A

Major Project

On

Predictive Analytics for Cyber Threats to Improve Cyber Supply Chain Security

(Submitted in partial fulfillment of the requirements for the award of Degree)

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

BY

KADARI GOHATHI (197R1A05L7)
KATHI VENKATA SAKETH REDDY (197R1A05L6)
KOSUMBA PRADEEP (197R1A05M0)

Under the Guidance of
M.MADHUSUDHAN
(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CMR TECHNICAL CAMPUS

UGC AUTONOMOUS

(Accredited by NAAC, NBA, Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi)Recognized Under Section 2(f) & 12(B) of the UGC Act.1956,
Kandlakoya (V), Medchal Road, Hyderabad-501401.

2019-2023

ABSTRACT:

Cyber Supply Chain (CSC) system is complex which involves different sub-systems performing various tasks. Security in supply chain is challenging due to the inherent vulnerabilities and threats from any part of the system which can be exploited at any point within the supply chain. This can cause a severe disruption on the overall business continuity. Therefore, it is paramount important to understand and predicate the threats so that organization can undertake necessary control measures for the supply chain security. Cyber Threat Intelligence (CTI) provides an intelligence analysis to discover unknown to known threats using various properties including threat actor skill and motivation, Tactics, Techniques, and Procedure (TT and P), and Indicator of Compromise (IoC). This paper aims to analyse and predicate threats to improve cyber supply chain security. We have applied Cyber Threat Intelligence (CTI) with Machine Learning (ML) techniques to analyse and predict the threats based on the CTI properties. That allows to identify the inherent CSC vulnerabilities so that appropriate control actions can be undertaken for the overall cybersecurity improvement. To demonstrate the applicability of our approach, CTI data is gathered and a number of ML algorithms, i.e., Logistic Regression (LG), Support Vector Machine (SVM), Random Forest (RF), and Decision Tree (DT), are used to develop predictive analytics using the Microsoft Malware Prediction dataset. The experiment considers attack and TTP as input parameters and vulnerabilities and Indicators of compromise (IoC) as output parameters. The results relating to the prediction reveal that Spyware/Ransomware and spear phishing are the most predictable threats in CSC. We have also recommended relevant controls to tackle these threats. We advocate using CTI data for the ML predicate model for the overall CSC cyber security improvement.

EXISTING SYSTEM :

A recent NCSC report highlights a list of CSC attacks by exploiting vulnerabilities that exist within the systems. Organizations outsource part of their business and data to the third-party service providers that could lead any potential threat. There are several examples for successful CSC attacks. For instance, Dragonfly, a Cyber Espionage group, is well known for targeting CSC organization. The Saudi Aramco power station attack halted its operation due to a massive cyberattack. There are existing works that consider CSC threats and risks but a lack of focus on threat intelligence properties for the overall cyber security improvement.

DISADVANTAGES:

- Those attacks on the cyber physical and cyber digital system components such as distributed denial of service (DDoS) attacks, IP address spoofing, and Software errors.
- The data is loss in the company server due to malware attack.

.

PROPOSED SYSTEM:

This paper aims to improve the cybersecurity of CSC by specifically focusing on integrating Cyber Threat Intelligence (CTI) and Machine Learning (ML) techniques to predicate cyberattack patterns on CSC systems and recommend suitable controls to tackle the attacks. The novelty of our work is threefold:

- Firstly, we consider Cyber Threat Intelligence (CTI) for systematic gathering and analysis of information about the threat actor and cyber-attack by using various concepts such as threat actor skill, motivation, IoC, TTP and incidents. The reason for considering CTI is that it provides evidence-based knowledge relating to the known attacks. This information is further used to discover unknown attacks so that threats can be well understood and mitigated. CTI provides intelligence information with the aim of preventing attacks as well as shorten time to discover new attacks.
- Secondly, we applied ML techniques and classification algorithms and mapped with the CTI properties to predict the attacks. We use several classification algorithms such as Logistic Regression (LG), Support Vector Machine (SVM), Random Forest (RF) and Decision Tree (DT) for this purpose. We follow CTI properties such as Indicator of Compromise (IoC) and Tactics, Techniques and Procedure (TTP) for the attack predication.
- Finally, we consider widely used cyberattack dataset to predict the potential attacks [6]. The predication focuses on determining threats relating to Advance Persistent Threat (APT), command and control and industrial espionage which are relevant for CSC. The result shows the integration of CTI and ML techniques can effectively be used to predict cyberattacks and identification of CSC systems vulnerabilities.

ADVANTAGES:

- our prediction reveals a total accuracy of 85% for the TPR and FPR.
- The results also indicate that LG and SVM produced the highest accuracy in terms of threat predication.
- Greater efficiency.

HARDWARE REQUIREMENTS:

- System : Pentium IV 2.4 GHz.
- Hard Disk : 40 GB.
- Floppy Drive : 1.44 Mb.
- Monitor : 15 VGA Colour.
- Mouse : Logitech.
- Ram : 512 Mb.

SOFTWARE REQUIREMENTS:

- **Operating System** **:** Windows
- **Coding Language** : Python 3.7

CONCLUSION :

The integration of complex cyber physical infrastructures and applications in a CSC environment have brought economic, business, and societal impact for both national and global context in the areas of Transport, Energy, Healthcare, Manufacturing, and Communication. However, CPS security remains a challenge as vulnerability from any part of the system can pose risk within the overall supply chain context. This paper aims to improve CSC security by integrating CTI and ML for the threat analysis and predication. We considered the necessary concepts from CSC and CTI and a systematic process to analyse and predicate the threat. The experimental results showed that accuracies of the LG, DT, SVM, RF algorithms in Majority Voting and identified a list of predicated threats. We also observed that CTI is effective to extract threat information, which can integrate into the ML classifiers for the threat predication. This allows CSC organization to analyze the existing controls and determine additional controls for the improvement of overall cyber security. It is necessary to consider the full automation of the process and industrial case study to generalize our findings. Furthermore, we are also planning to consider evaluating the existing controls and the necessary of future controls based on our prediction results.