Exploring Threats and Vulnerabilities in Hacker Web: Forums, IRC and Carding Shops

Victor Benjamin[1], Weifeng Li[1], Thomas Holt[2], Hsinchun Chen[1]

[1]Department of Management Information Systems
The University of Arizona
Tucson, AZ 85721
{vabenji, weifengli}@email.arizona.edu,
hchen@eller.arizona.edu

[2]School of Criminal Justice
Michigan State University
East Lansing, MI, 48824
holtt@msu.edu

Abstract— **Cybersecurity is a problem of growing relevance that impacts all facets of society. As a result, many researchers have become interested in studying cybercriminals and online hacker communities in order to develop more effective cyber defenses. In particular, analysis of hacker community contents may reveal existing and emerging threats that pose great risk to individuals, businesses, and government. Thus, we are interested in developing an automated methodology for identifying identify tangible and verifiable evidence of potential threats within hacker forums, IRC channels, and carding shops. To identify threats, we couple machine learning methodology with information retrieval techniques. Our approach allows us to distill potential threats from the entirety of collected hacker contents. We present several examples of identified threats found through our analysis techniques. Results suggest that hacker communities can be analyzed to aid in cyber threat detection, thus providing promising direction for future work.**

Keywords - Cyber security, Hacker forums, Hacker IRC, Carding shops, Threats, Vulnerabilities

I. Introduction

Over the past decade a growing amount of critical infrastructures have begun to rely on computer and information technologies in order to meet increasingly complex demands. While use of technology has helped achieve more advanced capabilities within infrastructure, an unfortunate consequence is that many of such systems are facing growing exposure and risk to cyber-attack. To further exacerbate the issue, advancing technologies are enabling hackers to commit cybercrime at a much greater scale now than in the past. The sheer number of emerging security threats necessitates further research and development for mitigating risk and exposure to vulnerabilities. As a result, researchers and practitioners have taken an increased interest in advancing current cybersecurity capabilities.

Traditional research in the security domain has often focused on improving security built directly into computing and networking. Often times this stream of research focuses scrutiny on vulnerabilities at the protocol and system levels, where incremental advancements can be made in order to thwart existing security threats. Conversely, very little work has been done to go beyond technological issues and instead focus investigation on the human element behind cybercrime. For example, much is unknown concerning hacker behaviors, the cybercriminal supply chain, underground hacker communities, etc. Specifically, the development of methods to model cyber adversaries is one of the critical but unfulfilled research need outlined in a 2011 report on cybersecurity by the National Science and Technology Council [1]. More research on “black hat hackers”, i.e., cybercriminals, would offer new knowledge on securing cyberspace against those with malicious intent, leading to the development of more effective countermeasures against security threats [2].

In particular, analysis of hacker community contents may reveal existing and emerging threats that pose great risk to individuals, businesses, and government. Thus we are motivated to develop automated methods for identifying, collecting, and analyzing hacker community contents in search of threats and vulnerabilities. Such a capability to extract tangible and verifiable evidence of threats would be of great asset to the security community.

II. Literature Review

We focus our literature review on previous hacker community studies, as past work provides context and helps guide the formulation of this research. In particular, we review research concerning three key types of hacker communities: (1) hacker forums, (2) hacker Internet-Relay-Chat (IRC), and (3) carding shops. These community types contain their own unique data that may aid in the identification of threats and vulnerabilities.

Hacker Forums

Hackers make extensive use of online communities such as forums to support cybercriminal activity. Primarily, participants in such forums openly disseminate tacit knowledge and share tools much like a legitimate electronic network-of-practice. Hacking tools, malware samples, source code, etc. are often freely distributed among forum participants by simply attaching them to posted messages [3, 4]. Additionally, knowledge and methodology is disseminated among hackers in the form of tutorials written as text files or embedded in images, and even instructional videos [5]. Many of these tutorials directly enable readers to launch cybercriminal attacks such as denial of service attacks, SQL injections, cross-scripting attacks, and more.

