Automate the Analysis of Online Underground Market

Mrwan BenIdris[1], Qing Ye[2], Sindhura Athuluri[3] and Abdelfattah Elbarsha[4]

[1] Department of Computer Science, University of Benghazi, Benghazi, Libya
[2,3] Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506, USA
[4] Health Informatics Department, University of Benghazi, Benghazi, Libya
{[1]mobenidris, [4]abdelfattah.elbarsha}@uob.edu.ly, {[2] qiye, [3]sa0060}@mix.wvu.edu

DOI: 10.29322/IJSRP.14.02.2024.p14630
https://dx.doi.org/10.29322/IJSRP.14.02.2024.p14630

Paper Received Date: 14th January 2023
Paper Acceptance Date: 17th February 2024
Paper Publication Date: 21st February 2024

Abstract- In the underground market for cyber-criminals, forums are extensively utilized to establish trade relationships and facilitate the exchange of various illegal items, resources, and crime-related services. Consequently, these underground forums contain a wealth of crucial resources for comprehending cybercrime. Given the significant role played by underground forums in the cybercrime ecosystem, analyzing these forums can yield valuable insights into cybercriminal activities. In this study, we focus on addressing three specific issues by utilizing data collected from two distinct underground forums: BlackHat World and HackForums. These issues include: (1) Identifying the product categories by extracting information from post titles. (2) Determining the economic model of the products and calculating their revenue. (3) Assessing the trustworthiness of sellers.

Index Terms- Cyber Criminal, Automate Analysis, Machine Learning, Fake Reviewers Underground Markets.

I. Introduction

Cybersecurity, a branch of information security in the realm of computers and the Internet, aims to safeguard networks, computers, programs, and data from unauthorized access, damage, or attack. In today's world, cybersecurity issues are prevalent as the number of cyber threats continues to rise rapidly. These threats are fueled by an underground economy of online crime that exploits vulnerabilities in personal information protection and network security. Cybercriminals operate with a clear division of labor and multiple value chains, accumulating significant illegal income while endangering public property safety. As the Internet's underground economy expands, cybercriminals establish hidden markets in the dark corners of cyberspace to facilitate trading and communication among various components of this illicit economy. Internet gangsters engage in covert operations to carry out online attacks for maximum profit. Unlike traditional crimes, the underground economy on the Internet relies heavily on the Internet itself for illegal transactions, communication, and theft. Illicit markets leverage legitimate Internet services and applications as a platform for their operations. Despite being perceived as hidden, these underground markets are relatively open to attract new participants and enhance efficiency. In this study, we gathered information from accessible underground forums to conduct extensive analysis. We obtained details about the goods exchanged on online forums and devised a method to assess the reliability of sellers. Additionally, we extracted product prices and computed the revenue generated by the underground economy.

