All Your Accounts Are Belong to Us

Vlad Bulakh[1(✉)], Andrew J. Kaizer[1], and Minaxi Gupta[2]

[1] Indiana University, Bloomington, IN 47405, USA
{vbulakh,akaizer}@indiana.edu
[2] Edmodo Inc., San Mateo, CA 94403, USA
minaxi@edmodo.com

Abstract. Over the last several years, there have been a number of high profile and well-publicized data breaches. These breaches led to the theft of personal, financial, and health information from users who are often only notified of such breaches well after they occur and the damage has already been done. Cyber criminals use account cracking tools, which are software programs that help miscreants gain access to users' online accounts, to perform credential stuffing attacks against the credentials exposed by these breaches.

In this paper, we study underground forums where intelligence related to popular account *cracking tools* is exchanged and investigate miscreants' motivations to use such tools to break into accounts. We also study six free and paid cracking tools used to steal user accounts and develop machine learning classifiers capable of detecting network packets generated by them. Organizations maintaining user accounts can utilize our classifiers to identify traffic related to cracking tools and defend against their attacks.

Keywords: Data breach · Underground forum · Credential stuffing
Account cracking · Credential verification · Cracking tools
Sentry MBA · Account Hitman · AIOHNB · Vertex · Classifier
Supervised machine learning · Random Forest

1 Introduction

Over the past several years, there has been an alarming increase in the number of data breaches throughout the world. The victims of these cyber criminals include prominent firms such as the Red Cross [17], Yahoo [21], ClixSense [56], Ubuntu Forums [41], Interpark [44], the Democratic National Committee [39], and Mossack Fonseca [12]. As a result of these breaches, millions of consumers' personal, financial, and medical information has been exposed to cyber criminals, who can use the information for financial, political, and social gains.

One key factor that has led to these breaches is the growing number of malicious tools that miscreants have at their disposal, including malware, credit card skimmers, and online account cracking tools such as Sentry MBA and Account Hitman. In this paper, we gain a better understanding of the online "cracker"

X. Lin et al. (Eds.): SecureComm 2017, LNICST 238, pp. 245–269, 2018.
https://doi.org/10.1007/978-3-319-78813-5_13

community and investigate defenses against such attacks. In particular, we concentrate our efforts on several popular underground forums specializing in cracking tools, which are computer programs that can be used to gain unauthorized access to other people's online accounts. The forums we analyzed contain configuration files, which are text files with website-specific settings for cracking tools. For example, a Facebook configuration file for the Sentry MBA tool might contain a custom *User-Agent* header field and an HTTPS address of the user login page, which helps the tool avoid being blocked by Facebook by making the traffic appear to be from a legitimate browser user. Configuration files also allow developers and users to keep their software up to date without modifying the source code and recompiling the program, providing an accessible approach for changing targets.

In addition to studying the configuration files exchanged in underground forums, we studied the cracking tools to examine their behaviors and identify defense mechanisms. This included fairly sophisticated cracking tools, including ones that could check the validity of existing/stolen credentials on popular websites such as Gmail, Amazon, eBay, PayPal, Steam, and others. Some tools could also be used to discover username and password combinations through brute force attacks, which use automated means to guess such information through trial and error. Criminals, however, appear to be using them primarily for checking the validity of breached credentials, also known as a *credential stuffing* attack [58].

Detecting these cracking tools then becomes a critical task to mitigate the risk they pose to an organization's users. Although identifying such attacks from the server side is no trivial matter, by analyzing the packets generated by both paid and free cracking tools, we devise a system capable of detecting up to 100% of attacks.