In general, a wealth of cybercriminal knowledge is distributed among hacker forum participants, with many individuals advancing their capabilities to commit cybercrime by utilizing such resources. It is very feasible for an individual with little to no hacking skills to gain knowledge and capability by simply visiting different hacker communities and consuming their contents. Research identifies the existence of such hacker communities to be common across various geopolitical regions, including the US, China, Russia, the Middle-East, and other regions where information technologies are either ubiquitous or growing rapidly in use [6][7]. This presents a growing problem of global significance. Research in this area has potential for high-impact on society.

Unfortunately, hacker forum data is often difficult and nontrivial to collect. Many hacker communities employ anti-crawling features, such as bandwidth monitoring, paywalls, restricted access based on forum tenure / reputation requirements, invitation-based access, and CAPTCHA images / verification codes to mitigate automated bot activity [3, 8, 9]. Additionally, there is potential for such forums to carry malware intended to infect visitors, as many cybercriminal-related web pages may be scripted to attempt exploitation of web browser vulnerabilities [10]. Such exploits could allow remote attacks to arbitrarily execute code on vulnerable machines, which is of concern to researchers and practitioners. Unlike many normal web contents, a strategic and security-minded approach must be considered when designing methods to automatically collect data from hacker forums.

Hacker IRC

As indicated previously, hacker communities commonly exist as Internet forums. However, they may also exist as Internet-Relay-Chat (IRC) channels. Unlike forums which can be accessed through a typical web browser, IRC exists on a separate protocol and can support real-time, synchronous chat among thousands of users simultaneously. Hackers commonly use IRC for real-time communication, as forums are more slow-paced. An IRC-specific client is needed to connect to IRC servers that host multiple channels open for participants to make use of. Each participant can choose to join one or many channels. Additionally, IRC messages are broadcast to all connected users. This differs from forums where participants may only browse threads of interest.

The majority of current hacker community research experiments with forum data rather than IRC channel data [7]. This may be perhaps due to easier accessibility of webpage-based forums than IRC channels requiring connection through a specific IRC client. Furthermore, forums act as natural archives of data, where threads and posts are stored and can be easily accessed years later. Conversely, IRC contents must be collected in real-time and are not normally archived anywhere for later retrieval. This difference in how data is stored by the two platforms also leads into differences concerning ease of identification; content shared on web forums can be indexed by search engines, while IRC channel contents are excluded. For this reason, researchers attempting to identify hacker communities for study are more likely to come across hacker forums than IRC channels.

Methods to identify and collect hacker IRC channels have been developed in previous studies. The major identification techniques used in hacker community research appear to revolve around keyword searches and scrutinizing known communities for hyperlinks and references to other potential hacker communities [3, 11]. Some hacker IRC can be identified as a result of this process. After identification, steps can be taken to collect data. Specialized IRC listener programs can be developed to utilize the IRC protocol and sit-in on known hacker IRC channels [3, 7]. The listener programs can passively log all data transmitted between channel participants.

Carding Shops

Carding shops are another important part of the global hacker and cybercriminal community. Carding shops help facilitate cyber carding crimes as they provide a supply chain for carders who wish to sell stolen cards. Unfortunately, there is very little academic literature concerning carding shops despite their importance.

Monitoring carding shops can allow card issuers and others to mitigate associated risks or losses by taking precautions at early stages. By leveraging stolen cards' metadata, we can infer useful information that allows us to identify emerging targets or victims [12]. In this research, we aim to identify emerging threats and targets and provide actionable intelligence for cybersecurity decision makers. To this end, carding shop data is of importance to our goals.

III. Technical Approach

Different methodologies must be employed for collecting and analyzing hacker contents from forums, IRC, and carding shops. Each platform has its own intricacies that require different strategies for research. We provide details of our work in regards to each type of hacker community.

Forums

We propose the development of an information retrieval framework for identifying emerging threats from the forums. The framework is based on the previous AZSecure framework and includes a hacker forum analysis pipeline accounting for collection and text analysis [12]. In particular, the framework proposed here performs comprehensive keyword weighting and search over our hacker forum corpus for identifying potential threats. Identified threats are categorized within three categories: (1) attack vectors and software vulnerabilities, (2) financial fraud threats, and (3) other notable threats. Within each threat category, postings are ranked based on the frequency of weighted keywords. We show our framework in Figure 1. Overall, the framework follows a traditional information retrieval pipeline.

