Document Title:	Examining the Structure, Organization, and Processes of the International Market for Stolen Data
Author(s):	Thomas J. Holt, Olga Smirnova
Document No.:	245375
Date Received:	March 2014
Award Number:	2010-IJ-CX-1676

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Examining the Structure, Organization, and Processes of the International Market for Stolen Data

Award Number: 2010-IJ-CX-1676

Thomas J. Holt
Michigan State University

Olga Smirnova
East Carolina University

Abstract

Over the last two decades, consumers have come to depend on computers and the Internet to engage in commerce and manage their finances. Businesses also rely on these technologies in order to process and maintain consumer data in massive databases. As a result, there has been a substantial increase in the risk of theft and fraud stemming from cybercriminals who can compromise these resources to their advantage. Recent evidence suggests that hackers who acquire sensitive consumer data sell this information to others in on-line forums for a profit. In turn, an underground economy has developed around the sale of stolen data, involving various resources that can be used to convert electronic data into real world currency and engage in various forms of cybercrime. There is, however, generally little research on the economics of these markets, or the distribution of goods sold. There is also minimal research on the social organization of market actors and the network structures present that support the stolen data market generally.

To address these questions, this study utilizes a sample of 1,899 threads generated from 13 web forums, 10 of which use Russian as their primary language and three which use English. These forums were hosted around the world, and act as online advertising spaces for individuals to sell and buy a range of products. The content of these forums were downloaded and translated from Russian to English to create a purposive, yet convenient sample of threads from each forum.

Qualitative content analyses were conducted to code the products and quantitative assessments demonstrate that the majority of products sold in these forums were some form of stolen data (84.3%). The majority of sellers offered dumps, referring to bank account or credit card data (44.7%), as well as CVV data from credit cards (34.9%) and various forms of

electronic data, such as eBay and PayPal accounts (1.4%). A number of sellers also offered resources to obtain currency from these accounts on or off-line (7.4%), and a small proportion sold malware and tools to engender cybercrimes. Products had a range of advertised prices, and the majority of stolen data came from the United States and Europe. These pricing structures were directly influenced by market forces, with differences evident in more legitimate forums with a higher degree of trust between participants.

A qualitative analysis was employed using grounded theory techniques and aspects of Best and Luckenbill's (1994) framework of deviant organization to examine the associations and working relationships present between participants at the micro and macro-level. The findings suggest that the markets are primarily collegial in nature at the individual level, enabling individuals to work together in order to facilitate transactions. There is also a distinct division of labor between participants on the basis of the products sold and skill sets available. At the macro-level, eight of the forums appear to operate as formal organizations based on managerial structures and long-term operations relative to other forms of on-line criminality.

Finally, quantitative social network analysis techniques were applied to explore the network structures present between participants within the forums in this sample. The density of the various networks were generally low, but the majority of networks had over 50 percent of users connected. Sellers were the most central actors, though buyers and neutral users posted more frequently and were critical to the facilitation of information sharing and communication of user reputation. In addition, there was a high correlation between the number of posts and the number of users in each forum. As a result, these networks appear to have substantive redundancies that make them difficult to disrupt through traditional external means of node removal.

Table of Contents

Executive Summary

Over the last two decades, consumers have come to depend on computers and the Internet to engage in commerce and manage their finances. Businesses also rely on these technologies in order to process and maintain consumer data in massive databases. As a result, there has been a substantial increase in the risk of theft and fraud stemming from cybercriminals who can compromise these resources to their advantage. Recent evidence suggests that hackers who acquire sensitive consumer data sell this information to others in on-line forums for a profit. In turn, an underground economy has developed around the sale of stolen data, involving various resources that can be used to convert electronic data into real world currency and engage in various forms of cybercrime. There is, however, generally little research on the economics of these markets, or the distribution of goods sold. There is also minimal research on the social organization of market actors and the network structures present that support the stolen data market generally.

To address these questions, this study utilizes a sample of 1,899 threads generated from 13 web forums, 10 of which use Russian as their primary language and three which use English. These forums were hosted around the world, and act as online advertising spaces for individuals to sell and buy a range of products. The content of these forums were downloaded and translated from Russian to English to create a purposive, yet convenient sample of threads from each forum.

The findings demonstrate that these forums act as advertising spaces where individuals could either sell illegally acquired data, utilize the funds to obtain liquid currency, or engage in various forms of fraud and cybercrime. Individuals would create a thread and list their products or request various items, indicating the cost of a good or service, preferred payment method, and

contact information. Most individuals preferred to communicate via ICQ, which is an instant messaging protocol, though a proportion also utilized email addresses from a variety of publicly accessible services. Payments were primarily accepted through electronic systems, including Liberty Reserve (16.8%) and WebMoney (11.1%), though some (19.5%) also accepted real world money transfer services through Western Union.

Qualitative content analyses demonstrate that the majority of products sold in these forums were some form of stolen data (84.3%). The majority of sellers offered dumps, referring to bank account or credit card data (44.7%), as well as CVV data from credit cards (34.9%) and various forms of electronic data, such as eBay and PayPal accounts (1.4%). A number of sellers also offered resources to obtain currency from these accounts on or off-line (7.4%), and a small proportion sold malware and tools to engender cybercrimes.

Products had a range of advertised prices, though most data was offered at a lower general cost than data manipulation services. In fact, the average advertised price for dumps was $102.60, while drops services that can be used to obtain funds from accounts had an average cost of $192.37. Given the substantial proportion of data sold in these forums, further consideration was given to the distribution of data by card type and country of origin. Of those sellers who listed the type of card in their advertisements, the majority came from Visa and MasterCard in keeping with their general market share around the world. The majority of data also originated from European nations (bank accounts, 40.3%; CVV, 29.0%; dumps, 29.6%; fullz, 41.3%), though Canada, the United States, and United Kingdom were also substantially victimized depending on the type of data sold. This pattern was somewhat consistent with the price paid for various types of data, as those nations with the largest representation in the market tended to have the lower average pricing. For instance, bank account data from the United Kingdom

($4.08) and the European Union ($4.12) were the least expensive overall, while U.S. data ($5.33) was the most costly. At the same time, the mean price for CVVs were the least expensive in the United States ($1.67). In addition, the cost for dumps ($2.81) and fullz ($3.34) were lowest in the UK, followed by the US ($3.04 and $3.47 respectively).

The social nature of the forums allows buyers to provide feedback about their experience with sellers in the thread. Positive feedback allows individuals to note successful encounters, promote successful sellers, and potentially increase the seller's reputation. Negative feedback is also critical so that prospective buyers avoid risky transactions and reduce the likelihood of loss. There are no formal dispute resolution services available for individuals who are dissatisfied with their experience with a seller, thus negative feedback is vital for individuals to affect disreputable sellers. In these instances, buyers typically used the term "ripper" to indicate that a seller is untrustworthy or attempting to cheat others. The application of this term is means to negatively affect the seller's reputation within the market.

In light of the role of customer feedback, two of the forums in this sample had a substantial number of complaints of ripping. Exploring the distribution of products on the basis of ripping demonstrates that there is a difference in the nature of data sold in forums. Specifically, ripping forums sold the majority of CVV data in this data set, though dumps were evenly distributed across both types. Non-ripping forums also offered a larger proportion of malware, data, and various mechanisms to remove funds from stolen accounts.

To further examine the role of forum dynamics on the advertised cost for stolen data, two logistic regression models were created using the log advertised price for dumps and eBay and PayPal accounts as the dependent variable respectively. The findings suggest that the use of Western Union payments were associated with higher prices for both products, as were dumps sold in forums that tested sellers' data. The use of free samples were, however, associated with lower prices for dumps which may be associated with ripping complaints and show a lack of trust in the sellers' products. Individuals who advertised in other people's threads were also associated with lower prices for both dumps and eBay/PayPal credentials. Dumps sold in ripping forums were also associated with lower advertised prices, as were products sold in Russian language forums.

A qualitative analysis was employed using grounded theory techniques and aspects of Best and Luckenbill's (1994) framework of deviant organization to examine the associations and working relationships present between participants at the micro and macro-level. The findings suggest that the markets are collegial in nature at the individual level, as buyers and sellers must work together in order to facilitate transactions. Though sales are structured between individuals outside of the forums, the market is also driven by mutual participation through the use of feedback in order to promote reputable sellers and identify rippers. A clear division of labor was also evident based on the range of products sold across all markets. At the macro-level, the forums also had a division of labor on the basis of moderators who managed the forum spaces through the use of bans and product testing in order to validate seller's advertisements. Eight of the forums also persisted for more than two years, suggesting that they may be formal organizations due to their operation over time.

Finally, quantitative social network analysis techniques were applied to explore the links and network structures present between participants within each forum. Generally, the majority of the networks were not dense, reflecting the redundancies present which may stem from the preponderance of service providers and vendors present within each site. The participants were also not well connected due to the generally low level of participation in the majority of threads sampled. Only three of the forums in this sample had more than 70 percent of all users connected through various threads. In addition, the most central users across the forums tended to be sellers due in part to the sales-oriented nature of the forums. Individuals who either made purchases or asked questions made the largest number of posts compared to sellers, ensuring the flow of information and development of reputations over time. In fact, there was a high correlation between the number of posts and the number of users in each forum. The relative lack of hierarchical structures or relationships observed suggests that these markets are inefficient and may be resistant to the removal of individual actors. The participatory and collegial nature of the forums may make it difficult to disrupt them through the arrest or elimination of central sellers.

As a whole, the analyses presented here demonstrate that the market for stolen data is a real threat to consumers and businesses alike. Victims from around the world can be harmed by the sale of personal information to facilitate identity theft, while financial service providers must reimburse victims for economic damages. The massive number of data sellers and the general pricing structures observed suggest that there is no easy or immediate way to disrupt or deter offenders engaged in these markets. There are low costs to entry and the variety of sellers enables individuals with various levels of skill to engage in transactions within the market. In light of the absence of key hierarchical management structures or evidence of corruption or physical violence, they do not appear to be structured in the same way as traditional ethnic organized crime groups like the Italian mafia or Yakuza. Instead, the market appears to be comprised of a network of international cybercriminals operating in a collegial fashion to further individual interests. As such, any attempt to affect the market requires a substantive law enforcement response along with increased consumer awareness on the risks of victimization.

Considering the structure of these networks and their organizational complexity, one of the most effective mechanisms to disrupt the sale of stolen data may be through the use of investigations against the payment systems used by participants. Investigations against WebMoney, Liberty Reserve, and other online payment systems may be able to affect the flow of money and reduce the efficiency of any transaction. In addition, it may be more prudent for law enforcement agencies to establish forums to foster the sale of stolen data using undercover identities. This tactic is a shift from traditional cases built against investigations of single actors, but may generate more substantial cases against entire networks of participants and sew mistrust among participants across the community of hackers and data thieves generally. In addition, there is a precedent for this type of investigation, as with the recent takedown of two forums operated in part by law enforcement agencies.

Finally, there is a need to increase the technological, investigative, and support resources within federal law enforcement agencies. This includes linguists and translators who understand both technical language and the jargon and slang common to market actors to correctly build the network structures between actors and document the markets generally. To accomplish this, there is a need to potentially increase funding to the Federal Bureau of Investigation, Secret Service, the Department of Treasury, the Department of Homeland Security and other federal agencies to ensure a more robust response to financially-driven cybercrimes and the actors responsible for attacks around the world.

There is also a need for careful revision and adjustment of cooperative agreements to facilitate the international investigation and prosecution of data thieves (Brenner 2008). The findings of this study demonstrate that participants are compromising banks, businesses, and citizens in the US and European Union. The forums themselves are hosted around the world, and the participants may be native to the Russian Federation or Russian speakers living abroad. As a result, it is vital that the Department of Justice and law enforcement agencies find ways to improve existing extradition treaties and cooperative frameworks with various agencies, such as the Russian FSB, to ensure that responsible actors may be detected and brought to justice.

There is also a need for improved awareness of the risks of electronic identity theft among home computer users who do not necessarily have a strong grasp of basic computer security principles. The risk for data theft may stem from individual behaviors such as responding to a phishing email or an active malware infection on one's home computer, or even downloading a rogue banking application for a smart phone or tablet. Users who are cognizant of these risks may still be affected through a mass data breach that affects millions of card holders.

Since there is no single way to reduce individual risk of harm, there is a need for public awareness campaigns to promote basic computer security principles and vigilance against identity theft. Consumers who understand the potential harm that can result from responding to unsolicited email, clicking on suspicious web links, and the need to run anti-virus and security tools may decrease their risk of victimization. At the same time, these campaigns should promote the need to regularly check bank and credit card statements for suspect charges and avoid making purchases through on-line vendors with no security features in place to protect personal information. Such information could be rolled out effectively through the Internet Crime Complaint Center and Federal Trade Commission, particularly during October's CyberSecurity Month, and may promote general security and diminish the scope of compromises.

To further promote security and increase corporate responsibility in the event of large scale data breaches that are beyond consumer control, there is a need for increased adoption of data breach legislation. Currently, 46 states require that companies disclose any loss of sensitive personal information to consumers in the event of a security breach. It is often difficult to determine the true scope of these beaches and determine how many customers actually face economic harm as a result of any such incident. Thus, greater transparency is needed on the part of both corporations and financial institutes to disclose the true number of customers affected and to what degree in as timely a fashion as is possible in order to reduce the risk of customer loss and economic harm generally.

Finally, there is a need for continuing research on the organizational and economic impact of stolen data markets. The network structures and relationships observed here may not be static, and many of the transactions between actors are hidden from view. Collaborations with law enforcement, financial institutions, and researchers are needed to gather private messages and ICQ communications that can be connected to forum data sources to create more accurate economic models and network structures between market actors. Longitudinal data sets are also necessary to understand the way that network structures change over time. Capturing multiple sub-forums within multiple forums and tracking seller and buyer usernames over time would allow for the development of complex network models of change and assessments of network centrality and density. In turn, this would allow researchers to identify if and when markets begin to transition from collegial structures to more organized and efficient marketplaces. Such information is vital to understand the nature of stolen data markets and their role in cybercrime and fraud globally.

I. Introduction

The Internet and World Wide Web have drastically changed the way businesses, government, and citizens communicate and conduct business globally (see Jewkes & Sharpe, 2003; Wall, 2001). Businesses now depend on the web to solicit customers and make sales. The banking and financial services sector utilizes these technologies to provide customers with full access to their funds and accounts with relative ease, at all hours of the day, from any location (James, 2005; Newman & Clarke, 2003). Home computer users can now use this technology around the clock with home-based high-speed dedicated Internet access through simple-to-use computers and mobile devices to connect to various resources (Brenner, 2008; Wall, 2007).

These innovations have significant benefits, but also create significant risks for fraud and theft as sensitive data, such as bank and credit card account numbers, personal information, and other electronic files (Allison, Shuck, & Learsch, 2005; Chu, Holt, & Ahn, 2010; Furnell, 2002; Holt & Turner, 2010; James, 2005; Morris, 2010; Newman & Clarke, 2003). The electronic databases, managed by businesses and financial institutions to store sensitive customer information, can be accessed and compromised by hackers to quickly and efficiently steal massive amounts of information (Newman & Clarke, 2003; Peretti, 2009; Wall, 2007). In fact, Heartland Payment Systems announced in 2009 that a small group of hackers were able to acquire information from 130 million credit and debit cards processed by 100,000 businesses (Verini, 2010). This was the largest breach of customer data in the United States, though the total number of high profile data breaches has increased significantly over the last four years (Verison, 2012).

The increased use of online retailers and banking services also increase the risk of theft by allowing consumers to transmit sensitive personal and financial information over the Internet

(James, 2005; Newman & Clarke, 2003). This information can be surreptitiously obtained by criminals through various methods, most notably phishing (James, 2005; Wall, 2007). In a phishing attack, consumers are tricked into transmitting financial information to fraudulent websites where the information is stored for later use or resale by the offender (see James, 2005; Wall, 2007). These crimes are particularly costly for victims and financial institutions alike (Anti-Phishing Working Group, 2012), as the Gartner Group estimates that phishing victims in the US lost $3 billion in 2007 alone (Rogers, 2007).

In light of the growing prominence of electronic data theft and the significant financial impact these crimes have for victims and compromised companies, it is critical that researchers consider the means by which cybercriminals dispose of the data they obtain. An emerging body of research has begun to examine this problem through the identification of online stolen data markets where computer criminals buy and sell information (Chu et al., 2010; Dhanjani & Rios, 2008; Franklin, Paxson, Perrig, & Savage, 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Holz, Engelberth, & Freling, 2009; Honeynet Research Alliance, 2003; Motoyama, McCoy, Levchenko, Savage, & Voelker, 2011; Thomas & Martin, 2006; Wehinger, 2011). These studies primarily focus on Internet Relay Chat, or IRC, channels and networks where hackers sell significant volumes of data obtained through phishing, database compromises, and other means (Franklin et al., 2007; Holz et al., 2009; Honeynet Research Alliance, 2003; Thomas & Martin, 2006). A small number of studies have also begun to explore the role of web forums in the formation of markets for stolen data, which operate through HTML-based web browsers rather than IRC clients (Chu et al., 2010; Holt & Lampke, 2010; Motoyama et al., 2011). In fact, forum-based markets may be more popular in supporting stolen data based on the recent large -

scale law enforcement operations against various forums operating in the United States (see Peretti, 2009; Symantec, 2012).

Regardless of the communications environment, evidence suggests that stolen data markets primarily facilitate the sale of credit card and bank account information, Personal Identification Numbers (PINs), and supporting customer information obtained from victims around the world, in batches of tens or hundreds of accounts (Chu et al., 2010; Franklin et al., 2007; Holt & Lampke, 2010; Honeynet Research Alliance, 2003; Thomas & Martin, 2006). Although financial service providers from around the world are compromised, the bulk of stolen data sold in these markets appears to come from the United States, followed by various European nations (Franklin et al., 2007; Holt & Lampke, 2010). Also, Visa and MasterCard products compose the largest percentage of financial products stolen and sold (see Franklin et al., 2007; Holt & Lampke, 2010). As a consequence, there are an overwhelming number of US citizens and financial institutions who are directly victimized as a result of their information appearing in stolen data markets.

Online data markets also enable individuals to directly market their services to engage in various forms of identity-based offenses and cybercrimes. Ads are regularly posted by hackers who will engage in spam and phishing campaigns for a fee, enabling criminals to directly profit from sophisticated forms of cybercrime with no direct application of skill or ability (see Chu et al., 2010; Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Honeynet Research Alliance, 2003). Sellers also offered cash-out services designed to obtain money from electronic accounts through direct withdrawals at Automatic Teller Machines in the real world or through fencing goods purchased online using stolen cards (Chu et al., 2010; Franklin et al., 2007; Holt & Lampke, 2010; Honeynet Research Alliance, 2003; Wehinger, 2011).

Additionally, identity documents, including passports and drivers license documents, are sold to facilitate identity theft on and off-line (Chu et al., 2010; Franklin et al., 2007; Holt & Lampke, 2010; Honeynet Research Alliance, 2003; Wehinger, 2011). Thus, online stolen data markets provide access to fraudulently obtained information and multiple methods to use this information for identity theft and crime.

Though these studies provide an important glimpse into the nature of the online market for stolen data, little research has examined the economic structure, social organization, and market dynamics that drive this criminal activity. In addition, the majority of studies utilize publicly accessible IRC channel data, which has led some to question whether these markets are actually reliable sources of data or an entry point for new participants in the market who are not well versed in the process of data theft (Herley & Florencio 2010; Wehinger, 2011). There is limited knowledge on the processes and social dynamics of closed web forum communities, whose operations are hidden from the general public. Furthermore, most active stolen data markets currently operate primarily via websites hosted in foreign nations, whose users communicate in Russian characters rather than English (Chu et al., 2010; Dunn, 2010; Symantec, 2012). As a result, there is a need for a systematic examination of stolen data markets to improve our knowledge of cybercrime and high-tech identity theft, and the law enforcement and computer security responses to these threats (Herley & Florencio, 2010; Wehinger, 2011).

Specifically, few studies explore the pricing structures for data and services within the market in general (Herley & Florencio, 2010). Though many studies estimate the number of products sold (Dhanjani & Rios, 2008; Franklin et al., 2007; Holz et al., 2009; Motoyama et al., 2011), few discuss the prices for data (except Holt & Lampke, 2010). In fact, Franklin et al. (2007) and Holz and colleagues (2009) utilize data from Symantec on the pricing for various data sold rather than the prices indicated in their data sources (see Herley & Florencio, 2010). Thus, there is a need for research examining the economics of the market for stolen data, including any social and economic factors that may affect the advertised price of data sold in the markets.

In addition, there has been generally little research on the organizational practices of actors within stolen data markets (Motoyama et al., 2012; Yip, Webber, & Shadbolt, 2013). This is due in part to the range of data sources used, and inconsistent operationalizations of the unit of analysis for organization (see Holt, 2013). Various researchers use the term "market" to describe the structure of the IRC channels and forums since they serve as a location and infrastructure for participants to engage in transactions (Franklin et al., 2007; Holt, 2013). Organizational researchers, however, define markets as a relational structure between actors that is generally uncoordinated with few commitments between participants (Apsers, 2011; Powell, 1990). As a result, it is difficult to situate the nature of social organization within the existing literature of stolen data markets generally (Wehinger, 2011).

To that end, Herley and Florencio (2010) argue that participants in stolen data markets will seek to organize their activities in order to maximize profits while minimizing risk of loss to outsiders and the uninitiated. This can create a two-tiered market, where lone actors have greater risks while organized groups create insulated resources that only they can use (Herley & Florencio, 2010). There is mixed support for this assertion in the research literature, as some find that the participants are simply seeking to sell or buy data from others with minimal commitments (Dhanjani & Rios, 2008; Franklin et al., 2007; Thomas & Martin, 2006). Others have found hierarchical management structures present to regulate exchanges between participants as well as informal social control mechanisms designed to affect trust between participants (Holt & Lampke, 2010; Motoyama et al., 2011; Peretti, 2009; Wehinger, 2011).

Further systematic inquiry is, however, necessary to identify any variations in the organizational composition and structure of stolen data markets and to produce targeted interventions to disrupt and dismantle these groups (Motoyama et al., 2011; Yip et al., 2013). Limited research has explored the social network structures and relationships between actors, primarily using small forums (Motoyama et al., 2011), or archival data of known stolen data markets that were disbanded due to law enforcement interdiction (Yip et al., 2013). As a result, there is a need for research using active markets of different sizes in order to identify any variations in the nature of social networks in stolen data markets.

In order to assess the economic, social, and organizational processes of stolen data markets as well as their financial impact on businesses and consumers in the global economy, this study utilizes a sample of threads generated from 13 Russian and English language stolen data markets operating in web forums online. This study addresses three key gaps in the literature using various analytic techniques. First, we explore the prevalence and prices of stolen personal financial information sold, as well as the process of payments and market forces that shape market interactions. The financial institutions and nations affected through the sale of data and financial products in the market are explored, along with the relationships between the price paid for different products in the market using quantitative analyses. The findings improve our knowledge of the costs and quantity of information sold, as well as the international impact that these markets have on individual victims and financial service providers. This study also provides some support for the idea that there are multiple tiers within the market which directly affect the price of goods and may shape the behavior of buyers and sellers (Herley & Florencio, 2010).

Second, a qualitative analysis is employed based on Best and Luckenbill's (1994) framework of social organization to examine the associations and working relationships present between participants at the micro and macro-level. This sociological model of organization identifies the actor and their encounters with others as the unit of analysis based on the presence or absence of associations with others, the existence of coordinated or purposive roles, managerial positions, and duration over time (Best & Luckenbill, 1994). In turn, we use this model to explore the relationships between actors and activities to understand the continuum of social organization within and across the forums. The findings suggest that the markets are primarily collegial in nature at the individual level, enabling individuals to work together in order to facilitate transactions. At the same time, several of the forums appear to operate as formal organizations based on managerial structures and long-term operations relative to other forms of online criminality.

Third, social network analysis techniques are applied to explore the links and network structures present between participants within and across the forums in this sample. The density and centrality of networks within each forum will be calculated and analyzed, along with visualizations of the network structures of participants. The effect that these networks have on relationships between participants will be explored, along with their implications for disruption and abatement strategies by law enforcement and security practitioners.

A. The Structure and Basic Economics of the Market for Stolen Data

The practice of acquiring and misusing personal identifiable information and financial data is sometimes referred to as carding (Moore, 2012). This activity dates back to the mid-

1990s when hackers would utilize statistical programs to randomly generate credit card numbers (Moore, 2012). The computer generated information would then be checked through payment processing systems to see if the number corresponded to active account. If so, the creators would then utilize the cards to engage in fraud.

This technique declined in popularity with the rise of e-commerce systems and online banking, which increased the quantity of consumer financial information that could be acquired through phishing and mass data breaches (James, 2005; Wall, 2007). For instance, phishing campaigns may generate a few hundred respondents that provide sensitive data in a matter of minutes (James, 2005). Capturing hundreds, if not thousands, of credit and debit card accounts gives an actor too much information to use relative to the short window of time they may have until the account is closed or a fraudulent transaction is detected. Thus, hackers and attackers began to sell the information they obtained to others in IRC channels and forums in order to make a profit (Dhanjani & Rios, 2008; Holt & Lampke, 2010; Honeynet Research Alliance, 2003; Franklin et al., 2007; Motoyama et al., 2011; Thomas & Martin, 2006; Wehinger, 2011).

Several studies demonstrate that hackers advertise data that they have stolen in a variety of way through advertisements in IRC channels or web forums (Holt & Lampke, 2010; Honeynet Research Alliance, 2003; Franklin et al., 2007; Motoyama et al., 2011; Thomas & Martin, 2006). These markets appear to be hosted and operated primarily out of Russia and Eastern Europe, though a small proportion exists in the US and parts of Western Europe (Dunn, 2011; Symantec Corporation, 2012). Regardless of the operating environment, market actors commonly sell credit card and debit card accounts, PIN numbers, and supporting customer information from around the world in bulk lots (Holt & Lampke, 2010; Franklin et al., 2007; Motoyama et al., 2011). Some also offer "cash out" services to obtain physical money from electronic accounts by hijacking electronic accounts to engage in electronic funds transfers established by a hacker (Holt & Lampke, 2010; Franklin et al., 2007; Motoyama et al., 2011; Thomas & Martin, 2006). Others offer "drops services," whereby individuals purchase electronics and other goods electronically using stolen cards, have them shipped to intermediaries who pawn the items, and then wire the cash to interested parties (Holt & Lampke, 2010). A limited number of sellers also offer spam lists and malicious software tools that can be used to engage in fraud (Franklin et al., 2007; Holt & Lampke, 2010).

Given the range of products sold, the majority of research on stolen data markets discuss the quantities of information sellers advertise as available for sale in the market, and use this data as a mechanism to calculate the prospective economic harm caused by these markets (Dhanjani & Rios, 2008; Franklin et al., 2007; Holz et al., 2009; Symanetc, 2008). In some cases, the number of personal credentials that are publicly posted in threads or posts in IRC chats are included in these counts (Dhanjani & Rios, 2008; Franklin et al., 2007; Holz et al., 2009; Symanetc, 2008). These figures may, however, grossly inflate the actual economic harm caused because that data may be invalid or falsified and cause no actual harm (Herley & Florencio, 2010).

The majority of studies do not also include any estimates for either the advertised price for data or the actual amount paid, regardless of whether the data is derived from a forum or IRC channel (Dhanjani & Rios, 2008; Franklin et al., 2007; Herley & Florencio, 2010; Holz et al., 2009; Motoyama et al., 2011). The lack of pricing data may stem from both the lack of information provided in some sellers' advertisements, and the range of discounts listed by sellers based on the age of data sold, or bulk discounts offered (Holt & Lampke, 2010). Only two studies list prices for data and are generated from forum data (Holt & Lampke, 2010) and IRC data respectively (Symantec, 2008). Despite the different data sources, the most common products advertised were dumps, defined as bank or credit card account details. The costs for this data ranged from $1.30 to $500 in the forums (Holt & Lampke, 2010), while IRC ads had a greater range from $0.10 to $1000 depending on the information included (Symantec, 2008). The next most common product were credit cards with CVVs (Credit Verification Values) ranging from $1 to $14 in forums (Holt & Lampke, 2010), and 0.50 to $12 in IRC data (Symantec, 2008).

Despite the lack of detail on the costs of data sold, there is sufficient evidence that the sales process in both forums and IRC channels begins when a seller posts an advertisement for a product or service including their preferred mode of contact and payment method (Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011). Typically, sellers accept online payments through various mechanisms depending on the market, including PayPal, PaySafeCards (Motoyama et al., 2011), e-Gold, Web Money (Franklin et al., 2007; Holt & Lampke 2010), and other online systems. Real world payments are also accepted by some sellers, though they must commonly be made through MoneyGram or Western Union (Holt & Lampke 2011; Motoyama et al., 2011). Interested buyers contact the seller and negotiate prices and complete transactions outside of the IRC channel or forum, typically through private messaging systems, ICQ, or email (Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011). As a result, it is difficult to determine how many cards or products are sold, in what quantities, and for what price (Herley & Florencio, 2010).

The hidden nature of the transactions is also complicated by the fact that most sellers require payment first, and will then deliver information or services (Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). The market is therefore structured to favor the sellers, as they dictate when and how information will be provided. Unscrupulous vendors can easily cheat or "rip off'" customers by accepting payments, and then either not delivering the purchased goods or sending invalid accounts and false information (Herley & Florencio, 2010; Holt & Lampke, 2010; Motoyama et al., 2012; Wehinger, 2011). There are also no formal dispute resolutions mechanisms that buyers can pursue due to the illegal nature of the transaction (Holt & Lampke, 2010; Wehinger, 2011).

To minimize the risk of loss, some forums utilize informal metrics designed to promote trust between participants and sanction those would otherwise cheat buyers (Holt & Lampke, 2010; Wehinger, 2011). For instance, some forums provide an escrow payment system, whereby a trusted party within the forum will hold payments on behalf of a seller until the buyer confirms they have received the items they ordered (Holt & Lampke, 2010; Wehinger, 2011). The use of escrow payments allows a seller to establish their reputation and demonstrate they can be trusted, though it adds to the complexity of any transaction and may not be viewed as necessary by some sellers (Wehinger, 2011). In addition, administrators in both IRC channels and forums may ban sellers who scam, or "rip off" customers by taking payments without delivering product (Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). Some forums also allow buyers to post feedback on their experience in order to establish the reputation of a seller (Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). The presence of positive comments about a seller and their services appear to affect a sellers' position within the market, as those with more positive reviews receive a higher number of contacts from prospective buyers (Motoyama et al., 2011).

Sellers offer customer service mechanisms designed to attract customers and maintain a client base over time, through the use of bulk discounts, samples, and real time customer support through various instant messaging clients (Franklin et al., 2007; Holt & Lampke, 2010; Wehinger, 2011). There is also some evidence that sellers in IRC channels post personal data, such as account numbers and victim names, which researchers argue is an attempt to demonstrate the validity of their products, or be viewed as a free sample (Dhanjani & Rios, 2008; Franklin et al., 2007; Holz et al., 2009; Symantec, 2008).

The various informal mechanisms identified in previous research suggest that there is a high degree of risk for any participant who wants to engage in a transaction. This calls to question why anyone would voluntarily engage in transactions with data sellers if they have the capacity to acquire or manipulate stolen data on their own (Herley & Florencio, 2010). Furthermore, since sellers offer data at extremely reduced prices suggests that they do not generate a profit equivalent to the actual value of the data (Herley & Florencio, 2010; see also Franklin et al., 2007; Holt & Lampke, 2010). As a result, it is likely that participants in open forums and IRC channels may be either unskilled hackers who do not have the ability to manipulate data or sellers attempting to rip off unsuspecting participants (Herley & Florencio, 2010; Wehinger, 2011).

