RESEARCH ARTICLE OPEN ACCESS

# A Data Analytical Approach to Cyber Crime Underground Economy

P.Subba Rao[1a], G.Lakshmi Konda[2b], Y.Nagasekhar Reddy[3c], G.Sai Durga Gnaneshwar[4d], Y.Bharath Kumar[5e], J.Bharath Kumar[6f]

[1]Assistant Professor, Department of Computer Science and Engineering, Santhiram Engineering College, Nandyal
[2,3,4,5,6]Department of Computer Science and Engineering, Santhiram Engineering College, Nandyal
P.Subba Rao: subbarao.cse@srecnandyal.edu.in

**ABSTRACT**
There hasn't been enough research done to create the framework for future academics and professionals in the area of information systems to follow despite the rapidly expanding cyber threats. Most people don't aware that hackers rely on a criminal economic structure called "CaaS." Due to a lack of prior work, we used a data analytics approach from a design science perspective to investigate cybercrime. We begin by providing a data processing framework for the study of the cybercrime underground, and we proceed to define CaaS and crime ware before introducing a classification algorithm to help us sort out the bad guys. Finally, to show how this infrastructure and classifications model might have been put into practice, we sketch out a possible organizational structure. In order to study the criminal underground market, this program initially uses a vast dataset collected from the internet security community. By adopting a design science approach, this study advances the design objects, underpinnings, and methods of the area. And on top of all that, it provides helpful guidance on how governments and businesses across a wide range of sectors should be ready for cybercrime attacks.

To be indexed with: It's a catchall word for the criminal underworld, black market, and hacker subculture.

## I. INTRODUCTION

Defenses against the growing threat of widespread cyber attacks (including malware and DDoS) have become a top priority for many businesses and other organizations. Likewise, criminal activity. The WannaCry ransomware was responsible for about 45,000 assaults in 2017 [1]. With the growing effect of cybercrime, authorities are under increasing pressure to increase expenditure on cyber security [2]. President Obama has proposed allocating over $19 billion on network security in the next fiscal year, an increase of nearly 35% from the previous year. Several recent assaults on a national or worldwide scale, like WannaCry and Petya, were executed by well-organized criminal gangs. In the underground black market, where hackers trade information about trying to hack, criminal organizations buy the vast majority of their hacking operations and equipment. The cybercrime sector of the underworld relies on this online black market, which is managed by organized crime groups [3]. A new kind of organization has emerged, one that oversees underground marketplaces and provides support for the activities of cybercriminals[11]. Cybercrime syndicates rely largely on shadowy underground networks for funding, support, and assaults (e.g., Hack forums and Crackingzilla). The vertical, concentrated, rigid, and stable design of cyber security organizations [4] sets them apart from traditional Mafia hierarchies, which rely on transparency and openness. Yet, this is not the case for criminal networks in cyberspace. Since cybercrime is a collection of interrelated issues, governments, organizations, and people often fail to notice the emergence of incredibly adept online fraudsters' marketing methods, such as Crimeware-as-a-Service (CaaS). Whereas many in the field of information security (IS) have taken an interest in cybercrime in response to the rising concerns generated by the exponential increase in cyber dangers, few have worked to provide the groundwork for this new interest or built adequate methodologies. Previous studies haven't delved far into the cybercrime "underground" economy. CaaS is one of the most important business models employed by the cybercrime underground, yet it is little understood. Both academic research and real-world experience have failed to shed light on the nature of this underworld or its underlying processes. The information vacuum and the real-world [5] difficulties that hackers encounter inspired our research. The cybercrime economy is examined using methods from design research and data analytics. The purpose of this paper is to (1) provide a framework for evaluating the cybercrime underground; (2) define CaaS and crime ware[14] in a way that correctly reflects both academic research and industry experience; (3) develop a classification method for CaaS and crime ware; and (4) develop an application to show how the recommended prototype and classifying idea may be applied in reality. By initially examining large datasets from the internet technical community, this approach of creating computational modeling (DSR) may be utilized to probe the ransomware ecosystem. Product of design science is an IT item created with the intention of resolving a particular issue. IT artifacts including as recommendation systems, ideas, structures, platforms, techniques, and applications are created as part of DSR [6]. The emphasis of behavioral science[8] is on establishing and justifying ideas that increase human or organizational capabilities, whereas the focus of DSR is on creating and justifying ways to understanding or forecasting human or organizational occurrences. DSR enriches the current body of literature and practice by shedding light on "design artefacts, planning and construct learning (e.g. framework), and/or design evaluation knowledge (e.g. methods). [7]. It follows these DSR guidelines and incorporates the creation of novel design objects, bases, and procedures. As DSR must demonstrate that design objects are "able to be implemented" in the economy to address a critical issue [7], we propose an efficient implementation approach rather than merely a theoretical one. A sample front-end application demonstrates the benefits and feasibility of using the suggested infrastructure and categorization strategy. So, this study added something to the field of design theory [9], 10]. In order to further our understanding of design science, we need novel formulations of ideas, theories, methods, and applications [10]. This investigation adds to the body of knowledge by providing essential ingredients such as constructs (classifications, structures, and programs), a modeling (complex implementation), an analytical approach, and instantiations (applications). DSR has advanced the field of methodology via the development and use of novel forms of assessment [12]. As a result, the classification model was tested in this study using dynamic analysis[15]. To round up the study, observational methods are used to analyze a front-end application (case examples). Moreover, this study provides useful recommendations for addressing the difficulties encountered by government organizations and enterprises of all sizes in their preparedness for cybercrimes[13].

