The Secondary Global Market for Hacked Data

Amit Rechavi[1], **Tamar Berenblum**[2]
The Hebrew University of Jerusalem, Israel

David Maimon[3]
Georgia State University, United States of America

Abstract

Cyber crime and hacking have become ubiquitous over the past decades. Although many studies have explored hacking communities, only a few have investigated hacking networks on the country and cross-country levels. We collected data on successful brute-force attacks (BFAs) and system-trespassing incidents (Sessions) on honeypots (HPs). Based on one million interactions in one month, we built a network of hackers and hacked data depicting the different roles of countries in the hacking scene. We depicted a suspected data exchange between the BFA and Session hackers and examined the network's topology considering this data transfer. Mapping IP addresses and countries, we found that only a few countries lead the hacking activities and are the network's core. Our contribution lies in studying and mapping the dynamics of hacking activity on the country level and in providing insights into the dynamic of the concealed trading in usernames and passwords. Due to the severe consequences of hacking activities, our findings carry practical implications.

Keywords: Hacking, Network topology, Cyber-sphere, Core-periphery, SNA.

Introduction

Cyber crime and hacking have been ubiquitous in cyberspace for many years. Although not new (Furnell, 2002), they have evolved in recent years to become more sophisticated and better organized (Goel, 2011; Lau et al., 2012; Grabosky, 2004, 2017). Cyber crime and hacking have grave consequences for business, government (Rantala, 2008; Shackelford, 2009), and individual users (Bossler & Holt, 2009; Internet Crime Complaint Center 2017; www.ic3.gov). Cyber-attacks are also increasingly becoming a political issue between countries (Shackelford, 2009; Kumar & Carley, 2016a,b). Nevertheless, due to their inherently concealed nature, obtaining insights into global

[1] Ruppin Academic Center, Emek Hefer, Israel 4025000, and The Federmann Cyber Security Center, The Hebrew University of Jerusalem, Jerusalem, 9112102 Israel
Email: amit.rechavi@gmail.com
[2] Academic Coordinator, The Federmann Cyber Security Center, The Hebrew University of Jerusalem, Jerusalem, 9112102 Israel. Email: tamar.berenblum@mail.huji.ac.il
[3] Associate Professor, Department of Criminal Justice and Criminology, Georgia State University, 7251 Preinkert Dr, College Park, MD 20742, USA. Email: dmaimon@umd.edu

hacking patterns is challenging. Despite some empirical studies on hacking activities (e.g., Lakhani & Wolf, 2003; Spitzner, 2003b; Young et al., 2007, Yip et al., 2012; Holt & Smirnova, 2014; Maimon et al 2014; Wilson et al 2015; Leukfeldt et al., 2016, Leukfeldt et al., 2017; Testa et al 2017), we have little understanding of hacking activity on the global scale.

The borderless nature of cyberspace and the exponential take-up of digital technology throughout the world guarantee that international cyber crime will remain a challenge (Grabosky, 2004, 2017). The high global heterogeneity of cyber crime laws, the existence of countries whose substantive criminal laws and procedures are still not attuned to the digital age, and the fact that some governments initiate, and back cyber-attacks add to this challenge (Kshetri, 2005). Though several studies have mapped cyber-attacks on the country level (Kigerl, 2012; Rughiniş & Rughiniş, 2014), the global process of two-step attacks which involves several countries and data transfer between them, has not been studied yet. In this research, we map global hacking dynamics on the country level and present and analyze the process of these twofold cyber-attacks.

Consistent with past criminological (Maimon et al., 2014) and technological studies (Salles-Loustau et al., 2011; Farinholt et al., 2017), we installed research honeypots (HPs) on computer networks of Israeli academic institutions. We gathered data on continuance attempts of guessing-process of usernames and passwords. These hacking activities are called brute force attacks (BFAs). In case the guessing was successful the hacker started to execute his/her intentions and these system trespassing incidents are called Sessions. During the research, we mapped the two activities, and the hackers' IP addresses we aggregated the attacks to country-level nodes and reconstructed their network. We then filtered the data and considered only Session activities that had not emerged from near-time previous BFA events. We suspected these sessions were initiated by hackers who obtained the HPs' usernames and passwords from a secondary (underground) market of hacked data. Accordingly, we cross-referenced these Sessions with previous BFAs and successful sessions to map the flow of hacked data and to reconstruct the interrelations at the country level. We found that only a few countries led the hacking activities and were sufficient to serve as the network's core.

The following section briefly reviews related studies on social networks and cyber crime. Then, we propose hypotheses on the roles and topology of the hacking network. Based on these hypotheses, we analyze the data, discuss our results, and suggest research and practical implications and future directions.

Related Studies

Hacking

Hackers and their behavior have been studied for several decades. There is vast research on their identity, motivations, culture, and ethics (e.g., Lakhani and Wolf, 2003; Spitzner, 2003b; Young et al., 2007; Yip et al., 2012; Holt et al., 2012; Holt & Smirnova, 2014; Maimon et al 2014; Wilson et al 2015; Leukfeldt et al., 2016; Leukfeldt et al., 2017; Testa et al 2017; Waldrop, 2016). These studies have focused both on the micro level, examining the individual hacker, and on the mezzo level examining the hackers as a community from a sociological perspective (Jordan & Taylor, 2008) Studying the

interaction between hackers (Dupont et al. 2016) and their social ties online and offline (Leukfeldt et al., 2016).

However, only a handful of studies have explored the national level in such hacking activities. In general, higher corruption and large internet bandwidth favor attacks origination (Kumar and Carley, 2016b). Wealthier countries with widespread internet use and better Information and Communication Technologies (ICT) infrastructure are targets for cyber-attacks (Rughiniş & Rughiniş, 2014; Kumar and Carley, 2016a) and societies with more internet users per capita have higher cyber crime activity (Kigerl, 2016; Kumar and Carley (2016a, b).

Previous studies on the hacking activities of specific countries suggest that the country level is of importance and should be further explored. Yegneswaran et al. (2003) exploring a large quantity (25b) and a wide variety of daily intrusion attempts, found that a small group of attackers was responsible for a significant share of hacking attempts and that this group served as the core of the hacking network. In 2008 Fossi et al. in their Honey Pot project analyzed over one billion spam messages and found China, Brazil, the United States, Germany, and Russia to be the top spamming countries, while Spain and Germany were found to be the top e-mail spammers. Four years later, Eastern Europe, Russia, Asia, and countries with fewer anti-cyber crime regulations and many internet users were found to be home to more offenders (Kigerl, 2012) and in 2013, Kshetri identified countries in Eastern Europe as hotspots for cyber criminal networks. Several years later, Symantec (2016) reported on more than 430 million new unique pieces of malware coming from China, seeing an 84 percent rise in bot-related activity in that country from 2014.

The current study continues this line of thought and tries to nuance these results and differentiate the cyber-attack roles hackers assign to different countries. We built a global network of hacking activities, which, once processed and analyzed, depict the global transfer of data within the hacking community at the international level with the country as the unit.

