Then and Now: On The Maturity of Cybercrime Markets

A study on regulation enforcement of forums in the online underground trading economy

Marco C. Corradin

A thesis presented for the degree of
Master of Computer Science

Distributed and Embedded Security Group (DIES)
Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS)
University of Twente
Netherlands
`m.c.corradin@student.utwente.nl`

UNIVERSITY OF TWENTE.

UNIVERSITÀ DEGLI STUDI DI TRENTO

Summary

Cybercrime is often in the news and at the attention of the scientific literature as the source of huge financial losses or the infection of large numbers of user machines becoming part of a botnet. These activities are often massive, and are supported by infrastructures and services that are reportedly served by an *active underground economy*. Yet, the current understanding of this phenomenon is that the markets underlying the cybercrime economy are by design fraught with problems and cannot possibly sustain the economy the effects of which we observe and read about everyday. This thesis presents a systematic analysis of an online underground black market, namely **Carders.CC** in which we assess the potential differences between markets that are susceptible to scammers (IRC markets) and markets that implement mechanisms to reduce this problem (forum markets). We find that cybercrime markets evolved from an equivalent of IRC markets to a strictly regulated state that may greatly favor market and trade efficiency does not hold for **Carders.CC**. This cybercrime market shows a total market failure; reputation mechanisms are not implemented correctly and hierarchy rewarding communities are not properly enforced. As a result scammers operate and move freely in the market making them indistinguishable from normal users. Despite the distrusting nature of criminals (seen as normal users in these markets), they are not able to distinguish “good” users from “bad” users resulting in a failed market. We therefore conclude that we virtually find no differences between IRC markets and badly regulation enforced forums.

Acknowledgements

I would like to express my appreciation to my supervisor at the University of Trento, Luca Allodi. He provided me with guidance during my research in terms of motivation, valuable feedback and expanding my current knowledge. I would also take the opportunity to thank my supervisor from the University of Twente, Damiano Bolzoni. He provided me with valuable additional thoughts on matters, critical feedback and eye-opening suggestions.

Additionally, I would like to thank the whole committee and staff that made it possible for me to conduct my research at the University of Trento. In particular prof. Fabio Massacci from the University of Trento and prof. Jan Schut from the University of Twente who made this research possible by means of their dedicated commitment.

Furthermore I owe particular thanks to Erasmus, providing me with a scholarship together with the Erasmus Social Network in Trento for organizing interesting and entertaining trips.

During the period of my research I experienced less inspiring times and difficulties regarding work pressure. Therefore I would like to thank my family and friends for providing me with an invaluable amount of love, care, understanding and support.

Finally, I would like to take the opportunity to thank everyone whom directly or indirectly supported me in any way by contributing to my master's thesis.

Gratefully yours,

Marco Claudio Corradin

Contents

List of Abbreviations

C&C	Command and Control
CaaS	Crimeware-as-a-Service
DGA	Domain Generation Algorithm
EaaS	Exploit-as-a-Service
GUI	Graphical User Interface
IRC	Internet Relay Chat
ISP	Internet Service Provider
OUSDM	Online Underground Social Dynamics Model
PHP	Hypertext Preprocessor
PSC	Pay Safe Card
RNG	Random Number Generator
SMF	Simple Machines Forum
SQL	Structured Query Language
UNIX	Uniplexed Information and Computing System
URL	Uniform Resource Locator
USD	United States Dollar
XAMPP	X (“Cross-Platform”), Apache HTTP Server, MySQL, PHP and Perl

List of Figures

List of Tables

Chapter 1

Introduction

Cybercrime is gaining more and more momentum as a source of threats for final users. Not only credit card, banking and financial frauds are continuously reported in the news and often studied in the literature [3, 14], but recent work has uncovered a whole infrastructure of services that are available to cybercriminals to deploy their attacks [15, 2, 26]. Exploitation tools, automated redirection of user connections to arbitrary domains [36], and trading of new malware or vulnerabilities are only examples of a multitude of *measured* effects of what is notoriously called "cybercrime". These infrastructures and services, on the other hand, must be sustained and provided by an underlying economy.

Market design is a problem of great interest in economics, as a successful market necessarily involves an equilibrium of forces that on one side encourages trading, and on the other discourages "cheaters". Obviously, a market where everybody cheats is not a sustainable market and is doomed to fail because nobody would initiate a trade. Cybercrime markets represent, intuitively, a fascinating case study for this problematic. This is not only because of the untrustworthy nature of *criminals* but also because these markets are typically run online which makes the criminals anonymous to a certain degree. Emphasizing on the untrustworthy nature of criminals, how can they trust other criminals in such a way that after the payment they will get access to the promised service? And even if the buyer gets *something*, how can he/she be sure that the requested product meets their initial expectation?

Questioning the previous results from [14], Herley and Florêncio showed that IRC cybercrime markets may be no different from the notorious *markets for lemons* theorized by Akerlof [1], where effectively the *asymmetry of information* between the seller and the buyer is such that "bad sellers" are incentive in participating in the market to the point that it makes no sense for the "good sellers" to remain active. In Akerlof's case, a "bad seller" is a seller that trades "lemons". A "lemon" is a defective car that is advertised as a good one. If the customer cannot assess the quality of the car before buying it (e.g. because she knows little about cars), then she will buy the cheapest she can find on the market. And because "lemons" are cheaper to the seller than good cars are, "good sellers" are ultimately forced out of the market. In Herley and Florêncio's case, a "lemon" was a credit card number with allegedly a certain amount of USD ready to be used by the buyer. Discerning "good sellers" from "bad sellers" is therefore a critical point of a market design. Herley and Florêncio clearly demonstrated that this is virtually impossible in the IRC cybercrime markets.

Yet, empirical evidence from numerous studies shows that attack tools traded in these markets [2, 15, 46] and economic/financial losses caused by cybercrime [3] are measurable and real. How can these observations be reconciled with the current understanding of the cybercrime markets? Current markets are run under a different structure than the IRC markets of Herley and Florêncio were: rather than anonymous, free-to-join, unregulated communities of criminals, modern cybercrime markets are run as virtual forums [26, 2, 52, 34]. Forums provide an easy way for the community administrators to control the flow of users into the community and to enforce a number of rules through moderation that can be aimed for mitigating the issues of information asymmetry [52] in a coherent market design structure. Therefore, proper regulation can be the key to a successful market.

In this thesis we reproduce the findings of Herley and Florêncio and expand on their research by testing whether a known online underground market **Carders.CC** is no different than an IRC market and how this related to the successfulness of scammers in terms of the amount of contracts they finalize. **Carders.CC** is a (failed) market for credit card numbers and other illegal goods, whose database leaked in 2011. We are able to reproduce and analyse the market in its entirety and we show how the systematic failure of its regulatory mechanisms led to a market where the so-called rippers – which is underground slang for scammers – and "legitimate users" are indistinguishable one from the other.

The thesis proceeds as follows: Chapter 2 discusses current relevant literature and sets the stage for the discussion of our research. The problematic nature Herley and Florêncio presents results in our main research question which we subsequently break into two sub questions in Chapter 3. Chapter 4 presents the *Carders.CC* data and describes the market and the designed regulatory mechanisms which is needed to gain a clear understanding on how the market works, operates and is designed. From this discussion, in Chapter 5 we formulate a number of hypotheses we would expect evidence for in the data if the regulatory mechanisms were properly enforced. The results of this analysis are given in Chapter 6. Finally, we discuss our findings in Chapter 7 where we emphasize on limitations and restrictions, provide suggestions for future research and conclude our thesis.

Chapter 2

Background & Literature Review

There has been much study conducted in the field of cybercrime markets. Several approaches have been proposed from a social, economic and technical background in order to understand more about the dynamics of cybercrime markets. Each of these studies contribute in their own way to the fighting of these specific kind of markets. Current literature on these underground markets can be clustered into two categories: studies that (indirectly) provide factual evidence of the workability of the underground markets, and studies that analyse the structure and economics of the markets.

In order to find corresponding literature we consulted Scopus, Elsevier and Google Scholar. Scopus is a bibliographic database containing abstracts and academic journal articles. The University of Twente has a subscription contract with Scopus which made it possible to search and view academic journal article abstracts and citations from their database. While Scopus is mainly a life, social, health and physical sciences related abstract and citation database, Sciencedirect is a full text journal article database. The company Elsevier that publishes academic literature combined these two databases into the SciVerse platform. The first searches were performed on both search engines but after noticing redundancy further searches were excluded since they search the same underlying database. Furthermore the non-subscription platform Google Scholar was consulted which indexes full text scholarly literature. Alternatively, Elsevier's Scirus free literature search engine has been taken into consideration. However, Scirus announced its closure for early 2014[1]. Therefore the reviewed literature is based on searches in one subscription and one free literature database since other subscription services were no option. After several searches SciVerse showed redundant results for the same given string input in Google Scholar. If there were any results on Scopus, Google Scholar was able to find these too. Therefore the remaining search strings were performed in Google Scholar.

In terms of categorizing and determining relevance of literature in underground black markets we base our first selection criteria on the subject of the literature. Literature on cybercrime markets can be divided into two main categories: studies that (indirectly) provide factual evidence of the workability of

[1] Announcement Scirus retires early 2014, http://www.scirus.com/

the underground markets, and studies that analyse the structure and economics of the markets. Trying to understand selling concepts, turnover values and economic dynamics of an online underground black market seems interesting into relating financial incentives of miscreants and the reason of their participation. In terms of market structure, the modus operandi of the participators learns us what kind of people are involved in cybercrime markets and how they operate. Here, studies regarding the accessibility, lifetime span, amount of black markets and other market structure studies are essential for understanding the infrastructure cybercriminals operate in. Studies in examining the illegal goods that are being traded in the markets give insight on what these goods mean, how they operate, spread, infect and sell. The quality of these goods also show that cybercrime markets are trading well-working products. In other words, by dividing the literature into three categories, namely Empirical Studies, Economics Studies and Social Studies, we learn more about cybercrime markets and their associated properties.

Finally, there are studies present that fully focus on fighting the underground market by manipulating the economy whereas other studies provide insights in cybercrime markets on one specific problem. All these studies can be used for fighting the underground market however, the researches that will purely focus in proposing a method to take down an underground market is categorized into fighting the underground market.

Results from search strings are evaluated by title selection. If a title indicates to cover a certain area in one of these three categories the study will be selected. The abstracts of the selected titles are being studied and if the authors seem to cover a problem in one of our pre-defined category we will continue to further examine the selected study by reading the full article. If in any way the selected study fails to meet our selection criteria by being unrelated or irrelevant the study will be unselected. For searching we consulted the article titles, abstract text and keywords fields in Scopus while for Google Scholar we have consulted every possible field, including an in-depth article search that searches also the content of an article.

The following papers are reviewed and selected through references of active members in the field of cybercrime analysis: [1], [2], [3], [4], [9], [10], [12], [15], [16], [17], [24], [25], [26], [27], [32], [33], [36] and [46]. All other selected literature can be found in Table 2.1.

Timeline Literature

The timeline in Table 2.2 shows the publication dates of the selected literature. This timeline shows that the majority of the selected literature has been published between 2008 and 2013. For one, this is in order to maintain novelty and to prevent researching outdated literature. Secondly, this shows the nature and popularity of cybercrime markets. In earlier years IRC markets arose on the internet and have changed over the years into more structural and hierarchical environments like forums. This shows that activities in the underground market remain but the needs of cybercriminals change. Since crybercrime markets were a problem then and still are we find it relevant to include literature from before 2008 into our selection. Additionally, these references have laid the fundamentals in the field of cybercrime markets and are one of the first researches that

SciVerse Scopus

Search String	Results	Papers Selected
Underground + internet + market	13	1

SciVerse ScienceDirect

Search String	Results	Papers Selected
Underground + internet + market	1892	2 ([7], [43])

Google Scholar

Search String	Results	Papers Selected
Underground + internet + market	61.800	7 ([13], [14], [18], [31], [39], [41], [54])
Classification + Underground + Markets + Vulnerabilities	21.900	2 ([38], [37])
Report + internet + underground + economy	42.700	1 ([47])
Internet + underground + economy	52.800	10 ([6], [11], [19], [23], [34], [35], [42], [44], [45], [49], [50])
Hackers + behaviour + in + the + underground + economy	13.300	6 ([5], [8], [20], [21], [40], [52])
Economics + Computer + Hacking	23.400	5 ([22], [29], [28], [30], [48])
Parser + for + underground + economy	4.700	1 ([53])
Malware + analysis + in + underground + black + markets	783	1 ([51])
Total	223.288	36

Table 2.1: Search strings for finding corresponding field literature

gave insight on the quality of traded goods, economics studies and social studies containing motivation, behaviour and social analysis. Many of the literature in later years is based on these theories and researches.

[54]						
[48]						
[38]					[53]	
[32]		[50]			[51]	
[30]		[37]			[40]	
[29]		[31]			[33]	
[16]		[28]	[49]	[46]	[24]	[52]
[14]	[39]	[27]	[41]	[45]	[22]	[47]
[12]	[36]	[20]	[35]	[44]	[15]	[43]
[10]	[25]	[18]	[19]	[34]	[7]	[42]
[9]	[23]	[13]	[11]	[21]	[5]	[26]
[1]	[3]	[8]	[6]	[17]	[4]	[2]
	2008	2009	2010	2011	2012	2013

Table 2.2: Timeline of the selected literature publishing dates

Methods and Data Set

Each one of the chosen literature that is being described in this chapter is summarized into one overview which can be seen in Table 2.3. This overview shows that the majority of the literature we discuss in this chapter is based on forum analysis and some on IRC markets. Observing the publishing dates of the literature in the timetable we conclude that older literature was more focused on IRC channels whilst more to date studies rely more on forums for their data set. Methods used in the literature are mostly unique and provide us with more information about what kind of methods are used to study cybercrime markets. Altogether this gives an overview of research that has been done in different areas, what kind of data they have used and how they tried to achieve to solve a specific problem in the field of cybercrime markets.

2.1 Efficient Markets

2.1.1 Empirical Studies

In 2008, Holz, Engelberth, and Freiling [23] conducted a first empirical study of collecting data from more than 70 keylogger dropzones. They make a distinction between two main contributions, namely a method to analyse a large scale of credential-stealing attacks in a highly automated fashion and analysing the results of this applied method. They analysed the dropzones of two different keylogger families: Limbo/ Nethell and ZeuS/ Zbot/ Wsnpoem. By infecting a virtualized and controlled environment they were able to extract more than 2,000 unique keylogger examples. Most of the dropzones were located in Asia or in Russia, but also some in the United States in which the average lifetime of a dropzone was approximately 61 days.

Reference	Data Set					Methods used
	IRC	Forum	C&C Channel	Social Media	Other	
2.1.1 Empirical Studies						
[23] Holz, Engelberth, and Freiling			X			
[15] Grier et al.					Malicious URLs	
[43] Sood and Enbody		X				CaaS Business Model
[17] Gundert and Berg					Exploit packs	
[2] Allodi, Kotov, and Massacci						
[44] Stone-Gross et al.			X			
[42] Sood, Enbody, and Bansal		X				
[11] Fallmann, Wondracek, and Platzer	X					Probing and Crawling
2.1.3 Social Studies						
[47] Symantec	X	X				
[13] Fossi et al.	X	X				
[49] Vömel, Holz, and Freiling		X				
[50] Yang et al.		X				Crawling
2.1.3 Social Studies						
[51] Yip, Shadbolt, and Webber		X				Social Network Analysis Techniques
[53] Zhao et al.				X		OUSDN
[34] Motoyama et al.		X				
2.1.4 Fighting the Underground Economy						
[31] Li, Liao, and Striegel						Uncertainty Economic Principles
[37] Radianti, Gonzalez, and Rich						Dynamic Simulation Model
[8] Cárdenas et al.						

Table 2.3: Review of literature with corresponding methods and data sets

Efficiency is key for underground markets to increase workability. One way to achieve this has been shown by Grier et al. [15] in which they introduced the Exploit-as-a-Service (EaaS) model. In the EaaS model the cybercriminal can rent a service in which the contractor provides a full service that supports all the necessities to infect computers for the buyer. They studied the impact of exploit-as-a-service by visiting 77,000 malicious URLs which led to more than 10,000 distinct binaries. By running these binaries in a contained environment they were able to analyse these binaries in a safe way. Their results show that many of the malware collected make use of drive-by downloads. Blackhole is responsible for 29% of the malicious URLs, followed by Incognito. Grier et al. showed that underground markets are adapting to the needs of participators by implementing refined business models. This means that underground markets show a clear change over the past last years. Paired with these changes comes an increased participation of the market due to the fact that the entry level of skills is considerably lower than before.

More efficient markets have been studied by Sood and Enbody [43]. They have shown evidence that also the Crimeware-as-a-Service (CaaS) model is applied to cybercrime markets. Where EaaS only provides a full service for exploitation and infection of machines, CaaS provides a full service that facilitates the cybercriminal with all the resources he/ she may need. This means that the service offered contains all necessary tools and services to commit the cybercrime like frameworks, settings, machine infections and identity masking. An underground market with the CaaS business model applied provides a more refined and automated trading process resulting in increased profits. Sellers will advertise a full service in which there are several crimeware services that a buyer can rent, like bot shops, DDoS service outlets, bulletproof services, code obfuscation services, plastique shops, credit card services, access logs, automated crimeware frameworks and phishing services. As we have seen with the EaaS business model also CaaS shows that underground markets are changing. Where in early years separate goods were being sold this business model shows us that the entry level for a person to commit a cybercrime is low. No knowledge or hacking skill is needed any more to commit a cybercrime. All that is required for a potential cybercriminal is to license a service with a CaaS provider and they are able to perform any attack as they fit. This refined and automated way of offering crimeware services with no need of knowledge attracts a bigger audience to the underground markets.

