Cyber supply chain threat analysis and prediction using machine learning and ontology

Cyber Supply Chain Threat Analysis and Prediction using Machine Learning and Ontology

Abel Yeboah-Ofori[1], Haralambos Mouratidis[2,3] ,Umar Ismail[1], Shareeful Islam[1], Spyridon Papastergiou[4]

[1]School of Architecture, Computing and Engineering, University of East London, UK

[2]Centre for Secure, Intelligent and Usable Systems, University of Brighton, UK

[3]Department of Computer and System Science, University of Stockholm, Sweden

[4] Department of Informatics ,University of Piraeus, Greece

abel2@uel.ac.uk, haralambos@dsv.su.se, u.ismail@uel.ac.uk,shareeful@uel.ac.uk,paps@unipi.gr

Abstract. Cyber Supply Chain (CSC) security requires a secure integrated network among the sub-systems of the inbound and outbound chains. Adversaries are deploying various penetration and manipulation attacks on an CSC integrated network's node. The different levels of integrations and inherent system complexities pose potential vulnerabilities and attacks that may cascade to other parts of the supply chain system. Thus, it has become imperative to implement systematic threats analyses and predication within the CSC domain to improve the overall security posture. This paper presents a unique approach that advances the current state of the art on CSC threat analysis and prediction by combining work from three areas: Cyber Threat Intelligence (CTI), Ontologies, and Machine Learning (ML). The outcome of our work shows that the conceptualization of cybersecurity using ontological theory provides clear mechanisms for understanding the correlation between the CSC security domain and enables the mapping of the ML prediction with 80% accuracy of potential cyberattacks and possible countermeasures.

Keywords: Cyber Security, Ontology, Cyber Supply Chain, Machine Learning, Threat Prediction, Cyber Threat Intelligence

1 Introduction

Cyber Supply Chain (CSC) security nowadays is more challenging due to the inherent system complexity and vulnerabilities among various system components and their cascading effect. Cybersecurity risks in CSC have increased exponentially leading to major security breaches in most organizations [1, 2]. The recent high profile cyberattacks such as Ukraine 2015 and Saudi Aramco 2017 smart grid attacks have brought diverse challenges, different threat landscape and unexpected challenges with unpredictable consequences [3]. Therefore, it has become imperative to have a comprehensive understanding of the CSC threat landscape. However, threat analysis in CSC is challenging due to a lack of understanding of the evolving threat landscape which often hinders the ability of organizations to analyze and effectively predict threats [1, 2].

This paper presents a novel threat analysis and prediction approach that uniquely combines work from Cyber Threat Intelligence (CTI), Ontology, and Machine Learning. In particular, this paper provides three main contributions. Firstly, we analyze CSC

threats using CTI and ontological theory. Ontologies provide semantic mapping and explicit knowledge necessary for threat analysis. Secondly, we present a systematic process to analyse and predicate cyber threats. The process includes activities related to cyber threat intelligence and machine learning techniques such as Random Forest (RF) and GBoost algorithms for threat analysis and prediction. ML is considered for mapping the relationships between cyberattack, cyber threat propagations and their cascading impact on the various supplier chain nodes. Thirdly, we integrate knowledge from datasets from the Microsoft Malware Prediction to support threat prediction [5]. The results show that the ontological approach provides mechanisms for understanding the correlation between the CSC security domain. Both RF and GBoost algorithms provide accuracy around 80%.