Hacking community: Topologies and their functionality

Hackers act as an online community (Whittaker et al., 1997), sharing hacked contents in "dark" or underground markets or social media forums (Goel, 2011). The nature of these web-based criminal activities has evolved to the point where it exceeds the volume of a closed group. The hacking community functions as a many-to-many marketplace where disgruntled employees and vendors, potential buyers, and intermediaries sell and resell data (Motoyama et al., 2011; Holt et al., 2013; Holt & Smirnova, 2014, Ablon et al., 2014, p.3). The topology of this network of sellers and buyers of stolen data was explored by previous studies (Yip et al., 2012; Holt & Smirnova, 2014) and the activities there range from rudimentary to extremely sophisticated and involve the exchange of cyber crime-related goods and services, supporting and being supported by both on- and offline criminal activities. Symantec (2016), for example, reports that a web toolkit, which includes updates and 24/7 support, can be rented for between 100 and 700 US$ per week, and one can order distributed denial-of-service (DDoS) attacks for a price ranging from $10 to $1,000 per day.

These hacking activities are the basis of a developed global cybercrime network (Benjamin & Chen, 2012; Rechavi et al., 2015). The topology of a network evolves or is designed to execute the network's functionality. A network's structure contributes to its

dynamics (Watts, 1999) and it is defined according to the nodes' roles (Faloutsos et al., 1999). Without understanding the processes leading a community to its current topology, its analysis is meaningless (Berger-Wolf & Saia, 2006). Network structure plays a significant role in the data diffusion process (Abrahamson & Rosenkopf, 1997), enables delivering messages or viruses (Watts, 2004); leverages financial functionality (Baum et al., 2003); help users navigate through the network (Liben-Nowell et al., 2005) and more. In dynamic social networks, network's topology evolves over time (Berger-Wolf & Saia, 2006; Hill & Braha 2010), thus making the topology-functionality interaction even more crucial, where topology-functionality interaction is not obvious (Banos et al., 2013), nor unidirectional, as the topology can affect the functionality as well (Braha & Bar-Yam, 2006; Kossinets & Watts, 2006; Morgan et al., 1997; Viswanath et al., 2009).

In communities of hackers, though sometimes small networks of hackers are decentralized (Lu et al., 2010), hackers tend to create a core of members who collaborate (Leukfeldt et al., 2016, 2017), and knowledgeable hackers distribute their work and know-how to other hackers (Odinot et al., 2016). In this structure, central nodes have a dense connection between themselves, they operate as the "core" of the network and connect to a sparse, non-central set of nodes – the periphery, which is not intra-connected (Borgatti and Everett, 2000). The core-periphery structure is a suitable topology where a small group of people holds a large body of knowledge and wishes to diffuse and share it with the rest of the community (Gomez- Rodriguez et al., 2010; Aral & Walker, 2012; Rombach et al., 2014). It was found that the most reputable and reliable actors in the hacking communities in the US and China were found to be hackers who contribute to their communities' developing skills (Benjamin and Chen, 2012).

However, contrary to capturing the topology as a derivative of the network's purpose with a clear goal, some view topology as the sum of multiple (sometimes contradictory) sub-topologies, each with its purpose and goals (Cross et al., 2001; Holme et al., 2004; Lickel et al., 2006), where the actions of the network's members are not coordinated nor agreed by all members. In these networks, the communication between members is not organized nor governed by a single entity. Global hacking activities might be a good example of such a network, where overlapping hacking activities could form a single hacking network with a structure of core-periphery (Yang & Leskovec, 2014). In this study, we wish to explore whether a core-periphery structure exists and supports the global hacking activities.

Hypotheses

BFA hackers excel in finding network vulnerabilities. Session hackers excel in malicious actions once they have found a weakness. When BFA hackers manage to capture the target's access keys (often username and password), they can use them to access the computer themselves; alternatively, they may exchange them with Session hackers in secondary markets. The results of the trade are OOTB attacks. We follow Kigerl's (2016) approach by which the number of computers that can translate into the number of potential malware hosts, such as spam botnets, is correlated with the size of the population. We follow Kigerl's (2016) suggestion that based on the preferences and skills of their population as well as on their regulations and infrastructures; certain countries might be selected by hackers to pursue particular hacking activities. Therefore, our first hypothesis is: *H1: Countries might excel in different hacking roles (BFA and Session).*

Since Session activity is impossible without a username and password, and since we have detected the BFAs responsible for capturing this data, we suggest that there is a linkage between BFAs specialists and Sessions executers. This linkage might be hidden, but its outcomes are revealed in the network of attacks. And so our second hypothesis is:

H2: There is a secondary market for usernames and passwords, and the results of the transfer of this hacked data are evident in attacks on HPs.

Mapping the attacks, we can reverse-engineer the network comprising BFA and Session countries. Though each country operates according to its motivation, the overlapping hacking activities form a global core-periphery structure (Yang and Leskovec, 2014), since this topology best serves the exchange of hacked data (Gomez-Rodriguez et al., 2010; Aral & Walker, 2012). Therefore, our third hypothesis is:

H3: Since the BFA hackers diffuse data to the rest of the network, the structure of the network will have a suitable topology, meaning a critical core of BFA specialists sharing data, with peripheral Session executers.

Summing the hypotheses, BFAs and Sessions are distinct hacking roles, and we can depict the hacked data (access keys) which BFA countries transmit to Session countries (H1). The topology of this network depicts the secondary market for hacked data (H2), and the topology is core-periphery to optimize the distribution of hacked data (H3).

Methodology

The current study models the relations between hackers' IP addresses and HPs' IP addresses and maps a directed network of hacker IPs to the HPs' IPs. An HP is a "security resource whose value lies in being probed, attacked or compromised" (Spitzner 2003a) and it facilitates data collection on real-life system trespassing incidents (Sessions). In December 2015, we deployed target computers on 60 public IP addresses in Israeli academic networks for 36 days. The period of 36 days was determined in accordance with the demands and limitation of the cyber-security department of the academic institution, where the HPs were installed.