Our contributions in this study are threefold:

1. *Characteristics of underground forums dealing with cracking tools*: We registered on four underground forums – webcracking.com, nethingoez.com, nulled.to and cracking.org – where members discuss cracking tools and exchange information about them, among other illicit discussions related to hacking tutorials or finding serial numbers to popular video games. We then analyzed these forums by scraping information about the number and length of threads and posts, user location, user join date, and user activity. Interestingly, we found that very few people actually ask for help on these forums. Instead, the majority of the posts are non-informative and made only because gaining access to the shared content required posting. Furthermore, judging from the users' browsing and posting habits, we find these forums to be niche places aimed at a fairly narrow, albeit loyal, audience.
2. *Comparison of popular paid and unpaid cracking tools*: We compare and contrast the features and performance of some of the most popular free and paid crackers, including Sentry MBA [1], Account Hitman [30], Vertex [11], AIOHNB [14], vCrack [16], and Multi-Hacker [25]. Surprisingly, we discovered that the free tools contained more features and performed in a similar capacity, indicating that miscreants who pay for crackers may not be deriving any

additional value beyond free tools. We also found both free and paid cracking tools to have a significant number of bugs and glitches, which is surprising considering how mature some of these tools are.

3. *Defending against identity theft*: Finally, we use the knowledge gained from our contributions to develop several machine learning algorithms that companies maintaining user accounts can use to detect when crackers are accessing their websites. Our classifiers rely on the features extracted from the network packets, such as packet size, HTTP version, and HTTP *Connection* and *Accept-Language* header fields, so companies can identify such threats before processing their requests.

2 Analysis of Cracking Forums

Analyzing the users, topics, and posts of these underground forums can provide valuable insight into some of the motivations and trends behind the cracking culture. In particular, we find that forums may share some high level properties – such as bursts of activity and a small core group of posters – but that the config files discussed on each website tended to focus on different targets – e.g. gaming versus file sharing websites. Before continuing with the analysis, a brief discussion of terminology and data collection is necessary to contextualize the problem space.

2.1 Terminology

An **administrator** (also called **admin**) is a forum member who has elevated privileges. Among other things, a typical forum administrator can: edit other members' posts, remove individual messages and complete threads, and issue warnings to and ban misbehaving forum members.

A **configuration file** (also called **config**) is a text file containing website-specific settings for a cracker. For example, an Amazon config file for Sentry MBA might contain a custom *Referer* field and an Amazon-specific timeout. A snippet from a Sentry MBA configuration file for Instagram can be seen below:

```
[Wordlist]
UserIndex=1
PassIndex=2
...
[Settings]
SiteURL=https://instagram.com/accounts/login/ajax/
Timeout=20
RequestMethod=2
Referer=1
...
```

Credential stuffing (also called **credential checking** and **credential verification**) is an attack in which cyber criminals load breached username/password combinations into a cracking tool like Sentry MBA and try to take over other people's online accounts by having the cracking tool check the supplied credentials against the target website.

A **forum** (also called a **message board**) is a website where people can communicate with each other by posting messages. The content of messages can include text, emotions, pictures, and videos.

An **original post** (often abbreviated as **OP**) is the first post in a continuous sequence of postings.

A **post** is a message in a form of text, emotions, pictures, etc. posted on the forum.

A **subforum** is located inside another forum. Subforums are often used to divide a single forum into specific discussion topics. For example, an underground forum might have a cracked programs subforum for cracking tools.

A **topic** (also called **thread**) is a sequence of posts/messages posted in the response to the original post.

A **topic starter** (often abbreviated as **TS**) is the person who posted the first message in a continuous sequence of postings (i.e. original post). A topic starter can also be called **original poster** and abbreviated as **OP**.

2.2 Data Collection and Methodology

The websites studied covered four of the most popular cracking forums: webcracking.com, nethingoez.com, nulled.to, and cracking.org. As of May 2017, all these websites are highly ranked by Alexa, with nulled.to, cracking.org, webcracking.com, and nethingoez.com having global ranks of 25K, 121K, 275K, and 300K, respectively.

We collected complete snapshots of webcracking.com, nulled.to, cracking.org, and nethingoez.com on December 19, 2015, July 8, 2016, September 12, 2016, and August 29, 2016, respectively. The scraping process focused on the subforums dealing with configuration files for the most popular cracking tools, such as Sentry MBA, Account Hitman, AIO Checker, and Vertex. For each snapshot, we collected the threads, posts, and users across all config file subforums. This ensures a *complete* overview of the subforums at that particular point in time.

