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

K. GOHATHI (197R1A05L7)

K.SAKETH (197R1A05L6)

K.PRADEEP (197R1A05M0)

Under the Guidance of

M.MADHUSUHAN

(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 UGCAct.1956,

Kandlakoya (V), Medchal Road, Hyderabad-501401.

2019-23

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the project entitled "**PREDICTIVE ANALYTICS FOR CYBER THREATS TO IMPROVE CYBER SUPPLY CHAIN SECURITY**" being submitted by **K. GOHATHI (197R1A05L7), K. SAKETH (197R1A05L6) & K. PRADEEP (197R1A05M0)** in partial fulfillment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering to the Jawaharlal Nehru Technological University Hyderabad, is a record of bonafide work carried out by him/her under our guidance and supervision during the year 2022-23.

The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

M.Madhusudhan
(Assistant Professor**)**
INTERNAL GUIDE

Dr. A. Raji Reddy
DIRECTOR

Dr. K. Srujan Raju
HOD

EXTERNAL EXAMINER

Submitted for viva voice Examination held on _____

ACKNOWLEDGEMENT

Apart from the efforts of us, the success of any project depends largely on the encouragement and guidelines of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project.

We take this opportunity to express my profound gratitude and deep regard to my guide **M.Madhusudhan,** Assistant Professor for his exemplary guidance, monitoring, and constant encouragement throughout the project work. The blessing, help, and advice given by him shall carry us a long way in the journey of life on which we are about to embark.

We also take this opportunity to express a deep sense of gratitude to the Project Review Committee (PRC) **Dr. Punyaban Patel, Ms. K. Shilpa, Dr. M . Subha Mastan Rao & J. Narasimharao** for their cordial support, valuable information and guidance, which helped us in completing this task through various stages.

We are also thankful to **Dr. K. Srujan Raju,** Head of the Department of Computer Science and Engineering, **Dr. Ashuthosh Saxena,** Dean R&D, and **Dr. D T V Dharmajee Rao,** Dean Academics for providing encouragement and support for completing this project successfully.

We are obliged to **Dr. A. Raji Reddy,** Director for being cooperative throughout the course of this project. We also express our sincere gratitude to Sri. **Ch. Gopal Reddy,** Chairman for providing excellent infrastructure and a nice atmosphere throughout the course of this project.

The guidance and support received from all the members of **CMR Technical Campus** contributed to the completion of the project. We are grateful for their constant support and help.

Finally, we would like to take this opportunity to thank our family for their constant encouragement, without which this assignment would not be completed. We sincerely acknowledge and thank all those who gave support directly and indirectly in the completion of this project.

K.GOHATHI (197R1A05L7)

K.SAKETH (197R1A05L6)

K.PRADEEP (197R1A05M0)

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.

LIST OF FIGURES/TABLES

LIST OF SCREENSHOTS

TABLE OF CONTENTS

1.INTRODUCTION

1. INTRODUCTION

1.1 PROJECT SCOPE

Cyber Supply Chain (CSC) security is critical for reliable service delivery and ensure overall business continuity of Smart CPS. CSC systems by its inherently is complex and vulnerabilities within CSC system environment can cascade from a source node to a number of target nodes of the overall cyber physical system (CPS). There are several examples forsuccessful CSC attacks. For instance, Dragonfly, a Cyber Espionage group, is well known for targeting CSC organization.

1.2 PROJECT PURPOSE

Due to the invincibility nature of cyber attacks on the cyber supply chain (CSC), and the cascading effects of malware infections, we use machine learning to predict attacks. As organizations have become more reliant on CSC systems for business continuity, so are the increase in vulnerabilities and the threat landscapes. Some traditional approach to detecting and defending malware attack has largely been antimalware or antivirus software such as spam filters, firewall, and IDS/IPS. These tools largely succeed, however, as threat actors get more intelligent, they are able to circumvent and affect nodes on systems which then propagates.

1.3 PROJECT FEATURES

In this project, we use ML techniques to learn the dataset and predict which CSC nodes have detection or no detection. The purpose is to predict which modes are venerable to cyberattacks and for predicting future trends. To demonstrate the applicability of our approach, we used a dataset from the Microsoft Malware Prediction website. Further, an ensemble is used to link Logistic Regression, and Decision Tree and SVM algorithms in Majority Voting and run on the training data and then use 10-fold cross validation to test the parameter estimation, accurate results and predictions. The results show that ML algorithms in Decision Trees methods can be used in cyber supply chain predict analytics to detect and predict future cyber attack trends.

.

2.LITERATURE SURVEY

2.LITERATURE SURVEY

2.1 Towards the Prediction of Renewable Energy Unbalance in Smart Grids

AUTHORS: R. D. Labati, A. Genovese, V. Piuri and F. Scotti

ABSTRACT: The production of renewable energy is increasing worldwide. To integrate renewable sources in electrical smart grids able to adapt to changes in power usage in heterogeneous local zones, it is necessary to accurately predict the power production that can be achieved from renewable energy sources. By using such predictions, it is possible to plan the power production from non-renewable energy plants to properly allocate the produced power and compensate possible unbalances. In particular, it is important to predict the unbalance between the power produced and the actual power intake at a local level (zones). In this paper, we propose a novel method for predicting the sign of the unbalance between the power produced by renewable sources and the power intake at the local level, considering zones composed of multiple power plants and with heterogeneous characteristics. The method uses a set of historical features and is based on Computational Intelligence techniques able to learn the relationship between historical data and the power unbalance in heterogeneous geographical regions. As a case study, we evaluated the proposed method using data collected by a player in the energy market over a period of seven months. In this preliminary study, we evaluated different configurations of the proposed method, achieving results considered as satisfactory by a player in the energy market.

2.2 Malware Attack Predictive Analytics in a Cyber Supply Chain Context Using Machine Learning

AUTHORS: A. Yeboah-Ofori and C. Boachie

ABSTRACT: Due to the invincibility nature of cyber attacks on the cyber supply chain (CSC), and the cascading effects of malware infections, we use machine learning to predict attacks. As organizations have become more reliant on CSC systems for business continuity, so are the increase in vulnerabilities and the threat landscapes. Some traditional approach to detecting and defending malware attack has largely been antimalware or antivirus software such as spam filters, firewall, and IDS/IPS. These tools largely succeed, however, as threat actors get more intelligent, they are able to circumvent and affect nodes on systems which then propagates. In our previous work, we characterized threat actor activities, including presumed

intent and historically observed behaviour, for the purpose of ascertaining the current threats that could be exploited. In this paper, we use ML techniques to learn the dataset and predict which CSC nodes have detection or no detection. The purpose is to predict which modes are venerable to cyberattacks and for predicting future trends. To demonstrate the applicability of our approach, we used a dataset from the Microsoft Malware Prediction website. Further, an ensemble is used to link Logistic Regression, and Decision Tree and SVM algorithms in Majority Voting and run on the training data and then use 10-fold cross validation to test the parameter estimation, accurate results and predictions. The results show that ML algorithms in Decision Trees methods can be used in cyber supply chain predict analytics to detect and predict future cyber attack trends.

2.3 Feasibility of Supervised Machine Learning for Cloud Security

AUTHORS: D. Bhamare, T. Salman, M. Samaka, A. Erba and R. Jain

ABSTRACT: Cloud computing is gaining significant attention, however, security is the biggest hurdle in its wide acceptance. Users of cloud services are under constant fear of data loss, security threats and availability issues. Recently, learning-based methods for security applications are gaining popularity in the literature with the advents in machine learning techniques. However, the major challenge in these methods is obtaining real-time and unbiased datasets. Many datasets are internal and cannot be shared due to privacy issues or may lack certain statistical characteristics. As a result of this, researchers prefer to generate datasets for training and testing purpose in the simulated or closed experimental environments which may lack comprehensiveness. Machine learning models trained with such a single dataset generally result in a semantic gap between results and their application. There is a dearth of research work which demonstrates the effectiveness of these models across multiple datasets obtained in different environments. We argue that it is necessary to test the robustness of the machine learning models, especially in diversified operating

conditions, which are prevalent in cloud scenarios. In this work, we use the UNSW dataset to train the supervised machine learning models. We then test these models with ISOT dataset. We present our results and argue that more research in the field of machine learning is still required for its applicability to the cloud security.

2.4 A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection

AUTHORS: A. L. Buczak, and E. Guven.

ABSTRACT: This survey paper describes a focused literature survey of machine learning (ML) and data mining (DM) methods for cyber analytics in support of intrusion detection. Short tutorial descriptions of each ML/DM method are provided. Based on the number of citations or the relevance of an emerging method, papers representing each method were identified, read, and summarized. Because data are so important in ML/DM approaches, some well-known cyber data sets used in ML/DM are described. The complexity of ML/DM algorithms is addressed, discussion of challenges for using ML/DM for cyber security is presented, and some recommendations on when to use a given method are provided.

