Bermudez Villalva *et al. Crime Sci* (2018) 7:17
https://doi.org/10.1186/s40163-018-0092-6

Crime Science

RESEARCH

Open Access

Under and over the surface: a comparison of the use of leaked account credentials in the Dark and Surface Web

Dario Adriano Bermudez Villalva[1*], Jeremiah Onaolapo[2], Gianluca Stringhini[2] and Mirco Musolesi[3]

Abstract

The world has seen a dramatic increase in cybercrime, in both the Surface Web, which is the portion of content on the World Wide Web that may be indexed by popular engines, and lately in the Dark Web, a portion that is not indexed by conventional search engines and is accessed through network overlays such as the Tor network. For instance, theft of online service credentials is an emerging problem, especially in the Dark Web, where the average price for someone's online identity is £820. Previous research studied the modus operandi of criminals that obtain stolen account credentials through Surface Web outlets. As part of an effort to understand how the same crime unfolds in the Surface Web and the Dark Web, this study seeks to compare the modus operandi of criminals acting on both by leaking Gmail honey accounts in Dark Web outlets. The results are compared to a previous similar experiment performed in the Surface Web. Simulating operating activity of criminals, we posted 100 Gmail account credentials on hidden services on the Dark Web and monitored the activity that they attracted using a honeypot infrastructure. More specifically, we analysed the data generated by the two experiments to find differences in the activity observed with the aim of understanding how leaked credentials are used in both Web environments. We observed that different types of malicious activity happen on honey accounts depending on the Web environment they are released on. Our results can provide the research community with insights into how stolen accounts are being manipulated in the wild for different Web environments.

Keywords: Cybercrime, Account compromise, Measurement

Introduction

Online services are popular among individuals and companies for personal, business, or academic purposes. Normally, users are required to create personal accounts which are protected by private credentials. A large amount of sensitive data is stored within such personal accounts and some of them, such as webmail accounts are primarily used to access further services. Consequently, users are victims of data theft by cybercriminals stealing account credentials for their own benefit. According to the Crime Survey for England and Wales Office (2016), one out of ten adults has been victim of some kind of personal data theft.

Cybercriminals use social engineering techniques such as phishing and spear phishing (Lynch 2005), malware on victims' devices (Stone-Gross et al. 2009) and also exploiting vulnerabilities in authentication databases (Newman and Clarke 2017; Wall 2007) to steal user credentials. After obtaining the credentials, criminals can monetise the accounts in different ways. They seek for sensitive information such as credentials to other online services, financial information and even intimate information that can be used to blackmail the victim. Similarly, they can be used to send spam or spear phishing emails to other victims. Finally, credentials can be used as goods which are traded or shared in underground outlets.

*Correspondence: uctzdbe@ucl.ac.uk
[1] Department of Security and Crime Science, University College London, London, United Kingdom
Full list of author information is available at the end of the article

It is a great challenge for researchers to determine what happens when an account has been compromised. Previous research focused on understanding the use of stolen accounts in the Surface Web, i.e., the portion of the Internet where websites are indexed in the search engines and it is accessible with any browser. Onaolapo et al. (2016) studies the activity of cybercriminals accessing compromised Google accounts leaked through different outlets. Lazarov et al. (2016) monitors criminal activity on leaked Google spreadsheets. Similarly, Bernard-Jones et al. (2017) investigates the effects of language on cybercriminals navigating on compromised webmail accounts.

However, at the same time, cybercriminals are becoming more sophisticated and continue to improve their methods and techniques in order to engage in outlets of compromised data without getting caught or blocked. For instance, the increasing use of the Dark Web and the anonymity that this platform provides has attracted cybercriminals who can commit various computer crimes and maintain their activities hidden from law enforcement agencies. The Dark Web refers to websites hosted on networks built on top of the Internet that are not indexed by conventional search engines and only accessible by specialized software such as The Onion Router (Tor) (Syverson et al. 1997).

The main feature of these networks is that they provide user privacy by obfuscating the traffic between a client and a website or online service; therefore, the user can access the hosted content anonymously (Marin et al. 2016). The Tor network offers encrypted communications through which content providers can anonymously distribute content. These features can hamper the attempts of law enforcement agencies to track illegal activities and ultimately stop criminals who, according to some studies are using hidden outlets in the Dark Web to find or trade stolen account credentials (Dolliver and Kenney 2016; Lacey and Salmon 2015). According to *Top10VPN.com*, the world's largest Virtual Private Network review site, someone's online identity is worth £820 to miscreants on the Dark Web as for February 2018 (Migliano 2018).

Although some research has investigated various types of illegal activities in the Dark Web (Dolliver and Kenney 2016; Christin 2013), very few studies have compared how the same crime unfolds in both environments: the Surface Web and the Dark Web. As such, this paper aims to address this gap by comparing the results of the experiment performed by Onaolapo et al. (2016) in the Surface Web with the results of a similar experiment performed in the Dark Web. The new experiment follows Onaolapo's methodology to leak and monitor honey accounts. These accounts resemble legit email accounts from common users and are leaked through several online services on the Internet. Data from both experiments was collected and analysed to provide some insights into the differences related to stolen credentials in both environments.