II. Related Work

The problem of cybersecurity and the underground economy is a global issue that is becoming more prevalent as online shopping and payment methods become increasingly popular worldwide. The underground economy encompasses activities such as credit card fraud, the leaking of private information, and the abuse of internet resources. Several studies have been conducted to understand and analyze the structure and workings of the underground economy. Tomes and Martin [1] exposed an underground market that aimed to finance fraud through the Internet Relay Chat (IRC) protocol. While they highlighted various trades and associated dangers, they did not provide a systematic analysis of the economy chain. Franklin et al. [2] focused on examining a thriving underground economy that primarily deals with the commercialization of various illicit activities, including credit card fraud, identity theft, spamming, phishing, online credential theft, and the trade of compromised hosts. Holz et al. [3] introduce a technique that enables them to observe the specific type and quantity of data stolen by attackers from compromised machines, rather than just the outcomes of trading this data. Their focus is on a new and emerging threat known as keyloggers, which establish communication with the attacker through dropzones. Dropzones are publicly writable directories on internet servers that serve as exchange points for keylogger data. The researchers' approach involves executing malware in a controlled environment, identifying the location of the dropzone, and collecting the keylogger data themselves if possible. This process is highly automated and was applied to two different categories of keyloggers to demonstrate its practical viability. Amir et al. [4] conducted an analysis of Malaysia's underground economy using a combination of quantitative and qualitative methods. The paper examines factors such as tax evasion, informal employment, and illegal activities to estimate the size and nature of Malaysia's underground economy. The results shed light on its significance within the country's overall economic landscape. Stone-Gross et al. [5] delved into the underground economy surrounding spam campaigns. The paper adopts a botmaster's perspective to understand how large-scale spam campaigns are coordinated within the underground economy. Through interviews with botmasters and analysis of their strategies, the paper sheds light on the economic motivations behind spam activities and provides insights into combating this pervasive issue. Capasso and Jappelli [6] explored the relationship between financial development and the size of the underground economy. The paper employs econometric techniques to analyze data from multiple countries or regions, investigating how financial development affects both formal and informal economic activities. The results provide insights into how financial systems can influence the growth or contraction of underground economies. Berger et al. [7] focused on estimating the size of the underground economy in Greece and its impact on fiscal policies. The paper employs various methods, including statistical analysis and econometric modeling, to estimate the size of the underground economy. The results reveal a significant underground economy in Greece, with a substantial fiscal impact. Ardizzi et al. [8] presented an alternative approach for measuring the size of an underground economy using currency demand analysis. The paper reinterprets existing methodologies and applies them specifically to Italy, providing new estimates for its underground economy based on currency demand patterns. Li and Chen [9] introduced a novel method for identifying top sellers in the underground economy using deep learning-based sentiment analysis. The paper collects data from online platforms associated with the underground economy and applies advanced machine-learning techniques to identify influential sellers based on sentiment analysis of customer reviews. The results offer valuable insights into the dynamics of the underground economy and its key players. Tahmasebi et al. [10] proposed a fuzzy model to estimate the size of the underground economy using structural equation modeling (SEM). The paper applies this model to a specific country or region and provides results that indicate the magnitude of the underground economy based on various indicators and factors. Tedds [11] focused on analyzing Canada's underground economy through various quantitative methods such as national account data, tax evasion indicators, and informal employment statistics. The paper provides insights into the size, composition, and trends of Canada's underground economy, highlighting its implications for policy-making and economic development. Our work stands out from related studies in several ways. Firstly, we focus on identifying product categories by extracting information from post titles. This approach allows us to accurately classify products into specific categories. Secondly, we determine the economic model of the products and calculate their revenue. By analyzing various factors such as pricing strategies and sales volume. Lastly, we assess the trustworthiness of sellers, addressing a crucial concern for online shoppers. Through a comprehensive evaluation process that considers seller ratings, reviews, and sellers' reputation, we aim to provide users with a reliable platform for making informed purchasing decisions.

III. Data Collection & Annotation

We gathered data for our project from two underground forums: BlackHat World and HackForums. BlackHat World primarily focuses on BlackHat search engine optimization (SEO) techniques and has been active since October 2005. HackForums covers a wide range of cybersecurity-related BlackHat topics and non-cybercrime subjects, starting in 2007. To obtain the data, we manually collected 1070 posts from HackForums and 430 posts from BlackHat World. Our focus was on the "Marketplace" sub-forums of both forums. On HackForums, all the posts were related to selling various products such as game-related items, social network-related products, and gift cards. On BlackHat World, most of the posts were offering services like website ranking boosting, SEO services, and link-related services. We collected the primary dataset by extracting details like product content, payment methods, and prices from the posts (shown in Figure 1). We manually excluded closed post data and removed irrelevant information that was not related to underground economy products. Additionally, we chose to ignore data that did not provide enough information, especially those lacking price details.

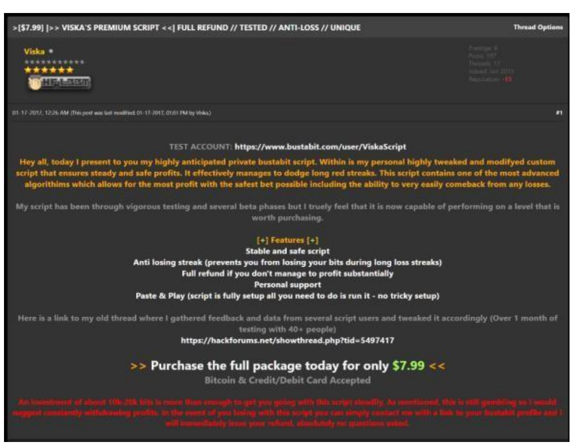


Figure 1. A screenshot showing the webpage used for collecting data.

