A DATA ANALYTICS APPROACH TO THE CYBERCRIME UNDERGROUND ECONOMY

Dr.M.Aravind kumar Professor, Department of CSE, West Godavari Institute of Science and Engineering,Affliated to jntuk, AndhraPradesh,India. drmaravindkumar@gmail.com
P. Sheela , Associate Professor, Department of CSE, West Godavari Institute of Science and Engineering,Affliated to jntuk, AndhraPradesh,India. sheela.pitta@gmail.com
M.Chilakaiah, Associate Professor, Department of CSE, West Godavari Institute of Science and Engineering, Affliated to jntuk, AndhraPradesh,India. m.chilakaiah@gmail.com
A.Mohan Manindranath, Assistant Professor, Department of CSE, West Godavari Institute of Science and Engineering,Affliated to jntuk, AndhraPradesh,India.aretimohan3@gmail.com
J.Bullemma, Student of CSE Department, 19PD1A0515,West Godavari Institute of Science and Engineering,Affliated to jntuk, AndhraPradesh,India. jeediguntabullemma166@gmail.com
V.Jyothsna, Student of CSE Department, 19PD1A0543, West Godavari Institute of Science and Engineering,Affliated to jntuk, AndhraPradesh,India. vagicherlajyothsna@gmail.com

Abstract

Despite the rapid escalation of cyber threats, there has still been little research into thefoundationsofthesubjectormethodologiesthatcouldservetoguideinformationsystemsresearchersand practitionerswhodealwithcybersecurity.Inaddition,littleisknownaboutcrime-as-a-service(CaaS),a criminal business model that underpins the cybercrime underground. This research gap and the practicalcybercrime problems we face have motivated us to investigate the cybercrime underground economy by taking adata analytic sapproach from a design science perspective. To achieve this goal, we:(1)proposea data analysis framework for analyzing the cybercrime underground; (2) propose CaaS and crimewaredefinitions; (3) propose an associated classification model, and (4) develop an example application to demonstrate how the proposed frame work and classification modelcouldbeimplementedinpractice.Wethenusethisapplicationtoinvestigatethecybercrimeundergroundeconomybyanalyzingalargedataset obtained from the online hacking community. By taking a design science research approach, this paper contributes to the designartifacts, foundations, andmethodologiesinthisarea.Moreover,itprovidesusefulpractical insights to practitioners by suggesting guidelines as to how governments and organizations in allindustries can prepare for attacks bythecybercrimeunderground
.

Keywords: Crime ware Service, crimeware, underground economy, hacking community, machine learning, design science research

1. INTRODUCTION

As the threat posed by massive cyber attacks (e.g.,ran-somware and distributed denial of service attacks (DDoS))and cybercrimes has grown, individuals, organizations, and governments have struggled to find ways to defend against them. In 2017, ransomware known as Wanna Crywasresponsible for nearly 45,000 attacks in almost 100 countries [1].The explosive impact of cybercrime has put governments under pressure to increase their cyber security budgets.

United States President Barack Obamaproposedspendingover $19billiononcybersecurityaspartofhisfiscalyear2017budget,anincreaseofmorethan35%since2016[2].Global cyberattacks (such as WannaCry and Petya) areexecuted by highly organized criminal groups, and organizedornational-levelcrimegroupshavebeenbehindmanyrecentattacks. Typically, criminal groups buyandsellhack-ingtoolsandservicesonthecybercrimeblackmarket,wherein attackers share a range of hacking-related information. This online under groundmarket is operated by groups of attackers, anditinturnsupportstheunderground cybercrime economy [3]. The cybercrime underground hasthus emerged as a new type of organization that both operates black markets and enables cybercrime conspiracies to flourish. Because organize dcyber crimere quiresan online network to existand toconductitsattacks,itishighlydependenton closed underground communities (e.g., Hackforums and Crackingzilla).The anonym it these closed groups offer means that cybercrime networks are structured differently than traditional Mafia-style hierarchies [4], which are vertical, concentrated, rigid, and fixed.