2.5 A Review of Cyber Security Dataset for Machine Learning Algorithms

AUTHORS: O. Yavanoglu and M. Aydos

ABSTRACT: It is an undeniable fact that currently information is a pretty significant presence for all companies or organizations. Therefore protecting its security is crucial and the security models driven by real datasets has become quite important. The operations based on military, government, commercial and civilians are linked to the security and availability of computer systems and network. The objective of this review is to explain and compare the most commonly used datasets. This paper focuses on the datasets used in artificial intelligent and machine learning techniques, which are the primary tools for analyzing network traffic and detecting abnormalities.

3.SYSTEM ANALYSIS

3.SYSTEM ANALYSIS

3.1 INTRODUCTION

System Analysis is the important phase in the system development process. The System is studied to the minute details and analyzed. The system analyst plays an important role of an interrogator and dwells deep into the working of the present system. In analysis, a detailed study of these operations performed by the system and their relationships within and outside thesystem is done. A key question considered here is, "what must be done to solve the problem?" The system is viewed as a whole and the inputs to the system are identified. Once analysis is completed the analyst has a firm understanding of what is to be done.

3.2 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 toa 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.

3.2.1 DISADVANTAGES OF EXISTING SYSTEM

- 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

3.2 PROPOSED SYSTEM

- 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 isfurther 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.

3.3 ADVANTAGES OF THE PROPOSED SYSTEM

The system is very simple in design and to implement. The system requires very low system resources and the system will work in almost all configurations. It has got following features

- 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.

3.4 FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- Economic Feasibility
- Technical Feasibility
- Social Feasibility

3.4.1 ECONOMIC FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available.Only the customized products had to be purchased.

3.4.2 TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

3.4.3 BEHAVIORAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

3.5 HARDWARE & SOFTWARE REQUIREMENTS

3.5.1 HARDWARE REQUIREMENTS:

Hardware interfaces specifies the logical characteristics of each interface between the software product and the hardware components of the system. The following are some hardware requirements.

- Operating system : windows, linux
- Processor : minimum intel i3
- Ram : minimum 4 gb
- Hard disk : minimum 250gb

3.5.2 SOFTWARE REQUIREMENTS:

Software Requirements specifies the logical characteristics of each interface and software components of the system. The following are some software requirements

- Python idel 3.7 version (or)
- Anaconda 3.7 (or)
- Jupiter (or)
- Google colab

4.ARCHITECTURE

4.ARCHITECTURE

4.1 PROJECT ARCHITECTURE

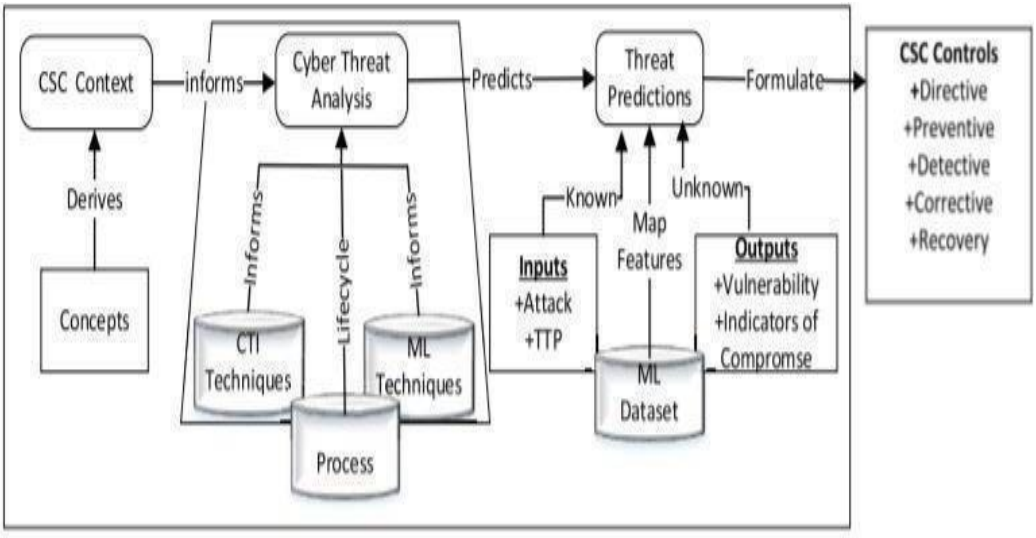


Figure 4.1: Project Architecture for Predictive Analytics For Cyber Supply Chain Security

4.2DESCRIPTION

PHASE 1: DETERMINE STRATEGY

CSC security strategy combines CTI and cybersecurity risk strategy including mechanisms, resources and plans to determine how security goals and controls will be formulated, implemented and achieved in line with organization goal and objectives. It includes identifying, analyzing, reviewing and evaluating organizational assets including infrastructures, resources and implementation procedures. CSC security strategy combines, CTI and cybersecurity risk assessment strategy to gather intelligence and formulate policies. Strategic, tactical and operational management roles and responsibilities are recursive and support each other to ensure security goals are achieved. Strategic management uses intelligence decision to support plans that determine security goals and assign responsibility including executive authorization of blueprints and budget allocation. It includes risk assessment, CSC requirements capturing and business function.

The risk strategy also considered implementation strategies and procurement policies for OT and IT acquisitions and integrations of assets.

PHASE 2: THREAT ANALYSIS

This threat analysis phase follows the CTI techniques to determine and analyse the threats of the CSC context. It requires the CSC strategy information for his purpose and includes three activities

- Identify and Gather Information:

This step identifies all vulnerable spots on the supply inbound and outbound chains on the meta-model that is used as indicators for an attack. For instance, in case of a malware attack, this activity looks for the relevant information such as the source of the attack, the tools, patterns and the attack vectors from the analysis of the malware attack that used as our indicator.

- Risk Assessments

The risk assessment activity includes the process to mitigate CSC risks by determining the probability and impact of CSC attacks and threats as well as the vulnerable spots that could be exploited within the cyber supply inbound and outbound chains and third-partyorganizations. It identifies all threats that may pose a risk on the system. Risk assesses the CSC security domain and analyse risks access spots that are capture captured. Develop mitigating techniques to control the risks by identifying risks posed by auditing the thirdparty organizations. Classify them based on their service provisions and levels of integration to the various supply chain network system.

- Analysis

This activity focuses on analysis of the threats to determine the actual source of the attack, the type of attack, the attack pattern, the TTP and attack vectors. This will assist to assign the IoC required and what controls are needed.

PHASE 3: THREAT PREDICATION

The phase considers CSC system nodes that are vulnerable to cyberattacks by integrating CTI and ML to obtain attack predictions of known and unknown attacks using three sequential activities: Determine Input Parameters, Predict Threats and Performance Evaluation

PHASE 4: CONTROL

This final phase aims to identify a list of controls that are to tackle the threat. The controls should ensure that the required security strategic and mechanism are put in place to mitigate the threats. This includes identifying security requirements, internal and external audit as well as threat monitoring and reporting. The process includes identification and review of existing controls, third-party audit and finally information sharing.

4.3 USE CASE DIAGRAM

In the use case diagram we have basically two actors who are the user and the administrator. The user has the rights to login, access to resources and to view thecrime details. Whereas the administrator has the login, access to resources of the users and also the right to update and remove the crime details, and he can also view the user files.

