IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125111

DARK-NET ECOSYSTEM CYBER THREAT INTELLIGENCE (CTI) TOOL

Abhishek R[1], Prof. Shreehari H S[2]

Department of ECE , SJC Institute of Technology, Chickballapur, Karnataka, India[1]

Assistant Prof**,** Department of ECE , SJC Institute of Technology, Chickballapur, Karnataka, India[2]

Abstract— The frequency and costs of cyber-attacks are increasing each year. By the end of 2020, the total cost of data breaches is expected to reach $2.1 trillion through the ever-growing online presence of enterprises and their consumers. The tools to perform these attacks and the breached data can often be purchased within the Dark-net. Many of the threat actors within this realm use its various platforms to broker, discuss, and strategize these cyber-threat assets. To combat these attacks, researchers are developing Cyber-Threat Intelligence (CTI) tools to proactively monitor the ever-growing online hacker community. This topic will detail in creating and using a CTI tool that leverages a social network to identify cyber threats across major Dark-net data sources. Through this network, emerging threats can be quickly identified so proactive or reactive security measures can be implemented.

Keywords— Dark-Net Ecosystem, Dark-Net Markets, Dark-Net Forums, Cyber-Threat Intelligence Tools, Dark-Net Network Visualization.

I. INTRODUCTION

The Darknet ecosystem refers to a collection of websites, forums, and marketplaces that are only accessible using specialized software or configurations that provide anonymity to users. The Darknet is often associated with illegal activities, including cybercrime, such as hacking, fraud, and the sale of illicit goods and services.

Cyber Threat Intelligence (CTI) tools are used to collect and analyze information related to cyber threats, including data from the Darknet. These tools help organizations to identify potential cyber threats and vulnerabilities, and to take appropriate measures to protect their assets. CTI tools use various techniques to gather data, including scraping the web, analyzing social media, and monitoring Darknet forums and marketplaces. Cyber-attacks have become an increasingly prevalent threat in recent years, with numerous high-profile incidents affecting individuals, organizations, and even entire nations. These attacks involve the unauthorized access, manipulation, or destruction of computer systems, networks, and data, often with malicious intent. 'DATA' breaches are becoming more frequent with the "rapid digitization of consumers' lives and enterprise records that will increase the cost of data breaches to $2.1 trillion globally by 2019" [1]. One of the most significant cyber-attacks in recent years was the WannaCry ransomware attack in 2017, which affected over 200,000 computers in 150 countries. Other notable incidents include the Equifax data breach in 2017, the Marriott data breach in 2018, and the SolarWinds supply chain attack in 2020. Cyber-attacks can have serious consequences, including financial losses, reputational damage, and even threats to national security. Various actors can carry them out, including criminal organizations, state-sponsored hackers, and even disgruntled employees.

II. LITERATURE REVIEW

The literature review of the topic will cover research efforts made within the areas of cyber-threat intelligence (CTI) and social network analyses in regard to the Dark-net, and the identified research gaps.

1. Social Network Analysis

B. Lane, et al. Examining the structure of the Dark-net ecosystem is an important exercise for gaining domain knowledge of these threat communities. The underlying processes across these sites such as trading, and vetting remain poorly understood. Using an Event Analysis of Systemic Teamwork (EAST) approach can be beneficial in exploring the implications in trying to understand the complex ecosystem while trying to identify vulnerabilities for potential disruption [2].

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125111

This approach for understanding the Dark-net ecosystem from a domain perspective can be useful when applying a social network-based analysis across multiple datasets. However, EAST's activity centric approach does not prioritize the content of these interactions, thus performing a social network analysis for this domain may expand upon these findings for the creation of actionable CTI [2]. The area of deep web social network visualization is a relativity untapped field. Researchers have mainly focused on plotting data more towards a surface level, such as mapping. onion sites via how they are connected to one another through URLs. An example of this is Hyperion Grey's Dark Web Map [3]. This overly expansive approach is a useful exercise for getting an idea of how this decentralized network operates in its entirety. However, such approaches rarely include the actual content of the sites. This lack of granularity prevents researchers from producing even moderately actionable CTI findings.

