Winter 12-11-2022

Network analysis of a darknet marketplace: Identifying themes and key users of illicit networks

Obi Ogbanufe
University of North Texas, obi.ogbanufe@unt.edu

Fallon Baucum
University of North Texas

Jasmine Benjamin
University of North Texas

Network Analysis of a Darknet Marketplace: Identifying Themes and Key Users of Illicit Networks

Obi Ogbanufe 1[1]
Ryan College of Business
University of North Texas,
Denton, Texas, USA

Fallon Baucum
Ryan College of Business
University of North Texas,
Denton, Texas, USA

Jasmine Benjamin
Ryan College of Business
University of North Texas,
Denton, Texas, USA

ABSTRACT

The global cost of cybercrime is estimated to reach $10 trillion by 2025. To perpetuate cybercrime, cybercriminals often use darknet markets, which are online platforms where cybercriminals sell, purchase, and trade stolen products and hacking tools. This study is a research in progress that focuses on analyzing darknet markets to identify key actors and understand their networks, interactions, and emergent themes. The study hopes to increase our understanding of the nature of criminal activities, add to the literature, and provide insights that may help stakeholders build tools for disrupting or preventing activities on the darknet.

Keywords: Darknet markets, cybercrime, social network analysis, centrality, topic modeling

[1] Corresponding author. obi.ogbanufe@unt.edu +1 940-565-3111

INTRODUCTION

The global cost of cybercrime is estimated to reach $10 trillion by 2025 (Morgan, 2020). According to the Federal Bureau of Investigations' annual Internet Crimes Report, cybercrime amounted to a $3.5 billion loss to organizations and individuals in the United States (FBI, 2020). To perpetuate cybercrime, cybercriminals use darknet markets, which are online platforms where cybercriminals sell, purchase, and trade stolen products and hacking tools. Strategies to analyze darknet markets are needed to address and mitigate these crimes. Given the prevalence and impact of cybercrime on individuals and organizations, there is a growing interest in understanding their structure and building tools to disrupt these activities (Mahmood et al., 2010). The disruption of these criminal activities is so crucial that it is a part of the Department of Homeland Security's (DHS) Criminal Investigations and Network Analysis Center's (CINA) mission.

This study focuses on the network analysis of darknet markets to understand their networks and on topic modeling to understand the content of their interactions. The body of research on darknet market analysis is small but growing (Benjamin et al., 2019). Darknet markets (DNM) and their transactions are viewed as digital traces that provide criminal investigations, law enforcement, and researchers with ample data to analyze (Aldridge & Askew, 2017).

In this research in progress, the objective is to analyze darknet networks and identify key actors and the content of their interactions. This study's analysis hopes to add to the literature, contributing to the understanding of darknet networks and the nature of criminal activities through the identified themes, potentially providing intelligence on emerging activities.

The rest of the paper is structured as follows; a review of online criminal networks is provided, followed by a research method description. Then we provide a preliminary data analysis and expected results.

RELATED WORK

DNMs are online platforms where criminals sell, purchase, and trade stolen products, illicit services, illicit drugs, and hacking tools. DNMs are also characterized by their use of hidden networks (e.g., Tor) and payment systems (e.g., bitcoin, escrow) that anonymize users and their transactions. There are primarily two streams of research in the online criminal networks' literature: forum-based hackers and darknet-based illicit traders. The literature stream on forum-based hackers examines actors that share tools and software used in breaches, phishing, and other related hacking of technology products. For example, Benjamin et al. (2015) provide an overview of the threats and vulnerabilities in hacker forums that can benefit cyber threat intelligence. Motoyama et al. (2011) empirically characterize six forums to examine their social network properties, services, and how trust is gained or lost. Further, Holt and Lampke (2010) analyze 300 threads from different forums used by cybercriminals, showing the wide range of stolen data assets being traded in hacker forums. More recently, Yue et al. (2019) analyzed hacker forums to understand the nature of cyber-attacks emanating from the forum.

The second stream is the darknet literature that examines underground marketplaces where illicit goods and services (i.e., drugs, firearms, trafficking) are traded. Researchers have explored criminals and their need to maintain a reputation level in order to increase their income and longevity (Décary-Hétu et al., 2012; Décary-Hétu & Dupont, 2013). Samtani et al.(2017) explore the connections and influence of criminals who trade illicit drugs on the darknet to identify their malicious assets and determine key individuals.