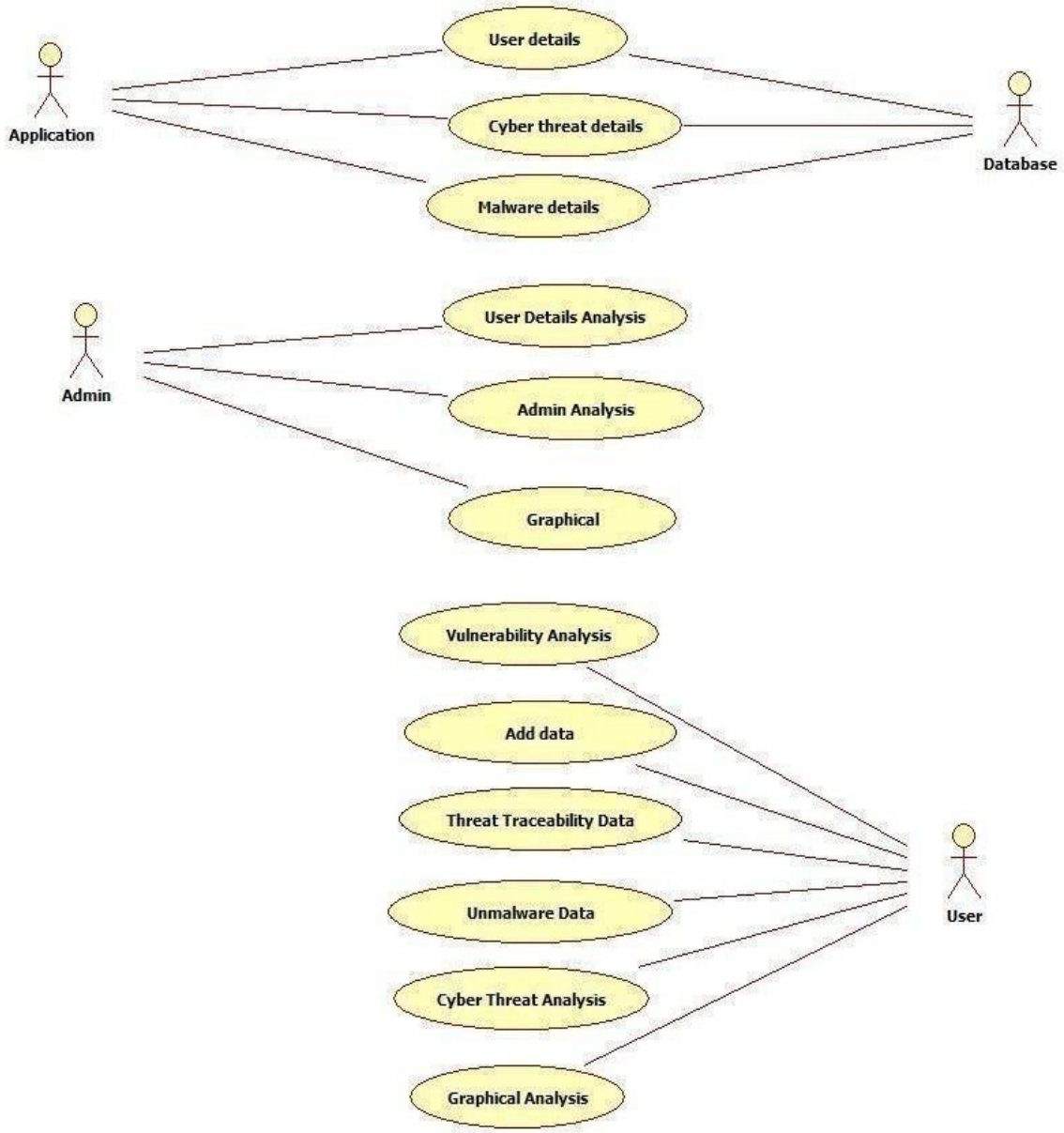


Figure4.3: Use Case Diagram for Predictive Analytics For CyberSupply Chain Security

4.4 CLASS DIAGRAM

The class diagram is used to refine the use case diagram and define a detailed design of the system. The class diagram classifies the actors defined in the use case diagram into a set of interrelated classes. The relationship or association between the classes can be either an "is-a" or "has-a" relationship. Each class in the class diagram may be capable of providing certain functionalities. These functionalities provided by the class are termed "methods" of the class. Apart from this, each class may have certain "attributes" that uniquely identify the class.

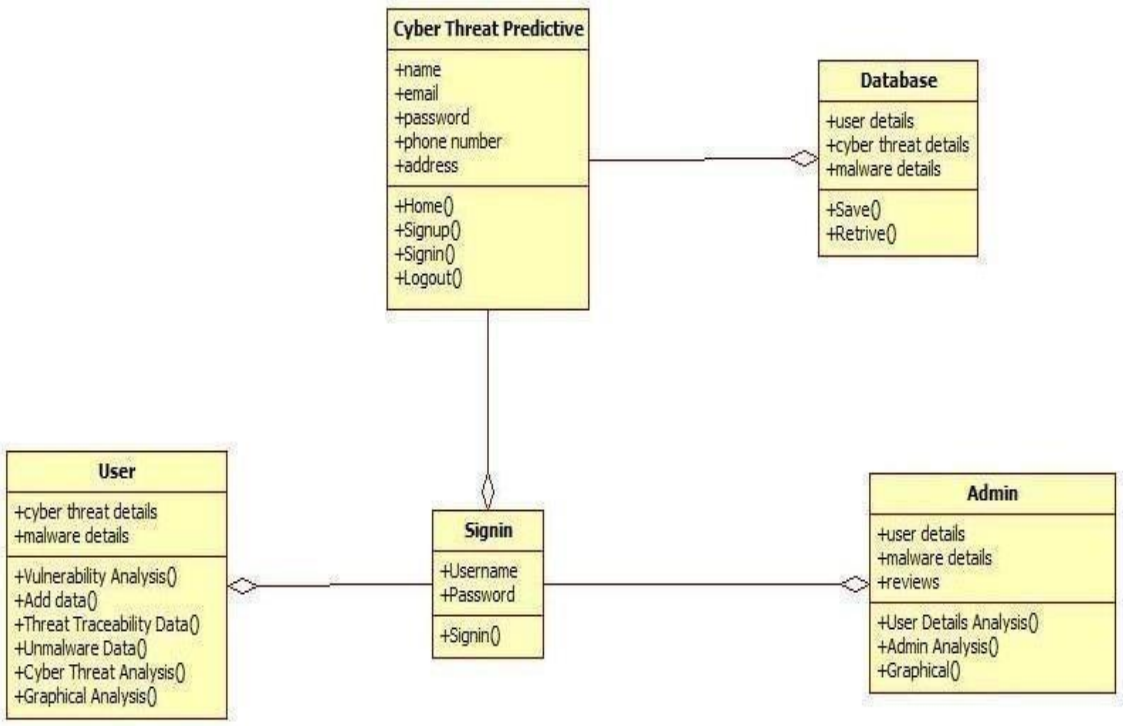


Figure 4.4: Class Diagram for Predictive Analytics For CyberSupply Chain Security

4.5 SEQUENCE DIAGRAM

A sequence diagram represents the interaction between different objects in the system. The important aspect of a sequence diagram is that it is time-ordered. This means that the exact sequence of the interactions between the objects is represented step by step. Different objects in the sequence diagram interact with each other by passing "messages".

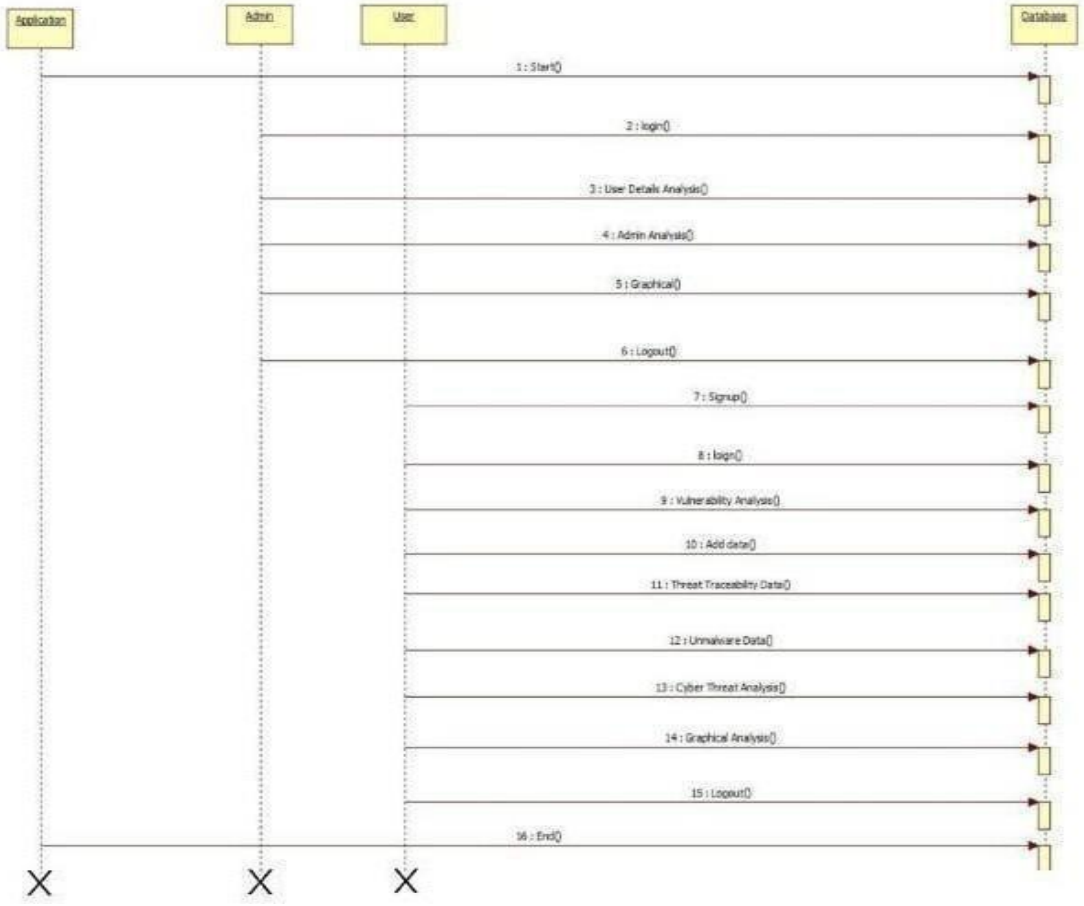


Figure 4.5: Sequence Diagram for Predictive Analytics For CyberSupply Chain Security

4.6 ACTIVITY DIAGRAM

The process flows in the system are captured in the activity diagram. Similar to a state diagram, an activity diagram also consists of activities, actions, transitions, initial and final states, and guard conditions.