With these so-called “as-a-service” models the technical knowledge needed to commit a cybercrime is very low. However, markets that do not apply these schemes turn out to offer products that do not need a high level of technical knowledge. Gundert and Berg [17] shed light on the installation procedure and buyer operations of exploit packs. They installed several exploit packs and found out that buyers require basic knowledge of Apache, MySQL, PHP and a general familiarity with UNIX . Many of these exploit packs turned out to have a GUI which makes usage even easier. By testing Eleonore, Crimepack, Icepack, Adpack, Gpack, Siberia, Blackhole, ZoPack, Sploit25, Fragus, Incognito and Yes they analysed administration pages, configuration files and infection processes. Another study on the quality of offered products in cybercrime markets has been conducted by Allodi, Kotov, and Massacci [2]. By analysing offered exploit kits they measure the resiliency and efficacy of cybercrime tools in delivering attacks. Numerous other studies showed the technicalities behind these infections processes [36, 26] and the creation of botnets [44, 16]. A similar line of research also gave precise insights on the mechanics of spam [25] and diffusion of attacks [9].

Quality of goods in terms of offered botnet infrastructures have been showed by Stone-Gross et al. [44]. They have published an article in 2011 performing a related botnet analysis in 2010 where they were able to control the Torpig botnet infrastructure for 10 days. This was possible by exploiting a technique due to the fact that Torpig was designed in such a way that it uses domain flux, meaning that a bot uses a domain generation algorithm (DGA) to compute a list of domain names. These domain names point to the C&C channel of the botnet master. This dynamic generation of domain names was being designed by botnet creators in order to prevent a shutdown of the botnet and to avoid a single point of failure. However, the Torpig controllers failed to register the future C&C channel domains making Stone-Gross et al. receive full control of all the Torpig bots by registering these domains in advance. From 25 January 2009 to 15 February 2009 they were able to collect 70 Gbytes of data. By analysing this data they discovered that form data is the most sent data item by bots followed by emails, windows passwords, pop accounts, HTTP accounts, SMTP accounts, Mailbox accounts and as last FTP accounts. By analysing submission header fields they estimated the live botnet size population at approximately 182,000 machines. They came to the conclusion that the Torpig's botnet goes after digital goods that are easy to monetize in the underground market. This was due to the fact that a typical Torpig configuration file consisted of roughly 300 domains belonging to the financial sector like banks. In these 10 days they were able to receive 8,310 financial institutes' accounts like PayPal, Poste Italiane, Capital One, E-Trade and Chase. Additionally, they obtained 1,660 unique credit- and debit-card numbers whereof 49% of these cards originated from the United States. The same authors, Stone-Gross et al., did more analysis on the impact of botnets in March 2011 [45]. Unique in their research is the fact that they analysed a large-scale botnet which is used for spam campaigns from the perspective of the botnet master. This gave them the opportunity highlight on quality of email address lists, the effectiveness of IP-based blacklisting and the reliability of bots. In this in-depth analysis of spam orchestrated by the Pushdol/ Cutwail botnet Stone-Gross et al. obtained access to 13 C&C servers used by actual botnet operators. They observed phishing, malware, diplomas, pharmacies, money mule and real estate spam campaigns good for over 500 billion spam sent. Additionally they observed a web forum spambot.biz known for its devotion to spam operations. This forum is for 91.3% Russian spoken and in order to gain access to the forum one needs to be recommended by at least two trusted members within the community, who are part of the top echelon of the spam community. By observing the traffic and botnet size in combination with the prices and information observed from spambot.biz they estimated the profit of the Cutwail botnet gang's somewhere between $1.7 million and $4.2 million from June 2009 until 2011. Again, these studies shows the efficacy of underground market products and how illegal goods are being supplied into the cybercrime markets.

More analysis on cybercrime markets has been done by Sood, Enbody, and Bansal [42] in 2013 where we clearly see the change of the underground market over the years. Whereas the cybercrime markets in the early days offered more standalone services, the frameworks described by Sood, Enbody, and Bansal show that a combination of services are being offered in frameworks. Still, IRC markets and web forums are mainly used as a trading channel in the underground market. They divide the underground related cybercrimes into three cycles. The first cycle contains the buying cycle. Complete frameworks are being rented like the ZeuS or SpyEye framework. In the second cycle they use this framework to infect victims. Finally, in the third cycle the owner will receive critical information from the bots to the C&C channel. In order to hide malicious code there are real-time underground websites like styx-crypt.com that guarantee obfuscation and morphing services. Some new forums found in this recent study are darkcode.com, madtrade.org and exploit.in. More online services like madc.su are now available to verify the authenticity of a credit or debit card. This provides the buyers to indicate the value of offered digital goods more accurately. Steklo.cc provides services for faking bills, IDs and even passports. Credit card skimming services are provided by validhshop.su and also through some web forums. The mitigation of different websites for a service is a new trend in cybercrime markets over the last year. Normally all these services were offered through advertisement on IRC markets and/ or web forums. This research confirms that the underground market keeps evolving and changing.

The infrastructure of the underground markets get more refined throughout the years. While in the past single item selling was more common we now see a shift in this business model by witnessing complete infrastructures being offered. These CaaS and EaaS models make it possible to ease the participation of people into cybercrimes since less knowledge is needed now.

Furthermore there are two main areas where cybercrime markets are being hosted, namely web forums and IRC channels. From the data that is available through either data leaks, probing by Fallmann, Wondracek, and Platzer [11] or other techniques, statistics result into showing several characteristics of the forum. These forums have different participators, ranging from rippers and spammers to verified vendors and VIPs. With the change from IRC channels to the need of a more hierarchical infrastructure such as the forum and a change of business models over the years from single item selling to CaaS and EaaS models it can be stated that the underground market is changing and evolving.

2.1.2 Economics Studies

Many studies try to analyse the underground markets from an economic perspective. The annual internet security threat report by Symantec [47] that was published in 2013 estimated the value of the goods offered throughout 2012 in cybercrime markets at an amount of $276 million. The statistics presented in this report show that cybercrime attacks are increasing and mitigated from different platforms. This dynamic change is due to the increase of mobile platforms in the last years. There were more vulnerabilities found on both mobile platforms and computer environments in 2012 than in 2011. Even though the growth of vulnerabilities in mobile platforms rises it is still a considerable lower amount (415) than the software vulnerabilities in computer environments (5,291). However, it is remarkable that mobile malware is increased by 58%. The amount of spam sent has decreased just like the amount of botnet zombies. However, the 0-day exploits in 2012 have almost doubled.

Fossi et al. [13] analysed the statistics from Symantec and observed several goods and services that were being sold in cybercrime markets. The most popular category was the offerings of credit cards in the underground market. The second most popular was financial accounts which also includes bank account credentials, magnetic stripe skimming devices, online payment services, online currency accounts and online stock trading accounts. Third came advertised goods and services for spam and phishing information. This includes email addresses, email accounts, passwords, scams and mailers. These reports confirm the undeniable fact that the underground market is active and offering working products.

A more in-depth study on credit cards offered in the underground market has been done by Vömel, Holz, and Freiling [49] in which they monitored an IRC channel for credit card advertisements. They monitored the channel #ccpower on the Undernet IRC network and recorded more than 675,000 messages. Unique to their data acquisition technique is the fact that they used a compromised honeypot. A honeypot is designed to be vulnerable and susceptible to attacks in order to get compromised. Once compromised, all activities are being logged and a honeypot owner can learn from the observation of this information. Even though public messages can be read from IRC channels, with a honeypot the researcher is able to monitor all the activity on the honeypot meaning all the public and private messages. They found that credit card related goods were the most popular goods advertised in the IRC channel, followed by cash-out, hacked hosts, bank logins, personal information, PayPal accounts, spam, hack- and scam, equipment and confirmation advertisements. This research gives also a good insight on the terminology used in the underground market.

Allegedly successful markets operating on IRC channels have been analysed by Fallmann, Wondracek, and Platzer [11] and Yang et al. [50]. Fallmann, Wondracek, and Platzer tried to identify underground marketplaces with a system design that uses probes. The so-called IRC sensor probes were used to gather information from IRC channels in combination with several evasion techniques to provide more certainty for the gathering of information. Hosted marketplaces in IRC channels were found by the probes with the strategy to match denoted patterns to channel names and topics. The web forum sensor was designed as a web crawler according to the approach described by Yang et al. which they expanded with several functionalities to adapt to the targeted web forums. One of the techniques implemented contained IP swapping. Some probes were able to only view a certain amount of pages, therefore they implemented a solution that randomly swaps probes with a different IP address. Another technique implemented was the registration to the web forum by the means of an authentication module in order to gain access to these web forums. They found that 4.7% of all IRC channels were related to underground economy operations. They did not elaborated on the collected messages (43 million) but only on the IRC channel commands statistics. The results from the messages collected by their web forum crawlers provided statistics for 11 underground web forum markets, good for 127GB with over one million posts. The main contribution of their research is the fact that they managed to automatically discover markets that are active in the underground world of the internet and to successfully obtain the messages monitored on these mediums.

These studies providing an economical approach towards well working markets provided evidence that there is a considerable amount of revenue and many working products including credit cards.

2.1.3 Social Studies

Studies with a social approach towards the underground markets analyse cybercriminals that operate in successful markets, like Yip, Shadbolt, and Webber [51] and Zhao et al. [53]. They showed that criminals prefer trading in a more secured and hierarchical system to further increase trading efficiency and stability of the market. Because of this increased need of a more structured hierarchy most markets moved towards forums, resulting into studies that try to infiltrate and analyse these forums [34].

Zhao et al. [53] contributed in the field of underground analysis by studying the social dynamics considering both social and user-generated contents. Their so-called Online Underground Social Dynamics Model (OUSDM) contained six fundamental entities and five basic types of unidirectional relationships between them. For adding ranking to the OUSDM they used SocialImpact which consisted of group indices, user indices and string & post indices. In order to implement their system they designed Cassandra. They evaluated Cassandra on livejournal.com, a Russian online social network. Cassandra found that users who talk more in the online social network do not make them more influential. They observed that influences can come from many different causes and their structured method is a new and unique contribution to the study field. In addition to the systematic analysis by Zhao et al. [53], structural analysis has been conducted by Yip, Shadbolt, and Webber [51]. They focused on the social dynamics of cybercriminals by looking at the relationship between cybercriminals. Their method was applied on four social network web forums: Carderplanet, Shadowcrew, Cardersmarket and Darkmarket. By examining the personal messages of these cybercriminals they found out that cybercriminals are willing to trade security for a certain level of efficiency. Cybercriminals demonstrated that they have a strong need for effective communication by repeatedly choosing for a hierarchical system such as a forum. More study on forums has been done by Motoyama et al. [34] in which they conducted a research on the social network make-up for six underground forums. They obtained SQL dumps of databases from the 6 underground forums BlackHatWorld, Carders, FreeHack, HackEl1te, HackSector, and L33tCrew through publicly available leaks. By analysing this data they managed to provide statistics on private messages, which is impossible for crawlers and other public data set gathering techniques to obtain. The way private messages are distributed among users' "associates", which they recall to as fellow members that they are linked too, were being observed for the L33tCrew forum. They found that 70% of their associates were responsible for 70% of their private messages. Group status turns out to be very important for forum members. Verified vendors and VIP members received 2-3 times more response personal messages than other members of the forum. Another finding was that many of the participants also participate in the other forums with the same nickname ranging from 7% to 17% in some forums by matching the same registration email address. This method leaves out the possibility to detect user overlaps when the same users do not use the same email address which means that the percentage given could be higher. The goods that they observed on the Carders forum were from popular to least popular: payments, game-related, credit cards, accounts, merchandise, software/ keys, services, victim logs, mail/ drop services and fraud tools. The other forums do not differ much, they all share the same top 4 goods in a slightly different order. Besides how social degree affects trading they also analysed banned users. On the Carders forum it seemed that more than 20% of their total members were banned. One of the most important reasons for banning a user was due to spamming, followed by duplicate accounts. Some other reasons were misuse, ripping, malware and trade-related reasons.

These social studies do not only approach the problem of emerging underground markets, providing us interesting findings from a different angle, but also show that these markets are offering well-working products. With respect to the structure of the market there is a noticeable shift from unstructured IRC markets towards a more hierarchical environment, namely forums.

2.1.4 Fighting the Underground Economy

There have been several methods proposed to influence the underground market on the internet. Li, Liao, and Striegel [31] assumed that money is the only determining force that influences the financial incentives to participate in the underground economy. With this assumption they proposed an economic approach to take away these financial incentives. Their model is based on injecting uncertainty into a botnet environment where they showed that the higher the uncertainty gets in the financial environment the lower the profits will get for the botnet masters and attackers. Their approach to the underground market is an interesting scheme and attempt to fight the botnet existence.

Another method to combat the growth and existence of the underground market is proposed by Radianti, Gonzalez, and Rich [37]. They attempted to answer different questions such as what affects the success and failure of the underground markets. With the use of their System Dynamics method they gained insights on these wanted observables. They tried to look if vulnerabilities can be disclosed from the black markets by patching. The limitation they encountered is that there are still users that do not apply updates automatically therefore there will be hackers who target these ignorant users. Underground markets that offer illegal goods in terms of software exploits and infections heavily depends on the life cycle of a vulnerability so if this vulnerability can be patched before exploits are being created will influence the market in such a way that will attract less actors and eventually deplete the underground market.

Some other proposals for fighting the underground market has been proposed by Cárdenas et al. [8] in 2009. They provided several proposals in which the supply of offences would be reduced, where one of them is to increase the public protection by sharpening the legislation regarding cybercrimes. Another proposal was private protection where the computer security industry contributes in order to lower the offences by tracking down domains that host malicious content. As last they believe that Internet Service Providers (ISP) can mean a great deal in lowering the offences.

All previously mentioned authors contributed by studying and showing that cybercrime markets are efficient and well-structured. One important factor in these markets is the quality of goods. When high quality products are offered, the stability and value of the market increases. Furthermore a broader scale of cybercriminals are attracted to the underground markets by applying highly refined and sophisticated business models which significantly lower the required level of technical knowledge in order to commit a cybercrime. It is undeniable that cybercrime markets are present, active, providing market stability and offering qualitative illegal goods.

2.2 Inefficient Markets

In Section 2.1 we discussed working and efficient markets. However, running an efficient underground economy in which criminals trade proper goods and services with other criminals is not a trivial exercise. Herley and Florêncio [19] showed that underground markets may mainly feature scammers who try to scam other members of the market. Herley and Florêncio showed that the underground economy is largely a “market for lemons”. This is clearly in contrast with the efficient markets described by the previous authors. The work of Herley and Florêncio was a first step in identifying the failing model of the underground IRC markets. Their findings showed three major shortcomings of those markets also relevant for our work:

1. Users could join the market freely and with an arbitrary identity. Feedback mechanisms (e.g. reputation) on the “reliability” of the users are not therefore enforceable.
2. There is no history of transactions available, meaning that it is impossible to look back at a users’ trades or community-provided feedbacks.
3. The community is largely unregulated and no assurance for the buyer or the seller exist that the trade they are engaging with is a “legitimate” one.

IRC markets are however an “outdated” model for cybercrime markets. Recent markets moved towards a forum-like environment [26, 52], which provides many advantages over the IRC model: first of all, users must register and are therefore assigned a unique ID. The forum structure provides a well-defined technological means for users to leave permanent and easily-searchable feedback on another user, and many forum platforms allow for the assignment of “reputation points” to different users which may directly reflect a members’ role in the community. Finally, a forum can be easily moderated and administered, meaning that an actual *regulation* of the market activities is possible. This makes the case of “forum markets” a completely different one from the IRC markets that have been shown to be irremediably flawed.

Chapter 3

Research Question

As described in Chapter 2 we classified currently conducted research on cybercrime markets into efficient and inefficient markets. Many authors showed that cybercrime markets are offering well-working products and applying well-defined business models. Besides these efficient markets we have also mentioned studies regarding inefficient markets. While there are not many of these studies present, Herley and Florêncio [19] showed that underground markets may mainly feature scammers. They systematically showed that the failure of a market is related with three of their defined features a market should meet when aiming for stability. The main problem is that there is little study conducted on failing markets. While Herley and Florêncio [19] mention three major features why IRC markets fail over forums they (nor others) do not apply this theory to cybercrime forums. Therefore this thesis will expand on their work by providing a detailed analysis of the failure of a "modern" forum community.

By being the first to apply Herley and Florêncio findings on IRC markets to a forum we answer the following question:

Are badly regulated cybercrime forum communities virtually no different than unregulated IRC communities?

To answer our main research question we divide the question into several sub questions. As we have discussed in Chapter 2 we expand on the research of Herley and Florêncio [19] to see whether the shortcoming of IRC markets apply on our chosen forum **Carders.CC** and how they relate to the possible failure of the market. The three main shortcomings of IRC markets are stated by Herley and Florêncio as follows:

1. There is no reputation mechanisms which represents one's reliability within the community.

2. There is no feedback mechanism where one could check the history of transactions from another user.

3. The community is largely unregulated.

From this perspective it is important to check whether these shortcomings are also represented in the forum we examine. To measure reliability of users within a community we try to answer the following sub question:

Sub 1. Is there a working reputation mechanism within the community that determines the reliability of one's trading?

This sub question is closely related to the second shortcoming in our list: no feedback mechanism. While this seems trivial to some, it is important for users that trade illegal goods to remain anonymous to some extend. Since we saw that forum markets erupted from the perspective of having a more hierarchical structure it is self-evident that to achieve this one needs to make a trade-off between anonymity and having a public identity. However, being able to see all the transactions of a user can be a great privacy issue for these cybercriminals since they trade in illegal goods. Therefore many reputation mechanisms are implemented and designed in such a way that they measure the trustworthiness of a user in terms of feedback from other users. Other users can publish information about a successful and satisfied trade with a user and higher their reputation. Of course, lowering one's reputation is also possible when a user is not satisfied. Because of this reason we answer the sub question stated above covering both shortcomings.

