After the Breach: The Monetization and Illicit Use of Stolen Data

Testimony by
Nicolas Christin, Ph.D.
Associate Research Professor
School of Computer Science, Institute for Software Research
College of Engineering, Department of Engineering and Public Policy
Carnegie Mellon University

Before the
Subcommittee on Terrorism & Illicit Finance
Committee on Financial Services
U.S. House of Representatives

The Honorable Stevan Pearce, Chairman
The Honorable Ed Perlmutter, Ranking Member

March 15, 2018

Chairman Pearce, Ranking Member Perlmutter, Members of the Subcommittee, thank you for hosting this important hearing today, and for giving me the opportunity to submit this testimony.

My name is Nicolas Christin. I am an associate research professor at Carnegie Mellon University, jointly appointed in the School of Computer Science and in the Department of Engineering and Public Policy. I am a computer scientist by training. My research focuses on computer security, and, for the better part of the last decade, I have been studying online crime. I have been focusing on the interface between technical and socio-economic aspects of computer abuse. In particular, I have conducted, with my research group, a series of measurement studies on online anonymous "dark web" marketplaces,[1,2] in an attempt to better understand the potential economic impact of these markets, including their role as retail channels for stolen data. This is the topic at hand today.

Monetizing stolen credentials and the asymmetry between societal costs and criminal revenue

The existence of an online market for stolen credentials – e.g., financial or personal data – can be traced back to the early 1990s. At the time, crooks were using dial-up forums (BBSes) to sell illicitly acquired credentials (e.g., credit card information directly stolen from postal mailboxes). As the Internet and the World Wide Web gradually rose to prominence in the mid-1990s, thieves started stealing credentials online (e.g., through "phishing" scams, tricking victims into revealing their financial information to miscreants), and dial-up forums moved to online chatrooms, primarily using the Internet Relay Chat protocol (IRC).

In an article based off seven months of data collected in early 2006,[3] Franklin et al. provided what is widely regarded as the first quantitative academic description of the online market for stolen credit card numbers. At the time, IRC chatrooms were still a very popular way for sellers and prospective buyers to transact. A vast majority of these chatrooms were open to the general public: All that was needed was freely available software (IRC clients) and the name and online location of the chatrooms used as marketplaces for purloined credentials.[4] This information could be easily found through simple web searches.

---

[1] Christin, Nicolas. "Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace." In *Proceedings of the 22nd International World Wide Web Conference (WWW'13)*, pages 213-224. Rio de Janeiro, Brazil. May 2013.
[2] Soska, Kyle and Nicolas Christin. "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem." In *Proceedings of the 24th USENIX Security Symposium (USENIX Security'15)*, pages 33-48. Washington, DC. August 2015.
[3] Franklin, Jason, Adrian Perrig, Vern Paxson, and Stefan Savage. "An inquiry into the nature and causes of the wealth of Internet miscreants." In *ACM conference on Computer and Communications Security*, pp. 375-388. 2007.
[4] Opening an IRC chatroom is an easy task, that requires little technical background. As a result, there exist IRC chatrooms dedicated to pretty much any topic one can imagine.

Franklin et al. described that the goods for sale included "bank logins and passwords, PayPal accounts, credit cards, and social security numbers (SSNs)."[5] Franklin et al. estimated that approximately 87,000 potentially valid credit card numbers were advertised in the chatrooms they monitored over their seven months of study, and concluded that the overall revenue for the market for stolen financial credentials sold on the chatrooms they monitored was somewhere between $37M and $95M.

In a subsequent 2013 study, Anderson et al. attempted to dimension the overall cost of cybercrime.[6] Extrapolating from estimates they produced for the United Kingdom, they projected that online card fraud (attempts to use stolen card numbers online, for instance, on electronic commerce websites) probably cost around $4.2B per year.

Anderson et al. caution that their projections "should be interpreted with utmost caution,"[7] and Franklin et al.'s study focuses on one specific distribution channel (IRC chatrooms). Furthermore, six years separate both studies. Nevertheless, the very strong disparity between both estimates, which differ by two orders of magnitude, suggests that the revenue criminals generate from the sale of stolen data is dwarfed by the costs data breaches impose on society. Later in this testimony, by considering more recent data, we will see that this asymmetry still exists today.