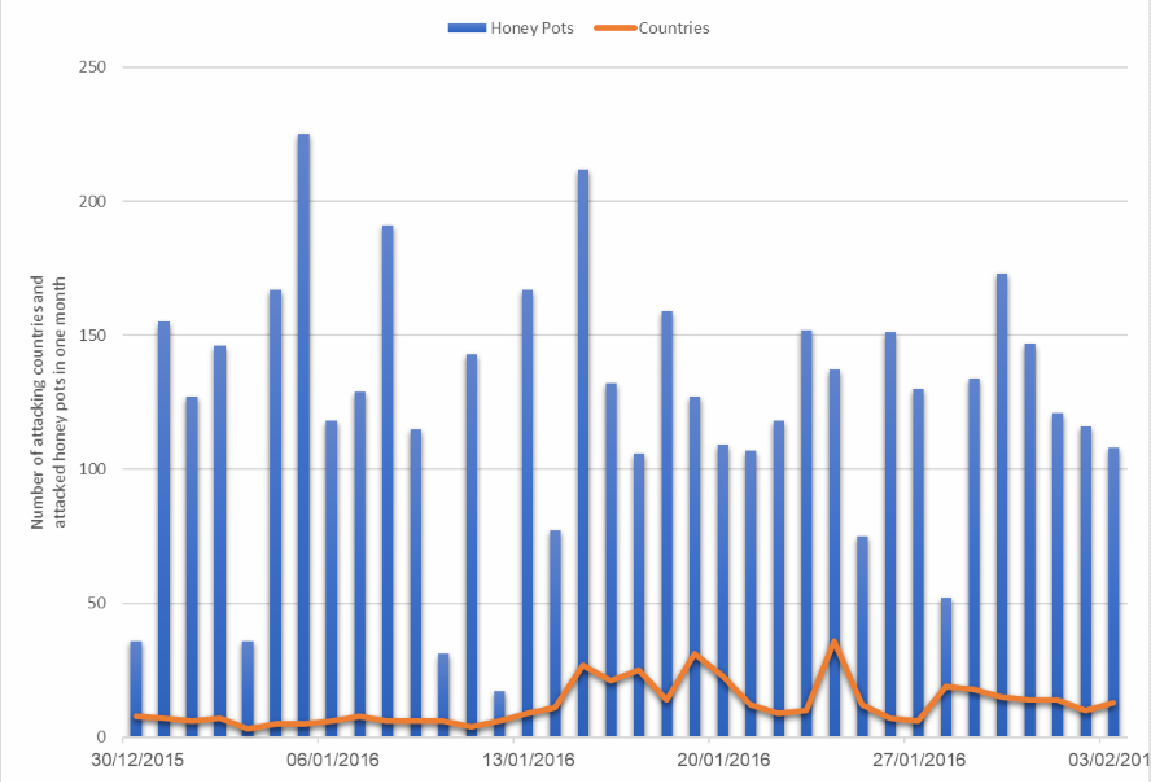
In this study, we differentiate between brute force attacks (BFAs) on HP and Sessions. In line with Keith et al. (2007) and the SANS Institute (2007), a victim of BFA is any target computer that has undergone a successful remote password-guessing attempt. These BFAs can occur several hundred times per minute, and their purpose is to reveal the correct username and password of computer systems. The second type of attack is a Session. A Session is defined as any unauthorized access facilitated by a previous BFA (Berthier & Cukier, 2009; Maimon et al., 2014). Session activity occurs once a hacker is inside a computer system and starts executing his/her malicious intentions such as copy, alter, delete or conceal data.

In the experiment, BFA hackers had to scan the network to penetrate an HP, identify the computers, and hack them through vulnerable entry points. The username and password of the HPs were randomly selected from a dictionary and screened for triviality and overuse. Having obtained access to the HP, an intruder could initiate Sessions for 30 days. Each keystroke of the intruder was logged and used for subsequent analysis. Graph 1 presents the volume of activity in one month of attacks.

We collected almost a million BFAs and several hundreds of Sessions. While exploring the data, we noticed several "out-of-the-blue" (OOTB) Sessions – those that began with

entering the correct username and password in the first attempt, without trial and error process, i.e., without related prior BFAs. Evidence suggests that the hackers in these Sessions knew the correct username and password without guessing, from a previous successful BFA (done by others) on the HP. We then looked for successful BFAs which have the highest chances to be the source of these correct username and password.

Graph 1. The volume of activity in one month of attacks



Our primary interest lay in identifying the transfer of hacked data between hackers who execute BFAs and hackers who use this data to execute Sessions. We used the data of these Session and their related hackers and linked them to the potential BFAs and their hackers. The relation between the BFAs and Session hackers created a global hacking network, which included all the connections between OOTB sessions and their related successful BFAs at the country level. Applying social network analysis (SNA) techniques, we analyzed the network's structure and behavior on the country level. In the 36-day period, 993K BFAs arrived from hundreds of IP addresses in multiple countries (See Appendix A); 389 attacks succeeded and turned into Sessions. Out of them, we manage to find 149 pairs of OOTB Sessions and successful BFAs, and these interactions formed the dataset we then analyzed. In our analysis, we did not distinguish between attacks made by hackers or bots. It is possible that at least some of the IPs of the BFAs belonged to zombie computers, part of a botnet activated by a third party located in a third country. The fact that bots or slave-computers can execute the attacks does not change the mapping and understanding of the transfer of hacked data between IPs and the two kinds of attacks.

Results

Figure 1. Attacks originated from 53 countries

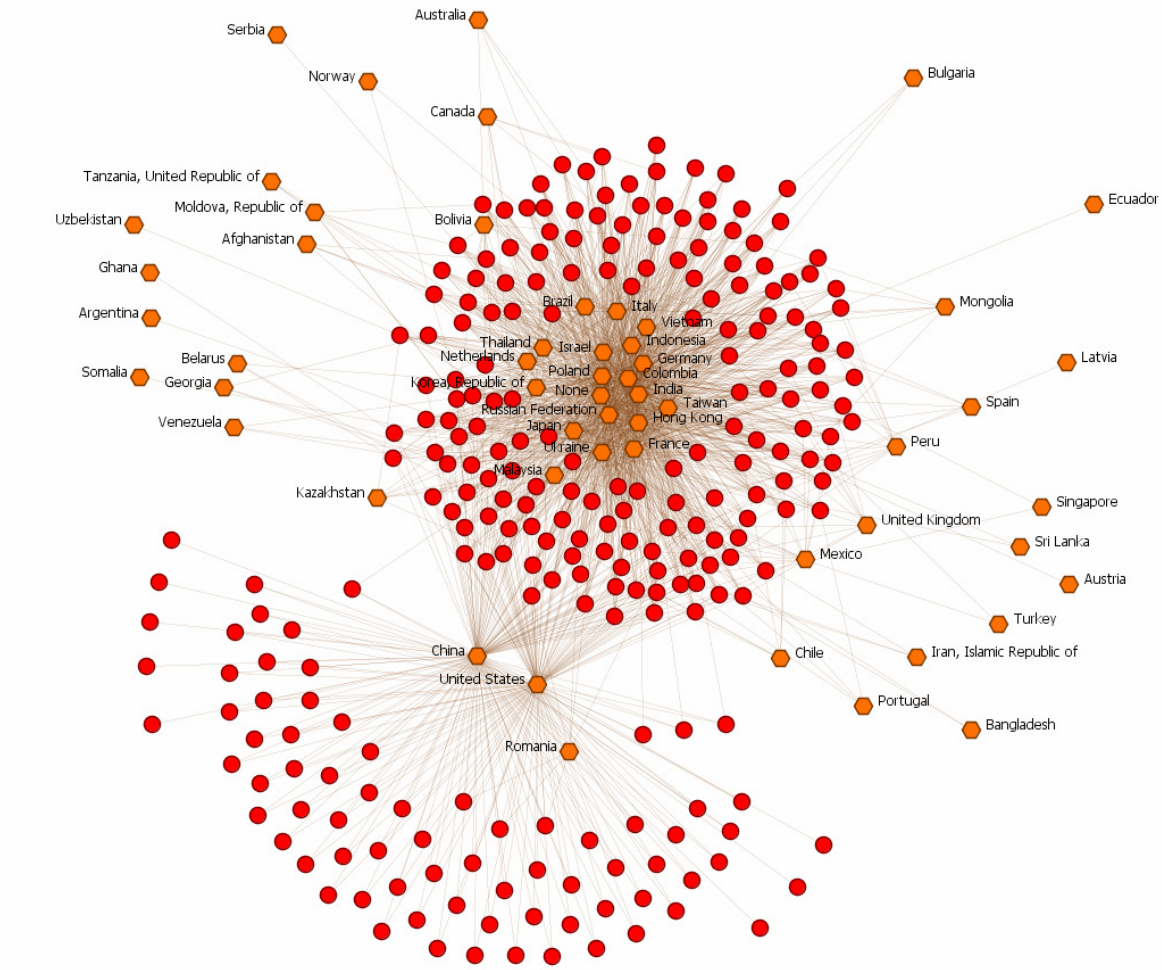
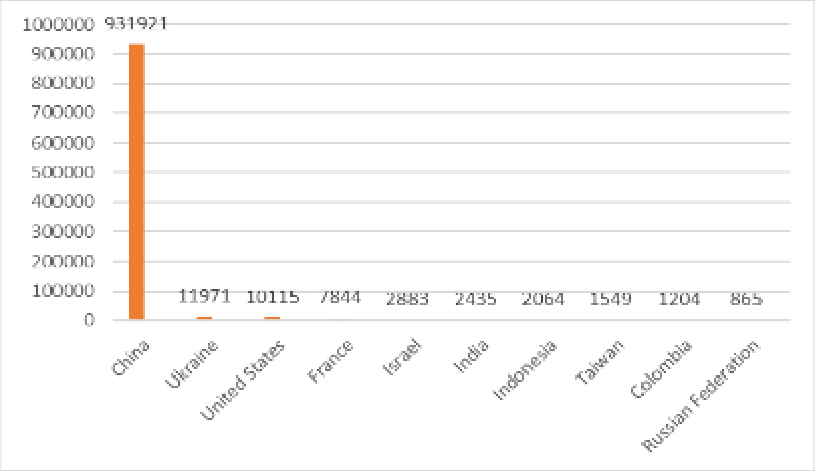