## II. RELATED WORK

### Crime software as a service: investigating the rise of criminal software as a marketable good

Criminals are increasingly turning to the dark web to buy and sell malware in the form of "crimeware as a service" (CaaS). CaaS not only tries to make cyber assaults more streamlined, automated, and accessible to persons with less technical knowledge, but it also provides substantial contributions to the problem. CaaS and the shadow economy that has developed around it are defined and explored in this article. In addition, the study discusses the many crimeware services available on the dark web.

### b. To what extent does cybercrime have a structure? What the Internet May Mean for Criminal Collaboration

This article speculates on the possible future developments of organized crime in the digital realm. The first part of the book examines criminal groups in the "real world." After a definition of "organized crime," the article explores the positive aspects of well-coordinated criminal teams in the "real world." The paper then discusses why the two established forms of organized crime in the "actual world"—the "crime" approach and the hierarchies American Mafia model—are unlikely to catch on in the virtual one. In this section, it is shown that the limitations inherent to accomplishing things in "real life" inspired the development of both of the aforementioned models, although such limitations are mostly absent in the virtual world. Next, the paper considers the possibility of a cyber manifestation of organized crime. For this purpose, it examines how the armed forces use Netware. The study concludes that cybercrime networks would be transient, horizontal, and fluid, in contrast to the permanent, hierarchical types of organization seen in the "real world." A combination of these factors might make it difficult for authorities to apprehend criminals.

### c. Cybercrime Organizations and Their Varieties

Improvements in emerging technologies for communication and information (ICT) are being used by three groups to violate the law: There are three types of criminal organizations that utilize information and communication technologies (ICTs) to do criminal acts: 1) traditional criminal gangs that use ICTs to better their unlawful operations on land; 2) organised harmful cyber groups that exclusively ecommerce platform; and 3) groups of philosophically and ideologically biased persons that use ICTs to commit criminal acts. Media technology is more responsible for the occurrence of disputes or part of the evidence that supports or disproves a dispute in court, so police departments would benefit from a deeper understanding of cyber forensics principles, recommendations, processes, tools, but rather procedures, along with pro fundamentals, instructions, methods, devices, and strategies. There seems to be a need for more study and innovative approaches to combating organized crime in cyberspace.

## III. METHODOLOGY

Ninety percent of people today rely on online services like financial services, healthcare, and traffic data to go about their daily lives, but this has given rise to the issue of cyber theft as well as attack, out of which malware developers can indeed detect data transfer to obtain information or to add erroneous numbers, resulting in the customer losing money or handling of their own mechanism. Passwords used by users may be brute-forced, and BOT assaults can boost fake ratings. Websites are vulnerable to a wide range of cyber attacks because programmers may hire themselves out on the Dark Web in exchange for illicit funds.Many attacks are documented in this article, and the researcher used Nave Bayes to categorize them based on information analysis, revealing which forms of cybercrime are more resource-intensive. Signatures of important attacks, such as spamming and Code injection, are used to train the classification system, which is being developed using data gathered from key logger attacks on enterprises.In order to put the idea into action, the stress reactions have been developed:

Dataset Upload and Analysis: At this time, we will load a database, do various analyses (such as counting cybercrimes), and then clean the data set by removing any outliers.