2 Related Work

2.1 Cyber Supply Chain and Threat Intelligence

Cyber supply chain (CSC) security provides secure integrated networks for various organizations. CSC attacks have increased exponentially, and its cascading impact is unquantifiable, causing collateral damage to organizations. Threat actors are using sophisticated attacks including advanced persistent threats (ATP) and command and control (C&C) methods to penetrate, manipulate and obfuscate in the supply inbound and outbound chains [7, 8, 9]. Cyber Threat Intelligence (CTI) provides technical indicators, context, and actionable advice relating to existing and emerging threat [9]. Pokorny 2018 proposed a CTI lifecycle approach required to identify intelligence goals [10]. Friedman & Buchanan, proposed a CTI approach based on organizational requirements, gathering information, analysis and dissemination to protect assets and documents [11]. Miller proposed a cyber supplier chain framework and attack pattern that provides a comprehensive view of supply chain attacks of malicious insertions across a full question life cycle [12]. The protection of the CSC is critical as it incorporates various embedded networks, software and computational algorithms for information flows and data structures in the live and mission-critical system. [13].

2.2 Ontology and Machine Learning for Cyber Security

Security ontology from the CSC perspective describes organizational security concepts, properties relationships and their interdependencies in a formal and structured manner [14]. The goal of security ontology is to extract relevant attack instances and information from data to ensure consistency and accuracy in the CSC security concepts and for knowledge reuse in the threat intelligence domain. Mozzaquatro et al proposed a model driven ontology-based cybersecurity framework for the internet of things that considers design time and run time concepts for knowledge reasoning [4]. Gao et al., proposed an ontology-based model of network and computer attacks for security assessment and standards classifications that establishes relationships among network security services, threats, vulnerabilities and causes of failures [15]. Gyrard et al. proposed an ontology for attacks and countermeasures for capturing and presenting

concepts of security requirements [16]. Machine Learning (ML) in cybersecurity uses various algorithms to learn and train datasets to determine their classifications and for threat predictions. ML algorithm is initially trained to allow the system to learn the data [17]. The purpose of using ML is to get the system to use past events to make an informed decision that can be used to predict future attacks [18].

3 Approach

This section provides an overview of the proposed approach and the underlying process for threat analysis and predication.

3.1 Integration of CTI, Ontology, and Machine Learning

The cyber threat intelligence is based on the threat actor profile, Tactic, Technique and Procedure (TTP), attack context and Indicator of Compromise (IoC) to provide an intelligence analysis about the threat. The proposed approach includes additional concepts related to CSC such as supply chain actor and controls. The ontology uses these concepts for a common understanding of the threat domain of CSC. Note that, due to the space limitations the details related to the concepts are not included in the paper. The ML techniques can effectively be used to analyse large data and discover the hidden patterns specifically relating to current and future threats. Such approach significantly assists organizations to gain situational awareness and understanding of the threat landscape.

Figure 1 shows the proposed approach that integrates CTI, ontology, and ML. The threat intelligence concepts are formalized using ontological theory and further used by the ML for the prediction. The CTI concepts provide information regarding threat actor intention, underlying techniques and indicators of the attack. The ontology concepts are then considered for the knowledge representation, semantic visualization and reusability of the knowledge which are useful for the CSC threat analysis [9]. Finally, the ML considers two classification models to determine the best performance and accuracy for the threat predication. We have considered Random Forest (RF) and GBoost classification algorithms for this purpose. RF is widely used for large and diverse datasets as it uses randomly selected subsets of samples to construct models to form a forest. Similarly, GBoost can also be used for large and diverse datasets and can train large datasets by gradually sequentially adding each subset in the optimization algorithms. A pipeline was used in an ensemble to link the RF and Gboost algorithms in a voting machine (VM) and ROC-AUC to plot the classifiers.