2. cyber-threat intelligence (CTI)

T. Kaskela, A. Oksanen, is mainly focused on the identification of cyber-threats through breach forensics across the Dark-net. However, the area of hacker-related assets on Dark-net Markets (DNMs) has received less attention than other parts of these digital black markets. This is due to the online economy for narcotics being often prioritized by law enforcement as it is believed to be a more widespread issue. Nevertheless, the relationship between vendors, buyers, and post authors remains consistent across these areas [4]. Details within product listings, feedback and threads can paint an accurate picture of the emerging threats and actors within this domain Through a text centric focus, several different approaches can be used to find relevant information on the prevalence of malicious tools, services, actors, and breach data [4].The resource that was the main influence for this area of the Topic was an analysis conducted by Ryan Compton in 2015 that aimed to examine co-occurrence relationships between threat actors and the products they offered [5]. This method used a stochastic block model-based hierarchal edge bundling to generate a visualization of the evolution product network.

3. Identified Research Gaps and Questions.

Two gaps were identified by the following above literature Survey. Within the scope of Dark-net analyses, the examination of cyber-threat assets and the communication surrounding them across multiple data sources has remained relatively unexplored through analytical approaches. The second gap that was found regarded the extent of pre- and post-obtainable data following breaches and their victims remains unclear.

III. RESEARCH DESIGN

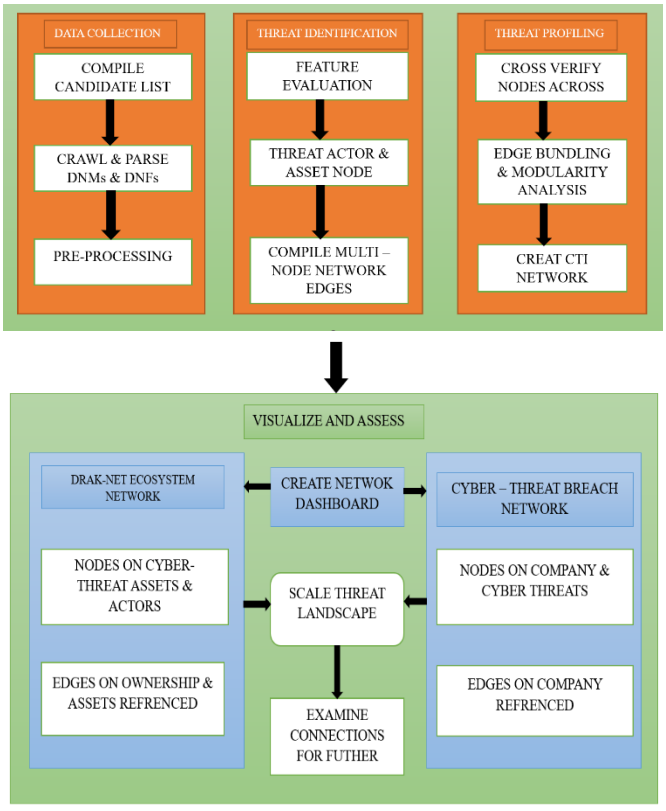

Figure 1 Dark-Net Ecosystem Network Analysis

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125111

A. Data Collection

The data collection phase began with an initial draft of potential Dark-net markets (DNMs), Dark-net Forums (DNFs) and exploit databases to be collected that was based on each site's number of listings, threads, threat actors as well as the number of cyber-threat related listings. From this list, python-based web crawlers were developed to collect each site. Our team's crawlers initially faced issues from several sites' crawling-prevention measures.