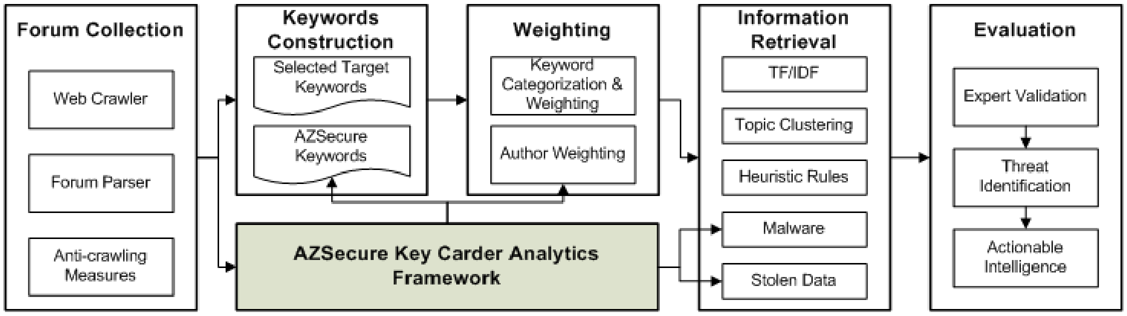


Figure 1 - Proposed Analytical Framework for Hacker Forums

Messages identified by our framework are assumed to contain evidence of potential threats. Generally, this includes information concerning the attack/threat details, as well as the potential victim targeted. Targets can be mentioned in the form of an IP address/URL, a Twitter hashtag/reply, a company, a product, a person, a street address, a serial number, or a phone number. The attack details contain information regarding the attack, including any specific details about vulnerabilities or methodology used.

One major output of the AZSecure Key Carder Analytics Framework is key hackers involved with the sale of malware or those that are associated with financial fraud threats [12]. Financial fraud threats includes the sale of stolen financial data, such as credit cards and bank accounts. The AZSecure framework possesses capability to assess seller quality based on feedback of other forum participants; sellers who are trustworthy and practice fair business are generally spoken well of, while scammers who rip off other participants are denounced within the forums. We use this information to help rank the framework's output. The intuition is that the postings of better quality sellers should rank higher because they are more likely to exhibit a real threat. Overall, we have extracted 373 malware sellers and 2,574 stolen data sellers.

The information retrieval component of our framework specifically focuses on identifying postings containing potential threats from our forum collection, and ranks postings based on relevance and urgency. For ranking, we developed a custom relevancy metric to order postings by their relevancy to a given input keyword. The relevancy metric is operationalized as the product of the input terms frequency within the post, the terms inverse document frequency, the terms relevance towards real-world entities that may experience threats (e.g., banking institutions, industry), the seller's weight, and the messages topical relevance.

Overall, the proposed information retrieval framework is an automated, scalable approach that aids us in identifying potential threats among hacker forums. Additionally, the framework can be easily tuned to reflect changes in hacker forums over time. Keywords can be added or removed, author weighting can be recalculated, topic categories can change, and so on.

IRC

IRC data is quite different from forum data and requires its own set of methodologies to collect and analyze. We first identify several hacker IRC channels through the use of keyword searches and by scrutinizing collected forums for information. After identification, we deployed automated IRC chat logging bots to identified IRC servers. As IRC data must be collected in real-time, multiple bots were issued from different hosts to avoid gaps in data collection caused by potential, bans, etc. We also practiced identity obfuscation by routing Internet traffic generated by our collectors through the Tor peer-to-peer anonymization network.