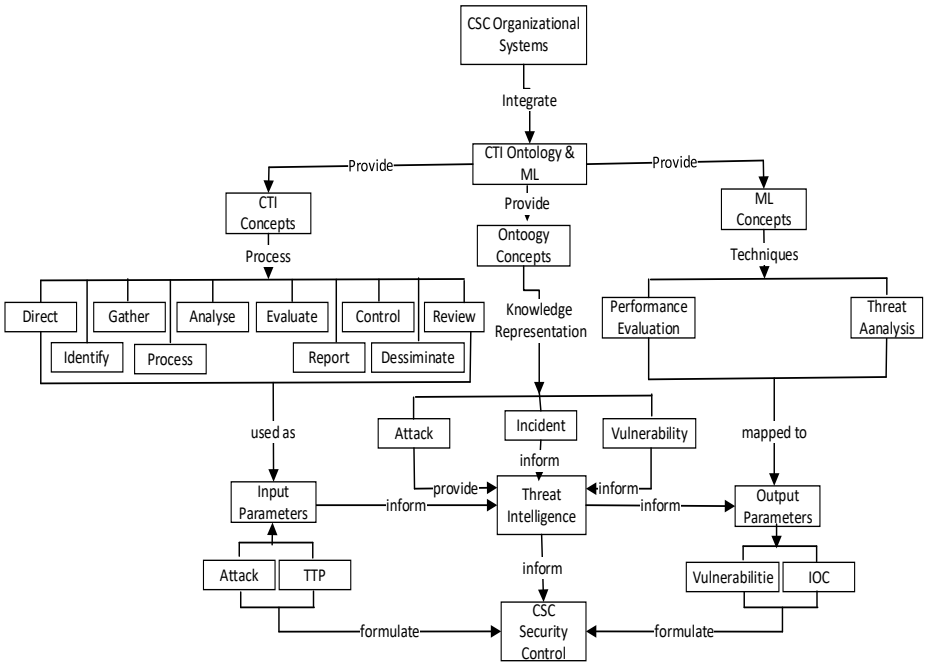

Fig. 1. Proposed Approach

3.3 Ontological View for Threat Analysis

An Ontology provides a formal language that enables the explicit specification and conceptualization of ideas that represent an abstract model of a phenomenon [12]. An Ontology enables the construction of knowledge and provides the advantage of knowledge representation in organized metadata of complex information resources. Ontology concepts are applied to address the challenges of hierarchical relationship, taxonomy and structured set of rules by facilitating, formalizing, decomposition and specification of the general categories of concepts in the CSC security domain. The ontology allows the provision of relationships and concepts representing the invariant conditions of CSC security. By using semantic rules and logical representation of the concepts, we modelled a graphical visualization of the concepts to aid the automated assessment, analysis, and data processing.

The ontological view of concepts presented in the previous section provides a taxonomy of the CSC concepts. However, it is imperative to use a knowledge representation technique to provide a general overview of the CSC concepts, including a detailed vocabulary that shows conditional obligations and logical reasoning that support ML reasoning. Hence, the ontological view of the concepts is explicitly transformed into their corresponding semantic rules to express the valid conditions that exist between the concepts. Such rules are vital for expressing any complex CSC domain knowledge, relationships and statements, as well as the reuse, extensibility, and sharing of the CSC concepts. Also, without such a precise formalization, the CSC knowledge representation may appear vague and ambiguous. Thus, the rules are vital for supporting machine reasoning. Figure. 2 explains the rules defining the ontological view of the CSC security

concepts and provides an overview of the ontological presentation and the underlying concepts. Figure 3 shows the rule set relevant for the threat and asset which are relevant for the threat analysis [20] and note that only a part of rules is added due to the space restriction.