To achieve this, we monitored honey webmail accounts leaked in the Dark Web for a period of a month using the infrastructure proposed by Onaolapo et al. (2016). For that purpose, we created fake Gmail accounts whose credentials were leaked in various outlets within online services of the Tor network such as paste sites (online outlets where users can store and share plain text) and underground forums. The intention of the experiment is to make cybercriminals interact with these credentials. Then, all events related to the emails in the accounts are recorded, namely when a mail is read, favourited, sent or a new draft is created. Similarly, we tracked the access to each account in order to obtain the system information and the origin of the login session.

The results suggest that stolen accounts are more likely to receive unwanted accesses when they are leaked in the Dark Web, especially on paste sites. The analysis of the activity performed on those accounts indicates that most access events are from curious actors who may be testing the credentials but do not perform any other activity. However, some of them repeatedly log in to the same account presumably to look for new relevant information. On the other hand, highly frequent use of unknown browsers suggests an attempt to hide the browser during the access. In summary, this paper makes the following contributions:

- We studied the activity generated on 100 email accounts whose credentials were leaked in different outlets of the Dark Web.
- We compare the results of this experiment with those obtained with one conducted with a similar methodology on the Surface Web (Onaolapo et al. 2016). Our results show that there are distinct differences between both Web environments in terms of malicious activity depending on the leakage outlet.
- Using the data collected, we publish a dataset containing the intrinsic characteristics of accesses to stolen accounts in a repository open to the public.[1]

Background and related work

Online accounts are valuable sources of personal information but they also usually gain a level of trust and reputation over time among contacts and other online services. There are several methods by which cybercriminals steal accounts credentials. Lynch (2005) analyses

[1] https://bitbucket.org/gianluca_students/surface_vs_dark_credentials_dataset.

phishing, where criminals send fake emails that seem to be official online services and make their victims type in their credentials on a fake site. Likewise, spear phishing attacks include fraudulent emails which are aimed at one or a specific group of users (Stringhini and Thonnard 2015). Another method used is to infect users with malware that steals information because their devices are not properly prepared to counter the threat (Stone-Gross et al. 2009). Finally, vulnerabilities in online databases can result in a massive leakage of credentials (Kontaxis et al. 2013). The aforementioned research describe stealing techniques but do not analyse what happens when an account has already been compromised.

Several studies has analysed the means by which cybercriminals dispose the information they possess. Criminal activities have led to a digital underground economy (Holz et al. 2009). The credentials of any account are goods that can be exchanged within this economy in several outlets. Holt and Lampke (2010) analysed the underground markets in which criminals release or trade the information obtained through malicious activities. In some cases, these accounts are freely released in order for the authors to build a reputation within the underground community (Butler et al. 2016). On the other hand, criminals seek some sort of financial gain and sell the stolen accounts to other criminals to monetise them. Ablon and Libicki (2015) argue that trading stolen data has become lucrative and easier to carry out than other types of illegal trade. Furthermore, a growing body of research has shown that personal and financial data can be obtained through markets for stolen data at a fraction of their true value (Holt and Lampke 2010). Therefore, there is a huge exchange rate of stolen credentials in the underground economy which are exposed in different outlets.

As a consequence, a small but growing body of research has focused on the actions taken by cybercriminals when obtaining access to the compromised online accounts. They can be used to send spam (Egele et al. 2013), find sensitive information or liquidate financial assets of the victim (Bursztein et al. 2014). Bursztein focuses on the stealing of credentials through phishing. However, compromised credentials can be obtained on several outlets. Onaolapo et al. (2016) analyses the activities cybercriminals carry out on compromised Gmail accounts. This work, which involves creating, populating and leaking fake Gmail accounts on paste sites, underground forums or by using malware, suggests that the attackers try to evade Google security mechanisms by using the location information of the account as the source of connection, if this information is provided.

Onaolapo et al. (2016) provide an analysis of the interaction between cybercriminals and hijacked accounts when stolen credentials are traded in outlets within the Surface Web. Based on the observations obtained from the accesses to the honey accounts, they identified a classification of the activity carried out by cybercriminals. There are four types of attackers according to the actions that they perform within the accounts:

- *Curious* log into the honey accounts and perform no further actions in them. They simply access the accounts to check the correctness of the credentials.
- *Gold Diggers* perform searchers on the emails contained in the account to find sensitive information that could be monetised in the underground economy.
- *Spammers* use the honey accounts to send spam messages by exploiting the trust that contacts have with the account owner.
- *Hijackers* change the account password to take full control of it, preventing the original owner of the account from having access.

Elsewhere, Stringhini et al. (2010) created 300 honey profiles on three major social networks to analyse how spammers operate. Similarly, Lazarov et al. (2016) leaked Google spreadsheets to understand what criminals do when they obtain illegal access to cloud-based documents. Dolliver and Kenney (2016) made a comparison of black markets in the Tor network using statistical analysis to determine significant differences among intrinsic characteristics of those markets.

The aforementioned research is performed on outlets positioned in the Surface Web which are those websites that are searchable and accessible using a web search engine such as Google, Bing, Yahoo, etc. On the other hand, the Deep Web refers to websites not indexed by a search engine but they can be directly accessed using a web address. As a part of the Deep Web, the Dark Web refers to websites on a darknet. Darknet is an encrypted network built on top of the Internet which has been designed specifically for anonymity and is accessible through specific software and tools. Examples of a Darknet are Tor, I2P, Freenet, DN42, etc. Therefore, the Dark Web contains websites whose content has been intentionally concealed (Weimann 2016). These websites are known as hidden services.