Evolution of the distribution channels and service professionalization

Distribution channels for stolen credentials have evolved in the past two decades. IRC chatrooms, while accessible to anybody with the proper software, only offered a relatively rudimentary text-based medium, primarily attractive to people with at least a modest amount of technological expertise. Easier to use web forums and websites with search functionality and better customer service thus became increasingly prominent since the early 2000s. These websites are generally referred to as "carding forums." Notorious examples include "carder.su," "shadowcrew.cc," among others. Original carding forums featured, in particular, forum moderators and embryonic reputation systems whose purpose was to provide better guarantees to prospective buyers and sellers. Each of these forums was run by a group of loosely affiliated criminals.

Business models also became increasingly complex. While a large number of vendors were simply selling credentials, some miscreants also started advertising "money mule recruitment services," destined to help with the transfer of funds acquired from stolen credentials to overseas accounts, or "confirmation services," which acted as external verification services to test the quality of the credentials being offered.

In short, these web-based distribution channels were designed to facilitate the sale and purchase of stolen data on a larger scale, by less sophisticated actors. Similar to industrial supply chains, the market for stolen data started to show increased specialization among its actors. Technically-savvy actors were in charge of procuring the "raw" data, i.e., causing the data breaches;

---

[5] Franklin et al., "An inquiry into the nature and causes of the wealth of Internet miscreants."
[6] Anderson, Ross, Chris Barton, Rainer Böhme, Richard Clayton, Michel J.G. van Eeten, Michael Levi, Tyler Moore, and Stefan Savage. "Measuring the cost of cybercrime." In *The economics of information security and privacy*, pp. 265-300. Springer, Berlin, Heidelberg, 2013.
[7] Ibid.

others were in charge of "commoditizing" these data, by breaking then down in lots suitable for individual resale; yet others were providing services surrounding stolen data (mule services, money laundering tutorials) without directly interacting with the data themselves.

This increased professionalization of the service, and the evolution of the distribution channels for stolen data continued with the next generation of retail channels for stolen data: "dark web" marketplaces.

The "dark web," online anonymous marketplaces, and cryptocurrencies

The World Wide Web can be fundamentally split between pages and websites that are indexed by search engines ("surface web"), and pages and websites that are not ("deep web"). While the deep web is thought to far exceed in size the surface web, most of the contents of the deep web is far from sinister. It includes, for instance, internal company pages, configuration pages (e.g., of properly configured home routers), and a large number of social network pages.

A very small portion of the deep web constitutes the "anonymous web," or "dark web."[8] All computers connected to the Internet are assigned an "IP address."[9] For instance, it is public knowledge that the computer with address 128.237.152.41 sits at Carnegie Mellon University. By looking up IP addresses, website operators and Internet Service Providers can thus easily determine who is browsing their websites, and likewise, one can usually learn on which server a given website is running. Websites in the anonymous web, however, are only accessible using freely available special-purpose software (e.g., Tor[10] or i2p[11]). This special purpose software allows the individual browsing the web to conceal their IP address, which is useful to bypass local censorship, maintain anonymity (often helpful for intelligence sources, or online investigators), or even, more mundanely, circumvent web tracking by online advertisers. The same software also allows webservers to conceal the IP address of the physical server on which they run if they so choose, yielding "end-to-end" anonymity: no one knows where the server or its visitors are located.

Starting in the mid-2000s, a number of forums promoting illicit contents started to appear in the anonymous web. These forums remained mostly confidential at the time, but the invention of the Bitcoin payment system in late 2008 drastically changed the picture. Until then, paying for illicit goods or services online was rather cumbersome: credit card payments and regular ACH transfers are very traceable and were thus a very poor fit for engaging in illicit transactions. Most miscreants instead relied on online payment systems like WebMoney or Liberty Reserve. However, those were

[8] We will refrain from using the "dark web" moniker, which is actually very confusing, given that most of these web sites are publicly accessible as long as the right software is used.
[9] The IP address may be public, in which case the computer is globally reachable, or private, in case a special purpose device with a public IP address, sitting on the same network as the computer, is required to allow the computer to connect to the Internet.
[10] Dingledine, Roger, Nick Mathewson, and Paul Syverson. Tor: The second-generation onion router. In Proceedings of the 13th USENIX Security Symposium, August 2004.
[11] I2P: The internet invisible project. http://www.geti2p.net.

also potentially problematic, as they were being run by centralized entities that could be pressured to intervene in cases of illicit activity or face legal consequences.[12] Bitcoin, on the other hand, is fully decentralized and pseudonymous, thereby offering superior privacy guarantees to its users, compared to these other, previously established payment systems.