The dataset would be split between training and testing halves, with 80 percent of the data used to train the Nave Bayes algorithm and 20 percent used to confirm the approach's predicted performance. The attacker classifications would be encoded using an integers ID.

Naive Bayes Classification Models Execution: In this course, classification algorithms would be trained using datasets with a high label agreement rate of 80% or more, and a recommended methodology would be developed.

Classification Performance Graph: We'll evaluate the efficacy of our suggested method by constructing a graph with high precision and naive Bayes accuracy using this component.

Cybercrime Prediction: We may use these modules to test a database of cybercrime networks and submit the results to a classification algorithm that can determine whether the database contains indicators of cybercrime.

## IV. RESULTS AND DISCUSSION

There are four distinct procedures for analyzing data under the proposed framework. Prospective hacking targets, as well as the various categorizations and market trends of CaaS and criminal ware, are discussed. That was a quick double-click as I was on the run. The following is the outcome of running the bat file:

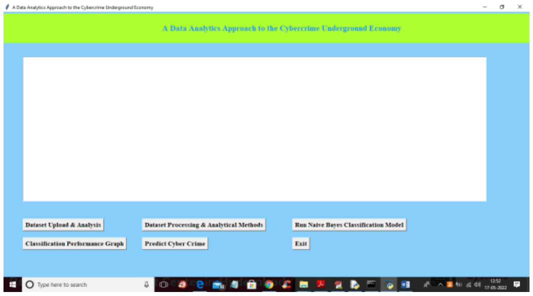

A Data Analytics Approach to the Cybercrime Underground Economy
A Data Analytics Approach to the Cybercrime Underground Economy
Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Run Naive Bayes Classification Model
Classification Performance Graph
Predict Cyber Crime
Exit
Type here to search
12:52
17-05-2022

Fig1: Dataset Upload & Analysis

If you want to upload a dataset for analysis, just click the corresponding icon in the upper right hand corner of the report.

The Naive Bayes retraining process is complete, and its prediction performance has been evaluated at 90%; to see the graph, choose the "Classification Methods" menu item.

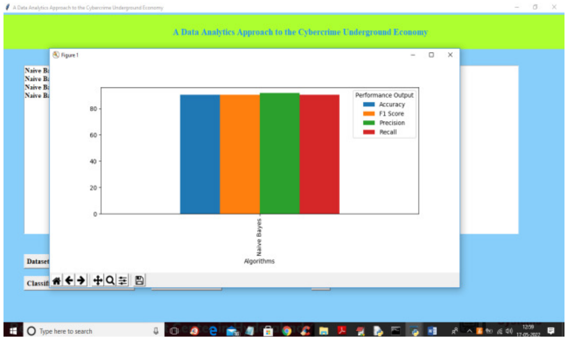
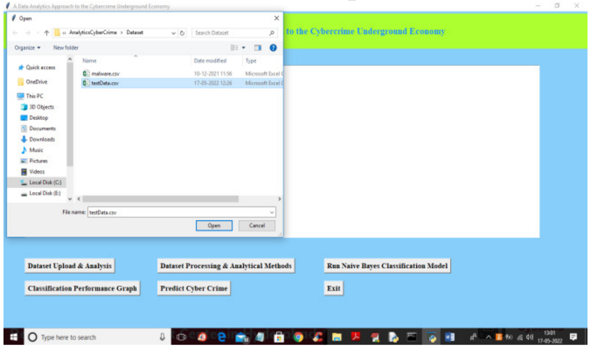


Fig2:Classification Performance Graph

The findings are broken down into individual categories denoted by different colors, such as Correct, Quality, etc. When testing data is submitted and shown visually, it can be seen that all parameters are functioning at or above 90% by viewing the graphs above and selecting on the 'Predict Computer Crime' tab.