According to some studies, since law enforcement agencies have improved their techniques to detect and catch offenders who perform illegal activities in the Surface Web, black markets or underground forums based on the hidden services have become more prominent over the last few years (Marin et al. 2016). Many cybercriminals are migrating their operations to the Dark Web. For instance, Hardy and Norgaard (2016) studied data from black markets in order to analyse this emergent

ecosystem of marketplaces. Unlike our work, this research only focused on markets such as Silk Road.

The Dark Web poses a major challenge since the identities of the actors involved in this platform remains largely unknown and law enforcement agencies do not have enough resources to stop or deter illegal activities. These facts represent strong incentives for criminals to use them. Thus, it is important to understand the behaviour of criminals trading stolen credentials in the Dark Web outlets. As there is no sound information available about this issue so far, this study shall provide some insight by measuring the activity on stolen email accounts in terms of unique accesses, type of activity performed, devices used for the access and its duration. Hence, we define our research question as: Does the Web environment affect cybercriminal activity?

Methodology

Using the honeypot infrastructure for the Surface Web experiment proposed by Onaolapo et al. (2016), we conducted a new experiment in the Dark Web. The aim of the experiment was to imitate the way of operating of cybercriminals releasing or trading stolen account credentials through some outlets in the Dark Web, specifically in some hidden services within the Tor network. The infrastructure tracked the actions performed by criminals who had the account credentials in their possession. The results of the experiment in the Dark Web are paired with the results of Onaolapo's experiment in the Surface Web to draw comparisons. For the sake of the comparison, we followed the same methodology used in the Surface Web experiment i.e., leaking the same number of accounts across the same type of outlets.

The first step of the experiment was the creation of Gmail accounts which are called honey accounts. These accounts resemble legitimate email accounts from common users. In the creation phase, 100 honey accounts were created manually on Gmail. The fictitious data to create the accounts was automatically generated using a database of random names for the accounts. All the accounts were populated with email messages from the Enron dataset to simulate a real email account belonging to a normal user. Enron was an energy company declared bankrupt in 2001 and the emails dataset from the company executives were made available to the public. This corpus contains a total of 517,431 messages from 150 users (Zhou et al. 2007). Each account received at least 200 emails which were sent in batches before and after the leak for it to resemble an active user account that handles a lot of information. The first names, last names and the name "Enron" were replaced in all the emails using the fictitious names.

In the next phase, the accounts were instrumented with scripts to monitor and register the activity of anyone visiting them. The monitoring infrastructure is based on the incorporation of Google Apps Scripts hidden in a Google Sheet as a normal document within each account. Google Apps Script is a JavaScript cloud scripting language used to automate different time-based and event-based tasks across Google products. The scripts were used to monitor all the actions over emails by scanning the emails to determine if an email has been read, sent, marked as important (Starred) or if a draft has been created.

Similarly, other scripts extracted more information from the 'Device activity and notifications' section within the Gmail account management dashboard from each account. This section uses the Google fingerprinting system to extract the data from the cookie generated for each log in to the accounts. A cookie is a small piece of data sent to a browser by a Web server while the user is browsing. Cookies are designed to be a reliable mechanism for websites to remember session information or to record the user's browsing activity. The cookie information includes: cookie identifier, public IP address, location, login time, browser and the operating system of the device where the login originated from. Each cookie found in our dataset is considered as an unique access to an account. As will be explained later, leaking the accounts in the Dark Web does not imply that the accounts will be accessed through Tor. In fact, this is very unlike because Gmail usually block login attempts from Tor.

Similar to the Surface Web experiment, the outlets chosen for the leaks were paste sites and underground forums. The idea behind leaking the accounts in different outlets is to compare malicious activity among them. A third type of outlet, black markets, was added to the Dark Web experiment for information purposes only but not used for the comparison as they were not used in the Surface Web experiment. The experiment was performed using 100 accounts for the leakage. They were divided into groups, each to be leaked on different hidden services within Tor.

The hidden paste sites chosen were `Insertor` and `Stronghold`. In terms of underground forums, the hidden services used were: `AlphaBay`, `Silk Road Forum` and `KickAss`, where there are many threads regarding illegal activities, such as data theft. The selection of these sites was due to the similarity they have with the outlets used for the Surface Web (`pastebin.com` and `pastie.org` for paste sites; `offensivecommunity.net`, `bestblackhatforums.eu`, `hackforums.net` and `blackhatworld.com` for underground forums) in terms of the degree of activity found, with many posts and messages exchanged daily by

members. Furthermore, the chosen sites do not have an account method allowing visitors to post without registration. While traffic is an important variable to consider in the experiment, we were unable to get statistics from these hidden services due to the nature of them in order to establish differences among the sites. We acknowledge the limitation and we discuss it later.

Activity on the honey accounts was recorded for a period of about seven months for the Surface Web and one month for the Dark Web, which was the period covered for our ethics approval. However, in order for the comparison to be homogeneous, we extracted the first month of observations in the Surface Web experiment. We chose the first month to replicate the same features in both environments as if the Surface Web experiment would have been performed for only one month to make sure not to introduce any statistical bias.