For the shortcoming regarding regulation we try to answer the following sub question:

Sub 2. Is regulation being enforced in the forum?

By answering this question we learn whether the forum differs from an IRC market in terms of regulation enforcement. Here we are explicitly interested in regulation that relate to hierarchical mechanisms in the forum that functions to distinguish the "elite" members from the "normal" members.

Should we be unable to confirm both questions this would result in the analysis of a forum that fails to meet the shortcomings stated by Herley and Florêncio and we answer our main research question by concluding that badly regulated forums are virtually no different than IRC markets. Additionally, the findings of this research will provide us with information about the scammers in failing markets. When forum communities are poorly regulated it will provide scammers to freely act within the community. This results in more successful scammers and eventually in failure of the market due to the lemon market principle discussed earlier.

Chapter 4

Market Data & Description

Before jumping into the methodology it is essential to understand certain characteristics and properties of the forum we will be examining. In order to answer our research questions we described in Chapter 3 we will test Herley and Florêncio [19]'s mentioned reputation and regulation features. To get a better understanding of these implemented mechanisms we describe the re-construction of our data, market regulation, user roles, user groups and the 3-Tier market structure that the forum is built upon.

4.1 Data Set

Data Collection

In 2010 an online underground market for credit cards and other illegal goods, **Carders.CC**, has been exposed by a hacking team named "inj3ct0r". The team has published the leaked database we base this work upon on public channels. The leaked package contains a Structured Query Language (SQL) dump of the database, a copy of the Owned and Exp0sed Issue no. 1 (documenting the leak) and an added text file containing all the private messages on the forum.

The structure of the database has a total of 68 tables. Figure 4.1 provides an overview of the tables relevant to this work categorized into several sections. An overview of all 68 tables can be found in Appendix B. All information regarding members like reputation, username, date registered, personal contact information and member roles can be found in one of the tables in the "Member Data" section. All the posts in the forum including topics and PMs can be found in one of the tables in the "Messages" section. In order for the forum to work properly some settings regarding regulation are being stored in one of the tables within the "Forum Structure Settings" section.

Data Re-construction

The data consists of forum posts and private message records spanning 12 months from 1 May, 2009 to May 1, 2010 containing a total of 215,328 records. In order to maintain the integrity of the information we created two environments: 1) A read-only replica of the forum and 2) a test environment in which writing to the database is allowed. The first environment has the sole purpose

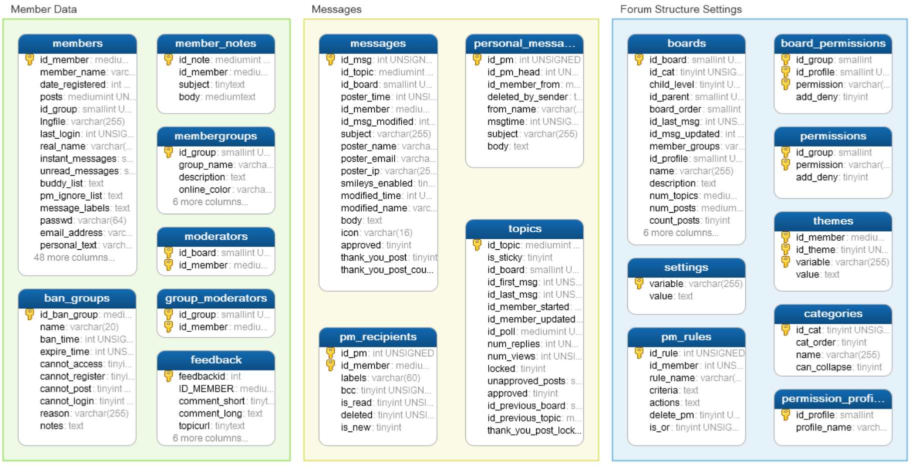
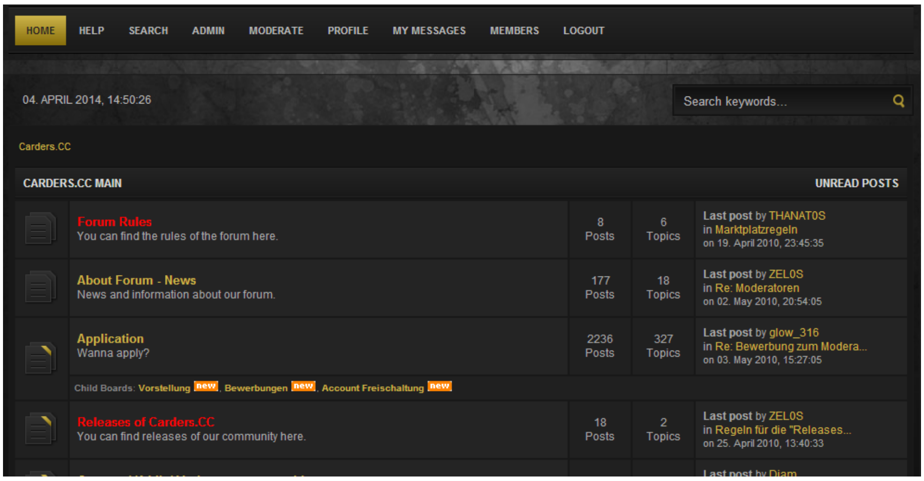


Figure 4.1: Database Diagram of Carders.CC



04. APRIL 2014, 14:50:26

Carders.CC

CARDERS.CC MAIN			UNREAD POSTS
Forum Rules You can find the rules of the forum here.	8 Posts	6 Topics	Last post by THANAT0S in Marktplatzregeln on 19. April 2010, 23:45:35
About Forum - News News and information about our forum.	177 Posts	18 Topics	Last post by ZEL0S in Re: Moderatoren on 02. May 2010, 20:54:05
Application Wanna apply?	2236 Posts	327 Topics	Last post by glow_316 in Re: Bewerbung zum Modera... on 03. May 2010, 15:27:05
Child Boards: Vorstellung new, Bewerbungen new, Account Freischaltung new			
Releases of Carders.CC You can find releases of our community here.	18 Posts	2 Topics	Last post by ZEL0S in Regeln für die "Releases... on 25. April 2010, 13:40:33

Figure 4.2: Screen capture of **Carders.CC** replica

to be used for data analysis and will therefore be left untouched and will remain exactly the same as it was when published. The second environment was mainly used for exploration of the forum by a more hands-on approach. We used XAMPP[1] as our server environment to (offline) revive the forum. This way we are able to freely explore the forum, read its content in an ordered matter, and most importantly emulate the role of an administrator that has access to all the mechanisms of interest for the analysis of the thesis. By examining the added notes Owned and Exp0sed Issue no. 1 we were able to create a what we believe can be considered a close to perfect replica of the original **Carders.CC** forum. It is important to recreate the original settings of **Carders.CC** in order to gain precise insights on the operations of the market, including the reputation mechanisms that were implemented at that time, users' posting history and dates.

[1] XAMPP, Apache Friends, https://www.apachefriends.org/download.html.

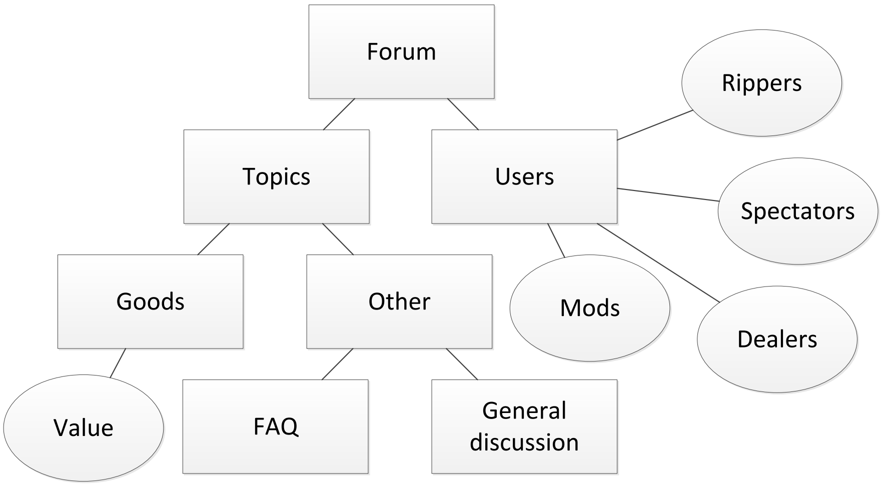


Figure 4.3: Object and attributes of **Carders.CC**

Following the notes provided by the data releasers we:

1. Used Simple Machines Forum[2] (SMF) as forum software allowing to browse the data in a structured way.

2. Implemented a MySQL back-end.

3. Imported the database twice, once for each environment, with the use of phpMyAdmin[3]– a tool to administrate the database.

A screenshot of the replica can be seen in Figure 4.2.

4.2 Market Structure

The market contains several characteristics that differ from IRC markets. The forum holds topics, in which advertisements for goods with a certain value are being made and users that operate in this market. Figure 4.3 shows all the characteristics of the forum. A more in-depth analysis gives us a grasp on how the forum looks like and operates which is needed for understanding the environment better. When further examining the forum in terms of the distribution of posts and members in the tier system we expected to notice that this distribution is gradually. In contrary, Figure 4.4 shows that over 96% of all posts in the trading market has been posted in the first tier. The majority of the users are active in Tier 1 meaning that over 85% of the members that are active in the trading market have participated in trading activities in tier 1. Tier 2 shows some activity in terms of members that posts and the total amount of posts made. More information about the Tier system will be given further on in this chapter, namely in Section 4.4.2.

[2] Simple Machines Forum, http://www.simplemachines.org/.
[3] phpMyAdmin, http://www.phpmyadmin.net/.

	Tier 1		Tier 2		Tier 3	
Posts	68,645	96%	2,565	3%	17	< 1%
Members	3,158	86%	510	13%	5	< 1%

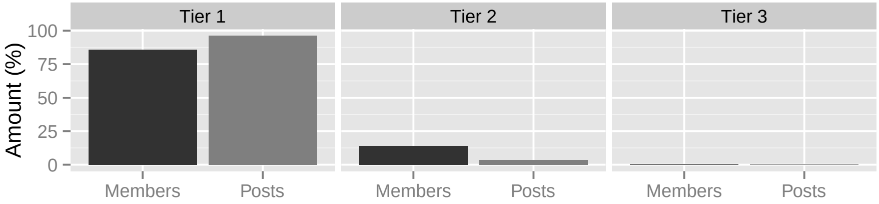


Figure 4.4: Distribution of posts and members in the market

While the figure shows that the majority of the posts are hosted by the first tier this does not exclude the possibility that tier 2 may contain a more refined and dedicated community. Tier 3 seems to have only 5 members and a total of 17 posts in which one of these members has an administrative role in the forum. We therefore will neglect these results and opt this data out from further analysis.

Upon further inspection of the market structure and the distribution of users and posts Figure 4.5 suggests that there is no to little activity in the first few months of the market. In the last 6 months of the forum there is a much higher activity in the forum.

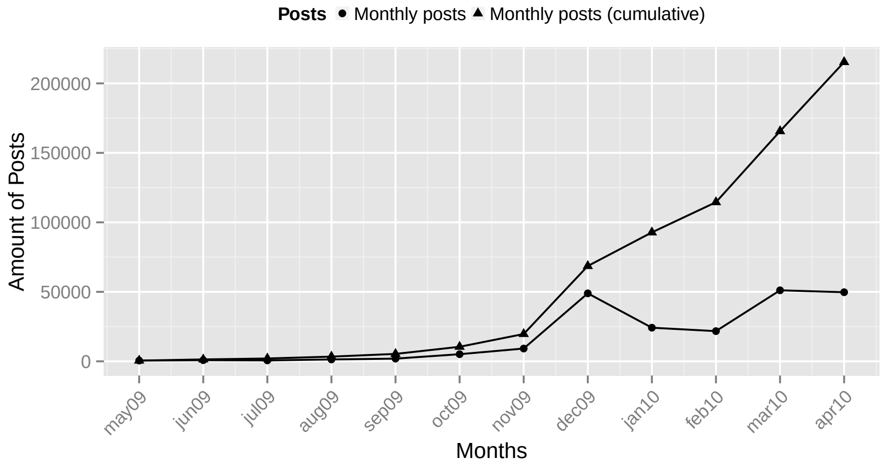
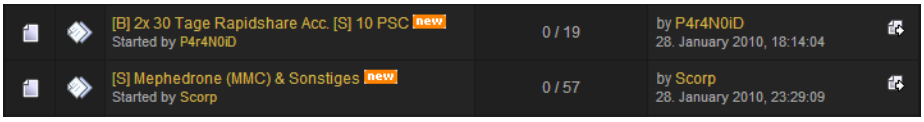


Figure 4.5: Posts on **Carders.CC** throughout the months

Regarding offered merchandise, **Carders.CC** offers several products on the market. From free tutorials, credit card samples and other free goods to more qualitative and trade related goods in the market. The market on **Carders.CC** contains several categories in which these offered illegal goods are offered. Since

Category	Tier 1	Tier 2
Hardware	×	
Services	×	
Cardable shops	×	
Tools/ Software	×	
Drops/ Packstation	×	×
Credit Cards	×	×
Accounts/ VPN/ socks	×	×
Intoxicants / drugs / medicines		×
Weapons / self-defence equipment		×

Table 4.1: Merchandise categories in the different tiers



[B] 2x 30 Tage Rapidshare Acc. [S] 10 PSC new
Started by P4r4N0iD
0 / 19
by P4r4N0iD
28. January 2010, 18:14:04

[S] Mephedrone (MMC) & Sonstiges new
Started by Scorp
0 / 57
by Scorp
28. January 2010, 23:29:09

Figure 4.6: Buying and selling advertisement

the market knows a tier system it should be a matter of course that higher tiers show transparency and adapt the same categorization as lower tiers. The contrary appears to be true since Tier 2 offers different categories that are not offered by Tier 1. These additional categories facilitate Tier 2 users to trade drugs and weapons. Some categories like hardware products, services, cardable shops and software are not being offered in Tier 2. This means that if one would only be interested in trading weapons he/ she needs to surpass Tier 1 and get access to Tier 2. This different offering of categories implicate two different markets and a tier system in which a hierarchy seems to be missing. However, regulation enforcement and policy settings should shed light on this matter in order to analyse the effects of the presence of two different markets. Table 4.1 shows the distribution of categories among the tiers. Tier 2 provides some overlapping categories where both markets offer the same illegal goods.

4.3 Advertisement

Users that join the community for selling or buying products are active in one of the market tiers within the forum. A user can advertise a product by creating a topic in the designated board in which this specific product falls. The advertisement post is labelled with a “buying” (”[S]” stands for ”Suche” in German which means buying) or a “selling” (”[B]” stands for ”Biete” which can be interpreted as selling) label. For example, Figure 4.6 shows a member that wants to sell 2 times a 30-day RapidShare account in exchange for 10 PaySafeCards (PSC) which is a popular prepaid payment option in the underground world. In these topics other users often discuss the product, ask questions and when a user shows interest as a potential buyer they contact the advertiser. According to the forum regulation, product trading has to be finalized via private messages between the two parties.

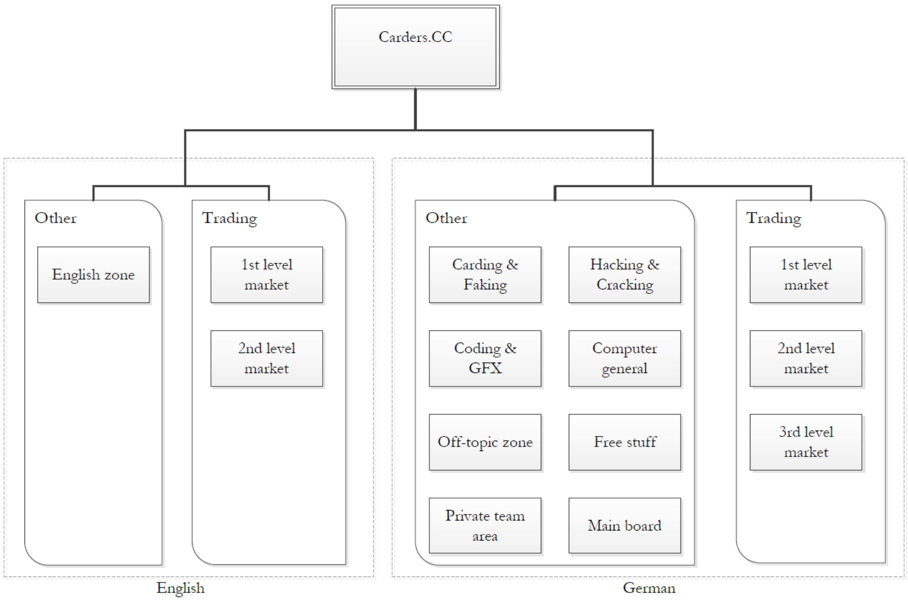


Figure 4.7: Categories of the forum

4.4 Market Regulation

Carders.CC allows both English and German speaking members on their forum. Figure 4.7 shows a schema of the two forum sections for English and German speakers. The German-speaking part of the community is clearly the most developed one: the English section has only 8% of all market posts while the remaining 92% are found in the German market. In this paper we therefore focus on the analysis of the German market. The forum knows a strict separation of trade related boards and non-trade related boards. Advertisement of (illegal) goods is permitted in the dedicated trading section. Members in this section are also allowed to request specific goods. The non-trade related boards serve the purpose of providing a discussion forum for the members where they can share thoughts, ask questions, publish tutorials and offer free goods on a specific subject. A third area of the forum, of little interest here, is dedicated to discussion of technical forum-related matters (e.g. maintenance).