The obtained data that is presented in Table 1 as an example of our collection process from the website.

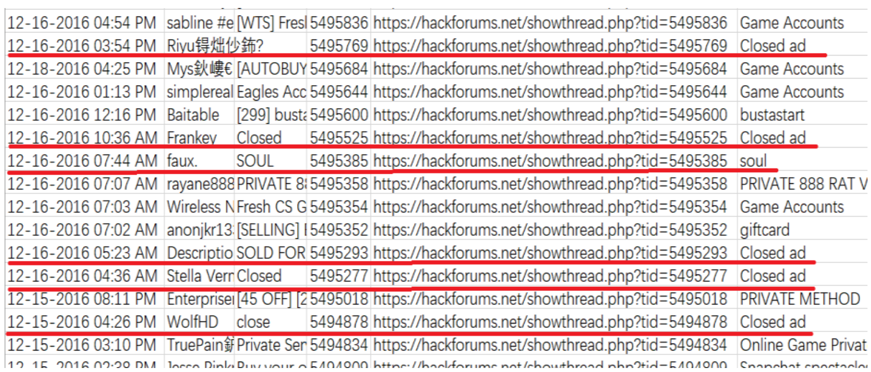
Table 1. Information Extracted from Figure 1.

Product content	Type of payment	Price
Viskas premium script	Credit/ Debit card	$7.99

We manually removed the data of closed posts. Then we find that there are some "meaningless" information, like the 7th line in Figure2, selling "SOUL" is not a really exchange of underground economy product, we choose to delete them. Then we decided to ignore the data which do not have enough information (i.e. no price information) for the future work.

IV. Data Preprocessing

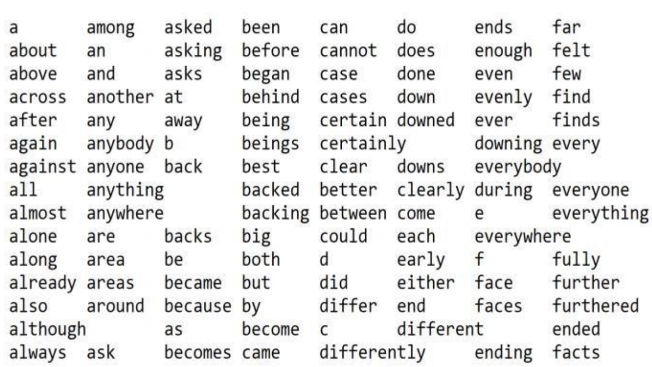
We manually eliminated the data from closed posts and identified certain information that was deemed "meaningless," such as the 7th line in Figure 2 where the sale of "SOUL" did not involve an actual exchange of underground economy products. Consequently, we opted to delete such entries. Additionally, we made the decision to disregard data lacking sufficient information, specifically those without price details, for our future work.



12-16-2016 04:54 PM	sabline #e	[WTS] Fresl	5495836	https://hackforums.net/showthread.php?tid=5495836	Game Accounts
12-16-2016 03:54 PM	Riyu锝焋仯鍗?		5495769	https://hackforums.net/showthread.php?tid=5495769	Closed ad
12-18-2016 04:25 PM	Mys欽嶁€	[AUTOBUY	5495684	https://hackforums.net/showthread.php?tid=5495684	Game Accounts
12-16-2016 01:13 PM	simplereal	Eagles Acc	5495644	https://hackforums.net/showthread.php?tid=5495644	Game Accounts
12-16-2016 12:16 PM	Baitable	[299] busta	5495600	https://hackforums.net/showthread.php?tid=5495600	bustastart
12-16-2016 10:36 AM	Frankey	Closed	5495525	https://hackforums.net/showthread.php?tid=5495525	Closed ad
12-16-2016 07:44 AM	faux.	SOUL	5495385	https://hackforums.net/showthread.php?tid=5495385	soul
12-16-2016 07:07 AM	rayane888	PRIVATE 8:	5495358	https://hackforums.net/showthread.php?tid=5495358	PRIVATE 888 RAT V
12-16-2016 07:03 AM	Wireless N	Fresh CS G	5495354	https://hackforums.net/showthread.php?tid=5495354	Game Accounts
12-16-2016 07:02 AM	anonjkr13:	[SELLING] I	5495352	https://hackforums.net/showthread.php?tid=5495352	giftcard
12-16-2016 05:23 AM	Descriptio	SOLD FOR	5495293	https://hackforums.net/showthread.php?tid=5495293	Closed ad
12-16-2016 04:36 AM	Stella Verr	Closed	5495277	https://hackforums.net/showthread.php?tid=5495277	Closed ad
12-15-2016 08:11 PM	Enterprisei	[45 OFF] [2	5495018	https://hackforums.net/showthread.php?tid=5495018	PRIVATE METHOD
12-15-2016 04:26 PM	WolfHD	close	5494878	https://hackforums.net/showthread.php?tid=5494878	Closed ad
12-15-2016 03:10 PM	TruePain翁	Private Ser	5494834	https://hackforums.net/showthread.php?tid=5494834	Online Game Privat