Similar to the AZSecure framework we utilize for our forum analysis, we base our IRC analysis on a list of keywords that may help identify potential threats. We first compute the term frequency of each keyword appearing in the provided list across our IRC collection. This provides us with a summary of how much each keyword was used by hacker IRC participants, thus helping identify most popular topics discussed. This type of ranking is useful as it can provide a quick summary of overall conversation occurring within hacker IRC communities. Additionally, we also tried to calculate the overall document frequency that each keyword appears in. This would provide us with information on how many distinct messages a keyword appears in, and not just overall frequency. However, IRC messages in our dataset are typically short-length and rarely contain the same keyword more than once per message, thus resulting in term frequency and document frequency possessing similar values.

Instead of document frequency, we found it to be more helpful to compute the number of different IRC participants that discuss each keyword. This provides us with some information on how widespread a keyword was discussed among all users. Additionally, this sort of ranking helps form a distinction between words discussed very frequently among only a limited number of users against words discussed more commonly among the broader hacker community.

Lastly, as all IRC messages are publicly broadcast to other participants, it is common etiquette for participants to often directly address one another within their messages in order to direct focus. We can use direct addressing to compute the social network among participants for each community. This is particularly helpful for identifying key actors of interest that may provide the most valuable evidence of potential threats.

Overall, we can use the different ranked lists to gain quick insight on the topics discussed within various hacker IRC channels. Additionally, top results of interest can be targeted for closer scrutiny to better understand hacker conversations regarding specific topics. .

Carding Shops

We demonstrate our proposed framework for collecting and analyzing carding shop metadata in Figure 2. The framework leverages data collection to gather carding shop metadata, ETL (extract-transfer-load) preprocessing to process the collected data into a uniform format, and reporting to conduct analysis from four perspectives: risk measure, location analysis, carder preference, and risk type.

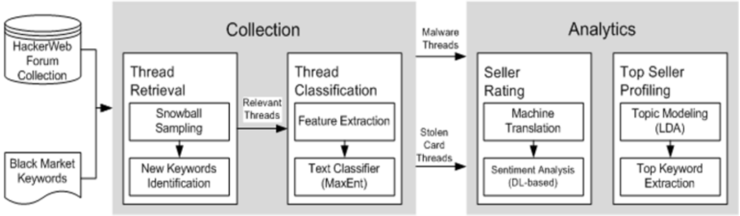


Figure 2. The Carding Shop Monitoring and Reporting Framework

There are three types of meta-information accessible to the buyers prior to their purchase: credit card information, data source information, and pricing information. However, carding shops differ in the granularity of detail they provide for each type of information. For example, few carding shops show detailed location information such as zip code; many of them list only the country or state.

IV. RESEARCH TESTBED

Data was collect from various sources for this research. For forums, we identified and collected recent contents (2014-2015) from several major hacker forums. The forums spanned both English and Russian communities. We utilized automated crawlers for collection, and circumvented anti-crawling measures through identity obfuscation techniques and adjusting crawling rates to conceal collection activities. We summarize the collection in Table I:

TABLE I. FORUM TESTBED

Forum	Language	Members	Threads	Posts	Time Start	Time Stop
A****t**	Russian	493	473	8,957	1/1/2014	8/19/2014
C*******e**	English	337	345	879	1/1/2014	12/29/2014
C*******m**	English	3,359	7,740	2,111	1/1/2014	1/2/2015
C**o**	Both	6,301	5,278	23,523	1/1/2014	1/20/2015
C***b**	Both	2,401	2,895	12,428	1/1/2014	1/22/2015
E***t**	Russian	1,205	1,090	10,071	1/1/2014	1/1/2015
H****e**	English	223	370	1,070	1/1/2014	12/29/2014
H*****d**	English	436	394	2,056	1/1/2014	12/30/2014
X**c**	Russian	3,099	27,936	34,422	1/1/2014	1/1/2015
Zy**	Russian	1,426	1,357	3,836	1/1/2014	1/3/2015

We also identified and collected chat data from several hacker IRC channels using automated collection programs from October 1st, 2014 until January 2nd, 2015. Collection for this analysis spanned hacker IRC channels from public IRC networks as well as IRC networks serving as hidden services, such as within the Tor anonymization network. During collection we observed that channels varied in volume of participant activity, thus we selected the most active channels for this analysis. Our collection can be seen in Table II.