Building on these technological developments, in February of 2011, a website called "Silk Road" opened on the anonymous web. Silk Road was the first to combine the network-level anonymity properties provided by the Tor network, with the superior privacy guarantees offered by the Bitcoin payment system, to offer a fully functional "online anonymous marketplace." Silk Road itself did not sell any product *per se*, but provided a venue where buyers and vendors could interact with each other anonymously, and with a certain level of trust. Similar to traditional electronic commerce marketplaces like eBay or the Amazon Marketplace, Silk Road offered a feedback-based review system, through which buyers could rate sellers (and, internally, sellers could also rate buyers, although this information was not public). Because Silk Road had very lax rules on what could and could not be listed, it very quickly become a haven for illicit activity.[13]

Buyers had to leave a public review for each purchase they made on the site. As a result, one could relatively precisely estimate the total number of sales taking place, and the associated revenue of the entire marketplace. Our original study[14] estimated that in the first months of 2012, Silk Road was on track to generate approximately $15M of yearly revenue and derived an overwhelming fraction of this revenue from drug sales – narcotics and prescription drugs.

Silk Road grew significantly in late 2012 and early 2013, reaching a revenue of more than $350,000 per day in the Summer of 2013.[15] The location of the server and identity of its operator were eventually discovered by authorities, which closed the site and arrested its owner in November 2013. Numerous "copycat" marketplaces immediately appeared in its stead (including a popular site named "Silk Road 2.0"), using a similar combination of network-level anonymization technologies and pseudonymous cryptocurrencies.

Most of these online anonymous marketplaces use English as the primary language. However, operators, vendors and buyers can be located anywhere in the world. Among notorious cases, Silk Road and Silk Road 2.0 were reportedly operated by U.S. nationals; AlphaBay was allegedly run by a Canadian citizen residing in Thailand; Hansa Market, by two German nationals; and Sheep Marketplace, by a Czech individual. Marketplaces in other languages do exist but tend to

[12] Liberty Reserve was eventually shut down by U.S. authorities under the Patriot Act, and its founder charged with and convicted of money laundering.
[13] Christin. "Traveling the Silk Road."
[14] Ibid.
[15] Soska and Christin. "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem."

be generally smaller. A notable exception was the Russian Anonymous Marketplace (RAMP), which was active for a few years and was a fairly sizeable site.[16]

By 2015, through research done at Carnegie Mellon, and partially supported by the Department of Homeland Security Science and Technology Directorate, we estimated that the ecosystem of online anonymous marketplaces was generating a revenue in excess of half a million US dollars per day, that is, approximately $200M a year.[17]

As this criminal ecosystem was growing, law enforcement orchestrated a number of takedowns. In particular, authorities managed to shut down a number of online anonymous marketplaces in 2014 (including Silk Road 2.0, and other less prominent bazaars) during a joint effort between U.S. and European law enforcement agencies dubbed "Operation Onymous." We nevertheless observed that the online anonymous marketplace ecosystem, as a whole, appeared highly resilient to such disruptions. When a leading marketplace is taken down (or absconds with its customers' money), consumers appear to move relatively quickly to another marketplace, and the long-term impact of online anonymous marketplace takedowns has yet to be convincingly observed.[18]

For instance, immediately after Operation Onymous, sales on the Evolution and Agora marketplaces, which had not been affected by law enforcement intervention, ramped up significantly. When these marketplaces disappeared in 2015, the AlphaBay marketplace, which had started operating in December 2014, rose to prominence, and went on collect a revenue exceeding $800,000 per day in early 2017.[19]

Remarkably, both Evolution and AlphaBay (and other lesser known online anonymous marketplaces) initially started as carding forums, before expanding their businesses to other areas. For instance, close to 50% of all revenue on AlphaBay in its first couple of months of existence (at a time when overall revenue was still low) came from "digital goods,"[20] a category that encompasses fraudulently obtained financial credentials, forged documents, hacking kits, and so forth.