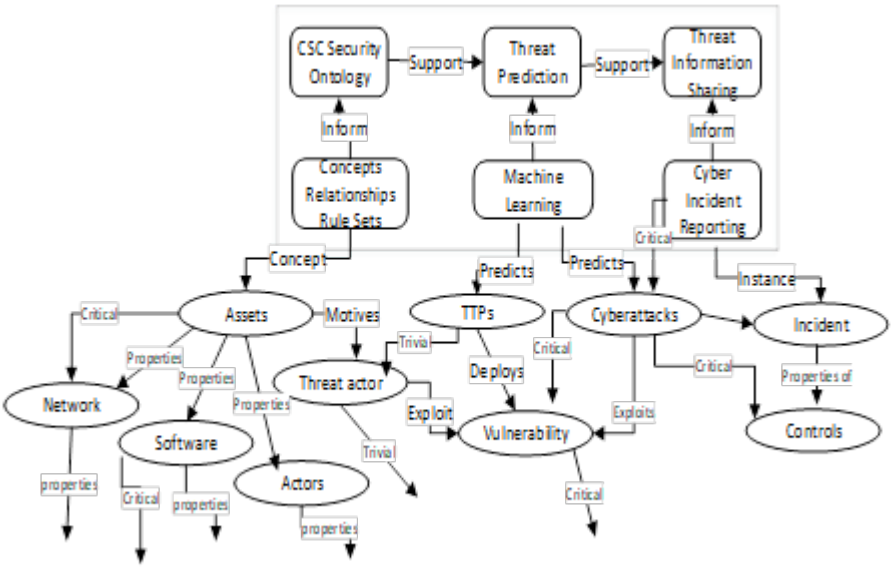

Fig 2. Ontology Concepts and Properties

[∀x(actor(x) → cambe(y) → internalActor(r) ∧ externalActor(x) ∧ threatActor(z)¬ canBe(r,x,z))]
[∀x(threatActor(x) → employs(y) → TTP(r) → toExploit(e) ∧ vulnerabilities(v)¬ canBe(r,x,z))]
[∀x(asset(x) → has(x) → assetType(q) → data(d) ∧ software(s) ∧ hardware(q) ↔ canBe (q,d,s,q,)]
[∀x(attack(x) → has(q) ∧ attackSeverity(s) ∧ attackVector(v) ∧ attackType(y) → for(f) → hardwareAsset(h)
∧ softwareAsset(w) ∧ dataAsset(d) ↔ canBe (s,v,y,h,w,d))]
[∀x(attackSeverity(x) → camBe(q) → lowSeverity(r) ∧ highSeverity(s) ∧ mediumSeverity(x) ¬ canBe(r.s.x))]
[∀x(attackVector(v) → camBe(q) → designFlaws(f) ∧ incorrectPermission(p) ∧ insufficientValidation(v)
∧ minsconfiguration(z) ↔ canBe (f,p,v,z,))]
[∀x(attackType(v) → camBe(q) → manipulationAttack(x) ∧ penetrationAttack(z) ↔ canBe (x,z,))]

Fig 3. Ontology Rule Set for Threat Analysis

3.4 Process

This section presents the overall process of threat analysis and prediction. It consists of two phases for the CSC threat analysis and predication.

Phase 1: Threat Analysis
The threat analysis phase considers the underlying CTI gathering process and ontology concepts for analyzing the CSC security domain. It considers a number of steps, i.e., CTI gathering, CSC ontology and ML. CTI provides information required for actionable decision makings to preventing cyber-attacks. CTI gathering follows steps to identify, gather and analyze the threat information for evaluate and controls. CTI provides evidence-based knowledge of threat actor's motives, intents, TTPs, and indicators of compromise (IoC) and control mechanisms relating to the existing and emerging threat. Once the threat data is consolidated then ontology is used to present the underlying concepts for the threat. This phase follows cybersecurity ontology learning to describe CSC security concepts, properties and the relationships required to model CSC security goals. Thi step considers all the relationships required to ensure control mechanisms are in place on the supply chain environment and their implementation standards.

Phase 2: Threat Prediction
The final phase of the process is to predicate the threat using ML techniques. It includes, data representation, feature selection, choosing and classification algorithm for the ML techniques to learn a dataset, input and output features for the prediction purpose.

- **Data Presentation:** the dataset was collected from the chosen data sets [5].
- **Feature Selection:** the feature selection process assists in normalizing the dataset. The feature selection identifies irrelevant column, removes duplicate columns, reduces dimensions, and prepares the dataset for training and testing.
- **Classification for Prediction:** RF and GBoost algorithms are run in MV to learn classification for accurate responses. The AUC-ROC distinguishes between probabilities and determines the right performance metrics to evaluate the algorithms.
- **Performance Evaluation**: The performance of the models is evaluated based on the TP, TN, FP and FN values and the elements of the confusion matrix [19].