This paper seeks to determine whether any of the characteristics of the accesses are associated with the environment they are coming from. The data gathered from both experiments may be useful for researchers to understand how attackers interact with stolen webmail accounts and how this malicious activity differs in the Surface Web and the Dark Web. Therefore, we will publicly release an anonymised version of the data for academic purposes.

Ethical considerations

The experiment was developed taking into account several ethical considerations in order not to affect actual Gmail users. First, the default `send-from` address of the honey accounts was altered so that when an email is sent from any of them, it was sent to a controlled SMTP mail server that was set up to receive and store these emails without forwarding them to the intended destination. The send-from address was changed using the settings menu within each Gmail account. This measure was taken to avoid abuse from cybercriminals. Similarly, we worked in collaboration with Google to ensure that accounts are suspended when they are hijacked or in case of problems beyond our control. In addition, the project was reviewed and obtained ethical approval by University College London.

Results

The Surface Web experiment identified 164 unique accesses to the accounts after the leak; on the other hand, 1092 unique accesses to the Dark Web accounts were recorded in our experiment (see Table 1). It is important to note that even though the credentials are leaked in Dark Web outlets, they are not always accessed from the Tor network. Thus, in our analysis, the Dark Web statistics refer to accounts which have been exposed but not accessed through Tor. In fact, only 378 accesses originated from the Tor network. In order to perform our statistical tests we coded the collected data into the following variables: cookie identifier, Web environment, IP address, outlet, taxonomy, login time, location browser and the operating system of the access.

Table 1 Unique accesses depending on the outlet

	Paste sites (%)	Underground forums (%)	Total
Surface	51.8	48.2	164
Dark	90.8	9.2	1092

Paste sites are more likely to be used by cybercriminals in the Dark Web ($\chi^2 = 177.587, p < 0.001$). Conversely, more logins come from underground forums in the Surface Web

We used a chi-square test (Agresti 1996) to determine whether a relationship exists between Web environment and outlet. The results showed that there is a significant relationship ($\chi^2 = 177.587$, $p < 0.001$). While most accesses from the Dark Web originate from the credentials leaked through paste sites, more logins in the Surface Web come from underground forums. This suggest that the exposure of stolen credentials is higher in Dark Web paste sites. On the contrary, underground forums in the Dark Web are less accessible since as we noticed, a great deal of them requires an invitation or referral to access them.

Taxonomy of account activity

Based on our observations on the honey accounts and the classification or taxonomy mentioned in previous sections, the following accesses were identified in the Surface Web: 103 *Curious*, 39 *Gold Diggers*, 2 *Spammers* and 20 *Hijackers*. On the Dark Web we registered 812 *Curious*, 227 *Gold Diggers*, 39 *Spammers* and 14 *Hijackers* (see Table 2).

We carried out a Fisher's Exact Test (FET) (Mehta and Patel 1983) to observe if there is a significant association between Web environment and taxonomy ($p < 0.001$, 99% CI). In this case, we are not using a chi square test to find significant differences because our contingency table has cells with expected frequencies of less than 5, which violates an assumption of this test. The test revealed that there is a significant association between Web environment and taxonomy ($p < 0.001$, 99% CI) but a Cramer's V statistic showed that the strength of the association is weak ($V = 0.233$). This result is for the overall analysis and a post-hoc is performed to find individual significances. We rely on a method that yields probability values for each combination of independent category levels and uses a Bonferroni correction to control for type I error inflation (Beasley and Schumacker 1995; MacDonald and Gardner 2000). The test reports the percentage

Table 2 Unique accesses depending on the taxonomy

	Curious (%)	Gold Digger (%)	Hijacker (%)	Spammer (%)	Total
Surface	62.8	23.8	12.2	1.2	164
Dark	74.4	20.8	1.3	3.6	1092

Hijacking is more likely to occur in the Surface Web (FET: $p < 0.001$)

contribution for each cell to the overall chi-square statistic. We found that there is significant association between the Web environment and *Hijackers* ($p < .001$). *Hijacking* is more likely to take place in the Surface Web (12.2%) compared to the Dark Web (1.3%) where this event is rare. Further analysis including the variable outlet (see Table 3) revealed that this association is significant only in paste sites ($p < 0.001$, 99% CI). This may be indication that attackers are stealthier in the Dark Web and try to go unnoticed without changing password in the accounts which in turn indicates a certain level of sophistication. Regarding the underground forums, the observed differences are not significant.

Device configuration of accesses

Google's system fingerprinting was used to collect information about devices accessing the honey accounts. Table 4 shows the distribution of web environment, operating system in each outlet where the credentials were leaked. There is a significant association between operating system and web environment when credentials are obtained in paste sites ($p < 0.001$, 99% CI). However this association is weak ($V = 0.198$). Although most of the accesses originate from Windows, our post-hoc analysis revealed that cybercriminals are more likely to use Android devices when using credentials gathered in the Surface Web than in the Dark Web (15.3% vs. 1.1%, $p < 0.001$). This may be an indication of a low level of sophistication as users are probably using their own mobile devices to access the accounts. On the other hand, Linux is more likely to be used in the Dark Web (22.5% vs. 7.1%, $p < 0.001$). It is reasonable to assume that Linux is used by more skilled criminals which is consistent with the evidence that there might be a higher level of sophistication in the Dark Web. In the case of underground forums, the observed differences are not significant.