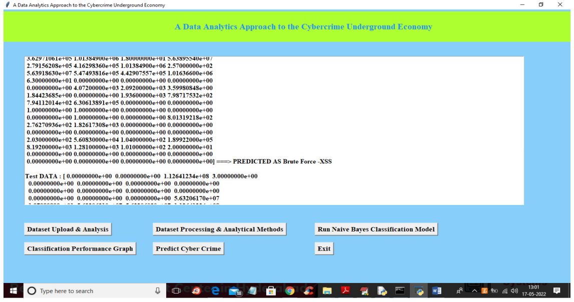

A Data Analytics Approach to the Cybercrime Underground Economy
Open
« AnalyticsCyberCrime › Dataset
Search Dataset
Organize ▾
New folder
Quick access
OneDrive
This PC
3D Objects
Desktop
Documents
Downloads
Music
Pictures
Videos
Local Disk (C:)
Local Disk (E:)
Name
Date modified
Type
Open
Cancel
to the 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

| Name | Date modified | Type |
|---|---|---|
| malware.csv | 10-12-2021 11:56 | Microsoft Excel C |
| testData.csv | 17-05-2022 12:26 | Microsoft Excel C |

File name: testData.csv

Fig3: Upload testData.csv

To see this result, load your dataset in the 'testData.csv' format by entering it and then selecting the 'open' button.



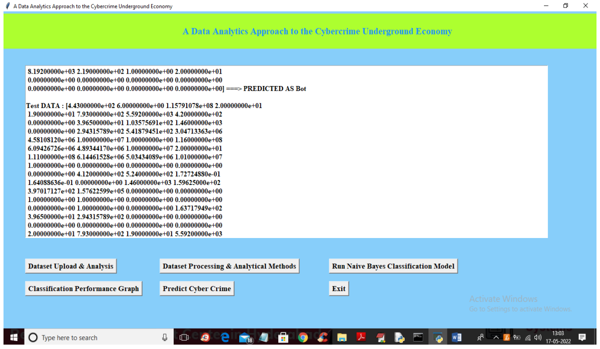
A Data Analytics Approach to the Cybercrime Underground Economy
A Data Analytics Approach to the Cybercrime Underground Economy
Dataset Upload & Analysis
Dataset Processing & Analytical Methods
Run Naive Bayes Classification Model
Classification Performance Graph
Predict Cyber Crime
Exit
Type here to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

Fig4: Predict Cybercrime

In the report seen above, data about Internet traffic appears in parenthesis, and the kind of cybercriminal behavior anticipated appears after the equal sign. This is a compilation of our predictions about future cyber assaults..



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
Activate Windows
Go to Settings to activate Windows.
Type here to search
13:03
17-05-2022

```
8.19200000e+03 2.19000000e+02 1.00000000e+00 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 Bot

Test DATA : [4.43000000e+02 6.00000000e+00 1.15791078e+08 2.00000000e+01
1.90000000e+01 7.93000000e+02 5.59200000e+03 4.20000000e+02
0.00000000e+00 3.96500000e+01 1.03575691e+02 1.46000000e+03
0.00000000e+00 2.94315789e+02 5.41879451e+02 3.04713363e+06
4.58108120e+06 1.00000000e+07 1.00000000e+00 1.16000000e+08
6.09426726e+06 4.89344170e+06 1.00000000e+07 2.00000000e+01
1.11000000e+08 6.14461528e+06 5.03434089e+06 1.01000000e+07
1.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
0.00000000e+00 4.12000000e+02 5.24000000e+02 1.72724880e-01
1.64088636e-01 0.00000000e+00 1.46000000e+03 1.59625000e+02
3.97017127e+02 1.57622599e+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 1.63717949e+02
3.96500000e+01 2.94315789e+02 0.00000000e+00 0.00000000e+00
0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00
2.00000000e+01 7.93000000e+02 1.90000000e+01 5.59200000e+03
```

Fig5: Predict Cybercrime

## V. CONCLUSION

Research in this area has focused more on building and analyzing artifacts than on developing and testing ideas since it is

often assumed that research in this area is primarily concerned with activities. They came up with a model for the flow of information and a technique for making predictions. We have also conducted retrospective analyses of these classifiers' effectiveness and post-implementation analyses of their performance using example programs. These four case studies illustrate the variety of potential industrial instances available to emerging academic institutions and businesses, as outlined by the DSR starting viewpoint. Ultimately, there are substantial societal implications of this research. Cybercrime and cyber war have been threatened for years by actors supported by nation states. As defined by Nice, cyber assaults are "the premeditated, politically motivated attack against information, personal computers, software engineers, and data that lead to crime against - anti aims by sub - national level organizations or covert agents." Political ideology drives cyber terrorism, whereas financial gain motivates the vast majority of cybercriminals. Hence, governments might, for instance, assist strengthen their capacity to safeguards their people in online digital realities by boosting their immediate remarks in reaction to assaults like cyber espionage or cyber terrorism. This problem significantly impacts the effectiveness of cyber defenses in maintaining a secure online environment.

