainty, including lack of knowledge or vagueness (Markowski and Mannan, 2009). This system provides a methodology for computing directly with the word. Fuzzy set theory is a generalisation of classical set theory that provides a way to absorb the uncertainty inherent to phenomena whose information is vague and supply a strict mathematical framework to ensure precision and accuracy, as well as the flexibility to deal with both quantitative and qualitative variables.

  Phase 1: Development of a Fuzzy Asset Criticality System (FACS): Criticality is the primary indicator used to determine the importance of assets to an organisation. After the different assets have been identified, we determine the criticality based on their relative importance using the Fuzzy Asset Criticality System (FACS).
  - **Fuzzification:** FACS determines asset criticality by using (C, I, A, CON and ACC) as the five fuzzy inputs for assessing the criticality of individual assets and assigning a level of criticality. Each input is assigned five fuzzy labels Very Low (VL), Low (L), Medium (M), High (H) and Very High (VH), for assessing the level of the fuzzy output Asset criticality (AC) value which is assigned five fuzzy labels Very Low Critical (VLC), Low Critical (LC), Medium Critical (MC), High Critical (HC) and Very High Critical (VHC) of individual assets. The details of fuzzy sets applied in the first step of the fuzzy inference system are presented in Table 1, which shows the numerical ranges in which fuzzy sets are selected based on them. The membership functions for AC also are depicted on a scale of 1 to 5. Figure 3 shows the structure of the FACS.

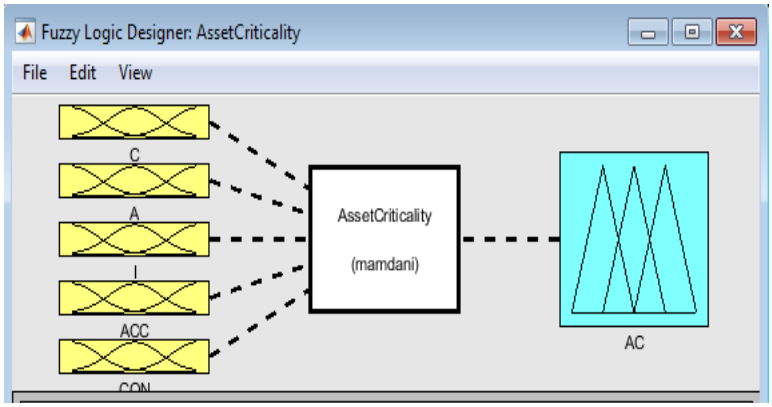

Figure 3: Structure of the Fuzzy Asset Criticality System (FACS)

Table 1: Fuzzy Ratings

Features	Asset Factors	Description	Linguistic Terms	Crisp Rating	Interpretation
Input	Confidentiality (C)	How much data could be disclosed, and how sensitive is it?	Very High (VH)	5	All data disclosed
			High (H)	4	Extensive critical data disclosed
			Medium (M)	3	Extensive non-sensitive data disclosed

			Low (L)	2	Minimal critical data disclosed
			Very Low (VL)	1	Minimal non-sensitive data disclosed
	Availability (A)	How many services could be lost, and how vital is it?	Very High (VH)	5	All services completely lost
			High (H)	4	Extensive primary services interrupted
			Medium (M)	3	Extensive secondary services interrupted
			Low (L)	2	Minimal primary services interrupted
			Very Low (VL)	1	Minimal secondary services interrupted
	Integrity (I)	How much data could be corrupted, and how damaged is it?	Very High (VH)	5	All data corrupt
			High (H)	4	Extensive seriously corrupt data
			Medium (M)	3	Extensive slightly corrupt data
			Low (L)	2	Minimal seriously corrupt data
			Very Low (VL)	1	Minimal slightly corrupt data
	Accountability (ACC)	Are the threat actors traceable to an individual?	Very High (VH)	5	Completely anonymous
			High (H)	4	Fully traceable
			Medium (M)	3	Highly traceable
			Low (L)	2	Possibly Traceable
			Very Low (VL)	1	Minimal Traceable
	Conformance (CON)	How much deviation from specified behaviour constitutes conformance?	Very High (VH)	5	Full variation
			High (H)	4	High profile variation
			Medium (M)	3	Clear variation
			Low (L)	2	Low variation
			Very Low (VL)	1	Very low variation
Output	Asset Criticality (AC)	How critical is the asset to the organisation?	Very Critical (VC)	5	Extremely critical and is of high value to the CI organisation, it requires an extreme level of protection
			Highly Critical (HC)	4	High importance to the organisation and requires a high level of protection.
			Medium Critical (MC)	3	The asset is moderately important to the organisation and requires moderate protection
			Low Critical (LC)	2	The asset is of minimal importance and does not require many levels of protection.
			Very Low Critical (VLC)	1	The asset non-critical and requires a very low level of protection

Phase 2: Rules: There are many fuzzy inference methods; however, this research uses the Min-Max fuzzy inference method proposed by Mamdani (Cordón, 2011).This research employs Mamdani's method due to several advantages (Cord, 2001):

- It is suitable for engineering systems because its inputs and outputs are real-valued variables
- It provides a natural framework to incorporate fuzzy IF-THEN rules from human experts
- It allows for a high degree of freedom in the choices of fuzzifier, fuzzy inference engine, and defuzzifier so that the most suitable fuzzy logic system for a particular problem is obtained. It provides a natural framework to include expert knowledge in the form of linguistic rules.

We used 125 IF-THEN rules to provide a database by mapping five input parameters (C, A, I, CON and ACC) and AC value. The rules are designed to follow the logic of the Asset criticality evaluator. A number of the IF-THEN rules of the developed system are shown in Figure 4.

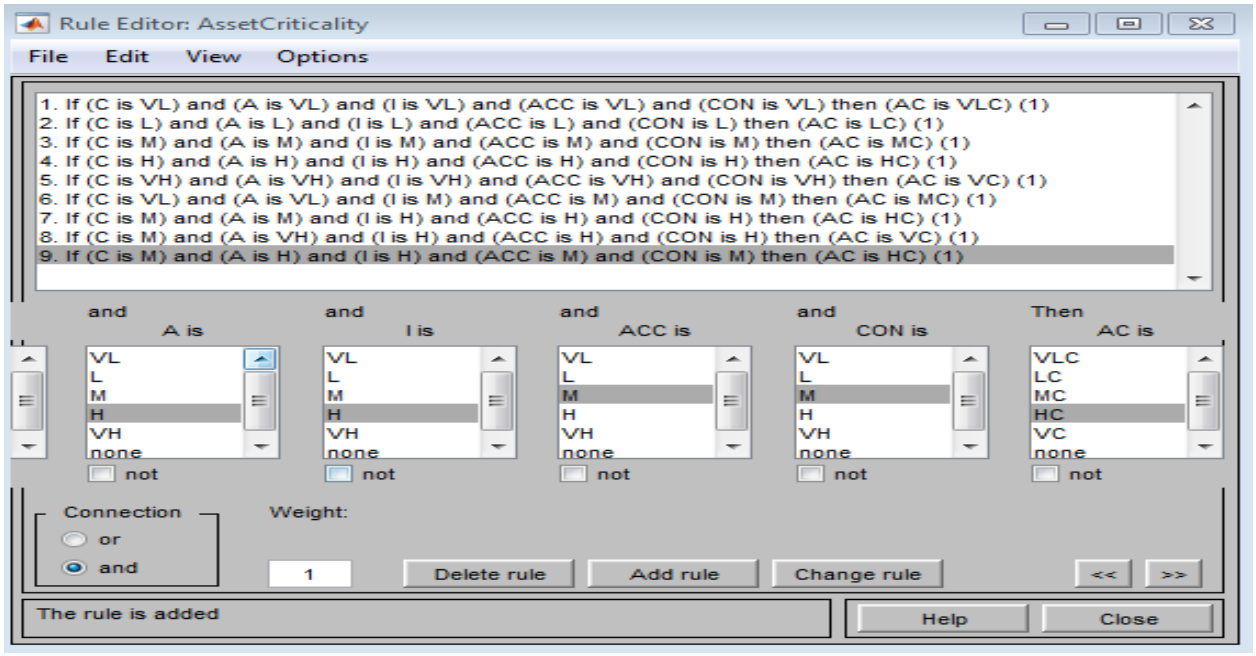

Figure 4: Rules Set for FACS

Phase 3: Inference Engine: An inference engine attempts to create solutions from the database. In this paper, the inference engine maps fuzzy input sets (C, A, I, ACC and CON) into fuzzy output set (AC). Figure 5 shows several IF-THEN rules to provide a more understanding of the proposed FACS model.

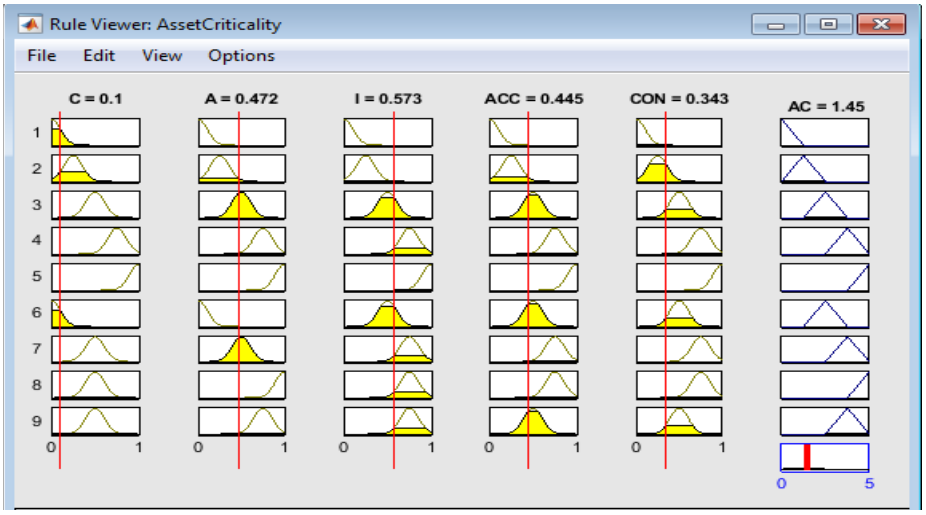

Figure 5: Sample of Rules

Phase 4: Defuzzification: Different methods for converting the fuzzy values into crisp values such as Centre of Gravity (COG), Maximum Defuzzification Technique and Weighted Average Defuzzification Technique. One of the most commonly used defuzzification methods is COG. The COG technique can be expressed as follows:

$$X^* = \frac{\int \mu_i(x) x dx}{\int \mu_i(x) dx} \textbf{(Equation 5.1)}$$

Where x* is defuzzified output, μi(x) is aggregated membership function, and x is the output variable.

Table 2 displays the asset inventory showing the critical level for each asset.

3.2.3. Activity 3: Threat Modelling

The Threat modelling activity focuses on identifying and measuring vulnerabilities and threats related to the assets. It is expected that this activity is performed by a security analyst or someone with similar knowledge. Based on the previous activity's assets, all possible threats that could impact the assets negatively are profiled in a register. However, effective identification and control of threats require an understanding of threat sources, threat actor behaviour, capability and intent (Workman, Bommer and Straub, 2008). This is possible when known attack patterns employed by the threat actor to exploit vulnerabilities are known to allow an organisation to understand and create a threat profile expansively. Because of these considerations, this activity is split into two steps for threat modelling: (i) the determination of vulnerability profile; and (ii) the determination of threat profile.

