IJMTST is UGC Approved Journal with Journal Id: 43137

International Journal for Modern Trends in Science and Technology, 8(S08): 120-129, 2022

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST08S0822
Available online at: https://www.ijmtst.com/vol8si08.html

Cyber Supply Chain Protection using Forecasting Insights for Cyber Threats

J.SagarBabu[1] | Ashok Kumar P S[2]

[1]Computer Science and Engineering, St.Peters's Engineering College, Hyderabad, TS, India
[2] Computer Science and Engineering, Don Bosco Institute of Technology, Banglore, KS, India

To Cite this Article
J.SagarBabu and Ashok Kumar P S. Cyber Supply Chain Protection using Forecasting Insights for Cyber Threats. International Journal for Modern Trends in Science and Technology 2022, 8(S08), pp. 120-129. https://doi.org/10.46501/IJMTST08S0822

Article Info
Received: 26 May 2022; Accepted: 24 June 2022; Published: 28 June 2022.

ABSTRACT

There are several subsystems in the Cyber Supply Chain (CSC) system that perform a variety of functions. The inherent vulnerabilities and threats from any element of the system, which may be exploited at any point in the supply chain, make supply chain security difficult. An interruption of this magnitude might have far-reaching consequences for business continuity. As a result, it is critical for businesses to identify and assess potential risks to their supply chains in order to implement effective countermeasures. Using multiple features such as threat actor skill and motivation, Tactics, Techniques, and Procedure (TT and P), and Indicator of Compromise (IOC), Cyber Threat Intelligence (CTI) delivers an intelligence study to find unknown to recognised dangers (IoC). This research examines and predicts risks to the cyber supply chain in an effort to increase its security level. Analysis and prediction based on the CTI attributes have been performed using Cyber Threat Intelligence (CTI) methodologies and Machine Learning (ML). This enables the identification of the inherent CSC vulnerabilities, allowing for the implementation of suitable control measures for the enhancement of cybersecurity generally. As a way to show our approach's viability, we collect a variety of CTI data and use a range of ML techniques, including as Logistic Regression, Support Vector Machine, Random Forest (RF), and Decision Tree (DT), using the Microsoft Malware Prediction dataset. As input parameters, the experiment analyses attacks and TTP, as well as vulnerabilities and indicators of compromise (IoC). Spyware/ransomware and spear phishing are the most predictable dangers in CSC, according to the predictions. In order to counter these dangers, we've proposed a variety of measures. CTI data should be included in the ML predicate model for CSC cyber security enhancement overall, according to our recommendations.

KEYWORDS: *Cyber Security, Cyber Attacks, Prediction, Malware, Vulnerabilities.*

1. INTRODUCTION

Smart CPS relies on a secure Cyber Supply Chain (CSC) to provide dependable service delivery and overall company continuity. By their very nature, CSC systems are difficult to understand, and vulnerabilities in the CSC system environment may spread across several nodes of the broader cyber physical system (CPS). The NCSC has released a list of CSC attacks that exploited system vulnerabilities [1]. Third-party service providers handle a portion of an organization's operations and data, increasing the risk of a security breach. There have been a number of successful CSC attacks in the recent past. CSC has been the target of cyber espionage groups such as Dragony [2], [3]. Due to a huge cyberattack, the Saudi Aramco power facility was forced to shut down. However, there is a lack of attention paid to the threat intelligence properties that can help improve overall cyber security.[4] Additionally, the organisation must be

able to forecast cyberattack trends so that it can plan its countermeasures in a timely manner [5]. It is not only the threat actors' motives and intentions that can be gleamed from predictive analytics, but also their current supply system vulnerabilities [6]. Using Cyber Threat Intelligence (CTI) and Machine Learning (ML) techniques, this paper aims to improve the cybersecurity of CSC systems by predicting cyberattack patterns on CSC systems and recommending appropriate controls to combat the attacks.By merging Cyber Threat Intelligence (CTI) and Machine Learning (ML) approaches to predict cyberattack trends on CSC systems and offer appropriate measures for combating assaults, this research intends to enhance the cybersecurity of CSC. Our work is innovative in three ways:

In the first place, we look to Cyber Threat Intelligence (CTI) for a methodical approach to acquiring and analysing data on the cyberthreat actor and the cyberattack itself. CTI gives evidence-based information about known assaults, which is why it is worth examining. This data is also utilised to uncover previously unidentified assaults, allowing security concerns to be better recognised and countered. CTI aims to prevent attacks and minimise the time it takes to uncover new ones by providing intelligence information.

For a second time, we used ML approaches and classication algorithms to forecast the assaults by mapping them to the CTI features. For this, we use a variety of classic techniques, including Logistic Regression (LG), Support Vector Machine (SVM), Random Forest (RF), and Decision Tree (DT). In order to predict an assault, we use CTI features like the Indicator of Compromise (IoC) and Tactics, Techniques and Procedure (TTP).