The well-structured nature of online forums allow for a set of rules to be enforced. The administrators of **Carders.CC** published the guiding rules of the community in the regulation section. What follows is an overview of the regulatory structure of the community that will be central to our analysis as it identifies rules to access the trading areas of the forum and provides a clear regulatory distinction between “good” and “bad” users.

4.4.1 User Roles

Each user in **Carders.CC** can assign positive or negative *reputation points* to other forum users. Higher reputation points should correspond to a higher level of trustworthiness for the user. A user’s status in the forum is also reflected

Role	Forum	Admins	Other
Newbie	×		
Normal User	×		
2nd Tier User	×		
3rd Tier User	×		
Verified Vendor	×		
Redaktion		×	
Moderator		×	
Global Moderator		×	
Administrator		×	
Scammers and Banned Users			×

Table 4.2: User roles

by its membership in one of 12 user roles that are identified by the forum administrators. Table 4.2 shows these roles with the category to which they belong. The entry rank Newbie labels a newly registered user in the forum. After surpassing the role of a newbie, the user gets the role of normal user. Further up in the hierarchy the user becomes a 2nd and 3rd tier user and gains access to more specialized and restricted marketplaces. A verified vendor sells goods that are verified by the administrative team and therefore tend to be more trusted by market participators. In contrary to the other forum roles a verified vendor does not require to climb up the rank ladder to achieve this entitlement.

The users with an administrative role manage, maintain and administer the forum. Members of the Redaktion are the editors of the forum. They publish news, events, regulation and other administrative information. The moderators maintain the forum and enforce regulation. Final, there is the administrator role which is the highest possible rank in which this member has access to all the features in the forum. The users of the forum in the administrative group are also responsible for banning users that have been reported for "ripping" another user in a transaction, or that have violated some sort of internal rule. Ripping means that the seller fails to deliver the requested goods to the buyer after receiving payment. This is one of the examples of users that can be banned from the forum. Some other reasons are spamming, double accounts, Terms of Service violation, etc. These will be further discussed in Section 4.5/.

4.4.2 3-Tier Market System

The forum's infrastructure knows a tier-based trading market in which the forum regulation clearly distinguishes three different *trading areas* (namely *Tiers*) in the forum. In each of these tiers access is constrained by an increasingly selective set of rules.

Tier 1
The lowest accessible tier is considered the public market on **Carders.CC**. Newly registered users on the forum (see Newbies in Table 4.2) are not permitted to join the public market in Tier 1. The forum regulation statement reports that users that have obtained the role of "normal user" can access this area. In order to become a "normal user" one needs to comply to the following access rule:

1. To become a normal user a newbie has to have posted *at least 5 messages*

on the forum.

Tier 2
This market section is intended to be reserved and dedicated for the "elite" users of the forum. More restrictive rules are declared for the access to the higher tiers. Access rules for Tier 2 are stated as follows:

1. Only users with at least 150 posts are allowed in Tier 2.
2. Users must have been registered to the forum for at least 4 months.

A user does not need to apply to gain the 2nd Tier User entitlement since the system will process this automatically. The rewards the user gets by gaining this role is more PM storage space and access to the 2nd Tier market.

Tier 3
This tier is an invitation-only section of the market. It is not possible to apply for this user role, meaning that only users selected and approached by users with administrative roles will gain the proposition to join the 3rd Tier. The regulation states clearly that it is not possible to buy either the 2nd nor the 3rd Tier entitlement. For Tier 3 the following rules hold:

1. The user has been selected by a team member of the forum to be granted access to Tier 3.
2. Access to Tier 2 is required.

We exclude Tier 3 from our analysis because it features only 5 users, including one administrator, and 17 posts. We therefore consider it a negligible part of the overall market and from this point refer to the 2-Tier System.

The division and implementation of a tier based system clearly aims at dividing the market into a more "elite" community the higher the tiers get. In this working scheme one would generally assume that more refined and "elite" users are more trustworthy, meaning that trading in higher tiers result in lowering the chance of dealing with a scammer.

4.5 User Groups

As noted by Herley and Florêncio [19] one of the main threats to the workability and stability of an underground market are "rippers". The classification of these rippers (users that try to scam other users) is therefore important for our study. To achieve the classification of normal users and rippers we divide the population into several different types of users. The forum is composed of multiple areas, some of which are not strictly related to trading. Because we are interested in the market characteristics of the forum, we exclude users that have never participated in the trading sections from the analysis. Rippers are banned users, therefore a classification of the banned users needs to be made. These banned users are excluded from the market for a variety of reasons. Banned users are usually assigned an (arbitrary) string tag that describes the reason of the ban (column `reason` under the SQL table `ban_groups`, 4.1). By manual inspection we identified five categories of banned users: *Rippers, Double accounts,*

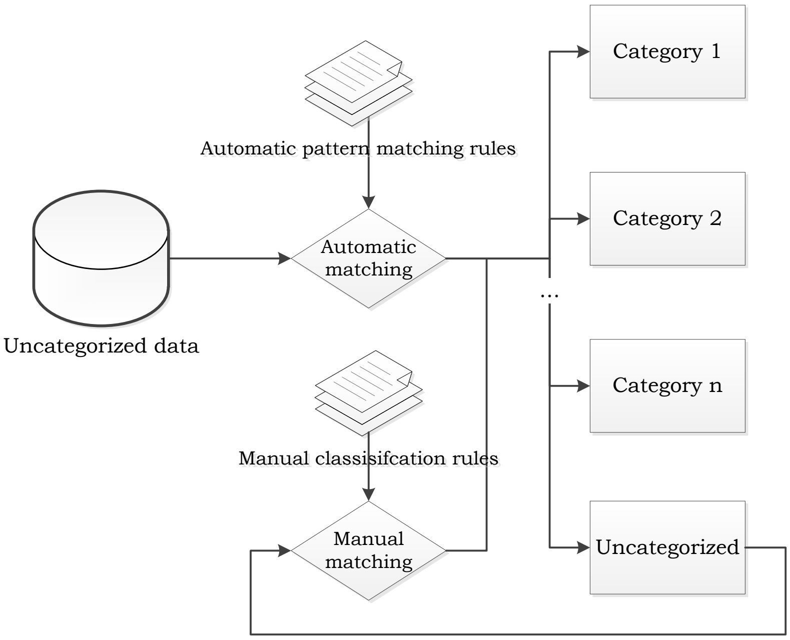


Figure 4.8: Scheme for the classification of user groups

Spammers, Terms of Service violators and an additional “Uncategorized” group for users banned without a reported reason. Unfortunately the string describing the reason of the ban is not in a standardized format, meaning that automated classification of users is not straightforward to implement. From our initial inspection of the data we created 20 regular expressions that we use to match the `reason` column. Each user is assigned to a category depending on the matching regular expression. Figure 4.8 shows the pattern matching scheme applied. While the uncategorised category leaves a set of users that cannot be identified due to incomplete arbitrary information we suggest manual classification for this data set. Another option is to use machine learning, which will be proposed for further research in Chapter **??**. After applying the automatic pattern matching scheme we distributed the population in five categories. In total around 20% of the population are banned users. Among these banned users we managed to categorize the five user groups in which rippers occupy roughly 5% of the whole population. Figure 4.9 shows the distribution of the population within the five categories.

3-Phase Filter Scheme

In order to solve several difficulties with the distribution of the population we applied a 3-phase filter scheme. This is a consecutive scheme in which each sequential filter is based on the previous filter results.

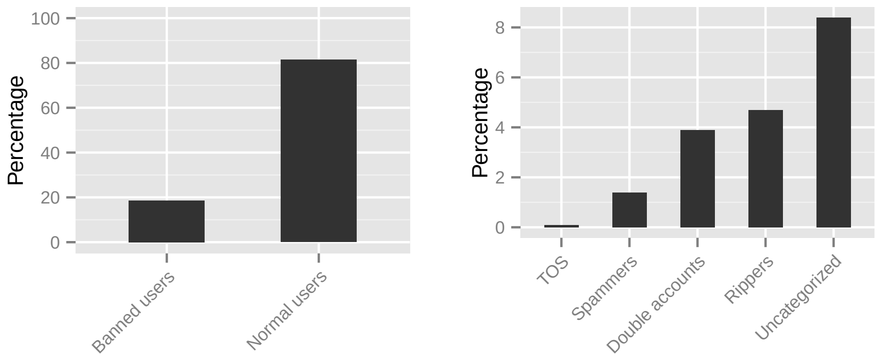


Figure 4.9: The distribution of the forum’s population

Filter 1: Regular Expression and Avoiding Duplicates
This filter is applied to the column reason in the table for the registered banned members ban_groups, which can be found in Figure 4.1. This column keeps track of the reason why members with administrative roles banned a certain member and is a free-text field meaning that the reason of the ban is described in a non-standardized way. Because user classification is critical to our work we divide the population in member groups by using regular expressions with over 20 patterns. Table 4.3 shows the patterns used for every member groups. Because of the arbitrary nature of the strings, there exists the possibility for a banned user to match multiple string patterns and therefore being assigned to multiple member groups. To avoid this problem we implemented a priority scheme mechanism in which every member group gets a pre-defined priority value assigned. By adding priority values to member groups the pattern matching scheme will choose to classify a banned member to only one member group with the highest priority value instead of multiple.

<table>
<tr><th>Category</th><th>Priority</th><th>Pattern</th></tr>
<tr><td>Rippers</td><td>5</td><td>”ripper|rippe|ripp|rip”</td></tr>
<tr><td>Double Accounts</td><td>4</td><td>”doppel|doppel-account
|double|double-account”</td></tr>
<tr><td>Empty</td><td>3</td><td>”ˆ$”</td></tr>
<tr><td>Spammers</td><td>2</td><td>”spammers|spam-
|spamming|spamm”</td></tr>
<tr><td>ToS</td><td>1</td><td>”tos|violation|rules
|swear|mouth”</td></tr>
</table>

Table 4.3: Classification of user groups by regular expression and priority

Filter 2: Minimizing Data Inconsistency
After applying the first filter we noticed an inconsistency in the data for each banned user. Some of these users showed no records of their reputation, registration date and other user related information. Since these records are needed for our research we filtered out all the banned users that did not meet this requirement. In total we classified 1263 banned users. Out of these we excluded

357 entries that were incomplete in the dataset. For example, no reason for the ban was reported or its posting history was not present in the dataset.

Filter 3: Trade Related Members

Finally, we are interested in identifying users that have been trading in the markets at some point in their posting history. This activity is measured by including only the users that have ever posted at least 1 post in the trading section of the forum. All members that have never participated in the trading section will therefore be discarded from the member group and excluded for further analysis. For this study we will explicitly be using the last highlighted Filter 3 set of members that remain after applying the 3-phase filtering. Table 4.4 shows the number of users for each group in each of the filtering process.

	Filter 1	Filter 2	Filter 3
Rippers	268	230	205
Double accounts	299	192	148
Uncategorised	78	59	40
Spammers	96	69	42
ToS	7	6	5
Normal users	3960	3960	2468
Total	5223	4866	3187

Table 4.4: Population distribution with use of the 3-phase filter scheme

Chapter 5

Research Methodology

In Chapter 4 we discussed the market properties and the main characteristics of a forum that distinguish it from a IRC market. These number of features play an important role in determining the stability of a market. From our analysis we derive a number of dimensions to measure the sustainability of the market. This market failed despite the explicit regulation of the forum aimed at distinguishing groups of "good" and "bad" users and at creating "safe trading places" where only experienced and trustworthy users participate. The systematic failure of these rules would intuitively re-create the same conditions Herley and Florêncio [19] identified for the IRC markets: information asymmetry would favour "ripping" behaviour and eventually lead the market to its failure. In this chapter we therefore systematically check the enforcement of the forum's regulation by formulating a number of hypotheses we derive from the description of the forum regulatory mechanisms reported in Chapter 4. If evidence for the validity of the hypotheses is not found in the data, we conclude that the regulation was not effectively enforced. Vice-versa, if most hypotheses are supported by the data, we would conclude that the forum administrators applied the stated rules. An overview of all hypothesis can be found in Table 5.1.

5.1 Reputation Mechanism

The forum identifies a hierarchy of user groups that each forum user can "escalate". Intuitively, in a functioning system a higher status should correspond to a higher assigned reputation in the market.

Hypothesis 1 *Banned users have on average lower reputation than normal users.*

If Hypothesis 1 is true it is evidence that the regulatory mechanism for reputation is effectively enforced, and provides the forum users an instrument to evaluate traders' historical trustworthiness. If the data does not support this, "reputation" in the forum is not a good *ex-ante* indicator of a users' trustworthiness.

Given the 2-Tier system, we would also expect the average reputation of users in Tier 2 to be higher than the average reputation in Tier 1.

Reputation Mechanism	
Hypothesis 1	Banned users have on average lower reputation than normal users.
Hypothesis 2	Users of Tier 1 have on average lower reputation than users of Tier 2.
Hypothesis 3	Banned users in the Tier 2 market have lower reputation than normal users in Tier 2.
2-Tier System Regulatory Enforcement	
Hypothesis 4	Users have at least posted 5 messages on the forum before their first post in Tier 1.
Hypothesis 5	Tier 2 users have at least 150 posts in the forum before posting their first message in Tier 2.
Hypothesis 6	Tier 2 users have been registered to the forum for at least 4 months before their first message in Tier 2.
Success of Rippers and Normal Users	
Hypothesis 7	Users finalize their contracts in the private messages market.
Hypothesis 8	Normal users receive more trade-initiation private messages than Rippers do.

Table 5.1: Overview of hypotheses

Hypothesis 2 *Users of Tier 1 have on average lower reputation than users of Tier 2.*

If Hypotheses 1 and 2 do not hold, it may as well be because moderators left Tier 1 by itself and concentrated all the regulatory effort on the higher market tiers. In this scenario, Tier 1 users may represent significant noise in the data. To verify this, we narrow our analysis to the sole Tier 2 market.

Hypothesis 3 *Banned users in the Tier 2 market have lower reputation than normal users in Tier 2.*

If Hypothesis 3 too does not hold, we conclude that the reputation mechanism provided no meaningful way for the forum users to distinguish between "bad traders" and "good traders".

5.2 2-Tier System Regulatory Enforcement

To test whether the 2-Tier system had any meaningful functionality in terms of dividing the market into more refined communities we measure this by verifying several hypotheses. These hypotheses test whether the published rules on the forum for access into the tiers are enforced, as discussed in Chapter 4.

Regulation Tier 1.

In order to participate in Tier 1 it is stated that a user has to have a user role higher than "Newbie". As from our analysis in section 4.4.1 Newbies are all non-banned users with less than 5 posts. To see whether this regulation is enforced we test the following hypothesis:

Hypothesis 4 *Users have at least posted 5 messages on the forum before their first post in Tier 1.*

If Hypothesis 4 is true this shows that anyone could enter Tier 1, including newly registered users with less than 5 posts. A possible consequence is that members participate in the market without having experience and knowledge about the community. In order to verify this hypothesis we measure for each user that has ever posted in Tier 1 the amount of posts they have posted before their first message in Tier 1.

Regulation Tier 2.

Access to Tier 2 of the Tier System is subject to several rules, as discussed in Section 4.4.2. In order to see whether the regulation of Tier 2 is enforced we verify the following hypotheses:

Hypothesis 5 *Tier 2 users have at least 150 posts in the forum before posting their first message in Tier 2.*

Hypothesis 6 *Tier 2 users have been registered to the forum for at least 4 months before their first message in Tier 2.*

If either one or both of the hypothesis fails we have verified that rules in Tier 2 are not properly enforced resulting in findings that tell more about the state in which Tier 2 is maintained. A badly maintained and administrated forum structure is in terms of hierarchical properties no different than IRC markets.

5.3 Success of Rippers and Normal Users

The role of the forum boards is to provide an asset to sellers and buyers to advertise their merchandise. The actual finalization of the trade however usually happens through the exchange of *private messages* between the trading parties [14, 19]. We inspect the existence of this “private market” by measuring if:

Hypothesis 7 *Users finalize their contracts in the private messages market.*

Given the unstructured nature of the data at hand, to test Hypothesis 7 we proceed with a manual inspection of a sample of 50 randomly picked threads in the PM market and classify them as “trade related” or “not trade related”. The goal is to understand whether the ratio of PM threads aimed at finalizing a trade supports Hypothesis 7 or not.

If Hypothesis 7 holds, than the exchange of private messages would be a good benchmark variable for us to measure the successfulness of “normal” users and “rippers” in closing trades. To check whether “normal users” are significantly more successful than “rippers” we test the following hypothesis:

Hypothesis 8 *Normal users receive more trade-initiated private messages than Rippers do.*

We measure this by counting the number of *unsolicited incoming private messages* a user receives i.e. the number of times a forum user initiates a trade with another forum user.

We would expect the results for Hypothesis 8 to be coherent with the results obtained so far. In other words, if the reputation mechanism works, the Tier System is properly enforced, and the exchange of private messages is used to conclude the trading process. Subsequently we would expect normal users to conclude more trades than rippers do. This is because the consistent enforcement of the forum rules would give the users an instrument to discern rippers from normal users. Otherwise, if the evidence gathered so far suggests a systematic failure in the market regulation, then we would expect rippers to be indistinguishable from normal users (because the user could do no better than randomly picking a seller from the whole population).