The evaluation criteria considers precision and recall in determining actual or predictive values in the feature extraction. The F-Score determine the harmonic mean for precision and recall [18].

4 Implementation

The objective of the implementation phase is to explore the applicability of our approach by using ML classification algorithms for threat predictions.

4.1 Cyber Threat Intelligence Gathering

The CTI gathering and process lifecycle steps include Direct, Identify, Gather, Process, Analyze, Evaluate, Report, Controls, Disseminate and Review.

- Direct: CTI goals are put together by strategic management to identify security goals and inform proper CSC security controls.
- Identify: Identify organizational goals, CSC requirements, assets, CSC network nodes, IP address, technical threats, and user threats actors.

- Gather: Data gather indicators of compromise from various endpoint's nodes, including firewalls logs, IDS/IPS reports, signatures, antimalware reports
- Analyze: Analyze IDS/IPS logs, firewall logs, and Antimalware intrusions to predict attacks that could be fed into the CTI.
- Evaluate: Evaluate threats, levels of risks, impacts on the CSC system and the effects on organizational goal.
- Report: Analysis and evaluations of Known and Unknown attacks for strategic management decision-making on threats levels.
- Disseminate: Designates cyber threat information sharing to all stakeholders on the CSC system.
- Review: Requires ad-hoc, periodical and annual reviews and updates to monitor current trends and alerts.

The application of ontology for the CSC security concepts enables the exchange, sharing and reuse of cyber threat information automatically, thereby providing a semantically stable structure of the underlying knowledge of CSC systems security. We use ontology to identify and map CSC concepts such as actors, assets, threats, attacks, vulnerability, TTPs and incident reporting that provide conceptual reasoning, relational knowledge and understanding of cyber threat intelligence required.

4.2 ML Threat Prediction

This section follows ML techniques to learn the dataset as discussed in section 3 for threat predictions. That include data preparation, description, feature extraction, choosing an optimization algorithm and determining the performance accuracies as follows.

Data Description: The dataset used is from a publicly available data source from a Microsoft Malware Prediction data website [5]. The data was collected by Microsoft Windows Defender with over 40,000 entries with 62 columns and each row represents different telemetry data entries. Each row in the dataset corresponds to a machine uniquely identified by a machine Identifier. The dataset integrates systems using other operating systems that do not represent Microsoft customer's machine only as it has been sampled to include a much larger proportion of malware machines.

Data Preparation: The format of the data is mainly csv files. We used the Kaggle API to downloaded the dataset. A NaN dictionary was created to handle all the unwanted duplicates and removed 8 duplicates leaving 54 columns as relevant from the 62 columns.

Feature Extraction: There are total 64 features in the data set . However, 32 features extracted which are relevant to understand the attack profile. The rest are features with duplicate samples and those with zero variances. Thus, we reduce the features by setting the selection criteria and the variance threshold in the code to remove all zeros to ensure dimensionality reduction during the training. Further, we derived the secondary data from the primary data by identifying features considered as probable threat in [2] for the ML.

Choosing an Optimization Algorithm for the Classifiers: The ML algorithms used for the work are RF and GBoost. A multiclass classifications approach was used in a AUC_ROC to model the selection metric for the multiclass classification problem. We used each classifier against all to distinguish between the probabilities of the classes to obtain the performance indices for Precision, Recall and F-Score. A pipeline was used to connect the algorithms in the loop to determine the best optimization algorithm. A 10-Fold cross-validation was used to determine the parameter estimation and validation for consistent and accurate result. We combine the algorithms using Majority Voting (MV) in the classifiers to determine the mean score of the results. Finally, the ROC-AUC distinguished between the accuracies of the binary classifiers for the predictions**.**

Determining the Performance Evaluation and Accuracies: Figure 4 depicts the results of the accuracies and combines 2 classification algorithms, RF and GB in a pipeline and run in a ROC curve to determine the true positive and false positive rates using the 10-Fold cross-validation. RF produces a performance result of 73% compared to GB 79% with a majority voting of 78%. The highest classifier from the performance model was GB as it can predict better performance in predicting attack. However, the results show a slight reduction in the overall score with the MV score of 78%. Further, it shows higher accuracy for the TPRs and FPRs as compared to Figure 2 where the performance went down when we included the RF algorithm.