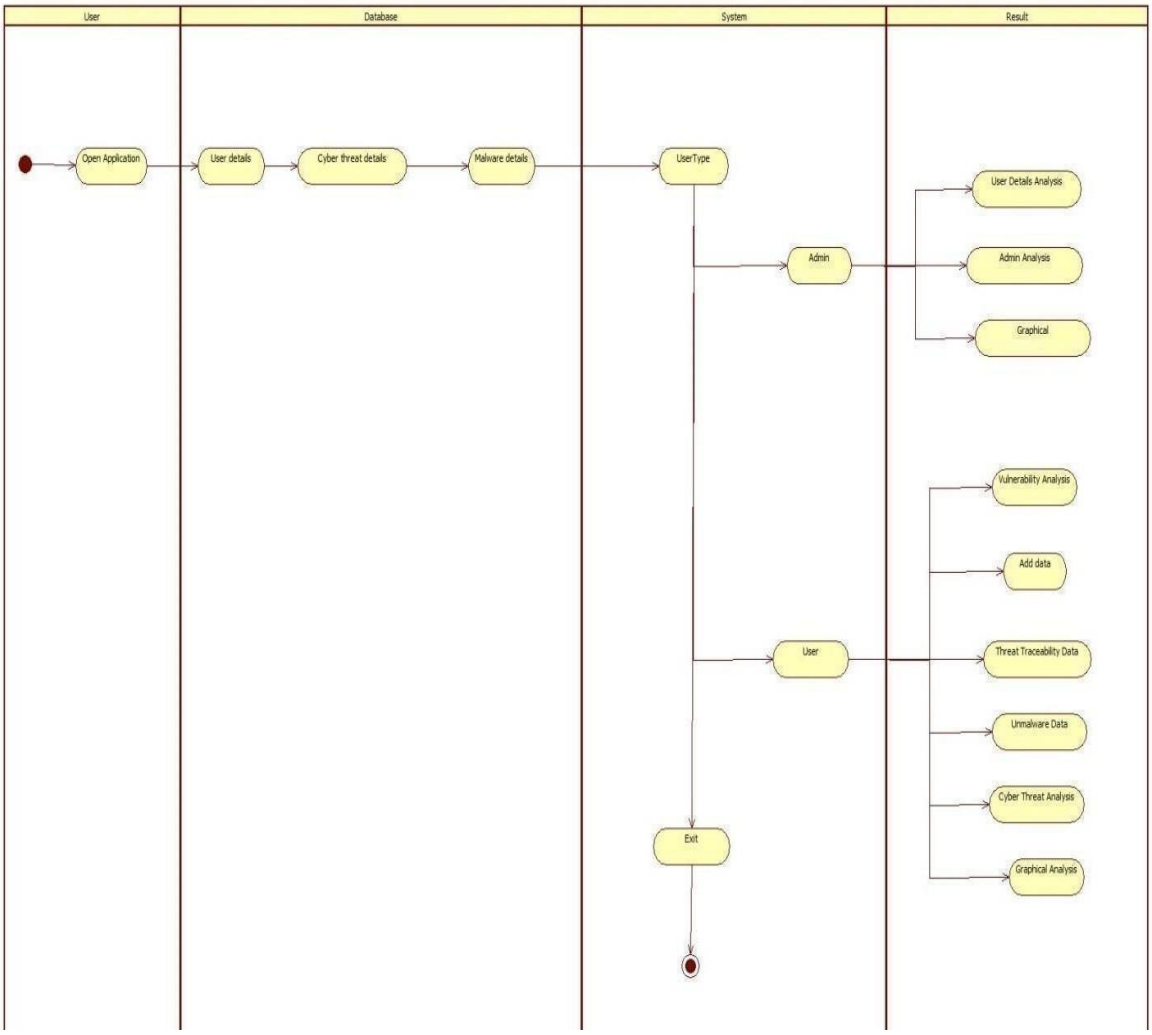


Figure 4.6: Activity Diagram for Predictive Analytics For CyberSupply Chain Security

5.IMPLEMENTATION

5.IMPLEMENTATION

5.1 SAMPLE CODE

```
#!/usr/bin/env python
import os
import sys

if__name__== "_main_":
   os.environ.setdefault("DJANGO_SETTINGS_MODULE", "Cyber_Hacking_Breaches.settings")
   try:
      from django.core.management import execute_from_command_line
   except ImportError as exc:
      raise ImportError(
         "Couldn't import Django. Are you sure it's installed and "
         "available on your PYTHONPATH environment variable? Did you "
         "forget to activate a virtual environment?"
      ) from exc
   execute_from_command_line(sys.argv)



import win32gui
import win32ui
from ctypes import windll
import Image

hwnd = win32gui.FindWindow(None, 'Calculator')

# Change the line below depending on whether you want the whole window
# or just the client area.
#left, top, right, bot = win32gui.GetClientRect(hwnd)
left, top, right, bot = win32gui.GetWindowRect(hwnd)
w = right - left
h = bot - top

hwndDC = win32gui.GetWindowDC(hwnd)
mfcDC = win32ui.CreateDCFromHandle(hwndDC)
saveDC = mfcDC.CreateCompatibleDC()
```

```
saveBitMap = win32ui.CreateBitmap()
saveBitMap.CreateCompatibleBitmap(mfcDC, w, h)

saveDC.SelectObject(saveBitMap)

# Change the line below depending on whether you want the whole window
# or just the client area.
#result = windll.user32.PrintWindow(hwnd, saveDC.GetSafeHdc(), 1)
result = windll.user32.PrintWindow(hwnd, saveDC.GetSafeHdc(), 0)
print result

bmpinfo = saveBitMap.GetInfo()
bmpstr = saveBitMap.GetBitmapBits(True)

im = Image.frombuffer(
    'RGB',
    (bmpinfo['bmWidth'], bmpinfo['bmHeight']),
    bmpstr, 'raw', 'BGRX', 0, 1)

win32gui.DeleteObject(saveBitMap.GetHandle())
saveDC.DeleteDC()
mfcDC.DeleteDC()
win32gui.ReleaseDC(hwnd, hwndDC)

if result == 1:
    #PrintWindow Succeeded
    im.save("test.png")
```

6.RESULTS

6.RESULTS

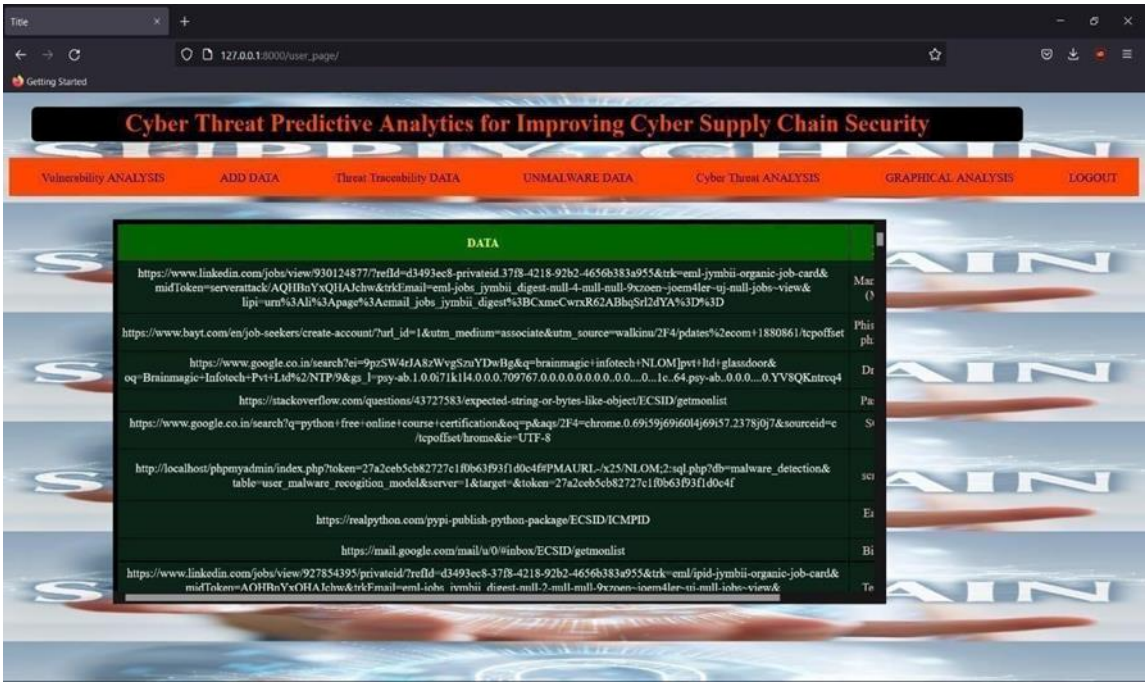
Login screen

In the below screen we can see the login screen where the user first need to sign up and after signing up the use can login with there details.