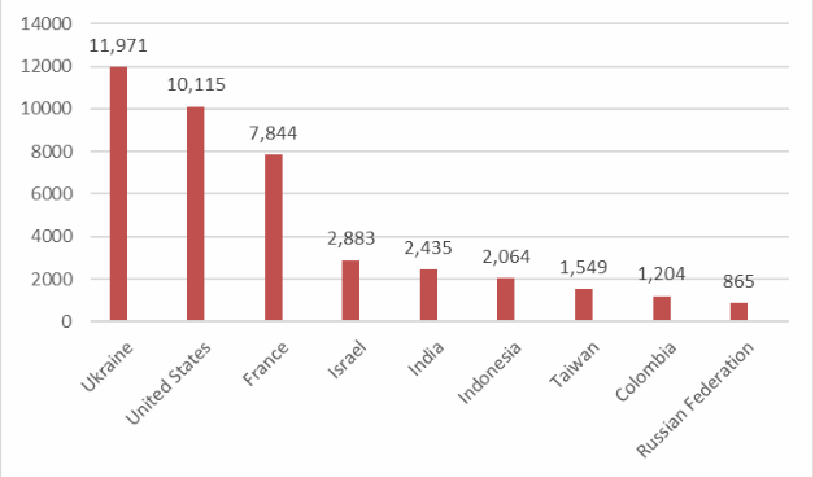
Figure 1 presents honeypots (circle), and attacks originated from 53 countries (hexagon) all over the globes. However, in a closer look, we analyzed who were the main players in this network. Graph 2a presents the distribution of the leading countries involved in BFA activity over the course of 36 days in late 2015. Graph 2b presents all countries except China, which is an outlier. Appendix B shows the activity of all 36 countries involved in BFA.

Graph 2. The distribution of BFAs and Sessions in the leading countries

Graph 2a – BFA distribution



Graph 2b. Sessions distribution

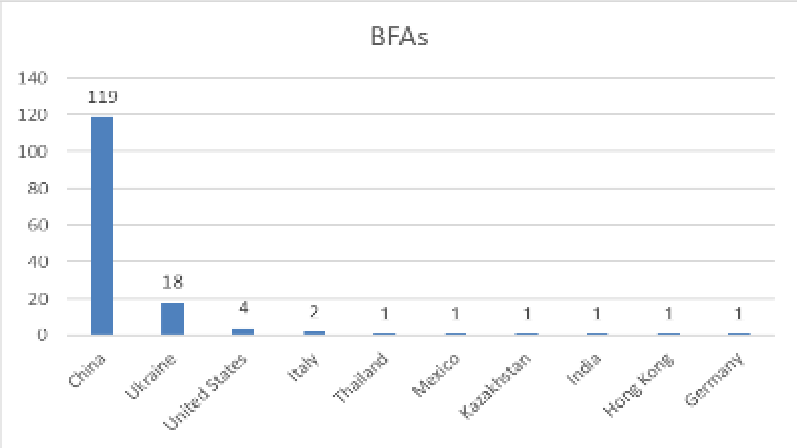


Our analysis revealed the attacks were highly dynamic. The network changed over time and was highly unstable. Since the HPs were open for a randomly-chosen period of 36 days, this period could indicate the overall pattern of attacks. We found that there was a slow start followed by a sharp increase in activity in the middle of the period, for about ten days. A sharp drop came right afterward, and moderate hacking activity continued after that. These findings resembled the findings in Yegneswaran et al. (2003).

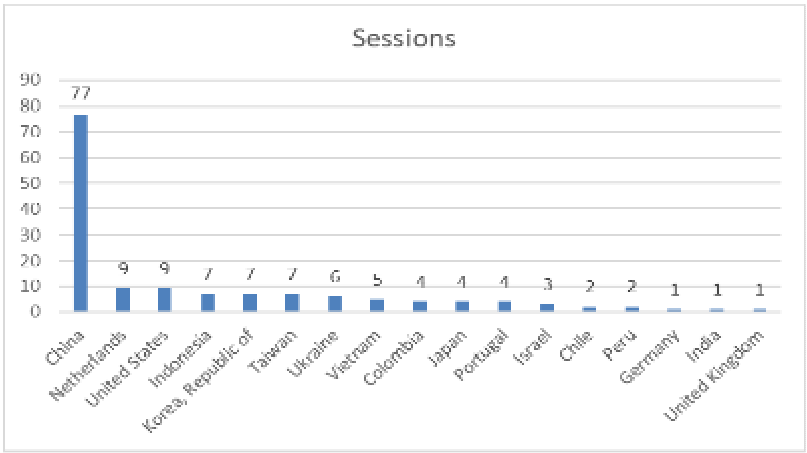
As mentioned before, our main interest was in the 149 OOTB Sessions, which were not the result of a known BFA on the specific HP. We gathered data concerning these Sessions and their potential originating BFA. Graph 3 presents the distribution of countries from which OOTB sessions were executed (3a) and the distribution of countries from

which the potential related BFAs originated (3b).Graph 3. The country-level distribution of BFAs and Sessions

Graph 3a – BFA distribution



Graph 3b – Sessions distribution



From Graph 3, we learn that China is the leading country in both activities. Although China's hacking activity is on the decline (Sanger, 2016), it is still the most powerful player in the hacking arena (Brownlee, 2015). When removing China from the dataset, Ukraine and the US are the second major players, and the third in importance are two European countries, the Netherlands and Italy, and East Asian countries – Indonesia, Hong Kong, North Korea, Taiwan, and Thailand.

Having identified the leading countries and the kinds of attacks originating from them, we turned to analyze the relations between the source countries of BFAs the source countries of Sessions. To do so, we built a network combining the BFAs and their possibly related 149 OOTB Sessions. The nodes of the network were countries, and the links were the possible transformations of hacked data. Figure 2 presents the network.