Social Network Analysis

Many of these studies use social network analysis to understand the social structure of actors in stolen data markets and identify leaders and their influence within the network (Décary-Hétu & Dupont, 2012; Holt et al., 2012; Lu et al., 2010). Social network analysis uses graph theory to explore the structural relationships among actors in a network, represented as nodes and edges (Samtani et al., 2017). Nodes are actors or entities, while edges are the communications between the entities. Analyzing the nodes and edges results in centrality measures, which are rankings of nodes within the network. Two main network centrality measures are often used to assess the structural position and ranking of members in a network; degree and betweenness centrality. Degree centrality identifies nodes (actors) that are leaders or experts in the network; these are actors who disseminate the most information in the network (Ogbanufe & Kim, 2018). Betweenness centrality is used to identify actors who are information brokers or gatekeepers in the network. Researchers have used these measures to identify prominent and influential hackers in a network setting. For example, Lu et al. (2010) assess the centrality of hackers in different forums and darknet sites. In the current study, identifying darknet actors with high centrality may serve a similar purpose; however, this study also examines topics and themes of interest in the darknet context.

RESEARCH METHOD

The data is archival data from a darknet marketplace called Wall Street Markets (WSM). Currently, a defunct site, WSM data provides a multi-year dataset from 2017 – 2019. The dataset has a sample of 45,371 posts and 5,533 users. The dataset contains messages posted and the threads in which they were posted. The Institutional Review Board (IRB) approved the use of the archival data. Following guidelines in the DNM literature on ethical research conducted with

data captured through the scraping of DNMs, this study anonymized usernames, paraphrased messages, and deleted URL links (Aldridge & Askew, 2017; Décary-Hétu & Aldridge, 2015). The data is anonymized by assigning each user a unique UserID.

PRELIMINARY DATA ANALYSIS

First, topic modeling is performed using SAS Enterprise Miner to understand the content or themes in their interactions. We extracted 10 topics from the dataset. The topics were then analyzed and labeled by the research team. We invited a darknet criminology domain expert to assist in verifying the labeling of the topics. In addition, the topic labels were verified using the DNM literature (Aldridge & Askew, 2017; Décary-Hétu & Aldridge, 2015; Morselli et al., 2017; Tzanetakis et al., 2016). Table 1 shows the 10 extracted topics, their initial labels, and their grouping. Group A is designated as topics where users (i.e., vendors) provide their profiles and external link locations so others can find them. We designate Group B as products sold on the darknet, like guides that provide other criminals with information on how to conduct a crime, stolen social security numbers, and bank account information for sale. Group C is designated as PGP-based conversations that reduce visibility and evade detection from law enforcement. Groups D and E are designated as ads and shout-outs to others on the site. These groups form a categorization of the emergent themes in the network.

Table 1: Topic Modeling Extraction of 10 Topics

Extracted Topic	**Label**	**Group**
[Link location], id, http, [link location]	External link location	A
guide, records, ssn, pack, extra	Guides and SSN records	B
room, chat, wallst3gi4a5wtn4, bank log, fraud chat room	Bank accounts for sale	B
pgp, pgp signature, begin, signature, sign	PGP key	C
9a7ae0b905, [link location]	External link location	A
order, vendor, write, know, buy	Vendor information	A
external, external contact, contact, allow link	External contact information	A
testicle, nun, eskimo, sand, smash	Advertisements	D
overdue, dont break, shoutouts, supermod	Shoutouts and feedback	E

profile, weed, onion, wallstyizjhkrvmj, order	Weed seller profiles	A

We then performed a network analysis of the data using R program. Figure 1 depicts the results, which show the network structure. The red dots are nodes showing a visual representation of each unique actor in the network. Each node is accompanied by a previously assigned UserID to identify the actor. There are 5,533 unique users (red nodes). The black areas are the UserIDs clustered together.