Figure 2. A screenshot displaying the manually removed data

The preprocessing method holds significant importance in text mining techniques and applications as it serves as the initial step in the text mining process. In this section, our focus was on preprocessing the product title data, with particular emphasis on removing stop words. "Stop words" typically refer to the most commonly used words that offer little value in text analysis. Removing these stop words is crucial as they are ubiquitous and contribute minimally to the analysis. The presence of stop words makes the text appear more cumbersome and occupies additional storage space. Eliminating stop words is essential during the preparatory stages of text data analysis. The most prevalent examples of stop words in text documents are short function words like "a," "the," "at," "which," and "that." These words are considered stop words since they lack meaningful content. While there is no universal list of stop words applicable to all languages, each specific research project typically has its own predefined list. Figure 3 presents a list of stop words.



a	among	asked	been	can	do	ends	far
about	an	asking	before	cannot	does	enough	felt
above	and	asks	began	case	done	even	few
across	another	at	behind	cases	down	evenly	find
after	any	away	being	certain	downed	ever	finds
again	anybody	b	beings	certainly		downing	every
against	anyone	back	best	clear	downs	everybody	
all	anything		backed	better	clearly	during	everyone
almost	anywhere		backing	between	come	e	everything
alone	are	backs	big	could	each	everywhere	
along	area	be	both	d	early	f	fully
already	areas	became	but	did	either	face	further
also	around	because	by	differ	end	faces	furthered
although		as	become	c	different		ended
always	ask	becomes	came	differently		ending	facts

Figure 3. Sections of the list containing stop words.

V. Proposed Method

Automation processing was handled by extracting three attributes from every post in a forum: the payment method, the product being discussed, and its corresponding price. These properties are not explicitly marked, so we employ a supervised learning method. We label a small portion of the data with ground truth and use it to classify the properties. The input to our tool is always a single post, but the output structure varies depending on the task. In this section, we discuss the development of our tools and present evaluation results that assess their effectiveness. We have extracted the product content from every data sample, but it varies significantly. To address this variability, we have categorized the content into different types using regular expressions in both forums. After pre-processing the annotated data, we obtained 1000 original post titles for analysis from both forums. The categories include "Black hat" (Link Related Service, SEO Service, Social Network Related Service, Website Ranking, Wiki Service, Writing Service) and "HackForum" (Accounts, Ewhore, Games, Gift cards, Methods, Service, Social Network Tutorial). To classify each product category mentioned above, we utilized separate Linear SVM classifiers. The input to each SVM classifier is the post title used to extract the corresponding product category: Feature extraction for each product category, training data for each product category, and Test data for each category. Based on our annotated data, we categorized post titles for each product and identified features for each category. For feature extraction, we focused on the most frequently repeated words in post titles as they provide higher information gain compared to other words. To obtain the features for each category of hackforum, a Python code has been developed to extract the top-used words from the 600 post titles.