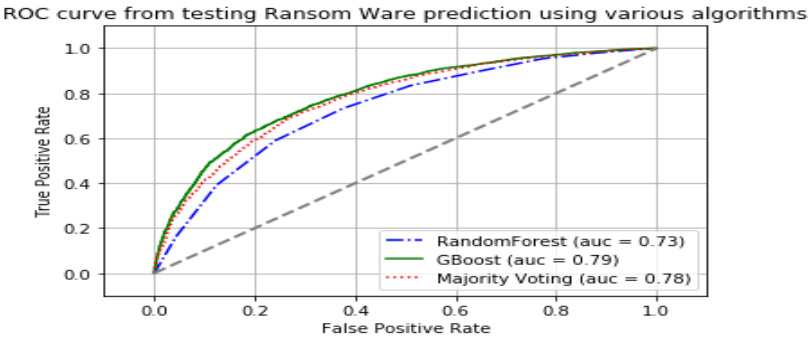

Fig. 4. Roc Curve for Prediction the RM and GBoost Algorithms in VM

10-fold cross-validation:
ROC AUC: 0.73 (+/- 0.01) [RandomForest]
ROC AUC: 0.79 (+/- 0.02) [GBoost]
ROC AUC: 0.78 (+/- 0.01) [Majority Voting]

5 Results

This section presents the results of using ML algorithms and presents the accuracies of the threats predication using ML evaluation using Precision, Recall, and F-Score.

Predicting the different types of responses based on the type of cyberattack: To predicting the different types of responses based on the type of cyberattack, we refer to

the probabilities of cyberattacks in our previous work [6], for the various accuracies of the attacks. Table 1 presents the performance of the classifications of RF and GBoost algorithms in identifying the various responses of cyberattacks based on the given malicious attack. From the table, RF achieved an accuracy of 80% and GBoost 78%. Comparing the performance of the classifiers, RF performed better for the Precision (P), Recall (R) and F-Score (F), whilst GBoost received a low precision, recall and F-score. Comparing that to the attack's categories signifies that Malware, Ransomware and spyware attacks provided different types of responses with 80% accuracy.

Table 1. Threat Predication on Endpoint Nodes Using RF and GBoost Classifiers

ACCURACY	RF 80%			GBOOST 78%		
ATTACKS	P	R	F	P	R	F
XSS/Session Hijacking	0.74	0.66	0.72	0.72	0.70	0.71
Ransomware	0.80	0.76	0.78	0.79	0.76	0.77
Spear Phishing	0.73	0.69	0.71	0.74	0.70	0.72
RAT/Island Hopping	0.76	0.74	0.75	0.72	0.70	0.71
Malware	0.79	0.76	0.77	0.79	0.76	0.77
Spyware	0.77	0.74	0.76	0.78	0.75	0.76
DDoS	0.71	0.69	0.70	0.78	0.75	0.76

Predicting TTPs deployed based on the response of the cyberattacks : To predicting the different types of responses based on the type of TTPs deployed, we determine the various accuracies of the attacks. Table 2 presents the performance of the classification algorithms in identifying the various TTPs deployed and the responses based on the given attack vectors. Comparing the TTPs deployed against the cyberattack such as XSS, session hijacking and RAT attack, RF achieved a low accuracy of 79% whereas GBoost achieved a higher accuracy of 82% for the precision, recall and F-score. Furthermore, ransomware, malware and spyware attacks identified different types of responses for the TTPs with 82% accuracy for the harmonic mean in identifying the attack vectors spear Phishing, email attachments, RAT and rootkit attacks.