- **Step 1: Determine the vulnerability profile:** Determining the vulnerability profile is vital because it allows for identifying and assessing vulnerabilities associated with critical assets. This step aims to identify potential asset vulnerabilities that a threat actor may leverage to exploit an asset. It is an essential and delicate task that has an impact on the successful operation of critical infrastructures. The Common Weakness Enumerator (CWE) methodology (Martin, 2007) is used to determine the vulnerability factors as a publicly known vulnerability source. Therefore, to estimate the likelihood of risk, it is necessary to estimate how a particular vulnerability is discovered and exploited. We adopt CWE, which allows for weaknesses to be characterised, allowing stakeholders to make informed decisions when mitigating risks caused by those weaknesses. Each related weakness is mapped to CAPEC and identified by a CWE identifier and the name of the vulnerability type. The CWE gives a general description, behaviour, likelihood of exploit, consequences of exploit, potential mitigation and related vulnerabilities. To apply the CWE methodology, a rating table is presented in Table 2 with corresponding values assigned to the different factors that can help organisations determine the likelihood of risk. Each option has a likelihood rating from 0 to 9, and the overall likelihood falls within high, medium and low, which is sufficient for the overall risk level. Although our work is based on these repositories, it is not bounded by them. A Security Analyst could explore other publicly available sources of vulnerability information, including internal experience, penetration test, catalogues of vulnerabilities available from industry bodies, national government, and legal bodies. The questions can also be extended to meet the organisation's need.

Table 2: Vulnerability Factor Rating

Vulnerability Factors	Vulnerability ID	Description	Likelihood rating	
			Weight	Value
Ease of discovery	EoD	How easy is it for vulnerability to be discovered?	1	Practically impossible
			3	Difficult
			7	Easy
			9	Automated tools available
Ease of exploit	EoE	How easy is it for vulnerability to be exploited?	1	Theoretical
			3	Difficult
			5	Easy
			9	Automated tools available
Awareness	Awa	How well known is this vulnerability to the threat actors?	1	Unknown
			4	Hidden
			6	Obvious
			9	Public knowledge
Intrusion detection	I_D	How likely is an exploit to be detected?	1	Active detection in application
			3	Logged and reviewed
			8	Logged without review
			9	Not logged

- **Step 2: Determine threat profile:** Determining the threat profile is essential because it allows for the identification and understanding of threat characteristics. To determine threats, a structured representation of threat information is required that is expressive and all-encompassing due to the dynamic and complex nature of a CPS. Therefore, this step effectively identifies the threat types, target assets, threat actor factors, TTP, and compromise indicators likely to affect a critical infrastructure's ability to deliver its services. As in previous steps, our work does not bound to a specific repository. Although we recommend using CAPEC (Common Attack Pattern Enumeration and Classification) (Barnum, 2008) and WASC (Web Application Security Consortium) (Consortium, 2009) to define the potential threat, provide context for architectural risk analysis, and understand trends and attacks to monitor, a security analyst could explore other available sources of threat information.
  Moreover, to better support this step, we propose the following procedures to support the creation of a comprehensive threat profile:
  - **Threat type:** To create a comprehensive threat profile, organisations need to identify the potential threats of assets that a threat actor may leverage to attack. The Security Analyst needs to back up his claim with a solid foundation of Information sources.
  - **Threat Actor factors:** Effective identification and control of threats require an understanding of threat sources, threat actor behaviour, skill, resources required, capability and intent (Workman, Bommer and Straub, 2008). Therefore, we adopt the OWASP methodology that considers various threat actor factors such as skill level, size, motivation, location, resources, and opportunity to understand the attack and its trend. Using these threat actor factors, a Security Analyst can determine the likelihood of an attack and the severity of the threat. This will provide the ability to create an impact rating for threats. We have developed a set of threat actor factors, and corresponding options of like hood rating as presented in Table 3.

Table 3: Threat Actor Factors Rating

Threat Actor factors	Description	Likelihood rating	
		Weight	Value
Skill level	How technically skilled is the threat actor?	1	No technical skills
		3	Some technical skills
		5	Advanced computer user
		6	Network and programming skills
		9	Security penetration skills
Location	Through what channel did the threat actor communicate to reach the vulnerability?	1	Internet
		8	Intranet
		8	Private Network
		7	Adjacent Network
		5	Local Network
		2	Physical
Motive	How motivated is the threat actor to find and exploit the vulnerability?	1	Low or no reward
		4	Possible reward
		9	High reward
Resources	What resources are required for the threat actor to find and exploit the vulnerability?	0	Expensive resources required
		4	Special resources required
		7	Some resources required
		9	No resources required
Opportunity	What opportunities are required for the threat actor to find and exploit the vulnerability?	0	Full access required
		4	Special access required
		7	Some access required
		9	No access required
Size	How large is the group of the threat actor?	2	Developers
		2	Systems administrators
		4	Intranet users
		5	Partners
		6	Authenticated users
		9	Anonymous internet users

- **Determine Tactics, Techniques and Procedures (TTP) and Indicator of Compromise (IOC):** TTP and IOC involve the pattern of activities used by a threat actor to plan and manage a cyber attack, thereby compromising critical assets. The different TTP types include *initial access, execution, credential access, persistence, privileged escalation, defence evasion, collection, lateral movement, exfiltration* and *command and control*. The different IOC include *network indicators, email indicators* and *host indicators.* Therefore, we adopt the ATT&CK (adversarial tactic, techniques and common knowledge) framework developed by MITRE to document standard TTP used to target, compromise and operate in an enterprise network. Using such framework makes our approach easier to adopt due to the wide usage of the ATT&CK framework. To calculate the risk level and know the appropriate controls to protect the organisation's assets, information about TTP must be known. We have defined possible TTP and IOC that are frequently employed when exploiting a vulnerability as shown in Table 4.

Table 4: TTP and IOC (Tactic, 2017)

Tactics type	Techniques	Procedure	IOC
Initial access	Spearphishing link	It employs links to download malware in an email by electronically delivering social engineering targeted at a specific individual or organisation.	Email, Network
	Drive-by compromise	A threat actor gains access to a system by visiting a website over the ordinarybrowsing course. The website is compromised where the threat actor has injected some malicious code.	Network
	Replication through removable media	The threat actor uses a tool to infect connected USB devices and transmit them to air-gapped computers when the infected USB device is inserted.	Host
	Spearphishing attachment	A threat actor attaches and sends a Spearphishing email with malicious Microsoft office attachment and requires user execution in other to execute.	Email
Execution	Command-line interface	The threat actor uses a command-line interface to interact with systems and execute other software during operation.	Host
	Dynamic data exchange (DDE)	Threat actor sends a Spearphishing containing malicious word document with DDE execution.	Host, Network
	Execution through module load	The threat actor uses this functionality to create a backdoor through which it can remotely load and call dynamic link library (DLL) functions.	Host
	Exploitation for client execution	Threat actor exploits a vulnerability in office applications, web browsers or typical third party applications to execute the implant into the victim's machines.	Network
Persistence	Account manipulation	Threat actor adds a created account to the local administrator's group to maintain elevated access.	Host, Network
	Accessibility features	The threat actor uses a combination of keys known as the sticky keys to bypass a user's windows login screen on remote systems during the intrusion.	Host, Network
	Component firmware	Threat actor overwrites the firmware on a hard drive by compromising computer components.	Host, Network
Privilege escalation	External remote services	Threat actors leverage legitimate credentials to log into external remote services	Host, Network
Defense evasion	Disabling security tools	Threat actor disables the windows firewalls and routing before binding to a port.	Host, Network
Credential access	Brute force	Threat actor brute forces password hashes to be able to leverage plain text credentials.	Host, Network
Discovery	Network sniffing	The threat actor uses a tool to capture hashes and credentials sent to the system after the name services have been poisoned.	Host, Network
	Network service scanning	Threat actor used BlackEnergy malware to conduct port scans on a host.	Host
	System information discovery	The threat actor uses tools such as systeminfo that obtains information about the local system.	Host
Lateral movement	Remote services	The threat actor uses putty secure copy client (PSCP) to transfer data or access compromised systems.	Host

	Third-party software	Threat actor distributes malware by using a victim's endpoint management platform.	Host
Collection	Data from information repositories	Threat actor collects information from Microsoft SharePoint services using a SharePoint enumeration and data dumping tool within target networks	Host, Network
	Email collection	The threat actor uses utilities to steal email from archived outlook files and exchange servers that have not yet been archived.	Email, Host, Network
	Man in the browser	The threat actor uses a Trojan spyware program to perform browser pivot and inject into a user's browser and trick the user into providing their login credentials on a fake or modified web page.	Network
Exfiltration	Data encrypted	The threat actor uses malware such duqu to push and execute modules that copy data to a staging area, compress it, and XOR encrypts it.	Host
Command and control	Commonly used port	The threat actor uses duqu, which uses a custom command and control protocol that communicates over commonly used ports and is frequently encapsulated by application layer protocols.	Network
	Remote file copy	The threat actor used Shamoon malware to download an executable to run on the victim.	Network

3.2.4. Activity 4: Risk Assessment

The output of threat modelling provides a list of vulnerabilities, related vulnerabilities, potential security threats, and assets' impact. The threat register serves as a help to the Security Analyst to orchestrate a risk register's creation and focus on the most potent threats. This activity allows for establishing the risk assessment context by following the threat register and formally approves the risk management activities within the organisation. The activity provides various additional estimations required for the risk evaluation by enabling the determination of risks that are likely to occur, the severity of the risks, and the steps to control or manage the risks. This activity prioritises the risk as high, medium and low.

Step 1: Predict Risk Types: This step proposes using machine learning techniques for predicting risk type, so that appropriate mitigation processes can be implemented. In this context, risk type prediction relies on a pioneering mathematical model such as machine learning for analysing, compiling, combining and correlating all incident-related information and data acquired from previous activities. The machine learning (ML) techniques automatically find valuable underlying patterns within i-CSRM concepts used as features, and then the patterns predict risk types. The i-CSRM features are considered input for the ML classifiers and ML classifiers to predicate the risk type. Therefore, we used well-known classifiers such as K-Nearest neighbours (KNN), Naïve Bayes (NB), the Naïve Bayes Multinomial (NB-Multi), Neural Network (NN) with Ralu activation function at activation layers and sigmoid function at the output layer, Decision Tree (DT), Random Forest (RF), and Logistic Regression (LR) for risk type prediction.

We present data extraction to generate a feature set, which is then further used on the ML classifiers for training purposes. Finally, the test data is used to check the accuracy of the prediction. Figure 6 shows how these features are used to train the classifiers and the step-by-step process of the risk prediction, i.e. the experiment in general. Data collection and extraction were considered from the dataset; feature extraction was carried out on those data and used to train the ML classifiers (NN, RF, LR, NB-Multi DT, KNN and NB). The data were further partitioned into 80% training and 20% testing. We used the widely known 10-fold cross-validation scheme to split the given data into testing and training set and reported the average results obtained over the ten folds. Predictions are carried out on the testing dataset, and accuracy measures the prediction. Also, risks types from multiple industry bodies can be considered because they maintain a regularly updated list of most pressing security risks. For example, the Common Attack Pattern Enumeration and Classification (CAPEC) provide a comprehensive list of risks that can be used for understanding and enhancing defense. All these sources can be used.