## VI. REFERENCES

[1] MV Subramanyam, K Soundararajan, J. Sofia Priya Dharshini“Adaptive Modulation and Coding With Incremental Redundancy Hybrid ARQ in MIMO Systems: A Cross Layered Design.”, International Journal of Engineering Research and Applications, Vol.3, no.5, pages. 503-7,2013.

[2] MV Subramanyam, K Satyaprasad, S L Prathapa Reddy“A hybrid genetic fuzzy approach for power control cross layer MAC protocol in wirelessnetwork”, International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT) , pages, 181-186, December2015.

[3] MVSubramanyam,RSumalatha“Imagede noisingusingSpatialAdaptiveMaskFilter for medical images”, International

Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), pages. 1-4, June2015.

**[4]** Makam Venkata Subramanyam, Kodati Satya Prasad, Bandani Anil Kumar"Anenergy efficient clustering using K-Means and AODV routing protocol in Ad-hoc networks" , Vol.12, no.2, pages. 125-134,2019.

**[5]** [5] Farooq Sunar Mahammad, M. Sharmila Devi**,** D Bhavana, D Sukanya, TV Sai Thanusha, M Chandrakala, P Venkata Swathi "Machine Learning Based Classification and Clustering Analysis of Efficiency of Exercise Against Covid-19 Infection" JOURNAL OF ALGEBRAIC STATISTICS, Vol.13, no.3, pages. 112-117, June2022.

**[6]** SunarmohammedFarooq,"StaticPeersfor Peer-to-PeerLiveVideoStreaming'',International journal of Scientific Engineering and Technology Research, Vol.05, No.34, Pages:7055-7064,October-2016.

**[7]** FarooqSunarMahammad,PalanisamyRamasamy,Karthik Balasubramanian "Comparative analysis of 3D-SVM and 4D-SVM for five-phase voltage sourceinverter", International Transactions on Electrical Energy Systems, Vol.31, No.12, Pages: e13138, December-2012.

**[8]** Farooq Sunar Mahammad, A Hemanth Kumar "Machine Learning Based PredictiveModelforClosedLoopAirFilteringSystem",JOURNALOFALGEBRAIC STATISTICS, Vol.13, no.3, pages. 609-616, July2022.

**[9]** M.Amareswara Kumar, FarooqSunarMahammad"Traffic Length Data BasedSignal
Timing Calculation for Road Traffic Signals Employing Proportionality Machine
Learning"JOURNAL OF ALGEBRAIC STATISTICS**,** Vol.13, no.3, pages. 40-45, June 2022.

**[10]** V Lakshmi Chaitanya, "Machine Learning Based Predictive Model for Data Fusion Based Intruder Alert System",journalofalgebraicstatistics,Vol.13,no.2,pages.2477- 2483, June2022.

.

**[12]**P Subba Rao. "Applications of machine learning in environmental engineering" JOURNAL OF Science,Technology and evelopment (2021): 643-648

**[13]**P SubbaRaol. "Using an Energy Efficient Extended Leach Work with MultilevelClustering Approach" from JOURNAL of Research Publication and Reviews (2020): 60-64

**[14]**L. E. Cohen and M. Felson, "Social Change and Crime Rate Trends: A Routine Activity Approach," Am. Sociol. Rev., vol. 44, pp. 588–608, 1979.

**[15]**M. Felson, "Routine Activities and Crime Prevention in the Developing Metropolis," Criminol., vol. 25, no. 4, pp. 911–932, 1987.

**[16]**F. Mouton, M. M. Malan, K. K. Kimppa, and H. S. Venter. “Necessity for ethics in social engineering research,” Comput. Security, vol. 55, 114–127, 2015.