This issue was solved by using ten crawlers all linked to different accounts. With these ten crawlers running in unison, each site's logs were congested enough to prevent the occurrence of captcha codes. Our team's crawlers initially faced issues from several sites' crawling-prevention measures. This issue was solved by using ten crawlers all linked to different accounts. With these ten crawlers running in unison, each site's logs were congested enough to prevent the occurrence of captcha codes. This collection spanned from October 2017 – January 2019 that includes the five largest DNMs that provided 224,270 product listings and 7,911 vendors, the two largest exploit databases (Exploit DB/0day.Today) providing 43,678 exploit listings, and three major DNFs providing 204,001 threads and 14,196 authors. A total of 112 categories are found across these ten datasets.

1) Data Sources

The two major data sources for this Social Network analysis of Dark-net communities are Dark-net forums as the primary data source for observing threat actor interactions and Dark-net markets as the cyber-threat asset product network. The second network produced is focused on the relationship between cyber threat assets and exploit database listings linked by companies to make a breach forensic network. Below are brief explanations of the data sources that were used to create these multi-node networks.

a) Dark-Net Forums

These community hubs are hot beds for threat actor communication and activity across various CTI data sources ranging from DNMs, carding shops, independent vendors, contractual hackers and much more. Dark-net forums are text rich and contain contextual and time series data that can be linked to other data sources for additional knowledge gain.

b) Dark-Net Markets

These deep-web-based e-commerce sites function primarily as black markets where users create profiles to buy or sell drugs, breach data, forged documents, currency, hacking tools, guides, as well as many other illegal and legal goods. In order to keep transactions anonymized and integrable, Dark-net markets typically adopt an escrow system where cryptocurrency is laundered through the site and is released following buyer feedback via product reviews.

c) Exploit Databases

Within both the surface web and deep web are databases dedicated to revealing exploits discovered across various operating systems, applications, sites, etc. These databases include exploits, shellcode, 0 days and much more. Most of these listings in their entirety are free for users while the lesser-known and more visceral listings are available for purchase via cryptocurrency. Table I details the DNF & DNM datasets used to create the multi-node social network.

Site	Type	Language	Total %	Actors
Rutor	DNF	RU	28.8% 8	8,318
Dream	DNM	EN	27.77%	2,958
Wallstreet	DNF	EN	20.66%	5,382
TradeRoute	DNM	EN	18.71%	4,201
FrenchDeepWeb	DNM	FR	1.33%	261
Silk3	DNF	EN	1.96%	429
Tochka	DNM	EN	0.83%	213

Table 1 Network Percentage by Site [5].

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125111

1. Threat Identification

The decision to use a multi-node approach was made in order to maintain a robust network that is granular and modular for any additional datasets. This made the edge creation stage exhaustive in order to identify any connections that were not immediately parent within the initial testbed. DNF data was used as the base of this network due to thread authors being able to provide versatile content that can add value were other data-sets may lack. An example that inspired this Topic was the universal absence of time series data across DNM listings that stunted the CTI credibility of the data. However, many DNF threads are used as a promotion or review platform for specific products for sale on their respective DNMs. By linking these datasets, not only can time data be added to product data, but data can also be derived via sentiment, credibility, buyer names as well as an estimated quantity sold. Table 2 details the data dictionary used to create the multi-node network between DNMs an DNFs. The breach forensics network aims to identify if the start and end points of recent data breaches can be linked to Dark-net data sources. Start points include potential exploits or user authentication means that could be used to execute a breach. Due to the obscurity and lack of consistency within listing names, classification rendered few results. Using SQL queries, text from each DNM product and DNF thread could be accurately pulled to discover company names. 132 major companies were identified across the three data sets, with notable industries such as banking, e-commerce, airlines and much more. Examples of some of the more prevalent companies are Amazon, PayPal, and Microsoft.