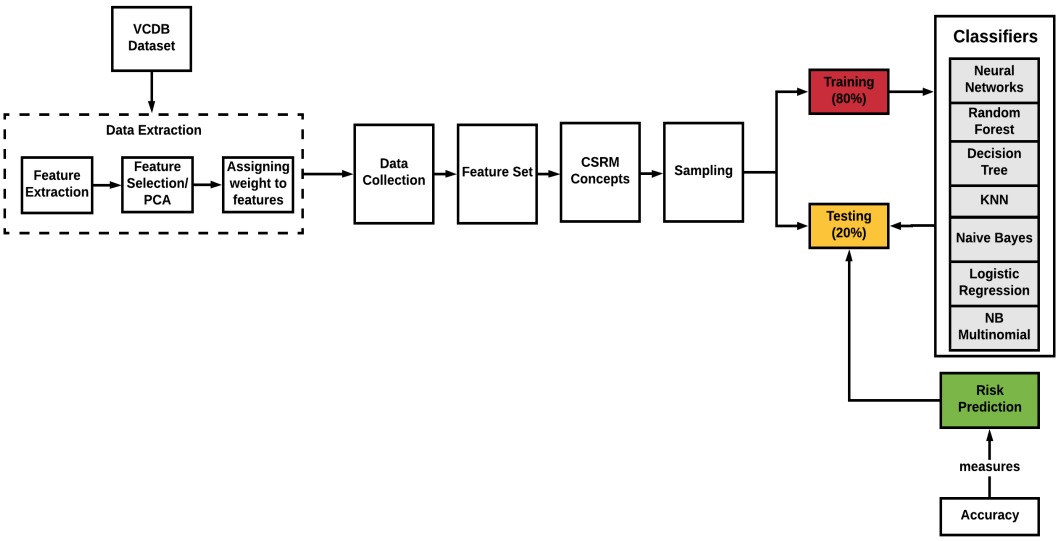

Figure 6: Classification process about the primary analysis

- **Step 2: Determine Risk Level:** After information about the potential risk types, threat, vulnerabilities and assets have been identified and gathered, the next step is to determine the risk level of all the possible risk types predicted. The risk level is usually not known and not estimated correctly. In essence, organisations need to rate security risks that have been identified. Therefore, for the risk level to be estimated, we used the technical impact factors. The technical impacts factors are inclined toward an asset's security goals that include; confidentiality, integrity, availability, accountability, and conformance. Also, iinformation about the threat actor and vulnerability factors needs to be gathered. The aim is to provide a rough estimate of the risk level's magnitude if a risk occurs.
- **Phase 1:** To estimate the overall (L) Likelihood of the risk, threat actor factors and vulnerability factors are put into consideration, as shown in equation 1. Each option has a likelihood rating from 0 to 9, as shown in table 3 and 4. The overall likelihood falls within high, medium and low, sufficient for the overall risk score. Table 5 shows the overall likelihood level.

$$L = \frac{TAF + VF}{2} \qquad (Equation\ 2)$$

Where:
L= Likelihood, TAF = Threat Actor Factors, , VF = Vulnerability Factors

TAF = SL + L + M + Res + Opp +S / n (Equation 3)

Where:
SL = Skill Level, L = Location, M = Motivation, Res = Resource, Opp = Opportunity
S = Size
n = total number of TAF factors (6)

VF = EoE + EoD + Aw + ID / n (Equation 4)

Where:
EoE = Ease of Exploit, EoD = Ease of Discovery, Aw = Awareness, ID = Intrusion Detection
n = total number of VF factors (4)

Table 5: Overall Likelihood Rating

Likelihood	Rating
Low	0.00 – 2.99
Medium	3.00 – 5.99
High	6.00 – 9.00

Phase 2: To Estimate the overall ($Impact_F$) impact of a successful attack, we consider the total loss of the asset's goals, as shown in equation 5. Each factor has a set of options with an impact rating from 0 to 9, as shown in table 6.

$$Impact_F = {}^{AF}/_{n} \qquad (Equation\ 5)$$

Where:
$Impact_F$= Impact Factor
AF = Asset Factors (L_C +L_ A +L_I +L_ ACC +L_ CON)
L_C = loss of Confidentiality, L_A = loss of Availability, L_I = loss of Integrity, L_ACC = loss of Accountability
L_CON = loss of Conformance, n = Total number of the Technical factors (5)

Table 6: Impact Factors

Impact Factors	0 to < 3 (Low)	3 to < 6 (Medium)	6 to 9 (High)
Loss of Confidentiality	Minor disclosure of critical assets	Critical assets are significantly affected	Highly critical assets are extensively affected
Loss of Integrity	Minor compromise of critical assets	Critical assets significantly compromised	All highly critical asset extensively compromised
Loss of Availability	Minor interruption of critical assets	Critical assets significantly interrupted	All critical assets extensively lost
Loss of Accountability	Threats are fully traceable	Threats are possibly traceable	Threats are completely untreatable
Loss of Conformance	A minor breach of compliance requirements	A significant breach of compliance requirements	All compliance requirements significant breached.

The overall impact level rating is three scales Low(0.00-2.99) , Medium(3.00-5.99), and High(6.00-9.00)
Phase 3: Determine Risk Severity: To determine the risk level, we estimate the likelihood and impact are combined to calculate the overall severity of risk using equation 6.

$$\boldsymbol{R_{Level} = L * Impact_F} \qquad (\boldsymbol{Equation\ 6})$$

Where;
R_{Level} = the risk level
I = the impact of the asset goals
L= the likelihood of the attack occurring within a given time-frame
Overall risk severity is rated as high (00-20), medium (21-45), High (46-65), and Critical(66-81)

3.2.5 Activity 5: Risk Controls

This final activity determines the control necessary to tackle the identified risk. This activity advocates to adopt CIS CSC that guides to identify the necessary control. The standard provides 20 control types which are categorized into three different classes. The primary objective of this activity is to specify relevant controls and evaluates the effectiveness of the existing control to determine whether new controls are required to tackle the identified risks. The activity consists of four steps to identify the controls and determine the effectiveness of existing controls.

- **Step 1: Identification of Existing Control Types:** This first step identifies the existing controls and categorises them based on the functionalities, i.e., corrective, detective and preventive, to mitigate the risk. The control categorisation supports the organisation to understand which control types are more implemented within the context. Once the controls are identified, it is necessary to determine the effectiveness of the existing controls.
- **Step 2: Evaluating the Effectiveness of Existing Controls:** This step involves assessing the effectiveness of existing controls and determines the level of effectiveness using criteria rating presented in table 7. This step also evaluates if existing controls are not adequate, then new controls would be recommended for the risk mitigation. If a control does not work as expected, this may cause vulnerabilities leading to risks. Consideration should be given to the situation where a selected control fails in operation, and therefore complementary controls are required to address the identified risk effectively. The controls are evaluated in terms of relevance, strength, coverage, integration, and traceability according to ISO 27005:2011 standard *(GOST, 2009).* For each criterion, a rating score from 1 to 5 is given to measure the effectiveness. Table 7 shows the five different criteria rating.

Table 7: Criteria Rating

Rating		Description
5	Adequate control	The control achieves the objectives intended to mitigate the risks.
4	Adequate control with some areas of improvement	The control achieves the objectives intended to mitigate the risks with evidence of some areas, though not critical, subject to improvement to meet sound controls' requisites.

3	Generally adequate control, with some critical areas	The control mostly mitigates the risks intended to mitigate the risks. However, the characteristics of some of the controls are not entirely consistent with basic sound controls
2	Inadequate control, subject to significant improvement	The control partially achieves the control objectives intended to mitigate the risks
1	Insufficient control	The control is not sufficient to achieve the control objectives intended to mitigate the risks.

The overall effectiveness of the controls is ranked into five scales : Insignificant(0-5), Minor(6-10), Moderate(11-15), Major(16-20), and Critical(21-25)

To find the overall evaluation of each control, equation 7 is given:

$$\mathbf{OE = R + S + C + I + T} \qquad \textbf{(Equation 7)}$$

Where:

OCE = Overall Control Effectiveness, R = Relevance, S = Strength , C = Coverage
I = Integration, T = Traceability

- **Step 3: Implement Control Measures to Determine New Risk Status:** This step involves performing appropriate analysis to measure the status of the risk by implementing the control. Each criterion helps the assessment; a rating score from 0 to 9 is given to measure which control addresses the specific control objective. It helps to further displays the current risk status for each risk type. It presents the risk events and their calculated risk values, and the control measures that can be used to mitigate the risk.
- **Create a Risk Register:** This final step produces a risk register as one of the main output of i-CSRM . The risk register provides a detailed about the risk, vulnerability and threat profile, assets and security goals. It displays the results of the risk calculation in terms of risk level and suitable controls to address the risks.

4. Integrated Cyber Security Risk Management Tool (i-CSRMT)

The i-CSRM tool is an implementation of the i-CSRM process. An automated tool designed to support i-CSRM activities that an organisation uses to perform security risk analysis and control for critical infrastructures. It provides a comprehensive workflow to guide the user through the individual activities, starting with identifying the actors and their roles within the organisation, identifying critical assets, revealing the particularly dangerous threats, risk calculation and finishing with control evaluation. This helps to minimise the efforts required to perform the risk management activities and provide accurate information about the risk level based on the CTI context to implement the proper controls. It is also designed to enable organisations to use threat intelligence report to predict a certain risk level. Another critical aspect of i-CSRMT is that it is formed based on the principles of renowned industry-standard. Also, the tool can be simultaneously accessed and used by multiple users and different organisations and allows managing multiple different projects simultaneously. The tool also provides a separate web interface for the different actors within the organisation (application administrators), giving them access to the user and project management.

4.1. i-CSMT Architecture

i-CSRMT is a three-tier web-based system comprising of a presentation layer, application layer and data layer. From a logical point of view, three-tier architecture is used to improve the tool's modularity and mainly allow for easy extension of features. Using client-server architecture, users can use any web browser to connect to the many services supported by the tool, such as initiating audit assessments. On the server-side, the webserver receives requests from the client, handles the request and generates an appropriate response to the client. The three-tier architecture role of three-tier architecture is explained as:

4.1.1. Presentation Layer

This layer manages the communication with the Web browser, renders the application Web pages, and controls the user access. The layer consists of a single module that represents the user interface. It is implemented using the Java Play Framework (Leroux and Kaper, 2014) and follows the Model-View-Controller (Enache, 2015) architectural pattern. The Views represent the contents of the application Web pages and are built using HTML, PHP, CSS and JavaScript. Some Views contain only parts of the user interface; either embedded into the Web pages or loaded dynamically using AJAX. The server's communication is managed using Controllers, which handle the HTTP requests and return responses Views.

4.1.2. Application Layer

The application layer is built using PHP, and it plays the role of linking together all the three layers by technically processing the various inputs and selections received at the presentation layer and interacting with the vast database in the third layer. Also, the layer houses the web server, scripting language and the scripting language engine of the tool. The Web server enables the processing of HTTP requests for initiating the activity process. The application layer provides the technical deal with dynamic content and streamlines the database's faster access to extract results.

4.1.3. Database Layer

The database provides a centralized place where data captured in the tool are stored, manipulated, and accessed. The layer comprises database management systems (DBMS) and the database, which is built using MySQL. The database layer's rationale is to centralize all data storage, store and retrieve the application data. In other words, it contains the methods for accessing the underlying database data. Fundamentally, the database layer is responsible for storing numerous types of data the tool will take as an input, generate as output and other external services that the tool may use. The database is accessible to the system administrators and employees as shown in Figure 7.