Final consideration for predicting prospective cyberattacks is a frequently used attack dataset [6]. Advance persistent threats, command and control, and industrial espionage are the primary emphasis of the forecast [7][9][10]. CTI and ML approaches may successfully be utilised to forecast cyberattacks and identify CSC systems vulnerabilities, according to the study's findings. In addition, our forecast shows that the TPR and FPR have a combined accuracy of 85%. LG and SVM were shown to be the most accurate in predicting threats, according to the data.

2. RELATEDWORK

In the field of cyber security, a number of popular CTI and ML models are in use [7]. Here you'll find examples of other people's work that we think is relevant to our own [8]. With the help of third-party service providers such as suppliers and distributors, the CSC security offers a safe foundation for the incoming and outgoing supply chains systems .[9] Suppliers and third-party vendors can safely exchange products and data via the internet in a secure supply chain context.[10] Operations technology (OT) and IT systems are integrated on Cyber Physical Systems infrastructures as part of this outsourced service.

[11]However, there are inherent threats, risks, and vulnerabilities that could be exploited by threat actors on the supply chain systems' operational technologies and information technologies. Data tampering, product delivery channel redirection, and data theft are all examples of outbound chain attacks.[12] Distributed denial-of-service (DDoS) attacks, IP address spoong, and software errors are just a few examples of the IT dangers that exist. It was proposed by NIST SP800 [13] that a four-tier framework for improving critical infrastructure cybersecurity incorporates the cyber supply chain risk management framework as one of its core components. [14]Tier 1 takes into account the CSC risk requirement strategy of the organisation. Supply chain risk identification, including products and services in the supply inbound and outbound chains, is taken into account at Tier 2. [15] Tier 3 implementation takes into account risk assessments, threat analyses, and the resulting impacts, and establishes the foundational requirements for a governance structure. Tier 3 implementation. It is Tier 4's responsibility to assess the supply chain risk associated with each product and service. The approach and tiers, on the other hand, focused on risk management rather than machine learning and threat prediction for the CSC domain's future trends. A supply chain attack framework and attack patterns were proposed by [14] to structure and codify supply chain attacks, as well Framework's goal was to provide a comprehensive view of supply chain attacks of malicious insertion across the acquisition lifecycle to identify the associated threat and vulnerability information.

For both known and undiscovered threats, cyber threat intelligence (CTI) has become one of the most important

actionable intelligences [4]. Cyberattacks and rising risks on CSC systems have had a disastrous effect on company processes, data, intellectual property, delivery channels, and the cost of recovery. Identification, threat analysis, and information dissemination to stakeholders are all part of CTI's methodology. Regarding threat intelligence platforms and CTI implementation processes, ENISA in [4] looked at the potential and constraints of existing threat intelligence platforms (TIP) from a strategic, tactical, and operational perspective while focusing on cybersecurity. The authors presented a threat intelligence programme model that gathers, normalises, enriches, correlates, analyses, and disseminates threat related information to stakeholders. Executive decision-making is a key consideration for strategic CTI objectives, while operational considerations include detecting information gaps in order to prioritise them for risk reduction. The operational objectives give a method for gaining a knowledge of the motivations, tactics, techniques, and capabilities of the threat actors. The procedures, on the other hand, do not make use of machine learning to forecast threats. [15] offers a threat intelligence-driven security model that incorporates six CTI stages and the procedures lifecycle necessary to establish intelligence objectives. Directing, gathering, processing, analysis, distribution, and feedback are all part of the CTI process.

For threat intelligence modelling, the author used data from a variety of sources, including network traffic, logs, and scans as well as data from outside sources like vulnerability databases and threat feeds. To make predictions, the threat intelligence-driven security approach relies on data gathered from network traffic, logs, and scans rather than machine learning techniques. In addition, establish cyber threat Intelligence metrics that take into account assets, requirements for corporate operations, opponent, and consumer intelligence. Intelligence needs; information collection; analysis; distribution; and use are all considered in the author's approach to threat intelligence. However, machine learning is not considered for anticipating unseen assaults in this strategy. A methodology provided by operationalizes and analyses adversarial behaviours across the lifespan of an organization's business process to identify the actions made by the attacker.

The author's method was based on the needs of the organization's intelligence, information collecting, analysis and dissemination in order to safeguard assets for strategic, tactical and operational knowledge and situational consciousness. There was a focus on the attacker's motivation and purpose, but not on ML for predicting threats. ML for CSC security is less important in the CTI functional process, which collects metrics and trend analysis for business risk assessment, prioritising, and decision support.

It has been shown that machine learning classifiers may be used to forecast cyberattack patterns in many cyber security application areas such as spam filters, antivirus, and IDS/IPS. Consider ML for Security [11], which proposes an algorithm for ML classification of HTTP attacks using a decision tree algorithm to learn a dataset for performance accuracy, and automatically label a request as valid or attack. Using a vector space model, which is often used to get data, the scientists created a classifier that flags any URL requests as potentially fraudulent. Compared to other methods, this one has a high level of accuracy and recall. In the CSC environment, however, the work was not focused on ML and threat prediction. They also investigated the potential of using machine learning models for cloud security testing in various operational situations and cloud scenarios.