Screenshot 6.1 Login screen

Vulnerability Analysis

The below screen can be seen after the user login when the user login to the website the threat analytics can be seen which is vulnerability analysis in which the previous data regarding threats can be seen we can also find the links witch are having more threat.



Screenshot 6.2 Vulnerability Analysis

Upload Data

In the below screen the user can add any of the new threat which is found in any type of organization which will be added to the data

Screenshot 6.3 Upload Data

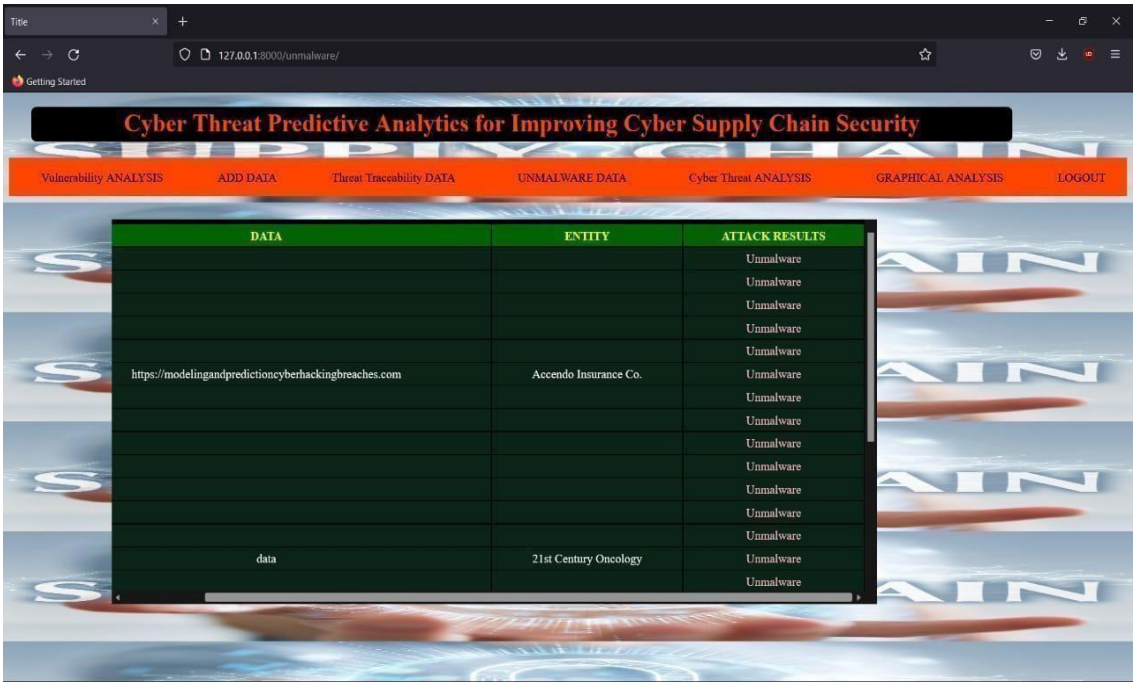
Threat Traceability Data

The below screen is the Threat Traceability Data where the user can find the entity, year, records and also the organization in which the threat is mostly is appearing

Screenshot 6.4 Threat Traceability Data

Unmalware data

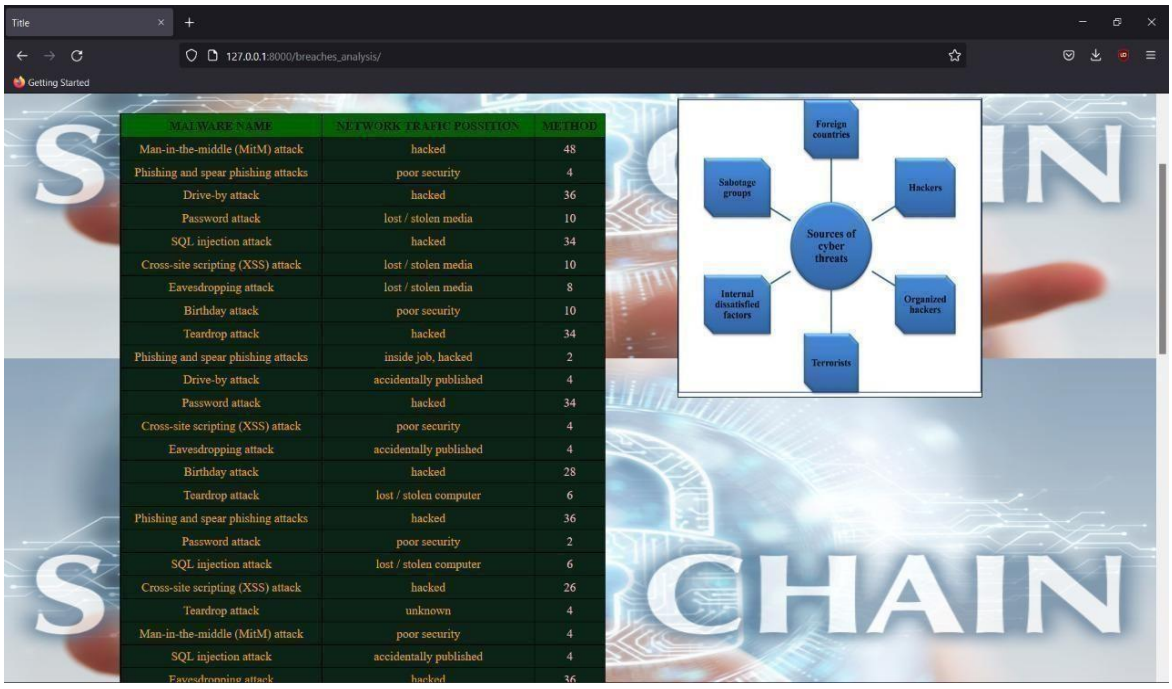
In the below screen the user can find the Unmalware data where the information is previously added by the user can be seen and also the whole data regarding the threat can be seen.



Screenshot 6.5 Unmalware data

Cyber threat analysis

The below screen is the Cyber threat analysis where the user can find the details of threat analysis and also in which type of attack is the threat and also the network trafic possition can be seen in this screen.



MALWARE NAME	NETWORK TRAFIC POSSITION	METHOD
Man-in-the-middle (MitM) attack	hacked	48
Phishing and spear phishing attacks	poor security	4
Drive-by attack	hacked	36
Password attack	lost / stolen media	10
SQL injection attack	hacked	34
Cross-site scripting (XSS) attack	lost / stolen media	10
Eavesdropping attack	lost / stolen media	8
Birthday attack	poor security	10
Teardrop attack	hacked	34
Phishing and spear phishing attacks	inside job, hacked	2
Drive-by attack	accidentally published	4
Password attack	hacked	34
Cross-site scripting (XSS) attack	poor security	4
Eavesdropping attack	accidentally published	4
Birthday attack	hacked	28
Teardrop attack	lost / stolen computer	6
Phishing and spear phishing attacks	hacked	36
Password attack	poor security	2
SQL injection attack	lost / stolen computer	6
Cross-site scripting (XSS) attack	hacked	26
Teardrop attack	unknown	4
Man-in-the-middle (MitM) attack	poor security	4
SQL injection attack	accidentally published	4
Eavesdropping attack	hacked	36

Screenshot 6.6: Cyber threat analysis

Spline Chart

In the below screen the user can find the Spline Chart in which the threat analysis can be seen in the graphical representation to understand it in a better way the graphical representation can be seen in the different forms like spline chart, bar chart, column chart.

Screenshot 6.7 Spline Chart

7.TESTING

7.1 INTRODUCTION TO TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, subassemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of tests. Each test type address the specific testing requirement.

7.1TYPES OF TESTING

7.1.1 UNIT TESTING

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

7.1.2 INTEGRATION TESTING

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfied, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

7.1.3 FUNCTIONAL TESTING

Functional tests provide systematic demonstrations that functions tested are available as system specified by the business and technical requirements, documentation, and user and manuals.

Valid Input :identified classes of valid input must be accepted.

Invalid Input :identified classes of invalid input must be rejected.

Functions :identified functions must be exercised.

Output :identified classes of application outputs must be exercised.

Systems/Procedures :interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes.

7.1.4 WHITE BOX TESTING

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

7.1.5 BLACK BOX TESTING

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot “see” into it. The test provides inputs and responds to outputs without considering how the software works.