Figure 2. Data transfer from BFA to Session countries with and without China

Figure 2a. Data transfer from BFA to Session countries with China

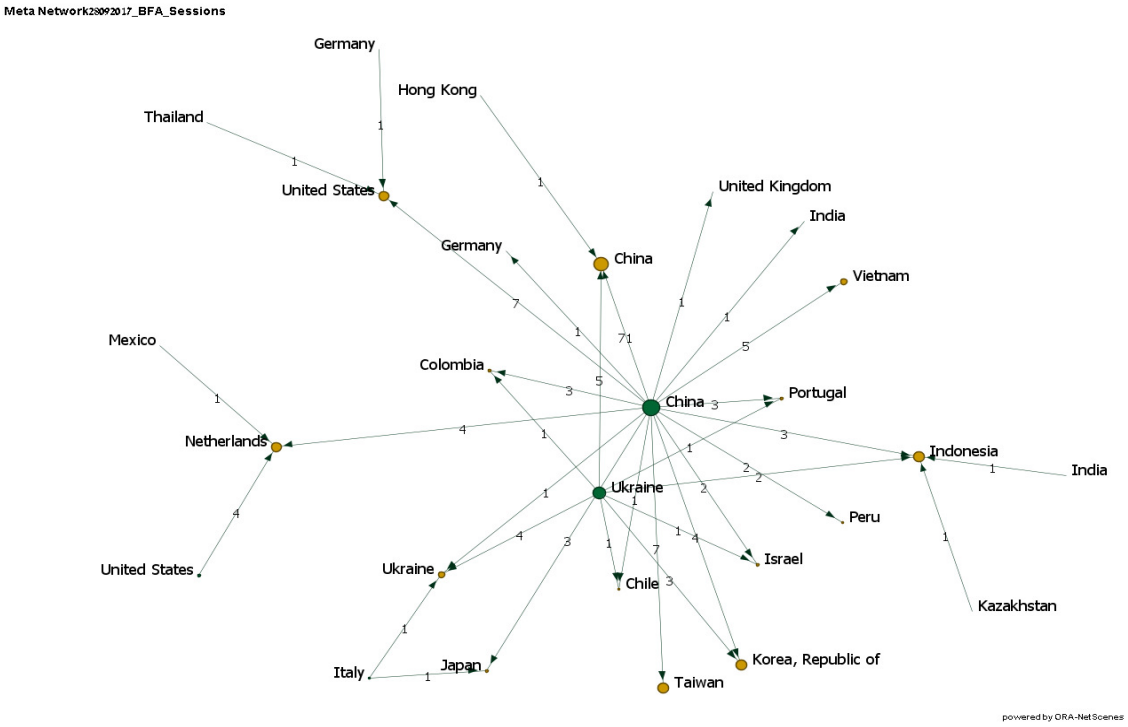
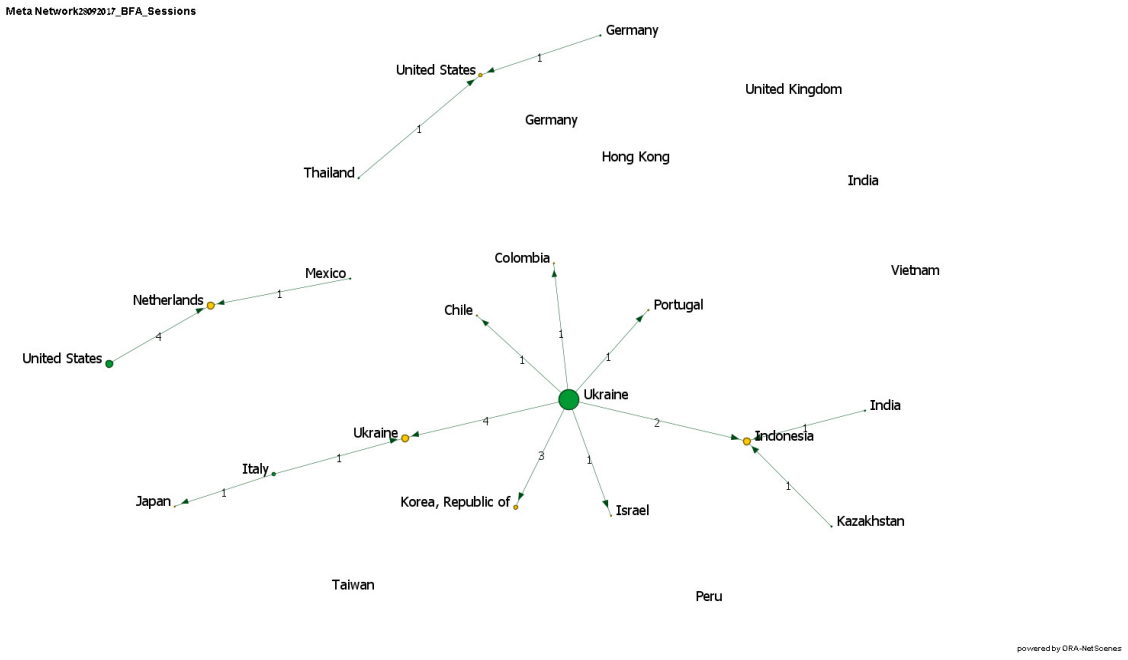


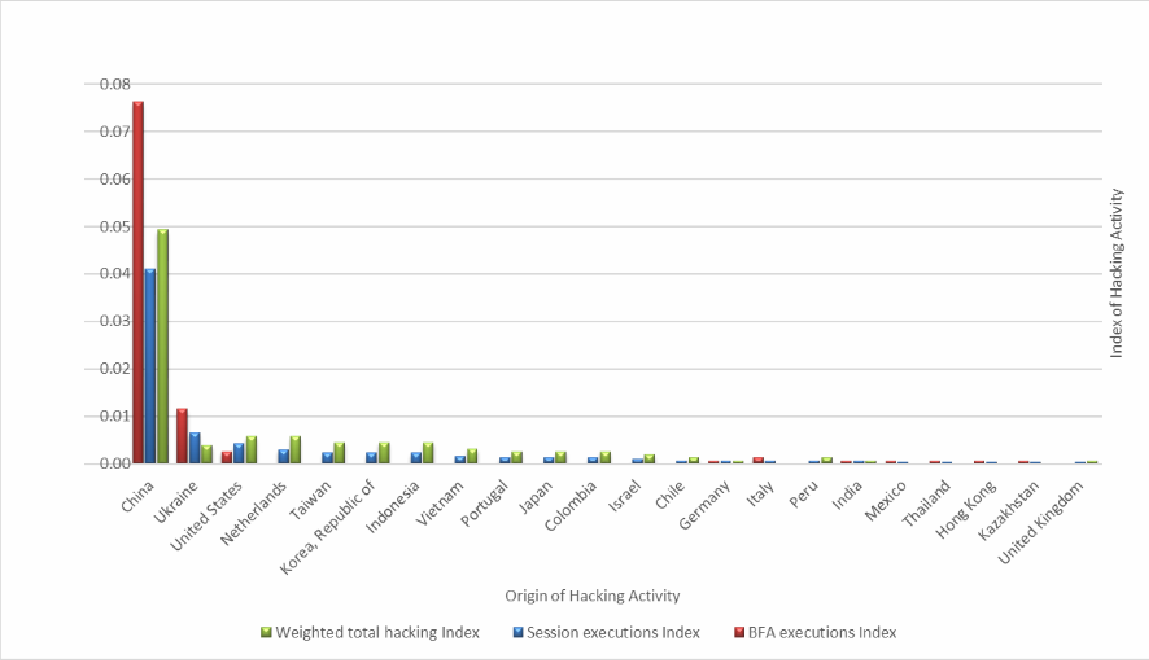
Figure 2b. Data transfer from BFA to Session countries without China



The BFA nodes are green, the Sessions nodes are yellow, and the values on the links represent the number of times we detected a possible data transfer from a BFA to a specific Session. For example, the Netherlands has initiated nine Sessions, linked to four BFAs from the US, four from China, and one from Mexico. The size of the nodes represents their total volume of activity. China appears twice in the figure since it is the primary node with the highest activity both as a BFA country and as a Session country and almost half (71 out of 149) of the attacks involve China as a BFA and Session country.