The authors compared the classification algorithms Logistic Regression, Decision Tree, Nave Bayes, and SVM to learn a dataset for accuracy in performance. The algorithms are used in network security and represent supervised schemes.The detection of anomalous packets was found to be 97% accurate. However, the research did take into account supply chain security threats.

Data mining and machine learning (ML) methods were also examined to support intrusion detection in cybersecurity applications. As a means of identifying attack types (misuse) as well as detecting an attack, the authors employed artificial neural networks, association rules, fuzzy association rules, and Bayesian networks classifiers in their research (intrusion). Instead of using ML models and threat prediction in the CSC environment, the techniques used and methods employed were not. Analyze the cybersecurity dataset for machine learning algorithms used to analyse network traffic and detect anomalies, as well as In order to compare the dataset, the author compared the machine learning techniques, evaluation methods, and baseline classifiers used in the experiments. There are substantial issues in certain datasets when it comes to feature

selection and these datasets don't apply to contemporary intrusion detection. There were no complaints about the current Microsoft Malware Threat Forecast website dataset that we utilised for the prediction. A decision tree method was used to learn a dataset that represents the correlation and normalisation of security logs, and the categorization of logs was carried out using ML techniques on that algorithm. Algorithms are tested to see whether they can anticipate how well a classifier performs as an attack after a training phase. Anomalies and known attacks are included in the dataset. An 80 percent accuracy rate was achieved by using internet data to simulate the DT algorithm, which was used to construct a framework for normalisation and correlation. As a result, the classification model did not evaluate additional classification methods that are significant for ML performance and threat analysis, such as SVM, LR and RF.

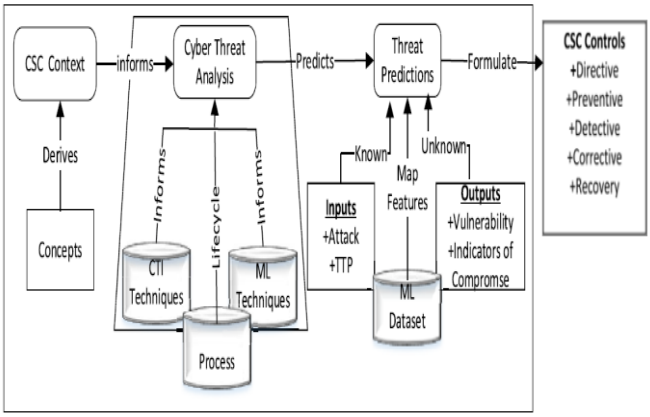
In another effort, machine learning algorithms are investigated for their usefulness in predicting power system disturbances and cyberattack discriminating classifiers and are focused specifically on identifying cyberattacks when deceit is the key tenet of the event. Deception is a key theme in cyber assaults, and therefore researchers in [14] analyzed the effectiveness of several algorithms to learn the dataset. They concentrated on identifying cyber attacks when deception is the primary objective. The authors of [12] for example suggested a technique for detecting SCADA power system cyberattacks that combined the CFS method with the K-Nearest-Neighbour (KNN) IBL algorithm of instance-based learning. The combination proved effective in reducing the enormous amount of characteristics and in increasing cyberattack detection accuracy while requiring little detection time. For SCADA-based IIoT platforms, [13] proposes an ensemble-learning approach for monitoring for cyberattacks. Random subspace (RS) learning and a random tree were used to create the model (RT). The authors in [10] suggested a semi-supervised deep-learning feature extraction approach for the protection of IIoT networks against cyberattacks at the trust border. The suggested method was flexible enough to learn about new threats.

CSC attacks from suppliers' inbound and outbound supply chains were not considered in the works. [12] predicted cybersecurity incidents by using ML algorithms to distinguish between different types of models based on ML predictive analytics on various datasets. Text mining techniques such as n-grams, bag of words, and machine learning were used by the authors to learn datasets for Naive Bayes and SVM algorithms.Predicting malware incident reaction and action response and classification accuracy was the goal of this investigation. The strategy failed to take into account CTI and ML in the context of the CSC system.As an additional risk-telling mechanism, the authors [15] advocated using an analysis of machine binary file appearance records to anticipate which computers are most likely to get infected with malware in the near future. An algorithm called random forest and semi-quantitative methodologies were utilised to produce a risk prediction model that captures use trends. The data show that 95% of the time, each degree of risk is associated with a machine infection occurrence. It's also important to note that [10] describe the amount to which an organization's network may be used to forecast cybersecurity events. As a way to predict future cybersecurity problems, the authors analysed Verizon's yearly data breach investigation report. Over 1000 event reports from multiple datasets were subjected to a random forest classifier. True positives accounted for 90 percent of the prediction result's accuracy. However, the research did not make any inferences or link the forecast to any current assaults. All of the works listed above are critical to the advancement of cyber security via the use of different machine learning methods. However, the security context of the CSC as a whole is not given enough attention. Only a few papers focus on attack prediction using threat intelligence data. It is possible that an assault on the CSC system might spread to other supply chain systems since cyberattacks are invisible. Because of this, ML analytics are needed in order to forecast cyberattacks as well as the hazards and vulnerabilities they pose. An organizational context for the threat analysis is also required.