These conditions have led to the suggestion that open markets are "lemon markets," in that the majority of products sold are unusable (Herley & Florencio, 2010; Wehinger, 2011). Instead, there is most likely a two-tiered market: one where unskilled buyers and sellers interact in the hopes of not getting ripped off, and a second more organized market operating with interconnected and trustworthy actors who insulate their activities from outsiders (Herley & Florencio, 2010; Wehinger, 2011).

The lack of empirical research exploring the nature of the quantities and cost of data sold, and the potential for varied costs based on market conditions makes it difficult to understand the economy of stolen data markets. Thus, there is a need for research that addresses these issues with multiple forums in order to identify whether there is a two-tiered market, and in what ways this shapes the costs for stolen data as a whole.

B. Considering the Social Organization of Stolen Data Markets

The issues inherent in the literature pertaining to the economics of the stolen data market also highlight the need for systematic inquiry on the social processes and organizational composition of participants. Since there may be multiple markets operating, the differences may be identified through the presence of complex organizational hierarchies that facilitate transactions, such as administrators who ban participants and the use of escrow agents (Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). At the same time, the prospective tiered structure of the market may create multiple organizational dynamics operating depending on the nature of the data source (Herley & Florencio, 2010; Wehinger, 2011). This gap in the literature requires substantive research in order to clarify the way in which actors involved in the sale of stolen data engage others and explore any variations in their organizational practices.

Sociological models of social organization provide valuable frameworks to operationalize and measure relationships between deviants, and how such relationships function on or off-line (Adler & Adler, 2006; Best & Luckenbill 1994; Decker et al., 1999; Holt, 2009; Mann & Sutton, 1998; Meyer, 1989). One of the most comprehensive and well applied social organization frameworks was developed by Best and Luckenbill (1994) to identify associations between individuals and groups, and the transactions they engage in. This framework can also be used to understand the way relationships affect individual positions within a clique or network as well as the role, or pattern of action they play in larger social networks and subcultures (Best & Luckenbill, 1994). In turn, social organization frameworks can be used to explore the presence or absence of collegial associations between actors, coordinated or purposive roles between participants, managerial positions, and duration over time.

Within the Best and Luckenbill (1994) framework of social organization, deviance is based on the concept of transactions, whereby behavior is focused toward a particular goal. In deviant transactions, participants are focused toward a specific goal which will bring a degree of gratification. Additionally, transactions have some form of a division of labor that can vary from an individual act to a multi-person scheme with distinct roles for each participant. Finally, transactions have "flexible coordination," such that individuals can adapt their behavior to meet a particular situation or disruption (Best & Luckenbill, 1994: 75). They argue there are three forms of transactions: individual deviance; deviant exchanges; and deviant exploitation. Individual deviance requires a single participant for the act to be completed. Exchanges require two or more actors working in collaborative, but distinctive roles to achieve an end (Best & Luckenbill, 1994). These individuals may be weakly tied and seeking only an immediate exchange, or more strongly tied and seeking long term relationships. Deviant exploitation, however, requires two actors who are working in conflicting roles such that one is an offender and the other is a target or victim of the offender (Best & Luckenbill, 1994: 75).

Best and Luckenbill also argue that deviants are organized in different ways based on the form of transaction they engage in over time. The structure of social relationships also vary based on any division of labor between participants, how frequently and successfully members of the group associate with one another, if they participate in deviance as a collective or alone, and how long their deviant activities extend over time and across virtual or real spaces (Best & Luckenbill, 1994) These characteristics create a continuum of organizational sophistication with five forms of deviant organization: loners, colleagues, peers, teams, and formal organizations (see Table 1).

Table 1. Best and Luckenbill's (1994) Social Organization Framework

Form of Organization	Characteristics			
	Mutual Association	Mutual Participation	Elaborate Division of Labor	Extended Organization
Loners	No	No	No	No
Colleagues	Yes	No	No	No
Peers	Yes	Yes	No	No
Teams	Yes	Yes	Yes	No
Formal Organizations	Yes	Yes	Yes	Yes

From Best & Luckenbill (1994): 12

Loners are the least sophisticated group, as they associate with one another infrequently and do not participate in deviant acts together. Colleagues are the next most sophisticated group, because individuals create a deviant subculture based on their shared knowledge. This provides a way for participants to share information and evaluate others associated with the subculture. Despite this connection, colleagues are not very sophisticated by measures of social organization: they do not offend together, have no division of labor, nor exist over time. Peers have all the characteristics of colleagues, and also offend together but are relatively short lived with no division of labor. Teams are more sophisticated than peers. They last for longer periods of time and have an elaborate division of labor for engaging in deviance. Teams tend to be relatively small in size, seek to garner money or power, and attempt to regularly operate while evading law enforcement. The formal organization is the most sophisticated deviant organization that Best and Luckenbill (1994) include in their framework. Formal organizations have all the elements of teams, as well as extended duration across time and space.

The Best and Luckenbill (1994) model provides a high degree of flexibility in the identification of organizational structures within deviant communities. In fact, the continuum of organizational behavior identified can encapsulate more traditional perspectives of organization, such as hierarchical organizations or less formal network models driven by normative relationships that involve reciprocal exchanges between participants. Best and Luckenbill (1994) also recognize that deviants involved in a specific activity can organize in different ways based on location or points in time. This provides researchers with a mutable framework that can adjust over time to better document the organizational practices of offenders.

Applying the Best and Luckenbill (1994) framework of social organization to stolen data sales suggests that this is a form of deviant exchange based on the flow of data or services from a seller to an interested buyer (Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Motoyama et al., 2012; Wehinger, 2011).   The products available enable individuals to engage in cybercrime based entirely on goods available within the market. Data sold in bulk lots can then be leveraged through “drops” or cashout services to convert data into liquid currency (Franklin et al., 2007; Herley & Florencio, 2010; Wehinger, 2011). Additionally, identity documents, including passports and drivers license documents, sold in the market further facilitate financial crimes on and off-line (Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Motoyama et al., 2012; Wehinger, 2011).

The range of products sold and the direct advertising process of data or services resembles the direct advertising of goods evident in street corner drug markets (Jacobs, 1996, 2000), or hawking markets for stolen goods (Schneider, 2005; Wright & Decker, 1994).   The sales process is communal and driven by interactions between participants to expand their ability to engage in hacking, identity theft, and cybercrime (Holt & Lampke, 2010; Motoyama et al., 2011). The sales process is also participatory since buyers provide feedback for sellers either through public comments or forum-based rating systems (Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). The presence of positive feedback appears to affect the number of contacts a seller has over time (Motoyama et al., 2011), and is driven by timely responses, competitive pricing structures, and customer service (Holt & Lampke, 2010).

There is mixed evidence of more complex organizational structures in forum-based markets (Herley & Florencio, 2010; Wehinger, 2011). Specifically, a small number of forum moderators manage exchanges between participants in order to reduce conflicts and block untrustworthy sellers or buyers (Chu et al, 2010; Holt & Lampke, 2010; Motoyama et al., 2011). Some forums also employ product testers or reviewers to provide impartial feedback on the quality of goods and services sold (Holt & Lampke, 2010; Peretti, 2009). The presence of a top-down management structure with some division of labor suggests that there may be teams or even formal organizations operating the sites or channels. The prospective mix of organizational complexity found in previous research requires further systematic inquire in order to identify any variations in the structure of stolen data markets generally.

C. Social Networks of Data Thieves and Hackers

In addition to qualitative research assessing the nature of relationships between participants in the market, there is a need for quantitative assessments considering the number, shape, and composition of networks in stolen data markets (see Motoyama et al., 2011; Yip et al., 2013). Research on legitimate organizations suggests that in order for the network to survive the removal of a key node or link, it must be balanced in terms of resilience and efficiency

(Bakker, Raab, & Milward, 2012). The more efficient dark networks must have differentiation through specialized roles that reduce redundancy and increase efficiency (Bakker et al., 2012). Differentiation allows organizations to reap the benefits of functional specialization, and also fosters integration through hierarchical structures. A top-down organizational structure enables easier coordination of individual components, and makes them more dense through coordinated relationships between individuals.

The benefits of this structure are, however, offset through the fact that they are less resilient to disruption. The removal of key nodes within hierarchical structures removes their connection to other nodes and can dismantle the entire network. As such Bakker et al (2012) stress two important characteristics of the dark networks: robustness capacity based on their ability to withhold strong shocks, and rebounding capacity after strong shocks. Those dark networks that are more resilient will have the capacity to both withhold strong shocks and rebound after them. This can be achieved through redundant relationships and roles in order to increase resiliency and replace lost connections (Bakker et al., 2012).

To that end, research on stolen data markets suggests that individual threads within the forums are designed to facilitate sales transactions, often with a high degree of connectivity between buyers and interested parties (Motoyama et al., 2011; Yip et al., 2013). Sellers appear to make fewer public posts relative to members who buy goods (Motoyama et al., 2011; Yip et al., 2013). Seller connectivity also appears to increase over time as individuals contact the seller to engage in a transaction and seller's reputation increases. The value of trust also has substantive impact on seller productivity, as approximately 10 percent of sellers across all of the sites in their sample accounted for almost 50 percent of all resources sold (Motoyama et al., 2011).

The larger population of forums consists of users who may buy or engage in trades and post more frequent than others. In fact, user social connectivity increases over time based on participation in either public threads or private messaging exchanges (Motoyama et al., 2011; Yip et al., 2013). When examining the network connections between participants based on posts made within a thread, individuals are often highly connected to one another. This is due to the nature of public exchanges, where one can assume that each post is a direct comment or tie to the previous users' posts (Motoyama et al., 2011; Yip et al., 2013). Participation in threads also appears to increase individual connectivity, such that involvement in multiple threads exponentially increases social ties over time (Motoyama et al., 2011).

The smallest proportion of forum users appear to be moderators and members of the managerial structure that enables the operations of the larger forum (Yip et al., 2013). Testers, moderators, and administrators provide oversight to validate a seller's reputation and in turn, increase their connectivity and engagement with others in the market (Yip et al., 2013). The use of validation services are not, however, consistent across all forums, leading to a small number of vendors with a verified reputation in the site (see also Holt & Lampke, 2010; Yip et al., 2013). As a result, unverified sellers appear to have a greater number of contacts and communicate with more experienced members who understand the process of the market and are willing to take certain risks in order to engage in commerce (Yip et al., 2013).

The limited body of research assessing the network structures evident in stolen data markets leaves multiple questions unanswered regarding the nature of these forums. Specifically, it is unclear if more structured management systems are common across all stolen data markets or if they are distinct to more organized forums. In addition, it is unclear how user centrality may vary based on the number of buyers and sellers present within any given forum.

In order to better understand these issues there is a need for research using active forum populations rather than the small forum sets (Motoyama et al., 2013) or historical data sources commonly used (Yip et al., 2013). Such information can improve our knowledge of the market, and identify more effective tactics to disrupt the activities of market actors (Bakker et al., 2012).

The existing body of research on stolen data markets demonstrates there is some variability in the price of data sold across open and closed markets. It is not clear what factors affect the costs for various items, and the potential impact of so-called lemon markets where unreliable data is sold. Furthermore, market participants appear to operate in collegial networks based on the participatory nature of the sales process in forums (Herley & Florencio, 2010; Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011; Yip et al., 2013). There also appear to be limited instances of more organized structures on the basis of moderators and place managers within forums (Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). There is, however, less research on the network structures that exist between market actors. Limited evidence suggest that sellers are well connected to the larger population of buyers and actors (Motoyama et al., 2011; Yip et al., 2013), while a smaller number of users play a managerial role in the market and engage the overall forum community (Yip et al., 2013).

II. Methods

A. Data Collection

Taken as a whole, stolen data markets serve as a critical resource to facilitate identity theft, hacking, and cybercrimes generally (Franklin et al., 2010; Herley & Florencio, 2010; Holt & Lampke, 2010; Motoyama et al., 2011). These markets enable actors to generate a profit through the sale of stolen data, or through the manipulation of data or by leveraging financial service providers and resources offered in the market. There is, however, generally little research on the influence that social and economic conditions in the market have on the impact of the price for goods (see Franklin et al., 2007; Herley & Florencio, 2010). Additionally, few have considered the social organization of market actors and the network structures present that support the stolen data market generally (Motoyama et al., 2011; Wehinger, 2011).

This study attempts to address these questions using qualitative and quantitative methods in order to expand our understanding of the market for stolen data generally. Using a sample of 1,889 threads from 13 public and private forums engaged in sale of stolen data in both Russian and English, the distribution of products are considered as well as regression models to explore the factors affecting the prices of products. Grounded theory techniques were applied to assess the social organization of these markets, while social network analyses were applied to consider the social connections between participants. The implications of these analyses for researchers, computer security, law enforcement, and policy makers are discussed in depth.

To explore the social organization of stolen data markets, this study utilizes a sample of 1,889 threads from 13 web forums where criminals and hackers buy, sell, and trade stolen financial and personal information. These forums act as online discussion groups where individuals can present issues or discuss problems. They are composed of threads which begin when a registered user creates a post within a forum, asking a question or making a statement (Holt, 2007, 2009; Holt & Lampke, 2010; Mann & Sutton, 1998; Motoyama et al., 2011). Other people respond to the remarks with posts of their own that are connected together to create threads. Thus, threads are composed of posts that centre on a specific topic under a forum's general heading (Holt, 2007, 2009; Holt & Lampke, 2010; Motoyama et al., 2011).

The content of threads provides concrete details on the number and types of data sold, and the financial services and resources used by participants, and the payment methods used. The composition of threads and the social nature of forums can provide invaluable information on the social ties between participants (Herring, 2004; Holt, 2009; Holt & Lampke, 2010; Motoyama et al., 2011). Most online communities operate within a relational J-curve, whereby a small number of users create the largest number of posts (Herring, 2004; Holt, 2009; Robinson, 1984). The same is true with this population of forum users, where a number of individuals made one post about the products being offered or give feedback about a seller's processes. This does not discount the value of a single post, particularly when it is the only post in a thread which signifies the interest and involvement of participants in that forum. Thus, forum data demonstrates the strength of associations between participants and provides information on the practices of market actors (Holt & Lampke, 2010; Mann & Sutton, 1998; Motoyama et al., 2011).

The sample of forums was developed via a snowball sampling procedure similar to those used in traditional qualitative field work in the real world (see Holt 2007; 2010; Holt & Lampke, 2010). Such a tactic is valuable as there is no immediate way to document the total number of stolen data markets operating around the world at any point in time. Thus, this sample began with the identification of three English language forums through Google.com using common terms in stolen data markets, including "carding dump purchase sale cvv" (Holt & Lampke, 2010; Motoyama et al., 2011). One of these sites was a sub-forum of a larger Russian language forum. After exploring the content of threads from these sites, three Russian language forums were identified via web links provided by forum users. Six additional forums were identified using the same processes to create a total of 10 Russian language sites and three English language forums.

Eight of the sites sampled were publicly accessible, in that the entire site could be accessed by anyone in the general public. The five remaining sites required that an individual create a registered user account within the site in order to access the content of the sub-forums related to data sales. Registration-restricted forums are thought to differ from that of publicly accessible forums since they add a layer of insularity and protection from outsiders and the general public (Holt 2010; Markham, 2011). Registration systems allow anyone to join by registering a username and password account with the forum. This is not as exclusive or secure as invitation only forums that completely exclude outsiders from access, though registration eliminates the potential for threads to be captured by search engines or identified by general public (Holt, 2010; Markham, 2011). In order to capture the forum content across all sites, the researcher created usernames for each forum but did not interact with other registered participants to reduce the potential for contamination (Holt, 2010; Markham, 2011). In addition, the forum names and weblinks for the sites included in this analysis are anonymized and pseudonyms are used for forum posters to provide a degree of confidentiality for the participants (Holt, 2010; Markham, 2011).

Examining the hosting locations of this sample of forums suggests they are similar to the larger composition of stolen data markets generally (see Table 2; Symantec, 2012). Four of the sites were hosted in Russia, and all of these sites utilized Russian as their primary language. Another was hosted in Latvia, a former member of the USSR, though its participants communicated in English and had a large proportion of ripping complaints. Europe was also a prominent host for these markets, as two were hosted in Germany, one in Luxembourg, and another in the Netherlands. One of the sites was hosted in the UK, though it utilizes the .ru country extension suggesting it is Russian. Another site was hosted in the British Virgin Islands, though the participants communicated entirely in Russian. Finally, two of the sites were hosted in the United States, though one forum's participants communicated entirely in English while the other was in Russian. It is also noteworthy that only three of the sites had their domain registered using a publicly identifiable persona including a name and email address. The rest used a private registration service to anonymize the identity of the hosting service. Thus, the mixed composition of the site hosting information relative to the participants' communication methods are reflective of the larger dynamics of the market for stolen data (Symantec, 2012).

Table 2: Hosting Detail For Each Forum

Forum Number	Hosting Country	Domain Registrant	Language	Registration Required
1	Germany	Private	RU	No
2	United States	Not Current	ENG	No
3	United States	Private	RU	No
4	British Virgin Islands	Private	RU	No
5	UK	Private	RU	No
6	Russia	Private	RU	Yes
7	Russia	Private	RU/ENG sub	Yes
8	Latvia	Private	ENG	No
9	Russia	Public	RU/ENG sub	No
10	Germany	Public	RU	Yes
11	Russia	Private	RU	No
12	Netherlands	Private	RU	Yes
13	Luxembourg	Private	RU	Yes

Using these forums, the research team captured all threads posted from carding or sales related sub-forums in order to develop a substantive volume of posts to better explore the organization of forum participants. The threads and content for each of these forums were saved as web-pages, then copied, cut, and pasted into MS Word documents. A certified Russian translator with substantive experience with technological jargon and forum communications translated the Russian language content from all forums.

Due to the availability of the translator, convenience samples of 25 threads from each Russian forum were selected to capture the most recently posted items for sale in each site (see Table 3). Additional samples of threads were translated from eight forums, particularly those that had active posting, in order to better assess the practices of actors and the network connectivity of participants. Due to a complication with the language encoding of threads from Forum 3, the translator was unable to completely translate all threads sampled. Thus, this forum content is excluded from both the social organization and social network analyses. Repeat threads were excluded from analysis, but translated to ensure reliability of content. The research team also oversampled threads from the English language forums in order to capture any variations in the nature of these markets and their organizational composition. This strategy provides a mix of user populations and duration over time, while at the same time creating a relatively matched sample of posts between English and Russian language threads across the forums.

Table 3: Forum Descriptive Statistics

Forum	# of Threads	First Post	Last Post
1	55	12/31/10	7/21/11
2	128	12/1/10	2/23/11
3	6	10/17/10	7/16/11
4	144	10/3/09	12/25/2011
5	89	6/6/08	12/11/11
6	48	2/5/09	11/14/11
7	202	12/26/10	7/9/11
8	590	4/1/09	11/1/11
9	312	4/1/11	7/21/11
10	35	4/10/10	3/7/11
11	60	5/9/07	2/25/12
12	71	11/7/07	11/9/11
13	153	6/6/07	7/25/11

The range of time included in this sample of threads provides a wealth of information concerning the social organization of stolen data markets (see Table 3 for detail). To that end, six of the forums had posts over more than a two year window, giving substantial insights into the products sold and organization of stolen data markets. Several forums contained only a few months of posts, demonstrating variations in the structure and duration of the market. As such, this sample provides a representation of the activities of buyers and sellers in stolen data markets using various languages over time.

We utilize both qualitative and quantitative methods to address the research questions of this study, with quotes from the data where appropriate. Specifically, quantitative analysis techniques are used to examine the economy of the market, while qualitative grounded theory analyses are used to explore the social organization of the market. Finally, quantitative social network analyses will be used to assess the relationships between participants in these markets. Each form of analysis and the measurement techniques employed are discussed in detail below.

B. Economic Data Coding

Content analysis techniques were applied to classify the various products, resources, and materials sold or sought out in these forums. The content of each ad was coded based on the detail provided. For example, a post was coded as a sale if an individual stated that they were "selling," "offering," or otherwise providing a service (see also Chu et al., 2010; Holt & Lampke, 2010; Motoyama et al., 2011). Requests for products were coded based on the language used, such as "need a," "buying," or "seeking." Each item was coded individually, such that an advertisement where an individual was selling credit card numbers as well as PayPal accounts were coded as a single instance of each activity (see also Holt & Lampke, 2010). In addition, any additional advertisements or updates in a thread were coded as new cases to capture variations in pricing and products over time. This created a total number of advertisements (N=12,844) that is larger than the overall number of threads where they appeared. Also, only 4.3% (n=554) of all advertisements involved requests to buy or exchange materials, thereby supporting the notion that these markets are oriented toward the sale of data (see also Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011). The usernames, email addresses, forms of payment accepted, and the sellers' terms of service were also captured to better document the practices of forum actors.

The products and services listed in each thread were coded into categories based on common aspects of the item or service offered, using common terms evident in the research literature on stolen data markets when possible (see Table 4 for detail; Franklin et al., 2007; Holt & Lampke, 2010; Holz et al., 2009; Honeynet Research Alliance, 2003; Motoyama et al., 2011). Any ad that provided access to dumps, or bank account/credit card numbers, were classified as *dumps* (see also Franklin et al., 2007; Holt & Lampke, 2010). An advertisement where the seller offered CVV data, which is a credit card number along with the Credit Verification Value included on the signature line of the card were coded as *CVVs* (Franklin et al., 2007; Holt & Lampke, 2010). Similarly, the code *Fullz* was used for any ad where a seller offered an account with all pertinent information included, such as the mothers' maiden name and contact information of the victim (Franklin et al., 2007; Holt & Lampke, 2010; Thomas & Martin, 2006). Any ad that offered access to either the username and password to electronically access a bank or credit card account or access to an account in general were coded as *Bank Accounts*. The sale of eBay or PayPal accounts and other service providers such as poker accounts or Amazon.com were coded as *eBay/PayPal*. Services that offered to crack email accounts or provided access to different personal accounts like Face book were coded as *Personal Info*. The final category *Other Financial Products* captures other resources that may be used by individuals such as gaming accounts that may be compromised.

Table 4: Distribution of Products Based on Buying and Selling

Product	Total	% of All Ads	Buying Posts	% of Total	Selling Posts	% of Total
Bank Accounts	205	1.6%	21	10.2%	184	89.8%
Cashout Services	235	1.8%	74	31.5%	161	68.5%
CVV	4481	34.9%	21	0.5%	4460	99.5%
Dedicated Servers	157	1.2%	0	0	157	100.0%
Drops for Laundering	165	1.3%	59	35.8%	106	64.2%
Dumps	5737	44.7%	68	1.2%	5669	98.8%
eBay/PayPal	183	1.4%	17	9.3%	166	90.7%
Equipment	198	1.5%	12	6.1%	186	93.9%
Fullz	122	0.9%	3	2.5%	119	97.5%
Identity Documents	89	0.7%	37	41.6%	52	58.4%
Malware	183	1.4%	31	16.9%	152	83.1%
Money Transfers	303	2.4%	20	6.6%	283	93.4%
Other Financial Products	10	0.1%	0	0.0%	10	100.0%
Other Products	277	2.2%	113	40.8%	164	59.2%
Personal Info and Accounts	`99	0.8%	39	39.4%	60	60.6%
Plastics	153	1.2%	13	8.5%	140	91.5%
Skimmers	125	1.0%	10	8.0%	115	92.0%
Spam and Scams	122	0.9%	15	12.3%	107	87.7%

In addition to the stolen financial data, there were a range of resources offered that would allow individuals to use this information to gain access to funds within the account. Specifically, ads were coded as *Cashout Services* when they offered resources to withdraw funds from compromised accounts through the electronic purchase of goods, making electronic payments to fictitious websites or transfers, or by withdrawing money from ATMS in the real world. The code *Plastics* was used whenever an advertiser mentioned credit-card shaped pieces of plastic with magnetic strips that can be read or written on in order to store data. Plastics can be encoded with data from a dump and then used in the real world to facilitate the cashout process (Franklin et al., 2007; Holt & Lampke, 2010; Thomas & Martin, 2006). For instance, plastic cards can be used at ATMS to withdraw funds or to make purchases in brick and mortar stores. This code also includes sellers who could make counterfeit cards through the application of holograms, logos, or embossing devices to produce raised numbers or names (Franklin et al., 2007; Holt & Lampke, 2010; Thomas & Martin, 2006).

The code *Drop* was used whenever individuals described services to either send or accept goods purchased using dumps or compromised accounts. Additionally, drops can refer to individuals who will cash checks or payments made from a dropped account and then wire the funds to another account (Franklin et al., 2007). Thus, drops serve a vital role in the process of acquiring funds from victim accounts. Related to drops are *Money Transfers*, referencing any service that would directly transfer funds between different forms of currency, such as paper money into electronic currency systems, or across electronic payment systems providers. Finally, individuals offering any and all materials that could be used to masquerade as another individual, such as passports or drivers licenses were coded as *Personal Identity Documents*.

A variety of resources were also sold that could be used before, during, or after data has been acquired to engage in various forms of cybercrime. Specifically, ads for malicious software such as Trojans, botnets, DDoS services, or other attack tools were coded as *Malware* (see also

Chu et al., 2010). The code *Spam/Scam Services* was used to capture any instance of individuals advertising email databases that could be used for the distribution of spam or fraudulent email scams. This code also includes schemes devised and marketed by hackers to engage in fraud or gather sensitive information.

The code *Dedicated Infrastructure* reflects any services for web hosting for websites and malicious content, VPN connectivity, or proxy connections through hacked or legitimate computers. The code *Skimmers* was applied to any instance where individuals offered so-called skimming devices, or magnetic strip reading devices that can be attached to ATMs or other card reading systems to capture data in the physical world (Holt & Lampke, 2010; Thomas & Martin, 2006). A number of ads also offered *Equipment,* or hard goods that may not be directly involved in the act of carding, such as computers, cell phones, televisions, printers. Finally, the code *Other Products/Services* was applied to any other resource that did not evenly fit into another category, such as hacking services for different websites, ICQ number sales, links to pornography, drugs, or accounts in file sharing sites.

Descriptive statistics regarding the advertised prices for each code will be presented, as well as any additional information concerning the country of origin for data and the financial institution harmed. Information will also be presented concerning the preferred payment systems used by buyers and sellers, as well as contact methods to communicate outside of the forums. The advertised price for data will also be detailed in U.S. dollars as this was the most common currency used. Any price listed in rubles was converted using the estimated value of the currency in US dollars on that day. Finally, logistic regression models will be presented to assess the impact of social and market forces on the advertised price for data.

C. Social Organization Coding

All forum threads were then printed and analyzed by hand using grounded theory techniques (Bryant & Charmaz, 2010; Charmaz, 2006; Corbin & Strauss, 1990; 2007) to derive concepts and information from the data, along with guiding questions from Best and Luckenbill (1994). This includes the ways that "deviant actors organize themselves to pursue their deviant activities" and how "these basic forms differ in organizational features, such as division of labor, coordination among the deviant actors, and objectives" (Best & Luckenbill, 1994: 12). In order to capture variations in organizational patterns over time, the following questions are addressed: "what conditions shape the development and transformation of organizational forms," and "how do organizational forms change over time, and what conditions account for these changes?" (Best & Luckenbill, 1994: 12). Since the data include various time points, it may not be possible to identify true behavioral and organizational changes (see also Holt, 2009). The general range of time covered across all the forums does, however, provide a basis of comparison that is generally in keeping with the tenets of grounded theory and enable a basis to discern prospective changes across sites or simply isolated patterns within a single forum.

These guiding concepts will be applied to the data along with questions used in social organization analyses of gang activity (Decker et al., 1998) and computer hackers (Holt, 2009; Meyer, 1989). These items were included due to the fact that computer hacking techniques are commonly employed in the course of stealing sensitive data (Holt & Kilger, 2012; Holt & Lampke, 2010; Motoyama et al., 2011). Also, these studies provide useful operationalizations of concepts in social organization research that can be applied to various forms of offending.

Specifically, the first series of questions used centers around the complexity of division of labor, considering if deviants offend together and any evidence of their division of labor.

This includes the presence of groups, their memberships, any relationships between group members, and any stratification and role specialization present (Decker et al., 1998). Second, the coordination of roles examines relationships between individuals, based on stated codes or rules on the regulation of relationships, and the way these rules are defined and enforced (Decker et al., 1998). Finally, purposiveness assesses relationships between groups and how they specify, strive toward, and achieve goals (Decker et al., 1998). This concept is addressed based on evidence of operations and crimes performed between multiple groups, and any leisure activities involving these groups.

These questions were used after the initial phases of data analysis to refine the concepts identified through the application of grounded theory techniques (Bryant & Charmaz, 2010; Charmaz, 2006; Corbin & Strauss, 1990; 2007). This method is particularly valuable for qualitative research as its procedures permit the researcher to develop a thorough, well-integrated examination of any social phenomena (Corbin & Strauss, 2007). Any concepts found within the data must be identified multiple times through comparisons to identify similarities. Specifically, grounded theory analyses begin with open coding where all data are placed into specific events or incidents, then labeled and grouped into categories and sub-categories using a specific identifying tag (Bryant & Charmaz, 2010; Charmaz, 2006; Corbin & Strauss, 1990; 2007). The second phase of axial coding involves testing the relationships between categories, subcategories, and the data itself to further develop the identified concepts. The final phase of analysis involves selective coding to determine how any categories or subcategories from previous stages could be linked to a "core category" of the phenomenon under study (Corbin & Strauss, 1990; 2007).

The inductive analyses generated from the grounded theory process were then compared against the Best and Luckenbill (1994) framework to explore the social organization of actors engaged in the sale of stolen data. Each of the four components of this model, mutual association, mutual participation, division of labor, and extended duration, will be discussed using respondents' comments or observations to help illustrate points where appropriate.

D. Social Network Analysis

In addition to the qualitative social organization analysis and economic regression models, a quantitative social network analysis was conducted. This type of analysis allows us to visualize and quantify the information on large networks and complex relationships. Network analysis allows both the visualization of users' communications, as well as the extraction of network connectivity. The exchanges between individuals in these forums allow for the identification of network structures within and across all the forums. This type of analysis allows for the identification of global patterns in the otherwise hidden networks of data market participants, and connectivity between participants. In addition, social network analysis enables researchers to consider connections between participants based on their role in both the forum and in the course of any sales or exchanges noted in a thread. Similar techniques have been employed with java forums (Zhang, Ackerman, & Adamic, 2007), hacker forums (Decary-Hetu & Dupont, 2012), social network profiles of malware writers and hackers (Holt, Strumsky, Smirnova, & Kilger, 2012), terrorist networks (Bakker et al., 2012), and simple analyses of stolen data forums (Motoyama et al., 2012). Thus, these analyses can be applied to various forms of both deviant and non-deviant behavior alike.

Applying social network analysis techniques to forum data, individual posters become network vertices (V), while their forum interactions (C) establish connections between them

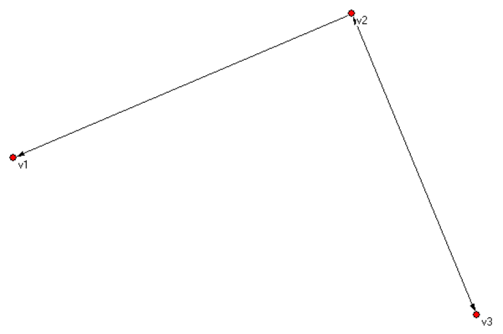
(Zhang, Ackerman , & Adamic, 2007). The username for each poster, regardless of what forum it appears in, serves as a basis to assess connections between individuals, and consider the flow of information from one agent, or vertex, to another. This allows us to build a set of arcs, or connections between hackers (C). These relationships (C) and participants (V) identify a given network N(V,C) where $C \subseteq V \times V$.