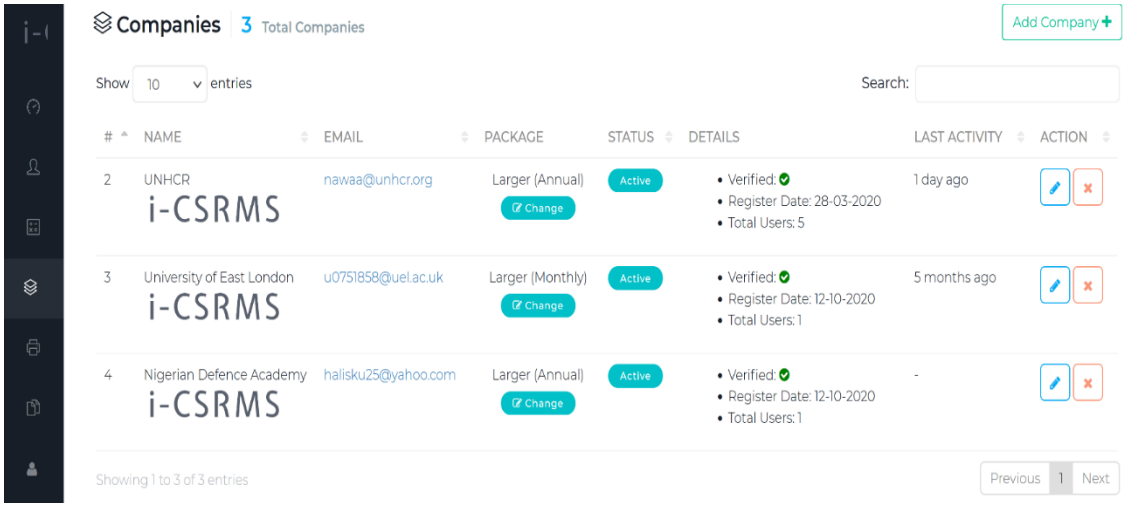

Figure 7: Admin dashboard for managing companies

4.2. i-CSRMT Features

This section provides an overview of i-CSRMT features as shown in Figure 8. These features follow the main activities of the i-CSRM process mainly asset identification and criticality, threat modelling, risk assessment and reporting. Therefore, the output of each activity is considered as a feature of i-CSRM. The features support interaction among multiple users and allows the users to split their work and delegate responsibilities. The application administrators can define dynamic user roles and assign them to the users to restrict their access to specific application parts. The primary purpose is to provide a general understanding of how the tool is decomposed and how the individual components work together to provide the desired functionalities. In general, the tool focuses on minimising the efforts required to perform the risk management activities and provide accurate information about the risks. The tool's main features include a main dashboard consisting of essential functions that can be performed. Each functionality contains essential components of a risk management process. The main features include Actor identification, Asset criticality, Threat modelling, Risk assessment, Control effectiveness, and Report dashboard.

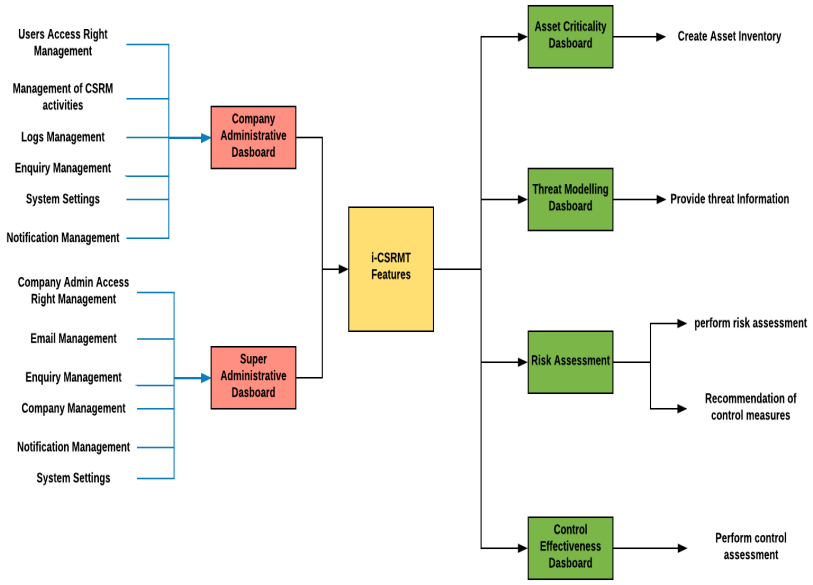

Figure 8: Features and components of i-CSRMT

5. Evaluation of i-CSRM

The primary purpose of the evaluation is to determine the applicability of the i-CSRM Framework in a real-world scenario. The evaluation comprises a set of associated methodologies and techniques with a distinct purpose of providing the means to establish the value, quality and relevance of research, and in some cases, provides feedback necessary for improvement (Boudreau, Gefen and Straub, 2001). There are many empirical evaluation methods and techniques that could be adopted, such as action research, experimental methods and descriptive methods.

5.1. Case study: Implementation of i-CSRM Framework

This presents the implementation of the i-CSRM framework process as well as i-CSRMT using the case-study. By following the i-CSRM process from start to end over some time, we systematically applied all the activities and steps within the i-CSRM process using i-CSRMT and the opportunity to collect feedback towards evaluating its validity. Therefore, a detailed description of the case study is provided by first presenting background information and implementing the existing system.

5.1.1. Study context

DisCos power holding company in Nigeria distributes electricity (Kemabonta and Kabalan, 2018) across the country, which serves at least 30,000 customers within a geographical area, with several branches and employees located in different states in Nigeria. The company is structured based on functional divisions, which include administration, support and IT. The company's first services are to provide last-mile services in the electricity supply value chain, transforming or stepping down electricity from the high voltage at the transmission level to lower voltage depending on the customer's category. They are responsible for the marketing and sale of electricity to customers, providing a tax to the government, collecting bills, handling electronic payments, exchanging information and providing customer care functions in its geographical area. In improving the continuity of service, timely recognition of faults, continuous monitoring and protection of the power systems, the company recently implemented a supervisory control system in all of its branches for sustainable service delivery.

5.1.2. The Workflow

The power distribution happens through a power distribution substation that comprises other components such as circuit transformers, breakers, and a bus bar. The bus bar splits and distributes power to distribution lines for reaching outto customers. The substation's whole distribution process and components are managed by a cyber-physical control system, consisting of a Supervisory Control and Data Acquisition (SCADA) system. In other words, the SCADA system monitors the entire power control system in real-time by performing automatic monitoring and controlling of

various equipment within the distribution lines. It also maintains the desired operation conditions, interrupts and restores power service during fault conditions. SCADA system also checks the status of various equipment continuously and sends control signals to the remote control unit accordingly. Further, it also performs operations such as bus voltage control, load balancing, circulating current control, overload control, and transformer fault protection

5.1.3. Recent Cyber Incident

DisCos is an official body with branches geographically split (Onochie, Egware and Eyakwanor, 2015); each has its workstations networked to allow personnel to perform their tasks. All branches deployed a new SCADA system to improve power reliability, cybersecurity, and resilience to disruption. They use a SCADA consisting of 5 generic machine types connected to a local Ethernet LAN to support their services. In a recent event, an employee monitoring the SCADA system in one of the branches received a carefully crafted spear-phishing e-mail message from a highly skilled anonymous organisation that contained a malicious Microsoft Word attachment and disguised as a medical report of his sick son. The employee clicked and opened the document, and malware was discovered to have spread across the network, operating systems, and targeting the SCADA system, which led to the unstable power system operation in the branch. The anonymous organisation gathered hashed credentials over a server message block (SMB) to identify information by downloading the word document. The anonymous organisation accessed workstations and servers on the corporate network that contained data output from control systems, accessed files about the SCADA systems, leaked network credentials, organisational design and control system information to a command-and-control server outside Discos organisation, and accessed e-mail accounts using outlook web access (OWA).

The anonymous organisation used a virtual private network (VPN) to maintain access to networks even with network proxies, gateways and firewalls. After the employee visited one of the compromised servers, a backdoor was installed on the machine, providing the anonymous organisation with remote access to the environment (networks, systems, databases). The anonymous organisation having available resources, disabled the host-based firewalls, obtained a foothold and the exploration activity primarily cantered on identifying the central host computer server with the highest volume of personally identifiable information (PII) script folder and file names from hosts. The anonymous organisation gained access to the database host computer server by leveraging its active directory information to identify database administrators and their computers. Passwords were cracked using password-cracking techniques, allowing the anonymous organisation to gain full access to those systems. This caused a loss of data and operational disruption as a result of network and computer security failure. This particular incident has resulted in an electrical power blackout that remained for up to 2 weeks, affecting around 30,000 customers and their businesses. As a result, DisCos has decided to use i-CSRM framework to assess future impact and control measures for similar incidents in the other branches. A brief description of a scenario allows us to exemplify how the DisCos could benefit from our proposed Framework.

5.2. Implementation of i-CSRM for the Study Context

In the context of DisCos we had the opportunity to determine i-CSRM relevance to a real-life context. As part of managing the entire evaluation process, the company assigned a team of professional stakeholders to guide the entire evaluation process and ensure necessary support to ensure evaluation is achieved in an ideal manner. This section provides a detailed description how the framework is applied to the case study. Before starting the activities, a meeting was organized where the evaluation plan's overall setting was decided, a project team was developed, and a first step was taken towards starting the activities. The project team comprised of representatives from senior management, the IT department and other stakeholders within the company.

5.2.1. Activity 1: Organisational Context

We started the activities defined in the proposed i-CSRM Framework with the organisational context, which allowed us to identify the organisation's key objectives and understand the key actors and their roles within the organisation. This enabled us to interact more effectively with key actors to gather information and implement the proposed i-CSRM Framework.

- **Step 1: Identification of Actors and their roles:** During the initial meeting and interaction with the implementation team, we were able to identify the key actors that support and influence the project and the different roles they play within the organisation. This enabled us to interact more effectively with the key actors to gather information and implement the proposed i-CSRM Framework. Table 8 provides a list of different actors and their roles.

Table 8: List of Actors and their Roles

Type	Actors	Role
Internal	Senior Management representatives	Comprises high ranking personnel of the company whose responsibility is to coordinate, plan, oversee and direct the overall project.
	IT Managers	In charge of the company's technology strategy and responsible for coordinating and leading the company's IT experts/IT department in implementing the Framework's process.
	System Analyst	Responsible for coordinating the development of systems, asset requirements, and control measures for ensuring the security of all assets.
	System Administrator	Responsible for the technical oversight of the entire content management system. He was also charged with installing, supporting and maintaining servers, responding to service outages and other problems.
	Security Analyst	Responsible for identifying cyber threats and establishing plans and controls to protect assets. Also responsible for performing vulnerability testing, risk analysis and security assessment activities
	Risk Manager	Risk Manager communicates risk policies and processes for an organisation. They ensure controls are operating effectively, provide hands-on development of risk models involving market, credit and operational risk and provide research and analytical support.
	Registered Users	Registered users who have permission to use the system

5.2.2. Activity 2: Asset Identification and Criticality

The project team embarked upon initial knowledge extraction through senior management support, and active involvements, were initial information that facilitated the identification of the organisation's critical assets. This enabled us to understand how things are done in the organisation regarding its activities, followed by identifying the security goals that are part of an essential component of the organisation's assets and identifying the most critical assets.

- **Step 1: Asset Profile:** The IT manager was involved in explaining and documenting the system and its components, which provided the basis to identify the organisation's critical assets and their security, needs to create a consistent asset profile. The IT manager also presented a comprehensive overview of the organisation's assets which are the target of analysis, from where we observed that the system comprises many different components as shown in Table 9.