CTI can assist you in achieving that objective. An important part of such effort may be found here. Using CTI and machine learning (ML), we've combined threat detection and analysis with threat forecasting to provide enterprises a better understanding of how to enhance CSC security.

3. PROPOSED SYSTEM ARCHITECTURE

Here, we'll go through the method we're thinking of taking to beef up CSC security as shown in Fig.1. It incorporates CTI and ML, as well as a methodical approach. The underlying concepts of the proposed approach are mentioned, such as the actor's goal and TTP as well as vulnerability, incident, and controls In order for CSC to concentrate on the probable system flaw, the technique evaluates both the incoming and outgoing chains of the vulnerability. Data is collected and analyzed using the CTI process, while machine learning is used to predict the threat. ML approaches are used to classification algorithms in order to learn a dataset for accuracy and predictive analytics purposes. Because the CTI lifecycle enables input parameters for detecting known assaults, as well as output parameters for forecasting future trends in both known and unknown attacks, it makes sense to integrate both CTI and ML for threat prediction.In order to do cyber threat assessments and make predictions, this method blends CTI operations with ML approaches. CSC network system nodes are to be scanned for vulnerabilities and signs of compromise based on previously discovered attacks. On CSC systems, we use CTI methods to collect threats (known attacks) and machine learning techniques to learn the dataset. When a threat actor infiltrates a system, the inputs they use are called assaults and TTP. It utilises attributes like attack type, pattern and vectors to figure out what kind of assault has been launched.



Predictive analytics for cyber threats may be achieved by integrating CTI procedures with ML approaches. CSC network system nodes are to be scanned for vulnerabilities and signs of compromise based on previously discovered attacks. Cyber threats (known attacks) may be gathered by using CTI methods, while machine learning techniques can be used to learn the dataset to predict cyber threats (unknown attacks). There are a number of inputs, including assaults and TTPs used by hackers to infiltrate a system. There are a number of attributes that the attack feature utilises to identify what kind of attacks were launched. The threat actor's TTP is a collection of attack patterns and attack vectors that they use. In addition to the threat actor's capabilities and threat indicators, TTP is a parameter that describes the threat. Using properties like user, system, and third-party vendors, the threat actor feature identifies the attack pattern by determining the vulnerable spots and the tools used in the attack. The threat actor's assault weapons and software programmes are known as "tools," and these are what they employ to conduct reconnaissance and launch an attack. Threat actors might utiliseNmap, Kali Linux and Metasploit to scan a network or exploit a network vulnerability, as an example. Vulnerabilities and signs of compromise are the output parameters, which are employed as threat intelligence. The threat actor's ability to penetrate a system and conduct an Advanced Persistent Threat (APT) attack and take command and control (C&C) is used to determine the indicators. Controls such as preventive, corrective, and recovery are also considered in order to secure the CSC system. Cyberattacks have a high degree of invincibility and uncertainty, which makes the threat environment unpredictable. This is the basis for our predictive analytics methodology. Similarly, forecasting cyberattacks in the context of the CSC organisation has proven difficult because of the shifting organisational needs, numerous interconnections, diverse business processes, and various delivery systems. To do so, the suggested method first evaluates important related studies and metamodel ideas to model the CSC assaults and CTI phases. Our supply chain attack indicators, for example, areidentified and integrated into the CTI stages. Our predictions are based on data that has been collected over the CTI process lifecycle and using machine learning (ML). To further enhance our threat prediction, we leverage the input and output parameters. Threat prediction findings are then assessed to offer accurate information on current and potential assaults and threats.

There are four stages to the procedure. Methodical and causal processes are used to determine strategy, threat analysis, threat prediction, and controls in each phase of

the strategy development process To accomplish the goals of the stages, each phase involves the necessary processes and activities. Input and output elements for our forecast include defining the CSC and security strategy of the enterprise, ML classifications, infrastructures, attack scenario, and so on. In order to identify the threat actor, threat profile, TTP, and IoC, participants in the threat analysis phase identify and collect threat information. They then assess and analyse that data. Input parameters for ML algorithms are considered, threats are predicted, and performance is evaluated using ML techniques to learn datasets during the threat prediction phase. In order to develop security policies and control mechanisms, the CSC conducts internal and external audits to identify the controls required. By following the process flow, we can better understand the phases and process.