In contrast, cyber crime net work sare lateral, diffuse, fluid, and evolving. Since cyber space is a network of networks[5], the threat posedby the rise of highly professional network-based cybercrimebusiness models, such as Crimeware-as-a-Service (CaaS),remains mostly invisible to governments, organizations, andindividuals.Even though Information Systems (IS) researchers and practitioners are taking an increasing interest in cybercrime, due to the critical issue sarising from therapid increasein cyberthreats, few have attempt edtoputthisnewinterestonasolidfoundationor developsuitablemethodologies.Previousstudies have not analyzed the underground economy behindcybercrime in depth. Furthermore, little is known about CaaS,one of the primary business models behind the cybercrimeunderground.Thereisanoveralllackofunderstanding,bothin research and practice, of the nature of this underground andthemechanismsunderlyingit.

This study takes a design science research (DSR) approach. Design science ‘‘creates and evaluates information technology artifacts intended to solve identified problems’’ [6].

2. RELATED WORK

Today, political and commercial entities are increasingly engaging In sophisticated cyber-warfare to damage, disrupt, or censor information content in computer networks. In designing network protocols, there is a need to ensure reliability against intrusions of powerful attackers that can even control a fraction of parties in the network.

The controlled parties can launch both passive (e.g., eavesdropping, nonparticipation) and active attacks (e.g., jamming, message dropping, corruption, and forging). Intrusion detection is the process of dynamically monitoring events occurring in a computer system or network, analysing them for signs of possible incidents and often interdicting the unauthorized access.Previous work done using data mining techniques

3. PROPOSED FRAMEWORK

Our data analysis framework's objective is to perform a big-picture examination of the cybercrime underground by encompassing all aspects of data analysis from start to finish. This structure is made up of four steps: (1) setting goals; (2) identifying sources; (3) deciding on analytical techniques; and (4) putting the application into action.

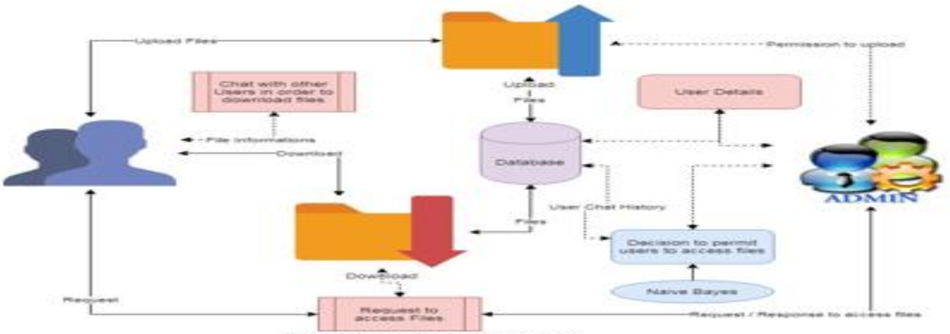


Fig1: System Architecture

Fig1.System Architecture

Step 1: Defining Goals

The first step is to identify the conceptual scope of the analysis. Specifically, this step identifies the analysis context, namely the objectives and goals. To gain an in-depth understanding of the current CaaS research, we investigated the cybercrime underground, which operates as a closed community. Thus, the goal of the proposed framework is to "investigate the cybercrimeunderground economy."

Step 2: Identifying Sources

The second step is to identify the data sources, based on the goals defined by Step 1. This step should consider what data is needed and where it can be obtained. Since the goal of this study is to investigate the cybercrime underground, we consider data on the cybercrime underground community. We therefore collected such data from the community itself and obtained a malware database from a leading global cyber security research firm. Because cybercriminals often change their IP addresses and use anti-crawling scripts to conceal their communications, we used a self-developed crawler that can resolve catches and anti- crawling scripts to gather the necessary data.