Table 9: Assets Identification

Asset Category	Sub-Asset category
Software assets	Microsoft office, Master boot record/files, Mail server, Service Manager, Windows/Android operating systems, UPS remote management interface, Computer security protection, Virtual machines, User identity access management
Hardware assets	Computer systems, Remote login systems, Windows machines, Keystroke Logger, Hard drives
Data assets	Skype messages, Internal domain names, Network/system information, Sensitive information, Admin credentials
SCADA systems	Industrial control systems (ICS), HMI computers, Remote terminal unit (RTU), Substations, ICS providers, SCADA database software, Programmable logic controllers (PLC), Firmware, Substations Ethernet devices, SCADA database software, Workstation, ICS software application and windows
Information and Communication Networks	Company's computer network, Virtual private network, Router/modem/ switches/proxy/gateways, Firewall UPS server, Network Internal server, Public-facing services, Command and control servers, Website, Remote access services, Operational network, Remote access services, URL, Bluetooth

- **Step 2: Identify Asset Security Goals:** After the asset inventory had been agreed and completed by the team, the next step was to identify each asset's goals. The security analyst conducted a high-level brainstorming exercise with the help of other team members to identify the most critical security goals for the assets identified in the previous step. At first, some representatives of DisCos emphasised that they are particularly worried about the privacy of data held by the CPS and availability of the services. However, the security analyst explained that the team had reviewed the information collected during the previous step and examined every functional requirement for the system through less important security goals such as confidentiality, integrity, and availability.
- **Step 3: Determine Asset Criticality:** Having identified the system's assets and its related security goals, the project team embarked on the next step of determining asset criticality on the identified assets in the previous step. The criticality level is determined and assessed in greater detail as part of the asset identification and criticality activity. An assessor team consisting of the security analyst and other experts prioritised assets in terms of the security goals by applying a novel asset criticality system using fuzzy logic proposed in i-CSRM process so that the most critical assets can be connected with top priorities. This step was conducted as a separate brainstorming exercise, and the primary goal was to determine the criticality of the assets formally approved by all project team members. The FACS allows experts to express their differences in the inference process with less bias and higher reliability. Therefore, asset criticality was determined using the method proposed in the process and each asset is assigned a level of criticality using fuzzy inputs and the crisp rating values. The result is shown in Table 10 and Figure 9.

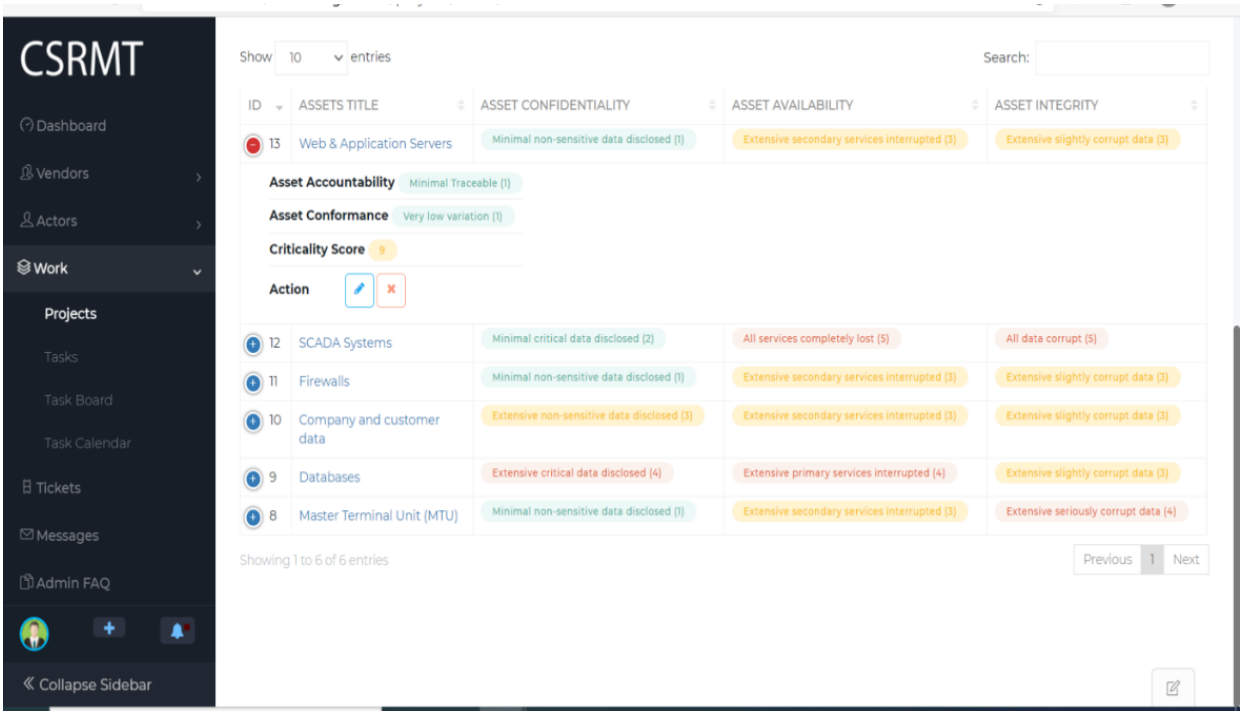

Figure 9: Asset criticality Result

Table 10: Asset criticality results

Asset Name	Asset Description	Asset Goals					Fuzzy output	Asset Criticality Level
		Fuzzy input						
		C	A	I	CON	ACC		
Routers, firewalls, intrusion detection systems	Monitor, analyse and filter any harmful signs, while being connected to the corporate network.	1	3	4	4	1	2.5	Medium Critical
Databases	Stores sensitive information about its customers,	4	4	3	4	5	4	Highly Critical

	personnel, marketing, landlords, tenants, transactions, assets, finances, and other information about the company's business process.							
Company and customer data	Represent sensitive and private information about employees, finances, assets.	3	3	3	4	4	3.5	Medium Critical
Web & Application Servers	Provides processes and delivers web contents such as images and assets information to employees and customers. The application server provides the platform for hosting frontend applications used by the company	1	3	3	1	1	2	Low Critical
SCADA Systems	Provides the user interface that allows employees and customers to visualise, access, and patronise the company's services.	2	5	5	1	4	3	Medium Critical

5.2.3. Activity 3: Threat Modelling

This activity aimed to identify the possible threats and vulnerabilities for the Discos. The activity was organized as a workshop, drawn from actors with expertise in risk management. The actors involved in this activity included the security analyst and a member of senior management. Also, various methodologies and standards were employed at different steps of performing the threat modelling activity. All participating actors were briefed about the parts of the standard/methodologies used and its benefit.

- **Step 1: Determine the Vulnerability profile:** The first step focused on identification of vulnerabilities and weaknesses by examining the attack surface and the relevant threat models. The analysis team moved on to create a vulnerability profile that contains the vulnerabilities that are exploited and affect assets. To direct this process, the project's team members, a security analyst and system administrator were brought together to conduct an informed brainstorming session to identify a detailed list of potential vulnerabilities. Secondly, a list of vulnerabilities compiled by CWE and CAPEC was presented to the team to understand by providing a standardized list of software weaknesses and the methods to exploit those weaknesses such that two or more people know they are talking about the same thing. By identifying the weak points, the security analyst documents the meeting's result by filling a vulnerability profile for the study context, which affected critical assets and caused a threat that led to risk.

- **Step 2: Determine Threat profile:** Having completed the asset inventory and identified vulnerabilities, the analysis team created a threat profile that identified the threats that can potentially affect the assets and compromise sensitive information. To direct this process, the project's team members, a security analyst and system administrator were brought together to conduct an informed brainstorming session to identify a detailed list of threats, threat actor factors, TTP and IOC. A list of security threats compiled by CAPEC and WASC was presented to the team. Firstly, the team started with identifying a combined list of 10 security threats that they perceived to be important to the organisation's assets. After a brief reconsideration, the list was updated with three additional threats. Secondly, the adoption of these two threat classification models proved helpful and straightforward in identifying, categorizing and determining the impact of potential threats, and it led to the participants having a better understanding of threat elements. With the adoption of CTI, a better understanding of threat actors, attack patterns, and TTP use is understood by the team. In this regard, the team considered all potential threats to document the threats, vulnerabilities, IOC and TTP associated with the assets; a template that shows several threat attributes is used. Figure 10 shows the threat modelling displays the threat actor factors, indicators of compromise (IOC), TTP, related attack patterns, execution flow and possible vulnerabilities.

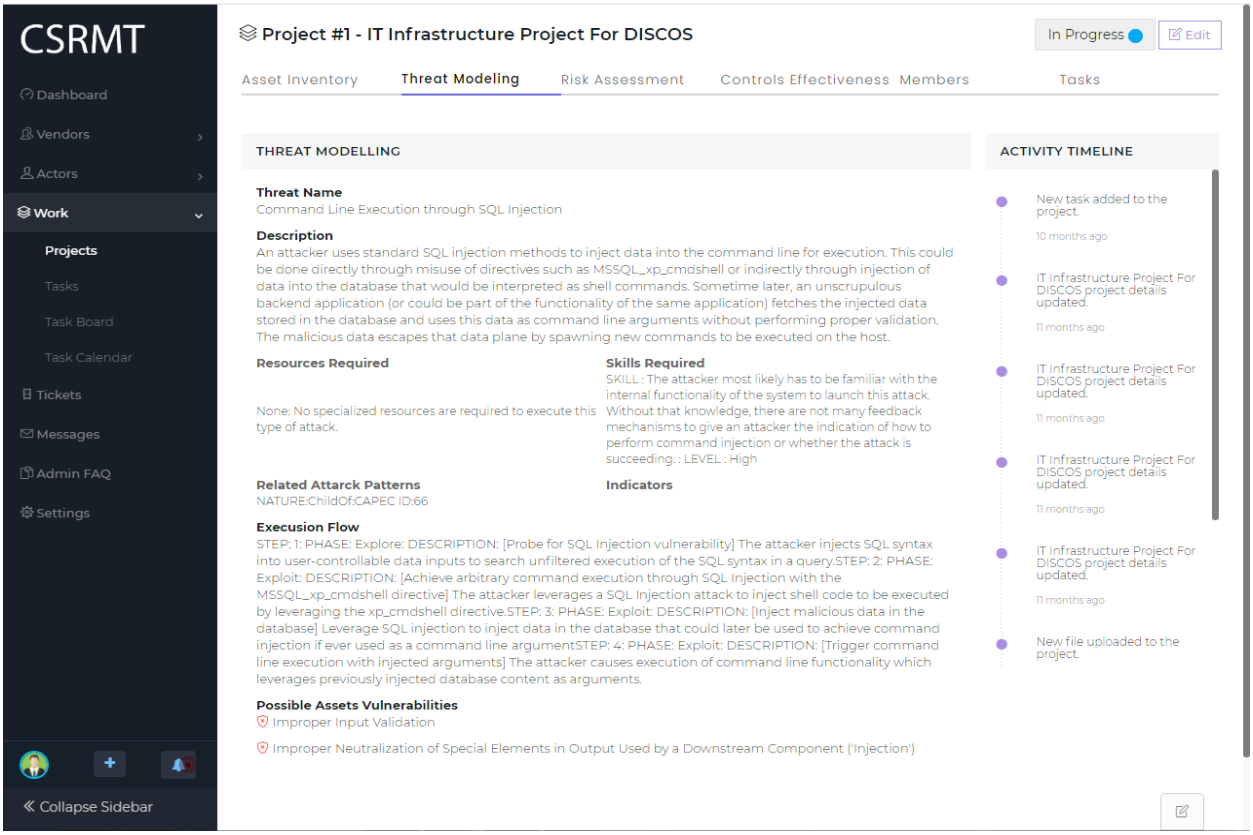

Figure 10: Threat and vulnerability profile

5.2.4. Activity 4: Risk Assessment

The next activity involved a risk management process whose goal was to identify as many potential threats, vulnerabilities and risks as possible. The activity was organized as a workshop drawn from stakeholders with expertise in risk management. The stakeholders involved in this activity include the security analyst, information security officer, and senior management member. Also, various methodologies, machine learning techniques and standards were employed at different steps of performing risk management. All participating actors were briefed about the parts of the standard/methodologies used and its benefit.