TABLE II. IRC TESTBED

Server	Channel	# of Messages	Start Date	End Date
A*****s.org	#a*****s	173,309	10/1/14	1/2/15
E******e.org	#e******e	83,546	10/1/14	1/2/15
H**5.org	#h**5	37,609	10/1/14	1/2/15
U******t.org	#h********m	111,455	10/1/14	1/2/15
U******t.org	#c******e	707,772	10/1/14	1/2/15
C**************t.onion	#a***a	11,012	10/1/14	1/2/15
I**************a.onion	#a****t	7,838	10/1/14	1/2/15
6**************f.onion	#f******s	35,217	10/1/14	1/2/15

For carding shops, we identified four major carding shops using our AZSecure key carder identification framework. The shops were collected using automated crawlers, similar to hacker forums. The products sold on these shops consisted of stolen credit card information, and “dumps,” which refer to the information encoded within the magnetic strip encoded on the back of a credit card. A summary of our collection can be seen in Table III:

TABLE III. CARDING SHOP TESTBED

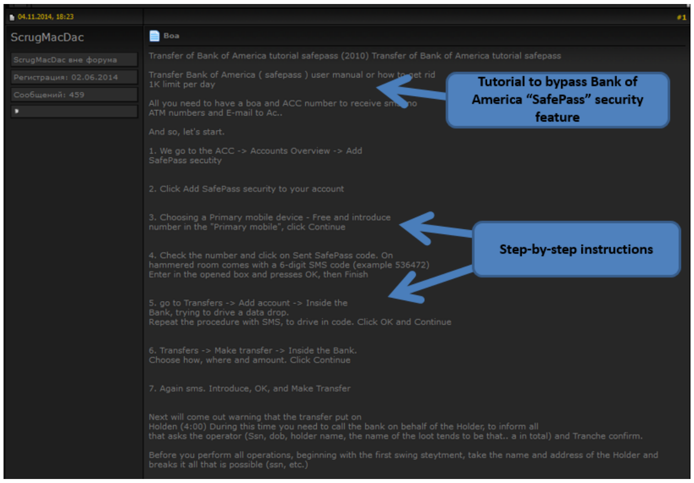
Carding Shop	Products	Date Collected	Collected Listings	Location Level
R******r	Credit Cards, Dumps	1/6/2015	9,055	Zip
G*******v	Credit Cards, Dumps	1/6/2015	32,755	City
C*******e	Credit Cards	1/6/2015	31,573	Country
S**********s	Dumps	1/6/2015	189,132	Country

V. SAMPLE RESULTS

We conduct our analysis of emerging threats and vulnerabilities within forums, IRC, and carding shops based on the described technical approach. Here we present and discuss some sample findings produced by our approach. We group findings based on their data source

Forums

We applied our proposed information retrieval system based on our AZSecure framework on collected hacker forums. As a result, we found various forum messages that contain information concerning real-world cyber threats. We provide two identified forum-based threats to demonstrate our findings. First, within the *C****o* forum, we see a participant posting a step-by-step guide to help others circumvent the "SafePass" security feature used on Bank of America accounts. The guide was posted on November 4th, 2014. The original forum posting can be seen in Figure 3.



04.11.2014, 18:23 #1

ScrugMacDac

ScrugMacDac вне форума

Регистрация: 02.06.2014

Сообщений: 459

Boa

Transfer of Bank of America tutorial safepass (2010) Transfer of Bank of America tutorial safepass

Transfer Bank of America (safepass) user manual or how to get rid
1K limit per day

All you need to have a boa and ACC number to receive sms no
ATM numbers and E-mail to Ac..

And so, let's start.