Category 1 (accounts) → *Features: 'acc', 'accs', 'account', 'accounts', 'email', 'emails', 'edu', 'gmail', 'google'*
Category 2 (ewhore) → *Features: 'ewhore', 'ewhoring', 'ewhores'*
Category 3 (games) → *Features: 'game', 'games', 'gaming', 'ingame', 'xbox', 'playstation', 'ps4', 'steam', 'steamworks', 'battlefield', 'csgo', 'league', 'leagends', 'lol', 'fifa', 'overwatch', 'minecraft', 'gta', 'halo', 'titanfall'*
Category 4 (giftcards) → *Features: 'giftcard', 'giftcards', 'gift', 'gifts', 'egift', 'card', 'cards', 'agc', 'gc', 'amazon', 'amazoncom'*
Category 5 (methods) → *Features: 'method', 'methods', 'bomber', 'bombing', 'script', 'bot', 'aimbot', 'tool', 'tools', 'project', 'projectx', 'turbo', 'traffic',, 'autopilot', 'module', 'modules', 'rat', 'autobuy', 'scraper', 'scrappers', 'source', 'cracker'*
Category 6 (service) → *Features: 'service', 'services', 'boost', 'boosting', 'homework', 'verification', 'proxy', 'proxies', 'proxysh', 'vpn'*
Category 7 (social network) → *Features: 'social', 'network', 'networking', 'twitter', 'instagram', 'facebook', 'follow', 'follower', 'followers' 'youtube', 'subscribe', 'subscription', 'friends'*
Category 8 (tutorial) → *Features: 'tutor', 'tutorial', 'tutorials', 'tutoring', 'ebook', 'ebooks', 'mentor', 'mentorship', 'mentoring', 'learn', 'guide', 'epub', 'book', 'pdf', 'how', 'start', 'business', 'easy', 'income'*

The obtained features for each category of Blackhat from the other 400 post titles:

Category 1 (link related service) → *Features: 'link', 'links', 'backlink', 'backlinks', 'linkwheel'*
Category 2 (seo service) → *Features: 'seo'*
Category 3 (social network related service) → *Features: 'social', 'network', 'networks', 'youtube', 'facebook', 'twitter', 'yahoo', 'blog', 'gmail', 'blogpost', 'blogposts'*
Category 4 (website ranking) → *Features: 'rank', 'ranking', 'boost', 'ranks', 'pagerank', 'search', 'increase', 'up', 'top'*
Category 5 (wiki service) → *Features: 'wiki', 'wikipedia', 'wicked'*
Category 6 (writing service) → *Features: 'write', 'writing', 'writer', 'written'*

Each product category was provided with training data, which consisted of post titles from preprocessed data. Examples of positively and negatively trained data were prepared for each category. Test data, comprising 10 post titles per category, was also collected. Trained data sample for Game Product Category:

[['xbox', 'game', 'sharelicense', 'transfer'],'positive']
[['starbucks', 'giftcards', 'free', 'method'], 'negative']
Trained data sample for Account Product Category:
[['jaruss', 'netflix', 'accounts', 'trusted', 'experienced'],'positive']
[['learn', 'money', 'online', 'easy'],'negative']
Trained data sample for tutorial Product Category:
[['learn', 'money', 'online', 'easy'],'positive']
[['jaruss', 'netflix', 'accounts', 'trusted', 'experienced'],'negative']
Trained data sample for gift cards Product Category:
[['selling', 'website', 'links', 'free', 'amazon', 'types', 'giftcards'],'positive']
[['starbucks', 'giftcards', 'free', 'method'], 'negative']
Trained data sample for methods Product Category:
[['bustabank', 'v13', 'private', 'method', 'btc', 'earned'],'positive']
[['xbox', 'game', 'sharelicense', 'transfer'],'negative']
Trained data sample for website ranking Product Category:
[['super', 'charged', 'wikis', 'boost', 'rankings', 'backlinks', 'google', 'loves'],'positive']
[['learn', 'money', 'online', 'easy'],'negative']
Trained data sample for wiki service Product Category:
[['wicked', 'wiki500', 'wiki', 'sites10k', 'wiki', 'links3', 'urlsfast', 'delivery20k', 'sb', 'blastonly'],'positive']
[['micro', 'niche', 'crusher', 'boost', 'rankings'], 'negative'] *etc…*