- **Step 1: Predict Risk Types:** In this step, a workshop was organised for the identification of risks types. The participants were presented with multiple risk types, usually associated with critical infrastructure and assets of all kinds. The risk sources are provided by industry bodies and are updated regularly, which means that they provide up-to-date information about the most pressing security issues in information systems and web applications. In particular, a list of risks provided by the VCDB dataset was presented in the workshop, and the participants were challenged to select those they perceive to be relevant threats previously identified. We have used ten output categories of risks, and the value range for the features is from (R1 = Crimeware, R2 = Cyber espionage, R3 = Denial of service, R4 = Everything else, R5 = lost and stolen assets, R6 = miscellaneous errors, R7 = payment card skimmers, R8 = point of sale, R9 = privilege misuse and R10 = web applications) with possible classes. This is a multi-class problem, and we have the following risk types as output features. A list of risks is therefore identified.

➢ **Phase 1: Prediction Result:** Table 15 presents the six classifiers' accuracy performance details in predicting the different risk types based on the given CSRM features (Assets, Controls, Threat Actor and TTP). Based on the Asset features, LR, DT and NB-Multi achieved 95%, 93% and 92% respectively for predicting risk type "Lost and Stolen Assets", "Everything Else", "Crimeware", "Cyber Espionage" and "Denial of Service". They failed to identify risk types "Point of Sale" and "Web Application". RF, KNN and NB achieved 87%, 86% and 71% respectively for predicting risk type "Crimeware", "Cyber Espionage", and "Lost and Stolen Assets". NN failed to predict any risk type and achieved 4%. Based on the TTP features, KNN, LR, NB-Multi, and DT achieved an accuracy of 80% for predicting risk type "Denial of Service", "Cyber Espionage" and "Everything Else". RF

achieved an accuracy of 72% for predicting risk type "cyber espionage" and "Everything Else" NN failed to predict any risk type and achieved 4%.

Based on the Threat Actor features, LR, NB-Multi and RF achieved 79% accuracy for predicting risk type "Everything Else", "Cyber Espionage" "Privilege Misuse", and "Crimeware". KNN could predict risk type "Everything Else", "Cyber Espionage", and "Privilege Misuse" while DT could predict risk type "Everything Else", "Cyber Espionage", and "Crimeware" both classifiers with 76% accuracy. The NB achieved 63% accuracy for predicting risk types "" Cyber Espionage" and "Privileged Misuse". NN achieved 3% accuracy and failed to predict any risk type. Lastly, based on the control features, KNN achieved the highest accuracy of 40% in predicting risk type "Everything Else". LR, DT, NB-Multi and RF achieved 39% for predicting risk type "Everything Else". NB and NN achieved an accuracy of 5% and 3% respectively. Both classifiers failed to predict any risk type. Asset and TTP features performed well on all the different classifiers except NN. Comparing the performance of all the features shows that NB failed to perform risk type prediction based on control features and NN achieved very low risk type prediction based on all the features. Therefore, for the risk types "Everything Else", "Privilege Misuse", "Denial of Service" and "Cyber Espionage" all the input features achieved high prediction. Table 11 shows that Asset and TTP are the best features to predict risk types presented in this work and associated graphical chart in Figure 11.

Table 11: Performance of the features on each of the classifiers for predicting risk types

Accuracy	**Risk Type Prediction Features**			
	Asset	**TTP**	**Threat Actor**	**Control**
LR	95%	80%	79%	39%
DT	93%	80%	76%	39%
NB-Multi	92%	80%	79%	39%
RF	87%	72%	79%	39%
KNN	86%	80%	76%	40%
NB	71%	56%	63%	5%
NN	4%	4%	3%	3%

- **Phase 2: Prediction Accuracy:** After predicting the possible risk types by feeding the CSRM features from VCDB dataset into our classifiers, the next step was to interpret the different classifiers' accuracy for various types of input features. Therefore, the predictive accuracy percentage of six different machine learning classifiers based on CSRM features was presented. However, each feature performed differently within classifiers. The best overall predictive accuracy including all input features was recorded with Decision Tree (DT) algorithm which is (92.92%) on Asset features, Controls (79.26%), TTP (62.73%), Threat Actor (61.32%), and Full features was (39.12%). The second best algorithm is NB Multi which gave us (91.90%) on asset features, control features (78.88%), threat actor (61.33%), TTP (59.54%) and full features gave us (39.05%). The third best algorithm is RF, it performed well on Asset features with (87.36%), control (78.75%), TTP (62.03%), Threat Actor (61.01%) and full features (38.93%). The fourth best algorithm is KNN, it performed well on almost all the input features, Asset features (85.77%), Controls (67.96%), TTP (58.07%), Threat Actor (56.80%) and the full features produced the least accuracy with (29.99%). The fifth best algorithm is the NB algorithm that performed well on the asset features with (71.03%), controls (55.90%), Threat Actor (19.85%), TTP (18.38%) and full features with (05.42%). The sixth algorithm which is NN didn't perform well on all the features, control features is (04.02%), Asset features is (03.51%), Full feature is (03.32%), TTP (03.13%) and threat actor (03.06%). This shows that the Asset features performed well with DT (92.92%), NB Mult (91.90%), RF (87.36%), KNN (85.77%) and NB (71.03%). NN did not perform well with (03.51%). The control features also performed well with DT (79.25%), NB Multi (78.88%), RF (78.74%) and KNN (67.96%). On the other hand, Neural Networks (NN) and Naïve Bayes (NB) did not make satisfactory prediction accuracy on all the features. It can be noted that the most prominent features to detect risk types are Assets and control features. The result clearly shows that DT outperformed other classifiers giving the highest satisfactory accuracy for the VCDB dataset for risk type prediction.

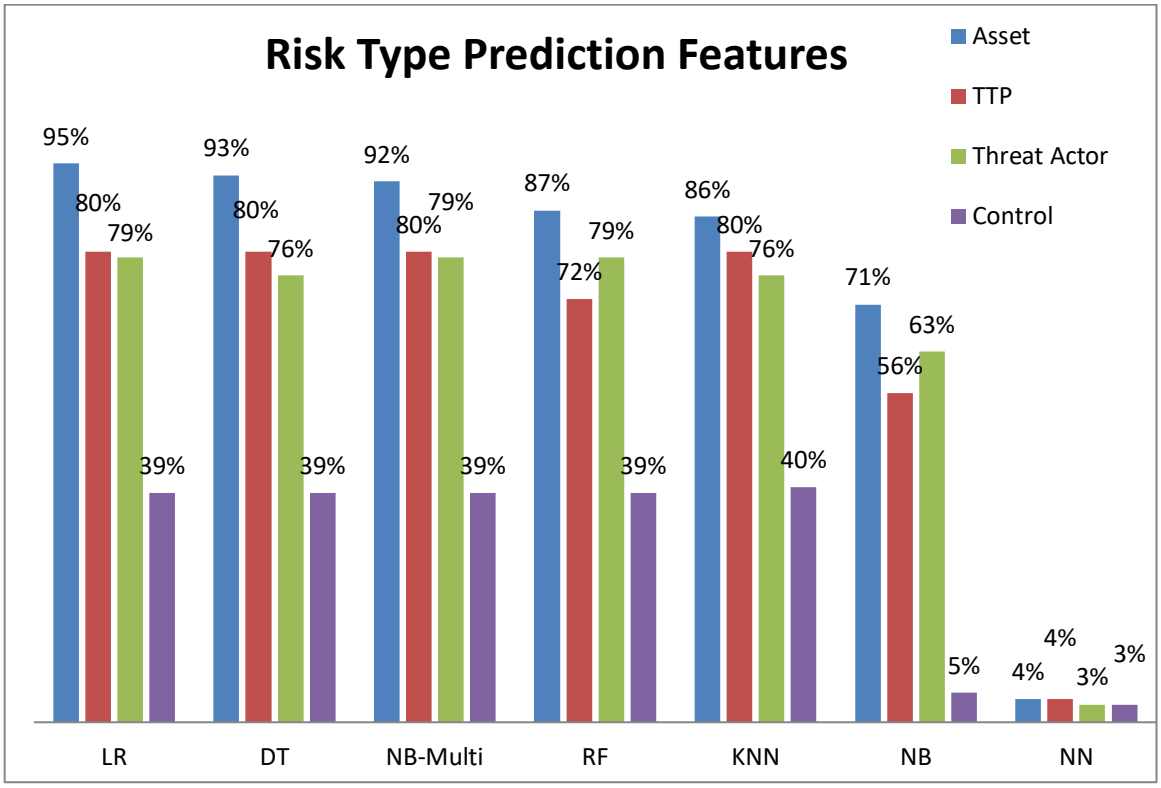

Figure 11: Performance of the features on each of the classifiers for predicting risk types

- **Phase 3: Results of the different classifier for the input features:** Figure 12 shows the accuracy results of different classifiers for the various kinds of input features. The most prominent features to detect the risk type are Assets and Controls, where accuracy is above 70%. From left to right (top to bottom), the X-axis denotes different classifiers and Y-axis denotes the corresponding accuracy for a given feature set. It can be seen from the descriptive result shown in figure 16 based on the asset features KNN, NB Multi, RF and DT have produced the most accurate predictions by giving the accuracy value of above 70% compared to NB and NN classifiers.

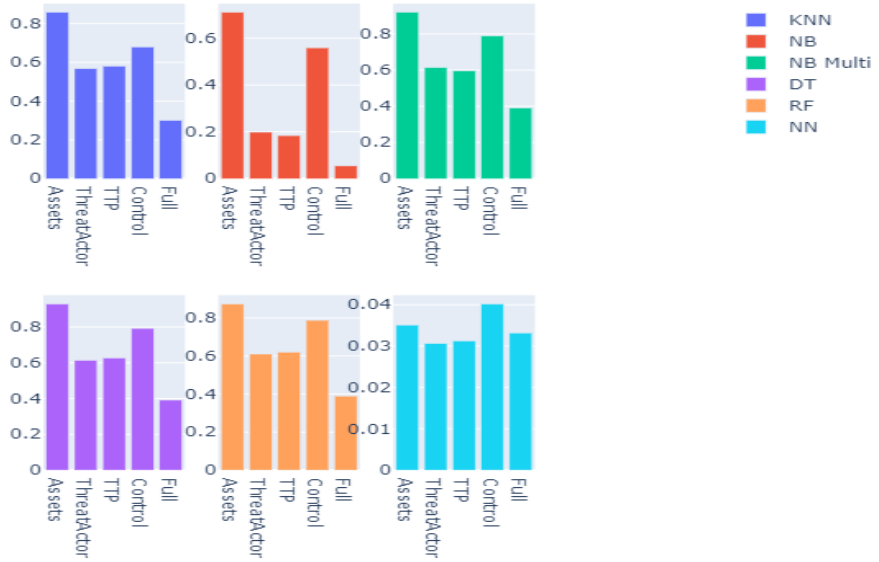

Figure 12.The accuracy of different classifiers for various types of input binary features

- **Phase 4: Results of Confusion Matrix:** This section describes the classifiers' performance on the test data for which the true values are known. This allows for the visualisation of the performance of an algorithm. In this case, the best overall predictive accuracy was recorded with KNN which produced better results than other classifiers as shown in Table 12.