At their peak, both Evolution and AlphaBay derived a majority of their revenue from commissions on drug sales. Nevertheless, sellers of illicitly acquired data had certainly taken notice that online anonymous marketplaces were a valuable distribution channel for their products.

---

[16] Greenberg, Andy. "How a Russian dark web drug market outlived the Silk Road (and Silk Road 2)." *Wired*. November 2014. https://www.wired.com/2014/11/oldest-drug-market-is-russian/
[17] Soska and Christin. "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem."
[18] Ibid.
[19] European Monitoring Centre for Drugs and Drug Addiction and Europol. "Drugs and the darknet: Perspectives for enforcement, research and policy." EMCDDA–Europol Joint publications, Publications Office of the European Union, Luxembourg. November 2017.
[20] Ibid.

Digital goods and online anonymous marketplaces

An overwhelming majority (more than 80%) of the revenue of the online anonymous marketplaces we monitored comes from the sale of narcotics or prescription drugs. However, a non-negligible portion of sales (~5-10%) concerns "digital goods," and this proportion appears to be growing (albeit slightly) over time.

In collaboration with researchers based at Delft Technical University in the Netherlands, we recently performed a thorough investigation of the market for digital goods in online anonymous marketplaces. Using data we had collected between 2011 and 2017, from most of the major online anonymous marketplaces operating during that time interval, we were able to observe the following.

First, the largest type of digital goods listings we observed (approximately 12,000 out of roughly 44,000 total offerings in the digital goods category) were "cash-out" schemes. These cash-out schemes primarily include 1) synthetic credit card numbers not associated with any real account, but that would pass rudimentary automated validity checks—those are usually not harmful to any specific individual, 2) "fullz," that denote comprehensive records, pairing for instance stolen credit card numbers, with the associated CVV codes, and in some cases the social security number or date of birth of the legitimate owner, and 3) various types of guides, including money laundering tutorials (e.g., how to recruit money mules). A smaller number of listings were for 4) bank and financial account credentials (e.g., PayPal logins) and 5) money laundering services (e.g., "Bitcoin deals," or cash payouts, such as vendors offering cash in the mail in exchange for Bitcoin).

In other words, similar to offerings on dedicated carding forums, we see a fairly large range of monetization techniques, ranging from the sale or purloined data, to integration of illicit profits in the legitimate financial system. We do not see, however, very high-value data leaks such as the OPM breach data. Highly valuable goods such as government personnel data are indeed more likely to be of value to nation-states (e.g., for intelligence purposes) than they are for financial purposes.

Second, the estimated revenue generated by cash-out schemes on anonymous online marketplaces remained low – altogether, digital goods, of which cash-out schemes are a subset, represented approximately a total revenue of $29M. We caution this is a conservative estimate, as 1) we measured only the most prominent generalist online anonymous marketplaces, and 2) due to technical difficulties inherent to large-scale data collection, the data we have may not be fully complete. Despite these caveats, these revenue numbers are strikingly low compared to the societal costs of breaches (e.g., costs to bank to reissue cards,[21] costs to individuals to restore their credit in case of a breach). This confirms the aforementioned trends observed in the late 2000s-early 2010s, which predated the use of online anonymous marketplaces as a retail channel for stolen data.

[21] Graves, James, Alessandro Acquisti and Nicolas Christin. Should Credit Card Issuers Reissue Cards in Response to a Data Breach?: Uncertainty and Transparency in Metrics for Data Security Policymaking. To appear in *ACM Transactions on Internet Technology*. 2018.

This overall low revenue can be partially attributed to the low value of most of the goods sold on online anonymous marketplaces. Overall, the median price for a cash-out listing is only around $60. Credit card numbers (without additional information) typically only sell for a few dollars apiece, and are often sold in lots (e.g., 100 Visa cards from UK banks); comprehensive records, including social security numbers, usually go for $100 or less.