Table 2. Identify the different TTP deployed based on the response of the cyberattacks

Accuracy	RF 79%			Gboost 82%		
ATTACKS	P	R	F	P	R	F
XSS/Session Hijacking	0.75	0.71	0.72	0.76	0.72	0.74
Ransomware	0.79	0.75	0.78	0.81	0.76	0.79
Spear Phishing	0.76	0.73	0.75	0.76	0.73	0.75
RAT/Island Hopping	0.77	0.74	0.76	0.77	0.75	0.76
Malware	0.79	0.75	0.77	0.82	0.77	0.81
Malware/Spyware	0.78	0.74	0.77	0.79	0.76	0.78
DDoS	0.72	0.70	0.71	0.73	0.72	0.71

6 Discussion

This section discusses the observations made from the evaluation of the accuracies and results of the CSC threat predictions. The study revealed several challenges facing organizations in securing the CSC systems as threat actors are executing arbitrary commands on the supply chain systems remotely and manipulating systems. For us to predict the accuracies of the threats, that identified vulnerable spots on the CSC network system that could be exploited, the percentage score of manipulation and the probability levels [2]. The vulnerable spots include network, vendor, website, firewall, IDS/IPS, software, IP, and the database system.

Table 3. Predicting Threat Indicators on Vulnerable spots

Attacks	Vulnerable Spots	Penetration	Manipulation%	Probability
Remote Access Trojan	Firewall	Y	80%	High
Island Hopping	IDS/IPS	Y	75%	High
Ransomware	Vendor	Y	90%	High
Session Hijacking	Network	Y	70%	Medium
DDoS	IP	Y	80%	Medium
Malware	Database	Y	75%	High
Malware	Software	Y	95%	High
XSS	Website	Y	90%	High

Comparing the features descriptions listed in Table 1 with the cyber-attack prediction Table 1 and TTP Table 2, we predict various cyberattacks that could be initiated on the vulnerable spots in Table 3. CSC systems integrate with other network systems using public-facing IPs. Threat actors could exploit the default browsers, websites and initiate attacks to penetrate the system and cause various manipulations as in Table 3.

6.1 Comparison with Existing Works

A number of works focus on using CTI, ML, and ontology for the threat analysis. For instances, [10] considered CTI approach for gathering intelligence goals, [11] considered CTI approach that protect assets, and [12] proposed a comprehensive CSC framework for attack pattern. Further, [4] proposed cybersecurity ontology framework for IoT and knowledge reasoning, [15] considered an ontology model that establishes relationships among networks, and [16] proposed a security ontology for capturing requirements. Furthermore, [17, 18, 19] used ML techniques on various algorithms to

learn datasets for performance accuracies and predictions. All the works are relevant and contributes to cyber security improvement, however none of the works considered integrating CTI, ontology and ML to extract relevant attack instances for knowledge representation and threat predictions in CSC security domain.

6.2 CSC Controls

This section discusses controls that support CSC security threats and predictions in line with the control ontology. The challenge of developing security controls, configuration settings with the best security properties is a complex task and goes beyond the ability of individual users as it requires analysis of potential threats and risks on the CSC system. It is therefore required to have an inventory of current control mechanism including audit trails of the other organizations and third-party vendors. This certainly supports to determine the existing security capabilities, prediction of TTP and indicators in order to determine the necessity of additional controls for the overall cyber security improvement. Establish, implement, and actively managing, tracking, reporting on, and correcting the security configuration using configuration management tools and change control process could prevent external penetration and manipulations that could lead to malware and ransomware attacks. There are different types of controls such as directive, preventive, and preventive need to select based on the nature of threat.