The browser distribution is outlined in Table 5. There is a significant association between Web environment and browser ($p < .001$). The post-hoc test shows that unknown browsers are more likely to be used in the Dark Web (60%) than in the Surface Web (39.9%) for paste sites ($p < .001$). While this may be an indication that criminals attempt to hide the browser user agent from the Google fingerprinting system when accessing the accounts, one could easily argue that any sophisticated attacker would use a common user agent in an effort to avoid triggering detection mechanisms when trying to log in. The

Table 3 Distribution of accesses for each outlet and taxonomy class

	Curious (%)	Gold Digger (%)	Hijacker (%)	Spammer (%)	Total (%)	Statistics (FET)
Paste sites						
Surface	63.5	17.6	18.8	0.0	85	$p < 0.001$
Dark	74.5	20.3	1.3	3.9	992	
Forums						
Surface	62.0	30.4	5.1	2.5	79	$p = 0.099$
Dark	73.0	26.0	1.0	0.0	100	

There are significant differences between the Surface Web and the Dark Web when credentials are leaked through paste sites and used to hijack an account (FET: $p < 0.001$)

Table 4 Distribution of accesses for each outlet and operating system

	Android (%)	Chrome OS (%)	iOS (%)	Linux (%)	Mac (%)	Unknown (%)	Windows (%)	Statistics (FET)
Paste sites								
Surface	15.3	0.0	0.0	7.1	4.7	9.4	63.5	$p < 0.000$
Dark	1.1	0.0	1.7	22.5	3.4	8.1	63.2	
Forums								
Surface	0.0	1.3	0.0	8.9	1.3	0.0	88.6	$p = 0.031$
Dark	7.0	0.0	2.0	7.0	5.0	1.0	78.0	

Most of the accesses originate from Windows; however, cybercriminals in paste sites are more likely to use Android devices when using credentials gathered in the Surface Web. On the other hand, Linux is more likely to be used in the Dark Web (FET: $p < 0.001$)

Table 5 Distribution of accesses for each outlet and browser

	Chrome (%)	Edge (%)	Firefox (%)	Iceweasel (%)	Internet Explorer (%)	Opera (%)	Safari (%)	Thunder-bird (%)	Unknown (%)	Vivaldi (%)	Statistics (FET)
Paste sites											
Surface	27.1	0.0	14.1	4.7	9.4	2.4	0.0	0.0	40.0	2.4	$p < 0.000$
Dark	11.4	0.3	25.9	0.0	1.0	0.8	0.5	0.0	60.1	0.0	
Forums											
Surface	31.6	0.0	19.0	0.0	0.0	0.0	0.0	0.0	49.4	0.0	$p < 0.000$
Dark	12.0	0.0	38.0	0.0	0.0	6.0	0.0	5.0	39.0	0.0	

There is a significant association between Web enviroment and the use of unknown browsers in the Dark Web (FET: $p < 0.001$)

collection of further data and an analysis of the accuracy of Google's fingerprinting system would be important to draw strong conclusions about this aspect. Similarly, there is a significant association between Web environment and Chrome for both outlets ($p < .001$). The use of Chrome is more likely to happen in the Surface Web for paste sites and underground forums. Interestingly, in the Dark Web we got five accesses from Mozilla Thunderbird clients. This indicates that several attackers, such as *Gold Diggers* or *Spammers*, are using the functionalities of this email application to abuse the accounts.

Duration of the accesses

When a new access occurs in a honey account, a cookie identifier is generated along with the timestamp of access. Indeed, each cookie in the dataset has a timestamp of the first access and a timestamp of the last known access to a honey account. We used these timestamps to determine the length of access of a cookie for each unique access (Onaolapo et al. 2016).

Figure 1 shows the Cumulative Distribution Function (CDF) of the lengths of accesses to the accounts in the Surface Web and the Dark Web. Most accesses were short, being less than a day, meaning that most visitors accessed the honey accounts only once and did not return. However, Dark Web accesses had a longer time between subsequent interactions with the accounts compared to the Surface Web for all taxonomies. Approximately 30% of Dark Web *Curious* logins connected to the accounts several days after the first login and only less than 5% did it in the Surface Web. For *Gold Diggers*, the trend is the same (approximately 20% vs. 5%). In the case of *Hijackers*, about 10% of accesses continued taking place during this period in both Web environments. However, this indication may not be entirely accurate because it represents the length of the access until the cookie was hijacked. The two *Spammers* in the Surface Web sent emails in bursts for a a short period (less than a day). Conversely, spam in the Dark Web occurred over almost ten days.

Discussion

Our findings show that accounts leaked through paste sites received more accesses in both Web environments but the scale of access is much larger for paste sites in the Dark Web. While it is true that paste sites are more likely to be used to leak credentials, there is a big difference in the exposure of the leaks between the Surface Web and the Dark Web. Normally in the Surface Web, content related to information leakage is removed from paste sites by administrators monitoring the site. On the contrary, paste sites are not monitored in the Dark Web and leaks tend to be published for longer. Therefore, credentials leaked in paste sites in the Dark Web are more exposed than in the Surface Web. Regarding underground forums, exposure is similar to paste sites in the Surface Web. On the contrary, credentials are less exposed in the Dark Web forums because they normally require the creation of an account and sometimes an invitation. One limitation of our work is that we were not able to establish whether the outlets used for our experiment are similar in terms of traffic. Therefore, the difference in the number of accesses between both Web environments may be due to the particular websites and hidden services we chose and not because of the environment itself.