Chapter 6

Data and Analysis

In Chapter 5 we described how to measure whether **Carders.CC** differs from IRC markets by verifying numerous hypotheses that determine market stability. In this chapter we measure and verify each of these hypotheses by analysing the data in a structured way. First, we examine **Carders.CC** and verify their corresponding hypotheses. This chapter contains box plots, histograms and density plots to present data and verify hypotheses in which each of them will be discussed below.

Box plots

We use box plot presentations in several cases due its simple and clear visuals of distributions. In a box plot the distribution that we are examining is being divided into four quartiles in which the first 25% contain the lower part, the next 25% the lower part of the box, followed by the mean (the mid value of the whole distribution), the upper part of the box 25% and finally the last 25% of the distribution in the upper part. It is important to notice the difference between vertical lines and dots in the box plots. The outliers in the box plots are indicated as dots and are presented as such because this observation point differs in terms of distance from other values. A line indicates that the values on this line do not differ much in terms of distance and are therefore close to each other.

Mann-Whitney Unpaired Test

For our analysis we use the Mann-Whitney unpaired test several times for verifying our findings in the box plots, histograms and other graphical representations. Due to the nature of our data at hand we chose for this specific test. Many of our data is non-parametric meaning that the data does not relate to any other distribution in any way. In verifying our hypotheses we mainly use numeric values in distributions where the data can be ranked by order. The numeric values do not relate to other values and therefore there is no other value that influences the numeric distribution, hence the non-parametric property. Furthermore our data at hand is unpaired, meaning that the amount of values in one distribution is not equal to the amount of values in the other distribution. Because of these specific properties of the data we chose the Mann-Whitney unpaired test.

For performing this test two distributions are being compared in order to determine whether a distribution is different from another. To achieve this we can use the Mann-Whitney unpaired test to either use one sided hypothesis tests or two sided. One sided tests are used to determine whether one distribution group tends to have larger response values than the other. We use one sided tests on the results presented by the box plots in Hypothesis 1, 2 and 3. The box plots give a visual image of the distribution, however to be more certain before drawing any conclusion the Mann-Whitney unpaired test gives us a more statistical approach towards the data of these findings. In order to test a distribution for equality we set a null hypothesis stating that both distribution are to be equal. Optionally, we state an alternative hypothesis (one-sided) to see if the null hypothesis fails. As a result we get a p-value which we use as a probability value to accept or reject our null hypothesis. Being able to do so, we set a threshold value of $\alpha = 0.05$. When the p-value is less than or equal to α we reject the null hypothesis and conclude that one groups tends to have a larger response value than the other. In a two-sided test this would mean that we conclude that there is a difference between the distribution, either larger or smaller.

6.1 Reputation Mechanism

To test our hypotheses for the reputation mechanism on **Carders.CC** we mainly analyse reputation values of users in the market in different sets for each hypothesis.

Hypothesis 1. ***Banned users have on average lower reputation than normal users.***

For the first hypothesis we test if banned users have a lower reputation than normal users in the market. Figure 6.1 is a box plot of the reputation levels for "banned" and "normal" users. Due to the nature of our data (many low reputation values and some high values) we emphasize on this strange distribution by using a box plot. As can been seen in Figure 6.1 the data is on a logarithmic scale. The distribution of outliers suggests that reputation points make little sense with respect to user categories. Reputation levels differ from 10 to 100 but there are many outliers between 100 and 65.000. With such a huge difference of reputation levels within the distribution we find it high likely that these numbers are accurate. For our hypothesis to be verified we would expect to see that both distributions differ and the reputation level for normal users are higher than banned users.

To see whether both distributions differ from each other we apply a Mann-Whitney unpaired test. Since the data at hand is non-parametric by nature due to the ranked order of reputation we made a choice to apply the Mann-Whitney unpaired test to see whether the visual representation of 6.1 is correct. We tested the reputation distribution of banned users and normal users to be the same by stating our null hypothesis: "*The difference in reputation between banned and normal users is zero*". Alternatively we set the hypothesis "*banned users have higher reputation than normal users*". Our alternative hypothesis should confirm the visual representation of Figure 6.1. Running the Mann Whitney

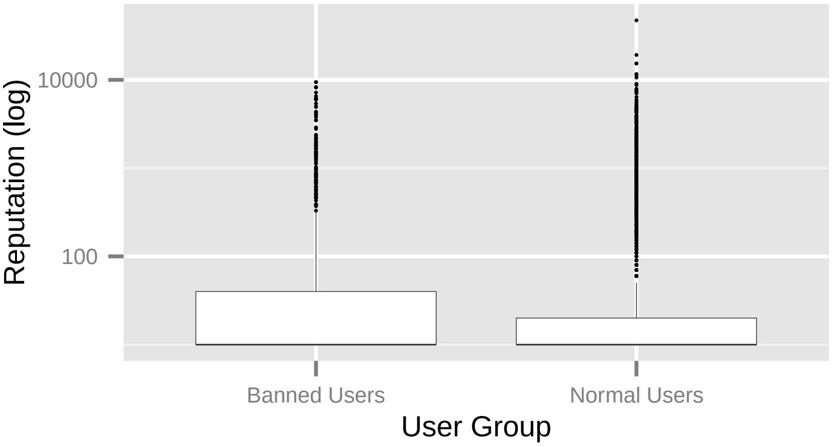


Figure 6.1: Box plot of reputation levels for normal users and banned users in the whole market (dots are outliers)

unpaired test results in a p-value of $p = 5.2e-15$. Since $p < \alpha$ we reject the null hypothesis and verify that banned users have a higher reputation than normal users. This results into rejecting Hypothesis 1.

Hypothesis 2. ***Users of Tier 1 have on average lower reputation than users of Tier 2.***

For the reputation of users in both tiers we verify whether users have a lower reputation in Tier 1 than users in Tier 2. For this hypothesis we do not distinguish between user groups since we would like to test the reputation values bound to a certain Tier community. Figure 6.2 shows the box plot distribution of reputation for users in Tier 1 and Tier 2. Again, we see that the distribution of reputation levels for users in Tier 1 have very high values, just as the outliers of users in Tier 2. However, more than 75% of the distribution has a reputation value below 100. This box plot indicates and confirms the findings of the box plot for Hypothesis 1, namely that the distribution of reputation levels make little sense. However, it seems that Tier 1 users have higher reputation levels than users in Tier 2.

To verify our findings in the box plot we once again apply the Mann-Whitney unpaired test. We interpret statistics of this data in the same way as for Hypothesis 1 with the use of a one-sided test. The null hypothesis we state reads as follows: "*Tier 1 and Tier 2 users have the same reputation distribution*". Our alternate hypothesis is stated as: "*Tier 1 users have a higher reputation than Tier 2 users*". Running the test yields us with a p-value of $p = 4.8e - 06$. In this case the p-value is smaller than our α-value, therefore we reject the null hypothesis and conclude that users in Tier 1 have a higher reputation than users in Tier 2. This results in rejecting Hypothesis 2 as well: reputation levels do not reflect membership in a "higher market level" and are effectively misleading.

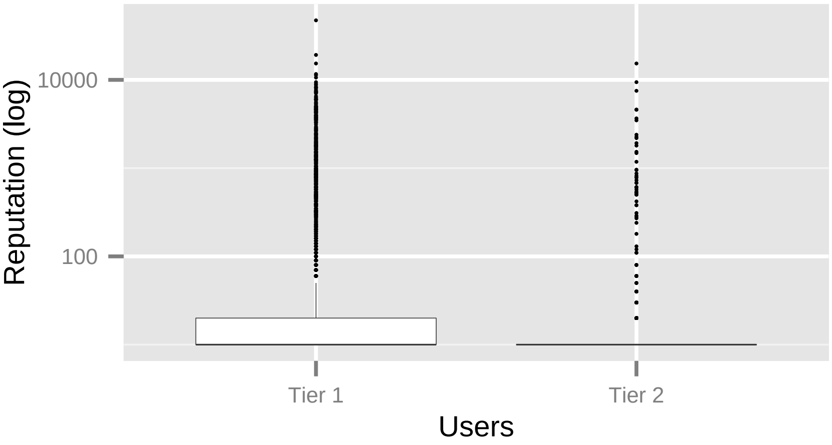


Figure 6.2: Boxplot of reputation levels for Tier 2 users vs Tier 1 users (dots are outliers)

Hypothesis 3: ***Banned users in the Tier 2 market have lower reputation than normal users in Tier 2.***

Finally, for testing the last hypothesis for the reputation mechanism we check whether reputation is at least a satisfactory indicator of user trustworthiness in Tier 2. Figure 6.3 reports the box plot distributions of reputation levels for banned and normal users in Tier 2. After having examined the previous two box plots here the result is readily obvious and straight-forward: normal users have on average a lower reputation than banned users. We clearly see that banned users have a higher mean and the majority of the normal users (over 75%) have a reputation level way below the average banned user in Tier 2.
Confirming these findings with the Mann-Whitney unpaired test yields a p-value of $p = 4.9e - 16$. For our null hypothesis we state "*Normal users and banned users in Tier 2 have the same reputation distribution*" and alternatively one-sided we state "*Banned users have a higher reputation distribution than normal users in Tier 2*". The p-value is smaller than α, therefore we reject the null hypothesis and confirm that banned users have a higher reputation than normal users in Tier 2. Consequently we also reject Hypothesis 3.

All evidence suggests that the reputation mechanism in the forum did not work. We therefore exclude that reputation could have been a significant and useful instrument in the hands of the user to identify trustworthy trading partners. This also means that cheaters, or rippers, had no "fear" of having their reputation level decreased by a disgruntled costumer, as reputation itself had no meaning whatsoever in the market.

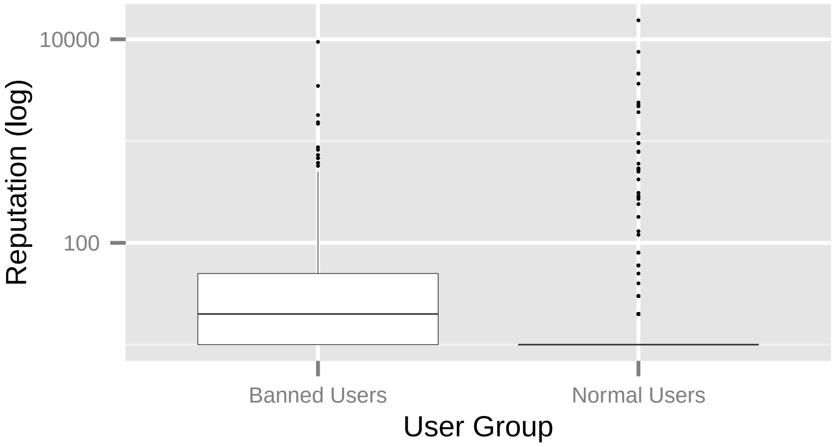


Figure 6.3: Box plot of reputation of tier 2 normal users vs tier 2 banned users (dots are outliers)

6.2 2-Tier System

To see whether the regulation in the 2-Tier System is enforced we verify Hypothesis 4, 5 and 6. These hypotheses determine whether regulation was enforced on **Carders.CC** and indicate whether the 2-Tier system was implemented and administered as stated on the published regulation announcement on the forum.

Regulation Tier 1

Access to Tier 1 has only one requirement, namely having reached a total amount of 5 posts within the forum. This access restriction will be assessed by the following hypothesis:

Hypothesis 4: ***Users have at least posted 5 messages on the forum before their first post in Tier 1.***

When rules are properly enforced in the first tier this would mean that no user with less than 5 posts is able to participate in Tier 1. Figure 6.4 shows a histogram. A histogram at this point is useful for showing a frequency distribution. In our case, we want to present the frequency of posts of all users that have ever posted in Tier 1 and categorize them into users that posted more and less than 5 posts. Doing so results in the presentation of the histogram that clearly shows that more than 50% of the users in Tier 1 gained access before their fifth post in the community. According to the rule enforcement we would have expected that no one of the users had posted less than 5 posts before they ever posted in Tier 1. Despite this being a very simple and straightforward rule to automate, there is no evidence of its implementation in the forum. By showing that users were able to post in Tier 1 regardless of their amount of posts we reject Hypothesis 4.

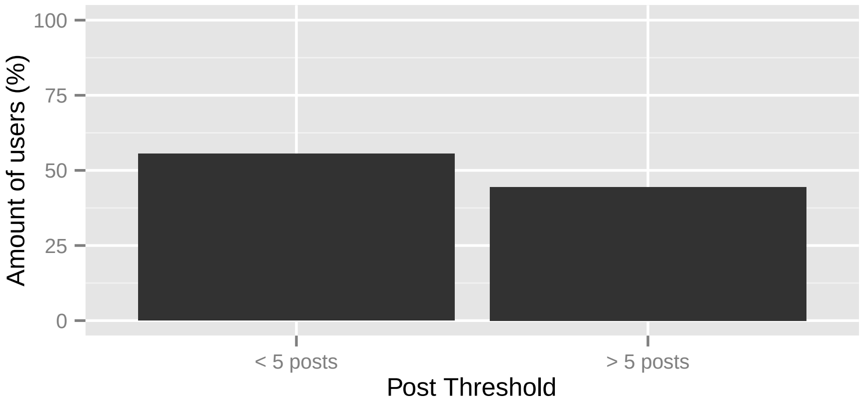


Figure 6.4: Amount of users in Tier 1 that have posted more and less than 5 posts

Regulation Tier 2

Tier 2 is subject to a set of rules that seem to be more difficult to achieve. To see whether these rules are enforced we test these rules by stating Hypothesis 5 and 6. In these hypotheses we verify whether regulation for access in Tier 2 is enforced.

Hypothesis 5: ***Tier 2 users have at least 150 posts in the forum before posting their first message in Tier 2.***

The first rule for access to Tier 2 states that users should have at least 150 posts before posting their first message in Tier 2. Figure 6.5 presents a stacked bar graph in which we show the posting frequency among each user group. For each of these distributions the stacked bar graph shows the amount of users of that user group that have posted more and less than 150 posts. In other words, this stacked bar graph shows a breakdown of the posting history for each user category. This way we gain a clear visual if there are any users in Tier 2 that posted less than 150 posts. In Figure 6.5 the majority of the users that have posted in Tier 2 for their first time clearly posted less than 150 posts before entering Tier 2. In particular *Double Accounts*, where nobody was restrained by the 150 post access rule for Tier 2. This may suggest that users already familiar with the forum (e.g. previously banned users) were accessing Tier 2 quicker than others, possibly purposely exploiting the lack of controls. In general, the great majority of users in Tier 2 gained access before the set limit of 150 posts. Therefore we also reject Hypothesis 5.

Hypothesis 6: ***Tier 2 users have been registered to the forum for at least 4 months before their first message in Tier 2.***

For Hypothesis 6 we verify the subscription rule for Tier 2. We apply a post density approach in order to show whether posts were made before the stated

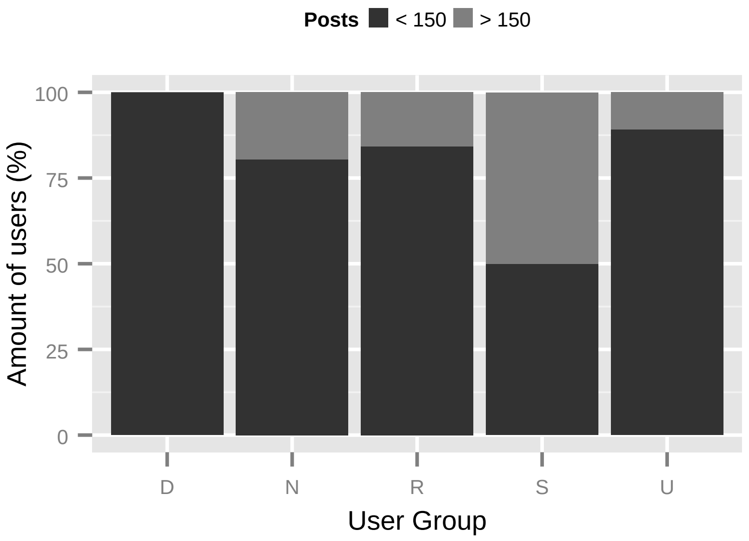


Figure 6.5: Users in Tier 2 with more and less than 150 posts. D=Double accounts; N=Normal Users; R=Rippers; S=Spammers; U=Unidentified banned users

4 months of registration in combination with a histogram on the background. The density plot applies a probability density function in order to determine the probability that a random value occurs at a given value. Translated to our needs we want to determine the probability that a post occurred at a specific given month. Figure 6.6 shows the density plot of posts in Tier 2 over the months in which a user is registered to the forum. Most of the posting activity of all users in Tier 2 is well before the 4 months threshold from registration. For banned users we see that in the first month after registration they peak the most in terms of the probability of posting. In the first three months the posting probability is high and gradually decreases. For the normal users the probability distribution does not differ much from banned users. Meaning that normal users also have the highest probability of posting in the first month after registration and gradually decreases during the upcoming months. However, if the registration threshold of 4 months would have been enforced we would expect to see no to little posting probability of a user in the first 4 months. Additionally this also supports the previous conclusion that users had immediate access to Tier 2 when registered. These findings result in the rejection of Hypothesis 6.