Using CTI and cybersecurity risk strategy, CSC may identify how security goals and controls can be created and executed in accordance with the organization's goals and objectives. Assessing an organization's assets entails looking at everything from its infrastructure to its resources to the methods used to deploy those resources. Gathering data and formulating policy is made easier with CSC's security approach, which incorporates CTI and an evaluation of cybersecurity risks. Recursive management roles and duties complement each other to ensure security objectives are met. Plans to identify security objectives and to allocate responsibilities, including executive approval of designs and budget allocation, are supported by intelligence decision making in strategic management. Control and validation measures are based on signs of compromise when making tactical decisions about strategic management blueprints such as security requirements capturing, third-party audit, and conguration management plans. It is the responsibility of operational level managers to monitor, identify TTPs and escalate threat warnings for remediation and controls on an ongoing basis. By gathering information on potential adversaries' motivations, tactics, and capabilities as well as their threats and attack methods and tactics-of-attack (TTP), CTI Strategy offers management with evidence-based knowledge. As part of the organization's risk assessment approach, an overarching CSC risk strategy is developed, which identifies the policies needed to manage the company's operations.Risk assessment, CSC needs capturing, and business functions are all included. OT and IT acquisitions and integrations were also included into the risk strategy's implementation and procurement regulations.

For an assault to be successful, it must identify all the points where a supply chain may be breached. As an example, in the event of a malware attack, this activity seeks for essential information like as the source of the assault, the tools used to carry it out, trends, and attack pathways. In order to identify possible signs of an attack, we employ threat activities, adversary behaviour, dangerous occurrences, or the current condition of the incident. It's possible that a threat actor may utilise these signs to find any weaknesses in the system.

It is possible to conduct penetration testing, vulnerability assessments, and threat propagation exercises to identify the supply chains on the OT (and IT) by following the following stages:. For an assault to be successful, it must identify all the points where a supply chain may be breached. While this activity searches for important information, such as the source of an attack, tools, trends, and attack pathways based on an examination of the malware that was used as an indication for this activity. We utilise threat activities, adversary behaviour, dangerous occurrences, or the current condition of the incident to identify possible signs of an attack. They may help discover any weaknesses that might be exploited by a malicious party. To establish the OT and IT supply inbound and outgoing chains, the following procedures [2] may be followed to conduct penetration testing, vulnerability assessment tests, and threat propagation exercises. To reduce CSC risks, risk assessors examine the likelihood and effect of CSC intrusions and threats, as well as the weak links in supply chains and third-party organisations that might be exploited. It identifies any and all dangers to the system that may be present. The CSC security domain is risk assessed, and the dangers of unauthorised access to secure areas are analysed. A third-party audit may expose a wide range of hazards that can be mitigated by recognising and addressing them. Classify them based on the services they provide and the degree to which they are integrated into the different supply chain network systems that they support. Analysis of the threats is what this activity is all about, and it concentrates on determining the source of an attack as well as its kind, its

pattern, its technique, and its vectors. This will make it easier to determine what kind of IoC and controls are necessary. Nodes susceptible to cyberattacks are considered by merging CTI and machine learning to forecast known and undiscovered assaults via three sequential actions in this phase. Vulnerabilities and IoC are intended to be output features of this activity. The IoC gives the indications of penetrations, cybercrime compromises, APTs, and C&Cs, while the vulnerabilities provide the company knowledge about places that may be exploited.

We use ML approaches and a dataset to predict the output characteristics based on the cyber threat assessments and the input features. Network nodes, firewalls, anti-virus, and anti-malware are all susceptible. On this list are all of the potential assaults and alterations that might be made to the system by the threat actor in order to get access to the system. This kind of assault is so difficult to predict that it can't be taken at face value. Using ML approaches to train and test datasets, we may collect different attack probabilities and their propagation impacts on the CSC. It is possible to use a confusion matrix to assess the performance of a model based on a particular test dataset. It's simply a comparison between the actual target values and the anticipated values from the machine learning model. A better knowledge of the values may be gained via the calculation and analysis of data in the matrix. When the dataset's occurrences are classified, there are four possible results.

TP, FP, FP, and FN rates are all examples of True Positive (TP), True Negative (FP), and False Positive (FP). For example, if an instance is positive and the result is positive, it is TP, or FP if the outcome is negative. It is a TN if the instance and result are both negative; otherwise, it is a FN [15]. The confusion matrix may be understood using the following way. The percentage of correct guesses in the total number of predictions is what we use to measure the accuracy of the confusion metric. Formulas for determining the entropy and the thermodynamic properties of matter are given below. The MAE calculates the difference vector's sum of the absolute mean and normal curve.

Whereas, MSE provides a rough estimate of the amount of the mistake by calculating the absolute value of the square root of the mean and reconverting the units to the original unit of the output variable. We utilised MAE and MSE to make predictions about actual numbers or regressions. Import AUC-ROC Function, for example, is one of the tasks involved. Set the entropy criterion and import the mean absolute error and the mean square error. In information theory, entropy is used to assess the degree of uncertainty in the source of data. Uniquely, it provides us with the level of disorganisation in our data by satisfying four uncertainties in our confusion matrix. Controls that can be used to mitigate the danger are identified in this last step. In order to reduce the dangers, the controls should guarantee that the necessary security strategy and mechanisms are in place Internal and external audits, threat monitoring, and reporting are all part of this process. Third-party audit, information, and identification of existing controls are all part of the process. During the phase of threat analysis, data is gathered using the CTI technique.