Specifically, a connection is represented by an arc or an arrow between two or more vertices, represented as dots, to visualize a thread where several forum users communicate. These can be simple interactions or exchanges (Best & Luckenbill, 1994): question and answer, advertisement and response, or information offered-offer received. A single dot represents a user who started a thread that no one else engages in or that did not generate any visible activity within that thread (see Figure 1). In the event that an individual replies to his/her own post in a thread, it becomes a loop or self-reference. In threads where individuals respond to an initial post, the direction of arcs (connections between users) goes from the topic starter to the rest of the thread (see Figure 2). The direction of advertised information moves from the ad to the people asking subsequent information or interested in the ad.[1]

[1] A different interpretation of the flow of information may be that those who replied to the thread direct their posts at the topic starter which would reverse the direction of the arrows. Even a more nuanced situation would be to disentangle who posts what information to whom in any thread. For example, in a particularly long thread the topic starter might provide an advertisement for their goods, while the second response from a new user would ask a question, and the third post from another user would be directed at both the thread starter and the second user. In order to provide an exploratory analysis, the direction of connections is kept constant in order to provide some comparative consistency across all forums.

Figure 1: Network Visualization of Threads With No Responses

Figure 2: Network Visualization of a Thread With Two Replies



Due to the nature of these forum, any user can see the content of a thread so long as they are either a registered user or able to access the forum. As a result, it may be difficult to truly disentangle the complete networks evident in any forum. Instead, the networks presented here are directed as the direction of information flows from the topic starter to the rest of the thread. In addition, the focus of this analysis is on the connections to the topic starter as the originating point of discussion in a thread. If a user has replied to multiple topics, then s/he will be connected to all those threads.

This framework will be applied to identify the basic network typology and structure of participants in the various forum data sets collected. Besides visualization, social network analysis allows to extract measure at two level of analysis: global and local (de Nooy, Mrvar, & Batagelj, 2005). The global level measures are extracted from all interactions on the network, while the local measures study the surroundings of particular nodes. Global network measures are necessary to understand how the whole of a network varies from its individual parts, and identify additional properties that emerge from interactions with large number of people.

Three major measures were calculated at the global level for each network in order to explore their resiliency, redundancy, and structure generally. First, network density was determined based on the ratio of existing ties to the number of all possible ties in the network. The higher the network density number, the higher the propensity of that network to transmit information between users (de Nooy et al., 2005).

Second, the average degree of connectivity was measured based on the average number of users that can be connected through the network. This measure identifies how far information can traverse a network through existing connections (de Nooy et al., 2005). Most networks that have higher than 1 average degree (e.g. where at least two users are connected) usually form a "giant component" where all participants are connected to one another (de Nooy et al., 2005). A component is a sub-set of network where each node is connected to one another. On one extreme, each thread may become a separate component (e.g. users post and reply only to separate threads) though at the other extreme everyone could participate in all threads on the forum. Most of the networks in these forums lay between these extremes.

Finally, the all degree centrality is calculated based on the general connectedness of users on the network based on associations with others in threads (de Nooy et al., 2005). At the global level, this measure indicates the overall connectivity on the network. At the local level, this information allows for the identification of central users (de Nooy et al., 2005) who may be a point of leverage to disrupt the larger network if they were removed (Bakker et al., 2012). The measures were calculated for both directed networks where information flows from the topic starter to participants in the thread, and undirected networks that assume everyone communicates with everyone.

Individual user-level measures were also calculated to compare against the global-level network measures. An all-degree centrality measure was calculated at the individual user level expressed as the proportion of vertices incident on a node from the total nodes on a network. Specifically, a user with the highest degree centrality will communicate with the largest number of other users on a forum. The local level measures can be easily depicted on the networks, and visualized on the basis of categorical (aka partitions) and continuous (aka vectors) variables. For example, the size of each node can be scaled based on its centrality to understand the flow of information from central users to other participants (de Nooy et al., 2005: 113). In each forum, the total number of posts made by each user over the life of the forum is provided in some section, whether under the date of their current post or below their username. This detail is vital to understand their overall involvement in the forum beyond what is evident in the threads sampled here. The number of posts is treated as a vector and represented as different sized vertexes, where larger nodes represent a larger number of posts.

In our specific case, the most central users may not be hubs of information but rather frequent posters that reply to a lot of threads to give feedback, or potentially attempt to disrupt threads by posting their own advertisements. In such situations, the centrality of a user (a measure derived from the network itself) will be highly correlated with the number of posts they make. This is a separate situation from instances where users connect to nodes (such as sellers) based on their reputation or knowledge. In fact, sellers with a solid reputation might not need to make frequent posts in a forum to facilitate economic activity. Instead, they may be more active in ICQ and other messaging systems to interact with customers (Motoyama et al., 2011; Yip et al., 2013). Thus, Pearson correlation coefficients are also calculated since both the number of posts and user centrality are continuous variables.

In addition, the networks are partitioned on the basis of the users' expressed interest in buying data, selling a product or exchanging information. These visualizations enable the comparison of network structures by economic activity: buying, selling, doing both buying and selling, exchanges, and neutral (no specific information as related to economic activity). The neutral color indicates that the node has made posts not specifically related to buying, selling, or exchange, as defined in economic analysis section of this report. This is the broadest category as it captures both positive and negative comments, feedback by users, thread hijacking, and administrative interferences.

Due to the nature of the data, economic activity is coded using an economic incidence as a unit of analysis and network data is coded with the users as unit of analysis and the ties between them define a network. The number of posts captures overall user activity on the forum and becomes a proxy for the overall activity of a user on a forum. User centrality is actually the measure derived from the networks themselves, and may demonstrate central positioning that is reflective of dark network structures overall. The Pajek software suite was used to create all network visualizations (de Nooy et al., 2005).

Though there are other forms of network analysis and measures that can be calculated, this analysis is restricted to basic measures in order to generally examine the structure of these markets. Additionally, the basic composition of the forums limits the ability to determine certain characteristics about their networks. Specifically, since forums allow anyone who is a participant to observe exchanges in threads, the indirect flow of information may be greater than what is captured by the number of participants in a thread (Zhang et al., 2007). The fact that most purchases and exchanges take place outside of the forums also limits our ability to fully represent the network structures of participants. The preliminary nature of these findings are, however, invaluable to identify any commonalities between these markets and other network structures in both legitimate (Zhang et al., 2007) and criminal groups alike (Bakker et al., 2012; Decary-Hetu & Dupont, 2012; Holt et al., 2012; Motoyama et al., 2012).

III. RESULTS

A. The Sales and Advertising Process

In examining the content of each forum, it was clear that each thread was created by an individual to serve as an advertisement for their products or services (see also Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011). The Thread Starter, or TS, would list their good or service, product detail, pricing, any rules regarding the sale, and their contact and payment information. This is best demonstrated in a post from Forum 4, where an individual sold dumps, or bank and credit card account information from around the world. The seller indicates the different types of dumps available, including classic Visa/MasterCards which have a lower general balance than gold and platinum accounts, and business/corporate cards. This may account for the price differences evident, and their availability by location:

> ***Dumps Fresh Base ... EU-USA-CANADA-ASIA-OTHER.. Best Valid..***
> Virgin dumps: Europe, Asia, Canada, Usa and other country! We offer very good quality. Thanks All.
> PRICE LIST:
> ************USA***************
> 1pcs CLASSIC/STANDARD= 20$
> 1pcs GOLD/PLATINUM = 25$
> 1pcs BUSINESS/SIGNATURE/PURCHASE/CORPORATE/WORLD = 30$
> 1pcs AMEX = 20$
> ************CANADA************
> 1pcs CLASSIC/STANDARD = 50$
> 1pcs GOLD/PLATINUM/BUSINESS/SIGNATURE/PURCHASE/CORPORATE/WORLD = 70-200$
> *******EUROPE & ASIA & LATIN & OTHERS*********
> ---[code 101 - non chip]--- [This indicates the dump contains only magnetic strip data, not the chip/pin data common in cards from other nations]
> 1pcs CLASSIC/STANDART = 110$
> 1pcs GOLD/PLATINUM = 130$
> 1pcs BUSINESS/SIGNATURE/PURCHASE/CORPORATE/WORLD = 150$
> 1pcs INFINITE = 200$
> **********************
> ----[code 201 - chip]---- [This indicates the dump contains the chip/pin data common in cards from other nations]
> 1pcs CLASSIC/STANDART = 50$
> 1pcs GOLD/PLATINUM = 65$
> 1pcs BUSINESS/SIGNATURE/PURCHASE/CORPORATE/WORLD = 120$
> 1pcs INFINITE = 150$
> RULES:
> (please read the rules carefully and follow all the steps, anyone breaking this rules shall expect to be fully ignored by service)
> 1. Contact with one of the our supports and choose dumps u want.
> 2. Calculate total price and submit your order.
> 3. Send us money and your e-mail.
> 4. We have 24 hours (maximum) to complete your order.(LR [Liberty Reserve Payment] INSTANT DELIEVERY)
> 5. We replace only Pickup/Hold Call Dumps with in 24 hours after time period we are not responsible
> PAYMENT INFO:
> LIBERTY RESERVE
> Support Icq: [removed]

This post demonstrates the range of prices for products offered, as well as the fact that sellers can set their own terms concerning the rules of a transaction. Examining the breakdown of products sold suggests there is substantial variation in the prices advertised (see Table 5 for detail). Dumps were the most common item sold in keeping with the larger body of research on data markets generally (see Herley & Florencio, 2010 for review). The average advertised price for dumps was much higher ($102.60) than that of the second most prevalent item, CVVs ($26.21) (see also Franklin et al., 2007; Holt & Lampke, 2010). In fact, CVV data was advertised at a slightly lower price than eBay and PayPal accounts ($27.25). In general, the average costs for data were lower than that of data manipulation services such as identity documents ($138.46), drops ($192.37), cashout services ($1,076.93), and money transfers ($1,424.59). Skimmers, used to capture data in the field, had the highest average cost at $2,382.60 (see also Holt & Lampke, 2010). Products related to data capture, such as spam ($96.33), dedicated servers ($100.97), and malware ($183.27) were also more expensive than financial data but in keeping with existing research (Dhanjani & Rios, 2008; Chu et al., 2010; Holz et al., 2009).

The majority of sellers also excluded pricing details from their advertisements, with the exception of data sellers (see also Chu et al., 2010; Holt & Lampke, 2010). This may be due to the time sensitive nature of stolen data, and the desire to easily inform interested customers of the costs for information by victim nation. Other products, such as bank accounts, malware, or services designed to manipulate data are not as time sensitive and may have more negotiable prices based on the needs of the seller or buyer. A number of service providers offering cashout services, drops, and money transfers did not list their prices. Instead, they described the costs for services as a percentage of the total amount of money they may asked to convert or move between accounts (see Holt & Lampke, 2010). This may be more sensible than offering specific costs because they can prorate their costs on the basis of the total money transferred.

Table 5: Pricing Information for Products Sold

Product	Min Price	Max Price	Average Price	Count With Price	%	Count With No Price	%	Percent Rather Than Price	%
Bank Accounts	5.00	700.00	187.44	63	30.7	142	69.3	2	1.0
Cashout Services	0.30	6000.00	1076.93	14	6.0	221	94.0	63	26.8
CVV	1.00	8000.00	26.21	4316	96.3	165	3.7	0	0
Dedicated Servers	0.20	700.00	100.97	42	26.7	115	73.3	0	0
Drops for Laundering	0.50	1000.00	192.37	27	16.4	138	83.6	62	37.6
Dumps	0.04	8000.00	102.60	5167	90.1	570	9.9	4	0.1
eBay/PayPal	0.20	800.00	27.25	118	64.4	65	35.6	3	1.6
Equipment	3.00	5000.00	549.51	61	30.8	137	69.2	0	0
Fullz	15.00	150.00	72.81	87	71.3	35	28.7	0	0
Identity Docs	0.50	500.00	138.46	32	40.0	57	60.0	0	0
Malware	2.00	1570.00	83.27	99	54.1	84	45.9	0	0
Money Transfers	10.00	38000.00	1424.59	37	12.2	266	87.8	85	28.1
Other Financial Products	6.00	15.00	10.75	4	40.0	6	60.0	0	0
Other Products	0.11	5000.00	177.26	82	33.2	195	66.8	6	2.2
Personal Info and Accounts	1.00	5025.00	197.19	44	44.4	55	55.6	3	3.0
Plastics	0.50	3000.00	261.47	47	30.7	106	69.3	0	0
Skimmers	200.00	9000.00	2382.60	23	18.4	102	81.6	0	0
Spam and Scams	8.00	600.00	96.33	24	16.4	98	83.6	4	3.3

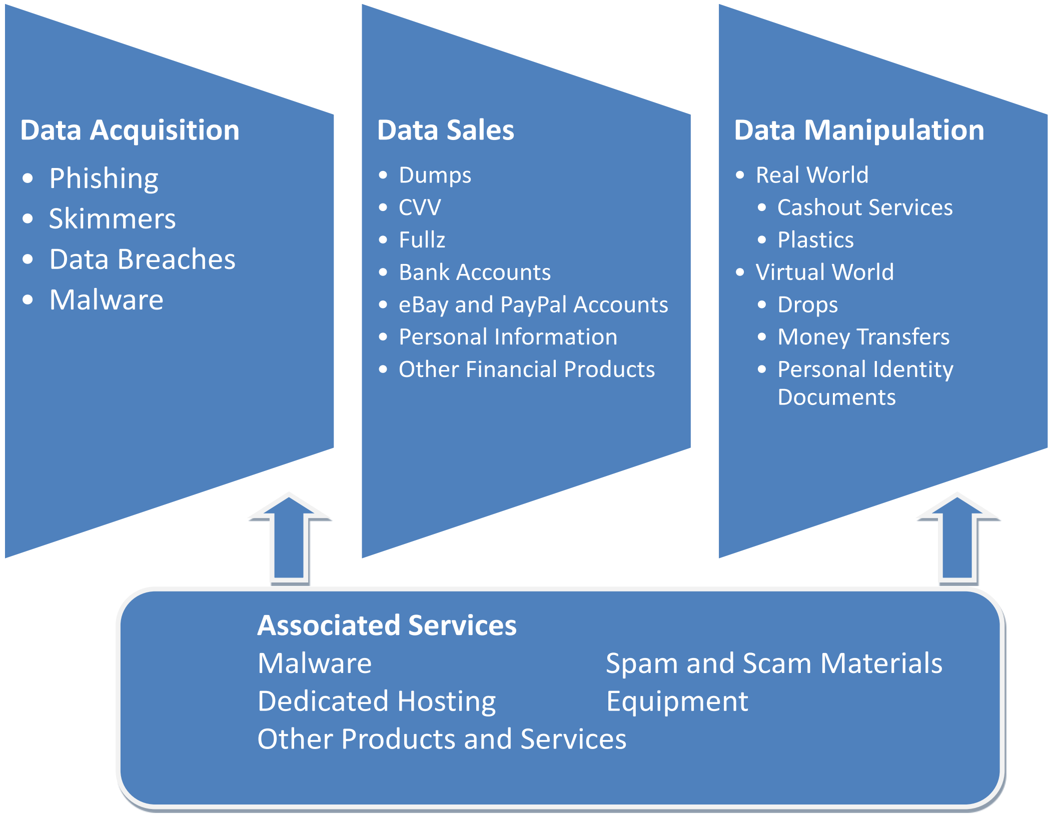
This issue was exemplified in a post from Forum 1 offering encashment services who indicated he would accept a percentage of the account value rather than request a set fee for service:

> We are the experience team [name removed] working in the area of banking innovations, and here on the site we are ready to offer you the following services:
>
> We in cash funds in the RF [Russian Federation] which we have received as electronic bank transfers.
>
> Help with encashment: direct scheme- no intermediaries
>
> -encashment of funds
>
> -Transmit of electronic funds
>
> -Diversion of funds
>
> -Work with accounts that have been seized by the authorities
>
> Work with dirty funds…
>
> Our advantages:
>
> -Low commissions
>
> -Speed (as a rule, funds are received on the day that they are credited to our company's account (usually on the day following the payment day in the RF), on the next day or maximum one day after crediting to our account depending on the amount.

The distribution of products sold across the markets demonstrates that every phase in the process of data theft is represented (see Figure 3). The majority of ads involved some type of stolen data (84.3%), particularly dumps and CVVs. Data manipulation products shared a smaller proportion of the market (7.4%), though this seems vital in order to obtain funds from the accounts or financial service products sold (Franklin et al., 2007; Herley & Florencio, 2010; Holt

& Lampke, 2010; Wehinger, 2011). A number of products were also available that could be used to support the acquisition or manipulation of data, including spam distribution, hosting services, and skimmers can be used for data acquisition on or off-line. As a whole, these markets enable individuals to engage in all forms of cybercrime and data theft from start to finish at relatively low prices (Herley & Florencio, 2010; Wehinger, 2011).

Figure 3: The Process and Products Available in Stolen Data Markets



Approximately half of all advertisements indicated the financial service providers and card issuing agencies where their data originated (see Table 6). The most prevalent providers harmed across all CVV, dumps, and fullz sellers were issued by MasterCard and Visa. These two institutions have the largest percentage of the financial payment market in the US (The Nilson Report, 2013), and an increasingly large market share internationally (Team, 2013).

American Express cards also comprised a large proportion of stolen data, followed by Discover in keeping with their much smaller market representation (The Nilson Report, 2013). A small number of related service providers were also impacted, including Alpha, Diners, JCB, and Maestro which may be a reflection of their overall market share.

Table 6: Card Service Providers Affected

Card Type	CVV	%	Dumps	%	Fullz	%
Alpha	0	0.0	2	0.0	0	0.0
Amex	424	9.4	347	6.0	1	0.8
Diners	0	0.0	7	0.1	0	0.0
Discover	384	8.5	250	4.3	1	0.8
JCB	1	0.0	6	0.1	1	0.8
Maestro	29	0.5	7	0.1	0	0.0
MasterCard	495	11.0	552	9.6	11	9.0
Visa	622	13.9	883	15.4	56	45.9
Missing	2526	56.4	3683	64.2	52	42.6
Total	4481	100.0	5737	100.0	122	100.0

Percent totals do not equal 100%

Many sellers also specified the country of origin for the data they sold (see Table 7 for detail). The majority of CVVs, dumps, and fullz come from Europe and the United States respectively. The UK and Canada are also commonly victimized by data thieves, though Australia, Russia, and Asian nations are a much smaller proportion of the market generally. The distribution of regions and nations is similar to that of other studies of stolen data markets (Franklin et al., 2007; Holt & Lampke 2010; Motoyama et al., 2011). There is, however, no immediate explanation as to why individuals in Europe and the United States may be more affected by data thieves. It may be a reflection of the large number of credit and debit card holders in these nations, as well as the large number of financial transactions that occur within the US via e-commerce and other systems on a daily basis (Verison, 2012). The large number of data breaches affecting consumers in these nations may also increase the risk of data exposure (Holt & Lampke, 2010; Symantec, 2012; Verison, 2012).

Table 7: Location of Stolen Data By Product

Location of Data	Bank Accounts	%	CVV	%	Dumps	%	Fullz	%
Asia	11	9.2	174	4.0	473	8.7	10	9.2
Australia/ New Zealand	4	3.3	284	6.3	152	2.8	8	7.3
Canada	23	19.3	411	9.2	675	12.5	13	11.9
Europe	48	40.3	1278	29.0	1598	29.6	45	41.3
Other	7	5.9	154	3.4	664	12.3	6	5.5
Russia	3	2.5	15	0.3	9	0.2	0	0.0
United Kingdom	3	2.5	1072	24.1	354	6.5	9	8.3
United States	20	17.0	1003	22.4	1481	27.4	18	16.5
Total	119	100.0	4481	100.0*	5406	100.0	109	100.0

Missing data excluded here; Percentage does not equal 100

To understand how product pricing relates to victim nations, the mean price for CVVs, dumps, and fullz were compared by country (see Table 8 for detail). Binary measures were computed for each geographic category. The mean price indicates that data from the US and UK are relatively inexpensive compared to other nations. A t-test with unequal variances was conducted to test whether there are significant differences in the average price by country.[2] The results suggest that the US is the least expensive country for CVVs (0=$2.63; 1=$1.67; T=29.861; sig.000), and follows closely behind the UK in terms of the price of dumps and fullz. This may be a consequence of the availability of data from the US and UK, as they have saturated the market and reduced their overall advertised price (see also Herley & Florencio, 2010; Holt & Lampke, 2010).

Bank account data from the United Kingdom (0=$4.82; 1=$4.08; T=.855) and Europe (0=$4.95; 1=$4.12; T=2.220; sig.05) are the least expensive, while US accounts are the most expensive (0=$4.68; 1=$5.33; T=-1.37; sig.672). There is no immediate explanation as to why this price differential exists, though it may be that individuals residing in Europe and Russia may be more easily able to access these accounts electronically. Dumps and fullz from Europe, Asia, and other nations such as the Middle East are the most expensive overall.

The price for bank accounts by country were not significantly different, with the exception of Europe, where they were less expensive relative to other nations (0=$4.95; 1=$4.12, T=2.220; sig. 05). The mean price for data from both Canada and the UK were not significantly different except in the case of dumps, where they were significantly lower than other nations. In fact, the mean price for dumps from the UK were $2.81 (T=13.000; sig.000), while Canadian dumps were $3.37 (T=5.907; sig.000). There is no immediate explanation for this variation, indicating the need for further research explore the pricing structures of stolen data markets.

[2] A t-test is a statistical hypothesis test that is used to compare the tendency of two populations to have similar means and assesses if a null hypothesis is supported. In this case, a significant T value indicates that there is a difference between the log price of data.

Table 8: Location of Data By Mean Log Price

	Bank Accounts			CVV			Dumps			Fullz		
	0	1	T	0	1	T	0	1	T	0	1	T
Asia	4.82	4.48	0.54	2.38	2.80	-5.36***	3.57	4.32	-12.31***	4.02	3.92	0.36
Australia and New Zealand	0	0	0	2.38	2.59	-3.40***	3.64	3.29	3.21***	4.00	4.08	-0.28
Canada	4.74	5.25	-1.06	2.39	2.38	0.32	3.67	3.37	5.90***	4.04	3.66	1.31
Europe	4.95	4.12	2.22*	2.16	2.98	26.19***	3.49	4.02	-14.17***	3.68	4.64	-6.83***
Other	4.76	5.09	-0.63	2.37	2.89	-6.14***	3.52	4.48	-19.09***	3.98	4.41	-1.31
Russia	0	0	0	2.39	3.49	-3.82***	3.64	3.85	-0.50	0	0	0
United Kingdom	4.82	4.08	0.85	2.40	2.37	0.99	3.69	2.81	13.00***	4.03	3.34	1.53
United States	4.68	5.33	-1.37	2.63	1.67	29.86***	3.85	3.04	22.05***	4.10	3.47	2.82**

*p≤.05 **p≤.01 ***p≤.001; 0= all other nations, 1= selected country

Notes: The binary measures were computed for each geographic category. That is, Bank Accounts sold in Asia (1) compared to all other accounts (0), and the T indicates the t-test measure.

Though the forums and threads serve as an advertising space, the actual sale of data and services took place outside of the forum (Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Motoyama et al., 2011). The most common and preferred method of contact listed in all advertisements was ICQ (N=8000; 58.2%), a sort of instant messaging protocol that is extremely popular among the Russian hacker community (Chu et al., 2001; Holt & Lampke, 2010; Holt et al., 2012). ICQ is currently owned by a Russian service provider, which may increase its attractiveness to users in this region because their records may not be readily accessed by US law enforcement.

A small proportion (N=231; 1.7%) of posters also indicated that they used Jabber, an instant messaging protocol. Finally, a number of individuals indicated their willingness to use email (N=9,121; 66.4%), most commonly Yahoo. These may, however, be a less valued communication method relative to instant messaging protocols like ICQ that can be anonymized while connecting buyers and sellers (Chu et al., 2010; Holt & Lampke, 2010). Several sellers indicated they used multiple communications methods, specifically ICQ and at least one email address (N=3911), while a small proportion used both ICQ and jabber (N=183).

Table 9: Personal Contact Details Indicated By Posters

Contact Method	N	%
Email	9121	66.4%
AOL	16	0.1%
Gmail	250	1.8%
Hotmail	645	4.7%
Other	1123	8.2%
Rambler	21	0.2%
Yahoo	7055	51.4%
Yandex	11	0.10%
ICQ	8000	58.2%
Jabber	231	1.7%

There was also some variation in the preferred payment systems accepted by sellers. The majority indicated they would use electronic payment systems that transfer funds directly between two parties (see also Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Wehinger, 2011). The most common payment systems advertised were Liberty Reserve (N=2,304; 16.8%) and Web Money (N=1,528; 11.1%), since they allow for direct deposits of money between two electronic accounts (Suroweicki, 2013). A very small proportion also accepted payments via Yandex (N=64; 0.4%), a subservice of the Russian Internet search engine and service provider. Approximately 20% of sellers accepted payments through Western Union (N=2,673; 19.5%) or Money Gram (N=864; 6.3%), which is somewhat risky for both the buyer and seller as this requires visits to a physical location to send and receive funds (Chu et al., 2010; Holt & Lampke, 2010). Thus, electronic payments may be preferred because they can be anonymized or sent through fraudulently created accounts to shield the identity of all participants. In fact, 1,067 ads indicated that the seller would accept at least two forms of electronic payments, and 2,332 (17%) of all sellers indicated that they would accept both on and off-line payment methods. Approximately 10,210 (74.2%) of the ads in this sample did not indicate a payment mechanism due to the prerogative of the seller or the fact that they accepted a proportion of a payment in the case of money laundering.

A small proportion of ads (N=373; 2.7%) across 11 forums also indicated that sellers accepted payments through guarantors or escrow systems. A guarantor acts as an intermediary in a transaction by holding money on behalf of the buyer until such time as the seller releases the requested merchandise (Herley & Florencio 2010; Holt & Lampke, 2010; Wehinger, 2011). Once the buyer confirms they have received their purchase, the guarantor releases the funds to the seller. This quote from Forum 7 details the process of guarantor and escrow agents:

> **Escrow service**
>
> Escrow only insures money at time (fixed time) transactions.
>
> All terms of deal negotiated between the parties. Escrow they spend is not necessary.
>
> Escrow doesn't check goods or services.
>
> The principle of insurance transactions:
>
> 1. Buyer pays Indemnitor amount of transaction and fees for escrow service. Reports icq number for which this sum is intended.
>
> P.S. Under arrangement escrow fee may pay any member of transaction.
>
> 2. Escrow confirms receipt of money to another party.
>
> 3. The seller (service) provides direct product (or service) to another party to transaction without participation of Escrow.

> 4. After receipt and verification of goods (providing services) buyer contacts the Escrow and to announce completion of transaction.
>
> 5. Escrow pays money to seller (service).
>
> Escrow service fee:
>
> >500$ - 8%
>
> <500$ - 6%
>
> 3000$ and more - 5%

The use of guarantors provides a valuable, but optional mechanism to reduce the risk of being cheated by unscrupulous sellers. The lack of internal regulatory mechanisms makes it difficult for buyers to recoup losses they may experience if a product or service is not delivered after payment is sent (Herley & Florencio, 2010; Holt & Lampke, 2010; Wehinger, 2011). Thus, the use of guarantors ensures a higher likelihood of success and fosters trust between participants in otherwise risky encounters.

In addition to the payment mechanisms preferred by a seller, they would also list any ways that they would assist prospective customers after a purchase. There was some variation in the practices of sellers based on the product or service they offered. For instance, sellers offering PayPal, eBay, or bank account login credentials specified their responsibility to customers concerning the use of the data as in this ad from Forum 4:

> We are the trusted sellers of the PayPal accounts. You'll find lots of USA/UK:
>
> Unverified + Credit Cards (confirmed) - 1 WMZ/LR
>
> Unverified + Bank Acc (confirmed) – 1 WMZ/LR
>
> Verified + Credit Cards (confirmed) + Bank accounts (confirmed) – 3 WMZ/LR
>
> Here are some of the rules of the service:

> Seller is not responsible for sm (security measures); we check all the accounts manually prior [to] giving them to you. You'll also get a clean socks5 [proxy connection to access the account online]
>
> Seller is not responsible for the unsuccessfull [SIC] usage of the account.
>
> We may exchange your account in case the password won't match. Please inform us promptly!
>
> Please provide us with the screenshot in all the weird situations...
>
> You're free to do whatever you want to do with the account that you've bought. We take no responsibility on your further actions.

This vendor clearly indicates that the use and management of account details rested solely with the buyer, barring an error in the seller's data. To that end, the replacement of products was one way to ensure customers were satisfied with their purchase (Franklin et al., 2007; Herley & Florencio, 2011; Holt & Lampke, 2010; Wehinger, 2011). In fact, 4,285 (31.2%) ads indicated the seller would replace non-functional products. The duration of time for replacements varied based on the individual seller, as one provider in Forum 2 indicated he would "exchange invalids within 1 hour." The majority of sellers allowed 24 hours for replacements, as in this post from a dumps seller in Forum 4 stating: "24 hour replace for major customers (more time could be added for regular customers), I replace only invalid card numbers."

In addition, a very small number of sellers (N=305; 2.2%) offered free tests or samples of their products to customers (see Franklin et al., 2007; Herley & Florenico, 2010; Holt & Lampke, 2010). There is some debate over the nature of free tests or samples of data as some sellers may do this in an attempt to validate their products or draw in customers (Franklin et al., 2007; Thomas & Martin, 2006). At the same time, this tactic may be a way to attract unskilled actors in ripping forums to cheat them (Herley & Florencio, 2010). Individuals in more reputable forums may, however, not offer such tests because it would draw time away from paying customers and reduce both their profit margin and productivity (Herley & Florencio, 2010; Wehinger, 2011).

A small proportion of sellers (N=883; 6.4%) also indicated that they operated specialized customers service lines via ICQ or email, or noted that they would assist anyone after a purchase (Holt & Lampke, 2010; Wehinger, 2011). Such a measure may prove vital to maintain customers over time, particularly when customers are unable to use their services to the best of their ability. Customer service mechanisms may also be important for money laundering services so that buyers can readily obtain status updates on any transaction from the service provider.

A very small proportion of sellers (N=357; 2.6%) also had their products tested by the forum moderators. If a product is tested, it means that the seller must provide a sample to forum moderators or operators who then determine the validity of the sellers' claims (Holt & Lampke, 2010; Wehinger, 2011). Testing ensures that customers can trust the seller and know that they will not be cheated by an unscrupulous vendor (discussed further in Section C). These services may be more prevalent in insulated markets where there is a decreased likelihood of being cheated due to the organized (Herley & Florencio, Wehinger, 2011).