6.3 Unregulated Trading

In the previous sections the data showed that mechanisms are possibly faulty implemented and not working properly. As a result, we now measure the effects of these regulatory inefficiencies within the market to see the impact on the

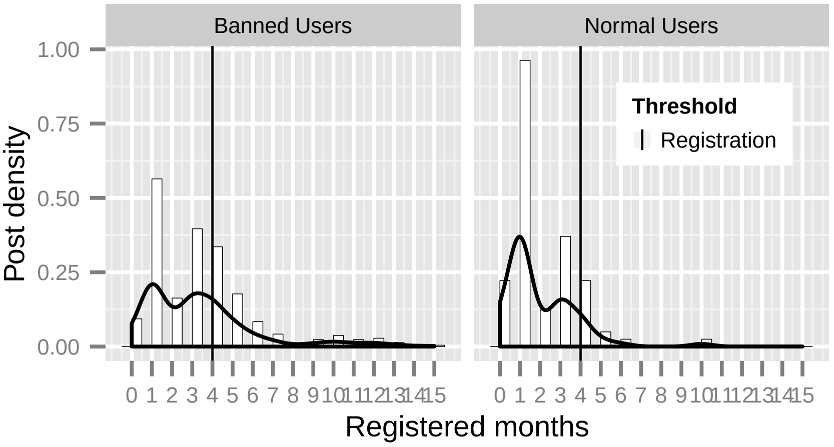


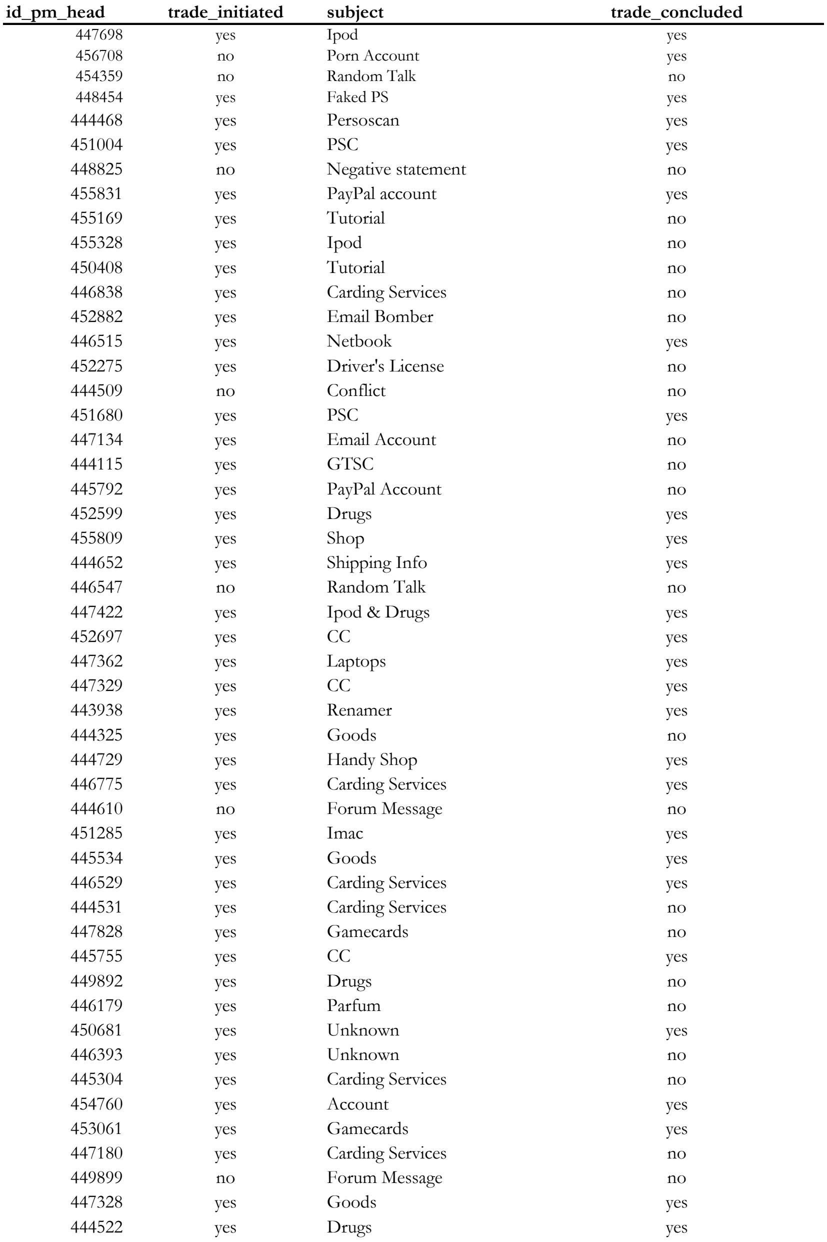
Figure 6.6: Post density of posts for Banned and Normal users in Tier 2

trades of the users.

Hypothesis 7: ***Users finalize their contracts in the private messages market.***

In Hypothesis 7 we verify whether users finalize their contracts in the private market. In order to do so we manually analyse a randomly picked sample of 50 threads sent as private messages among users to see whether we can safely assume that private messages were effectively used to finalize trades. These threads can be interpreted as conversations between two parties in a private channel. To receive randomness in our sample we have used a Random Number Generator (RNG) service from *random.org*[1] which is based on atmospheric noise. Figure 6.7 shows a list of 50 randomly selected PM samples in the unregulated market. Each PM is a message between two or more users. This means that the *id_pm_head* can be seen as an unique identifier for this conversation. Therefore each record of these 50 samples can contain 1 or more PMs. By diving into these conversations we manually decided whether they were trade related, what the subject was and if contact information was exchanged for finalizing the contract. Figure 6.8 shows that 86% of the manually examined sample conversations are trade initiated. This means that the majority of the PMs we manually examined contained content that indicated they are interested in trading illegal goods. In some cases, subjects like acquisitions, random talk and alerts from the system (Forum Messages) were discussed and therefore were not trade initiated. Over 50% of all examined PMs contained exchanged contact information between two parties. This occurs when two parties agree on the terms which results into exchanging contact information in order to pay and deliver the product. Exchanged contact information in our manually selected sample involves e.g. ICQ, Post Address and PayPal information. In some of

[1] True Random Number Generator, Randomness and Integrity Services Ltd., http://www.random.org



id_pm_head	trade_initiated	subject	trade_concluded
447698	yes	Ipod	yes
456708	no	Porn Account	yes
454359	no	Random Talk	no
448454	yes	Faked PS	yes
444468	yes	Persoscan	yes
451004	yes	PSC	yes
448825	no	Negative statement	no
455831	yes	PayPal account	yes
455169	yes	Tutorial	no
455328	yes	Ipod	no
450408	yes	Tutorial	no
446838	yes	Carding Services	no
452882	yes	Email Bomber	no
446515	yes	Netbook	yes
452275	yes	Driver's License	no
444509	no	Conflict	no
451680	yes	PSC	yes
447134	yes	Email Account	no
444115	yes	GTSC	no
445792	yes	PayPal Account	no
452599	yes	Drugs	yes
455809	yes	Shop	yes
444652	yes	Shipping Info	yes
446547	no	Random Talk	no
447422	yes	Ipod & Drugs	yes
452697	yes	CC	yes
447362	yes	Laptops	yes
447329	yes	CC	yes
443938	yes	Renamer	yes
444325	yes	Goods	no
444729	yes	Handy Shop	yes
446775	yes	Carding Services	yes
444610	no	Forum Message	no
451285	yes	Imac	yes
445534	yes	Goods	yes
446529	yes	Carding Services	yes
444531	yes	Carding Services	no
447828	yes	Gamecards	no
445755	yes	CC	yes
449892	yes	Drugs	no
446179	yes	Parfum	no
450681	yes	Unknown	yes
446393	yes	Unknown	no
445304	yes	Carding Services	no
454760	yes	Account	yes
453061	yes	Gamecards	yes
447180	yes	Carding Services	no
449899	no	Forum Message	no
447328	yes	Goods	yes
444522	yes	Drugs	yes

Figure 6.7: Manual classification of PMs in the unregulated market

	True		False	
Trade Initiated	43	86%	7	14%
Trade Concluded	27	54%	23	46%

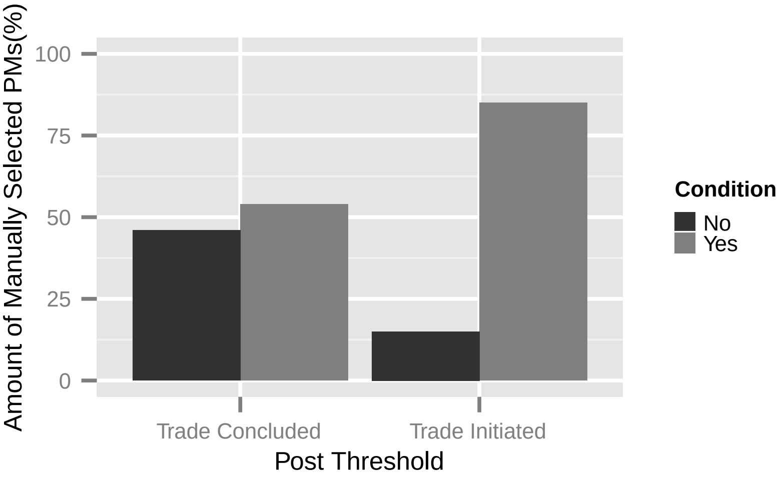


Figure 6.8: Amount of trade related PMs

these cases both parties did not finalize the contract due to the fact that there was no response given or some disagreement between the two users occurred. Furthermore it is not evident that a trade concluded PMs are also classified as trade initiated. In some cases terms and other trade initiated matters were not discussed in the PMs but rather in the forum. Since they probably do not want to share contact information they only sent this information as a PM. Therefore it is possible for PM conversations to not contain any information that suggests that the PM is trade related but is only being used to conclude a trade. It is arguable whether there is a highly active unregulated market but in terms of our hypothesis the data shows that it is undeniable that the unregulated market contains trade initiated PMs and contracts are finalized in this private channel. Therefore our findings of the manual classification of PMs supports Hypothesis 7.

Hypothesis 8: ***Normal users receive more trade-initiation private messages than Rippers do.***

We are now interested in seeing whether users that have been banned for explicitly *ripping* other users are more or less successful than normal users. Given the results we obtained so far, we expect the two to be indistinguishable: if there is no available mechanism to distinguish between “good” and “bad” users (as the findings indicates up to here), then choosing with whom to trade can be no better than randomly picking from the population of traders. Figure 6.9 is a box plot containing the representation of received PMs for Rippers and Normal users in the forum. By inspecting the distributions it is clear that the two distributions overlap significantly. The box plot shows that the means are almost

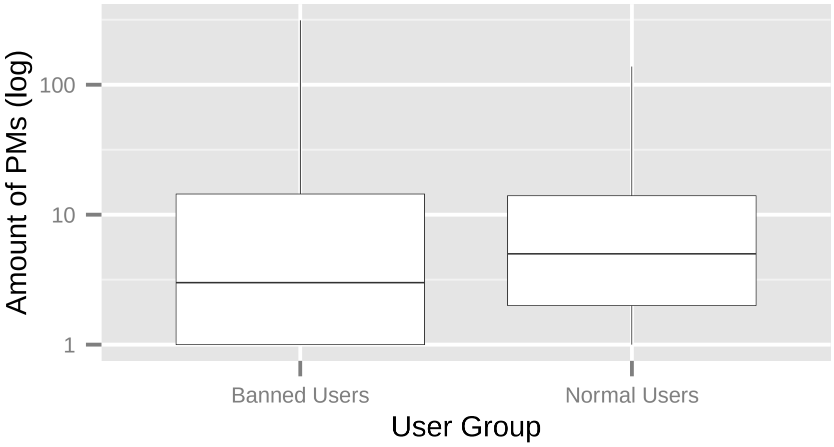


Figure 6.9: Initiated trades for Ripper users and Normal users

identical, meaning that Normal Users and Rippers receive on average the same amount of PMs.

To verify the representation of our box plot we apply the Mann-Whitney unpaired test, again we chose this test due to the nature of our non-parametric and rank ordered numeric amount of received PMs for each user. At this point we are interested in knowing whether the populations are identical, as Figure 6.9 indicates. In order to see if the PM distribution of Normal users and Rippers are equal we apply a two-sided approach since we do not know which distribution would be greater or smaller. We therefore state our null hypothesis as follows: *"There is no difference in the averages of received private messages for Rippers and Normal users"*. Testing the null hypothesis two-sided yields a p-value of $p = 0.98$. Since the p-value is larger than our α-value we accept the null hypothesis and conclude that Normal users and Rippers have identical distributions. As expected in light of the evidence so far, the systematic failure of the forum mechanisms made rippers and normal users effectively indistinguishable to the trade initiator.

After verifying the hypotheses we see that the majority is rejected except for Hypothesis 8. Hypothesis 8 suggests that there is a private market in which contracts are being finalized. Even though this may seem a trivial understanding there are a couple of reasons that contradicts the fact that there is an unregulated market present. First, the forum regulation state that finalizing contracts should be done in private channels, however showing that all regulation (rejection of Hypothesis 4, 5 and 6) is not enforced this makes it unlikely that the main reason for an unregulated market is due to the enforcement of this rule. For our research we find it interesting whether there is an unregulated market present and if rippers are indistinguishable from normal users. Rejecting almost all of our hypotheses suggests that there is no regulation enforced and that the reputation mechanism is not working properly. Making these key elements fail in a forum makes it no different than an IRC market and therefore a clear

		True	False
Reputation			
Hypothesis 1	Banned users have on average lower reputation than normal users.		×
Hypothesis 2	Users of Tier 1 have on average lower reputation than users of Tier 2.		×
Hypothesis 3	Banned users in the Tier 2 market have lower reputation than normal users in Tier 2.		×
Regulation			
Hypothesis 4	Users have at least posted 5 messages on the forum before their first post in Tier 1.		×
Hypothesis 5	Tier 2 users have at least 150 posts in the forum before posting their first message in Tier 2.		×
Hypothesis 6	Tier 2 users have been registered to the forum for at least 4 months before their first message in Tier 2.		×
Unregulated Market			
Hypothesis 7	Users finalize their contracts in the private messages market.	×	
Hypothesis 8	Normal users receive more trade-initiation private messages than Rippers do.		×

Table 6.1: Overview of verified hypotheses

distinguish between rippers and normal users misses. An overview of all the hypotheses verified can be found in Table 6.1.

Chapter 7

Discussion & Conclusions

“Regulation” is the main advantage that a forum-based community has over an IRC-based community: it provides the forum users with a set of rules and mechanisms to assess the information they can collect on a particular trade. The analysed market attempted to enforce this by providing a regulatory mechanism for a) user reputation; and b) access to “elite” market tiers. This may be not sufficient as an instrument for the user to have complete information on the transaction. However, it could provide the user with some baseline information of with whom the user is initiating the trading with, ruling out part of the *information asymmetry* problem identified for other markets [19], and precisely by mitigating the *adverse selection* problem [10].

Our analysis showed that each and everyone of these mechanisms has been faultily implemented, with the result that the only potential means for a user to assess ex-ante a trade are pointless or even misleading. The systematic failure of the regulatory mechanisms clearly led to a market were users had no disincentives in ripping others, and where users had no means to distinguish “good traders” from “bad traders”. As a result, we showed that there is in fact no difference in the number of trades initiated with a ripper and trades initiated with a normal user. This could not lead nowhere but to the failure of the market, which we show being effectively of the same nature of Herley and Florêncio’s.

During the research several limitations and restrictions arose from the data. For one, it was not possible to analyse the difference of the two markets in the Tier system due to the structure of the forum. Since the categories were not redundant this resulted into two different environments in which Tier 2 offered goods that were not offered in Tier 1. A comparison between market properties of Tier 2 and Tier 1 was hence not possible.

Our replica of the forum is limited in its resemblance with the original version. We were able to obtain information about most of the settings, software packages and versions from the data. However, we were not able to detect possible plug-ins, let alone custom written code and alteration to the software.

The regular expression method we used for classifying member groups in the population was not able to classify all banned members. During the process of pattern matching the data showed to contain a lot of records that made it impossible to be used for our approach. In our pattern matching scheme we saw that after applying our manual classification we were able to distribute the population into five categories. However, there still remains an uncategorised set of banned users that cannot be categorized with regular expression techniques. We therefore propose to fully distribute the population with use of other techniques. We believe that machine learning is a possibility that delivers a result that leaves a lower amount of uncategorised users. We therefore leave the improvement of our distribution method for future research.

Furthermore we have showed that there is an unregulated market in terms of finalized contracts. We are able to monitor and classify finalization of contracts until the point of the exchange of contact information but not beyond this point. This leaves out an essential part of the contract, namely the transaction. The forum does not provide a payment option and only serves as a medium to advertise illegal goods. Being able to monitor the payment would have provide us with a more solid analysis on finalizing contracts. In current studies it has yet to be achieved to analyse transactions in an accurate way. Since IRC and forum markets are being used as an advertisement board potential buyers leave the environment upon agreement between two parties. Therefore transactions being made, quantities and contact information is only measurable when discussed on the advertisement board. It would be more accurate to be able to measure directly these transactions since all current economical studies use estimations of underground market revenues.

Concerning threats to validity we noted several cases in the data that posed a threat. All of the private messages contained a timestamp with the date set in the last month which seemed unlikely. All of the 12.000 PMs had a posting interval of a couple of seconds. By checking PMs of members that were not active at the moment the PM was sent suggested that the timestamp were incorrect. We therefore avoided using the timestamp of the PMs in our analysis.

Furthermore we have showed that rippers are indeed an undeniable presence in our analysed underground market **Carders.CC**. Due to the fact that we observed that rippers are indistinguishable from normal users and the tier system does not work properly it would be interesting to research a comparison between rippers in failed markets and rippers in successful markets. In collaboration with the University of Trento we have roughly touched this surface by comparing results of a successful underground market and our research of a failed underground market. More information can be found in Appendix C. The contribution of this Appendix is to provide an example of regulation in a successful underground community in which (indirect) effects are daily reported in security news and industry reports. We leave the continuation of this research for future research.