As threat indicators, we look for network node vulnerabilities, IP address threats, IEDs, and other forms of cyberthreats related to the organization's mission. This comprises the threat actors' methods of operation and their TTPs. We use the meta-attack model's concepts and attributes to identify the attack pattern and TTP used against the CSC in our investigation. During this stage, you'll need to identify potential attack vectors, weak areas, and potential TTP threats. All of the endpoints' antimalware logs are collected, as well as a signature, threat indicators, and event logs from the IDS/IPS. For the implementation, we utilised a dataset from a Microsoft Malware website. Microsoft's endpoint systems have been the target of malware attacks, according to this dataset. There were more than 40,000 items in the MicrosoftWindows Defender database, with 64 columns and each row representing separate telemetry data entries. Multiple malware assaults have been identified on endpoint nodes located in various regions, each having a unique machine identity, time-stamped timestamp, organisational identifier, and default browser identifier. Accordingly, the dataset has been sampled to contain a bigger percentage of malware-infected devices than a typical Microsoft customer's PC.

Since the CSC systems incorporate diverse network infrastructures for business process and interoperability, we utilised this dataset for predictive analytics.This is what the feature description has to say about it: Geography-specific machine identifiers, such as GeoNameIdentier, use MachineIdentier to determine a

machine's geographic location. DefaultBrowsersIdentier is a class that identifies the machine's installed default web browsers. Provides a unique identifier for the company that owns the computer. A determined eld based on the Spynet Report's AV Products eld is provided. The operating system's process architecture is taken into account by the processor. True if the computer has TPM is indicated by HasTpm (Trusted Platform Module). Look at the current operating system version. OsBuild, the current operating system's build information.

"Census Device Family" The DeviceClass property identifies the desktop or mobile device for which an OS version is designed. True for Windows 8.1 and higher with Windows rewall enabled if this property has been set. Uploading data from a website APIs or HTML le, choosing the data we want, and saving it as a CSV le are the steps involved in this activity. The average of the dataset's columns is used to prepare the data. It was also possible to import already generated datasets using the machine learning identifyer's categories. As a result, 40,000 training datasets with 62 variables were created. The most important aspects of the source dataset have been identified that are pertinent to our research. The main data had 62 characteristics, with an emphasis on our past work's attack ideas, tools, and vulnerabilities. With the goal of identifying potential vulnerabilities, we analysed the behaviours of threat actors to determine their intent and previous patterns of behaviour. As a result, we identified eight susceptible places and the likelihood that a cyber attacker may exploit those locations, which included the following: Firewall, IDS/IPS, Vendors CSC system, Network, IP Addresses and Databases, as well as Software and Websites. The attackand TTP are addressed as input factors for the predictions.In order to do this, we must first identify the attack kinds, tools, vectors, and input capabilities. Following the dataset's current feature description in [6], we developed the new features. The current datasets and characteristics [6] are used to identify additional features for predicting attack inputs and outcomes [7, 8]. All of these characteristics and variables may be found in the dataset we used in this study. Observed attack types may be described using attack patterns, which are abstract mechanisms for expressing how an attack is carried out. Using attack patterns, TTPs, and vulnerabilities as compromise indicators, we calculate the output parameters.

Additionally, the ML prediction dataset includes built-in attack roles. In order for our ML to be able to accurately forecast which node is susceptible and hence more likely to be attacked, we need to be able to create attack proles. For our future predictions, even if we can't utilise specific features, we may look at traits that are linked to and useful in describing how attacks start and vulnerabilities are exploited in the real world. Many of the parameters we examined were selected to indicate the stakeholders' CTI and security awareness. We utilised a pipeline to link the several classications in order to pick the classifiers. To estimate the parameters, we make use of 10-fold cross-validation. For each method, we perform the 10-Fold cross-validation and check the parameter ten times to ensure that the values don't change and that the results are correct. During the test, we employed a 10-fold cross validation method to get the most accurate prediction data possible. An estimator's parameter values may be searched exhaustively using the GridsearchCV. In order to get the average score of all the outcomes, we use the Majority Voting (MV) method in the classifiers. Our last method for distinguishing between binary classifications is ROC-AUC.

4. RESULTS AND DISCUSSION

Threat prediction findings are presented and analysed in this section. As a result, we evaluate a variety of factors, such as attack likelihood, TTP, weak points, and IoC. A probability scale of 0.00000 percent is used to determine the propagation. Calculations were used to determine how severe each modification was. low (15%), middle (16%) or high (16%) (above 60 percent).Predicting the likelihood of an assault.

To see how these algorithms fared in classifying cyber attacks in response to the specific malicious assault, look at Fig.2 (LR, SVM, RF). LR, DT, SVM, and RF all attained accuracy levels of 66%, 63%, 62%, and 66%, respectively, in the table. In terms of Precision, Recall, and F-Score, LR and RF outperformed the other classifiers, however DT and SVM had subpar results. An 85 percent accuracy rate in identifying various forms of reactions to Malware, Ransomware, and spyware assaults can be shown by comparing it to the attack categories.Forecasting TTP deployment based on cyberattacks' responses.