The trained and test data for post titles in each category were used to train SVM classifiers. The output of this process was the extracted product. The accuracy achieved in extracting the product was 55.6%. Various types of payment methods were identified through manual annotation such as PayPal, bitcoin, BTC, Amazon gift card, ETH, PM, cash, etc. Most posts contained payment methods in text format. Python code was utilized to extract relevant words from the article text column using search terms like PayPal, BTC, Bitcoin, ETH, etc. To determine if a seller is trustworthy, reviewers' comments (shown in Figure 4) play a crucial role as positive feedback on products or buying experiences indicates trustworthiness. In the underground economy forum, a reputation system can assist in assessing a seller's trustworthiness. Additionally, user ratings (shown in Table 2) can be used for this purpose.

Table 2. The Most Teen Rating User Name

Ranking	User name	Rating	Ranking	User name	Rating
1	Credit/ Debit card	55898	6	mickyfu	14389
2	WizGizmo	34453	7	t0mmy	14207
3	Asif WILSON	31607	8	meathead1234	13705
4	BassTrackerBoats	24331	9	popzzz	13326
5	peter2002	21250	10	Zwielicht	13200

However, it is important to consider the trustworthiness of reviewers themselves and whether they are genuine or fake. Establishing a connection between the reviewer and the user is necessary to ensure reliability. It is possible that sellers may have multiple accounts and write favorable reviews about themselves. For that reason, we used Bipartite Matrix (Two-mode Network) to identify the fake reviewers.

VI. A Case Study

In this case study, we collected reviews from 292 reviewers who had written reviews for 12 sellers. We organized this data into a matrix to create a Two-mode network. Using R programming, we visualized the relationship between the reviewers and the sellers in Figure 4. We noticed that there was a specific group of reviewers who frequently commented together, but it was difficult to identify them. Therefore, we decided to remove any reviewers who did not raise suspicion while keeping the others who might be fake in order to further investigate their connection with the sellers. These reviewers could potentially be users themselves. After classifying the reviewers by clustering them, we observed that many of them exhibited similar behavior.

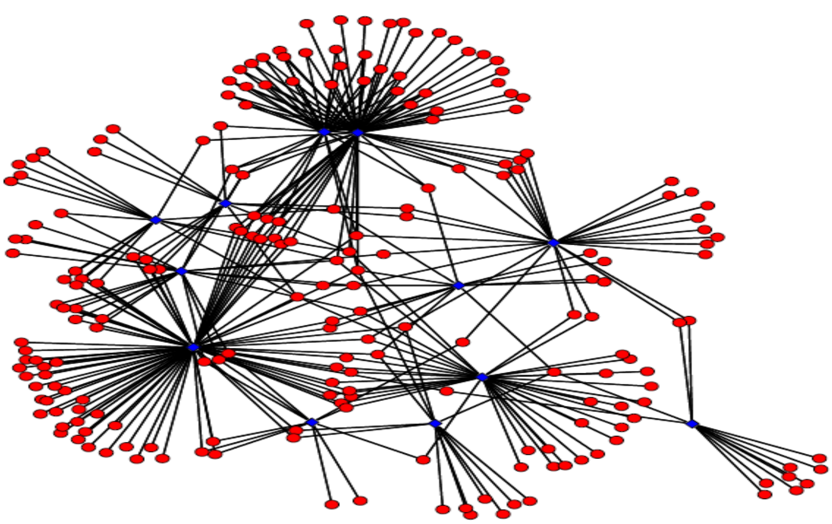


Figure 4. The network connecting reviewers and sellers.

The hierarchical connection between reviewers is demonstrated through the use of a dendrogram diagram. In Figure 5, a group of reviewers is depicted within a red oval, engaging in identical activities, which raises the possibility that they might be fraudulent reviewers. This discovery intrigued us and led us to conduct a more comprehensive examination of this particular group. After performing the Bipartite Matrix and clustering analysis, it became clear that all twenty-seven reviewers enclosed by the red oval had reviewed the exact same posts. This observation suggests that they could potentially be fake reviewers. Nevertheless, additional investigation is necessary to validate this suspicion.