Table 12: Performance measure for KNN classifier

Output	Precision	Recall	F1-Score
1	1.000	0.525	0.689
2	0.700	0.687	0.693
3	0.729	0.501	0.694
4	0.766	0.578	0.659
5	0.735	0.561	0.636
6	0.614	0.340	0.438
7	0.820	0.432	0.566
8	0.815	0.373	0.512
9	0.950	0.710	0.813
10	0.264	0.711	0.385
Accuracy	**0.576**	**0.576**	**0.576**

- **Step 2: Determine Risk Level:** After identifying the various IOC, TTP, vulnerabilities, threats, and predicted the risk types using dataset, we identified and assessed the risks by estimating the assets' likelihood and impact. The Web pages allow the organisation to adapt various aspects associated with risks and their relations. This includes risk types, risk impact, risk likelihood and control measures. As stated previously the risk calculation considers the risk likelihood and impact. The web page displays the results of the risk calculation. Each risk event is evaluated and presented separately with the elements used in the calculation and the calculated risk value. Figure 21 presents the calculated risk value which represents how dangerous the risk is to the organisation.

5.2.5. Activity 5: Risk Controls

The final activity involved identifying and evaluating existing controls using four seeps.

- **Step 1: Identification of Existing Control Types:** We first identified DisCos existing controls to ensure that the controls are working correctly. The organisation detected the controls; some are shown in table 13 to address the identified risks. The outcome determines the security control budget for the organisation, and decisions are optimised.

Table 13: Existing Control Types

Control type	Attack Techniques	Control description
Preventive	Brute Force	After a certain number of a failed login attempt to prevent passwords from being guessed, set account lockout policies.
	Disabling security tools	The proper process, registry, and file permission should be in place to prevent the anonymous organisation from disabling or interfering with the Disco's security services.
Detective	Account discovery	Identify unnecessary system utilities or potentially malicious software that may be used to acquire information or data about system and domain accounts, and block them by using whitelisting tool or software restriction policies where appropriate.
	System Network Configuration Discovery	
	File and Directory Discovery	
	Data from the local system	
	Spear-phishing attachment	Network intrusion prevention systems should be put in place to scan and remove malicious e-mail attachments.

Corrective	External Remote Services	Limit access to remote services through centrally managed VPNs, and other managed remote access internal systems through network proxies, gateways and firewalls.
		Use strong two-factor or multi-factor authentication for remote service accounts to mitigate the anonymousorganisation's ability to leverage stolen credentials.
	Credential Dumping	Ensure that administrator accounts have complex, unique passwords across all systems on the network.
	E-mail Collection	Use of two-factor authentication for public-facing webmail servers is recommended as a best practice to minimise the use of usernames and passwords to the anonymousorganisation.
	Forced Authentication	Use strong passwords to increase the difficulty of credential hashes from being cracked if they are obtained.
	User Execution	Training is required for the Disco employees to raise awareness on raising suspicion for potentially malicious events.
	Spear-phishing attachment	Antivirus can also be used as it automatically isolates suspicious files

Step 2: Evaluating the Effectiveness of Existing Controls: It was proposed that control effectiveness should be specified according to five fundamental categories namely: relevance of the control, strength of the control, coverage of the control, integration of the control and traceability of the control. The participants became involved and based on their expert opinion; effectiveness of the existing controls is specified in Figure 13.

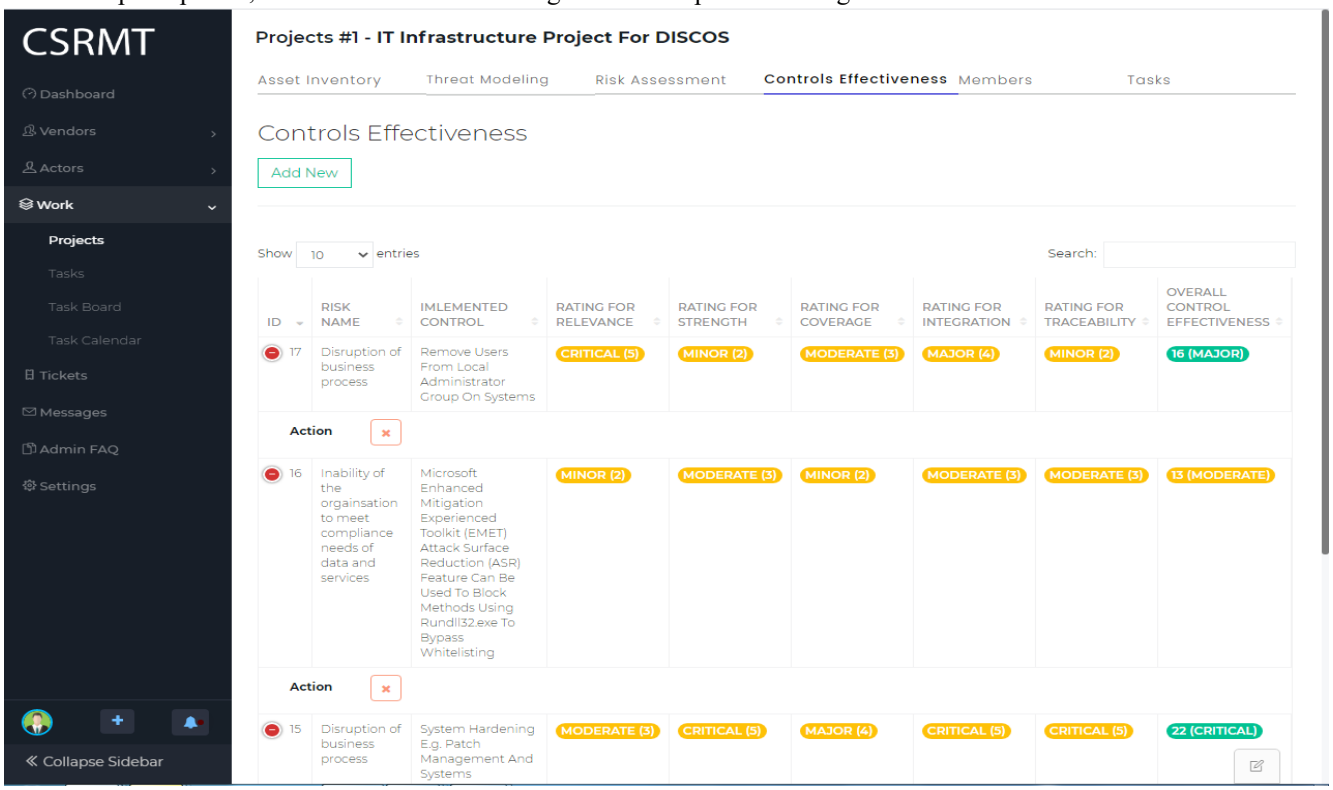

Figure 13: Control Effectivenes Result

- **Step 3 and 4: Implement Control Measures to Determine New Risk Status and risk register:** We first identified existing controls, to ensure that the controls are working correctly. The Web page allows the organisation to define a list of available controls. The user can select the control measure using the control rating: None, partial and full as shown in Figure 14 to address the identified risks. Finally, the step 4 creates the risk register to record all identified risks and controls.

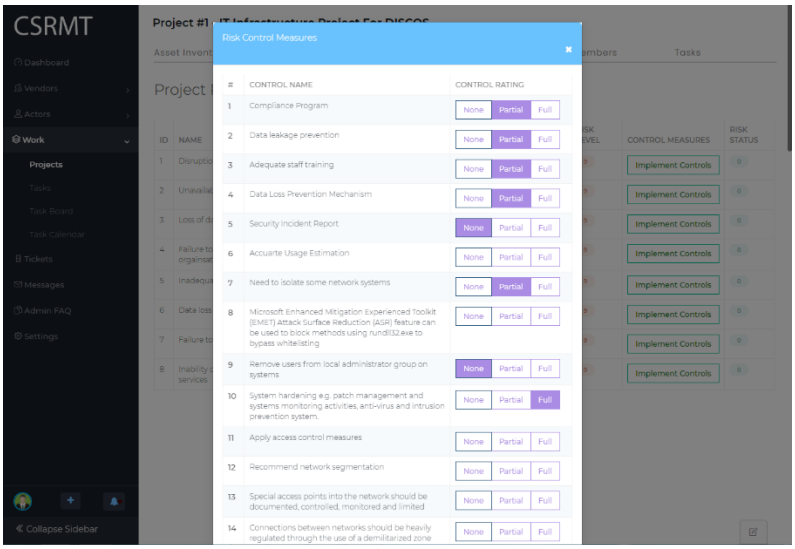

Figure 14: Control measure implementation

6. Discussion

The users of the studied context observed that the i-CSRM framework is very effectives in terms of performing the risk management activities. The approach provides a detailed about the assets and traces the vulnerabilities and threats based on the identified assets which makes it easy to understand the potential risks. It provides a comprehensive and holistic analysis of the risk taken into account the asset, threats and vulnerabilities so that suitable control actions can be evaluated. The integration of existing standards and ML models certainly provides a wider adaption of i-CSRM.

The i-CSRM framework is a practical approach to assess and manage cyber security risk, specifically the activities under the process are operational. The integrated risk management framework lays out the basics for defining critical assets, evaluating their weaknesses and risks for determining the appropriate controls. This approach has made stakeholders aware of the possible threats and predicate risk types that could impact their critical services and business operations, therefore taking the necessary actions to control threats and risk events from occurring. Furthermore, gaining a better view of Disco's existing risk control practices, evaluating them, and suggesting changes raised the overall visibility.

The outcomes of our case study were compared to those of other research reported in the literature. Compared to other works in the literature, the applied cyber-security risk management framework is a systematic solution. A previous author (Abouzakhar, 2013)) identified a range of security risks and events through different critical infrastructure domains. The work incorporates specific mitigation steps for critical infrastructures, such as vulnerability assessments and penetration testing approaches; however, this paper's emphasis was not just on vulnerability evaluation but also on how danger can be measured, mitigated, and managed. Because of the interdependency between properties, asset detection and cascading vulnerabilities were not taken into consideration. Authors of a previous paper (Hokstad, Utne and Vatn, 2012) suggested a risk and threat analysis approach for critical infrastructure that focuses on severe incidents while emphasising critical infrastructure business dependencies. However, no systematic study has been performed to define essential assets and weaknesses specific to such assets or identify the specific chains of events (cascading vulnerabilities). The authors of a previous paper (Cherdantseva et al., 2016) stressed the need for a holistic risk management system that includes all phases of the risk management process; our work reflects this to enhance the CPS's cyber-security. In comparison to the writers of a previous paper (Sridhar, Hahn and Govindarasu, 2012a), who suggested a layered method for assessing risks based on protection, our work evaluated risks cyber-attacks databases, as well as risk level and proper controls. Although the writers of a previous paper (Cardenas et al., 2009)explored a

framework for avoiding, detecting, and restoring attacks for protecting CPS, our study presented a mechanism for recognising sensitive properties, evaluating cascading weaknesses, creating cyber-attack scenarios, determining the effect of an attack happening, and providing preventive controls to better protect the CPS.