Classication algorithms' effectiveness in recognising different TTPs launched and responses depending on the provided attack vectors. DT and SVM performed poorly in comparison to XSS, session hijacking, and RAT assault in terms of low accuracy recall and F-score, respectively. For malware attacks, LR obtained the greatest precision and F-score with an accuracy of 83%. With an 83 percent accuracy rate for harmonic mean, ransomware and spyware assaults identified numerous forms of replies for the TTPs, including rootkit, email attachment and RAT.

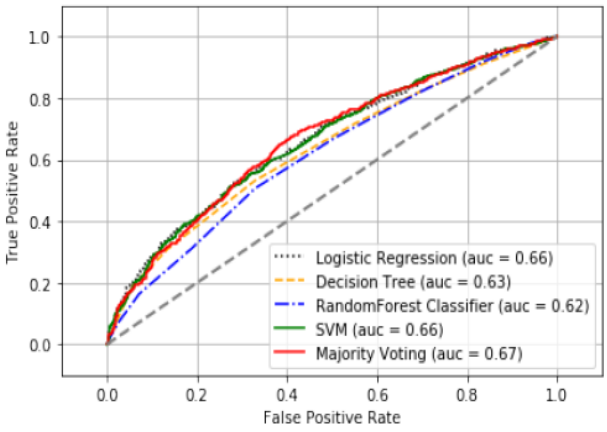


Fig. 2. Plot the LG's ROC curve accuracy for all algorithms SVM, DT, and RF in MV.

There are various findings from the CSC assaults that may be used to enhance the overall security of the Smart CPS by leveraging CTI lifecycle procedures and ML for predictive analysis. According to the findings, the supply chain systems are vulnerable to remotecommand execution and system manipulation by attackers, posing a number of security difficulties for the company. When it comes to evaluating and forecasting future assaults and attack tendencies, the machine learning approach to cyber security has shown to be beneficial. Threat intelligence systems that may forecast which CSC system nodes are vulnerable to assaults are developed using machine learning (ML) techniques using classification algorithms such as SVM, DT, RF, and MV. To calculate the true positive and false negative rates, we plot the accuracy of all methods in ROC. AUC ROC. With an accuracy of 0.66, SVM was shown to be the best parameter in the study.

We can mix algorithms using ML to see which ones offer the greatest accuracy and output for the parameters we're trying to predict. However, it does not allow us to get insight into the motivations and intentions of the threat actor. Intrusion detection security apps use network classifiers for cyber security detection and analytics. While [12] studied datasets for network traffic analysis, [13] looked at datasets that may be used to identify current intrusions. Furthermore, [13] used a decision tree approach to represent the correlation and normalisation of security logs to study log classification. Similarly, [14] examined the performance and ML predictions for power system disturbance and cyber attack discriminations of NNge. LF, DT, Nave Bayes, and SVM classication algorithms. For detection of smart grid hacks, [15] employed an instance-based learning classication algorithm to learn the dataset. To identify assaults on Industrial IoT networks, [12] employed an ensembled learning model that combines a random subspace with a random tree. To counteract IoTcybersecurity risks in a smart city, researchers [10] used ML approaches to train datasets for anomaly detection models such as LR and SVM classifiers and deep learning models such as DT and RF.Using deep learning on a semi-supervised model to identify unknown cyberattacks, [9] also suggested an innovative industrial IoT network trust border security.

A down sample encoder cooperative data generator was utilised to train the algorithm to capture the real distribution of the assault model on the industrial IoT attack surface. In addition, researchers in [11] used Naive Bayes and SVM algorithms to study and analyse different datasets obtained from SMEs to forecast cybersecurity events. A risk teller system that uses machine learning to identify whether machines are at risk of becoming infected or clean and forecasts if a company will encounter cybersecurity issues in the future is also modelled [6]. Despite the fact that each and every one of these projects contributes to the advancement of cyber security. A lack of attention is paid to the entire CSC security, and ML classifiers are utilised to predict threats. The proposed studyintegrates ideas from the CSC and CTI domains to give a conceptual picture. It uses CTImethods to do a thorough threat analysis and then incorporates ML classifiers to provide predictions about the severity of the danger. In order to learn the malware threat prediction dataset, we also evaluated LG, DT, SVM, and RF algorithms in Majority Voting.