Figure 2b presents data transfer from BFAs to Sessions without China. The network disintegrates without Chinas, and although Ukraine is a leading player in the remaining network, the network as a whole, cannot function as such. To conclude the data analysis, we gathered for each country the aggregation of hacking activities executed using its IP addresses. Graph 5 presents the distribution of hacking activities.

Graph 5. Total Hacking index



Graph 5 makes it very clear that attacks coming from IPs in China play a huge role in both BFA and Session activities. Figure 1 demonstrates the critical role of Chinese IPs in the attacks and depicts the fact that all other attacks led by IPs from Ukraine and the US are secondary to China. (see also Appendix B). Attacks coming from China create and consume a substantial portion of the hacked data, and in this sense, they are the heart of the network and function as its core. Our findings are in line with previous research concerning the geography of hacking, though we have not found activities from Romania (Rechavi et al., 2015) or Bulgaria (Rughiniş & Rughiniş, 2014).

Discussion and Conclusion

Certain countries play unique roles as origins of BFA or Session attacks. Mapping the attacks has enabled us to depict the hacked access keys used in the sessions (H1). These OOTB sessions helped us create a country-level map of relations between BFAs coming from certain countries and Sessions coming from other countries. The topology of this network has a primary node (Chinese IPs), which is responsible for most of the BFA and Session activities. This central position in our network meets Mislove et al.'s (2007) definition of network core as a set of strongly connected nodes crucial for network connectivity (removing it breaks the network into many loose clusters). BFA hackers capture the access data and share it with the Session' countries.

We collected data concerning two-phase attacks and performed a unique analysis based on the process of attack. Connecting the attacks on the country level suggests there is a hidden market of hacked data. The evidence suggests a possible global market where valuable hacked data change hands (H2). Whether the hackers freely share or exchange it for other malicious skills is beyond the scope of this research. We argue that the network we depict has a core-periphery topology where a small number of IPs is responsible for the BFAs and sends the data to many less active IPs that use the data for Session activities (H3). This pattern, which divides the job between BFAs and Sessions, enables higher specialization and more efficient data transmission from core countries who execute to the periphery countries who execute Sessions.

Previous studies found the stolen-data market to be a network of global cyber criminals who operate in various shared ways to execute their individual interests (Holt & Smirnova, 2014). The current study suggests evidence that this stolen data is being used to attack computers all over the globe.

This result can inform a localized effort to put an end to most BFA hacking activities. We know from previous studies that the core's rich connection structure to all other countries enables its survival (Csermely et al., 2013) and that shutdown attempts will not affect its survival (Ablon et al., 2014, p. 36; Holt & Smirnova, 2014). Hence, we argue that this kind of topology poses a new opportunity for cyber security researchers, policymakers, and practitioners. Calculating the activity volume without looking at the whole (network) picture is not enough. Understanding the overall network topology, as well as the centrality of each country in the hacking network and its role, can focus the efforts in the right direction.

Nowadays, the correlations between strengthening cyber-policy enforcement and a positive change in the volume of cyber attacks are not clear (Kumar et al. 2016). We argue that countries must take different means, to tackle not only BFAs and their subsequent Session attacks separately, but the relations between the two as well. Knowing which hacking activity is central for each country, its role in the network and its relations with other countries is the first step towards a comprehensive solution. Approaching the problem by focusing on the links between the countries in addition to law enforcement in the countries themselves, thus preventing the data from traveling from BFA to Session hackers can be the new and refreshing point of view.

Limitations and Future Research

Hacking networks are dynamic. In 2006-2007, most hackers were from Russia, while by 2013, the USA hosted almost a fifth of the market, and another third originated from Ukraine, and Romania combined (Ablon et al. 2014, p. 26). Hence, our hacking topology may be related to the period in which we studied the network (2016). Second, we assumed that the usernames and passwords of the HPs were strong enough for hackers not to be able to guess them. If hackers managed to access the HP without obtaining credentials from another source, our experiment did not detect that. Third, the same hackers might change IPs and act one day as BFA hackers from China and the next as Session hackers from Chile. The above potential activity pattern is the reason that our study explored the IPs that were in use and not the people or the countries involved. Fourth, our algorithm, which matched BFAs and Sessions, is based on the assumption that a BFA hacker wants to get rid of the hacked credentials as soon as possible. A probabilistic model where the probability is calculated for every possible connection between BFAs and Sessions might provide a more nuanced perspective on the frequencies and volumes of data transfer and chart a different map of connected countries. This calculation should be based on big data of interactions, which we did not have. Finally, the scope of our experiment is limited to an academic institution in Israel. More institutions and different facilities (governmental, industrial) might evince different hacking topologies. Beyond addressing those limitations, future studies should not only collect more data and specify hacking network structures, but also assess ways of controlling and combating them.

Acknowledgments

This study was supported by the Israeli Ministry of Science, Technology, and Space (Grant No. 3-10888).

References

Ablon, L., Libicki, M. C., & Golay, A. A. (2014). *Markets for cybercrime tools and stolen data: Hackers' bazaar.* Santa Monica: Rand Corporation.

Abrahamson, E., & Rosenkopf, L. (1997). Social network effects on the extent of innovation diffusion: A computer simulation. *Organization Science*, *8*(3), 289-309.

Aral, S., & Walker, D. (2012). Identifying influential and susceptible members of social networks. *Science*, *337*(6092), 337-341.

Baños, R. A., Borge-Holthoefer, J., & Moreno, Y. (2013). The role of hidden influential in the diffusion of online information cascades. *EPJ Data Science 2* (6): 1-16.

Baum, J. A. C., Shipilov, A. V. & Rowley, T. J. (2003). Where do small worlds come from? *Industrial and Corporate Change*, *12*(4), 697.

Benjamin, V., & Chen, H. (2012, June). Securing cyberspace: Identifying key actors in hacker communities. In *Intelligence and Security Informatics (ISI), 2012 IEEE International Conference* (pp. 24-29). IEEE.

Berger-Wolf, T. Y., & Saia, J. (2006). A framework for analysis of dynamic social networks. *Proceedings of the 12th ACM SIGKDD*, 523-528.

Berthier, R., & Cukier, M. (2009). An evaluation of connection characteristics for separating network attacks. *International Journal of Security and Networks*, *4*(1), 110-124.