None of these works provides a structured risk assessment mechanism that considers the asset criticality before evaluating vulnerabilities. Our research identifies and contrasts current risk reduction solutions for CPS in critical infrastructure, allowing critical infrastructure organisations to do an in-depth cyber-security study on CPS. There are certain similarities between our research and other works in terms of risk assessment and reduction. In a previous paper (Bialas, 2016b), the authors discussed danger by addressing interdependencies and risk monitoring. These results are fully or partially close to what we observed in our study. However, specific threats found (Gai et al., 2016), such as energy waste and deploying mobile cloud computing problems, are not strictly comparable to our studied background. Lack of contingency planning, emergency response, reporting systems, robust risk assessment, and the use of machine learning tools to assess the risk level and analyse the efficacy of current controls are some of the specific risk factors not listed in other reports. We urged consumers and operators not to shirk their IT obligations, since the threats of essential infrastructure vary depending on the organization's background. It is also important to raise knowledge of cyber security threats through the whole enterprise and the supply chain climate, as well as to continue to improve and use innovative cyber security capabilities to exercise risk assessment and risk evolution.

7. Conclusion

Risk management is a continuous process for maintaining the effective functioning of critical assets for any organisational context. In particular, Critical Infrastructures need resilience for the service delivery and risk management is an essential component to achieve this. The threat landscape is constantly evolving with new techniques and more sophisticated organised attacks. Therefore, it is necessary for the risk management activities to consider the threat context to assess and manage the risks. This research proposes the Integrated Cyber Security Risk Management Framework (i-CSRM) that adopts various existing standards and cyber threat intelligence data for risk management. i-CSRM also includes Machine Learning (ML) models to predicate the risk types so that organisations can undertake the necessary proactive measures to tackle the risks. The framework also includes a tool support to automate some of the risk management activities. Finally, i-CSRM is applied in a CI-based industrial context and the results of applying the framework are very promising. Specifically the studied context was able to identify and assess risks using i-CSRM and determine the right level of control for the overall business continuity. The participants' observation is that i-CSRM is a practical approach for the risk management, and integration of CTI make the risk management activities more effective. We believe that the proposed i-CSRM framework, its process and supporting tool will significantly impact the cybersecurity domain and state of the art in general. The i-CSRM framework focuses only on the supervised learning method, which requires labelled dataset. As a part of our future research, we would like to deploy the i-CSRM in different CI context and implement different data sets for the risk type predication. Additionally, it is necessary to develop a checklist to make the process easy to use for risk assessment and management.

Acknowledgements This work was partially supported by the AI4HEALTHSEC EU project, funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883273 and Cybersane project with grant agreement No 833683

Declaration of Conflict of Interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abu, M. S. et al. (2018) 'Cyber threat intelligence–issue and challenges', Indonesian Journal of Electrical Engineering and Computer Science, 10(1), pp. 371–379.
Baldoni, R. (2014) Critical infrastructure protection: threats, attacks, and counter-measures. Technical Report. Available online: http://www. dis. uniroma1. it/~ tenace ….

Barnum, S. (2008) ‘Common attack pattern enumeration and classification (capec) schema description’, Cigital Inc, http://capec. mitre. org/documents/documentation/CAPEC_Schema_Descr iption_v1, 3.
Bialas, A. (2016) ‘Risk Management in Critical Infrastructure—Foundation for Its Sustainable Work’, Sustainability. doi: 10.3390/su8030240.
Boudreau, M.-C., Gefen, D. and Straub, D. W. (2001) ‘Validation in information systems research: A state-of-the-art assessment’, MIS quarterly, pp. 1–16.
Castro, J., Kolp, M. and Mylopoulos, J. (2002) ‘Towards requirements-driven information systems engineering: the Tropos project’, Information systems, 27(6), pp. 365–389.
Chen, P. P.-S. (1976) ‘The entity-relationship model—toward a unified view of data’, ACM transactions on database systems (TODS), 1(1), pp. 9–36.
Cherdantseva, Y. et al. (2016) ‘A review of cyber security risk assessment methods for SCADA systems’, Computers & security, 56, pp. 1–27.
Consortium, W. A. S. (2009) ‘Web application security consortium threat classification’.
Conti, M., Dargahi, T. and Dehghantanha, A. (2018) ‘Cyber threat intelligence: challenges and opportunities’, in Cyber Threat Intelligence. Springer, pp. 1–6.
Cord, O. (2001) Genetic fuzzy systems: evolutionary tuning and learning of fuzzy knowledge bases. World Scientific.
Cordón, O. (2011) ‘A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems’, International journal of approximate reasoning, 52(6), pp. 894–913.
Enache, M. C. (2015) ‘Web Application Frameworks.’, Annals of the University Dunarea de Jos of Galati: Fascicle: XVII, Medicine, 21(3).
Evans, E. (2004) Domain-driven design: tackling complexity in the heart of software. Addison-Wesley Professional.
Gandhi, R. et al. (2011) ‘Dimensions of cyber-attacks: Cultural, social, economic, and political’, IEEE Technology and Society Magazine, 30(1), pp. 28–38.
Goodpaster, K. E. (1991) ‘Business ethics and stakeholder analysis’, Business ethics quarterly, pp. 53–73.
GOST, R. (2009) ‘ISO/IEC 31010-2011 Risk management. Risk assessment methods’.
Gupta, R. et al. (2020) ‘Machine learning models for secure data analytics: A taxonomy and threat model’, Computer Communications, 153, pp. 406–440.
Husák, M. et al. (2018) ‘Survey of attack projection, prediction, and forecasting in cyber security’, IEEE Communications Surveys & Tutorials, 21(1), pp. 640–660.
Islam, S. et al. (2017) ‘A risk management framework for cloud migration decision support’, Journal of Risk and Financial Management, 10(2), p. 10.
Izuakor, C. and White, R. (2016) ‘Critical Infrastructure Asset Identification: Policy, Methodology and Gap Analysis’, in Critical Infrastructure Protection X: 10th IFIP WG 11.10 International Conference, ICCIP 2016, Arlington, VA, USA, March 14-16, 2016, Revised Selected Papers 10. Springer, pp. 27–41.
Kemabonta, T. and Kabalan, M. (2018) ‘Using What You Have, to Get What You Want–A Different Approach to Electricity Market Design for Local Distribution Companies (DISCOs) in Nigeria’, in 2018 IEEE Global Humanitarian Technology Conference (GHTC). IEEE, pp. 1–2.
Knight, S. and Burn, J. (2005) ‘Developing a framework for assessing information quality on the World Wide Web.’, Informing Science, 8.
Kure, H. and Islam, S. (2019) ‘Cyber Threat Intelligence for Improving Cybersecurity and Risk Management in Critical Infrastructure’, Journal of Universal Computer Science, 25(11), pp. 1478–1502.
Leroux, N. and Kaper, S. de (2014) Play for Java: Covers Play 2. Manning Publications Co.
Lilly, B. et al. (2019) ‘Applying Indications and Warning Frameworks to Cyber Incidents’, in 2019 11th International Conference on Cyber Conflict (CyCon). IEEE, pp. 1–21.
Machado, L., Filho, O. and Ribeiro, J. (2009) ‘UWE-R: an extension to a web engineering methodology for rich internet applications’, WSEAS Transactions on Information Science and Applications, 6(4), pp. 601–610.
Markowski, A. S. and Mannan, M. S. (2009) ‘Fuzzy logic for piping risk assessment (pfLOPA)’, Journal of loss prevention in the process industries, 22(6), pp. 921–927.
Martin, R. A. (2007) ‘Common weakness enumeration’, Mitre Corporation.
Mbanaso, U. M., Abrahams, L. and Apene, O. Z. (2019) ‘Conceptual Design of a Cybersecurity Resilience Maturity Measurement (CRMM) Framework’, African Journal of Information and Communication, 23, pp. 1–26.
Onochie, U. P., Egware, H. O. and Eyakwanor, T. O. (2015) ‘The Nigeria electric power sector (opportunities and challenges)’, Journal of Multidisciplinary Engineering Science and Technology, 2(4), pp. 494–502.
Rød, B. et al. (2020) ‘From risk management to resilience management in critical infrastructure’, Journal of Management in Engineering, 36(4), p. 4020039.

Sapori E, Sciutto M and Sciutto G (2014) 'ScienceDirect A quantitative approach to risk management in Critical Infrastructures', Transportation Research Procedia, 3(3), pp. 740–749. doi: 10.1016/j.trpro.2014.10.053.
Singh, S. K. et al. (2020) 'Machine Learning-Based Network Sub-Slicing Framework in a Sustainable 5G Environment', Sustainability, 12(15), p. 6250.
Straub, D., Boudreau, M.-C. and Gefen, D. (2004) 'Validation guidelines for IS positivist research', Communications of the Association for Information systems, 13(1), p. 24.
Strom, B. E. et al. (2017) 'Finding cyber threats with ATT&CK-based analytics', The MITRE Corporation, Bedford, MA, Technical Report No. MTR170202.
Tactic, A. (2017) 'Techniques and Common Knowledge (ATT&CK)'.
Tanwar, S. et al. (2019) 'Machine learning adoption in blockchain-based smart applications: The challenges, and a way forward', IEEE Access, 8, pp. 474–488.
Tounsi, W. and Rais, H. (2018) 'A survey on technical threat intelligence in the age of sophisticated cyber attacks', Computers & security, 72, pp. 212–233.
Workman, M., Bommer, W. H. and Straub, D. (2008) 'Security lapses and the omission of information security measures: A threat control model and empirical test', Computers in human behavior, 24(6), pp. 2799–2816.
Zadeh, L. A. (1988) 'Fuzzy logic', Computer, 21(4), pp. 83–93.
Abouzakhar, N. (2013) 'Critical infrastructure cybersecurity: A review of recent threats and violations', in European Conference on Information Warfare and Security, ECCWS, pp. 1–10.
Hokstad, P., Utne, I. B. and Vatn, J. (2012) 'Risk and vulnerability analysis of critical infrastructures', Springer Series in Reliability Engineering, 64, pp. 23–33. doi: 10.1007/978-1-4471-4661-2_3.
Cherdantseva, Y. et al. (2016) 'A review of cyber security risk assessment methods for SCADA systems', Computers & security, 56, pp. 1–27.
Bialas, A. (2016b) 'Risk Management in Critical Infrastructure—Foundation for Its Sustainable Work', Sustainability. doi: 10.3390/su8030240.
Gai, K. et al. (2016) 'Dynamic energy-aware cloudlet-based mobile cloud computing model for green computing', Journal of Network and Computer Applications. Elsevier, 59, pp. 46–54.
ISO 27005:2018 (ISO 27005) Information technology — Security techniques — Information security risk management, https://www.iso.org/standard/75281.html,
ISO 31000: 2018 (ISO 31000) Risk management — Guidelines, https://www.iso.org/standard/65694.html
ISO/IEC 27001,(ISO 27001) Information technology - Security techniques - Information security management systems - Requirements, https://www.iso.org/isoiec-27001-information-security.html
NIST Special Publication 800-39 , (NIST 800-39)Managing Information Security Risk, Organization, Mission, and Information System View, https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-39.pdf
NIST(NIST CSF) , Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1 , 2018, https://www.nist.gov/cyberframework
Centre of Internet Security (CIS), 2020. https://www.cisecurity.org/