In terms of the type of activity (taxonomy), there is a higher concentration of *Hijackers* in the Surface Web. *Hijacking* can be considered as malicious but the absence of it can mean that cybercriminals are more sophisticated and try to go unnoticed when using credentials. Thus, there is a higher level of malicious activity in the Surface Web but miscreants tend to be more stealthy in the Dark Web. Interestingly, our data shows that there is high concentration of *Curious* in the Dark Web. Even though no activity is performed on the honey accounts, it is reasonable to assume that more skilled attackers would not interact with the accounts to avoid detection. Unfortunately, we are not able to detect these "sophisticated" *Curious* users. Furthermore, the high level of *Curious* activity in the Dark Web can be explained by sophisticated miscreants crawling websites searching for stolen data and using bots to just perform the login in order to build a credentials database for further inspection.

We showed that a variety of operating systems and browsers were used to access the honey accounts. Android is more likely to be used in the Surface Web showing a low level of sophistication as personal devices may be used to log into the accounts. On the other hand, the use of Linux is a sign that high-skilled attackers are accessing the Dark Web accounts. It may be the case that sophisticated attackers are using Windows bots for accessing the accounts, yet we are not able to measure automatic accesses with our infrastructure.

With regards to the type of browsers used, accesses from unknown browsers are more likely to happen in the Dark Web: this fact indicates that attackers try to hide their browser user agent information, suggesting some degree of sophistication. However, the use of browser extensions to change or hide the browser the user agent is common among users nowadays. Moreover, it could be argued that skilled users are prone to use known or typical user agents as an attempt to avoid being flagged as malicious users. In the Surface Web, Chrome is more likely to be used to log in to the accounts. The use of this common browser suggest a low level of sophistication in this environment. Our data was collected using