7.1.6 ACCEPTANCE TESTING

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.
Test Results: All the test cases mentioned above passed successfully. No defects encountered.

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

- All field entries must work properly.
- Pages must be activated from the identified link.
- The entry screen, messages and responses must not be delayed.

Features to be tested

- Verify that the entries are of the correct format
- No duplicate entries should be allowed
- All links should take the user to the correct page.

7.2 TESTCASES

7.2.1 USER REQUIREMENTS

1. Home

Use case ID	Cyber Threat Predictive Analytics for Improving Cyber Supply Chain Security
Use case Name	Home button
Description	Display home page of application
Primary actor	User
Precondition	User must open application
Post condition	Display the Home Page of an application
Frequency of Use case	Many times
Alternative use case	N/A
Use case Diagrams	
Attachments	N/A

8.CONCLUSION

8.CONCLUSION & FUTURE SCOPE

8.1 PROJECT 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 analyse 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.

8.2FUTURE SCOPE

The project can be further developed into the real time application for which can be applied in the organizations on the networks. The, organizational implementation plays an important role in the finding of the infected systems and to show the probable infected systems for which can be vulnerable.

9.BIBILIOGRAPHY

9.BIBILIOGRAPHY

9.1 REFERENCES

1. National Cyber Security Centre. (2018). Example of Supply Chain Attacks.
2. A. Yeboah-Ofori and S. Islam, ‘‘Cyber security threat modelling for supply chain organizational environments,’’ MDPI. Future Internet, vol. 11, no. 3, p. 63, Mar. 2019.
3. B. Woods and A. Bochman, ‘‘Supply chain in the software era,’’ in Scowcroft Center for Strategic and Security. Washington, DC, USA: Atlantic Council, May 2018.
4. Exploring the Opportunities and Limitations of Current Threat Intelligence Platforms, Version 1, ENISA, Dec. 2017.
5. C. Doerr, TU Delft CTI Labs. (2018). Cyber Threat Intelligences Standards—A High Level Overview.
6. Research Prediction. (2019). Microsoft Malware Prediction.
7. A. Yeboah-Ofori and F. Katsriku, ‘‘Cybercrime and risks for cyber physical systems,’’ Int. J. Cyber-Secur. Digit. Forensics, vol. 8, no. 1, pp. 43–57, 2019.
8. CAPEC-437, Supply Chain. (Oct. 2018). Common Attack Pattern Enumeration and Classification: Domain of Attack.
9. Open Web Application Security Project (OWASP). (2017). The Ten Most Critical Application Security Risks, Creative Commons Attribution-Share Alike 4.0 International License.
10. US-Cert. (2020). Building Security in Software & Supply Chain Assurance.

9.2 GITHUB LINK

https://github.com/Gohathi/predictive-analytics-for-cyber-threats-to-improve-cyber-supply-chain-security

10. PAPER PUBLICATION

PREDICTIVE ANALYTICS FOR CYBER THREATS TO IMPROVE CYBER SUPPLY CHAIN SECURITY

[1] Mudimela Madhusudhan, [2] Kadari Gohathi ,[3] Kathi Venkata Saketh Reddy, [4] Kosamba Pradeep
[1] Assistant Professor of Dept of CSE,CMR Technical Campus,Medchal, Hyderabad
[2,3,4] B.Tech Student,Dept of CSE, CMR Technical Campus, Hyderabad
[1]madhusudhan.cse@gmail.com,
kadarigohathi@gmail.com,[3]sakethkathi147@gmail.com,[4]pradeepkosamba@gmail.com

Abstract: In this paper, the aim is 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. 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.

Words — Cyber Threat Intelligence, Cyber supply chain, Cyber threat intelligence

I. INTRODUCTION

Cyber Supply Chain (CSC) security is critical for reliable service delivery and ensure overall business continuity of Smart CPS. CSC systems by its inherently is complex and vulnerabilities within CSC system environment can cascade from a source node to a number of target nodes of the overall cyber physical system (CPS). There are several examples for successful CSC attacks. For instance, Dragonfly, a Cyber Espionage group, is well known for targeting CSC organization.

Due to the invincibility nature of cyber attacks on the cyber supply chain (CSC), and the cascading effects of malware infections, we use machine learning to predict attacks. As organizations have become more reliant on CSC systems for business continuity, so are the increase in vulnerabilities and the threat landscapes. Some traditional approach to detecting and defending malware attack has largely been antimalware or antivirus software such as spam filters, firewall, and IDS/IPS. These tools largely succeed, however, as threat actors get more intelligent, they are able to circumvent and affect nodes on systems which then propagates. we use ML techniques to learn the dataset and predict which CSC nodes have detection or no detection. The purpose is to predict which modes are venerable to cyberattacks and for

predicting future trends. To demonstrate the applicability of our approach, we used a dataset from the Microsoft Malware Prediction website. Further, an ensemble is used to link Logistic Regression, and Decision Tree and SVM algorithms in Majority Voting and run on the training data and then use 10-fold cross validation to test the parameter estimation, accurate results and predictions. The results show that ML algorithms in Decision Trees methods can be used in cyber supply chain predict analytics to detect and predict future cyber attack trends.

II. LITERATURE SURVEY

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.

Towards the Prediction of Renewable Energy Unbalance in Smart Grids

AUTHORS: R. D. Labati, A. Genovese, V. Piuri and F. Scotti

ABSTRACT: The production of renewable energy is increasing worldwide. To integrate renewable sources in electrical smart grids able to adapt to changes in power usage in heterogeneous local zones, it is necessary to accurately predict the power production that can be achieved from renewable energy sources. By using such predictions, it is possible to plan the power production from non-renewable energy plants to properly allocate the produced power and compensate possible unbalances. In particular, it is important to predict the unbalance between the power produced and the actual power intake at a local level (zones). In this paper, we propose a novel method for predicting the sign of the unbalance between the power produced by renewable sources and the power intake at the local level, considering zones composed of multiple power plants and with heterogeneous characteristics. The method uses a set of historical features and is based on Computational Intelligence techniques able to learn the relationship between historical data and the power unbalance in heterogeneous geographical regions. As a case study, we evaluated the proposed method using data collected by a player in the energy market over a period of seven months. In this preliminary study, we evaluated different configurations of the proposed method, achieving results considered as satisfactory by a player in the energy market.

Malware Attack Predictive Analytics in a Cyber Supply Chain Context Using Machine Learning

AUTHORS: A. Yeboah-Ofori and C. Boachie

ABSTRACT: Due to the invincibility nature of cyber attacks on the cyber supply chain (CSC), and the cascading effects of malware infections, we use machine learning to predict attacks. As organizations have become more reliant on CSC systems for business continuity, so are the increase in vulnerabilities and the threat landscapes. Some traditional approach to detecting and defending malware attack has largely been antimalware or antivirus software such as spam filters, firewall, and IDS/IPS. These tools largely succeed, however, as threat actors get more intelligent, they are able to circumvent and affect nodes on systems which then propagates. In our previous work, we characterized threat actor activities, including presumed intent and historically observed behaviour, for the purpose of ascertaining the current threats that could be exploited. In this paper, we use ML techniques to learn the dataset and predict which CSC nodes have detection or no detection. The purpose is to predict which modes are venerable to cyberattacks and for predicting future trends. To demonstrate the applicability of our approach, we used a dataset from the Microsoft Malware Prediction website. Further, an ensemble is used to link Logistic Regression, and Decision Tree and SVM algorithms in Majority Voting and run on the training data and then use 10-fold cross validation to test the parameter estimation, accurate results and predictions. The results show that ML algorithms in Decision Trees methods can be used in cyber supply chain predict analytics to detect and predict future cyber attack trends.

Feasibility of Supervised Machine Learning for Cloud Security

AUTHORS: D. Bhamare, T. Salman, M. Samaka, A. Erba and R. Jain

ABSTRACT: Cloud computing is gaining significant attention, however, security is the biggest hurdle in its wide acceptance. Users of cloud services are under constant fear of data loss, security threats and availability issues. Recently, learning-based methods for security applications are gaining popularity in the literature with the advents in machine learning techniques. However, the major challenge in these methods is obtaining real-time and unbiased datasets. Many datasets are internal and cannot be shared due to privacy issues or may lack certain statistical characteristics. As a result of this, researchers prefer to generate

datasets for training and testing purpose

in the simulated or closed experimental environments which may lack comprehensiveness. Machine learning models trained with such a single dataset generally result in a semantic gap between results and their application. There is a dearth of research work which demonstrates the effectiveness of these models across multiple datasets obtained in different environments. We argue that it is necessary to test the robustness of the machine learning models, especially in diversified operating conditions, which are prevalent in cloud scenarios. In this work, we use the UNSW dataset to train the supervised machine learning models. We then test these models with ISOT dataset. We present our results and argue that more research in the field of machine learning is still required for its applicability to the cloud security.