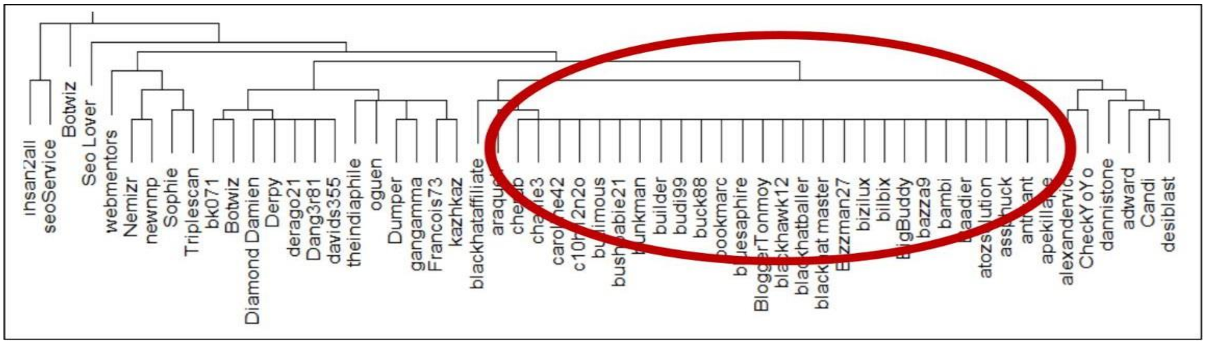


Figure 5. The hierarchical connection between reviewers

We developed a Python program that could extract prices from 500 links related to SEO-Links building. The SEO-Link building category had a total of 3100 posts between June 29, 2008 and November 27, 2017. Table 3 shows that the majority of products being sold belonged to Website Ranking.

Table 3. Distribution of Products Sold in the SEO-Link Building Category

Product	# of views	User	Product	# of views	User
Ranking	1197414	Sarkark21	Ranking	177438	D3t0x
Ranking	880738	Ianist	Ranking	140589	Unknown Zero
Writing	651344	Amarande	Ranking	140114	Hardcorebiker
Google Master	477075	PauloPaul	Ranking	114604	Take Action
Ranking	287664	DX-Generation	Ranking	37660	The profit Bird

We considered two scenarios and learning models: one with a single price and another with multiple prices. For posts with a single price, we used it as is, but for those with multiple prices, we calculated the average. This average helped address the issue of bargaining. For example, if a user had four packages priced at $50, $70, $90, and $100 respectively, we had two cases as shown in Figure 6.

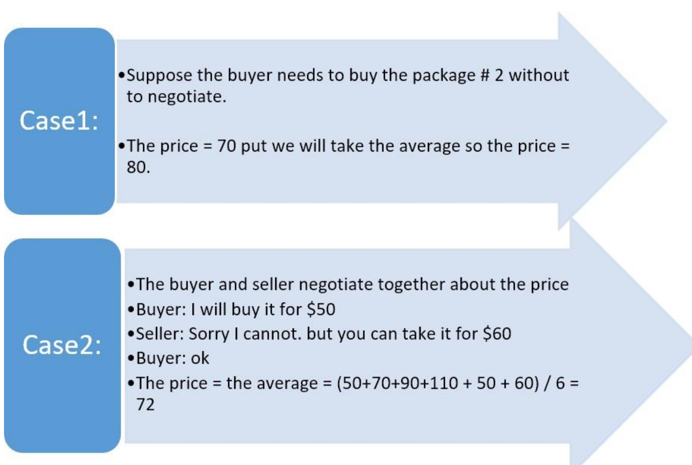


Figure 6. Instances where multiple prices are observed in many cases.

We tested our pricing results by manually evaluating 30 posts from BlackHat World. Unfortunately, we encountered 12 misclassifications resulting in a 40% error rate. We were unable to calculate revenue for BlackHat World and HackForums due to various reasons. Firstly, the prices were not consistent as many posts had multiple prices. Additionally, our links from BlackHat World were all from one marketplace - SEO-Link building - which is one of fourteen other marketplaces mentioned in Table 4. Lastly, bargaining between sellers and buyers took place in different parts of the post, making it difficult to determine the exact price.