B. Threat Profiling

1) Dark-Net Ecosystem Social Network

From working extensively with the data to create every possible relationship, an initial outline of the graph was made that would better align with the CTI relevant features that needed to be included for this Topic's analysis. Because of this, the threat profiling stage acted as a secondary data preprocessing phase to eliminate any unneeded noise that arose following the node and edge creation tasks. This preprocessing was done by plotting the node and edge data into Gephi and running initial network statistics to generate the degree, rank, closeness and betweenness for each node. Through focusing on relevant threat actors while using each node's statistics, the initial graph was filtered down from 450,378 to 167, 63 nodes. Most of the eliminated nodes stemmed from products, or threads with minimal community involvement (few authors connected) or having no apparent CTI value (category associated with narcotics or legal goods). This final dataset created three networks, each with a varying focus.

2) Cyber-Threat Breach Forensic Network

To create a hierarchical network, nodes and edges were created based on categorical queries that pulled only relevant CTI data that provided 97 nodes across the 5 DNMs, 3 DNFs and 2 Exploit Databases.

C. Visualization

In order to maintain the desired finesse of a practical and actionable CTI tool, the ability to add any missed or new data would need to be a vital feature. Through this consideration, an interactive network dashboard was built using Gephi. This powerful open sourced tool is the best user-friendly network analysis and visualization tool for big data. Through creating a thorough dashboard, additional data can be streamed directly into a network and exported for fast targeted CTI analyses.

IV. ANALYTICAL APPROACH

1. Dark-Net Ecosystem Social Network

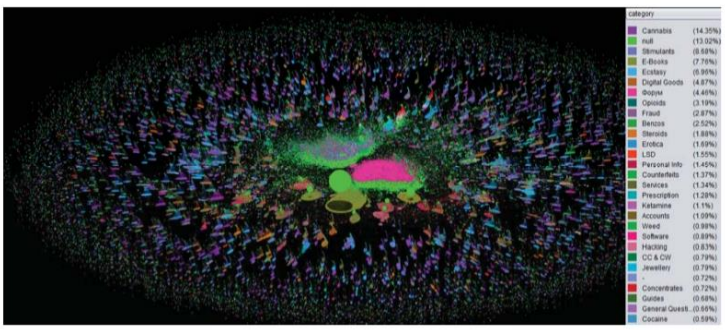

Figure 2 Analytical Approach: Darknet Ecosystem Social Network. - (Green nodes are threat-actors; other colors are categories) [11].

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125111

2. Dark-Net Cyber-Threat Assets Network

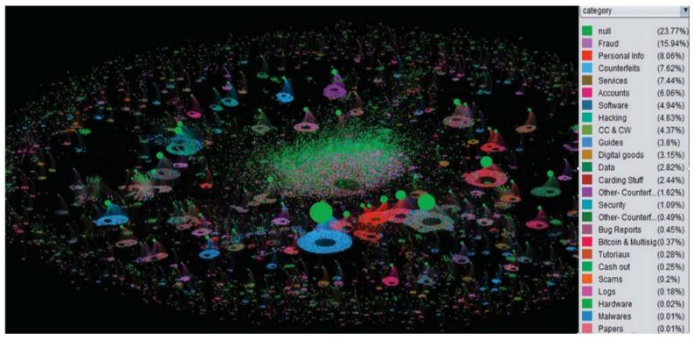

Figure 1 Analytical Approach: Darknet Ecosystem Cyber-Threat Network (Green nodes are threat-actors; others are security categories) [11].

V. RESULTS & CONCLUSION

Fraud is the most populated category through its connection to 47 of the 67 companies used within this edge bundled network. Breached accounts and hacking tools follow closely behind as cyber-threats that are currently available to be used to attack companies' and their customers. Through looking at a notable breached company, PayPal has a weight 238% larger than the next highest company node with listings spanning across all 12 DNM categories and five exploit categories.

The information conveyed across these networks creates a coherent depiction of this major facet within the Dark-net ecosystem. By examining the connections within each network, researchers will be able to find CTI related findings that may have been overlooked through initial analyses that take an overly expansive or focused approach.