7 Conclusions

Predicting cyber supply chain threats has proved daunting due to the various network integrations and the complexities involved in the different configurations. Further, the sophisticated and stealthy nature of cyberattacks on CSC systems has made a threat analysis of CSC security threat analysis very challenging. In this work, we have used CTI and CSC security ontology concepts to analyse the threat and ML techniques to predict threats. CSC ontology concepts provided us with knowledge reuse in the CTI domain and an understanding of the attack instances. The ML predictions indicate the precision of 80% accuracy for cyberattacks such as malware and ransomware on systems without regular antivirus updated. The ontology provided knowledge of security controls that systematize all security phenomenon. Future works include data sets from other sources to generalize and improve our prediction results. Further, it is also required to automate the process so that CSC organizations can feed the data into the tools and obtain the prediction results.

Acknowledgments

This research has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952690. The results of this paper reflect only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

References

1. Woods, B., Bochman, A.: Supply Chain in the Software Era. Atlantic Council: Washington, DC, USA, (2018). Doi: 10.1109/5254.920602
2. Yeboah-Ofori, A., Islam, S.: Cyber Security Threat Modeling for Supply Chain Organizational Environments. Future Internet, 11, 63, (2019). doi: 10.3390/611030063.
3. Department for Business Innovation and Skill. Information Security Breaches Survey. Technical Report. PWC and InfoSecurity. (2013).
4. Mozzaquatro, B. A., Agostinho, C., Goncalves, D, Martins, J., Jardim-Goncalves, R.: An Ontology-Based Cybersecurity Framework for the Internet of Things. MDPI. Sensors, 18, 3053; (2018) doi:10.3390/s18093053
5. Microsoft Malware Prediction. Research Prediction. Kaggle Dataset. (2019). https://www.kaggle.com/c/microsoft-malware-prediction/data. Last Accessed 2021/01/28.
6. Maedche, A., and Staab. S.: Ontology Learning for the Semantic Web. IEEE Intell. Syst., 16, 72–79. (2001).
7. CERT-UK.: Cyber-Security Risks in the Supply Chain. TLP Whitepaper.
8. US-Cert. "Building Security in Software & Supply Chain Assurance. https://www.us-cert.gov/bsi/articles/knowledge/attack-patterns. Last Accessed 2020/11/24
9. ENISA "Exploring the Opportunities and Limitations of Current Threat Intelligence Platforms" Version 1. (2017).
10. Porkorny, Z.: What Are the Phases of The Threat Intelligence Lifecycle? The Threat Intelligence Handbook. (2018).
11. Freidman, J., Bouchard. M.: Cyber Threat Intelligence Guide: Using Knowledge About Adversary to Win The War Against Targeted Attacks. iSightPartners. (2018).
12. Miller, J. F.: Supply Chain Attack Framework and Attack Pattern. Mitre. (2013).
13. Boyens. J.: Integrating Cybersecurity into Supply Chain Risk Management. RSA; Moscone Center: San Francisco, CA, USA, (2016).
14. Arbanas, K. and Cubrilo, M.: Ontology in Formation Security. JIOS. Vol 39. No.2. PP. 107-136. (2015).
15. Gao, J., Zhang, B., Chen and X., Luo, Z.: Ontology-based model of network and computer attacks for security assessment. 18(5):554–562, (2013). DOI: 10.1007/s12204-013-1439-5
16. Gyrard, A., Bonnet, C., Boudaoud, K.: The STAC (Security Toolbox: Attacks & Countermeasures) Ontology. Conference, Pages 165–166, Rio de Janeiro, Brazil, (2013).
17. Villano. E. G. V.: Classification of Logs Using Machine Learning. Norwegian University of Science and Technology. (2018).
18. Boschetti, A., Massaron. L: Python Data Science Essentials. (2016). 2nd Edition. UK. ISBN 978-1-78646-213-8.
19. Mohasseb, B. Aziz, J. Jung, and J. Lee. Predicting Cyber Security Incidents Using Machine Learning Algorithms: A case study of Korean SMEs. (2019). University of Portsmouth.
20. Martimiano L. A. F., Moreira, E. S.: An OWL-based Security Incident Ontology Protégé. Conference, pages 1-4, Madrid, Spain. (2005). Semantic Scholar.