1. We go to the ACC -> Accounts Overview -> Add
SafePass secutity

2. Click Add SafePass security to your account

3. Choosing a Primary mobile device - Free and introduce
number in the "Primary mobile", click Continue

4. Check the number and click on Sent SafePass code. On
hammered room comes with a 6-digit SMS code (example 536472)
Enter in the opened box and presses OK, then Finish

5. go to Transfers -> Add account -> Inside the
Bank, trying to drive a data drop.
Repeat the procedure with SMS, to drive in code. Click OK and Continue

6. Transfers -> Make transfer -> Inside the Bank.
Choose how, where and amount. Click Continue

7. Again sms. Introduce, OK, and Make Transfer

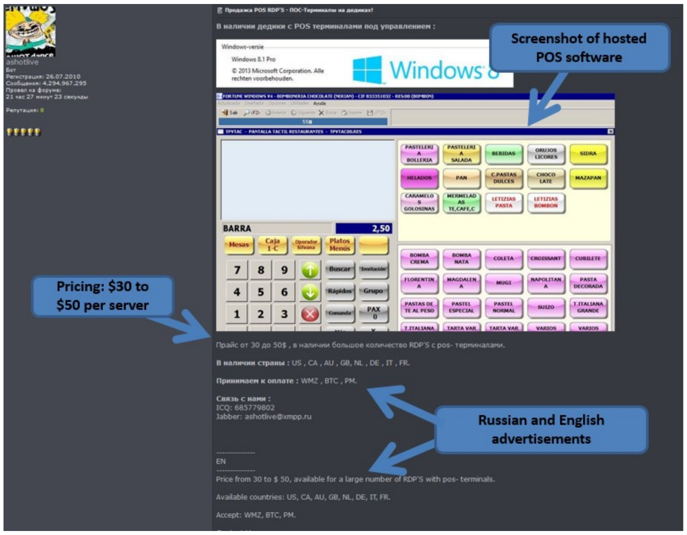
Next will come out warning that the transfer put on
Holden (4:00) During this time you need to call the bank on behalf of the Holder, to inform all
that asks the operator (Ssn, dob, holder name, the name of the loot tends to be that.. a in total) and Tranche confirm.

Before you perform all operations, beginning with the first swing steytment, take the name and address of the Holder and
breaks it all that is possible (ssn, etc.)

Tutorial to bypass Bank of America "SafePass" security feature
Step-by-step instructions

Figure 3 – *C****o* forum post containing guide to bypass Bank of America "SafePass"

A second example involves the Russian *A******t* forum. In this message, a Russian hacker posts an advertisement inviting others to rent servers running various point-of-sales software. Other hackers could rent servers to analyze available point-of-sales software for vulnerabilities that may be exploited to steal data from businesses. This example can be viewed in Figure 4.



ashotlive
Бот
Регистрация: 26.07.2010
Сообщения: 4,294,967,295
Провел на форуме:
21 час 27 минут 23 секунды

Репутация: 0

Продажа POS RDP'S - ПОС-Терминалы на дедиках!

В наличии дедики с POS терминалами под управлением :

Windows-versie
Windows 8.1 Pro
© 2013 Microsoft Corporation. Alle rechten voorbehouden.

Прайс от 30 до 50$, в наличии большое количество RDP'S с pos- терминалами.

В наличии страны : US , CA , AU , GB, NL , DE , IT , FR.

Принимаем к оплате : WMZ , BTC , PM.

Связь с нами :
ICQ: 685779802
Jabber: ashotlive@xmpp.ru

EN

Price from 30 to $ 50, available for a large number of RDP'S with pos- terminals.

Available countries: US, CA, AU, GB, NL, DE, IT, FR.

Accept: WMZ, BTC, PM.

Figure 4 – A Russian hacker of the *A******t* forum offers servers to rent for the purposes of analyzing and developing vulnerabilities for point-of-sales software.

IRC

Analysis of IRC contents helped extract the most relevant conversations between hackers that contained potential threats. As IRC is plaintext in nature, participants often share hyperlinks to external resources that provide more information and context to supplement conversations within chat. These external resources are often of great interest; we document two examples here.