Table 4. Marketplace Overview

Product	# of views	User	Product	# of views	User
BHW Marketplace rules and how to post	9	523	Miscellaneous (MISC)	830	196223
Affiliate programs - CPA networks	119	11450	Proxies For Sale	117	70793
Content / Copywriting	324	112459	**SEO - Link building**	**3121**	**884940**
Domains & websites for sale	181	7189	SEO - Other	115	44510
Hosting	143	28477	SEO - Packages	218	84020
Hot Deals	55	2531	Social Media	553	227303
Images, logos & videos	136	21200	Web Design	211	83644

VII. Limitation

One limitation of our study is the reliance on manual evaluation of posts from BlackHat World. This approach introduced subjectivity and potential human error in the classification process, as evidenced by the 40% error rate in misclassifications. Another limitation is the inability to calculate revenue for BlackHat World and HackForums. The inconsistent pricing information, with multiple prices mentioned in many posts, made it challenging to determine accurate revenue figures. Furthermore, the presence of bargaining between sellers and buyers in different parts of the post posed a challenge in determining the exact price. This ambiguity made it difficult to accurately assess pricing patterns and calculate revenue. Overall, these limitations highlight the need for more robust data collection methods and standardized pricing information to improve accuracy and reliability in future studies on pricing analysis in online forums like BlackHat World and HackForums.

VIII. Conclusions

This study aimed to address three specific concerns regarding underground forums: identifying different types of products, understanding the economic model of these products, and evaluating the reliability of sellers. SVM classifiers were used to categorize product types based on post titles, resulting in an accuracy rate of 55.6%. The Bipartite Matrix analysis was employed to identify potential fake reviewers who displayed similar patterns of behavior. However, further investigation is necessary to confirm this suspicion. Additionally, we extract prices from 500 links associated with SEO-Link building, revealing that the majority of products being sold were related to Website Ranking. Overall, this research provides valuable insights into the underground forum ecosystem and emphasizes the need for ongoing exploration in this field.

References

[1] T.Cymru, “The underground economy: priceless,” login, vol. 31, no. 6, December 2006.

[2] Franklin, J., Perrig, A., Paxson, V., & Savage, S. (2007). An inquiry into the nature and causes of the wealth of internet miscreants. Ccs, 7, 375-388.

[3] Holz, Thorsten, Markus Engelberth, and Felix Freiling. "Learning more about the underground economy: A case-study of keyloggers and dropzones." Computer Security–ESORICS 2009: 14th European Symposium on Research in Computer Security, Saint-Malo, France, September 21-23, 2009. Proceedings 14. Springer Berlin Heidelberg, 2009.

[4] Amir, A., Masron, T. A., & Ibrahim, H. (2010). An analysis of the underground economy in Malaysia. In ICBER 2010 Conference Proceedings Paper (Vol. 101).

[5] Stone-Gross, B., Holz, T., Stringhini, G., & Vigna, G. (2011). The Underground Economy of Spam: A Botmaster's Perspective of Coordinating {Large-Scale} Spam Campaigns. In 4th USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET 11).

[6] Capasso, S., & Jappelli, T. (2013). Financial development and the underground economy. Journal of Development Economics, 101, 167-178

[7] Berger, W., Pickhardt, M., Pitsoulis, A., Prinz, A., & Sardà, J. (2014). The hard shadow of the Greek economy: New estimates of the size of the underground economy and its fiscal impact. Applied Economics, 46(18), 2190-2204.

[8] Ardizzi, G., Petraglia, C., Piacenza, M., & Turati, G. (2014). Measuring the underground economy with the currency demand approach: A reinterpretation of the methodology, with an application to i taly. Review of Income and Wealth, 60(4), 747-772.

[9] Li, W., & Chen, H. (2014, September). Identifying top sellers in underground economy using deep learning-based sentiment analysis. In 2014 IEEE joint intelligence and security informatics conference (pp. 64-67). IEEE.

[10] Tahmasebi, M., & Rocca, M. (2015). A fuzzy model to estimate the size of the underground economy applying structural equation modeling. Journal of applied economics, 18(2), 347-368.

[11] Tedds, L. M. (2018). The underground economy in Canada. In Size, causes and consequences of the underground economy (pp. 157-176). Routledge.

Correspondence Author – *Mrwan BenIdris, mobenidris@uob.edu.ly*