Though there were various mechanisms that sellers could use to attract customers, the majority of sales took place outside of the forums. In order to inform others about their experiences, customers could provide feedback about the transactions within a seller's thread (Holt & Lampke, 2010). Prospective buyers could use this information in order to identify reputable sellers and reduce their risk of loss. In fact, negative comments were identified in a small proportion of ads (N=2467; 18%), while positive comments were present in almost 12% of ads (N=1576; 11.5%)

In reading the feedback provided by customers, one of the most important comments made involves the use of the term "ripper" or "scammer" (Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Honeynet Research Alliance, 2003; Motoyama et al., 2012). These terms were commonly used by a customer when a seller's goods did not match their advertisement or sent nothing to the buyer at all. For instance, an individual in Forum 8 created a thread saying:

> I have been ripped off three times, so bad and I don't have any money left right now. what should I do? one of scammer named [removed] i was trust him, i bought 3 cvv from EU for test him, he did give to me, the second time, I ordered wu transfer. he never give to me, just took my money away.

This comment exemplifies the issue of ripping because the buyer did not receive what they ordered. As a result, the buyer has to spend much more than they intended because of their frequent losses (Herley & Florenico, 2010).

The use of the term "ripper" or "ripped off" enables prospective buyers to know who is unreliable, and thereby reduce their risk of loss or negative experiences (Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Wehinger, 2011). There is, however, no immediate way for individuals to determine when someone is a ripper from an advertisement. In fact, there were a number of complaints regarding rip offs in Forum 8 and one of the users tried to explain how one might identify rippers:

> - ripper wants to receive the money as fast as possible and he doesn’t care of the final [outcome] of deal; so the first main sign of ripper – desiring to receive the money fast, he thinks out a lot of reasons for this – pregnant wife, blocked keeper, drop’s worrying etc
>
> - ripper wants to be shown as well-knowing guy so he uses a lot of terms and specific words;
>
> - nickname; often greed and fieriness [fireiness] can be read in ripper’s nick which are changing like a gloves, so asa [SIC] we see Ecspress, Fast, Easy etc and almos[t] with the words “money”, “cash” etc we should be careful already and begin to verify this person.
>
> - Number of posts – potential ripper usually has too little posts for his registration date or too much posts – tries to make it’s number more. It’s also recommended to read what the person posts about on forums and make conclusion about his mind, if there are stupid posts or posts without any meanings – make conclusion yourself.
>
> - A lot of rippers usually post at the and [end] of there [SIC] posts “escrow accepted”. But when you talk that you want to work through escrow he usually finds lots of reasons don’t work through it.
>
> Conclusion: none of these things can tell you that this guy is ripper. But in combination it gets you the information about him and it’s better don’t deal with such guy.

As this quote demonstrates, there is no single way to identify a ripper based on their advertisement. Engaging in a transaction and posting the experience is the only way to determine the veracity of a seller (Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). Seven forums in this sample had comments related to ripping, and within four of these forums there was only a complaint lodged against a single seller. Such a small number of complains may be expected due to the difficulty in ensuring all sellers in any market are reputable (Herley & Florencio, 2010; Holt & Lampke, 2010; Wehinger, 2011).

The remaining three forums had a much larger proportion of ripping complaints. In fact, Forum 4 had approximately 16% of threads involving complaints about rippers. Forums 2 and 13 had more than 30% of all threads featuring complaints of ripping. Due to the large percentage of complaints evident in these two forums, they will be considered ripping forums in the remainder of this analysis. In fact, excluding the products sold in these two forums from the rest demonstrate substantive differences in the products sold (see Table 7). While dumps are the most prevalent product across all forums, CVVs were more common in forums with ripping complaints (see Herley & Florencio, 2010). In addition, removing ripping forums demonstrates the diversity of products used to remove money from stolen accounts or acquire information generally.

Table 10: Top Ten Products Sold By Forum Type

Including All Forums			Excluding Two Forums		
Product Type	N	%	Product Type	N	%
Dumps	5735	44.7	Dumps	2748	63.6
CVV	4481	34.9	Cashout Services	196	4.5
Money Xfer	303	2.4	Other Products	170	3.9
Other Products	277	2.2	Malware	151	3.5
Cashout Services	235	1.8	Dedicated Hosting	139	3.2
Bank Accounts	205	1.6	Drops	136	3.1
Equipment	198	1.5	Money Transfers	127	2.9
Malware	183	1.4	eBay and PayPal	108	2.5
Drops	165	1.3	Spam/Scam Materials	104	2.4
Dedicated Hosting	157	1.2	Plastics	86	2.0

While each thread was generally thought to be an ad space for the individual seller who created it, some sellers also attempted to advertise their products in others' threads. This sort of advertising is generally viewed as unacceptable, and was referred to as hijacking a thread. This is because the hijacker might draw customers away from the thread starter, and was specifically banned in three of the forums (see additional detail in Section C). This behavior is viewed as generally unacceptable in legitimate forums because it disadvantages the thread starter and creates competition (Wehinger, 2011). There were limited instances of hijacking present in this sample (10.4%), and the majority were found in Forum 8. In fact, 14.7% of all threads in this forum involved hijacking. In addition, there was a significant difference in the number of hijacking found in the ripping and non-ripping forums (Chi Square=31.104; sig.000), with 7.2%

of hijacking in ripping forums versus 3.2% in non-ripping forums. As a result, the number of hijacking incidents appears to be a substantive measure of the legitimacy of the forum.

To further explore any differences in the nature of the forums, the distribution of products across the forums were separated by the language used (see Table 8). While dumps and CVVs were the most common products sold in all forums, they were most prevalent in English language forums. This may be due to the issue of ripping, as all ripping forums in this sample used English as their primary language. Unfamiliar buyers may be more likely to purchase dumps or account data due to the perception that the data is inexpensive, and easy to use. Other products were, however, more evenly distributed across the two language sets, including eBay and PayPal accounts, personal accounts, other products, and cashout services. Russian language forums had a much larger proportion of products related to the acquisition of data, including hosting services, malware, and spam and scam materials. This provides initial support for the notion that there are multiple markets operating, that may be insulated based on the language used by participants (Herley & Florencio, 2010; Wehinger, 2011).

Table 11: Product Distribution by Forum Language

Product	English	%	Russian	%	Total
Bank accounts	171	83.4	34	16.6	205
Cashout service	102	43.4	133	56.6	235
CVV	4456	99.4	25	0.6	4481
Dedicated Hosting	29	18.5	128	81.5	157
Drops	96	58.1	69	41.9	165
Dumps	5381	93.2	356	6.8	5737
eBay/PayPal accounts	94	51.3	89	48.7	183
Equipment	126	63.6	72	36.4	198
Fullz	121	99.2	1	0.8	122
Identity Documents	68	76.4	21	23.6	89
Malware	33	34.4	150	65.6	183
Money Transfer	264	87.1	39	12.9	303
Other Financial Product	10	100.0	0	0.0	10
Other Product	126	45.5	151	54.5	277
Personal Accounts	40	40.4	59	59.6	99
Plastics	112	73.2	41	26.8	153
Skimmer	125	100.0	0	0.0	125
Spam/Scam Materials	26	21.3	96	78,7	122
Total	11380	88.4	1464	11.6	12844

B. Economics

The range of stolen data sold in the various forums sampled suggests that financial institutions are regularly being harmed by cybercriminals. The pricing structures noted suggest that there is significant variation in the advertised price for stolen data. In order to explore the influence of social and market forces on the advertised price for data, two linear regression models were created for dumps and eBay/PayPal account credentials using price as the dependent variable. These two products were selected as they are the most commonly identified products sold in samples of both IRC (Dhanjani & Rios, 2008; Franklin et al., 2007; Holz et al., 2009; Thomas & Martin, 2006) and forum-based research (Holt & Lampke, 2010; Motoyama et al., 2011). Both of these products were also sold in both ripping and non-ripping forums, making them ideal to explore the influence of market dynamics on advertised price (Herley & Florencio, 2010; Wehinger, 2011). Because of the positive skew in pricing distributions for both products, the log advertised price was used to reduce any measurement error (Oliver & Norberg, 2010).

Multiple social and market-related variables are included to consider how they may be associated with the advertised prices (see Table 12). Specifically, four binary measures (0=no; 1=yes) are included for the payment method a seller accepted: *Western Union*, *WebMoney*, *Liberty Reserve*, and *Escrow* payments. Sellers who accept Western Union payments may have higher prices because of the difficulty associated with accepting paper currency at a physical location (Herley & Florencio, 2010; Holt & Lampke, 2010; Wehinger, 2011). In addition, accepting Western Union payments may be a reflection of actor sophistication as they must operate with others to acquire, transfer, and accept payments (Herley & Florencio, 2010; Holt, 2013; Wehinger, 2011). The use of escrow payments may be associated with higher prices because the use of an intermediary increases the likelihood of successful transactions. The reduced risk of loss through escrow payments may increase the legitimacy of the market and trust between participants. In much the same way, Western Union may also be associated with more legitimate markets due to the inherent trust sellers must have that a physical payment will arrive from a buyer.

Table 12: Descriptive Statistics for Dumps and eBay/PayPal Credentials Regressions

	Dumps (N=5167)				eBay/PayPal (N=118)			
Variables	Mean	SD	Min	Max	Mean	SD	Min	Max
Dump	3.640	1.218	-3.22	8.99	----	----	----	----
eBay/PayPal	----	----	----	----	1.966	1.504	-1.61	6.68
Western Union	.422	.494	0	1	.127	.334	0	1
WebMoney	.251	.434	0	1	.144	.352	0	1
Liberty Reserve	.366	.481	0	1	.211	.410	0	1
Escrow Payment	.066	.249	0	1	.025	.158	0	1
Customer Service	.15	.352	0	1	---	---	---	---
Test/Free Samples	.03	.181	0	1	.05	.221	0	1
Free Replacements	.52	.499	0	1	.03	.158	0	1
Product Tested	.04	.195	0	1	.02	.130	0	1
Positive Comments	1.03	3.524	0	32	.05	.221	0	1
Negative Comments	.51	1.55	0	15	---	---	---	---
Hijackers	.08	.371	0	4	.24	.649	0	2
Ripper Forum	.523	.499	0	1	---	---	---	---
Russian	.04	.192	0	1	.64	.481	0	1

Electronic payment methods like WebMoney and Liberty Reserve may be associated with lower prices because they allow immediate money transfers between the buyer and seller (Motoyama et al., 2012). Rippers may prefer to use these methods in order to quickly acquire payments, but they are also standard mechanisms in the larger marketplace (Franklin et al., 2007; Holt & Lampke, 2010; Wehinger, 2011).

Four binary measures (0=no; 1=yes) for customer service are also included to understand any influence they may have on the advertised price for products. First, *customer service* is included based on the sellers' use of specialized customer service support lines through ICQ or email to aid customers in case of questions or issues (Holt & Lampke, 2010; Wehinger, 2011). This variable also measures whether the seller provides support for buyers after a purchase in order to facilitate the use of data. Either form of support may be associated with higher pricing because of the perceived legitimacy of the seller and their reputation (Holt & Lampke, 2010; Wehinger, 2011). This measure is not included in the model for eBay and PayPal credentials due to missing data.

Second, a measure was included for sellers offering *tests or free samples* of data in order to attract customers. Sellers who offer free data may do this to validate their products (Dhanjani & Rios, 2008; Franklin et al., 2007; Holt & Lampke, 2010), but it may also serve as trap to rip off unsuspecting customers (Herley & Florencio, 2010; Wehinger, 2011). Sellers who do not give out tests may be more legitimate as they do not waste time distributing data without payment or reduce their profit margin (Herley & Florencio, 2010). Thus, the use of free samples should correspond to lower advertised prices in order to draw in customers (Herley & Florencio, 2010).

The third measure of customer service was free *replacements* for invalid or expired accounts. Some sellers indicated that they would replace accounts that were inactive within a certain time period after purchase (Franklin et al., 2007; Holt & Lampke, 2010). This may be a meaningful measure of customer service, though it limits sellers' profit margins through the free distribution of information (Herley & Florencio, 2010). As a result, sellers offering free replacements may also have generally lower prices in an attempt to attract customers and possibly rip them off (Herley & Florencio, 2010).

The final measure of customer service is a binary measure based on whether a seller has had their *product tested* by the forum moderators. Reputable forums provide resources for sellers to have their products validated and then publicly reviewed (Holt & Lampke, 2010; Wehinger, 2011). Sellers give a sample of data to a tester, who then assesses the validity of the sellers' claims. Testing ensures that customers can trust the seller and their claims, and know that they will not be cheated (Holt & Lampke, 2010; Wehinger, 2011). These services are not present in all forums, and may be absent in forums where rippers are active (Wehinger, 2011). As such, these services may increase prices because they minimize the risks for participants, and should generally be found in more insulated and secured markets (Herley & Florencio, 2010; Wehinger, 2011).

To assess the relationship between pricing and customer comments, two continuous variables were created based on the number of *positive* and *negative feedback* posted in a given thread about a seller's products. Since transactions between buyers and sellers take place outside of the forums, customers could discuss their experiences within sellers' thread to describe their encounters with the seller (Holt, 2013; Holt & Lampke, 2010; Motoyama et al., 2012 Wehinger, 2011). If a customer did not feel satisfied, either because the goods were not as advertised or were not sent at all, they could state their dissatisfaction publicly. In much the same way, those who were pleased by their interactions could make a post about the seller's practices or data (Holt, 2013; Holt & Lampke, 2010; Motoyama et al., 2012 Wehinger, 2011).

The presence or absence of feedback in any thread may not necessarily be a negative reflection on the seller. They may have recently posted their ad, or had minimal customer interest in their products. In addition, positive feedback could be fraudulently created by a seller in the hopes of generating interest in their products (Herley & Florenico, 2010; Holt, 2013). Both measures are included in the model for dumps, though only positive feedback is used in the eBay and PayPal credential model due to multicollinearity issues.

A continuous variable was also included to assess the number of instances of *hijacking*, where individuals place an ad for their products in an existing sellers' thread. This behavior is viewed as generally unacceptable in legitimate forums because it disadvantages the thread starter and creates competition (Wehinger, 2011). Forums with substantial management and oversight by moderators do not allow such practices and can ban a user for this activity (Holt, 2013). Less reputable forums may allow hijacking to occur with no punishment for users. Taken as a whole, the presence of hijackers in a thread should be correlated with reduced prices for stolen data.

An additional binary measure was created to examine the relationship between product pricing and the forum as a potential *ripping forum* (0=no; 1=yes) (Herley & Florencio, 2010). This variable was created based on the number of comments involving complaints about ripping or scamming in each forum. There is no established metric to determine whether a forum is saturated with rippers, but only seven of these forums had comments related to ripping. Four of these forums (57.1%) involved complaints about a single seller. Such a small number of complaints may be expected due to the difficulty in ensuring all sellers are reputable (Herley & Florencio, 2010; Holt & Lampke, 2010).

The remaining three forums had a much larger proportion of ripping complaints. In fact, Forum 4 had approximately 16 percent of threads involving complaints about rippers. Forums 2 and 13 had more than 30 percent of all threads featuring complaints of ripping. Due to the large percentage of complaints evident in these two forums, they were coded as ripping forums (1=Yes), and the remainder were coded as non-ripping. It is expected that prices in these forums will have lower advertised prices to attract participants (Holt & Lampke, 2010). Those with fewer instances of ripping should have higher costs because the participants recognize that they can trust one another (Holt & Lampke, 2010). This measure is excluded in the model for eBay and PayPal data due to the small proportion of these data advertised in ripping forums (22%) and multicollinearity issues.

A final binary variable was included for the primary language used by forum participants (0-Eng; 1-Rus). English speakers who do not have any familiarity with Russian will have difficulty communicating or participating in the market. Rippers might be attracted to English-language forums because it increases their pool of buyers globally, and access to inexperienced buyers. In addition, ripping forums used English as their primary language. As such, language serves as a way to explore price differentials across multiple markets. Russian-language forums might also have lower prices because of market insulation and increased trust between participants (Herley & Florencio, 2010; Holt & Lampke, 2010; Wehinger, 2011).

An OLS regression model was conducted to explore the factors affecting the log advertised price for dumps and eBay and PayPal credentials respectively. Multicollinearity diagnostics indicated that no variance inflation factor (VIF) exceeded 4 (with critical cut off point of 10) and no tolerance was below 0.2 in both models. The model overall explains nearly fifth (about 20%) of the variance in our dependent variable.

The first regression model, examining the log advertised price for dumps, finds some support for the notion that there may be multiple markets operating with different pricing structures (Herley & Florenico, 2010; Wehinger, 2011). Since our dependent variable is log-transformed, we have to interpret regression coefficients as percent changes. The coefficients were exponentiated in order to be interpreted as percent changes, and are only reported here for significant coefficients (Olivier & Norberg, 2010). To that end, those who accepted Western Union payments advertised dumps at nearly 75 percent higher prices than others (see Table 13, Model 1). Interestingly, WebMoney payments were associated with lower prices, while Liberty Reserve payments were associated with higher prices. The acceptance of escrow payments had the highest effect of 297 percent difference in price.

The presence of customer service measures was also related to the advertised price of dumps. Sellers who offered customer service lines or support were associated with higher prices. The use of free replacements was associated with generally lower prices, though the tests and free samples were non-significant. Those with products tested by forum administrators had increased overall prices that were almost 101 percent higher than non-tested products. This provides limited support for the assertion that sellers' behavior affects price, and that more organized forums may have differential prices for products based on trust (Herley & Florenico, 2010; Wehinger, 2011).

Sellers who received positive comments were associated with a three percent decrease in price. The presence of hijacking in an ad was also associated with lower pricing, as were those advertised in ripping forums. In fact, ripping forum products had prices that were 44 percent lower than non-ripping forums. Thus, reduced prices may draw in unsuspecting customers but increase their risk of loss (Herley & Florencio, 2010).

Table 13: Regression Models of Log Price for Dumps and eBay/PayPal Credentials

	Model 1 Dumps N=5167				Model 2 eBay/PayPal N=118			
Variables	B	S.E.	Beta	% Change	B	S.E.	Beta	% Change
Western Union	.564	.044	.299***	75	1.658	.401	.369***	424
WebMoney	-.319	.043	-.114***	-27	1.600	.477	.375***	395
Liberty Reserve	.177	.044	.070***	19	-1.619	.478	-.442***	-80
Escrow Payment	1.379	.077	.282***	297	-.885	1.064	-.093***	-58
Customer Service	.096	.049	.028*	10	---	---	----	---
Test/Free Samples	-.085	.086	-.013	---	.475	.633	.070	---
Free Replacements	-.409	.038	-.168***	-33	1.631	.955	.171	---
Product Tested	.699	.081	.122***	101	4.784	1.431	.412	---
Positive Comments	-.033	.006	-.095***	-3	2.459	.615	.361***	1069
Negative Comments	-.026	.014	-.034	---	---	---	---	---
Hijackers	-.112	.045	-.034*	-11	-.576	.2235	-.249***	-43
Ripper Forum	-.585	.050	-.240	-44	---	---	---	---
Russian	-.824	.094	-.130***	-56	-1.515	.313	-.484*	-78
Intercept	3.895	.063***			2.731	.291***		

Model 1: F=109.143; .000, R2= .216, Adjusted R2=.214; F=17.204; .000, R2= .617, Adjusted R2=.581
*p≤.05 **p≤.01 ***p≤.001

To that end, Russian language forums had prices that were 56 percent lower than English language forums. These forums may be somewhat insulated from outsiders and inexperienced buyers, thereby allowing sellers to offer lower prices for goods (Herley & Florencio, 2010; Wehinger, 2011). There might be also more trust on a smaller Russian-speaking network, with potentially larger repercussions for rippers that might lead to lower prices.

An additional regression model was created to assess the factors affecting the log advertised price for eBay and PayPal credentials (see Table 13, Model 2). This model explains over 50 percent of variance in the dependent variable, but the findings need to be viewed with caution due to the smaller number of transactions of this type.

The findings suggest there are some differences in the factors affecting costs for data relative to those of dumps. Those who accepted Western Union had higher advertised prices, as did those who accept WebMoney. In fact, sellers accepting these payment types advertised their data at nearly four times higher prices. Sellers accepting Liberty Reserve and escrow payments had significantly lower advertised prices, in opposition to the model for dumps.

Customer service metrics had generally little impact on the overall advertised price. Only those who received positive feedback had generally set much higher prices for their products. The magnitude of percent change also need to be viewed with caution in this model due to small sample. This finding supports the idea that positive feedback provides a measure of trust in the market, though it is different from the relationship observed for the price of dumps.

Hijackers in the thread were associated with 43 percent lower advertised prices in keeping with the relationship observed in the price for dumps. Advertised prices were also 78 percent lower in Russian language forums in keeping with the price of dumps. Thus, these relationships suggest there are different markets operating with differential pricing structures based on insularity and the presence of prospective rippers (Herley & Florencio, 2010; Wehinger, 2011).

C. Social Organization

1. Mutual Association and Participation

The economic models demonstrate the presence of multiple markets with differential pricing structures based in part on the organizational dynamics of the market. In the following section, the social organization of market actors is explored using qualitative analyses to assess the relationships between participants. In examining the posts, it is clear that the forums facilitate connections between interested individuals to readily sell or acquire data. Services are advertised to interested parties. The language of ads was relatively consistent across all the forums, and began with an initial post by the seller or prospective buyer. In addition, there was some variation in the organization of participants in stolen data transactions based on the product or service offered.

Sellers offering credit cards or specific financial products appeared to involve two individuals: the buyer and seller. Individuals offering encashment or money laundering services as in the above example operate through partnerships and small groups. This is due in part to the fact that encashers cannot obtain funds without data. For instance, a seller in Forum 13 offered their services to place funds onto ATM cards which could then be used to obtain money from stolen accounts, and clearly used the term "we" in the course of their ad and indicated that a degree of cooperation was needed to complete a transaction, stating:

> Provision of services for the identification of electronic payment systems and preparation of ATM cards for drops [RU].
> **WM [WebMoney] registration process:**
> We provide you with a photocopy of the passport. Using this information you register a new WMID [WebMoney Identification Number]. You receive a formal certificate under the data of the photocopied passport, you pay for the application to receive a personal

> certificate. You give us the data of the newly registered WMID used during registration. We write this data into the application and send them to the Certification Center. After a while you receive the personal certification for this WMID.
> ONLY you will have access to this WMID, and no one else.
> **Terms of service:**
> - The transaction will be conducted through a guarantor, or 100% prepayment.
> This is how we see that a person has a serious attitude, that he has money and that he will not go missing when the order is completely ready.
> - We have the right to require profiles on forums in our topic in order to further check on the financial ability.
> - Service has the right to refuse a client without explaining the reasons!
> - We have the right to refuse to work after accepting payment in case of force majeure circumstances. If we refuse, we will return the money which we have received within a 2-day time period, after advising of the refusal. Take into account that we mean 2 BUSINESS days.
> - The use of accounts (atm cards) means encashment of funds, and not their storage. It is prohibited to store money.
> - Pouring dirt from CIS countries into accounts (atm cards and payment systems) is STRICTLY FORBIDDEN! If such cases are discovered, the account will immediately be blocked, no payments will be made, and the account owner will be listed as a black.
> - You must understand that a drop is a person, and a person can be greedy, piggy and the thirst for illicit gain from easy money, remember that however it might be, the card is not yours, it belongs to another person and at any moment he can block it. Our task is not to allow this, and up to now we have done this well.
> - Our service is not liable if your account is closed by the payment system for fraud and various fraudulent actions.
> - We have the right to REFUSE to restore control over the accounts if they are lost by fault of the client. We control our drops, but we are not required to take him around the banks in order to restore your ATM card, or other things. Be careful with things.

This example demonstrates that some service providers are very technologically sophisticated and understand how to circumvent security protocols. It is not clear why they offer their services to others, when these skills could easily be employed for themselves to gain a profit. Regardless, this demonstration of specialized knowledge suggests participants within these forums are colleagues or peers based on their minimal mutual participation in offending (see Herley & Florencio, 2010; Holt & Lampke, 2010).

Though the acquisition of data or services may be more individual in nature, the sales processes within the forums are participatory and extremely social in nature. Individuals can post positive or negative feedback on the basis of their experience with a seller (Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). Such information is critical in helping prospective buyers find reputable sellers, and to establish the reputation of good sellers. For example, an individual selling credit cards from various countries in Forum 1 received several positive posts for his products, including “everything went well. You can trust this person.”; “I got 2 car[d]s. Everything okay! My trust!”, and “the carton [card] is in good working order! I bought it and I will only buy from him in the future. There were no problems. Thanks!” This sort of feedback helps demonstrate a sellers’ level of trust and build out their reputation over time if positive feedback persists (Holt & Lampke, 2010).

Buyers who post positive feedback also help the other actors within the market determine the most reliable operators to work with. Those who receive multiple positive comments from a seller are likely to be trustworthy and give quality service. For instance, an individual offered a service to withdraw cash from Western Union, Web Money, or Yandex and then place those funds either on ATM cards or transfer them to the client via WebMoney. He received multiple positive responses from customers over time:

> Killz: I withdrew cash everything was okay
>
> Cypher: Everything was excellent and online.
>
> Iglio: I worked with this person everything was ok
>
> Horvath: I laundered wm without any problems. Everything was good.
>
> Nash: This is not the first time that I have worked with the TS, everything was tiptop!
>
> Nod31: I use the services... twice and everything was precise and the payment was on time according to the contract.. service +1

The seller received no negative feedback in the course of the thread, and comments from repeat clients suggests that he is tremendously reliable. Thus, buyers can utilize feedback to seek out providers with strong reputations to have a more satisfactory experience (Holt & Lampke, 2010; Wehinger, 2011).

Sellers who provide data that is incomplete, cards that are not valid, or simply take payments and do not deliver any products typically receive negative feedback from buyers (Holt and Lampke, 2010). Public recognition of a bad sale or transaction is critical so that other forum users can understand when and how they may have be swindled. For instance, an individual from Forum 4 named Jackson posted an ad selling credit card data from various countries and received substantial negative feedback because of his use of bad data and poor response times. Prospective and actual customers noted this in their posts, stating:

> **Nickly**: registration [on the sellers personal shop website is] temporary disable[d] so what is the essence or advantage of this advert here when someone cannot make reg[istration]
>
> **Stan**: They have been selling the same dump since April, very stale, support is no help!
>
> **Nickly**: yes Stan bulshit [SIC] garbage dumps in his shop…all claims dumps there are dead since April and is still claiming 90 percent valid… lol buyer be careful
>
> **Vendor**: dumps extremely low working % You will be lucky to card yourself a happy meal with one of these shit ass dumps… Thumbs down.

This feedback illustrates that the seller is offering poor quality data and poses a risk to market participants. In turn, he may experience lower sales relative to actors with more positive reviews and products (see Holt & Lampke, 2010; Motoyama et al., 2011).

Since the market enables individuals to purchase products that they may not be able to use effectively use on their own, some sellers were careful to quantify or debate negative feedback. Sellers who can demonstrate that a poor experience with a product or service stems from the buyers error may be able to maintain a share of the market and minimize fallout from negative reviews (Holt & Lampke, 2010). For instance, a customer named DILS in Forum 5 purchased a dedicated server space from a seller. He posted a negative comment stating: “it's not worth getting a dedicated from him. 2 times I got one from him and even for a week they didn't work he shuts off accounts specially.” The seller then posted a relatively detailed response in an attempt to negate this feedback, including evidence of their chat logs:

> Dear smarty DILS! Let's dot all the i's so that everyone understands what's going on. First of all, I only sold one dedicated to you. and second (according to your own words) .your friend supposedly for your friends supposedly for you, here is an excerpt from the correspondence:
>
> "DILS_Host-Portal (21:12:34 15/08/2009)
>
> it simply already the second thing that happened
>
> Prof (21:13:22 15/08/2009)
>
> we mean the second time?
>
> Prof (21:17:13 15/08/2009)
>
> in this asya I have no such events
>
> Prof (21:17:19 15/08/2009)
>
> in the history
>
> DILS_Host-Portal (21:17:43 15/08/2009)
>
> I tossed beans over to webmoney my account is r298.......

DILS_Host-Portal (21:18:01 15/08/2009)

but we came to an agreement my friend ...."

I'm not telepathic and can't see who is giving dedicateds to whom or who is reselling them and even more so I can't bear any liability for this.

Furthermore - if in your words the ded did not work for a week already - this is not in any way associated with me disconnecting accounts, as you said yourself, already you can't call your story anything more than bullshit nonsense.

Well and that's the main portion of our communication with you:

[/b]DILS_Host-Portal (21:05:45 15/08/2009)

Hi!

Prof (21:05:59 15/08/2009)

hey

DILS_Host-Portal (21:06:07 15/08/2009)

I have a question for you

DILS_Host-Portal (21:06:22 15/08/2009)

I bought a ded from you

Prof (21:06:31 15/08/2009)

well

DILS_Host-Portal (21:06:34 15/08/2009)

it hasn't been working for weeks now what should I do

DILS_Host-Portal (21:06:52 15/08/2009)

I haven't been able to do anything it

Prof (21:06:57 15/08/2009)

if a week already - then I think nothing

DILS_Host-Portal (21:06:57 15/08/2009)

even download

DILS_Host-Portal (21:07:22 15/08/2009)

maybe you can give another ded ?

DILS_Host-Portal (21:07:45 15/08/2009)

because I just can't use it

Prof (21:11:31 15/08/2009)

well alas, the guarantee is only 5 days, and what you did there or didn't do there I have no way of checking

Prof (21:11:34 15/08/2009)

is that logical?

DILS_Host-Portal (21:12:17 15/08/2009)

ok tks [thanks]

[b]

Thanks and I went off to go scribble some backbiting lines. I'm freaking out over you. really :)

Negotran: DILS but for me it didn't work not only for a week, that even for one day. The replacement question took 3 days to decide, and then they shoved some shady deds at me, almost giveaways – and wanted me to say thank you, $16 is not in amounts for which you stir up a fight and further I simply don't recommend service, and with such an attitude it will die off by itself soon.

This example illustrates the depth of exchanges that occur outside of the forums, and the length that participants may go to in order to legitimate claims about a seller and their experience. In addition, the posts demonstrate that some sellers will attempt to avoid any tarnish on their reputation or status within the forum (Holt & Lampke, 2010). As noted in the third post by Negotran, too much defensive commentary about one's products may actually be viewed as an unusual behavior. As such, prospective buyers may be better served seeking out encounters with other vendors who received positive feedback to avoid being ripped off (Holt & Lampke, 2010).

Taken as a whole, stolen data markets are peer driven, similar to illicit online markets for prostitution where customers discuss and rate the services of sex workers (Holt & Blevins, 2007; Milrod & Weitzer, 2012; Sharp & Earle, 2003), and other more legitimate business markets around the world (Aspers, 2011). The use of peer review comments in on-line markets enable individuals to identify reliable providers and assess prospective risks in advance of a purchase, whether in the case of eBay or Amazon (Aspers, 2011) or prostitution forums (Holt & Blevins, 2007; Milrod & Weitzer, 2012). Stolen data markets are also similar to prostitution forums in that users who are thought to provide reliable reviews and have participated in the site for some time may be given greater respect and status (Holt & Blevins, 2007). The use of public comments serve as a key risk reduction mechanism for actors within stolen data markets. The ability to engage those who appear to be more legitimate providers gives buyers a prospective edge when dealing with potentially unreliable actors. Should they feel the risk is worthwhile, they can also utilize guarantors or escrow agents in the course of a transaction. (see Sections A and B for further detail). The addition of a third party as a guarantor adds a layer of organizational complexity to any transaction. The guarantor must fulfill their duties and ensure delivery of payment to the seller or return funds to the buyer depending upon the outcome of the transaction (Herley & Florencio, 2010; Holt & Lampke, 2010; Wehinger, 2011). This decreases the efficiency of the sales process due to the need for contact between all parties to ensure smooth delivery of products and payments. In addition, two of the sites indicated that guarantors could charge a fee for their services based on the amount being paid in the transaction. This was exemplified in a post explaining guarantor services in Forum 4:

> Guarantor services
>
> The guarantor of a forum has been created so that you will not be deceived… By conducting a transaction through a guarantor, you can be sure that you will not be deceived.
>
> Terms for working through a guarantor:
>
> 1. The buyer and the seller must reach agreement on working through a Guarantor.
> 2. The buyer and the seller must contact the Guarantor using icq.
> 3. One of the Parties to the transaction gives money to the Guarantor, and the other goods.
> 4. The guarantor's services are free up to $30. [After this amount, the guarantor will take a variable percentage of the total for their efforts. The rates are described below]
>
> up to 500 wmz - 8%
>
> from 500 wmz - 6%
>
> from 3000 wmz - 5%

The use of guarantors or escrow agents directly increase the transaction costs for buyers, but reduce the prospective economic harm that stems from unsuccessful exchanges (Herley & Florencio, 2010; Wehinger, 2011). Thus, the use of guarantors increases the number of participants within any transaction, while promoting positive associations between buyers and sellers (Holt and Lampke, 2010). In addition, these comments demonstrate that the process of buying and selling is social in nature and a peer-driven process.