Through this dashboard, CTI researchers can perform:

- Fast integrations of additional data.
- Targeted analyses across multiple datasets.
- Automated approaches to digital targeting process.

The effectiveness of this CTI tool needs to be evaluated outside of this Topic. However, by examining the results within this study, the economy of the Dark-net and breach victims can be better scaled and examined for granular analyses. The versatility of this tool will guarantee its relevancy and efficiency for any future research within the often-hectic area of cyber-threat intelligence.

REFERENCES

[1] J. Moar, et al. Cyber Crime and the Internet of Threats, Juniper, 2017.
[2] B. Lane, D et al. The Dark Side of the Net: Event Analysis of Systemic Teamwork (EAST) Applied to Illicit Trading on a Darknet Market. Proceedings of the Human Factors and Ergonomics Society Annual Meeting Vol.62 Iss.1. Sage Journals, (2018).
[3] M. Hasse, Dark Web Map, Hyperion Grey, (2019).
[4] T. Kaskela, A. Oksanen, Seller's reputation and capacity on the illicit drug markets: 11-month study on the Finnish version of the Silk Road,Research Gate, (2017).
[5] R. Compton, Darknet Market Basket Analysis (2015).
[6] J. Robertson, A. Diab, E. Marin, E. Nunes, V. Paliath, , J. Shakarian, Darkweb Cyber Threat Intelligence Mining. Cambrige University Press, (2017).
[7] T. Holt, O. Smirnova, A. Hutchings, Examining signals of trust in criminal markets online. Oxford Journal of Cyber Security, (2016).
[8] H. Lawrence, A. Hughes, R. Tonic, C. Zou, D-miner: A Framework for Mining, Searching, Visualizing, and Alerting on Darknet Events IEEE. University of Central Florida, (2017).
[9] T. Reksna, D. Garlaschelli, F. Takes, Complex Analysis of Darnket Black Market Forum Structure. Huygens-Kamerlingh Onnes Lab. Leiden University, (2017).

DOI: 10.17148/IJARCCE.2023.125111

[10] https://www.google.com/url_Analytical-Approach-Darknet-Ecosystem-Breach-Network-Green-nodes-are-threat-actors (accessed on 19/04/2023).
[11] https://www.google.com/Dark-Net-Ecosystem-Cyber-Threat-Intelligence,Arnold Ebrahimi - (accessed on 19/04/2023).
[12] Keim, Y., & Mohapatra, A. K. 2022. Cyber threat intelligence framework using advanced malware forensics. International Journal of Information Technology Singapore, 14, 521–530.
[13] Sakellariou, G., Fouliras, P., Mavridis, I., & Sarigiannidis, P. (2022). A Reference Model for Cyber Threat Intelligence (CTI) Systems. Electronics (Switzerland), 11(9).
[14] Evangelista, A. (2018). Darknet markets - competitive strategies in the underground of illicit goods. Eindhoven University of Technology.
[15] Cilleruelo, C., De-Marcos, L., Junquera-Sanchez, J., & Martinez-Herraiz, J. J. (2021). Interconnection between Darknets. IEEE Internet Computing.
[16] Z. Mador, Keep the dark web close and your cyber security tighter, Computer Fraud & Security, vol. 2021, no. 1, pp. 6–8, 2021.
[17] P. Shakarian, Dark-web cyber threat intelligence: from data to intelligence to prediction, Information, vol. 9, no. 12, p. 305, 2018.
[18] T. Miloshevska, Dark web as a contemporary challenge to cyber security, Kriminalisti˘cke Teme, vol. 5, pp. 117–128, 2019.
[19] S. Samtani, M. Abate, V. Benjamin, and W. Li, Cybersecurity as an industry: a cyber threat intelligence perspective, in 'e Palgrave Handbook of International Cybercrime and Cyberdeviance, T. J. Holt and A. M. Bossler, Eds., Palgrave Macmillan, London, USA, 2020.