Borgatti, S. P., & Everett, M. G. (2000). Models of core/periphery structures. *Social Networks, 21*(4), 375-395.

Brownlee, L. (25 September 2015). New Report of Malicious Chinese Cyber Attack On A U.S. Government Agency. *Forbes.com*. Retrieved from Forbes.com/sites/lisabrownlee/2015/09/25/new-report-of-malicious-chinese-cyber-attack-on-a-u-s-government-agency/#517381d7309b

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.

Braha, D., & Bar-Yam, Y. (2006). From centrality to temporary fame: Dynamic centrality in complex networks. *Complexity*, 12(2), 59-63.

Cross, R., Borgatti, S. P., & Parker, A. (2001). Beyond answers: Dimensions of the advice network. *Social Networks, 23*(3), 215-235.

Csermely, P., London, A., Wu, L. Y., & Uzzi, B. (2013). Structure and dynamics of core/periphery networks. *Journal of Complex Networks, 1*(2), 93-123.

Dupont, B., Côté, A. M., Savine C., & DécaryHétu D. (2016) The ecology of trust among hackers, *Global Crime*, *17*(2), 129-151. doi: 10.1080/17440572.2016.1157480.

Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999, August). On power-law relationships of the internet topology. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, 262.

Farinholt, B., Rezaeirad, M., Pearce, P., Dharmdasani, H., Yin, H., Le Blond, S., McCoy, D., & Levchenko, K. (2017). "To catch a ratter: Monitoring the behavior of amateur dark comet rat operators in the wild". *Security and Privacy (SP) 2017, IEEE Symposium* (pp. 770-787).

Furnell, S. (2002). *Cyber crime: Vandalizing the information society*. Boston: Addison-Wesley.

Fossi, M., Johnson, E., Turner, D., Mack, T., Blackbird, J., McKinney, D. & Gough, J. (2008). Symantec report on the underground economy. Symantec Corporation, 51.

Goel, S. (2011). Cyberwarfare: connecting the dots in cyber intelligence. *Communications of the ACM, 54*(8), 132-140.

Gomez Rodriguez, M., Leskovec, J. & Krause, A. (2010, July). Inferring networks of diffusion and influence. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1019-1028). ACM.

Grabosky, P. (2004). The global dimension of cybercrime. *Global Crime*, *6*(1), 146-157.

Hill, S. A. & Braha, D. (2010). Dynamic model of time-dependent complex networks. *Physical Review E, 82*(4), 046105.

Grabosky, P. (2017). The Evolution of Cyber crime, 2006-2016. In T. J. Holt (ed.), *Cyber crime through an Interdisciplinary Lens* (pp. 15-36). New York: Routledge.

Holme, P., Edling, C. R., & Liljeros, F. (2004). Structure and time evolution of an internet dating community. *Social Networks, 26*(2), 155-174.

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(1), 891-903.

Holt, T. J., Chua, Y. T., & Smirnova, O. (2013). An exploration of the factors affecting the advertised price for stolen data. In *eCrime Researchers Summit* (eCRS)(pp. 1-10).

Holt, T. J., & Smirnova, O. (2014). *Examining the Structure, Organization, and Processes of the International Market for Stolen Data*. Washington, D.C. : US Department of Justice.

Isaksson, J., & Enbom, J. (2015). News Management in the Swedish School Sector. International Conference on Management, Leadership & Governance, 127.

Jordan, T., & Taylor, P. (1998). A sociology of hackers. *The Sociological Review*, *46*(4), 757-780.

Keith, M., Shao, B. & Steinbart, P. J. (2007). The usability of passphrases for authentication: An empirical field study. *International Journal of Human-Computer Studies, 65*(1), 17-28.

Kigerl, A. (2012). Routine activity theory and the determinants of high cybercrime countries. *Social Science Computer Review*, 30(4), 470-486.

Kigerl, A. (2016). Cyber Crime Nation Typologies: K-Means Clustering of Countries Based on Cyber Crime Rates. *International Journal of Cyber Criminology, 10*(2), 147.

Kshetri, N. (2005). Pattern of global cyber war and crime: A conceptual framework. *Journal of International Management*, 11(4), 541-562.

Kshetri, N. (2013). Cybercrimes in the Former Soviet Union and Central and Eastern Europe: current status and key drivers. *Crime, Law and Social Change*, 60(1), 39-65.

Kumar, S., Benigni, M., & Carley, K. M. (2016, November). The impact of US cyber policies on cyber-attacks trend. In *Intelligence and Security Informatics (ISI),* 2016 IEEE Conference on (pp. 181-186). IEEE.

Kumar, S., & Carley, K. M. (2016a, November). Understanding DDoS cyber-attacks using social media analytics. In *Intelligence and Security Informatics (ISI),* 2016 IEEE Conference on (pp. 231-236). IEEE.

Kumar, S., & Carley, K. M. (2016b, November). Approaches to understanding the motivations behind cyber attacks. In *Intelligence and Security Informatics (ISI),* 2016 IEEE Conference on (pp. 307-309). IEEE.

Kossinets, G. & Watts, D. J. (2006). Empirical analysis of an evolving social network. *Science, 311*(5757), 88.

Lakhani, K. R. & Wolf, R. G., (2003). Why Hackers Do What They Do: Understanding Motivation and Effort. *Free/Open Source Software Projects (September 2003). MIT Sloan Working Paper No. 4425-03*. Available at SSRN: https://ssrn.com/abstract=443040

Lau, R. Y., Xia, Y., & Li, C. (2012). Social Media Analytics for Cyber Attack Forensic. *International Journal of Research in Engineering and Technology (IJRET), 1*(4), 217-220.

Leukfeldt, E. R., Kleemans, E. R., & Stol, W. P. (2016). Cybercriminal networks, social ties and online forums: social ties versus digital ties within phishing and malware networks. *British Journal of Criminology*, 57(3), 704-722.

Leukfeldt, E. R., Kleemans, E. R., & Stol, W. P. (2017). Origin, growth and criminal capabilities of cybercriminal networks. An international empirical analysis. *Crime, Law and Social Change*, 67(1), 39-53.

Liben-Nowell, D., Novak, J., Kumar, R., Raghavan, P. & Tomkins, A. (2005). Geographic routing in social networks. *Proceedings of the National Academy of Sciences of the United States of America, 102*(33), 11623-11628.

Lickel, B., Rutchick, A. M., Hamilton, D. L. & Sherman, S. J. (2006). Intuitive theories of group types and relational principles. *Journal of Experimental Social Psychology*, *42*(1), 28-39.

Lu, Y., Luo, X., Polgar, M., & Cao, Y. (2010). Social network analysis of a criminal hacker community. *Journal of Computer Information Systems*, 51(2), 31-41.

Maimon, D., Alper, M., Sobesto, B. & Cukier, M. (2014). Restrictive deterrent effects of a warning banner in an attacked computer system. *Criminology*, *52*(1), 33-59.

Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P., & Bhattacharjee, B. (2007). Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement,* 29-42.

Morgan, D. L., Neal, M. B. & Carder, P. (1997). The stability of core and peripheral networks over time. *Social Networks*, *19*(1), 9-25.

Motoyama, M., McCoy, D., Levchenko, K., Savage, S., & Voelker, G. M. (2011, November). An analysis of underground forums. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement*, 71-80.