Though the market is participatory in nature, there is one form of interaction that is unacceptable in reputable forums: hijacking another person's thread. Each new thread is generally viewed as the space for that individual and their product or service. As a result, anyone who attempted to advertise their products in another person's thread violated the basic norms of the forum. This sort of behavior is disruptive because it may draw customers away from the thread starter and create competition in the market. In fact, Forum 13 specifically stated in its rules that it is strictly forbidden to "offer your goods/services in someone else's sales topic."

There were limited instances of hijacking present in this sample of forums, as only 10.4% involved any potential hijacking. The majority were present in Forum 8, one of the two ripping forums in this sample, with 14.7% of all threads involving hijacking. In addition, there was a significant difference in the number of hijacking between the ripping and non-ripping forums (Chi Square=31.104; sig.000), with 7.2% of hijacking in the ripping forums versus 3.2% in the non-ripping forums. Thus, an abundance of hijacking across multiple threads may be an indication of the general reputation and status of a forum. Those with a higher proportion of hijacking may be less credible than others, as well as more poorly organized and managed.

2. *Division of Labor*

The content of these forums demonstrate that actors have substantive specialization for roles based on the variety of products and services available to either obtain or manipulate stolen data and financial credentials (see Table 4). A substantial proportion of sellers offered different forms of illegally acquired financial information ranging from dumps (44.7%), cvvs (34.9%) to specific bank accounts (1.6%) and even eBay and PayPal account information (1.4%). Stolen data represents the common resource offered across all forums (84.3% of all ads), suggesting that hackers and data thieves utilize this space to obtain a profit from data they have acquired (Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011; Thomas & Martin, 2006).

Resources to access and remove funds from accounts comprised a much smaller proportion of the market, with cashout (1.8%), money transfers (2.4%) and drops services (1.3%) providing a niche service in the market. A small proportion offered resources to use data in the real word through plastic credit card creation (1.2%), as well as passport scans and identity documents (0.7%) to support false identities. Finally, the tools available to support fraud through spam (0.9%), malware (1.4%), and infrastructure to either acquire data or manipulate it (1.2%) demonstrate that individuals could purchase virtually any service and move sequentially from vendor to vendor in order to obtain funds from stolen accounts. Thus, these markets engender fraud and cybercrime from end to end (Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011).

The content of ads also demonstrated clear segmentation in the services offered by providers. Across all the forums sampled, no dumps seller advertised cashout services, and no drops service offered bank logins or web hosting. There was also no evidence of vendors working together, or recommending a specific vendor so that a customer could utilize their product. Individuals would discuss when a drop service or encasher was needed, but did not suggest someone for this process. As a result, the markets engender a clear division of labor and specialized knowledge to facilitate identity theft and cybercrime generally. Those unskilled actors can utilize all service providers while those with sufficient skill can simply pay for the individual services they need at any point in time with fewer costs to maximize profit (Herley & Florencio, 2010; Wehinger, 2011).

Beyond the individual division of labor, the forum moderators and administrators play a critical role in managing user behavior and regulating encounters between participants (Chu et al., 2010; Holt and Lampke, 2010; Motoyama et al., 2011). Administrators were noted across all the forums, though their involvement in each thread was very limited. Administrators have the ability to dictate the rules for participation in the market as a buyer or seller, but sellers can specify their terms and conditions of any given sale. For instance administrators in eight of the forums could review and test products sold by any vendor, as noted in the following post from the administrator of Forum 7:

> **Checking Rules**
>
> Checking your goods will take place voluntarily or if the administration of the forum requires it.
>
> Checking of goods is done by **ICQ [number removed]**
>
> The check last from one to three days.
>
> After the check, the moderator guarantees that there will not be any stupid flames in the topic and that the quality of the goods will not be discussed. The moderator will write a review on this and close the topic. If a requirement to provide your product for testing is refused, you risk being banned, and your announcement will be erased. No money is taken for testing.

You provide the product for the test in the same configuration in which you sell it

These posts demonstrate the inherent value and benefits of checking processes for sellers and buyers because they can trust the reputation and credentials of the seller (Holt & Lampke, 2010). The process also helps to reduce infighting within the forum, but adds a layer of organizational complexity through the need to engage administrators in the process. Thus, checking services add to the division of labor within any forum and support the notion of market segmentation and insulation that reduces risk of loss (see Herley & Florencio, 2010; Wehinger, 2011).

A number of sellers would also indicate when their products had been tested on a separate site to help demonstrate their reputation and prospective trust in the market. For example, a seller from Forum 5 noted:

> I would like to offer you the services for filling up your mailbox with pinodosy [spam to people from NATO countries, esp. The USA]... I HAVE UNDERGONE A CHECK ON THE CLOSED FORUM [name removed] I am also ready to undergo a check on your platform. I work both through the guarantor of any forum as well as through protection.

Similarly, a seller in Forum 4 placed the following comment at the end of his ad: "Checks have been undergone with the guarantors of the forum [name removed], and [forum name removed]- checks have been undergone with the guarantor of the forum." Such comments help to demonstrate a seller's prospective reputation and trustworthiness.

The use of checking or tests helps to reduce the risk of loss, but does not guarantee satisfaction due to the voluntary nature of the process. This was exemplified in Forum 10 where there were two incidents where actors lodged complaints against sellers despite having their products checked by administrators. In particular, an individual posted an ad for a scheme to infect users with malware in order to make money. Potential buyers complained about the seller's scheme and product, stating:

> Goldburg: The check doesn't mean anything. The majority of cases they check theoretically (or for the presence of such), and at the same time neither in the post of the TS [Thread Starter] nor in the post of the checker is their word said about the labor input, needed [SIC] for certain skills, knowledge etc.

> TS, it would be easier to describe into words in the topic and not waste your time or anyone else's.
>
> ALT: well this is just what means absolutely nothing. The asky [ICQ number] is a crooked nine-number [nine digit number], no info has been provided on undergoing a check on other forums.. In general I trust the moderator..
>
> TS:I can undergo a check on any board, but I don't see any sense in this because I'm selling to 10 people, two are thinking about it, another three will buy within the next few days, so whoever is interested, they will buy it, and I don't want to bother with tests on other boards and waste time on that.

Thus, sellers who do not feel it necessary to pursue checks do not have to bend to pressures depending on what they have to offer. Those who do so can increase their prospective level of trust in the market, and may be able to increase their sales (Herley & Florencio, 2010; Holt & Lampke, 2010).

Interestingly one of the ripper forums in this sample (Forum 1) had no structure in place for testing products, though they were attempting verify sellers. This process was largely ineffectual due to the fact that the sellers were unwilling to work with the forum administrator. He wrote about this issue, stating:

> I protect everyone from any unverified seller and write about it. It says many that these guyz don’t want to pay some fee like $50 (may be they have no any money.^))) LOL and also don’t want to give me their stuff 4 checking))) IT SAYS MANY! Also you can find these rippers/sellers on other forums with fake comments. THERE WILL NOT BE RIPPERS!!! AS SOON AS ANY WANTS TO BE VERIFIED THERE WILL BE CATEGORY WITH VERIFIED SELLERS (also verified seller can rip you off with good

> amount) you can use escrow. Also I've all stuf [SIC] whats market can provide anyone (dumps, cc, skimmers, drops)…but I don't sell anything and don't provide any service. I'm only admin and provide you with forum 4 deals and talks.

This demonstrates that forums whose members are unwilling to have their products tested to establish a reputation may be more risky than others. Such markets may be less insulated from outsiders attempting to cheat buyers, thus careful actors should attempt to avoid these sites in order to reduce their risk of loss (Herley & Florencio, 2010; Wehinger, 2011).

An alternative to testing or checking services was to buy from individuals with a sales reputation provided by site administrators. The verification process helps to demonstrate a level of trust in the seller based on their market performance and customer feedback (Holt and Lampke, 2010; Motoyama et al., 2011). Two of the forums in this sample allowed individuals to become verified sellers. In Forum 9, a user could become a verified vendor by contacting the forum administrators, as in the following post:

> Support VPro Celler provides access to sellers account only after:
>
> 1. Verification via PM [Private Message] to one of forums
>
> 2. Talking with contacts listed inside promotional theme [advertisement] of service/vendor…

A similar comment was noted in Forum 8, where the supermoderator posted:

> **Attention!** All the Members here are advised not to deal or buy anything from Unverified Sellers! If you do so and you get scammed! We are not responsible! Only buy from those sellers who are verified! For all the Sellers who want to get themselves verified contact me through pm!😛

This forum was, however, heavily populated with instances of ripping and complaints from customers. Thus, the use of verification mechanisms neither guarantees a satisfactory experience nor much protection for users. In turn, verification should not be viewed as a metric of organization because it is insufficient in reducing the likelihood of ripping (Herley & Florencio, 2010).

Administrators can also play a pivotal role in the management of forum encounters based on the comments of participants and the behavior of users. The administrative techniques noted help to promote trust between participants, but does not eliminate the risk that participants face from being cheated or swindled. The feedback individuals post describing their experience within a seller's thread helps serve as a warning and informal social sanction. These claims can, however, be disruptive and lead to infighting, reducing the economic viability of a thread and the perceived legitimacy of a forum. (Franklin et al., 2007). To ensure that only the most serious claims are given public recognition, administrators in four of the forums required that complainants provide proof that they had been cheated or ripped off, such as chat logs or proof of payment. For instance, moderators from Forum 3 stated that claims of cheating or bad reviews were "Strictly Forbidden" and specified the consequences for such activities in their rules document:

> To leave fictitious rules in topics. The fact that a transaction has been carried out must be confirmed by the appropriate proof upon the first request of the administration. If this cause is violated, the user leaving a fictitious review will be banned, and perhaps even permanently..

> To leave the following type of message: “TS [Thread Starter] is a burner [rip off artist] you shouldn’t have anything to do with him.” We will immediately post the blogs proving guilt..
>
> Reviews from users with 1-10 messages who have not been on the forum for long will be deleted at the discretion of the moderator.

The fact that administrators could delete posts from users with less than 10 messages demonstrates the limited barriers to entry for participants. Should a user be banned for complaints, they could make additional accounts in order to continue to disrupt the market with bad reviews. Since forum administrators can utilize bans to sanction and edit user content when necessary and ensure orderly communications within a thread, this supports their prospective position at the top of a forum's division of labor.

The use of banning was differentially enforced since there were banned user accounts present across all of the forums.[3] This may stem from variations in the administrative capabilities provided, since some forums had posts with notes of administrative edits and deletions, while others simply listed the word banned next to a username in a thread. Regardless of the way that bans are retroactively applied to user posts, the power to ban users is a key mechanism for administrators to regulate transactions and exchanges within a forum.

There was also evidence of purposive relationships between forums based on buyers and sellers referencing ads and content from other sites through external links. Some sellers noted when their products were sold or reviewed on other sites as a means to validate their reputation. For instance, an individual in Forum 2 sold log files obtained from compromised computers that contained email address information, bank and financial account data, and Facebook

[3] There was no clear evidence to account for the variations observed. The structure of the forum may allow administrators to completely remove any appearance by a user or simply keep them from posting in the future.

information. In their ad they wrote: “sale verified at the following forum” and provided three external links, as well as a second post indicating “passed testing [by forum administrators] on [link removed].” External links were found in seven of the legitimate forums, and were largely absent from the sites with a preponderance of rippers. This small measure further demonstrates there are multiple markets operating as cross-listed content can aid in dispute resolution or identity construction. In addition, these ads should not be interpreted as proof of forum actors working together, but of concurrent participation in multiple markets at any time.

3. Extended Duration

The final element of organizational sophistication within the Best and Luckenbill (1994) framework is extended duration. This concept reflects the basis of the length of time that a group has been operating. With that in mind, there are two different data points that can be used to assess duration within the forums. The first is the date of the first post captured in the sample of threads. To that end, three of the forums had posts dating back to 2007, meaning that the site had been active for five years based on the last active date of posts. The substantial duration of time is unusual to many hacker forums (see Holt, 2007; Meyer, 1989), suggesting that these groups are established resources in the market for stolen data. In fact, the oldest advertisement in the data set was posted in Forum 13 dated June 6, 2007. The seller in this ad offers passport scans and identity documents made to order with stolen personal information. Both the ad and the seller were still active and operating at the time of data collection, though this example may be an isolate in the larger data set since the majority of ads were posted within the last two years (see Table 14). In addition, one forum had posts made over a two and half year period, and three had two years of data. As a result, more than half of the sample appears to have an extended duration consistent with more formal organizations (Best & Luckenbill, 1994; Holt, 2009).

Table 14: Duration of Forum Participation By Join Date of User and Post Dates*

Forum Number	First Join	Last Join	First Post	Last Post
Number	Date	Date	Date	Date
1	NA	NA	12/31/2010	7/21/2011
2	NA	NA	12/1/2010	2/23/2011
4	8/31/2009	2/28/2011	10/3/2009	12/25/2011
5	11/10/2005	12/11/2010	6/6/2008	12/11/2010
6	12/1/5/2009	8/2/2011	2/5/2009	11/14/2011
7	12/1/2010	12/20/2011	12/26/2010	7/9/2011
8	8/1/2008	12/1/2011	4/1/2009	11/1/2011
9	4/1/2011	7/1/2011	4/1/2011	7/21/2011
10	4/13/2009	3/8/2011	4/10/2010	3/7/2011
11	2/23/2007	2/25/2012	5/9/2007	2/25/2012
12	5/1/2007	6/10/2011	11/7/2007	11/9/2011
13	12/18/2004	8/1/2011	6/6/2007	7/25/2011

*Forum 3 excluded from this analysis due to limited data.

The second data point captured in this analysis is the date an individual joined the forum. The date that they joined the forum is usually posted under their username in each post to provide insight on the length they have belonged to the group. Using this data point reinforces the idea that four of the forums have an extended duration over time, as individuals have belonged to the site since 2004, 2005, and 2007 respectively. An additional four forums have had members join in 2008 and 2009, suggesting that the site has been active for several years relative to the post dates captured in this convenience sample. The remaining four sites appear to have been created more recently indicating they are not as well established as the other stolen data markets.

In total, 61.5% of our sample demonstrate an extended duration over time and may meet the criteria of a formal organization. The remaining forums may either be in their early phases of operation or are perhaps more short lived in keeping with the general duration of forums in the underground generally (Holt, 2009; Holt & Lampke, 2010; Meyer, 1989). In addition, Forum 8 was one of the two ripping forums in this sample yet had over two years of posts. This forum had limited managerial oversight, but an extended duration over time reducing its likelihood of being a formal organization. As such, this provides support for the argument that a range of markets are currently operating with varied organizational complexity (Herley & Florencio, 2010; Wehinger, 2011). Within this sample, ripping markets appear to be less organized and managed, a proportion of legitimate markets have minimal oversight, and a proportion have substantive managerial roles and an extended duration over time making them more formal organizations.

D. Social Network Analysis

Both the quantitative and qualitative analyses suggest that there are variations in the structure of the markets affecting participants. To further explore the relationships between participants, social network analysis techniques were applied to identify the network structures present between forum users. First, global level analyses were conducted and individual forum networks created to compare all networks simultaneously followed by local analyses to understand relationships around each node. For example, degree centrality can be calculated for the network of users as a whole, and as a measure by individual node where the degree centrality is "expressed as a proportion of maximum degree, which is the number of other vertices on the network" (de Nooy et al., 2005: 115). Since any member of the forum could read the threads posted but not participate, the global and local level of analysis show different properties of the networks.

At the global level, the basic network descriptions by forum are presented in Table 15, including the number of threads, users, ties between users, loops, multiple lines and loops. Individual loops maybe viewed as individual deviance or self-adds. Multiple lines indicate multiple contributions of the same user to the thread, and reflect the role of the forums in facilitating deviant exchanges between participants. The network is built from individual users as nodes, and their participation in the same thread as the tie or arc between the users. The findings suggest substantive variation in the nature of each forum. In Forums 1 and 5, users may not personally know one another due to the lack of interactions and presence of self-loops where the creator is primarily updating the post or attempting to attract attention to their thread. There are also a number of forums (e.g. Forums 7 and 8) where there is more conversation and engagement between participants. The relationships observed may be from either positive or negative reviews and commentary based on the number of ties relative to the number of threads, users and loops. These forums may demonstrate the social learning process evident in general interest hacker forums (e.g. Holt et al., 2010) and online networks (Zhang et al., 2007).

Table 15: Network Metrics By Forum

Forum	# of Threads	# of Users	# of Ties	# of loops	Multiple lines	Multiple loops	Percent in the largest component
1	55	81	49	48	0	55	18.5
2	128	160	196	101	11	25	86.25
4	144	170	210	120	103	225	50.59
5	89	88	7	77	0	9	4.54
6	48	416	295	39	0	8	58.89
7	202	157	160	68	13	134	71.98
8	590	471	350	278	121	470	55.29
9	312	650	762	286	2	26	73.69
10	60	237	392	40	85	56	60.61
11	35	66	50	33	10	85	97.01
12	71	119	95	53	3	18	62.19
13	153	293	240	136	23	127	55.63

To further explore the network structure of the forums, modified networks were created by removing loops and multiple lines to calculate the density and centrality of the forums (see Table 16; de Nooy et al., 2005). The majority of forums have low network density, suggesting that there are redundancies present that may actually facilitate network resiliency. For instance, the presence of multiple dumps sellers in a single forum may increase the redundancies present and make it difficult to disrupt the network through the removal of one or two sellers. In addition, the average degree reflects how many nodes on average can be connected through threads on the forum. For instance, an average degree of two implies that at least two ties are related on the node. Self loops are excluded to clearly reflect actual connections between different users. The removal of these loops dramatically reduces the average degree connectivity for the majority of networks. In addition to all-degree centrality, we have calculated all-degree centralization which expresses to which extent a network has a center (de Nooy et al., 2005).

Table 16: Network Density and Centrality Metrics By Forum

Forum	Original network		No loops, no multiple lines		Original network	No loops, no multiple lines	
	Network Density	Average Degree	Network Density	Average Degree	All Degree Centrality	All Degree Centrality	All Degree Centralization
1	0.016	2.568	0.008	1.210	7	7	2.456
2	0.013	4.163	0.008	2.450	8	7	35.215
4	0.023	7.741	0.007	2.471	15	16	18.482
5	0.012	2.114	0.001	0.159	11	12	1.453
6	0.002	1.644	0.002	1.418	11	13	45.007
7	0.015	4.777	0.007	2.038	17	19	8.084
8	0.006	5.266	0.002	1.512	54	57	22.843
9	0.003	3.346	0.002	2.345	10	13	26.409
10	0.010	4.835	0.007	3.308	1	5	33.630
11	0.041	5.394	0.012	1.515	19	18	5.406
12	0.012	2.840	0.007	1.597	7	10	8.850
13	0.006	3.283	0.003	1.488	13	16	18.885

In light of the low density evident across these forums, it appears that they are inefficient at the distribution of knowledge and information due to multiple redundancies or less connectivity between different subsets of the networks. The myriad data sellers and service providers demonstrate that removing a single individual from the network will have no true impact on the network as a whole. Others can easily replace that individual seller or actor, and maintain the general status of the market (see also Motoyama et al., 2012). This is particularly evident in Forum 8, which is one of the two ripping forums within this sample. There is a large user population as well as a number of isolates within each thread suggesting that there are minimal replies to any post (Herley & Florencio, 2010).

This relationship is not consistent in the other ripping forum, Forum 2, which has a smaller number posts and a high average degree. There were no consistent differences between ripping and non-ripping forums based on network measures. This may stem from the different timeframes sampled across these two forums: Forum 8 had posts starting in February 2009 while Forum 2 began in December 2010. Forum 2 may be in the initial phases of its development and have less reputation or recognition among users relative to the more established population in Forum 8. Over 80 percent of forum participants in the surveyed threads communicate with each other in Forum 2 compared to about 50 percent in Forum 8. Thus it is possible that Forum 2 may have a larger proportion of isolates over time as the negative consequences of ripping affect users (Herley & Florencio, 2010; Wehinger, 2011).

For the local level analysis, each network can be broken down into individual components. A fully connected network, for example, will have each node connected through different paths. This is not the case with the forums included in this sample, though some have a large component structure where a majority of nodes are connected through their participation in different threads (see Table 15). This structure is formed by including weak ties between participants where users participating in multiple threads form the center of these components. Weak ties might have less redundant information and form bridges between more connected components of the network (Granovetter, 1973). Three forums in this sample (Forums 7, 9, and 11) have more than 70 percent of all nodes connected through various threads. Users in these forums may be more familiar with one another and therefore operate in a more insulated network with reduced risk of compromise (see Herley & Florencio, 2010; Wehinger, 2011). Forum 5, however, has very little connectivity, and the average degree for this forum is influenced by the only component formed on this network.

At the same time, the largest forum in this sample (Forum 8) is also a ripping forum and has the lowest percentage of users in the main component. This disconnect between users may be a reflection of ripping, as individuals may create identities just to post ads and cheat users (see Herley & Florencio, 2010; Wehinger, 2011). This provides support for the assertion that ripping markets may have limited ties between users while reputable forums have greater connections between users and stronger connections which help to establish user reputations (Motoyama et al., 2011). In turn, this further reinforces the concept of multiple markets operating online to sell data with varying degrees of trust and user insularity (Herley & Florencio, 2010; Wehinger, 2011).

At the local level, we are able to calculate measures specific to individual nodes in each forum. Each network will be explored in turn, comparing economic activity by centrality and by the number of posts (the color legend in each graph is sales=green, buying=yellow, exchange= red; both buying and selling=violet; neutral=blue). The number of posts measure captures overall user activity on the forum which is in most cases larger than the total number of posts made within this sample of threads. Thus, the total number of posts becomes a proxy for the overall activity of a user in any given forum. The usernames are removed from each network in order to provide a measure of anonymity for the participants. Visualizations for Forum 3 will not, however, be presented due to the small sample size.

For Forum 1, all the users are recorded as neutral on the forum (see Figure 4a). The size of the isolates in the Figure 4a is the same, as they all have the same centrality of zero since they do not have any other nodes connecting them. The loops are included in the measures of individual centrality to identify isolates or users who increase their centrality through self-promotion and self-reference. In Figure 4b the size of these isolates differ by the number of posts they created. The number of posts and the user's centrality (a proportion of nodes incident on the node under study from all nodes on the network) are significantly correlated with each other with a Pearson correlation of 0.5(p<0.001), suggesting that an individual's node centrality is correlated with the frequency of the participant's posts.

Figure 4a: Forum 1 Economic Activity by Centrality

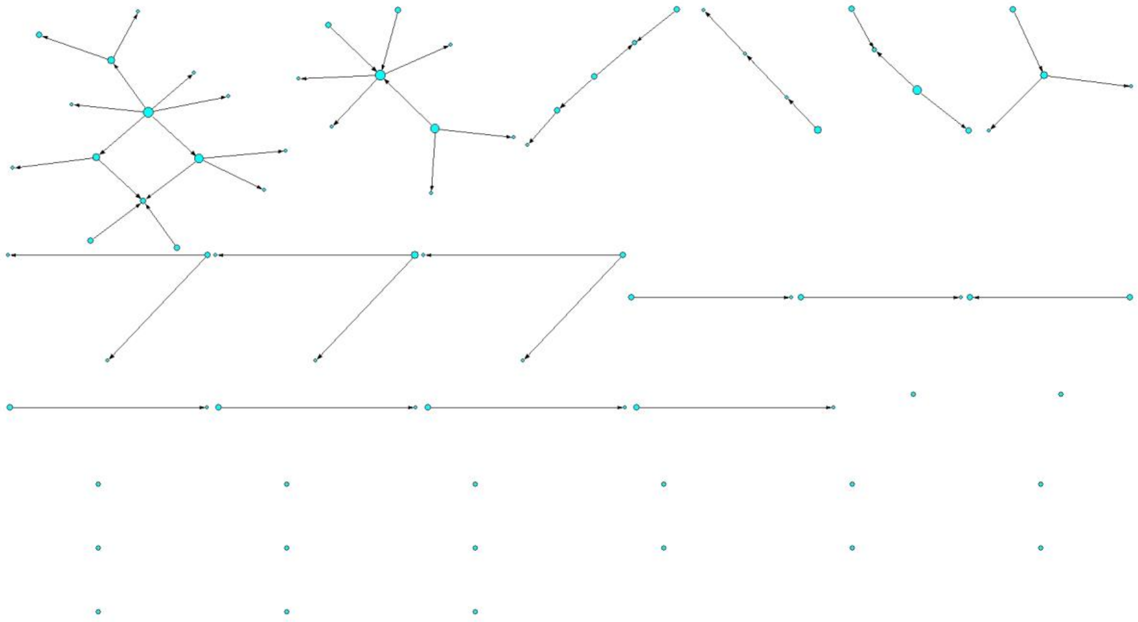
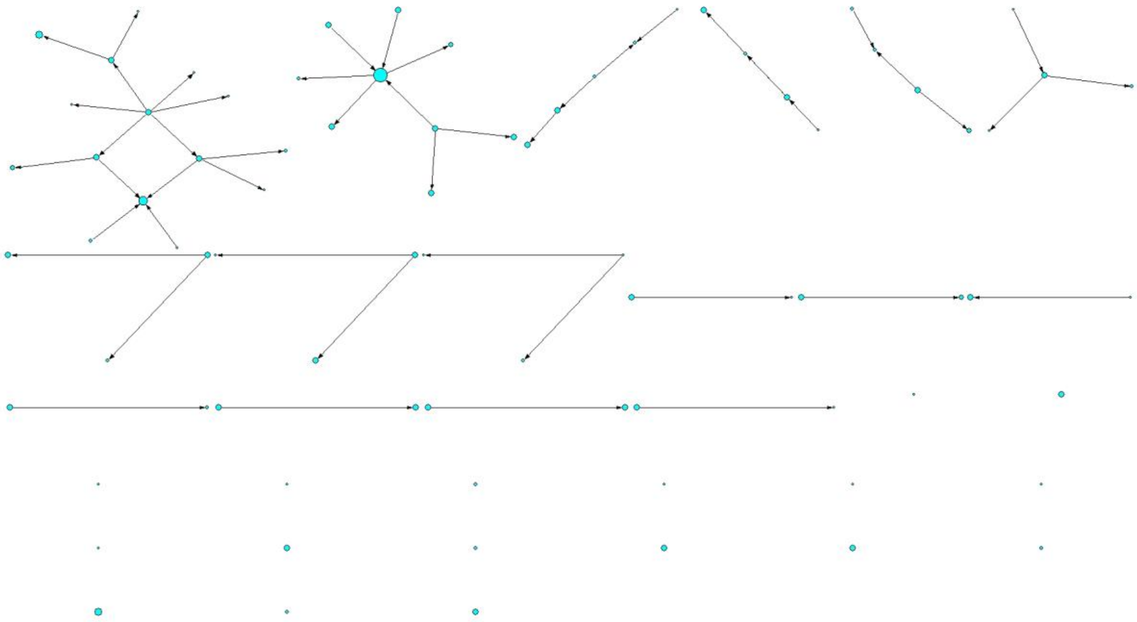


Figure 4b: Forum 1 Economic Activity by Number of Posts



The visualizations of Forum 2 suggests that sellers are more central (see Figure 5a), while the buyers and neutral users post more often (see Figure 5b). Neutral users appear to be the most important hubs for network connectivity overall, suggesting that they are key players within this forum. There is also very high correlation between the number of posts and users' centrality (0.935, p<0.001). This forum is one of the two ripping forums in this sample, which may account for this relationship as neutral users are critical to help identify who are rippers.

Figure 5a: Forum 2 Economic Activity by Centrality

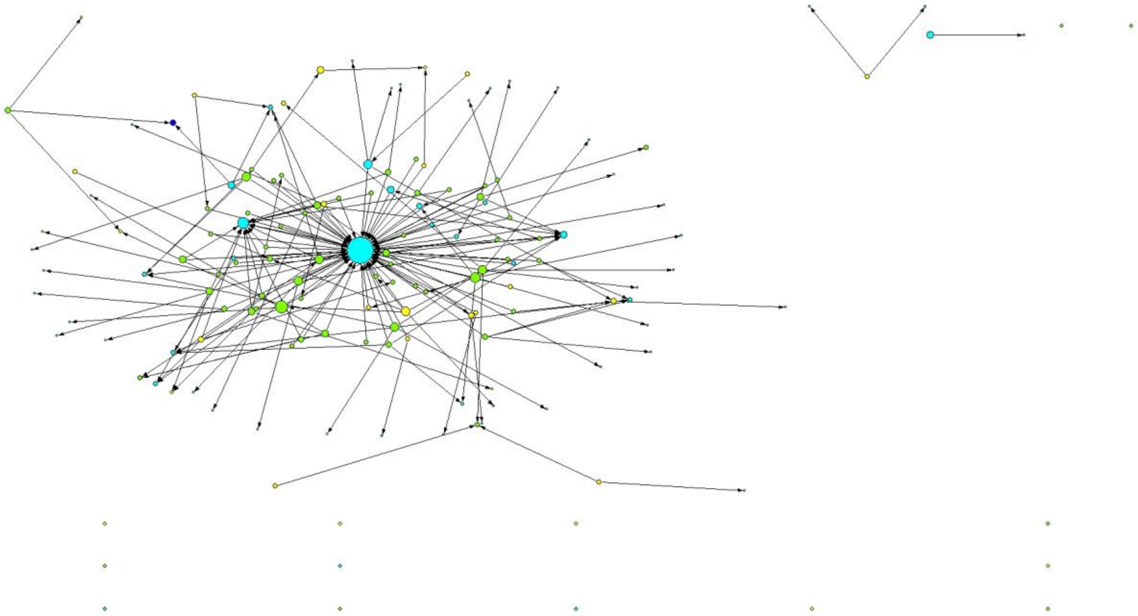
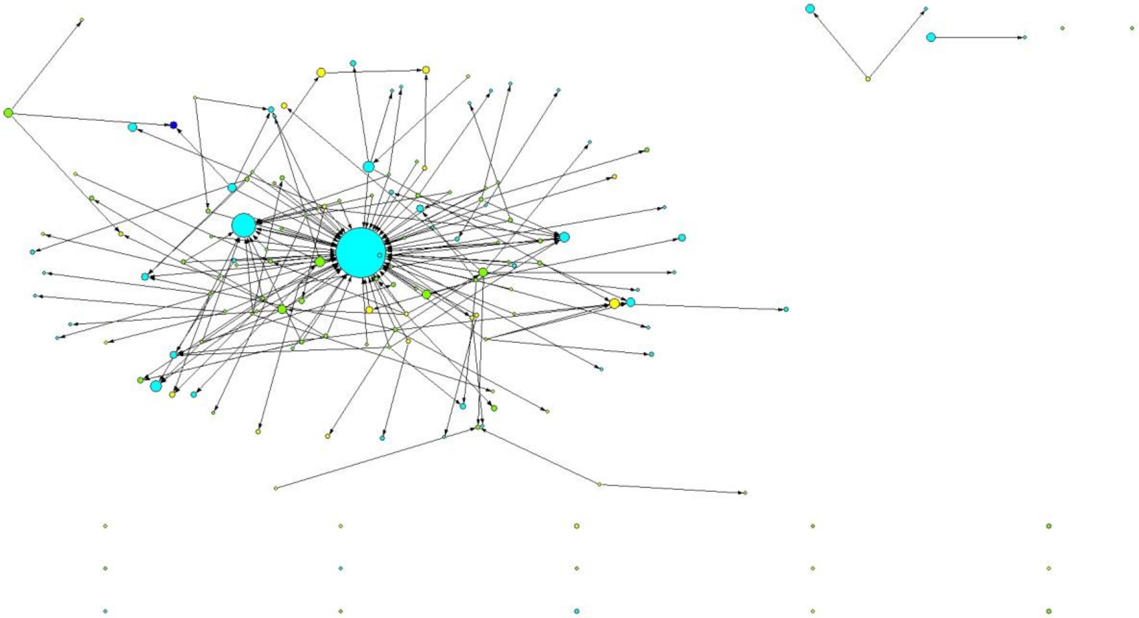


Figure 5b: Forum 2 Economic Activity by Number of Posts



Forum 4, on the other hand, shows a modest correlation coefficient between centrality and the number of posts (r=0.398; p<0.05; see Figures 6a and b). But it also underscores the fact that sellers are central to this forum despite the fact that neutral users or buyers create a larger number of posts (see also Motoyama et al., 2012; Yip et al., 2013).