A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection

AUTHORS: A. L. Buczak, and E. Guven.

ABSTRACT: This survey paper describes a focused literature survey of machine learning (ML) and data mining (DM) methods for cyber analytics in support of intrusion detection. Short tutorial descriptions of each ML/DM method are provided. Based on the number of citations or the relevance of an emerging method, papers representing each method were identified, read, and summarized. Because data are so important in ML/DM approaches, some well-known cyber data sets used in ML/DM are described. The complexity of ML/DM algorithms is addressed, discussion of challenges for using ML/DM for cyber security is presented, and some recommendations on when to use a given method are provided.

A Review of Cyber Security Dataset for Machine Learning Algorithms

AUTHORS: O. Yavanoglu and M. Aydos

ABSTRACT: It is an undeniable fact that currently information is a pretty significant presence for all companies or organizations. Therefore protecting its security is crucial and the security models driven by real datasets has become quite important. The operations based on military, government, commercial and civilians are linked to the security and availability of computer systems and network. The objective of this review is to explain and compare the most commonly used datasets. This paper focuses on the datasets used in artificial intelligent and machine learning techniques, which are the primary tools for analyzing network traffic and detecting abnormalities.

III. PROPOSED METHODOLOGY

CSC security strategy combines CTI and cybersecurity risk strategy including mechanisms, resources and plans to determine how security goals and controls will be formulated, implemented and achieved in line with organization goal and objectives. It includes identifying, analyzing, reviewing and evaluating organizational assets including infrastructures, resources and implementation procedures. CSC security strategy combines, CTI and cybersecurity risk assessment strategy to gather intelligence and formulate policies. Strategic, tactical and operational management roles and responsibilities are recursive and support each other to ensure security goals are achieved. Strategic management uses intelligence decision to support plans that determine security goals and assign responsibility including executive authorization of blueprints and budget allocation. It includes risk assessment, CSC requirements capturing and business function.

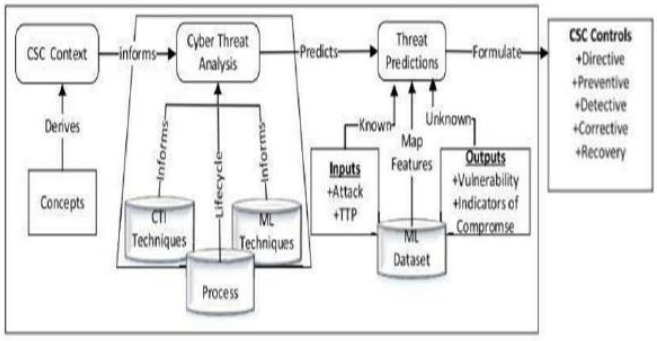


Fig 3.1 Project functionality

IV. ELEMENTS REQUIRED

Cyber Threat Intelligence (CTI)

Cyber Threat Intelligence 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).

Cyber Supply Chain (CSC)

Cyber Supply Chain 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 securit

V. WORKING PROCEDURE OF THE MODEL

Identify and Gather Information:
This step identifies all vulnerable spots on the supply inbound and outbound chains on the meta-model that is used as indicators for an attack. For instance, in case of a malware attack, this activity looks for the relevant information such as the source of the attack, the tools, patterns and the attack vectors from the analysis of the malware attack that used as our indicator.

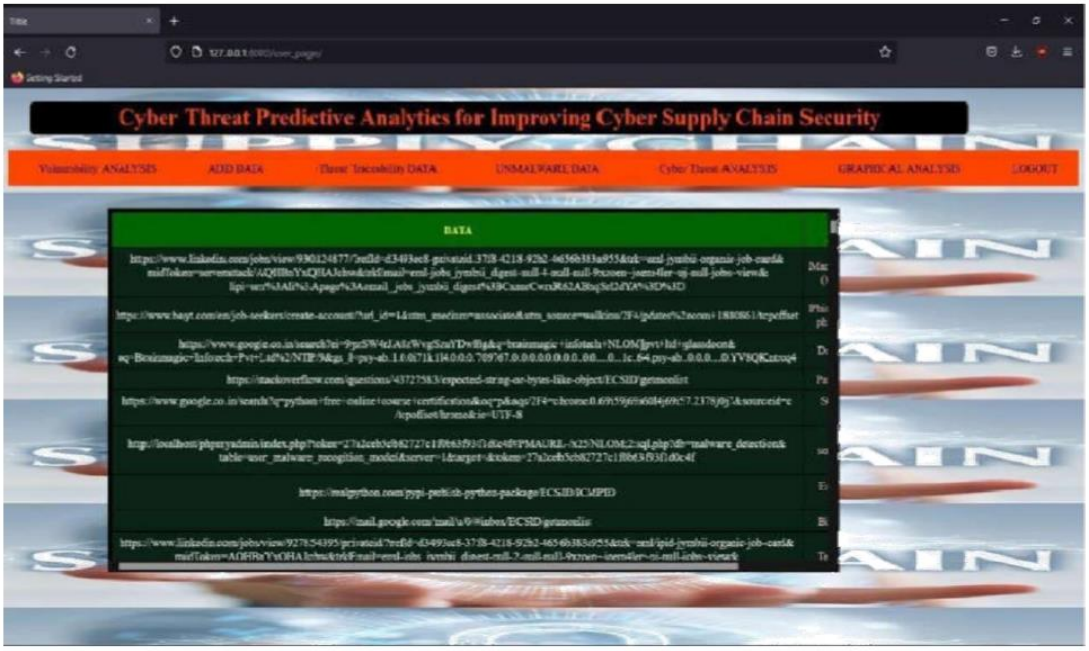
Risk Assessments
The risk assessment activity includes the process to mitigate CSC risks by determining the probability and impact of CSC attacks and threats as well as the vulnerable spots that could be exploited within the cyber supply inbound and outbound chains and third-party organizations. It identifies all threats that may pose a risk on the system. Risk assesses the CSC security domain and analyse risks access spots that are capture captured. Develop mitigating techniques to control the risks by identifying risks posed by auditing the thirdparty organizations. Classify them based on their service provisions and levels of integration to the various supply chain network system.

Analysis
This activity focuses on analysis of the threats to determine the actual source of the attack, the type of attack, the attack pattern, the TTP and attack vectors. This will assist to assign the IoC required and what controls