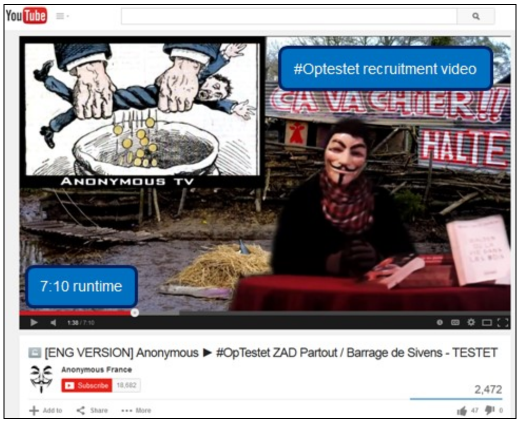
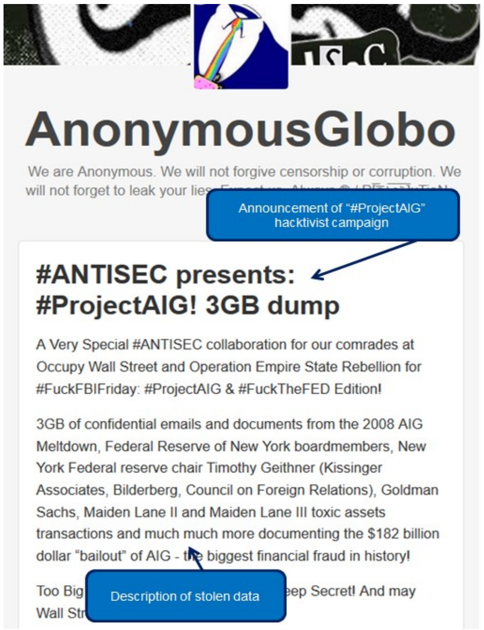


Figure 5 - Recruitment video for #Optestet, an environmental campaign. Viewable at https://www.youtube.com/watch?v=rO7uo6Wy--0

We collect and analyze the *#A******s* IRC channel, which claims affiliation with the *Anonymous* hacking group. *Anonymous* is widely considered to be one of the most active and popular hacking groups currently operating. Within their IRC community, they routinely post recruitment messages and videos for various hacktivist campaigns. Figure 5 contains an example of a recruiting video for the *#OpTestet* campaign, which is an environmentalist campaign targeting the French Ministry of Defense over a construction project.



AnonymousGlobo

We are Anonymous. We will not forgive censorship or corruption. We will not forget to leak your lies

#ANTISEC presents:
#ProjectAIG! 3GB dump

A Very Special #ANTISEC collaboration for our comrades at Occupy Wall Street and Operation Empire State Rebellion for #FuckFBIFriday: #ProjectAIG & #FuckTheFED Edition!

3GB of confidential emails and documents from the 2008 AIG Meltdown, Federal Reserve of New York boardmembers, New York Federal reserve chair Timothy Geithner (Kissinger Associates, Bilderberg, Council on Foreign Relations), Goldman Sachs, Maiden Lane II and Maiden Lane III toxic assets transactions and much much more documenting the $182 billion dollar "bailout" of AIG - the biggest financial fraud in history!

Too Big ... eep Secret! And may
Wall Str

Announcement of "#ProjectAIG" hacktivist campaign
Description of stolen data

Figure 6 - *#ProjectAIG*, a hacktivist campaign targeting and releasing stolen data from the AIG insurance group.

Another threat identified by our research involves attacks against financial and insurance firms. In particular, attacks were intended to steal data that could then be made publicly available with the goal of exposing corruption. In one case, an incident referred to as *#ProjectAIG* resulted in 3GB of confidential e-mails and documents to be stolen from the AIG insurance group and leaked publicly (Figure 6).

Carding Shops

As mentioned previously, carding shops provide a mechanism for hackers to sell stolen credit cards and other financial data. Monitoring carding shops can allow card issuers and others to mitigate associated risks or losses by taking precautions at early stages. We identify tens of thousands of credit cards for sale in just the four carding shops we monitor (Table IV). Carding shops outside of our current focus would possess even more stolen data. Overall, analysis of carding shop data can provide insights into market activities and potential breaches. For example, if the carding shops experience a sudden surge in cards for sale, it may correlate with an ongoing data breach.