Step3: Selecting analytical methods

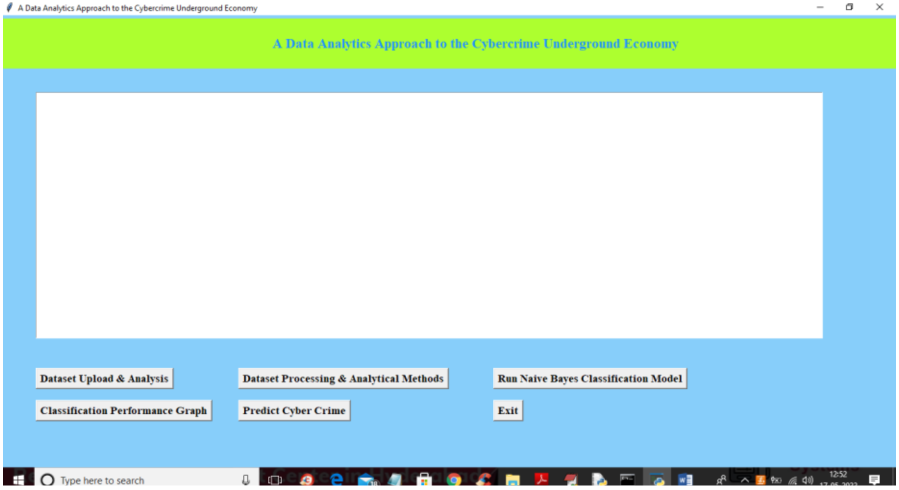
A diverse range of items are sold in the cybercrime underground, with different degrees of associated risk. For this study, we focused mainly on items critical to hacking. We first filtered the messages to select only those that carried significant risks

Step4: Implementing an application

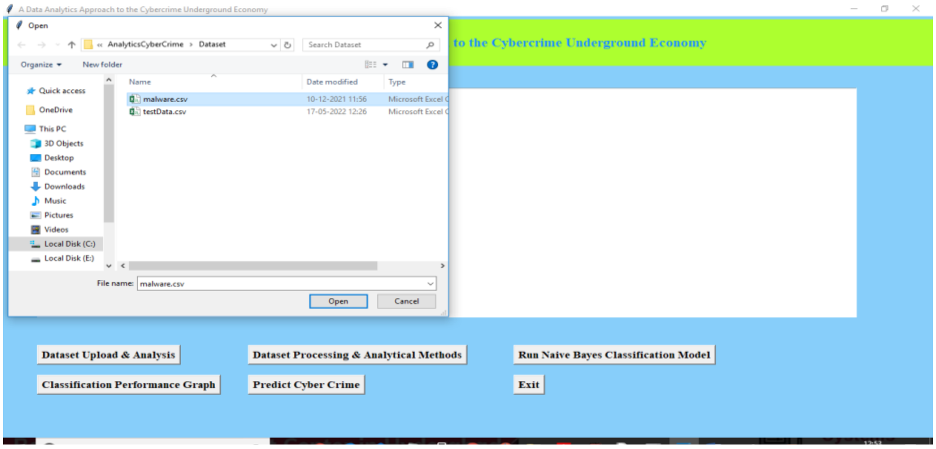
Although organizations emphasize the measures they take to prevent cybercrime, their overall effectiveness has yet to be empirically demonstrated in practice. In the last step of our framework, we demonstrate the use of the proposed CaaS and crimeware definitions, classification model, and analysis framework.