By replicating and expanding the findings of Herley and Florêncio [19] we were able to systematically verify whether our chosen failed forum **Carders.CC** meets the hierarchical properties set by Herley and Florêncio in which they discuss the difference between forums and IRC channels. We found that the forum is poorly administrated and maintained. The reputation mechanism is not properly implemented and findings show that reputation of users make no to little sense. When one cannot distinguish rippers from normal users in terms of trustworthiness it is evident that this contributes in failing the market. According to the lemon market principle when the market is flooded by scammers the "good" users will eventually leave the market, leaving only scammers in the market. Since it is hard to assess whether an offered product has value it is key to trust the selling party. If this is not properly implemented and working, like in our findings, eventually the market will fail as happened with **Carders.CC**.

Reputation is not the only mechanism in which **Carders.CC** fails. By failing all our hypotheses that test whether the 2-Tier System is enforced we conclude that regulation concerning access to higher tiers is not enforced. The Tier System, consisting of two significant tiers, should facilitate a more “elite” community in Tier 2. A result of a bad regulated Tier System is that users can move freely within the community. This means that rippers, spammers and other unwanted user groups can participate in higher tiers. In other words, users that can access Tier 2 do not have gained any added benefits like one would for example expect to deal with more “elite” and trustworthy users.

To conclude our findings we see that rippers and normal users have no clear distribution in the higher tiers. This means that one is not able to distinguish “good” users from “bad” users. In the unregulated market (private market) we saw that rippers and normal users are finalizing almost an equal amount of contracts. This high possibility of being ripped, eventually leading to flooding the market with scammers, is one of the results Herley and Florêncio state when dealing with a failing market.

We are the first to apply Herley and Florêncio’s findings on IRC markets to a different embodiment, namely a forum environment. From all the evidence from our findings we come to the conclusion that we confirm the findings of Herley and Florêncio by showing that a badly regulated cybercrime *forum* community is virtually no different from an unregulated *IRC* community. As a result, users participating in those markets have no means to safely assess the characteristics of the user they are trading with. As predicted by Herley and Florêncio, this leads to a chaotic market where rippers and legitimate sellers and buyers are indistinguishable, and therefore there is no disincentive for the rippers in scamming other users.

Bibliography

Peer reviewed

[1] G. A. Akerlof. “The Market for ”Lemons”: Quality Uncertainty and the Market Mechanism”. In: *The Quarterly Journal of Economics* 84.3 (1970), pp. 488–500.

[2] L. Allodi, V. Kotov, and F. Massacci. “MalwareLab: Experimentation with Cybercrime Attack Tools”. In: *Presented as part of the 6th Workshop on Cyber Security Experimentation and Test*. Washington, D.C.: USENIX, 2013. URL: https://www.usenix.org/conference/cset13/workshop-program/presentation/Allodi.

[3] R. Anderson and T. Moore. “Information security economics – and beyond”. In: *ADVANCES IN CRYPTOLOGY - CRYPTO 2007, LNCS*. Springer Verlag, 2008.

[4] R. Anderson, C. Barton, R. Böhme, R. Clayton, M. van Eeten, M. Levi, T. Moore, and S. Savage. “Measuring the Cost of Cybercrime.” In: *WEIS*. 2012.

[5] V. Benjamin and H. Chen. “Securing cyberspace: Identifying key actors in hacker communities”. In: *Intelligence and Security Informatics (ISI), 2012 IEEE International Conference on*. IEEE. 2012, pp. 24–29.

[6] H. Binsalleeh, T. Ormerod, A. Boukhtouta, P. Sinha, A. Youssef, M. Debbabi, and L. Wang. “On the analysis of the zeus botnet crimeware toolkit”. In: *Privacy Security and Trust (PST), 2010 Eighth Annual International Conference on*. IEEE. 2010, pp. 31–38.

[7] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu. “Design and analysis of a social botnet”. In: *Computer Networks* (2012).

[8] A. Cárdenas, S. Radosavac, J. Grossklags, J. Chuang, and C. Hoofnagle. “An economic map of cybercrime”. In: TPRC. 2009.

[9] D. Dagon, C. Zou, and W. Lee. “Modeling Botnet Propagation Using Time Zones”. In: *In Proceedings of the 13 th Network and Distributed System Security Symposium NDSS*. 2006.

[10] K. M. Eisenhardt. “Agency Theory: An Assessment and Review”. English. In: *The Academy of Management Review* 14.1 (1989), pp. 57–74. ISSN: 03637425. URL: http://www.jstor.org/stable/258191.

[11] H. Fallmann, G. Wondracek, and C. Platzer. "Covertly Probing Underground Economy Marketplaces". In: *Proceedings of the 7th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*. DIMVA'10. Bonn, Germany: Springer-Verlag, 2010, pp. 101–110. ISBN: 3-642-14214-1, 978-3-642-14214-7. URL: `http://dl.acm.org/citation.cfm?id=1884848.1884856`.

[12] E. L. Feige. "Defining and estimating underground and informal economies: The new institutional economics approach". In: *World development* 18.7 (1990), pp. 989–1002.

[14] J. Franklin, A. Perrig, V. Paxson, and S. Savage. "An inquiry into the nature and causes of the wealth of internet miscreants". In: *ACM conference on Computer and communications security*. 2007, pp. 375–388.

[15] C. Grier, L. Ballard, J. Caballero, N. Chachra, C. J. Dietrich, K. Levchenko, P. Mavrommatis, D. McCoy, A. Nappa, A. Pitsillidis, N. Provos, M. Z. Rafique, M. A. Rajab, C. Rossow, K. Thomas, V. Paxson, S. Savage, and G. M. Voelker. "Manufacturing Compromise: The Emergence of Exploit-as-a-service". In: *Proceedings of the 2012 ACM Conference on Computer and Communications Security*. CCS '12. Raleigh, North Carolina, USA: ACM, 2012, pp. 821–832. ISBN: 978-1-4503-1651-4. DOI: `10.1145/2382196.2382283`. URL: `http://doi.acm.org/10.1145/2382196.2382283`.

[16] J. B. Grizzard, V. Sharma, C. Nunnery, B. B. Kang, and D. Dagon. "Peer-to-peer Botnets: Overview and Case Study". In: *Proceedings of the First Conference on First Workshop on Hot Topics in Understanding Botnets*. HotBots'07. Cambridge, MA: USENIX Association, 2007, pp. 1–1. URL: `http://dl.acm.org/citation.cfm?id=1323128.1323129`.

[19] C. Herley and D. Florêncio. "Nobody Sells Gold for the Price of Silver: Dishonesty, Uncertainty and the Underground Economy". English. In: *Economics of Information Security and Privacy*. Ed. by T. Moore, D. Pym, and C. Ioannidis. Springer US, 2010, pp. 33–53. ISBN: 978-1-4419-6966-8. DOI: `10.1007/978-1-4419-6967-5_3`. URL: `http://dx.doi.org/10.1007/978-1-4419-6967-5_3`.

[20] T. J. Holt. "The Attack Dynamics of Political and Religiously Motivated Hackers". In: *Cyber Infrastructure Protection* (2009), pp. 161–182.

[21] T. J. Holt and B. H. Schell. *Corporate Hacking and Technology-Driven Crime: Social Dynamics and Implications*. Information Science Reference, 2010.

[22] T. J. Holt, D. Strumsky, O. Smirnova, and M. Kilger. "Examining the Social Networks of Malware Writers and Hackers". In: *International Journal of Cyber Criminology* 6.1 (2012), pp. 891–903.

[23] T. Holz, M. Engelberth, and F. Freiling. *Learning more about the underground economy: A case-study of keyloggers and dropzones*. Springer, 2009.

[25] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxson, and S. Savage. “Spamalytics: An Empirical Analysis of Spam Marketing Conversion”. In: *Proceedings of the 15th ACM Conference on Computer and Communications Security*. CCS ’08. Alexandria, Virginia, USA: ACM, 2008, pp. 3–14. ISBN: 978-1-59593-810-7. DOI: `10.1145/1455770.1455774`. URL: `http://doi.acm.org/10.1145/1455770.1455774`.

[26] V. Kotov and F. Massacci. “Anatomy of Exploit Kits. Preliminary Analysis of Exploit Kits as Software Artefacts”. In: *Proc. of ESSoS 2013*. 2013.

[27] N. Kshetri. “Positive externality, increasing returns, and the rise in cybercrimes.” In: *Commun. ACM* 52.12 (Dec. 8, 2009), pp. 141–144. URL: `http://dblp.uni-trier.de/db/journals/cacm/cacm52.html#Kshetri09`.

[28] N. Kshetri. “Positive externality, increasing returns, and the rise in cybercrimes”. In: *Communications of the ACM* 52.12 (2009), pp. 141–144.

[29] N. Kshetri. “The simple economics of cybercrimes”. In: *Security & Privacy, IEEE* 4.1 (2006), pp. 33–39.

[30] P. T. Leeson and C. J. Coyne. “Economics of Computer Hacking, The”. In: *JL Econ. & Pol’y* 1 (2005), p. 511.

[31] Z. Li, Q. Liao, and A. Striegel. “Botnet economics: uncertainty matters”. In: *Managing Information Risk and the Economics of Security*. Springer, 2009, pp. 245–267.

[34] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker. “An Analysis of Underground Forums”. In: *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*. IMC ’11. Berlin, Germany: ACM, 2011, pp. 71–80. ISBN: 978-1-4503-1013-0. DOI: `10.1145/2068816.2068824`. URL: `http://doi.acm.org/10.1145/2068816.2068824`.

[35] T. Ormerod, L. Wang, M. Debbabi, A. Youssef, H. Binsalleeh, A. Boukhtouta, and P. Sinha. “Defaming botnet toolkits: A bottom-up approach to mitigating the threat”. In: *Emerging Security Information Systems and Technologies (SECURWARE), 2010 Fourth International Conference on*. IEEE. 2010, pp. 195–200.

[36] N. Provos, P. Mavrommatis, M. A. Rajab, and F. Monrose. “All Your iFRAMEs Point to Us”. In: *Proceedings of the 17th Conference on Security Symposium*. SS’08. San Jose, CA: USENIX Association, 2008, pp. 1–15. URL: `http://dl.acm.org/citation.cfm?id=1496711.1496712`.

[37] J. Radianti, J. J. Gonzalez, and E. Rich. “A quest for a framework to improve software security: Vulnerability black markets scenario”. In: *Proceedings of the the 27th International Conference of the System Dynamics Society*. 2009.

[38] J. Radianti, E. Rich, and J. J. Gonzalez. “Using a Mixed Data Collection Strategy to Uncover Vulnerability Black Markets”. In: *Workshop for Information Security and Privacy*. Citeseer. 2007.

[39] J. Radianti and N. Ulltveit-Moe. “Classification of Malicious Tools in Underground Markets for Vulnerabilities”. In: *Norsk informasjonssikkerhetskonferanse (NISK), Kristiansand, Norway* (2008).

[40] W. Shim, L. Allodi, and F. Massacci. “Crime Pays If You Are Just an Average Hacker”. In: *Cyber Security (CyberSecurity), 2012 International Conference on*. 2012, pp. 62–68. DOI: 10.1109/CyberSecurity.2012.15.

[41] A. Shulman. “The underground credentials market”. In: *Computer Fraud & Security* 2010.3 (2010), pp. 5–8.

[42] A Sood, R. Enbody, and R. Bansal. “Cybercrime: Dissecting the State of Underground Enterprise”. In: (2013).

[43] A. K. Sood and R. J. Enbody. “Crimeware-as-a-service—A survey of commoditized crimeware in the underground market”. In: *International Journal of Critical Infrastructure Protection* 6.1 (2013), pp. 28 –38. ISSN: 1874-5482. DOI: http://dx.doi.org/10.1016/j.ijcip.2013.01.002. URL: http://www.sciencedirect.com/science/article/pii/S1874548213000036.

[44] B. Stone-Gross, M. Cova, B. Gilbert, R. Kemmerer, C. Kruegel, and G. Vigna. “Analysis of a botnet takeover”. In: *Security & Privacy, IEEE* 9.1 (2011), pp. 64–72.

[45] B. Stone-Gross, T. Holz, G. Stringhini, and G. Vigna. “The underground economy of spam: A botmasters perspective of coordinating large-scale spam campaigns”. In: *USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*. 2011.

[48] M. Van Eeten, J. Bauer, J. Groenewegen, and W. Lemstra. “The economics of malware”. In: TPRC. 2007.

[49] S. Vömel, T. Holz, and F. Freiling. *I’d like to pay with your Visa Card: an illustration of illicit online trading activity in the underground economy.* Universität Mannheim/Institut für Informatik, 2010.

[50] J.-M. Yang, R. Cai, Y. Wang, J. Zhu, L. Zhang, and W.-Y. Ma. “Incorporating Site-level Knowledge to Extract Structured Data from Web Forums”. In: *Proceedings of the 18th International Conference on World Wide Web*. WWW ’09. Madrid, Spain: ACM, 2009, pp. 181–190. ISBN: 978-1-60558-487-4. DOI: 10.1145/1526709.1526735. URL: http://doi.acm.org/10.1145/1526709.1526735.

[51] M. Yip, N. Shadbolt, and C. Webber. “Structural analysis of online criminal social networks”. In: *Intelligence and Security Informatics (ISI), 2012 IEEE International Conference on*. 2012, pp. 60–65. DOI: 10.1109/ISI.2012.6284092.

[52] M. Yip, N. Shadbolt, and C. Webber. “Why forums? An empirical analysis into the facilitating factors of carding forums”. In: (2013).

[53] Z. Zhao, G.-J. Ahn, H. Hu, and D. Mahi. “SocialImpact: Systematic Analysis of Underground Social Dynamics”. In: *Computer Security ESORICS 2012*. Ed. by S. Foresti, M. Yung, and F. Martinelli. Vol. 7459. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012, pp. 877–894. ISBN: 978-3-642-33166-4. DOI: 10.1007/978-3-642-33167-1_50. URL: http://dx.doi.org/10.1007/978-3-642-33167-1_50.

[54] J. Zhuge, T. Holz, C. Song, J. Guo, X. Han, and W. Zou. *Studying malicious websites and the underground economy on the Chinese web*. Springer, 2009.

Others

[13] M. Fossi, E. Johnson, D. Turner, T. Mack, J. Blackbird, D. McKinney, M. K. Low, T. Adams, M. P. Laucht, and J. Gough. “Symantec report on the underground economy”. In: *Symantec Corporation* (2008).

[17] L. Gundert and M. van den Berg. *A Criminal Perspective On Exploit Packs.* Team Cymru Business Intelligence Team. 2011.

[18] M. Hanley, A. P. Moore, D. M. Cappelli, and R. F. Trzeciak. *Spotlight On: Malicious Insiders with Ties to the Internet Underground Community.* 2009.

[24] F. Howard. *Sophos Technical Paper: Exploring the Blackhole Exploit Kit.* SophosLabs. 2012.

[32] M. McCormack. “Swiss police raid underground bulletin boards”. In: *COMPUTER FRAUD AND SECURITY* (1996), p. 4.

[33] Microsoft. *Security Intelligence Report Volume 14.* Tech. rep. Microsoft, 2012.

[46] Symantec. *Analysis of Malicious Web Activity by Attack Toolkits.* Online. Accessed on June 1012. Symantec. 2011.

[47] Symantec. *Symantec Corporation Internet Security Threat Report 2013.* Tech. rep. 18. 2012 Trends, 2013.