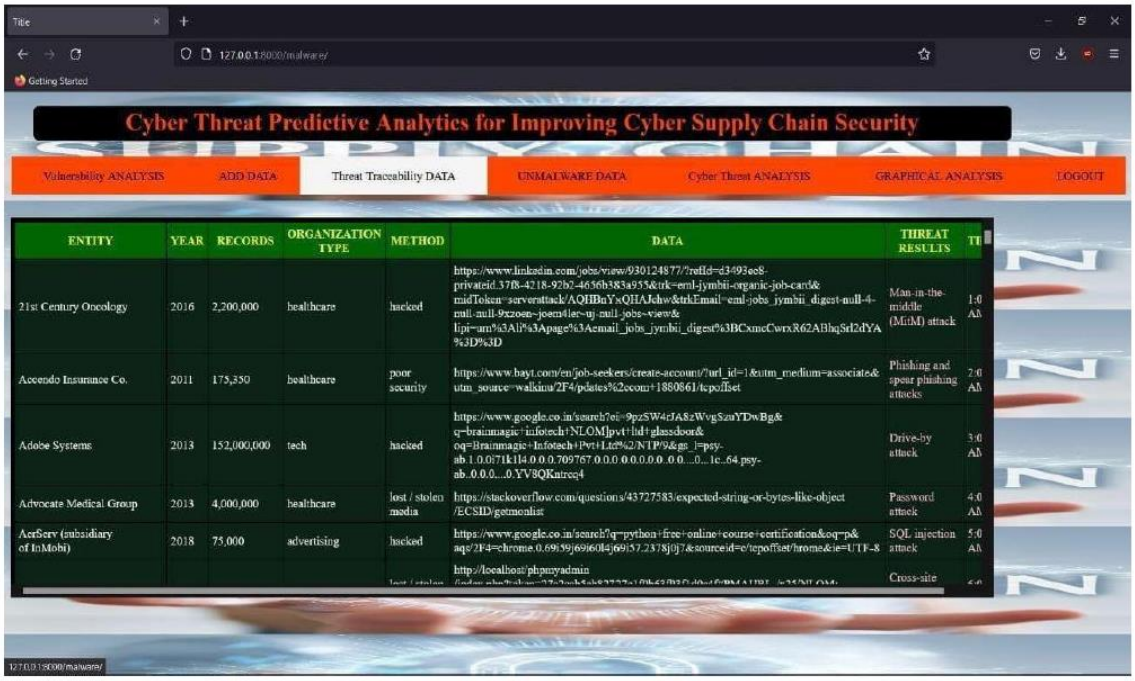
VI.RESULTS

Login screen

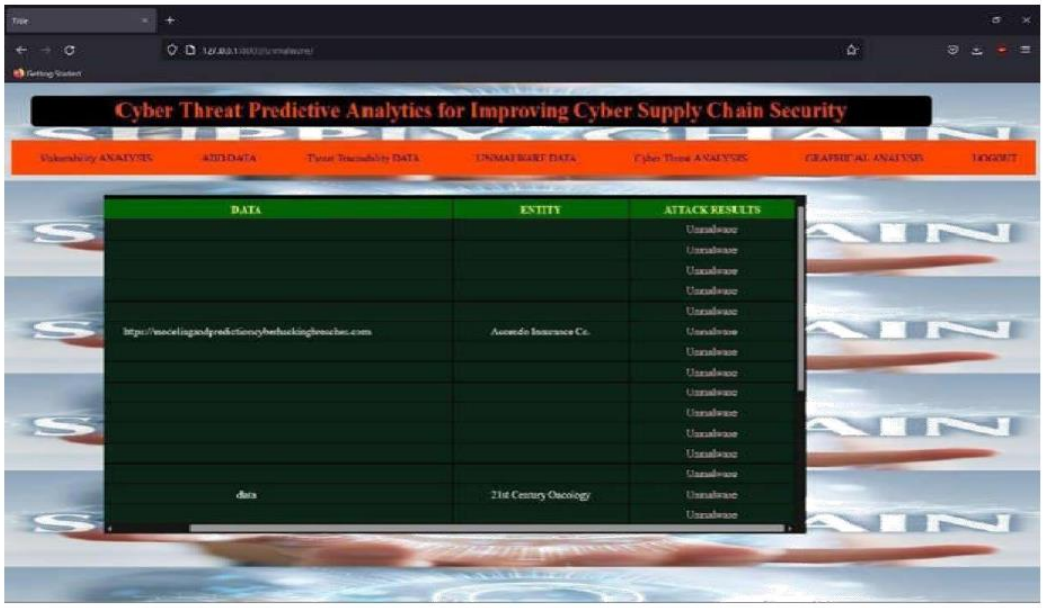
Upload Data



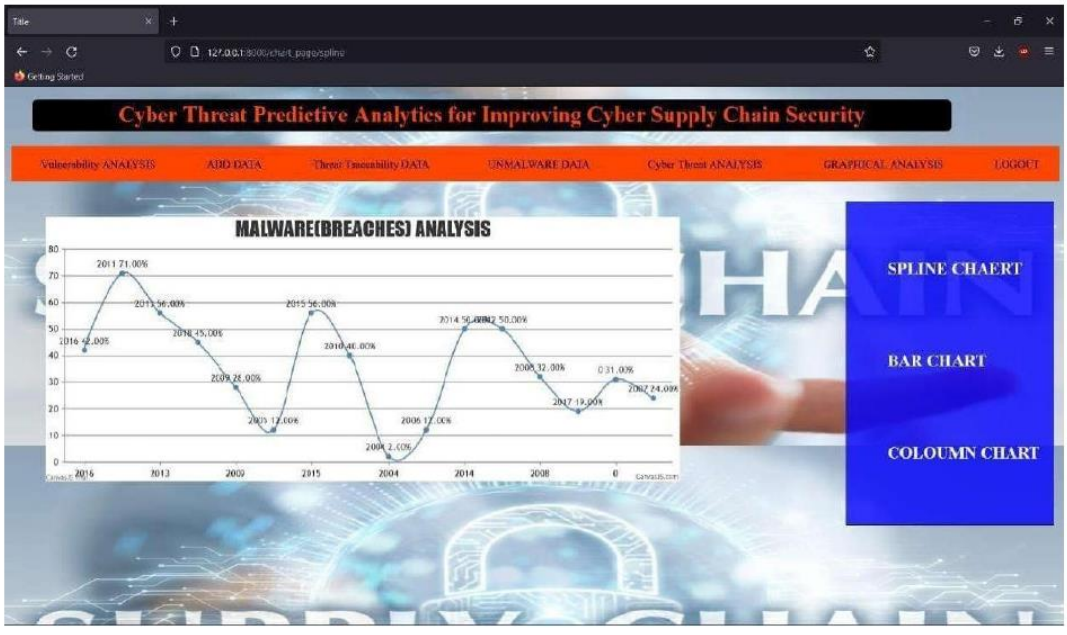
Vulnerability Analysis



Threat Traceability Data



Unmalware data



Spline Chart

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 analyse 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

FUTURE SCOPE

The project can be further developed into the real time application for which can be applied in the organizations on the networks. The, organizational implementation plays an important role in the finding of the infected systems and to show the probable infected systems for which can be vulnerable.

REFERENCES

1. National Cyber Security Centre. (2018). Example of Supply Chain Attacks.
2. A. Yeboah-Ofori and S. Islam, ‘‘Cyber security threat modelling for supply chain organizational environments,’’ MDPI. Future Internet, vol. 11, no. 3, p. 63, Mar. 2019.
3. B. Woods and A. Bochman, ‘‘Supply chain in the software era,’’ in Scowcroft Center for Strategic and Security. Washington, DC, USA: Atlantic Council, May 2018.
4. Exploring the Opportunities and Limitations of Current Threat Intelligence Platforms, Version 1, ENISA, Dec. 2017.
5. C. Doerr, TU Delft CTI Labs. (2018). Cyber Threat Intelligences Standards—A High Level Overview.
6. Research Prediction. (2019). Microsoft Malware Prediction.
7. A.Yeboah-Ofori and F. Katsriku, ‘‘Cybercrime and risks for cyber physical systems,’’ Int. J.Cyber-Secur. Digit. Forensics, vol. 8, no. 1, pp. 43–57, 2019.
8. CAPEC-437, Supply Chain. (Oct. 2018). Common Attack Pattern Enumeration and Classification: Domain of Attack.
9. Open Web Application Security Project (OWASP). (2017). The Ten Most Critical Application Security Risks, Creative Commons Attribution-Share Alike 4.0 International License.
10. US-Cert. (2020). Building Security in Software & Supply Chain Assurance

11.CERTIFICATION

JAC : A JOURNAL OF COMPOSITION THEORY

ISSN NO: 0731-6755 / Impact Factor - 5.7
web : www.jctjournal.com, e-mail : jctjournal@gmail.com

Certificate of Publication

This is to certify that the paper entitled

"PREDICTIVE ANALYTICS FOR CYBER THREATS TO IMPROVE CYBER SUPPLY CHAIN SECURITY"

Authored by

Kadari Gohathi

from

CMR Technical Campus,Medchal, Hyderabad.

Has been published in

JCT JOURNAL IN VOLUME XVI ISSUE III MARCH 2023

SEBASTIAN ZEKI
EDITOR IN CHIEF
JCT JOURNAL

JAC : A JOURNAL OF COMPOSITION THEORY

ISSN NO: 0731-6755 / Impact Factor - 5.7
web : www.jctjournal.com, e-mail : jctjournal@gmail.com

Certificate of Publication

This is to certify that the paper entitled

"PREDICTIVE ANALYTICS FOR CYBER THREATS TO IMPROVE CYBER SUPPLY CHAIN SECURITY"

Authored by

Kathi Venkata Saketh Reddy

from

CMR Technical Campus,Medchal, Hyderabad.

Has been published in

JCT JOURNAL IN VOLUME XVI ISSUE III MARCH 2023

SEBASTIAN ZEKI
EDITOR IN CHIEF
JCT JOURNAL

JAC : A JOURNAL OF COMPOSITION THEORY

ISSN NO: 0731-6755 / Impact Factor - 5.7
web : www.jctjournal.com, e-mail : jctjournal@gmail.com

Certificate of Publication

This is to certify that the paper entitled

"PREDICTIVE ANALYTICS FOR CYBER THREATS TO IMPROVE CYBER SUPPLY CHAIN SECURITY"

Authored by

Kosamba Pradeep

from

CMR Technical Campus,Medchal, Hyderabad.

Has been published in

JCT JOURNAL IN VOLUME XVI ISSUE III MARCH 2023

SEBASTIAN ZEKI
EDITOR IN CHIEF
JCT JOURNAL

JAC : A JOURNAL OF COMPOSITION THEORY

ISSN NO: 0731-6755 / Impact Factor - 5.7
web : www.jctjournal.com, e-mail : jctjournal@gmail.com

Certificate of Publication

This is to certify that the paper entitled

"PREDICTIVE ANALYTICS FOR CYBER THREATS TO IMPROVE CYBER SUPPLY CHAIN SECURITY"

Authored by

Mudimela Madhusudhan, Assistant Professor

from

CMR Technical Campus,Medchal, Hyderabad.

Has been published in

JCT JOURNAL IN VOLUME XVI ISSUE III MARCH 2023

SEBASTIAN ZEKI
EDITOR IN CHIEF
JCT JOURNAL