4. RESULTS AND DISCUSSION

Below is the dataset screen used in this project

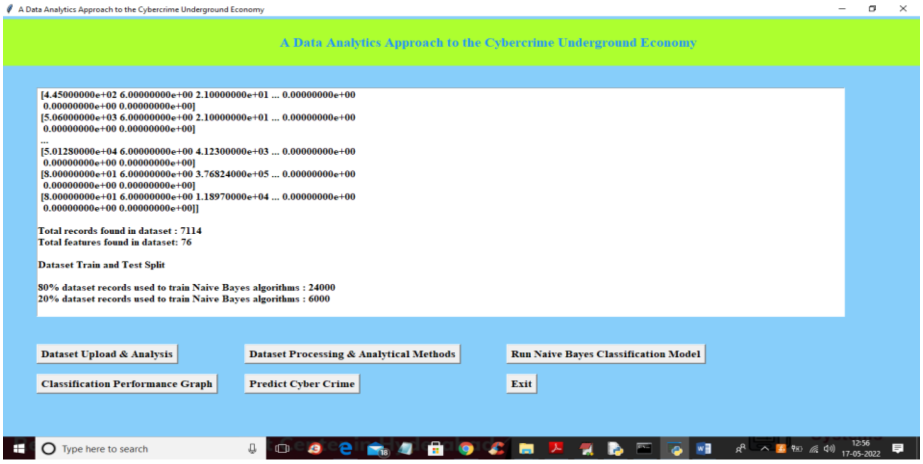

A Data Analytics Approach to the Cybercrime Underground Economy
A Data Analytics Approach to the Cybercrime Underground Economy
Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Run Naive Bayes Classification Model
Classification Performance Graph
Predict Cyber Crime
Exit
Type here to search
12:52

In above screen click on 'Dataset Upload & Analysis' button to upload dataset and get below Output.



In above screen selecting and uploading 'malware.csv' file and then click on 'Open' button to load dataset and get below analysis output

In below screen we can see all data is converted to numeric format and then we can see total dataset and then we can see 80% dataset used for training and 20 for testing and now dataset is ready and now click on 'Run Naive Bayes Classification Model' button to train classification model and get below output.

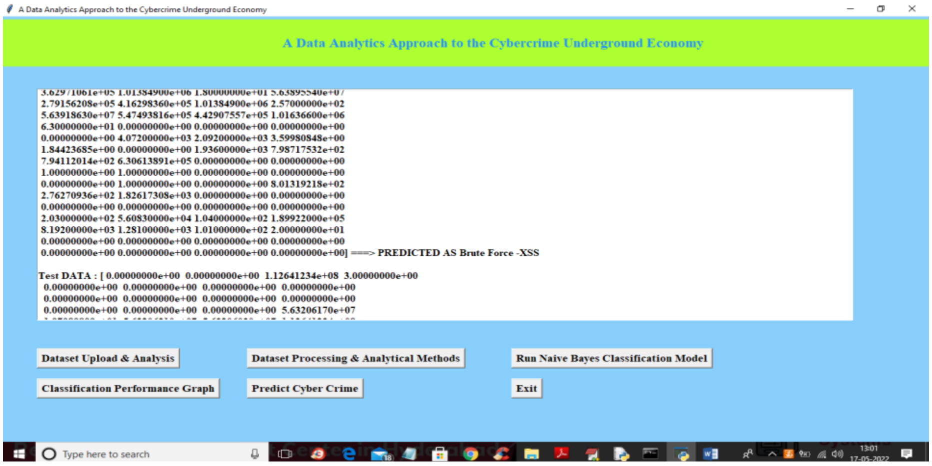

A Data Analytics Approach to the Cybercrime Underground Economy
A Data Analytics Approach to the Cybercrime Underground Economy
Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Run Naive Bayes Classification Model
Classification Performance Graph
Predict Cyber Crime
Exit
Type here to search
12:56
17-05-2022

```
[4.45000000e+02 6.00000000e+00 2.10000000e+01 ... 0.00000000e+00
 0.00000000e+00 0.00000000e+00]
[5.06000000e+03 6.00000000e+00 2.10000000e+01 ... 0.00000000e+00
 0.00000000e+00 0.00000000e+00]
...
[5.01280000e+04 6.00000000e+00 4.12300000e+03 ... 0.00000000e+00
 0.00000000e+00 0.00000000e+00]
[8.00000000e+01 6.00000000e+00 3.76824000e+05 ... 0.00000000e+00
 0.00000000e+00 0.00000000e+00]
[8.00000000e+01 6.00000000e+00 1.18970000e+04 ... 0.00000000e+00
 0.00000000e+00 0.00000000e+00]]

Total records found in dataset : 7114
Total features found in dataset: 76

Dataset Train and Test Split

80% dataset records used to train Naive Bayes algorithms : 24000
20% dataset records used to train Naive Bayes algorithms : 6000
```