Odinot, G., Verhoeven, M. A., Pool, R. L. D., & De Poot, C. J. (2016). Cyber crime, organised crime and organised cybercrime in *the Netherlands: Empirical findings and implications for law enforcement*.

Rantala, R. (2008). *Bureau of Justice Statistics Special Report: Cyber crime against Businesses, 2005*. Washington, DC: Bureau of Justice Statistics.

Rechavi, A., Berenblum, T., Maimon, D., & Sevilla, I. S. (2015). Hackers topology matter geography: Mapping the dynamics of repeated system trespassing events networks. In Advances in *Social Networks Analysis and Mining (ASONAM), 2015 IEEE/ACM International Conference on* (pp. 795-804). IEEE.

Rombach, M. P., Porter, M. A., Fowler, J. H. & Mucha, P. J. (2014). Core-periphery structure in networks. *SIAM Journal on Applied Mathematics*, *74*(1), 167-190.

Rughiniş, C., & Rughiniş, R. (2014). Nothing ventured, nothing gained. Profiles of online activity, cyber-crime exposure, and security measures of end-users in *European Union*. *Computers & Security*, *43*, 111-125.

Salles-Loustau, G., Berthier, R., Collange, E., Sobesto, B. & Cukier, M. (2011, December). Characterizing attackers and attacks: An empirical study. In *Dependable Computing (PRDC), 17th Pacific Rim International Symposium on IEEE*, 174-183.

Sanger, D. E. (June 20, 2016). Chinese Curb Cyberattacks on U.S. Interests, Report Finds. *NewYorkTimes.Com*. Retrieved from Nytimes.com/2016/06/21/us/politics/china-us-cyber-spying.html?_r=0

SANS Institute (2007). *SANS Top-20 2007 Security Risks* (2007 Annual Update). Retrieved from http://www.eweek.com/c/a/Security/SANS-Top-Internet-Security-Risks-of-2007.

Shackelford, S. (2009). From nuclear war to net war: analogizing cyber-attacks in international law. *Berkley Journal of International Law (BJIL)*, *25*(3).

Spitzner, L. (2003a). The honeynet project: Trapping the hackers. *IEEE Security & Privacy*, *1*(2), 15-23.

Spitzner, L. (2003b). Honeypots: Tracking hackers: Vol. 1. Reading, MA: Addison-Wesley.

Symantec, (2016). *ISTR: Internet Security Threat Report* Volume 21 (April 2016). Retrieved from Symantec.com/content/dam/symantec/docs/reports/istr-21-2016-en.pdf.

Testa, A., Maimon, D., Sobesto, B., & Cukier, M. (2017). Illegal roaming and file manipulation on target computers: Assessing the effect of sanction threats on system trespassers' online behaviors. *Criminology & Public Policy*, *16*, 689–726.

Viswanath, B., Mislove, A., Cha, M. & Gummadi, K. P. (2009). On the evolution of user interaction in Facebook. *Proceedings of the 2nd ACM Workshop on Online Social Networks*, 37-42.

Waldrop, M. M. (2016). How to hack the hackers: The human side of cybercrime. *Nature*, *533*(7602). Retrieved from https://www.nature.com/news/how-to-hack-the-hackers-the-human-side-of-cybercrime-1.19872.

Watts, D. J. (1999). Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*, *105*(2), 493-527.

Watts, D. J. (2004). The "new" science of networks. *Annual Review of Sociology*. *30*, 243-270.

Whittaker, S., Isaacs, E. & O'Day, V. (1997). Widening the net: Workshop report on the theory and practice of physical and network communities. *SIGCHI Bulletin, 29*, 27-30.

Wilson, T., Maimon, D., Sobesto, B., & Cukier, M. (2015). The effect of a surveillance banner in an attacked computer system: Additional evidence for the relevance of restrictive deterrence in cyberspace. *Journal of Research in Crime and Delinquency*, 829-855.

Yang, J., & Leskovec, J. (2014). Overlapping communities explain core-periphery organization of networks. *Proceedings of the IEEE, 102*(12), 1892-1902.

Yegneswaran, V., Barford, P. & Ullrich, J. (2003). Internet intrusions: Global characteristics and prevalence. *ACM SIGMETRICS Performance Evaluation Review*, *31*(1), 138-147.

Yip, M., Shadbolt, N., & Webber, C. (2012). Structural analysis of online criminal social networks. At *2012 IEEE International Conference on Intelligence and Security Informatics*, United States. pp. 60-65.

Young, R., Zhang, L., & Prybutok, V. R. (2007). Hacking into the minds of hackers. *Information Systems Management*, *24*(4), 281-287.

Appendix A: BFA and Session Distribution

BFAs Sessions	Chile	China	Germany	Hong Kong	India	Italy	Kazakhstan	Mexico	Thailand	Ukraine	United States	Total
Chile	0	1	0	0	0	0	0	0	0	1	0	**2**
China	0	71	0	1	0	0	0	0	0	5	0	**77**
Colombia	0	3	0	0	0	0	0	0	0	1	0	**4**
Germany	0	1	0	0	0	0	0	0	0	0	0	**1**
Indonesia	0	3	0	0	1	0	1	0	0	2	0	**7**
Israel	0	2	0	0	0	0	0	0	0	1	0	**3**
Japan	0	3	0	0	0	1	0	0	0	0	0	**4**
Korea, Republic of	0	4	0	0	0	0	0	0	0	3	0	**7**
India	0	1	0	0	0	0	0	0	0	0	0	**1**
Netherlands	0	4	0	0	0	0	0	1	0	0	4	**9**
Peru	0	2	0	0	0	0	0	0	0	0	0	**2**
Portugal	0	3	0	0	0	0	0	0	0	1	0	**4**
Taiwan	0	7	0	0	0	0	0	0	0	0	0	**7**
Ukraine	0	1	0	0	0	1	0	0	0	4	0	**6**
United Kingdom	0	1	0	0	0	0	0	0	0	0	0	**1**
United States	0	7	1	0	0	0	0	0	1	0	0	**9**
Vietnam	0	5	0	0	0	0	0	0	0	0	0	**5**
Total	**0**	**119**	**1**	**1**	**1**	**2**	**1**	**1**	**1**	**18**	**4**	**149**

Appendix B: BFA Distribution of all 939K attacks

IP Country	BFA		IP Country	BFA
China	931,921		Japan	52
Ukraine	11,971		Chile	50
United States	10,115		Peru	49
France	7,844		Spain	41
Israel	2,883		Mexico	38
India	2,435		Kazakhstan	37
Indonesia	2,064		United Kingdom	29
Taiwan	1,549		Australia	28
Colombia	1,204		Romania, Malaysia	23
Russian Federation	865		Bolivia	21
Italy	798		Portugal	16
Germany	486		Afghanistan	14
Hong Kong	422		Mongolia	13
None	235		Bangladesh, Iran	8
Canada	231		Tanzania, Norway, Georgia	7
Poland	167		Belarus, Moldova,	5
Thailand	148		Bulgaria, Venezuela	4
Brazil	107		Sri Lanka, Turkey	3
Korea, Republic of	71		Ecuador, Ghana, Serbia, Singapore	2
Netherlands	62		Argentina, Austria, Latvia, Somalia, Uzbekistan	1
Vietnam	60			