Figure 6a: Forum 4 Economic Activity by Centrality

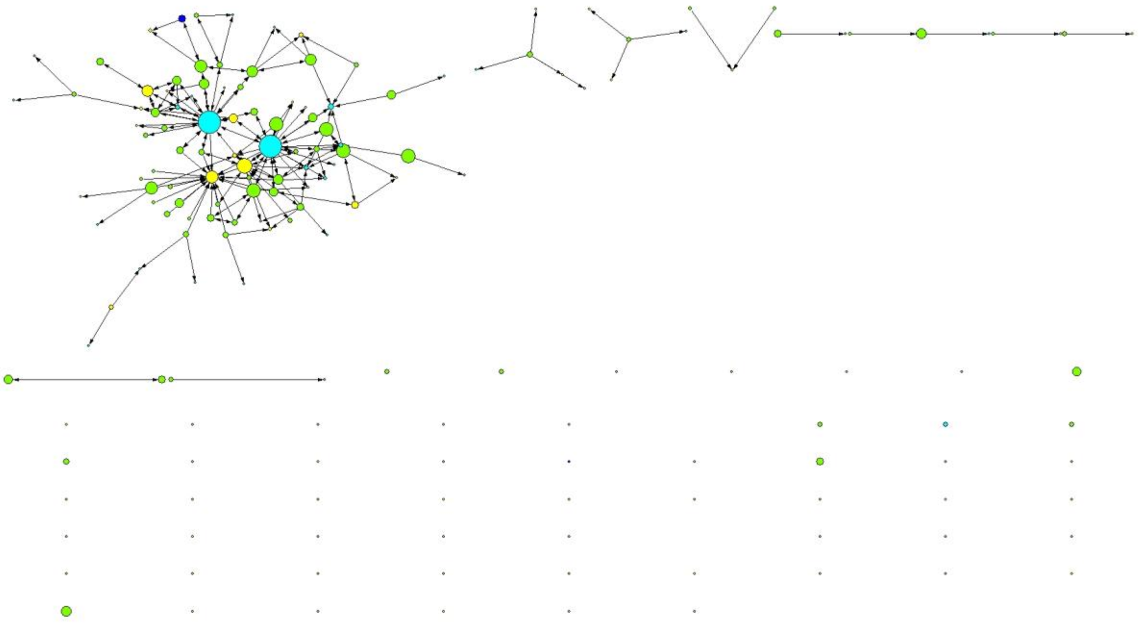
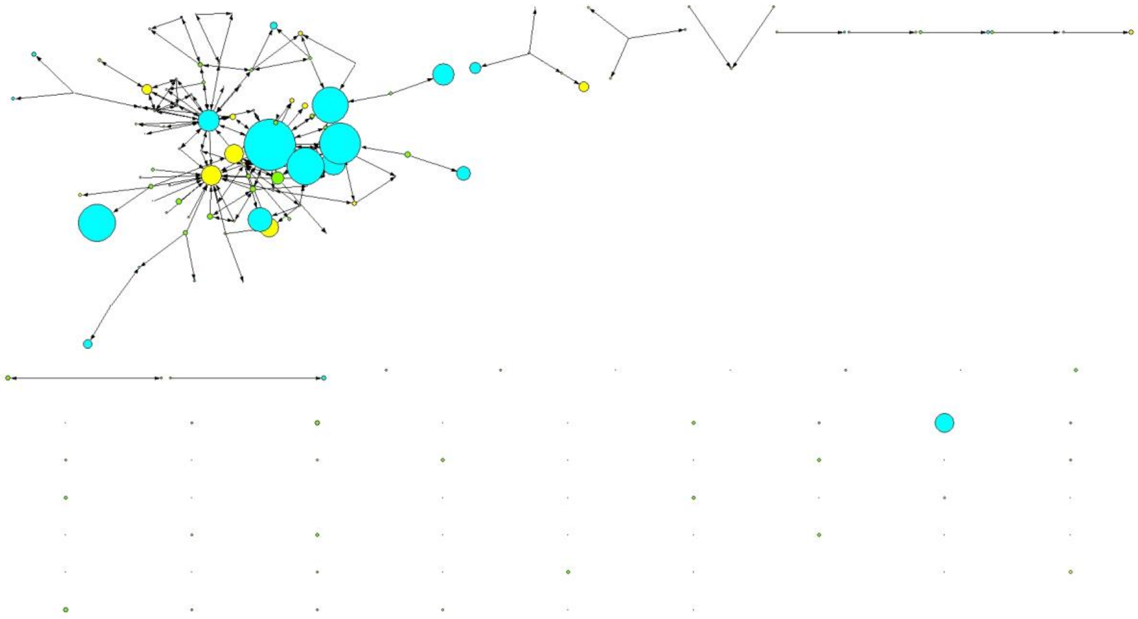


Figure 6b: Forum 4 Economic Activity by Number of Posts



The visualization of Forum 5 demonstrates that there are a large number of isolates (Figures 7a and b). There is no correlation between the number of posts and the user's centrality (r=0.06; p>0.05). Examining the number of posts further demonstrates this trend, though one of the sellers (large green isolate) becomes more visible on the network based on the number of posts (comparing Figure 7a to 7b). This might be due to the user's activity, which may have occurred in other parts of the forum not captured in this sample.

Figure 7a: Forum 5 Economic Activity by Centrality

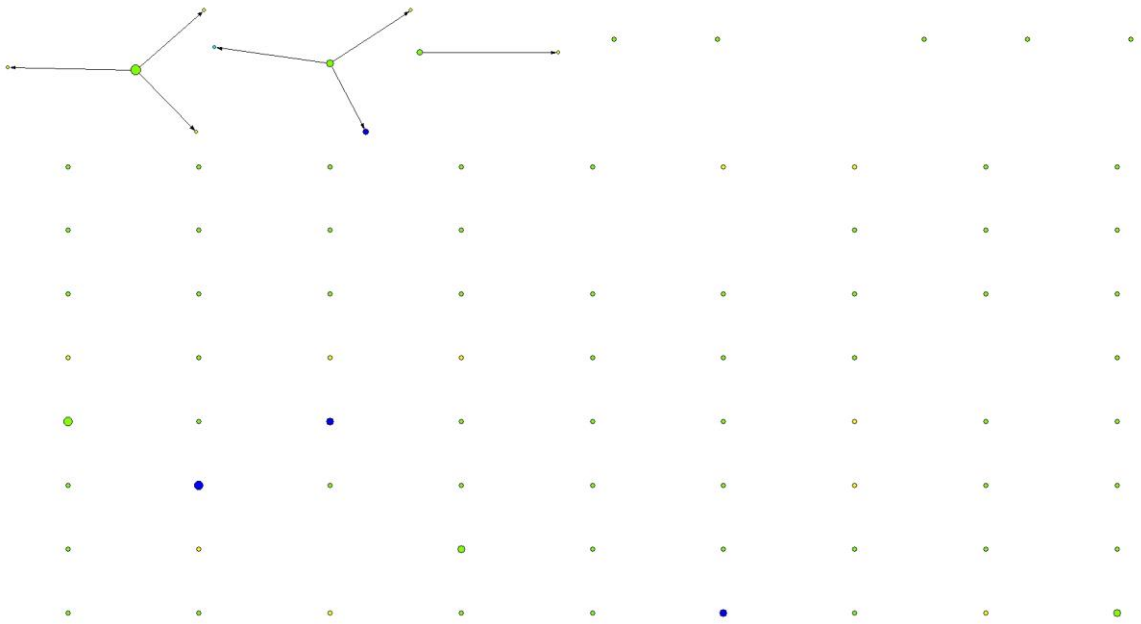
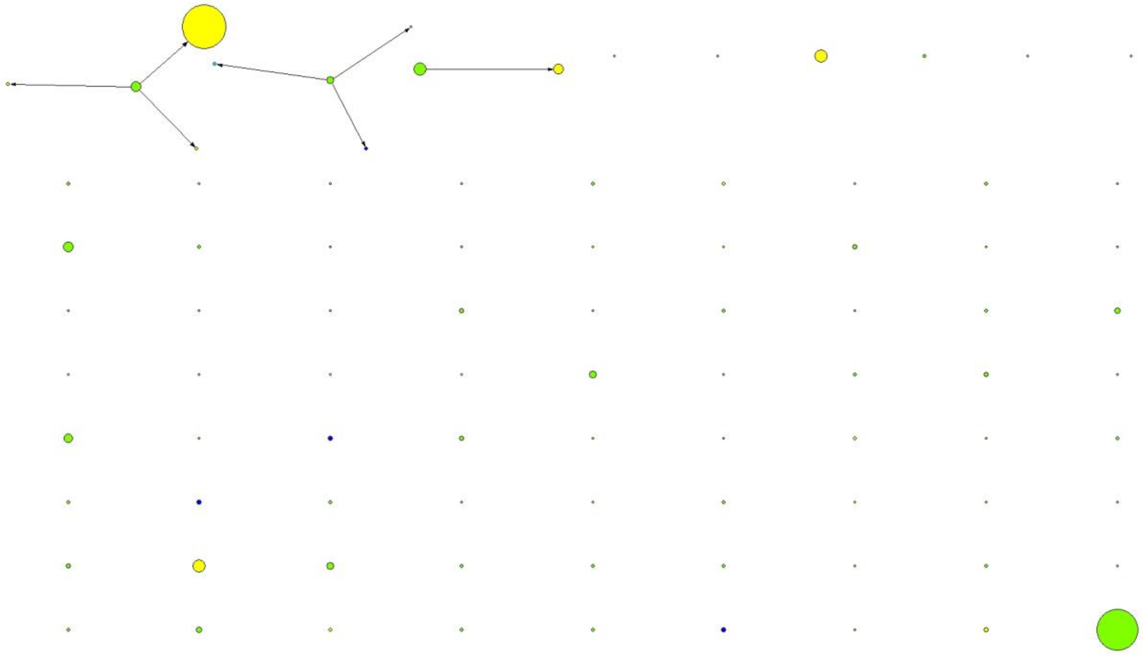


Figure 7b: Forum 5 Economic Activity by Number of Posts



Forum 6 again demonstrates that buyers post frequently, while sellers are more central in the network. There is no significant correlation between the frequency with which a user posts and the number of other nodes this user is connected to in this sample (r=0.067; p>0.05). The one user who both buys and sells product is both central and "vocal" on the network.

Figure 8a: Forum 6 Economic Activity by Centrality

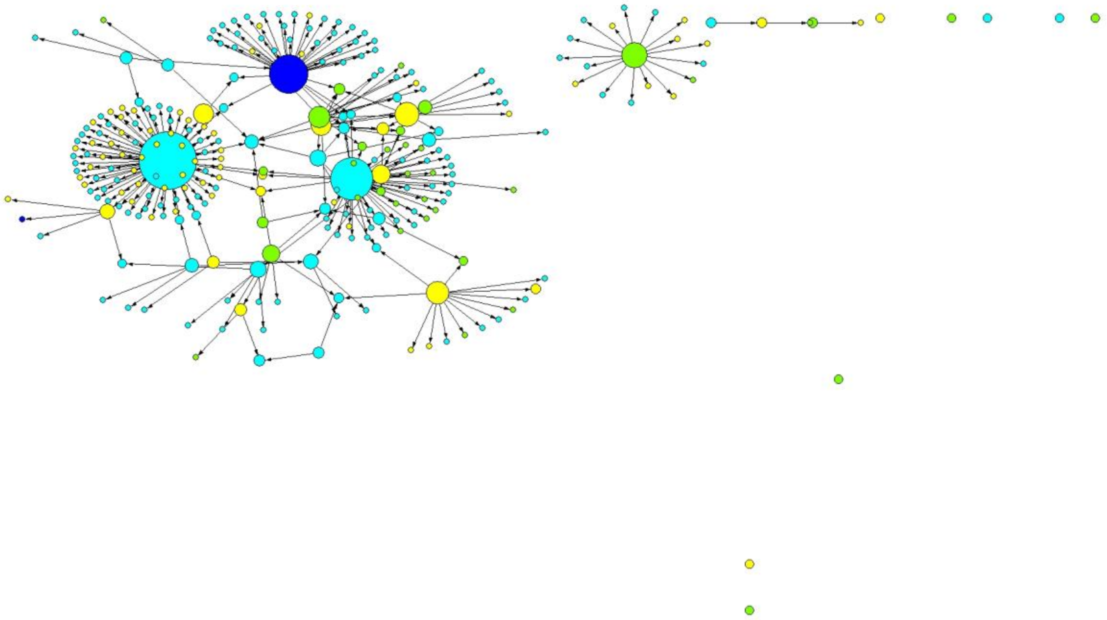
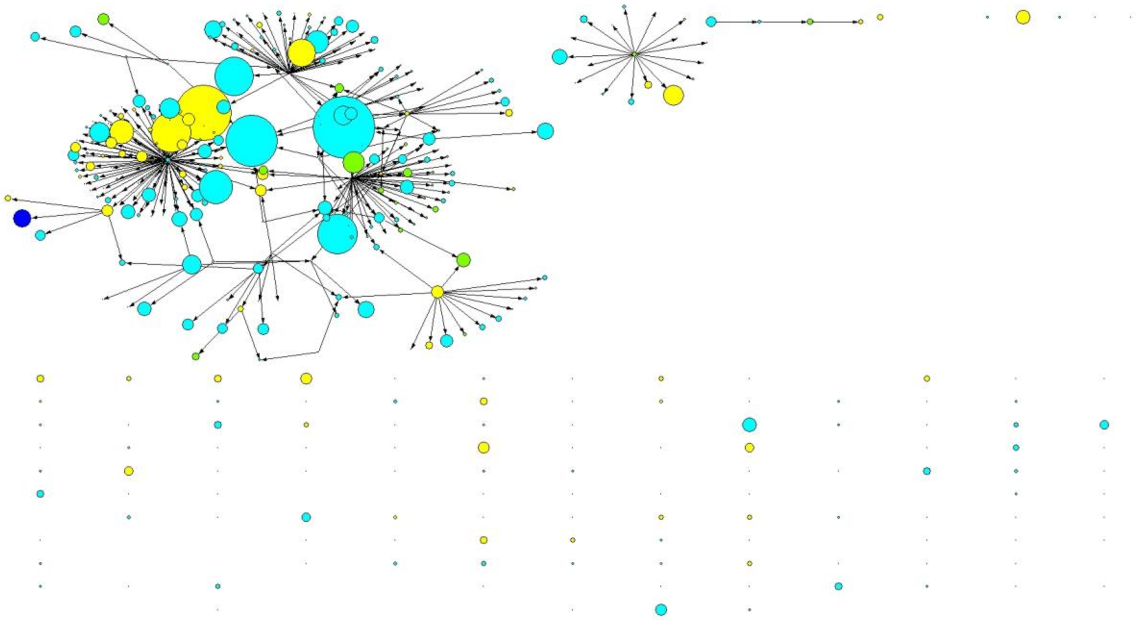


Figure 8b: Forum 6 Economic Activity by Number of Posts



Forum 7 has a smaller number of users than some of the other forums sampled, but has a larger number of threads and is a more connected network overall (see Figures 9a and b). The number of posts and users' centrality are correlated at r=0.457(p<0.001).

Figure 9a: Forum 7 Economic Activity by Centrality

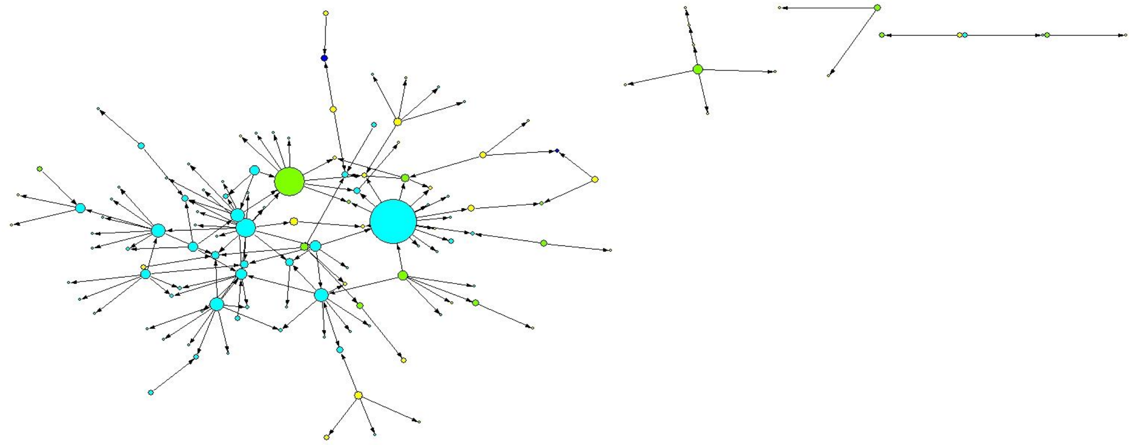
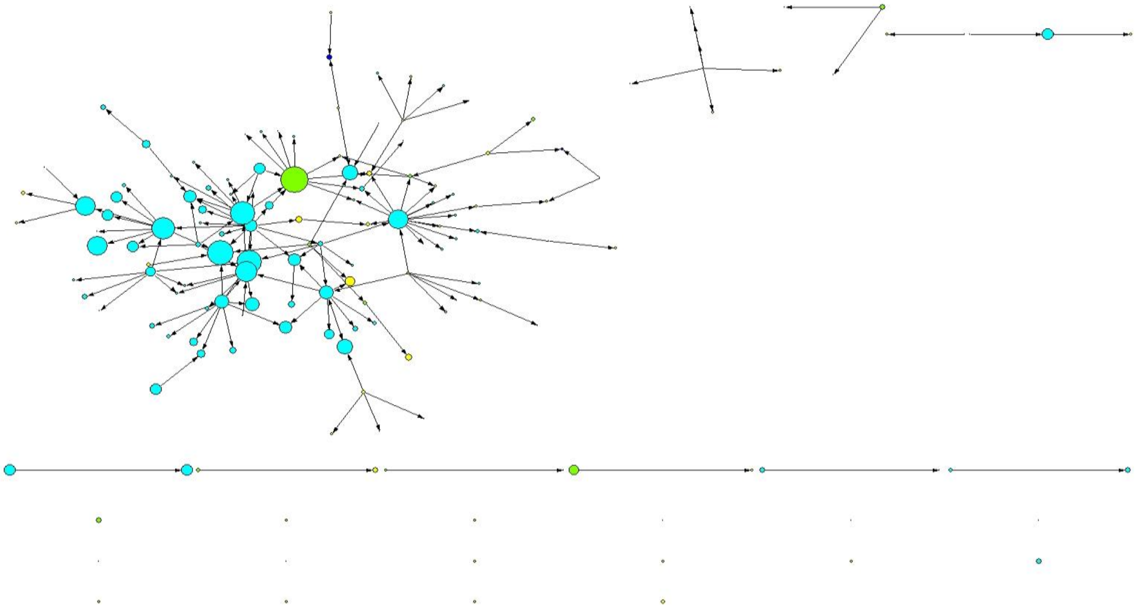


Figure 9b: Forum 7 Economic Activity by Number of Posts



The visualizations of Forum 8, one of the two ripping forums, suggest that there are a large number of sellers present (see Figures 10a and b). The sellers in this site are also more central than other users , but unlike the other ripping forum, the centrality on this forum is not highly correlated with the number of posts a user makes (r=0.15; p<0.05).

Figure 10a: Forum 8 Economic Activity by Centrality

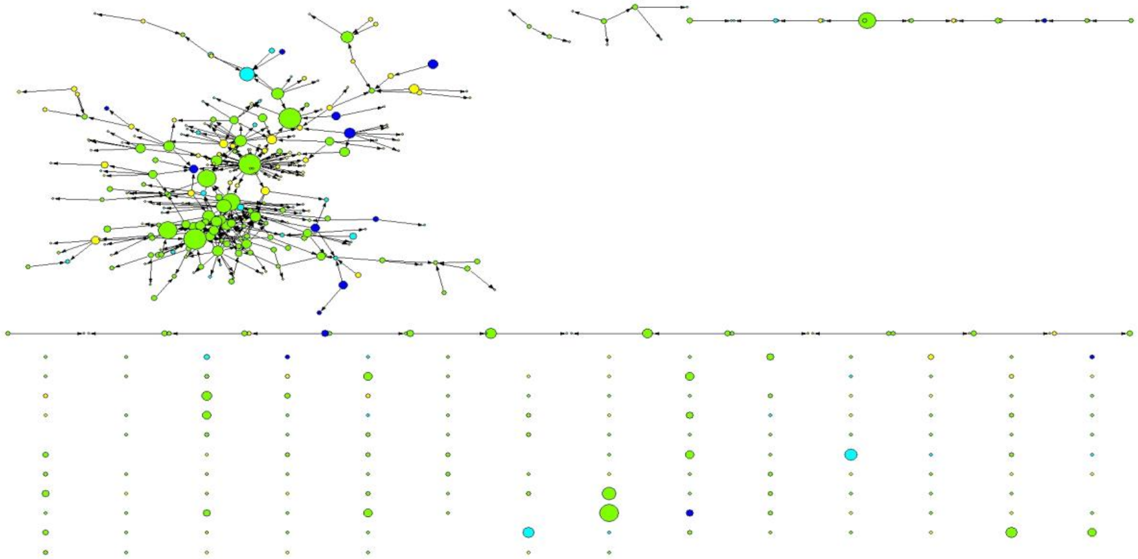
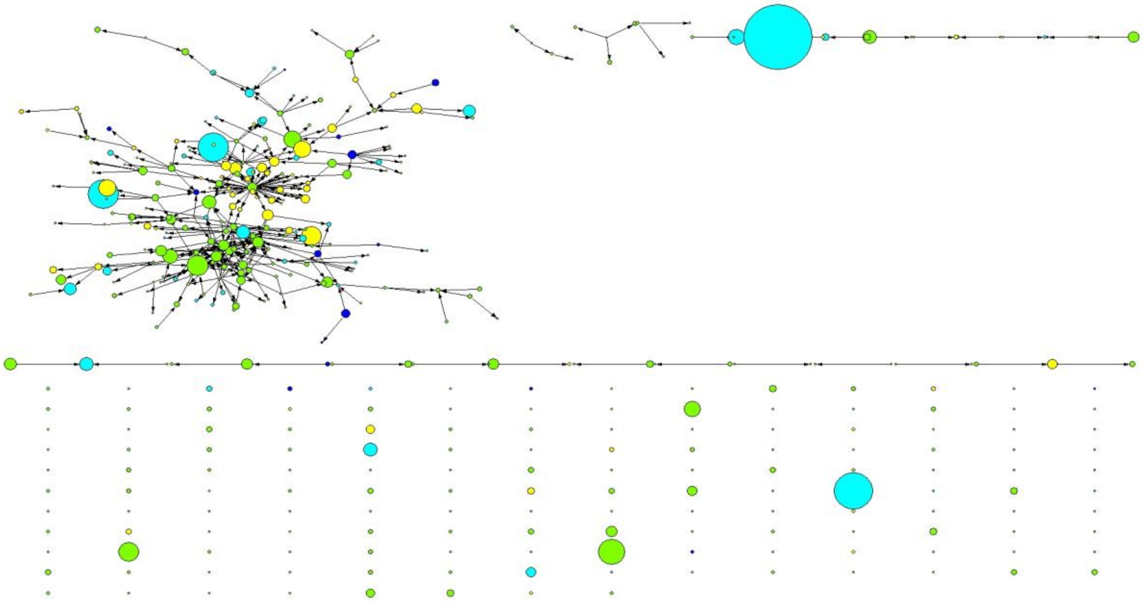


Figure 10b: Forum8 Economic activity by number of posts



The networks in Forum 9 again suggest that neutral users and buyers create a larger number of posts, while sellers are more central (see Figures 11a and b). At the same time, the correlation between the number of posts and the centrality of users is relatively high r=0.74(p<0.001).

Figure 11a: Forum 9 Economic Activity by Centrality

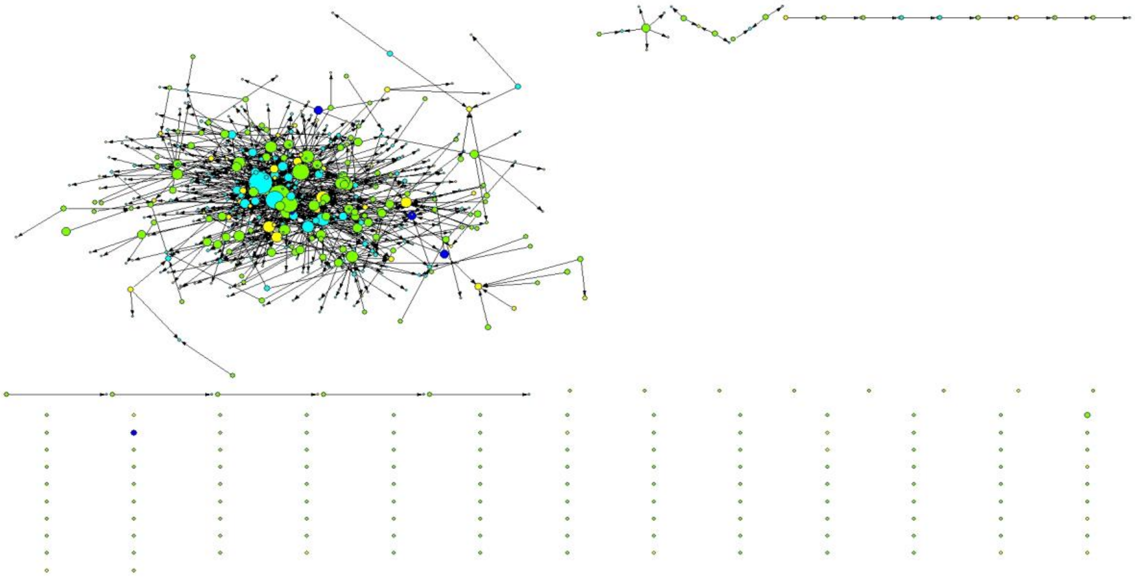
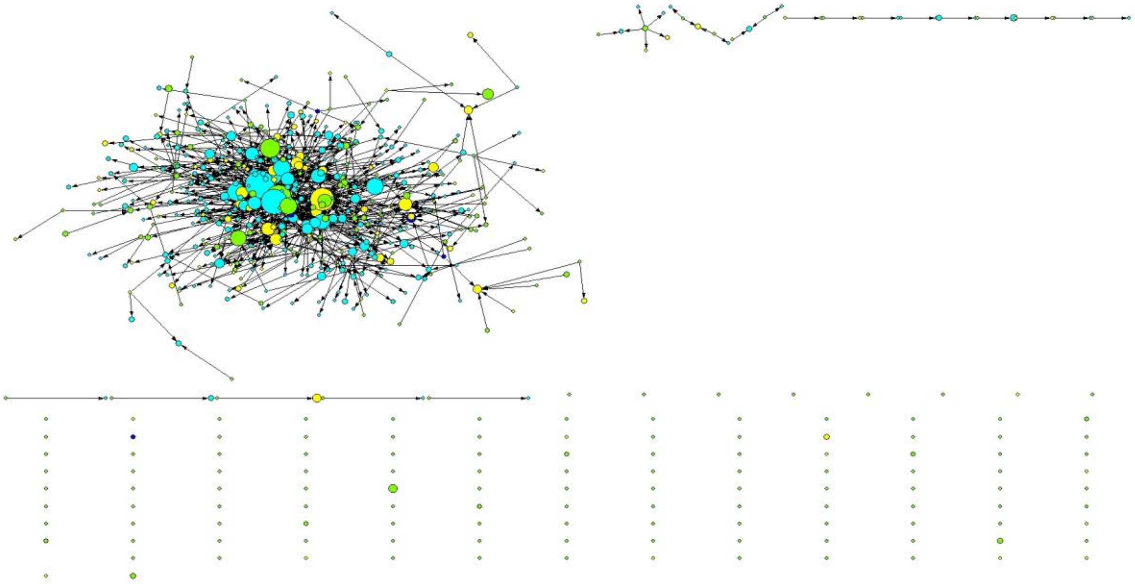


Figure 11b: Forum 9 Economic Activity by Number of Posts



Forum 10 was relatively smaller compared to the other forums in this sample as it had only 66 users. This forum, however, had a larger average degree of connectivity at 5.39. In addition, these visualizations support the general framework that all sellers are central players while neutral users and buyers make the largest proportion of posts (see Figures 12a and b). The number of posts and centrality are not correlated in this forum (r=-0.11; p=n.s.)