Appendices

Appendix A

Automatic pattern matching rule for identifying Rippers

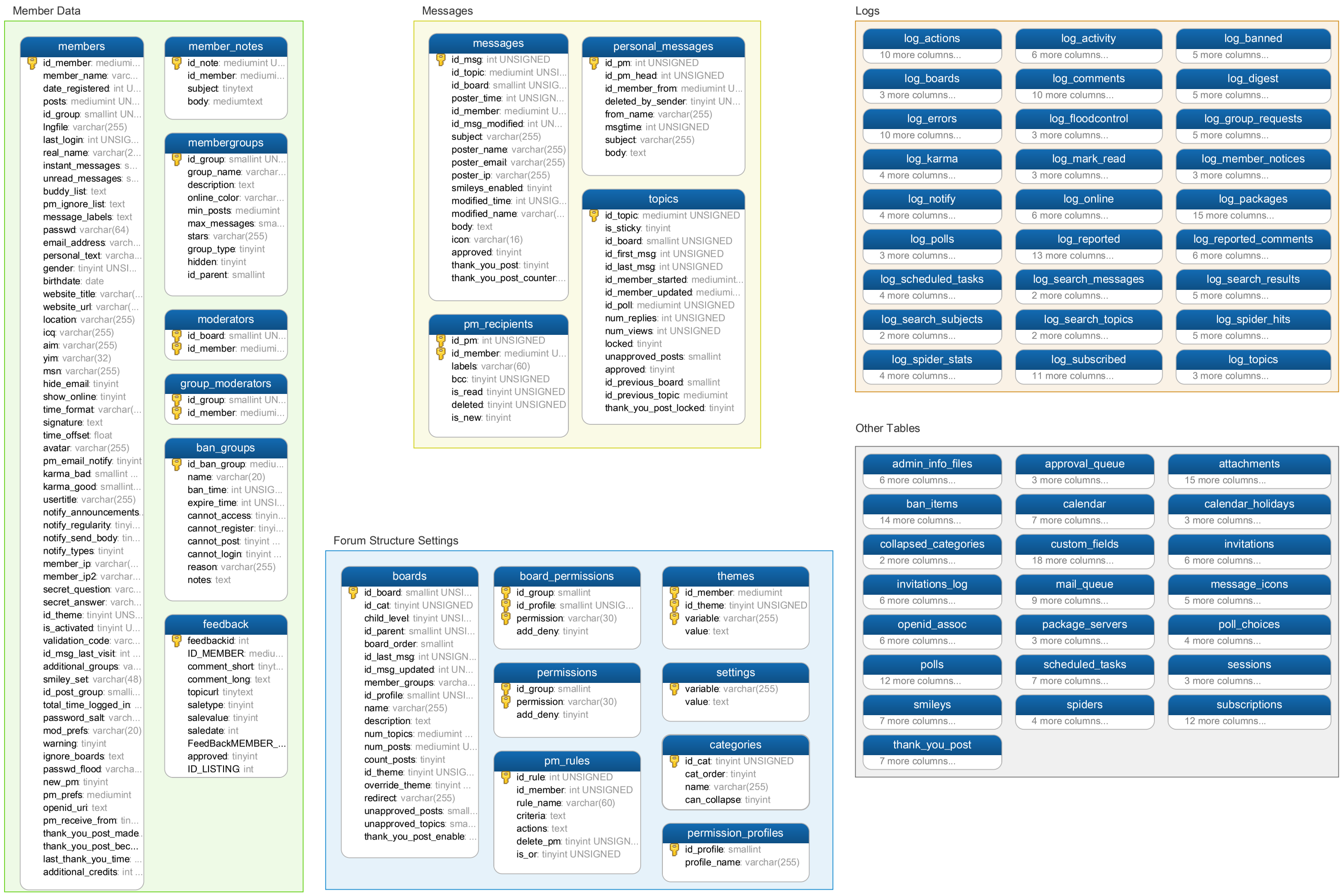
Pattern matching rule: Rippers

Identify rippers by looking at banned users

<table>
<tr><td>Database:</td><td>smf_carders_cc</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Table:</td><td>carders_smf_ban_groups</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Columns:</td><td>reason</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Categories:</td><td>Rippers</td><td>Spammers</td><td>Double account</td><td>Forum rules violatio</td><td>No reason</td></tr>
<tr><td>Patterns:</td><td>ripper</td><td>spammers</td><td>doppel</td><td>child</td><td></td></tr>
<tr><td></td><td>rippe</td><td>spam</td><td>doppel-account</td><td>children</td><td></td></tr>
<tr><td></td><td>ripp</td><td>spamming</td><td>double</td><td>porn</td><td></td></tr>
<tr><td></td><td>rip</td><td>spamm</td><td>double-account</td><td>swear</td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td>mouth</td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Query:</td><td>SELECT COUNT(*) FROM carders_smf_ban_groups WHERE reason REGEXP 'ripper|rippe|ripp|rip'</td><td>SELECT COUNT(*) FROM carders_smf_ban_groups WHERE reason REGEXP 'spammers|spam|spamming|spamm'</td><td>SELECT COUNT(*) FROM carders_smf_ban_groups WHERE reason REGEXP 'doppel|doppel-account|double|double-account'</td><td>SELECT COUNT(*) FROM carders_smf_ban_groups WHERE reason REGEXP 'child|children|porn|swear|mouth'</td><td>SELECT COUNT(*) FROM carders_smf_ban_groups WHERE reason=""</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Result:</td><td>270</td><td>109</td><td>301</td><td>4</td><td>83</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Total banned user</td><td>1296</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Total pattern matc</td><td>767</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Uncategorized:</td><td>529</td><td></td><td></td><td></td><td></td></tr>
</table>

Appendix B

Diagram Database of Carders.CC


Member Data
members
id_member: mediumi...
member_name: varc...
date_registered: int U...
posts: mediumint UN...
id_group: smallint UN...
lngfile: varchar(255)
last_login: int UNSIG...
real_name: varchar(2...
instant_messages: s...
unread_messages: s...
buddy_list: text
pm_ignore_list: text
message_labels: text
passwd: varchar(64)
email_address: varch...
personal_text: varcha...
gender: tinyint UNSI...
birthdate: date
website_title: varchar(...
website_url: varchar(...
location: varchar(255)
icq: varchar(255)
aim: varchar(255)
yim: varchar(32)
msn: varchar(255)
hide_email: tinyint
show_online: tinyint
time_format: varchar(...
signature: text
time_offset: float
avatar: varchar(255)
pm_email_notify: tinyint
karma_bad: smallint ...
karma_good: smallint...
usertitle: varchar(255)
notify_announcements...
notify_regularity: tinyi...
notify_send_body: tin...
notify_types: tinyint
member_ip: varchar(...
member_ip2: varchar...
secret_question: varc...
secret_answer: varch...
id_theme: tinyint UNS...
is_activated: tinyint U...
validation_code: varc...
id_msg_last_visit: int ...
additional_groups: va...
smiley_set: varchar(48)
id_post_group: smalli...
total_time_logged_in: ...
password_salt: varch...
mod_prefs: varchar(20)
warning: tinyint
ignore_boards: text
passwd_flood: varcha...
new_pm: tinyint
pm_prefs: mediumint
openid_uri: text
pm_receive_from: tin...
thank_you_post_made...
thank_you_post_bec...
last_thank_you_time: ...
additional_credits: int ...
member_notes
id_note: mediumint U...
id_member: mediumi...
subject: tinytext
body: mediumtext
membergroups
id_group: smallint UN...
group_name: varchar...
description: text
online_color: varchar...
min_posts: mediumint
max_messages: sma...
stars: varchar(255)
group_type: tinyint
hidden: tinyint
id_parent: smallint
moderators
id_board: smallint UN...
id_member: mediumi...
group_moderators
id_group: smallint UN...
id_member: mediumi...
ban_groups
id_ban_group: mediu...
name: varchar(20)
ban_time: int UNSIG...
expire_time: int UNSI...
cannot_access: tinyin...
cannot_register: tinyi...
cannot_post: tinyint ...
cannot_login: tinyint ...
reason: varchar(255)
notes: text
feedback
feedbackid: int
ID_MEMBER: mediu...
comment_short: tinyt...
comment_long: text
topicurl: tinytext
saletype: tinyint
salevalue: tinyint
saledate: int
FeedBackMEMBER_...
approved: tinyint
ID_LISTING: int
Messages
messages
id_msg: int UNSIGNED
id_topic: mediumint UNSI...
id_board: smallint UNSIG...
poster_time: int UNSIGN...
id_member: mediumint U...
id_msg_modified: int UN...
subject: varchar(255)
poster_name: varchar(255)
poster_email: varchar(255)
poster_ip: varchar(255)
smileys_enabled: tinyint
modified_time: int UNSIG...
modified_name: varchar(...
body: text
icon: varchar(16)
approved: tinyint
thank_you_post: tinyint
thank_you_post_counter...
personal_messages
id_pm: int UNSIGNED
id_pm_head: int UNSIGNED
id_member_from: mediumint U...
deleted_by_sender: tinyint UN...
from_name: varchar(255)
msgtime: int UNSIGNED
subject: varchar(255)
body: text
pm_recipients
id_pm: int UNSIGNED
id_member: mediumint U...
labels: varchar(60)
bcc: tinyint UNSIGNED
is_read: tinyint UNSIGNED
deleted: tinyint UNSIGNED
is_new: tinyint
topics
id_topic: mediumint UNSIGNED
is_sticky: tinyint
id_board: smallint UNSIGNED
id_first_msg: int UNSIGNED
id_last_msg: int UNSIGNED
id_member_started: mediumint...
id_member_updated: mediumi...
id_poll: mediumint UNSIGNED
num_replies: int UNSIGNED
num_views: int UNSIGNED
locked: tinyint
unapproved_posts: smallint
approved: tinyint
id_previous_board: smallint
id_previous_topic: mediumint
thank_you_post_locked: tinyint
Forum Structure Settings
boards
id_board: smallint UNSI...
id_cat: tinyint UNSIGNED
child_level: tinyint UNSI...
id_parent: smallint UNSI...
board_order: smallint
id_last_msg: int UNSIGN...
id_msg_updated: int UN...
member_groups: varcha...
id_profile: smallint UNSI...
name: varchar(255)
description: text
num_topics: mediumint ...
num_posts: mediumint U...
count_posts: tinyint
id_theme: tinyint UNSIG...
override_theme: tinyint ...
redirect: varchar(255)
unapproved_posts: small...
unapproved_topics: sma...
thank_you_post_enable: ...
board_permissions
id_group: smallint
id_profile: smallint UNSIG...
permission: varchar(30)
add_deny: tinyint
permissions
id_group: smallint
permission: varchar(30)
add_deny: tinyint
pm_rules
id_rule: int UNSIGNED
id_member: int UNSIGNED
rule_name: varchar(60)
criteria: text
actions: text
delete_pm: tinyint UNSIGN...
is_or: tinyint UNSIGNED
themes
id_member: mediumint
id_theme: tinyint UNSIGNED
variable: varchar(255)
value: text
settings
variable: varchar(255)
value: text
categories
id_cat: tinyint UNSIGNED
cat_order: tinyint
name: varchar(255)
can_collapse: tinyint
permission_profiles
id_profile: smallint
profile_name: varchar(255)
Logs
log_actions
10 more columns...
log_activity
6 more columns...
log_banned
5 more columns...
log_boards
3 more columns...
log_comments
10 more columns...
log_digest
5 more columns...
log_errors
10 more columns...
log_floodcontrol
3 more columns...
log_group_requests
5 more columns...
log_karma
4 more columns...
log_mark_read
3 more columns...
log_member_notices
3 more columns...
log_notify
4 more columns...
log_online
6 more columns...
log_packages
15 more columns...
log_polls
3 more columns...
log_reported
13 more columns...
log_reported_comments
6 more columns...
log_scheduled_tasks
4 more columns...
log_search_messages
2 more columns...
log_search_results
5 more columns...
log_search_subjects
2 more columns...
log_search_topics
2 more columns...
log_spider_hits
5 more columns...
log_spider_stats
4 more columns...
log_subscribed
11 more columns...
log_topics
3 more columns...
Other Tables
admin_info_files
6 more columns...
approval_queue
3 more columns...
attachments
15 more columns...
ban_items
14 more columns...
calendar
7 more columns...
calendar_holidays
3 more columns...
collapsed_categories
2 more columns...
custom_fields
18 more columns...
invitations
6 more columns...
invitations_log
6 more columns...
mail_queue
9 more columns...
message_icons
5 more columns...
openid_assoc
6 more columns...
package_servers
3 more columns...
poll_choices
4 more columns...
polls
12 more columns...
scheduled_tasks
7 more columns...
sessions
3 more columns...
smileys
7 more columns...
spiders
4 more columns...
subscriptions
12 more columns...
thank_you_post
7 more columns...

Appendix C

Overview of an alternative successful market

In this appendix we provide an introductory overview of the organization of another, still active and arguably well-functioning cybercrime market. Because the market is running, we refrain from explicitly state its name in this manuscript. It is a market for exploits, botnets and malware. It is also one of the main markets that introduced exploit-as-a-service [15] in the cyberthreat scenario, as we find there the main players and products that the industry reports be driving the majority of reported web-attacks [47]. Indirect evidence of this markets' efficacy is the recent burst in cyberattacks driven by means of tools, services and infrastructures traded or rented in these markets [15, 31, 44, 2]. In this case we do not have an SQL dump of the market, but we will provide instead first-hand evidence that the problems we highlighted for **Carders.CC** are not present here. For the purpose of this thesis we will only focus on a fraction of the characteristics of this market, that will serve as a comparison to our analysis on **Carders.CC**: the reputation mechanism, and the punishment mechanism. All this is documented and referenced to in the format [$CODEn$], with $CODE$ being an internal code we used to classify the evidence and n begin the document number. Interested researchers can contact the authors to have access to any document referenced in the following.

C.1 Reputation

Reputation points are attributed to users by other users after a positive or negative interaction between the two [DMN 6]. Of course, such system is subject to abuse; for example, a user may want to lower his competitors' reputation level to improve the competitiveness of their own business, or create fake accounts on the market to provide "collective" negative feedback. This adversarial behavior is limited by the mechanism's implementation rules: *"Only users with more than 30 posts can change reputation. Only 5 +/- reputation points per day can be assigned by any user to other users."* [DMN 6]. This effectively places an upper bound in the number of reputation points one may assign in a given day and decreases one's influence over the overall distribution of reputation points in the market. This by itself may largely overcome the obvious problem **Carders.CC**

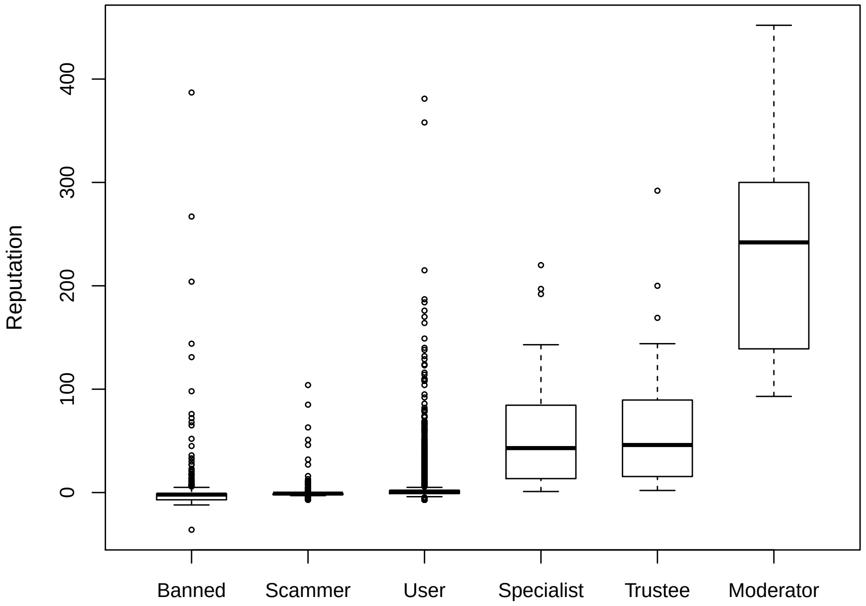


Figure C.1: Boxplot representation of reputation distribution among categories

had.

The forum regulation outlines seven "reputation groups" [DMN 5]. The following presents these groups in descending order of trustworthiness, i.e. those on top of the list belong to the most reliable users in the community.

1. Admin.
2. Moderator.
3. Trustee: Here belong members of the community that *"own important services, or are moderators or administrators of other forums"* [DMN 5].
4. Specialist: Users elected in this group are considered "advanced" users" with a "high level of literacy".
5. User: Normal users.
6. Rippers: In this group belong users that have been reported and have been found guilty of "scamming". It is explicitly recommended *"to not have deals (business, work) with users of this group"* [DMN 5].
7. Banned: Users that have been precluded access to the forum.

Figure C.1 reports a boxplot representation of the distribution of reputation scores among user categories. Categories are listed in ascending order as in [ADM 5]. It is here clear that higher rankings are reflected in higher reputation levels of the users. We run a Mann-Whitney unpaired test to check if the difference in reputation levels between categories is significant, and we find that reputation levels significantly increase with higher categories. The only exception is for the Trustee and Specialist categories, for which no difference is found (which is explained by the elective nature of these categories). While this does not mean that higher reputation results in a higher ranking (as a number of endogenous factors other than reputation may be related to the inclusion in a user group - i.e. there is a self-selection problem), it does show that the reputation mechanism is effectively enforced and results in coherent distributions among users. The difference with the same analysis for **Carders.CC** is clear.

C.2 Punishment

Users can effectively report other users to the board of administrators when they think they have been scammed. The administrators remark that *"We make [cheaters] public with pleasure."* [ADM 6]. The inclusion of a user in the list of cheaters is a fairly refined process, that requires a report to be filed, an investigation to be carried, and that allows the "alleged scammer" the right to defend himself before the decision by the moderators. The whole phase takes place in a dedicated sub-community of the market, a sort of "*court of justice*" where the offended reports the (alleged, at this point) offender.

The reporting is to be filed according to a specific procedure established in the market regulation, that includes the *"name, contacts, a proof of the fact (log, screenshot of correspondace, money transfers,..) and a link to the user's profile."* Following the filing, an actual **trial** takes place. The defendant is given the opportunity to reply to the accusation. The investigation can be carried both by moderators and administrators, while the final decision usually belongs to the administrator. The community is also often active in the discussion, reporting further evidence or personal experience with the accused, or helping in the investigations. We observed and documented many of those trials. Examples are [BOU 2], [SEL 1], [INT 12]. A full description of those is out of the scope of this thesis, but on a qualitative note what we observed is that:

1. The defender always reports detailed information on the accused user and on the case of complaint.

2. Many witnesses appear in "court" giving opinions on the evolution of the case, or providing supporting evidence for either the accuser and the defender.

3. The moderators and the administrators are always present in each report, and actively regulate the discussion.

4. When the defender does not show up within the time limit specified by the administrator [DMN 6], the case always goes to the defender.

5. When the defender shows up, he/she always publish evidence of his case, being those screenshots of chats with the accuser or "webmoney" transaction logs.

6. Some cases last several months, with all parties actively participating in the discussion and new evidence being examined or asked for iteratively.

7. When the evidence provided by either of the defender or the accuser is not conclusive, the case goes to the opponent or a “null” is thrown: when neither of the two is convincing, nobody wins.

8. Users that end up being found guilty are *always* exposed in the list of cheaters and/or are banned from the forum.

C.3 Discussion

The organizational and structural differences of this operating market with respect to **Carders.CC** is evident. The reputation and punishment mechanisms generate meaningful information for the user to use when he/she needs to decide with which user to trade:

1. Evidence supports the hypothesis that reputation points are meaningfully assigned to users. This arguably results in a useful tool for the user to asses what users to trade with.

2. The punishment mechanism is a well-regulated one and direct evidence from the market suggests that the “trials” are conducted in a fair manner. This is known to boost market activity and clearly incentives honest behaviour.

3. Users that have been found guilty are, if not banned, publicly exposed and regularly assigned to the “scammers” group associated with their name on the board. This allows other users to clearly assess a scammer’s trading history and make an informed decision with whom to trade.