A Data Analytics Approach to the Cybercrime Underground Economy
A Data Analytics Approach to the Cybercrime Underground Economy
Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Run Naive Bayes Classification Model
Classification Performance Graph
Predict Cyber Crime
Exit
Type here to search
13:01
17-05-2022

```
3.62971061e+05 1.01384900e+06 1.80000000e+01 5.63895540e+07
2.79156208e+05 4.16298360e+05 1.01384900e+06 2.57000000e+02
5.63918630e+07 5.47493816e+05 4.42907557e+05 1.01636600e+06
6.30000000e+01 0.00000000e+00 0.00000000e+00 0.00000000e+00
0.00000000e+00 4.07200000e+03 2.09200000e+03 3.59980848e+00
1.84423685e+00 0.00000000e+00 1.93600000e+03 7.98717532e+02
7.94112014e+02 6.30613891e+05 0.00000000e+00 0.00000000e+00
1.00000000e+00 1.00000000e+00 0.00000000e+00 0.00000000e+00
0.00000000e+00 1.00000000e+00 0.00000000e+00 8.01319218e+02
2.76270936e+02 1.82617308e+03 0.00000000e+00 0.00000000e+00
0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
2.03000000e+02 5.60830000e+04 1.04000000e+02 1.89922000e+05
8.19200000e+03 1.28100000e+03 1.01000000e+02 2.00000000e+01
0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00] ===> PREDICTED AS Brute Force -XSS

Test DATA : [ 0.00000000e+00  0.00000000e+00  1.12641234e+08  3.00000000e+00
  0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
  0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
  0.00000000e+00  0.00000000e+00  0.00000000e+00  5.63206170e+07
```

In above screen in square bracket we can see network traffic data and after arrow =➔symbol we can see the type of cybercrime attack prediction. Scroll down above screen to view all cybercrime prediction

5. CONCLUSION

Finally, this study also has important implications for society. Over the last few years, the world has been facing cyberterrorism and cyberwar threats from nation-sponsored attackers. Pollitt defined cyberterrorism as ‘‘the premeditated, politically motivated attack against information, computer systems, computer programs and data which results in violence against non-combatant targets by subnational groups or clandestine agents.’’ Unlike most cybercrime, which is primarily motivated by monetary gain, cyberterrorists are politically motivated. As a result, governments should, for example, strengthen their ability to protect their citizens in online virtual environments by enhancing their immediate responses to threats such as cyberespionage and cyberterrorism. This issue therefore has profound implications in terms of the need for a global cyber defense to maintain a cyber-safe environment

REFERENCES

[1] K. Graves, Ceh: Official certified ethical hacker review guide: Exam 312-50. John Wiley & Sons, 2007.

[2] R. Christopher, “Port scanning techniques and the defense against them,” SANS Institute, 2001

[3] M. Baykara, R. Das¸, and I. Karado ˘gan, “Bilgi g ¨uvenli ˘gisistemlerindekullanilanarac¸larinincelenmesi,” in 1st International Symposium on Digital Forensics and Security (ISDFS13), 2013, pp. 231–239.

[4] S. Staniford, J. A. Hoagland, and J. M. McAlerney, “Practical automated detection of stealthy portscans,” Journal of Computer Security, vol. 10, no. 1-2, pp. 105–136, 2002.

[5] S. Robertson, E. V. Siegel, M. Miller, and S. J. Stolfo, “Surveillance detection in high bandwidth environments,” in DARPA Information Survivability Conference and Exposition, 2003. Proceedings, vol. 1. IEEE, 2003, pp. 130–138.

[6] K. Ibrahimi and M. Ouaddane, “Management of intrusion detection systems based-kdd99: Analysis with lda and pca,” in Wireless Networks and Mobile Communications (WINCOM), 2017 International Conference on. IEEE, 2017, pp. 1–6.