Figure 12a: Forum 10 Economic Activity by Centrality

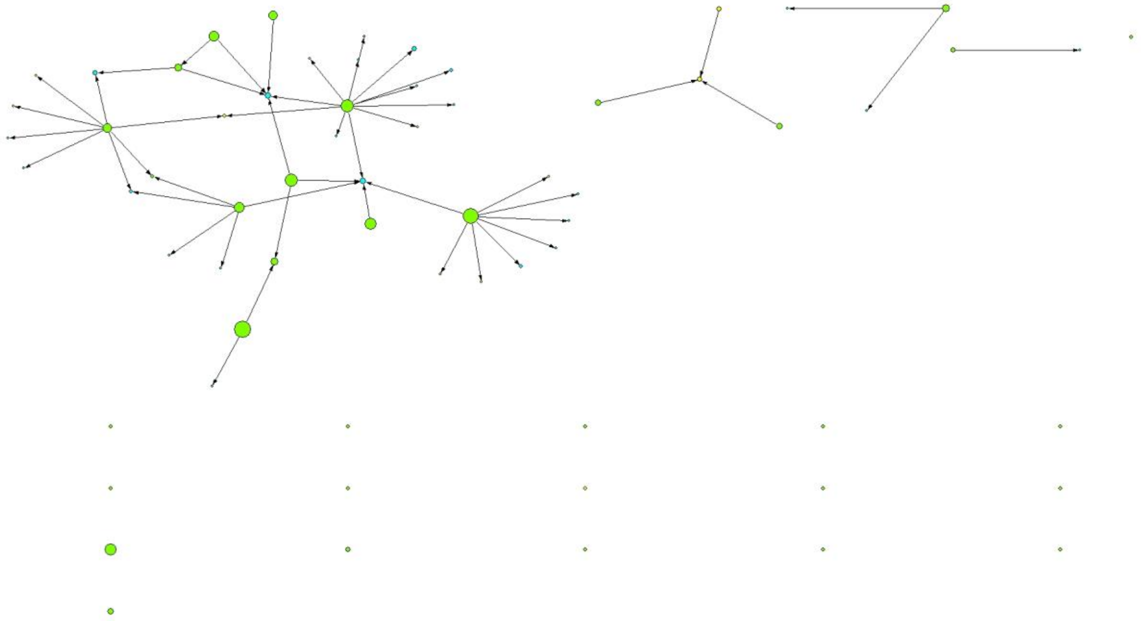
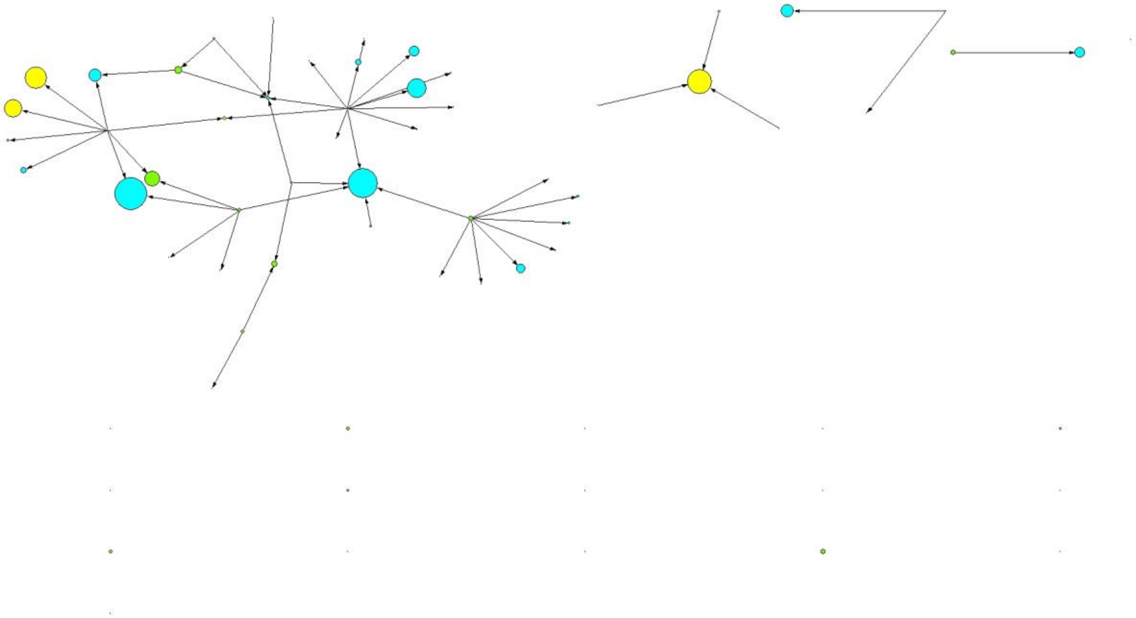


Figure 12b: Forum 10 Economic Activity by Number of Posts



Forum 11 is characterized by an emergent 'giant component' where 97% of forum users can be connected with each other through threads (see Figures 13a and b). Unlike the previous forums, all central users are neutral posters in this forum. This forum also registers a relatively rare economic activity – exchange (visualized by a red dot). This may be due to the fact that most of this forum population discussed methods and practices to facilitate the acquisition or sale of data, rather than the actual sale of products and services. Thus, this network may share more with in common with traditional hacker forums (e.g. Decary-Hetu & Dupont, 2012) rather than the stolen data markets previously presented. The correlation between users' centrality and the number of posts is r=0.46 (p<0.001).

Figure 13a: Forum 11 Economic Activity by Centrality

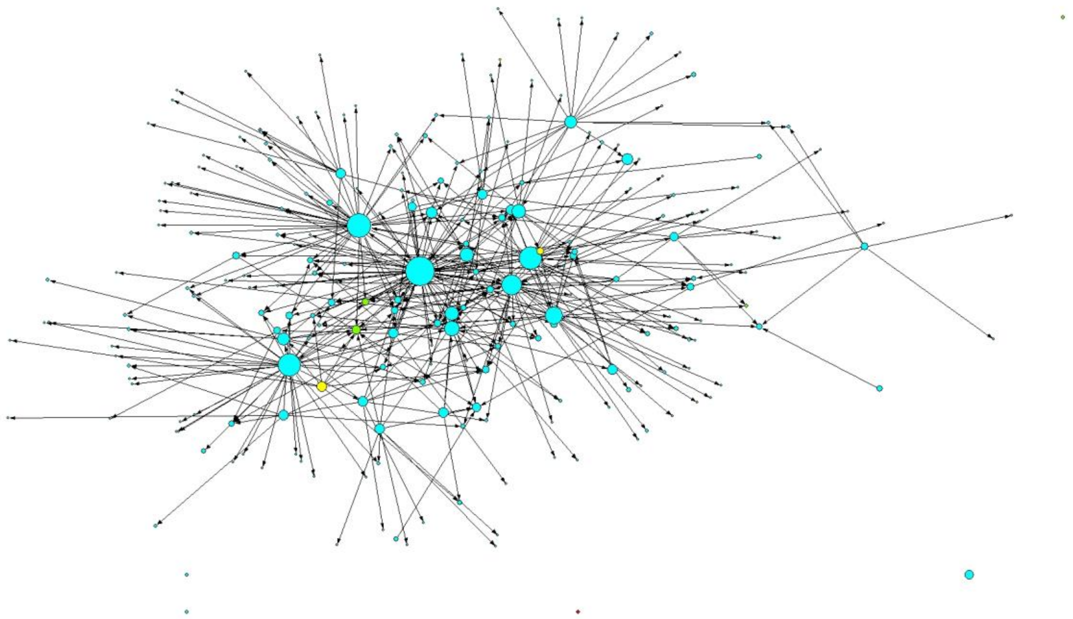
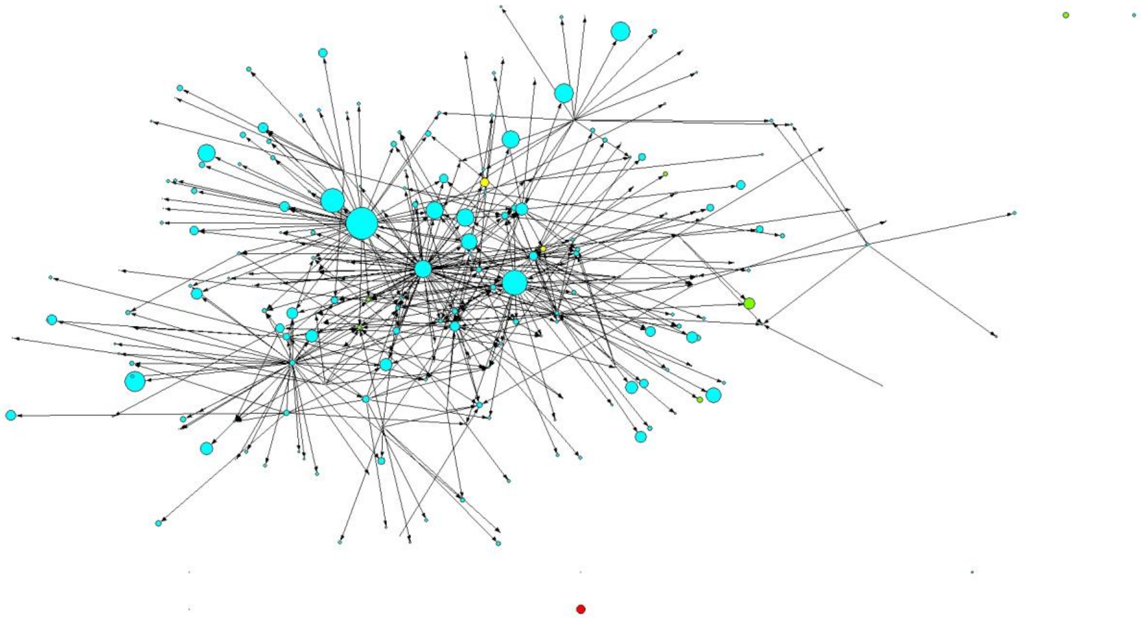


Figure 13b: Forum 11 Economic Activity by Number of Posts



Forum 12 is, however, more heavily engaged in the sale of data with a very low Pearson correlation coefficient between users' centrality and the number of posts (r=-0.009; p<0.05). In this network, two neutral users have the greatest centrality, followed by sellers, though neutral users had the largest proportion of posts overall (see Figures 14a and b).

14a: Forum 12 Economic Activity by Centrality

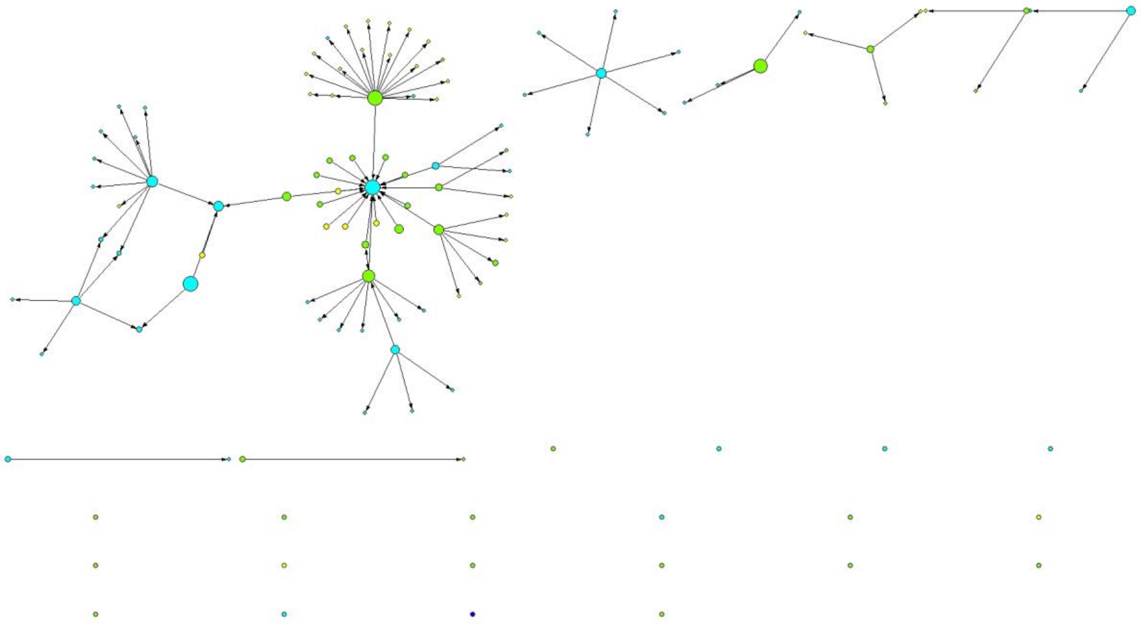
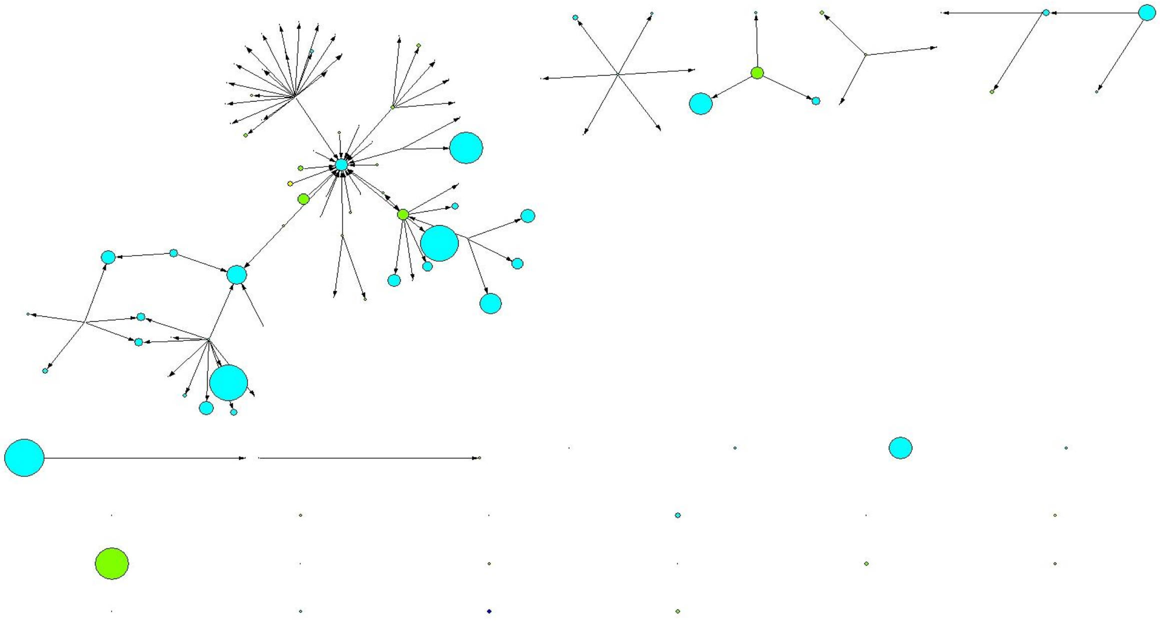


Figure 14b: Forum 12 Economic Activity by Number of Posts



The final forum in this sample, Forum 13, conforms to the larger pattern of sellers being the most central actors within each forum, with buyers and neutral players posting more frequently (see Figures 15a and b). There are also a number of isolates evident, regardless of their role in the economy of the site (r=0.04; p<0.05).

Figure 15a: Forum 13 Economic Activity by Centrality

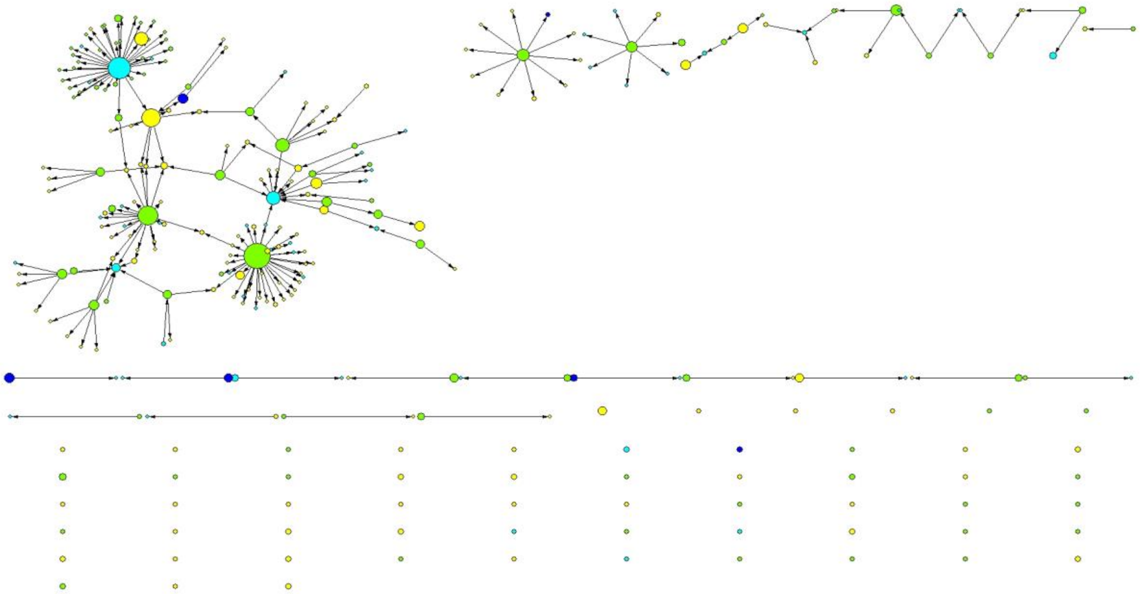
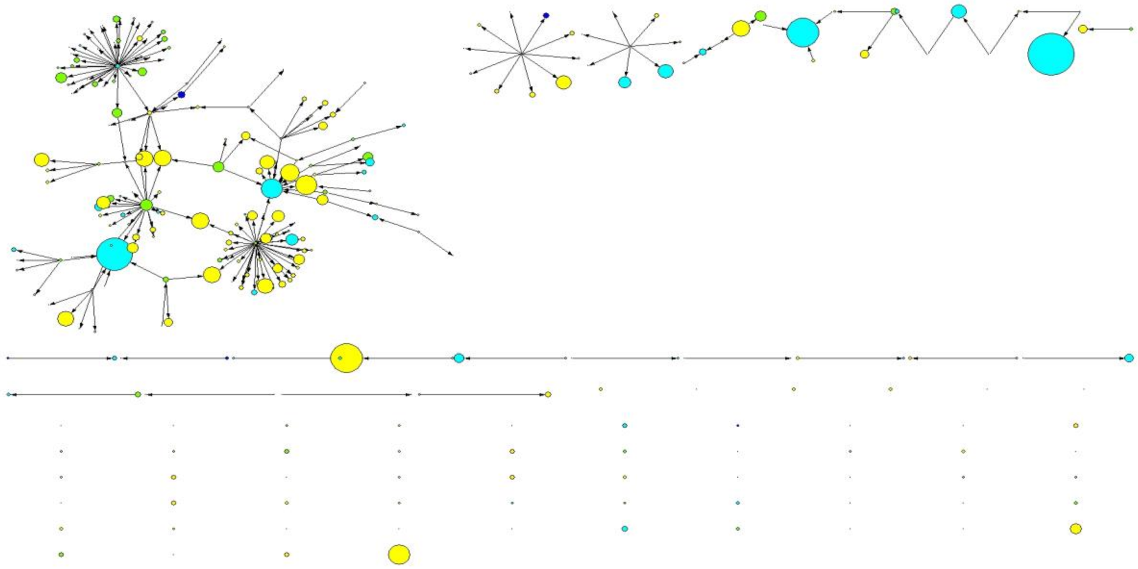


Figure 15b: Forum 13 Economic activity by number of posts



As a whole, the visualizations of centrality measures demonstrate two issues. First, neutral users in the market are a critical resource to communicate normative values to others. Sellers offer stolen data and services that engender cybercrimes which drive the market, but the users who ask questions, provide feedback, or engage others help to disseminate information about individual reputation and the general norms of the market. This provides additional support for the qualitative finding that the market is participatory in nature (see Best & Luckenbill, 1994). The lack of strong connections between participants also demonstrates that the market is collegial in nature rather than highly structured (Best & Luckenbill, 1994). Sellers are central hubs in some networks or isolates in others despite not making as many posts in keeping with the J-curve relationship observed in various research on forums (Herring, 2004; Holt, 2009; Holt & Blevins, 2007; Robinson, 1984). The low correlation between centrality and the number of posts indicates that users appear central to the network based on the proportion of other users that participate in the same thread. Thus, the frequency with which a user posts in response to a thread is vital to their centrality.

In order to further validate the network structures identified here, a series of random networks were generated using the same number of participants as in these networks with a similar low probability of making a tie of 0.01 (results not shown). The simulated networks had higher densities than the actual networks and larger core components where almost all users were connected. The key differences between the simulated networks and those identified here suggest that certain underlying mechanisms make the networks of stolen data markets relatively unconnected. Since the observed networks could not be randomly generated and are most likely a reflection of underlining dark network processes, the markets appear to be resilient to external disruption. The removal of one individual will not affect others because of the redundancy of services and user knowledge. These forums may also not be efficient as individuals must find ways to determine who is a legitimate seller or poster and minimize the risk of being ripped off. This finding provides further support for the presence of multiple markets that vary in structure and insularity from outsiders (see Herley & Florencio, 2010; Wehinger, 2011).

IV. CONCLUSIONS

A. Discussion of Findings

As computers and the Internet increasingly become the vehicle for commerce and financial management around the world, there is a need to understand the various threats consumers and corporations face from hackers and data thieves (James, 2005; Newman & Clarke, 2003; Wall, 2007). A small but growing body of research suggests that personal information acquired through various methods are sold in forums and IRC channels to individuals around the world (Chu et al., 2010; Dhanjani & Rios, 2008; Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Holz et al., 2009; Honeynet Research Alliance, 2003; Motoyama et al., 2011; Wehinger, 2011). These studies primarily focus on the products sold in the market, though there has been limited focus on the social nature of these markets and the factors that affect the structure of relationships between participants (Holt & Lampke, 2010; Motoyama et al., 2011; Yip et al., 2013). As a result, there are myriad questions that must be addressed in order to improve our knowledge of the market for stolen data and the social dynamics between actors.

Using a sample of 1,889 threads from 13 Russian and English language forums, this study utilized both quantitative and qualitative research methods to examine the economics and organization of the market for stolen data. The findings suggest that individuals interested in engaging the market can obtain virtually any resource needed to engage in financially-driven cybercrimes (see Mann & Sutton, 1998). Sellers create threads advertising their products in any given forum, and provide information about their product, pricing, payment preferences, and contact information. The majority of sellers utilized electronic payment such as WebMoney or Liberty Reserve, though a small proportion also accepted real world payments through Western Union and MoneyGram (Franklin et al., 2007; Holt & Lampke, 2010; Wehinger, 2011). A number of forums in this sample also offered escrow payment systems, where a trustworthy user in the forum acts as an intermediary to hold funds on behalf of the seller until the buyer received the product they purchased (Herley & Florencio, 2010; Holt & Lampke, 2010). All contacts took place outside of the forums, primarily through IRC, though buyers could post about their experiences with the seller in the forum. This sort of public review enabled prospective buyers to research a vendor and gain some insights into the reputability of their claims.

The findings suggest that the majority of products sold in this sample of forums are some form of illegally acquired personal financial data (see also Franklin et al., 2007; Holt & Lampke, 2010; Honeynet Research Alliance, 2003; Motoyama et al., 2011). Dumps were the most common product sold, followed by CVVs, though a proportion of sellers also offered resources to acquire virtual or real money from these accounts using cashout services or money transfers. The information sold also came from victim nations around the world, with the majority from US and European customers (Holt & Lampke, 2010). The average price for data also varied substantially, and had significant differences based on the country of origin, suggesting that the quantity of available data affected its price in the open market (see also Franklin et al., 2007; Herley & Florencio, 2011; Holt & Lampke, 2010). The pricing structure for data suggests that vendors offered their services at prices below that of the true value of the account, in much the same way as real world stolen goods vendors (Schneider, 2005; Stevenson et al., 2001). In fact, the forums may be analogous to stolen goods markets as sellers directly hawk their goods to interested parties at competitive prices (Cromwell et al., 1991, 1993; Schneider, 2005; Wright & Decker, 1994).

There were, however, substantive variations in the types of products sold based on the nature of the forum itself. Two of the forums in this sample had a large number of complaints against sellers because they either accepted payment and sent no product or gave the buyer inactive or incomplete data of no value. There are no formal dispute resolution services that a buyer can use to affect unscrupulous data sellers, making stolen data markets an ideal environment to cheat others in a similar fashion to robbing drug dealers in the real world (Cross, 2000; Jacobs, 1996, 2000; Jacobs, Topalli, & Wright, 1996). As a result, individuals used the term ripper in order to identify untrustworthy vendors and reduce others' risk of loss by calling attention to their activities (see Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010).

Comparing the distribution of products in the markets based on complaints of ripping suggests these unreliable forums sold a much larger percentage of data that could be fraudulently advertised. In fact, removing ripping forums from the sample reduced the number of advertisements for dumps by half, and almost completely eliminated CVV ads. This provides support for the notion that there are substantial differences in the behavior of markets based on the presence of rippers (Herley & Florencio, 2010; Wehinger, 2011). In fact, Herley and Florencio (2010) argued that there are multiple markets for stolen data, and those that are publicly accessible are populated with rippers and therefore less representative of the true costs for data.

In order to explore this issue further, two linear regression models were created to explore the influence of various factors on the advertised price for dumps and eBay/PayPal accounts. The findings support the idea that there are multiple markets operating with different pricing structures for products (Herley & Florenico, 2010; Wehinger, 2011). Those sellers using more organized payment systems such as Western Union had higher advertised prices, as did those whose dump data was tested by forum moderators. Escrow payments were also associated with higher prices for dumps, but lower advertised prices for eBay and PayPal data. Individuals who hijacked a thread, posting their own products in another's' advertising space, were associated with lower prices for both dumps and eBay/PayPal credentials. Dumps sold in ripping forums were also associated with lower prices, while both dumps and eBay credentials sold in Russian language forums were offered at a lower cost.

Taken as a whole, these models provide initial support for the presence of multiple markets with different pricing structures based on the quality of data and the risk of being ripped off. Disreputable markets may have lower general pricing in order to attract inexperienced buyers to make a purchase (Herley & Florencio, 2010; Wehinger, 2011). Sellers may make comments with regard to the quality of their data or suggest that they will offer services to compensate buyers who receive a poor product, but there is no way that these claims can be enforced (Herley & Florencio, 2010; Wehinger, 2011). More insulated markets may have somewhat higher costs though there is lower risk of losing money and a greater degree of trust between participants (Herley & Florencio, 2010; Wehinger, 2011). As such, buyers can engage in transactions with greater ease and acquire what they need without difficulty.

In addition to the economic conditions present in the market, this study also explored the organizational structure of participants in these forums. The findings indicate that the participants in stolen data forums operate at various stages of deviant sophistication, similar to that of gangs (Decker et al., 1998) and computer hackers (Holt, 2009). Those who sell and buy data appear to operate as colleagues within the market to facilitate the exchange of data. Individuals do not have to work with others, but the collegial environment provides access to those who can facilitate partnerships to achieve a specific goal (Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011). An individual could buy cards from one seller, and then seek out an encasher or provider who will liquidate an account. They may use these sellers again, or seek out others based on the availability of products and access to resources. Thus, the forums foster a substantive division of labor between participants based on the range of products and services available (Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011).

At the same time, the buying and selling process is peer-driven since actors can engage one another and influence action through recommendations posted in a thread. Buyers can discuss their experiences and interactions with sellers, and those who receive extremely positive feedback may be more likely to gain multiple clients over time (Holt & Lampke, 2010; Motoyama et al., 2011). Forum administrators can provide reviews of products or influence the status of a seller which may also affect their share of the market. Additionally, administrators can ban users on the basis of fraudulent claims in order to moderate user activity. These mechanisms help to reduce the risk of loss for buyers, though the relatively low barriers to enter and participate in a forum allow unscrupulous vendors to take advantage of prospective buyers (Herley & Florencio, 2010; Holt & Lampke, 2010; Motoyama et al., 2011). Individuals may ignore clear warning signs based on personal interests or needs and lose money with no formal recourse for compensation.

This study also demonstrates that these forums vary in their organizational complexity based on extended duration over time and the presence of purposive relationships between groups. Eight of the forums sampled constitute formal organizations, while the others appear to be driven by teams due to their short duration and generally limited organizational complexity. Interestingly, one of the forums with an extended duration was also a ripping forum. This suggests less reliable forums may operate for as long as more trustworthy markets, and supports the assertion that multiple markets operate at any point in time (Herley & Florencio, 2010; Wehinger, 2011). A similar dynamic has been noted in computer hacker forums, where groups that persist over time operate concurrently with short term groups of differing quality and skill (Holt, 2009; Meyer, 1989). These common organizational forms may be a reflection of the general nature of on-line communities, as well as a consequence of the common applications of hacking in the course of data theft (Holt, 2009; Holt & Kilger, 2012).

The social network analyses presented also support the findings of both the economic and qualitative organizational analysis. The visualizations and basic network characteristics suggest that the networks present between actors in the market are generally inefficient, as they are neither dense nor well connected. These conditions may actually make the networks more challenging to disrupt, as they have a greater potential to be robust and rebound from any threat to their integrity (Bakker et al., 2012). The redundancy of sellers and buyers mean that the removal of one key actor may be replaced by another, making them more resistant to external disruption, unlike traditional hierarchical structures.

The most central users across most all of these networks were sellers, though in some cases buyers and neutral players were almost equal in terms of centrality. This is sensible given that sellers drive the market and engender the creation of threads within most of these forums.

Buyers and neutral users create more posts, making them an integral part of network formation as they enable the flow of information and development of seller reputations over time.

Taken as a whole, the exploratory findings of this analysis suggest that the participants in these markets are generally collegial in nature (Best and Luckenbill, 1994). The hidden network structures identified suggest that there is a great deal of redundancy within the market and reducing the efficient flow of information between participants (see also Decary-Hetu & Dupont, 2012; Holt et al., 2012). As such, researchers must continue to explore these networks to identify any changes in their resilience and efficiency and understand when and how they may transform into more complex organizational structures.

B. Implications for Policy and Practice

As a whole, the analyses presented here demonstrate that the market for stolen data is a real threat to consumers and businesses alike. Victims from around the world can be harmed by the sale of personal information to facilitate identity theft, while financial service providers must reimburse victims for economic damages. The massive number of data sellers and the general pricing structures observed suggest that there is no easy or immediate way to disrupt or deter offenders engaged in these markets (Franklin et al., 2007; Holt & Lampke, 2010; Wehinger, 2011). Thus, there are a range of policy implications that must be considered in order to increase the efficacy of law enforcement responses and consumer protections.

To begin, there is a need to carefully consider any estimates made as to the scope of financial harm caused by the sale of stolen data. The findings of this analysis suggest that there are substantive price differences in the advertised cost for data based on the nature of the market. Economic models produced by data generated from open forums may not truly reflect the actual costs for data and the forces that affect participant behavior (Herley & Florencio, 2010). Instead, further study is needed using data generated from multiple markets that are open to the public, closed by registration, and registration only accessible to better estimate the distribution of products sold and the prospective return on investments that buyers may receive. Such a study would require cooperation with law enforcement agencies engaged in undercover operations in order to create a vetted underground identity that can pass through any background check established by forum administrators.

It is also necessary for policymakers and the research community to recognize that the participants in these forums and the markets themselves are not organized crime groups as per traditional research on the mafia, Yakuza, and other entities (see Abadinsky, 2007; Jamieson, 2000; US Department of Justice, 2008). The markets have a clear division of labor present and an interest in economic gain by leveraging weaknesses in international security protocols and systems. There is, however, no evidence that the participants attempt to engage in violent behavior or corruption in order to further their goals, nor do they have a specifically insulated leadership structure (US Department of Justice, 2008).

As a whole, the forums operate on a continuum of structure on the basis of managerial engagement to provide a space where individuals can complete transactions on a one-to-one basis. There was little evidence of a truly insular hierarchical management structure present within or across the various forums' administration. Administrators, moderators, and testers are present in certain forums but make a small number of posts and may not engage in micro-management of transactions. It is possible that the organizational behavior of participants may change over time to become more hierarchical and efficient. At present, the markets do not appear to be a form of organized cybercrime, but rather a network of international cybercriminals operating in a collegial fashion to further their individual goals.