5. FUTURE SCOPE AND CONCLUSION

For both national and global contexts, the integration of complex cyber physical infrastructures and applications in a CSC environment has resulted in economic, commercial, and social impacts. Since any weakness in any area of the system might affect the total supply chain environment, CPS security remains an issue. By merging CTI and ML for threat analysis and prediction, this research intends to enhance the security of CSCs. We used CSC and CTI principles, as well as a methodical approach, to assess and forecast the danger. Majority Voting accuracy was shown using the LG, DT, SVM, and RF algorithms, and a list of potential risks was compiled as a consequence of the research. In addition, we found that CTI is successful in extracting threat information that can be integrated into ML classifiers for threat forecasting. In this way, the CSC organisation is able to analyse the current controls and develop new measures to increase overall cyber security. We can't generalise our findings without considering the process as a whole, as well as a real-world industrial case study. In addition, we'll be looking at the current controls and the ones that may be needed in the future, depending on our predictions.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] M. Swann, J. Rose, G. Bendiab, S. Shiaeles and F. Li, "Open Source and Commercial Capture The Flag Cyber Security Learning Platforms - A Case Study," 2021 IEEE International Conference on Cyber Security and Resilience (CSR), 2021, pp. 198-205, doi: 10.1109/CSR51186.2021.9527941.

[2] A. M. Kanca and Ş. SAĞIROĞLU, "Sharing Cyber Threat Intelligence and Collaboration," 2021 International Conference on Information Security and Cryptology (ISCTURKEY), 2021, pp. 167-172, doi: 10.1109/ISCTURKEY53027.2021.9654328.

[3] A. Aigner and A. Khelil, "A Security Scoring Framework to Quantify Security in Cyber-Physical Systems," 2021 4th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS), 2021, pp. 199-206, doi: 10.1109/ICPS49255.2021.9468168.

[4] G. Langner, J. Andriessen, G. Quirchmayr, S. Furnell, V. Scarano and T. J. Tokola, "Poster: The Need for a Collaborative Approach to Cyber Security Education," 2021 IEEE European Symposium on Security and Privacy (EuroS&P), 2021, pp. 719-721, doi: 10.1109/EuroSP51992.2021.00058.

[5] M. ÖZARAR, A. Akansu and B. Hasbay, "Impact of Cyber Maturity Level on Health Sector," 2021 International Conference on Information Security and Cryptology (ISCTURKEY), 2021, pp. 127-131, doi: 10.1109/ISCTURKEY53027.2021.9654395.

[6] H. F. Al-Turkistani and H. Ali, "Enhancing Users' Wireless Network Cyber Security and Privacy Concerns during COVID-19," 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA), 2021, pp. 284-285, doi: 10.1109/CAIDA51941.2021.9425085.

[7] N. Sun et al., "Defining Security Requirements With the Common Criteria: Applications, Adoptions, and Challenges," in IEEE Access, vol. 10, pp. 44756-44777, 2022, doi: 10.1109/ACCESS.2022.3168716.

[8] P. Lau, L. Wang, Z. Liu, W. Wei and C. -W. Ten, "A Coalitional Cyber-Insurance Design Considering Power System Reliability and Cyber Vulnerability," in IEEE Transactions on Power Systems, vol. 36, no. 6, pp. 5512-5524, Nov. 2021, doi: 10.1109/TPWRS.2021.3078730.

[9] Y. Kawanishi, H. Nishihara, H. Yoshida and Y. Hata, "A Study of The Risk Quantification Method focusing on Direct-Access Attacks in Cyber-Physical Systems," 2021 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), 2021, pp. 298-305, doi: 10.1109/DASC-PICom-CBDCom-CyberSciTech52372.2021.00059.

[10] Ö. Durmuş and A. Varol, "Analysis and Modeling of Cyber Security Precautions," 2021 9th International Symposium on Digital Forensics and Security (ISDFS), 2021, pp. 1-8, doi: 10.1109/ISDFS52919.2021.9486345.

[11] X. Ning and J. Jiang, "Design, Analysis and Implementation of a Security Assessment/Enhancement Platform for Cyber-Physical Systems," in IEEE Transactions on Industrial Informatics, vol. 18, no. 2, pp. 1154-1164, Feb. 2022, doi: 10.1109/TII.2021.3085543.

[12] N. Koroniotis, N. Moustafa, F. Schiliro, P. Gauravaram and H. Janicke, "The SAir-IIoT Cyber Testbed as a Service: A Novel Cybertwins Architecture in IIoT-Based Smart Airports," in IEEE Transactions on Intelligent Transportation Systems, doi: 10.1109/TITS.2021.3106378.

[13] M. Wan, J. Li, Y. Liu, J. Zhao and J. Wang, "Characteristic insights on industrial cyber security and popular defense mechanisms," in China Communications, vol. 18, no. 1, pp. 130-150, Jan. 2021, doi: 10.23919/JCC.2021.01.012.

[14] J. Diakoumakos, E. Chaskos, N. Kolokotronis and G. Lepouras, "Cyber-Range Federation and Cyber-Security Games: A Gamification Scoring Model," 2021 IEEE International Conference on Cyber Security and Resilience (CSR), 2021, pp. 186-191, doi: 10.1109/CSR51186.2021.9527972..

[15] Z. Sun and S. Zhang, "Modeling of Security Risk for Industrial Cyber-Physics System under Cyber-attacks," 2021 4th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS), 2021, pp. 361-368, doi: 10.1109/ICPS49255.2021.9468202.