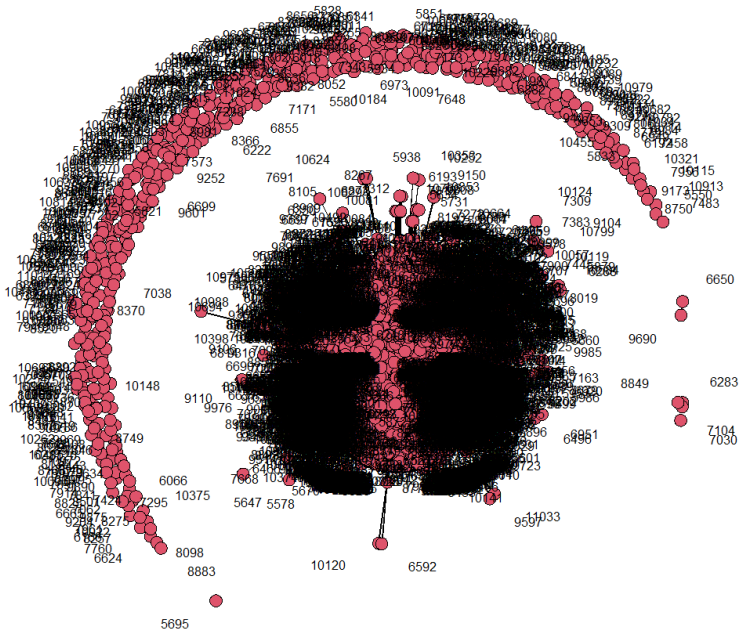

Figure 1. Network Analysis Diagram.

In this diagram, there are two main communities and a few smaller communities; a community is a cluster of nodes (Leskovec et al., 2009). The first community is the cluster of nodes shaped in a curve. This community is full of regular actors that participate within the network but are not interconnected with key actors. The second community, which is central in the diagram, consists of actors with high centrality that are viewed as the key players within the network (Abbasi et al., 2014). These actors are connected through links in which they distribute information. The length of each link between actors varies in distance based on how close the

relationship is. In addition, we present the analysis of key actors in the network. Tables 2 and 3 show degree centrality (leaders) and betweenness centrality (information brokers). Table 3 shows the combined key leaders based on the highest degree and betweenness centrality values. The results show that UserID 8671 is a leader in both degree and betweenness centrality categories.

Table 3: Degree Centrality

UserID	Degree
5595	403188
8671	366810
5534	2330
5535	3
5536	64
5537	3
5538	4
5539	3
5540	22

Table 4: Betweenness Centrality

UserID	Between
8671	48.16
6626	3.901
9468	3.845
6172	3.106
6491	2.016
7121	1.895
8344	1.831
7767	1.690
6769	1.547
5595	1.434

Table 6: Forum Leaders

Centrality	UserID
Degree	5595
	8671
	5534
Between	8671
	6626
	9468

PRELIMINARY RESULTS AND CONCLUSION

We identified 10 major themes (e.g., PGP for visibility reduction) and key actors in the network in the preliminary analysis. The analysis is expected to contribute to the literature (e.g., Benjamin et al., 2019; Yue et al., 2019) and illuminate the effectiveness of social network analysis to identify key actors and topic analysis to identify the major themes and topics of their interactions. Given the constraints on law enforcement resources (Horton-Eddison et al., 2021), it also allows practitioners to identify effectively actors. Given that the disruption of criminal activities is a crucial part of law enforcement, we hope this research expands current

understandings of darknet networks and the nature of criminal activities through the identified themes and provides intelligence on emerging activities. By identifying and categorizing emergent themes in the darknet, we hope that it helps stakeholders understand the techniques actors use to communicate and avoid visibility. Since this work is a research in progress, more analyses are planned (e.g., statistical, topic modeling with more themes, and comparison of different darknet markets).

ACKNOWLEDGEMENTS

This study was developed under an appointment to the DHS Summer Research Team Program for Minority Serving Institutions, administered for the U.S. Department of Homeland Security (DHS) by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between DHS and the U.S. Department of Energy (DOE). ORISE is managed by Oak Ridge Associated Universities (ORAU) under DOE contract number DE-SC0014664. This document has not been formally reviewed by DHS. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of DHS, DOE, or ORAU/ORISE. DHS, DOE and ORAU/ORISE do not endorse any products or commercial services mentioned in this publication.

REFERENCES

Abbasi, A., Li, W., Benjamin, V., Hu, S., & Chen, H. (2014). Descriptive Analytics: Examining Expert Hackers in Web Forums. In *IEEE Joint Intelligence and Security Informatics Conference*. The Hague: IEEE.

Aldridge, J., & Askew, R. (2017). Delivery dilemmas: How drug cryptomarket users identify and seek to reduce their risk of detection by law enforcement. *International Journal of Drug Policy*, *41*, 101–109. https://doi.org/10.1016/j.drugpo.2016.10.010

Benjamin, V., Li, W., Holt, T., & Chen, H. (2015). Exploring threats and vulnerabilities in hacker web: Forums, IRC and carding shops. *Intelligence and Security*.