TABLE IV. TOP 10 INSTITUTIONS WITH STOLEN CARDS

Rank	Financial Institution	Frequency (Cards for Sale)
1	MasterCard	36,704
2	JPMorgan Chase	27,406
3	Bank of America	21,394
4	American Express	16,845
5	Wells Fargo	16,509
6	U.S. Bank	8,491
7	Citibank	7,216
8	Citizens Bank	6,012
9	Banco do Brasil	3,711
10	Fifth Third Bank	3,636

Typically cards are listed for sale in bundles referred to as *"*bases.*"* Cards bundled together for sale within bases typically have location information associated with them, designating what zipcode or city they are associated with. Based on this location data, we identify 6 cities to be the most active for stolen card activity (Table V). Such information may be useful to financial institutions for implementing policy or additional security at locations where carding activity seems very active.

TABLE V. MOST ACTIVE CITIES FOR STOLEN CARD ACTIVITY

City	# of Bases Involved
Austin, TX	14
Cincinnati, OH	14
Houston, TX	14
Newark, NJ	14
Phoenix, AZ	14
Rochester, NY	14

VI. CONCLUSION

Hacker forums, IRC channels, and carding shops all appear to contain a variety of contents relevant to discovering current and emerging cyber threats. From our work, we identify several examples of evidence related to threats against financial institutions and government. We will continue analysis of hacker community and carding shop data, and will look to expand the scope of our collection in the future.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under Grant No. SES-1314631 and also under Grant No. DUE-1303362.

REFERENCES

[1] National Science and Technology Council, "Trustworthy Cyberspace: Strategic Plan for the Federal Cybersecurity Research and Development Program," pp. 1–19, 2011.

[2] M. Siponen, D. Straub, H. R. Rao, and T. S. Raghu, "Moving Toward Black Hat Research in Information Systems Security An Editorial Introduction to the Special Issue," *MIS Q.*, vol. 34, no. 3, pp. 431–433, 2010.

[3] H. Fallmann, G. Wondracek, and C. Platzer, "Covertly Probing Underground Economy Marketplaces," *Proc. 7th Int. Conf. Detect. intrusions malware, vulnerability Assess.*, pp. 101–110, 2010.

[4] V. Benjamin and H. Chen, "Securing Cyberspace : Identifying Key Actors in Hacker Communities," *IEEE Intelligence and Security Informatics*, pp. 24–29, 2012.

[5] T. J. Holt, D. Strumsky, O. Smirnova, and M. Kilger, "Examining the Social Networks of Malware Writers and Hackers," *Int. J. Cyber Criminol.*, vol. 6, no. 1, pp. 891–903, 2012.

[6] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker, "An analysis of underground forums," *Proc. 2011 ACM SIGCOMM Conf. Internet Meas. Conf. - IMC '11*, p. 71, 2011.

[7] V. Benjamin and H. Chen, "Time-to-event Modeling for Predicting Hacker IRC Community Participant Trajectory," in *IEEE Intelligence and Security Informatics*, 2014.

[8] A. Abbasi, W. Li, V. Benjamin, S. Hu, and H. Chen, "Descriptive Analytics : Examining Expert Hackers in Web Forums," in *IEEE Intelligence and Security Informatics*, 2014.

[9] M. T. Qassrawi and H. Zhang, "Client Honeypots : Approaches and Challenges," *4th Int. Conf. New Trends Inf. Sci. Serv. Sci.*, pp. 19–25, 2010.

[10] M. Cova, C. Kruegel, and G. Vigna, "Detection and analysis of drive-by-download attacks and malicious JavaScript code," *Proc. 19th Int. Conf. World wide web - WWW '10*, p. 281, 2010.

[11] T. J. Holt and E. Lampke, "Exploring stolen data markets online: products and market forces," *Crim. Justice Stud. A Crit. J. Crime, Law, Soc.*, vol. 23, no. 1, pp. 33–50, Mar. 2010.

[12] Li, W., & Chen, H. (2014). Identifying Top Sellers In Underground Economy Using Deep Learning-based Sentiment Analysis. IEEE Joint Intelligence and Security Informatics Conference, 64-67.