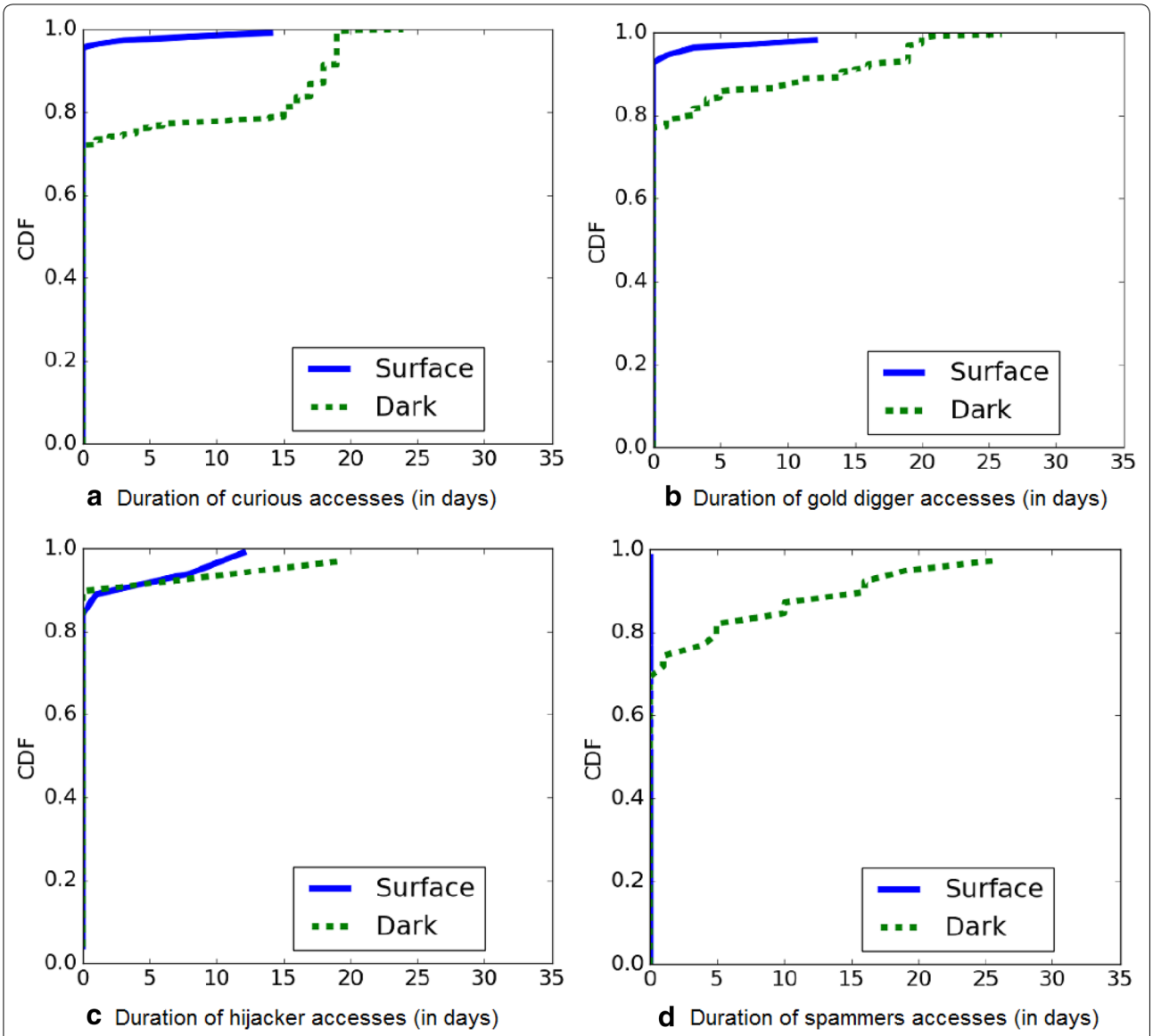

Fig. 1 CDF of the length of unique accesses on the honey accounts for: **a** Curious, **b** Gold Diggers, **c** Hijackers and **d** Spammers. X axis represents the duration of the access in days. Most accesses in all categories occured only once

Google's fingerprinting system, thus the reliability of the results depends on the accuracy of the system. Nevertheless, the observed differences suggest that a considerable percentage of sophisticated users attempt to be stealthy in the Dark Web when credentials are obtained through paste sites. Also, the comparison shows that attackers in the Dark Web are more likely to connect several times to look for new information in the accounts.

The comparison shows us that although the differences in terms of the type of activity are not substantial in some cases, the Dark Web attracts individuals who seek to discover the secrets of the dark side of the Web. The high number of accesses through hidden services suggests that there is a great interest in the information contained in the Dark Web outlets. It is reasonable to assume that this information could lead many users to use it in a malicious way and end up becoming cybercriminals.

We believe that security systems for account logins can be improved with the help of behavioural detection systems which are capable of finding activity patterns that seem to be different to those commonly used in the accounts. Therefore, information about accesses to compromised accounts can be useful in building algorithms that allow early detection of malicious

activity. We observed malicious activity taking place on accounts leaked in the Dark Web suggesting an increasing use of this environment as a platform to perform illegal activities, especially as far as the trade of stolen information is concerned. For this reason, data gathered from this project may support the development of policies focused on disabling hidden outlets dedicated to those activities.

One of the important limitations of this comparison is that the experiment for the Surface and the Dark Web were performed in different spaces of time. Therefore, the level of activity in both Web environments could have changed from one experiment to the other. Thus, the data of the experiments may not be enough to generalize our results. Our future agenda includes setting up honeypot infrastructure for both environments on other online service to establish a more accurate comparison. Another limitation was the number of Gmail accounts that we were able to create for our experiment. The creation of an account requires the registration of a phone number and any automatic approach is flagged as spam by Gmail; therefore, we were not able to create a large number of them.

Conclusion

In this paper, we compared the data from two similar experiments in which credentials of honey email accounts were leaked in the Surface Web and the Dark Web. We collected and performed a comparison based on different variables in our observations. Compromised accounts received more unauthorised accesses in the Dark Web than in the Surface Web, especially when credentials are released in paste sites due to level of exposure of this type of outlet. We found that there is a relationship between the Web environment and the type of activity performed in the honey accounts as well as the configuration of the devices used to log in to the accounts. We believe that our findings can help the research community to get a better understanding of the different types of malicious activity on stolen accounts. This comparison will contribute to the development of behavioural rules than can be included in detection systems aiming to protect users from attackers in different layers of the Internet.

Abbreviations

Tor: The Onion Router; FET: Fisher's Exact Test; CDF: Cumulative Distribution Function.

Authors' contributions

Experiments, data collection, analysis and interpretation of data and drafting of manuscript: DBV. Honeypot infrastructure: JO. Critical revision: GS, MM. All authors read and approved the final manuscript.

Author details

[1] Department of Security and Crime Science, University College London, London, United Kingdom. [2] Department of Computer Science, University College London, London, United Kingdom. [3] Department of Geography, University College London, London, United Kingdom.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Received: 20 June 2018 Accepted: 13 November 2018
Published online: 23 November 2018

References

Ablon, L., & Libicki, M. (2015). Hacker's bazaar: The markets for cybercrime tools and stolen data. *Def. Counsel J.*, *82*, 143.

Agresti, A. (1996). *An introduction to categorical data analysis* (Vol. 135). New York: Wiley.

Apps Script-Google Developers. https://developers.google.com/apps-script

Beasley, T. M., & Schumacker, R. E. (1995). Multiple regression approach to analyzing contingency tables: Post hoc and planned comparison procedures. *The Journal of Experimental Education*, *64*(1), 79–93.

Bernard-Jones, E., Onaolapo, J., & Stringhini, G. (2017). Email babel: Doeslanguage affect criminal activity in compromised webmail accounts? CoRR abs/1704.07759. 1704.07759.

Bursztein, E., Benko, B., Margolis, D., Pietraszek, T., Archer, A., Aquino, A., et al. (2014). Handcrafted fraud and extortion: Manual account hijacking inthe wild. *Proceedings of the 2014 conference on internet measurement conference*. IMC '14, pp. 347–358. , New York, NY, USA: ACM. https://doi.org/10.1145/2663716.2663749.