Benjamin, V., Valacich, J. S., & Chen, H. (2019). DICE-E: A Framework for Conducting Darknet

Identification, Collection, Evaluation with Ethics. *MIS Quarterly*, *43*(1), 1–22.
Décary-Hétu, D., & Aldridge, J. (2015). Sifting through the Net: Monitoring of Online Offenders by Researchers. *The European Review of Organised Crime*, *2*(Iccc), 1–19. http://sgocnet.org/site/wp-content/uploads/2014/07/07_DecaryHetuAldridge_pp122-141.pdf
Décary-Hétu, D., & Dupont, B. (2012). The social network of hackers. *Global Crime*, *13*(3), 160–175.
Décary-Hétu, D., & Dupont, B. (2013). Reputation in a dark network of online criminals. *Global Crime*, *14*(February), 2–3.
Décary-Hétu, D., Morselli, C., & Leman-Langlois, S. (2012). Welcome to the Scene: A Study of Social Organization and Recognition among Warez Hackers. *Journal of Research in Crime and Delinquency*, *49*(3), 359–382.
FBI. (2020). FBI 2019 Internet Crime Report. *2019 Internet Crime Report Released*. https://www.fbi.gov/news/pressrel/press-releases/fbi-releases-the-internet-crime-complaint-center-2019-internet-crime-report. Accessed 20 February 2021
Holt, Thomas, & Lampke, E. (2010). Exploring stolen data markets online: products and market forces. *Criminal Justice Studies*, *23*(1).
Holt, TJ, Strumsky, D., & Smirnova, O. (2012). Examining the social networks of malware writers and hackers. *International Journal of Cyber Criminology*, *6*(1), 891–903.
Horton-Eddison, M., Shortis, P., Aldridge, J., & Caudevilla, F. (2021). Drug Cryptomarkets in the 2020s: Policy, Enforcement, Harm, and Resilience, 1–12.
Leskovec, J., Lang, K. J., Dasgupta, A., & Mahoney, M. W. (2009). Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, *6*(1), 29–123. https://doi.org/10.1080/15427951.2009.10129177
Lu, Y., Luo, X., Polgar, M., & Cao, Y. (2010). Social Network Analysis of a Criminal Hacker Community. *The Journal of Computer Information Systems*, *51*(2), 31.
Mahmood, M. A., Siponen, M., Straub, D. W., Rao, H. R., & Raghu, T. S. (2010). Moving Toward Black Hat Research in Information Systems Security: An Editorial Introduction to the Special Issue. *MIS Quarterly*, *34*(3), 431–433.
Morgan, S. (2020). Cybercrime To Cost The World $10.5 Trillion Annually By 2025. *Cybercrime Magazine*. https://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/
Morselli, C., Décary-Hétu, D., Paquet-Clouston, M., & Aldridge, J. (2017). Conflict Management in Illicit Drug Cryptomarkets. *International Criminal Justice Review*, *27*(4), 237–254. https://doi.org/10.1177/1057567717709498
Motoyama, M., McCoy, D., Levchenko, K., Savage, S., & Voelker, G. M. (2011). An analysis of underground forums. In *ACM SIGCOMM Internet Measurement Conference* (pp. 71–79).
Ogbanufe, O., & Kim, D. J. (2018). "Thanks for Sharing": Using Hacker Forum Data for Prediction of Knowledge Sharing and Withholding Behaviors. In *ICIS 2017: Transforming Society with Digital Innovation*.
Samtani, S., Chinn, R., Chen, H., & Nunamaker, J. F. (2017). Exploring Emerging Hacker Assets and Key Hackers for Proactive Cyber Threat Intelligence. *Journal of Management Information Systems*, *34*(4), 1023–1053.
Tzanetakis, M., Kamphausen, G., Werse, B., & von Laufenberg, R. (2016). The transparency paradox. Building trust, resolving disputes and optimising logistics on conventional and online drugs markets. *International Journal of Drug Policy*, *35*(2016), 58–68. https://doi.org/10.1016/j.drugpo.2015.12.010
Yue, W. T., Wang, Q. H., & Hui, K. L. (2019). See no evil, hear no evil? Dissecting the impact of online hacker forums. *MIS Quarterly: Management Information Systems*. https://doi.org/10.25300/MISQ/2019/13042