The low retail value of stolen data is partially due to improvements in fraud detection, and proactive blacklisting of financial credentials. These defenses make a significant fraction of credit card numbers and other credentials being sold online at a retail level likely to be worthless. This also explains why sellers often package stolen data in lots. Interested buyers purchase lots with the hope that some of the information purchased is not yet blacklisted, and/or would give them access to sizeable revenue. But, by and large, for a buyer, engaging in such transactions is akin to purchasing a lottery ticket from a less-than-reputable outlet.

Third, we found that, since 2014, on any given day, approximately 3,000 to 4,000 vendors were simultaneously actively offering cash-out listings. While different vendors may have been active at different times, the overall population of active vendors increased slightly between 2014 and 2017. More interestingly, the top 10% of vendors were responsible for 80% of the revenue. In other words, the business of selling purloined data (or services surrounding it) appears to be truly profitable only for the most successful criminals on these forums.

These most successful vendors manage to establish a solid reputation, and either have a steady influx of goods to sell, or a large number of goods to sell over relatively short periods (a couple of months). We see few "bulk" sales, which hints that large breaches are first broken down into smaller lots, and those smaller lots are subsequently sold independently. This commoditization process, unfortunately, generally does not play out in public marketplaces. This makes traceability challenging: Retail-level vendors on online anonymous marketplaces rarely advertise the provenance of the credentials they are offering, and in fact, may not even know how these credentials were acquired in the first place.

Summary and moving forward

With the increased digitization of records, online financial fraud and data breaches are becoming a critical problem. Our recent measurement studies of online anonymous marketplaces, or "dark web markets," allow us to get a relatively precise idea of some of the business models in use, and of the economics of stolen data.

We find that revenue generated by criminals engaged in monetizing these breaches pales in comparison to the potential costs of the remedies. For instance, stolen credit card and identity details are often sold in lots, at low retail prices. However, the owner of a stolen banking credential

has to invest potentially considerable time, effort and money to attempt to repair the damage incurred by the theft.

There is also a noticeable level of activity in the sale of "services" surrounding data breaches, such as verification of the data, or money laundering and integration of the profits generated. Further, these marketplaces are international in nature, and, even if certain actors might be identified (e.g., through undercover operations), jurisdiction issues may complicate prosecution and/or arrest of individual vendors.

In addition, the online anonymous marketplace ecosystem as a whole has shown strong resiliency to law enforcement takedowns. Shutting down one or more marketplaces has so far mostly seemed to result in criminals moving to different marketplaces, and long-term impacts on the ecosystem are uncertain. Takedowns also may potentially lead some of the members of that ecosystem to move their activity to less publicly observable forums (e.g., private vendor forums).

We also observe that most of the revenue is generated by a small fraction of all criminals. This is a constant in cyber-crime, beyond stolen data markets.[22,23] A few highly successful criminals seem to attract relatively large numbers of amateurs that do not profit much, if at all, from their activities. Unsuccessful retail-level vendors nevertheless contribute to the overall problem, by making the market for stolen data larger, and more complex.

All of these findings indicate that focusing on preventing breaches from happening in the first place is likely to be more economically efficient than attempting to disrupt distribution channels, or recovering from a data breach once it has happened.

Finally, measurements of "dark web" marketplaces focus on the retail end of the stolen data ecosystem, and are thus an imperfect signal, particularly when it comes to tracing stolen data back to a specific breach. Nevertheless, such measurements give us important information on the health and evolution of the market for illicitly acquired data, and on the monetization techniques in use.

It is thus important to continue supporting these documentation efforts, so that we can decipher the evolving economic and business models that support stolen data markets. Indeed, understanding the criminals' economic incentives is critical to determine which combination of defensive strategies (technical, legal, economic) are most likely to disrupt criminal business models, thereby creating adverse incentives for miscreants to engage in these activities in the first place.

---

[22] Clayton, Richard, Tyler Moore, and Nicolas Christin. Concentrating Correctly on Cybercrime Concentration. In the *Proceedings (online) of the 14th Workshop on Economics of Information Security (WEIS 2015)*. Delft, Netherlands. June 2015.
[23] Leontiadis, Nektarios. *Structuring Disincentives for Online Criminals*. PhD Thesis. Carnegie Mellon University. August 2014.