Butler, B., Wardman, B., & Pratt, N. (2016). Reaper: an automated, scalable solution for mass credential harvesting and osint. 2016 APWG symposium on electronic crime research (eCrime), pp. 1–10. https://doi.org/10.1109/ECRIME.2016.7487944

Christin, N. (2013). Traveling the silk road: A measurement analysis of a largeanonymous online marketplace. *Proceedings of the 22nd international conference on World Wide Web*. WWW '13 pp. 213–224. New York, NY, USA: ACM.https://doi.org/10.1145/2488388.2488408.

Dolliver, D. S., & Kenney, J. L. (2016). Characteristics of drug vendors on the tor network: A cryptomarket comparison. *Victims & Offenders*, *11*(4), 600–620. https://doi.org/10.1080/15564886.2016.1173158.

Dolliver, D. S., & Kenney, J. L. (2016). Characteristics of drug vendors on the Tor network: A cryptomarket comparison. *Victims & Offenders*, *11*(4), 600–620.

Egele, M., Stringhini, G., Kruegel, C., & Vigna, G. (2013). Compa: Detecting compromised accounts on social networks. In: NDSS.

HTTP Cookie. https://en.wikipedia.org/wiki/HTTP_cookie

Hardy, R. A., & Norgaard, J. R. (2016). Reputation in the internet black market: An empirical and theoretical analysis of the deep web. *Journal of Institutional Economics*, *12*(3), 515–539. https://doi.org/10.1017/S1744137415000454.

Holt, T. J., & Lampke, E. (2010). Exploring stolen data markets online: products and market forces. *Criminal Justice Studies*, *23*(1), 33–50. https://doi.org/10.1080/14786011003634415.

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

Kontaxis, G., Athanasopoulos, E., Portokalidis, G., & Keromytis, A.D. (2013). SAuth: Protecting user accounts from password database leaks.

Lacey, D., & Salmon, P. M. (2015).It's dark in there: Using systems analysis to investigate trust and engagement in dark web forums. International conference on engineering psychology and cognitive ergonomics. pp. 117–128. New York, Springer.

Lazarov, M., Onaolapo, J., & Stringhini, G. (2016). Honey sheets: What happens to leaked google spreadsheets?

Lynch, J. (2005). Identity theft in cyberspace: Crime control methods and their effectiveness in combating phishing attacks. *Berkeley Technology Law Journal*, *20*, 259.

MacDonald, P. L., & Gardner, R. C. (2000). Type i error rate comparisons of post hoc procedures for i j chi-square tables. *Educational and Psychological Measurement*, *60*(5), 735–754. https://doi.org/10.1177/00131640021970871.

Marin, E., Diab, A., & Shakarian, P. (2016). Product offerings in malicious hacker markets.

Mehta, C. R., & Patel, N. R. (1983). A network algorithm for performing fisher's exact test in r x c contingency tables. *Journal of the American Statistical Association*, *78*(382), 427–434. https://doi.org/10.1080/01621459.1983.10477989.

Migliano, S. (2018). Dark Web Market Price Index (UK Edition). Top10VPN. https://www.top10vpn.com/privacy-central/cybersecurity/dark-web-market-price-index-feb-2018-uk/

Newman, G. R., & Clarke, R. V. (2017). *Superhighway robbery*. United Kingdom: Willan Publishing (UK).

Office of National Statistics. (2016). https://goo.gl/G9Wg6F

Onaolapo, J., Mariconti,E., & Stringhini, G. (2016). What happens after you are pwnd:Understanding the use of leaked webmail credentials in the wild. *Proceedings of the 2016 internet measurement conference. IMC '16*, pp. 65–79. New York, NY, USA: ACM.https://doi.org/10.1145/2987443.2987475.

Stone-Gross, B.,Cova, M., Cavallaro, L., Gilbert, B., Szydlowski, M., Kemmerer, R., Kruegel, C., & Vigna, G. (2009). Your botnet is my botnet: Analysis of abotnet takeover. *Proceedings of the 16th ACM conference on computer and communications security. CCS '09*, pp. 635–647. New York, NY, USA: ACM. https://doi.org/10.1145/1653662.1653738.

Stringhini, G., & Thonnard, O. (2015). That ain't you: Blocking spearphishing through behavioral modelling. *Proceedings of the 12th International conference on detection of intrusions and malware, and vulnerability assessment*. Vol. 9148. DIMVA 2015, pp. 78–97. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-319-20550-2_5.

Stringhini, G.,Kruegel, C., & Vigna, G. (2010). Detecting spammers on social networks. *Proceedings of the 26th annual computer security applications conference*. ACSAC '10, pp. 1–9. New York, NY, USA: ACM.https://doi.org/10.1145/1920261.1920263.

Syverson, P.F., Goldschlag, D.M., & Reed, M.G. (1997). Anonymous connections and onion routing.

Wall, D. (2007). *Cybercrime: The transformation of crime in the information age* (Vol. 4). United Kingdom: Polity.

Weimann, G. (2016). Terrorist migration to the dark web. *Perspectives on Terrorism*, *10*(3), 40–44.

Zhou, Y., Goldberg, M., Magdon-Ismail, M., & Wallace, A. (2007). Strategies for cleaning organizational emails with an application to enron email dataset. Conference of North American association for computational social and organizational science.