In light of the structure and dynamics of the markets investigated, there is a need to find ways to efficiently disrupt the market for stolen data. Previous research has called for the use of slander attacks against forums, where threads are flooded with posts claiming that a seller is offering bad data or attempting to cheat customers (Franklin et al., 2007; Herley & Florencio, 2010). This sort of strategy may be effective against unregulated markets, such as open forums, as this may cause confusion among participants. More organized and regulated forums with observant administrators would, however, be able to diffuse and disrupt slander attacks shortly after they begin. The range of informal mechanisms available in structured markets make them insulated from basic disruption tactics. For instance, escrow services enable participants to have a satisfactory exchange, or engage in transactions with those who have gone through checking services. Prospective buyers could also examine advertisements and reviews posted on other sites to vet a sellers' reputation. Finally, administrators may ban those users and edit the posts of those actors who attempt to disrupt the market with false posts and information. As a result, slander attacks may only be effective against less organized forums, which may not be the most pertinent markets for law enforcement to target.

The general resiliency of the network structures observed in these forums suggest that there may be no easy or immediate way to disrupt them through external shocks to participants like slander attacks. Instead, there may be greater value in affecting the tools used by market actors to engage in commerce, including WebMoney, Liberty Reserve, and Western Union. Law enforcement agencies may be able to initiate investigations against these service providers which may, in turn, have a short term impact on the practices of stolen data vendors and buyers. Such a tactic would also equally impact less organized and organized markets because of the reliance on electronic payments in both markets. For instance, US law enforcement prosecuted the payment service e-Gold on four charges of money laundering due in part to its use by data thieves in the early 2000s (Holt & Lampke, 2010; Peretti, 2009; Surowiecki, 2013). The service provided a digital gold-backed currency that was known to be used by members of the carding group the Shadow Crew to send and receive payments for stolen data (Peretti, 2009). This investigation forced market actors to identify other payment systems and disrupted the flow of money for a short period of time.

At present, the payment processor LibertyReserve is being prosecuted in the US for its role in money laundering (Surowiecki, 2013). This was a popular payment mechanism in the forums in this sample, though the disruption took place after data collection for this study ended. Thus, it is unlikely that their activities were affected, but may cause vendors to accept other payments to offset the risk of detection. Should law enforcement agencies begin to target the other payment systems advertised in these forums, such as WebMoney, they may be able to successfully throttle the number of transactions that are completed and reduce the efficiency of the market generally. The impact of such an investigation may also affect other forms of deviance or crime that utilize these payment systems in order to complete financial transactions.

Attempts to target on-line payment systems may also engender cooperative agreements and multinational cooperation to influence money laundering generally. The lack of truly international strategic partnerships to track and hinder money laundering strategies of organized crime groups and other forms of offending has been recognized as a key weakness (Van Duyne & Levi, 2005; Unger, 2007). Thus investigating international payment services would help to increase both the transparency of investigation and increased connections between otherwise compartmentalized enforcement agencies. In fact, WebMoney's proprietor and administrative headquarters are in London which may provide a through-point to prosecute the company using existing working relationships between the US, UK, and European Union generally. Such a tactic is not a panacea for market disruption as it will only slow the flow of money while buyers and sellers adapt to different strategies. In addition, the resulting displacement of payments into alternative processing systems requires constant observation to catalogue changes in offender behavior (see Holt, 2010; Holt & Lampke, 2010).

Given the range of markets operating and the insularity evident in more organized forums, it appears that undercover investigations against single forum actors may be largely ineffectual at impacting the larger community of data sellers. Federal agencies have infiltrated several forums through participation as data buyers, or in some cases, by turning market participants into confidential informants (Poulsen, 2012). The success of these tactics depends on an implicit understanding of the formal and informal mechanisms between participants to manage relationships and transactions. This information can only be generated through constant observation of participant behaviors across multiple forums to discern differences in subcultural and market forces.

As a result, there may be greater value in using undercover identities established by law enforcement agencies to create forums to facilitate the sale of stolen data. This tactic is a divergence from traditional case building against single individuals, though it may be necessary as investigations against lone actors and agents in illicit drug and prostitution markets in the real world are largely unsuccessful (Harocopos & Hough, 2005; Jacobs, 1996). Instead, creating forums would allow for the implementation of intelligence led policing strategies that allow agencies to track the behavior of participants, identify key buyers and sellers, and build cases against entire networks of individuals (see also Poulsen, 2012). This technique was successfully employed by law enforcement during the "Dark Market" case where an undercover operative established a forum and eventually pursued prosecutions against the buyers and sellers engaged in the site (Poulsen, 2012). Not only would such an investigative strategy ensure sufficient evidence against the participants, but also create massive distrust between market actors due to difficulty in determining whether an actor was legitimate or a member of law enforcement. In turn, this would affect both the supply and demand side of the market in ripping and non-ripping forums alike.

There is also a need for additional support resources within federal law enforcement agencies in order to facilitate these investigative strategies. For example, the Russian language forums examined here require some degree of linguistic and technical competency in order to effectively understand the content of these ads and identify the relationships between participants. Thus, linguists and support analysts are needed to ensure cases can be effectively built and established over time. Resource allocations are also needed to support the computer and communications necessary to properly develop and host websites and materials related to undercover identities in various markets. Thus, there is a need to increase the financial allocations to the Federal Bureau of Investigation, Secret Service, the Department of Treasury, the Department of Homeland Security and other federal agencies to more effectively combat the problem of cybercrime. In turn, this may increase the efficacy of the federal response to financially-driven cybercrimes and the networks of actors that compose the market for stolen data.

There is also a need for careful revision and adjustment of cooperative agreements to facilitate the international investigation and prosecution of data thieves (Brenner 2008). The findings of this study demonstrate that participants are compromising banks, businesses, and citizens in the US and European Union. The forums themselves are hosted around the world, and the participants may be native to the Russian Federation or Russian speakers living abroad. As a result, it is vital that law enforcement agencies find ways to improve existing extradition treaties and cooperative frameworks to ensure that responsible actors may be detected and brought to justice (Brenner, 2008).

There is also a need for improved awareness of the risks of electronic identity theft among home computer users who do not necessarily have a strong grasp of basic computer security principles (Holt & Lampke, 2010; James, 2005; Newman & Clarke, 2003; Wall 2007). The risk for data theft may stem from individual behaviors such as responding to a phishing email (James, 2005) or having an active malware infection on one's home computer (Holt & Turner, 2010), or even downloading a rogue banking application for a smart phone or tablet. Users who are cognizant of these risks may still, however, be affected through a mass data breach that affects millions of card holders (Peretti, 2009). Since there is no single way to reduce individual risk of harm, there is a need for public awareness campaigns to promote basic computer security principals and vigilance against identity theft. Consumers who understand the potential harm that can result from responding to unsolicited emails, clicking on suspicious web links, and the need to run anti-virus and security tools may decrease their risk of victimization (Bossler & Holt, 2009; Holt & Turner, 2010).

These campaigns should promote the need to regularly check bank and credit card statements for suspect charges and avoid making purchases through online vendors with no security features in place to protect personal information. Educating consumers on how to identify fraudulent email requests, or phishing messages, along with untrustworthy retail sites online may also help to reduce their risk of victimization (James, 2005; Ngo and Patternoster, 2011). In addition, basic computer security protocols, such as the use of anti-virus software and regular system updates, and safe web surfing habits must be discussed, and demonstrated so they can be implemented at home. Promoting mobile security is also useful so that individuals understand that personal data and financial applications on phones must be as carefully managed as their personal computers (Symantec, 2013). Such information could be rolled out effectively through the Consumer Financial Protection Bureau, Internet Crime Complaint Center, and Federal Trade Commission, and implemented locally through financial institutions on-line and in-person through schools and libraries where individuals of various income levels use the internet and computers generally (Zickuhr, Rainie, Purcell, & Duggan, 2013). Such initiatives could be launched bi-annually, particularly during April's tax season and October's Cyber Security Month to promote general security and diminish the scope of compromises.

To further promote security and increase corporate responsibility in the event of large scale data breaches that are beyond consumer control, there is a need for increased adoption of data breach legislation. Currently, 46 states require that companies disclose any loss of sensitive personal information to consumers in the event of a security breach (National Conference, 2012). It is often difficult to determine the true scope of these beaches and determine how many customers actually face economic harm as a result of any such incident. Thus, greater transparency is needed on the part of both corporations and financial institutes to disclose the true number of customers affected and to what degree in as timely a fashion as is possible in order to reduce the risk of customer loss and economic harm generally.

C. Implications for Further Research

Though this study provides initial insights into the economics, organization, and network structure of stolen data markets, there is a need for substantial data collection and research to address the limitations of this work. Specifically, this study demonstrated the scope of personal information sold, and demonstrated some differences in the advertised price for data based on the type of forum where the ad is placed. The advertised price may not, however, be the true amount an individual pays for data given that all transactions took place outside of the forums (Franklin et al., 2007; Holt & Lampke, 2010; Holz et al., 2009; Honeynet Research Alliance, 2003; Motoyama et al., 2011; Wehinger, 2011). The only way to acquire such information is to capture the personal exchanges between participants in the markets, or engage in covert transactions with prospective sellers which create ethical challenges for academic researchers (see Holt, 2010; Markham, 2011).

As such, there is a need for greater collaboration with both law enforcement and financial institutions in order to gather data that can provide more accurate economic models to estimate the scope of harm and return on investment for data thieves. For instance, the ability to analyze communications between actors that take place either through private messaging systems within the forum or ICQ would be invaluable as this is where the majority of negotiations appear to take place (Franklin et al., 2007; Motoyama et al., 2011; Yip et al., 2013). The exploratory networks identified here suggest that neutral players are critical to the formation of networks, while sellers are more central to the network as a whole. Collecting information on the number of contacts and transactions made outside of the forum may directly affect the centrality of a user, and change the shape of the network entirely. Thus, further study is needed to determine any similarities in the networks evident in and out of the forums.

Capturing data on the number of cardholders impacted in any breach and the total number of fraudulent or disputed charges brought as a result of the sale of data would also prove invaluable. Such information is only available from banks and financial institutions, and would allow researchers to provide more accurate estimates of the financial losses victims incur as a consequence of their information being sold. Also, researchers could calculate the return on investments that participants in these markets receive from purchased data relative to its costs (Stevenson et al., 2001). This information would be essential to improve our understanding of the economy of stolen data markets overall.

Research is also needed to better understand the impact that the social structure of the market has on actor behavior and economics generally. These findings support the idea that there are multiple markets operating, such that those with more organized operational spaces through forum management and customer engagement may have differential pricing structures (Herley & Florencio 2010; Wehinger, 2011). This analysis, however, utilized data from open and registration-only forums which may be different from that of closed and vetted membership forums that may be more insulated from outsiders. In addition, the limited presence of administrators within the various forums analyzed made it difficult to determine their position within the network structure of the market. This may be a function of the forums sampled, as they may not be as organized as other markets within the larger cybercrime underground. Future research must find ways to access these forums in order to improve our knowledge of the tiered structure of stolen data markets and the practices of actors.

There is also a need to examine what factors shape individual reputation over time. The number of vendors offering the same products across multiple forums suggests that it may be difficult to differentiate oneself from other players in the market and garner a proportion of sales (Holt & Lampke, 2010; Motoyama et al., 2011). Thus, quantitative and qualitative assessments of the factors affecting individual trust and reputation in the market may be invaluable to determining who are key players across large networks of data sellers. In turn, these studies may engender more effective investigations against market actors.

Further research is also needed with larger data sets to understand the way that network structures change over time. While some of these forums operated over several years of time, the use of single sub-forums does not provide further context for an actor's evolution through participation in multiple sites. Their participation in multiple threads can be determined, though there may be missing data involving encounters with other individuals in separate sections of the site. Capturing multiple sub-forums within a site and tracking users over time would allow for the development of complex network models of change and assessments of network centrality and density. In turn, this would allow for the identification of key points in the evolution of a forum from legitimate to ripping, or vice versa.

Research is also needed to further explore and refine the Best and Luckenbill (1994) framework of social organization. Their continuum of organizational sophistication allowed for differentiation between forms of organization based on the relationships of actors within the market for stolen data. Not all facets of actor organization fit within Best and Luckenbill's (1994) ideal types. For example, the categorization of web forums is complicated by their two-population composition of users and moderators who interact in unique ways (see also Holt, 2009). Thus research is needed to clarify and operationalize the concepts that structure Best and Luckenbill's (1994) classification scheme with an emphasis on virtual relationships. Such clarification is critical to identify the social relationships between deviants and criminals on and off-line (see Adler & Adler, 2006; Holt, 2009).

In addition, the majority of the networks explored here appear to be largely inefficient, but also resilient to external shocks. This structure may change over time, particularly as a consequence of researcher and law enforcement interventions that attempt to disrupt the markets (see Chu et al., 2010; Holt 2010). Longitudinal research utilizing surreptitious data collection strategies in various forums could improve our knowledge of key points in the evolution of the market over time. In fact, this may enable researchers to identify if and when markets begin to transition from collegial structures to more organized and efficient marketplaces. Such information is vital to understand the nature of stolen data markets and their role in cybercrime and fraud globally.

V. References

Abadinsky, H. (2012). *Organized Crime, 10th edition.* New York: Cengage Learning.

Adler, P. A., & Adler, P. (2006). Self-injurers as loners: The social organization of solitary deviance. *Deviant Behavior*, *26,* 345-378.

Allison, S. F.H., Schuck, A.M., & Learsch, K.M. (2005). Exploring the crime of identity theft: prevalence, clearance rates, and victim/offender characteristics. *Journal of Criminal Justice*, *33,* 19-29.

Anti-Phishing Working Group. (2012). *APWG Phishing activity trends report, 2nd quarter 2012*. http://docs.apwg.org/reports/apwg_trends_report_q2_2012.pdf.

Aspers, P. (2011). *Markets.* Cambridge, Polity Press.

Bakker, R. M., Raab, J., & Milward, H. B., (2012). A Preliminary Theory of Network Resilience. *Journal of Policy Analysis and Management*, *31,* 33-62.

Best, J., & Luckenbill, D.F. (1994). *Organizing Deviance, 2nd edition*. New Jersey: Prentice Hall.

Bossler, A. M., & Holt, T. J. (2009). On-line activities, guardianship, and malware infection: An examination of routine activities theory. *International Journal of Cyber Criminology, 3,* 400-420.

Brenner, S. W. (2008). *Cyberthreats: The emerging fault lines of the nation state.* New York: Oxford University Press.

Bryant, A., & Charmaz, K. (2010). *The Sage Handbook of Grounded Theory.* Thousand Oaks, CA: Sage.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis.* Thousand Oaks, CA: Sage.

Chu, B., Holt, T. J., & Ahn, G. J. (2010). *Examining the Creation, Distribution, and Function of Malware On-Line.* Technical Report for National Institute of Justice. NIJ Grant No. 2007-IJ-CX-0018. [Online] Available at http://www.ncjrs.gov/pdffiles1/nij/grants/230112.pdf

Corbin, J. & Strauss, A. (1990). Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. *Qualitative Sociology*, *13,* 3-2.

Corbin, J., & Strauss, A. (2007). *Basics of Doing Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* Thousand Oaks, CA: Sage.

Cromwell, P. F., Olson, J. N., & Avary, D. W. (1991). Breaking and Entering: An ethnographic analysis of burglary. *Studies in Crime, Law, and Justice*, 8. Newbury Park: Sage.

Cromwell, P. F., Olson, J. N., & Avary, D. W. (1993). Who buys stolen property? A new look at criminal receiving. *Journal of Crime and Justice*, *16,* 75-95.

Cross, J. C. (2000). Passing the buck: Risk avoidance and risk management in the illegal/ informal drug trade. *International Journal of Sociology and Social Policy,* 20, 68-94.

de Nooy, W., Mrvar, A., & Batagelj, V. (2005). *Exploratory social network analysis with Pajek.* Cambridge: Cambridge University Press.

Decary-Hetu, D., & Dupont, B. (2012). The Social Network of Hackers. *Global Crime, 13,* 160–175.

Decker, S. H., Bynum, T., & Weisel, D. (1998). A Tale of Two Cities: Gangs as Organized Crime Groups. In J. Miller, C. L. Maxson, & M. W. Klein (Eds.) *The modern gang reader* (pp. 73-93). Los Angeles, CA: Roxbury Publishing Co.

Dhanjani, N., & Rios, B. (2008). Bad sushi: Beating phishers at their own game." Presented at the Annual Blackhat Meetings, Las Vegas, Nevada.

Dunn, J. E. (2012). Russia cybercrime market doubles in 2011, says report. *IT World Today*. [Online] Available at http://www.itworld.com/security/272448/russia-cybercrime-market-doubles-2011-says-report

Franklin, J., Paxson, V., Perrig, A., & Savage, S. (2007). An Inquiry into the nature and cause of the wealth of internet miscreants. Paper presented at CCS07, October 29-November 2, 2007 in Alexandria, VA.

Furnell, S. (2002). *Cybercrime: Vandalizing the Information Society.* Boston, MA: Addison-Wesley.

Goodin, D. (2007). TJX breach was twice as big as admitted, banks say. *The Register.* [Online] Available at http://www.theregister.co.uk/2007/10/24/tjx_breach_estimate_grows/

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology, 78,* 1360-1380.

Harocopos, A., & Hough, M. (2005). Drug Dealing in Open-Air Markets. *Problem oriented guides for police: Response guide series (31).* Washington, DC: U.S. Department of Justice, Office of Community Oriented Policing Services.

Herley, C. & Florencio, D. (2010). Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy. In T. Moor, D. J. Pym, & C. Ioannidis, (Eds.) *Economics of Information Security and Privacy*, (pp. 35-53). New York: Springer.

Herring, S. C. (2004). Slouching toward the ordinary: Current trends in computer-mediated communication. *New Media & Society*, *6(1),* 26-36.

Holt, T. J. (2007). Subcultural evolution? Examining the influence of on- and off-line experiences on deviant subcultures. *Deviant Behavior, 28,* 171-198.

Holt, T. J. (2009). Lone hacks or group cracks: Examining the social organization of computer hackers. In F. Smalleger & M. Pittaro (Eds.), *Crimes of the Internet* (pp. 336-355). Upper Saddle River, NJ: Pearson Prentice Hall.

Holt, T. J. (2010). Exploring strategies for qualitative criminological and criminal justice inquiryusing on-line data. *Journal of Criminal Justice Education*, *21,* 300-321.

Holt, T. J. (2013). Exploring the social organization and structure of stolen data markets. *Global Crime*, *14,* 155-174.

Holt, T. J., & Blevins, K. R. (2007). Examining sex work from the client's perspective: Assessing johns using online data. *Deviant Behavior, 28,* 333-354.

Holt, T. J., & Lampke, E. (2010). Exploring stolen data markets on-line: Products and market forces. *Criminal Justice Studies*, *23,* 33-50.

Holt, T. J., Strumsky, D., Smirnova, O., & Kilger, M. (2012). Examining the social networks of malware writers and hackers. *International Journal of Cyber Criminology, 6,* 891-903.

Holt, T.J., & Turner, M.G. (2012). Examining risks and protective factors of on-line identity theft. *Deviant Behavior, 33,* 308-323.

Holz, T., Engelberth, M., & Freiling, F. (2009). Learning more about the underground economy: A case-study of keyloggers and dropzones." In M. Backes, & P. Ning (eds.), *Computer Security-ESCORICS*, (pp. 1-18). Berlin and Heidelberg, Springer.

Honeynet Research Alliance. (2003). Profile: Automated Credit Card Fraud. *Know Your Enemy Paper* series. [Online] Available at http://www.honeynet.org/papers/profiles/ccfraud.pdf.

Jacobs, B. (1996). Crack dealers apprehension avoidance techniques: A case of restrictive deterrence. *Criminology, 34,* 409-431.

Jacobs, B. (2000). *Robbing Drug Dealers: Violence Beyond the Law*. New York: Aldine de Gruyter.

Jacobs, B. A., Topalli, V., & Wright, R. (2000). Managing Retaliation: Drug robbery and informal sanction threats. *Criminology, 38,* 171-198.

James, L. (2005). *Phishing Exposed.* Rockland: Syngress.

Jamieson, A. (2000). *The Antimafia: Italy's Fight Against Organized Crime.* New York: Palgrave Macmillian.

Jewkes, Y., & Sharp, K. (2003). Crime, deviance and the disembodied self: transcending the dangers of corporeality. In Y. Jewkes (ed), *Dot.cons: Crime, deviance and identity on the Internet* (pp.1-14). Portland, OR: Willan Publishing.

Mann, D., & Sutton, M. (1998). Netcrime: More changes in the organisation of thieving. *British Journal of Criminology*, *38,* 201-229.

Markham, A. N. (2011). Internet Research. In D. Silverman (Ed.), *Qualitative Research: Issues of Theory, Method, and Practice, 3rd Edition* (pp. 111-127). Thousand Oaks, CA: SAGE Publications.- check date on this ref and either correct in text or in cite.

Meyer, G. R. (1989). *The social organization of the computer underground.* Master's thesis, Northern Illinois University.

Milrod, C. & Weitzer, R. (2012). The intimacy prism: Emotion management among the clients of escorts. *Men and Masculinities, 15,* 447-467.

Moore, R. (2012). *Cybercrime: Investigating high-technology computer crime.* London: Elsevier.

Morris, R. G. (2010). Identity thieves and levels of sophistication: Findings from a national probability sample of American newspaper articles 1995-2005. *Deviant Behavior, 31,* 184-207.

Motoyama, M., McCoy, D., Levchenko, K., Savage, S., & Voelker, G. M. (2011). An Analysis of Underground Forums. *IMC'11*, 71-79.

National Conference of State Legislatures. (2012). *State Security Breach Notification Laws.* http://www.ncsl.org/issues-research/telecom/security-breach-notification-laws.aspx

Newman, G., & Clarke, R. (2003). *Superhighway robbery: Preventing e-commerce crime.* Cullompton: Willan Press.

Olivier, J. & Norberg, M. N. (2010). Positively skewed data: Revisiting the box-cox power transformation. *IJPR, 3,* 68-75.

Peretti, K. K. (2009). Data breaches: What the underground world of "carding" reveals. *Santa Clara Computer and High Technology Law Journal, 25,* 375-413.

Poulsen, K. (2012). *Kingpin: How one hacker took over the billion dollar cybercrime underground.* New York: Broadway.

Powell, W. W. (1990). Neither market nor hierarchy: Network forms of organization. *Research in Organizational Behavior, 12,* 295-336.

Robinson, M. (1984). *Groups.* New York: John Wiley and Sons.

Rogers, J. (2007). Gartner: Victims of online phishing up nearly 40 percent in 2007. *SC Magazine.* [Online] Available at http://www.scmagazineus.com/Gartner-Victims-of-online-phishing-up-nearly-40-percent-in-2007/article/99768/

Schneider, J.L. (2005). Stolen-goods markets: Methods of Disposal. *British Journal of Criminology, 45,* 129-140.

Sharp, K., & Earle, S. (2003). Cyberpunters and cyberwhores: prostitution on the Internet. In Y. Jewkes, (Ed.), *Dot Cons. Crime, Deviance and Identity on the Internet*. (pp. 36-52). Portland, OR: Willan Publishing.

Stevenson, R. J., Forsythe, L. M. V., & Weatherburn, D. (2001). The stolen goods market in New South Wales Australia: An analysis of disposal avenues and tactics. *British Journal of Criminology, 41,* 101-118.

Surowiecki, J. (2013). "Why did criminals trust liberty reserve," The New Yorker, May 31, 2013. [Online] Available at http://www.newyorker.com/online/blogs/newsdesk/2013/05/why-did-criminals-trust-liberty-reserve.html

Symantec Corporation. (2008). *Symantec: Symantec Internet Security Threat Report XII.* [Online] Available at http://eval.symantec.com/mktginfo/enterprise/white_papers/b-whitepaper_internet_security_threat_report_xii_04-2008.en-us.pdf

Symantec Corporation. (2012). *Symantec Internet security threat report, Volume 17.* [Online] Available at http://www.symantec.com/threatreport/

Team, T. (2013). Visa and MasterCard Battle For Share in Global Shift To Plastic. *Forbes Online* May 23, 2013. [Online] Available at http://www.forbes.com/sites/greatspeculations/2013/05/03/visa-and-mastercard-battle-for-share-in-global-shift-to-plastic/

The Nilsen Report. (2013). *Purchase Volume Market Shares for US General Purpose Brands 2007 to 2012.* [Online] Available at http://www.nilsonreport.com/publication_chart_and_graphs_archive.php

Thomas, R. & Martin, J. (2006). The underground economy: Priceless. *;login: The Usenix Magazine, 31,* 7-17.

Unger, B. (2007). *The scale and impacts of money laundering.* Cheltenham: Edward Elgar.

U.S. Department of Justice. (2008). *Overview of the law enforcement strategy to combat international organized crime.* U.S. Department of Justice, Washington D. C. [Online] Available at http://www.justice.gov/ag/speeches/2008/ioc-strategy-public-overview.pdf

Van Duyne, P. C., & Levi, M. (2005). *Drugs and money: Managing the drug trade and crime money in Europe.* London: Routledge.

Verini, J. (2010). The Great Cyberheist. *The New York Times*. November 14, 2010. [Online] Available at http://www.nytimes.com/2010/11/14/magazine/14Hacker-t.html?_r=1

Verision. (2012). *2012 Data Breach Investigations Report and Executive Summary*, 2012. [Online] Available at http://www.verisionbusiness.com/resources/reports/rp_data-breach-investigations-report-2012_en_xg.pdf

Wall, D.S. (2001). Cybercrimes and the Internet. In D. S. Wall (Ed.), *Crime and the Internet* (pp.1-17). New York: Routledge.

Wall, D. S. (2007). *Cybercrime: The transformation of crime in the information age.* Cambridge: Polity Press.

Wehinger, F. (2011). The Dark Net: Self-Regulation Dynamics of Illegal Online Markets for Identities and Related Services. *Intelligence and Security Informatics Conference* 209-213.

Wright, R.T., & Decker, S.H. (1994). *Burglars on the Job: Streetlife and Residential Break Ins.* Boston: Northeastern University Press.

Yip, M., Webber, C., & Shadbolt, N. (2013). Trust among cybercriminals? Carding forums, uncertainty, and implications for policing. *Policing and Society, 23,* 1-24.

Zhang, J., Ackerman, M. S., & Adamic, L. (2007). Expertise networks in online communities: Structure and algorithms. *WWW'07: Proceedings of the 16th International Conference on World Wide Web*, 221-230.

Zickuhr, K., Rainie, L., Purcell, K., & Duggan, M. (2013). *How Americans value public libraries in their communities.* Pew Internet and American Life Project. Retrieved from http://libraries.pewinternet.org/

VI. Dissemination of Research Findings

A. Papers

Accepted, In Press

Holt, Thomas J., Olga Smirnova, and Yi-Ting Chua. 2013. "An exploration of the factors affecting the advertised price for stolen data." Proceedings of the eCrime Research Summit, September 17-18, 2013.

In Print

Holt, Thomas J. 2013. “Exploring the social organization and structure of stolen data markets.” *Global Crime*, 14: 155-174.

Holt, Thomas J. 2012. “Exploring the Social Organization and Structure of Stolen Data Markets.” Proceedings of the 4th Annual Illicit Networks Workshop. Vancouver B.C., Canada, October 1-2, 2012.

B. Presentations

Holt, Thomas J., Olga Smirnova, and Yi-Ting Chua. 2013. "An exploration of the factors affecting the advertised price for stolen data." Presented at the eCrime Research Summit, San Francisco, California, September 17-18, 2013.

Holt, Thomas J. 2013. "Exploring the Role and Structure of Stolen Data Markets in the Cyber Underground." Invited Presentation at the Home Office, London, England, April 29, 2013.

Holt, Thomas J. 2013. "A Qualitative Analysis of the Social Organization of Stolen Data Markets." Presented at the annual meeting of the Academy of Criminal Justice Sciences meetings, Dallas, Texas.

Holt, Thomas J. 2012. "Exploring the Social Organization and Dynamics of the Online Black Market." Presented at the annual meeting of the American Society of Criminology meetings, Chicago, Illinois.

Holt, Thomas J. 2012. "Transnational Organized Crime: Impact and Outcomes." Panelist and Presenter at the National Institute of Justice Conference, Arlington, VA, June 20, 2012.

Fitzgerald, Sarah, Amelia Levett, and Thomas J. Holt. 2011. "Examining the economics and products available in stolen data markets." Presented at the annual meetings of the American Society of Criminology, Washington D. C.

Appendix

Glossary

Cashout Services: service offered in stolen data markets to enable actors to access, remove, and drain funds from bank accounts on and off-line for either a percentage of the total amount or a flat fee.

CVV: Credit Verification Value imprinted as a three to four digit number on the signature line of credit cards that enables the cardholder to make purchases without being physically present at the time of the transaction

Data Breach: an unintentional release of sensitive information, including personally identifiable information, often through some application of computer hacking.

Drop: a service offered in stolen data markets to either send or accept goods purchased using dumps or compromised accounts. May also refer to service where unwitting victims will cash checks or payments made from a compromised account and then wire the funds to another account.

Dump: stolen credit card or bank account number and the associated customer data that can be obtained through different means that is commonly sold across stolen data markets.

Escrow: form of payment offered by trusted agents within stolen data markets where a third party holds payments on behalf of a seller until the buyer confirms they have received the items they ordered. The use of escrow payments enables trust between actors, but adds to the complexity of any transaction. May also be referred to as a guarantor service.

Fullz: dumps that contain all of the information associated with the account and account holder sold in stolen data markets.

Hijack: any attempt to disrupt a seller's thread by posting an advertisement for their products or services.

ICQ: an instant messaging computer program designed in Israel and now owned by the Russian company Mail.ru Group.

Internet Relay Chat (IRC): a form of computer-mediated communication that sends messages in near real time between users via a client that is installed locally on user systems. Discussions are separated into channels based on interests and groups, but individuals can also send message to individual users.

Liberty Reserve: an electronic payment system using digital currency that operated out of Costa Rica until its dissolution in May 2013 due to law enforcement investigations.

Malware: malicious software that takes multiple forms in order to harm computer systems and compromise sensitive data.

Moderator: an individual within forums who has the responsibility to manage discussions and block users for violations of forum rules.

Phish/phishing: an attempt to acquire sensitive personally identifiable information, including bank account details, usernames, and passwords, through the use of fraudulent on-line communications.

Plastics: blank credit cards with unwritten magnetic stripes that can be developed into fraudulent cards through the use of holograms, embossing equipment, and dumps

Private Message (PM): a direct message sent between participants in forums through internal messaging systems managed within the forum.

Ripper: individuals who attempt to steal money from buyers in the stolen data market through the sale of invalid data or non-delivery of goods. This term is used derogatorily in order to warm others in the market.

Web Forum: an asynchronous form of computer-mediated communication that allows for ongoing, in-depth discussions of specific topics and issues organized into discrete categories (also called a bulletin board or message board). Forums can regulate access to content through the use of registration systems. Those forums who do not allow users to view content without registering are referred to as closed, while those that provide access to anyone are typically called open forums.

Web Forum Post: the basic building block of web forums. Individuals provide their opinions or pose questions in a post to the forum. A series of posts on a single topic is referred to as a thread.

WebMoney: an on-line payment system designed to provide real time transactions between participants through direct transfers based on WebMoney (WM) units. This system does not require users to enter bank account or credit card details, providing a level of anonymity to any transaction.