Data Cleaning

We saw a number of inconsistencies in the collected data, even when that data was from the same underground forum, that could undermine data analysis if not accounted for. For example, the cracking tool field could say "SentryMBA," "Sentry MBA proxyless," "SentryMBA proxylexx," "Sentry," "Sentary MBA," "SMBA," "SenMBA," and "S. MBA," all of which refer to the same cracking tool – Sentry MBA[1]. We also saw a number of incorrect entries and labels. For example, a thread might be located in the Vertex subforum, but have tags corresponding to other cracking tools, e.g. "Account Hitman."

Due to these factors, considerable effort was spent on data sanitization. About 10% of the data we collected had to be cleaned, which involved standardizing the names of the cracking tools (e.g. both "Hitman" and "Acc Hitman"

[1] Although it is possible that some of those could be referring to different Sentries, such as the original Sentry [46], which is the predecessor of Sentry MBA [19], a manual analysis of 25 threads revealed that all of them were about Sentry MBA.

became “Account Hitman”), inferring missing information (e.g. determining the cracking tool from the first post in the thread) and ignoring invalid entries. Overall, approximately 2% of all threads and 4% of all posts have been discarded through this process.

2.3 Users

Looking at the number of users who posted at least one message in the config file subforums across all websites, we observe that nulled.to leads with 14,446 unique users and is followed by cracking.org, webcracking.com, and nethingoez.com with 7,500, 2,720, and 1,719 unique users, respectively. This indicates that the degree of popularity for cracking activities varies widely across various underground forums.

Interestingly, we see that all forums have small-to-large gaps between user registration dates that could last from several days to several weeks – except cracking.org which did not show the registration date at the time of data collection. For example, although 721 people created new accounts on nulled.to between April 25–May 5, 2016, with no day having fewer than 24 new registrations, no new accounts were created from May 6–June 23, 2016. Such large gaps could be either due to the websites’ doing user registrations in batches or service availability issues.

Additionally, webcracking.com showed a user-supplied location during our data collection period, which, admittedly, could be falsified. Of those users who did specify their location, most came from the United States, followed closely by the United Kingdom, Germany, France, Canada, Italy, Spain, Turkey, India, and Brazil. Notably absent from this list are some of the well-known countries that engage in more insidious forms of consumer-oriented cybercrime, such as Russia or China. A focus on the top countries may indicate a proclivity towards a less tech-savvy, more “script kiddie”-oriented audience.

Also, when we cross-reference user account names across the config file subforums of each underground forum, we see that the overwhelming majority of account names can only be found in one of the four forums, with 3.2% instances of an exact account name match on two different forums, 0.4% matches on three different forums, and only 0.08% matches across all four forums. This implies that either miscreants utilize separate identities on each forum or that they tend to use only one source for their cracking needs.

Furthermore, looking at the average number of active users across all four underground forums, we see that, on a per-hour basis, there are 682 members and 604 guests active on nulled.to, 50 members and 95 guests active on nethingoez.com, and 19 members and 180 guests active on cracking.org. For webcracking.com, we were only able to get the daily statistics, which showed that an average of 271 members and 1,169 guests are active on any given day. Compared to popular, legitimate forums such as reddit.com and 4chan.org, which can have hundreds of thousands of *active* users at any given time with many of them having posted dozens and even hundreds of thousands of messages, these underground forums appear to be niche places aimed at a very narrow audience.

2.4 Threads

When looking at the number of active threads that share and discuss configuration files, cracking.org takes the first place with 3,197 threads – despite having the second-fewest active members at any given hour amongst the four forums. It is followed by nulled.to (833 threads), nethingoez.com (708 threads) and webcracking.com (698 threads), which each have a comparable number of threads. Also, the overall number of threads is disproportionately large compared to the number of websites targeted by the config files. This is due to the fact that once a config-breaking change is made to the target website, some forum users tend to post a new config file thread instead of updating the old one.

The situation is slightly different when we look at the number of views that each thread receives. A typical config file thread on nulled.to gets 1,044 views, while threads created on cracking.org, nethingoez.com, and webcracking.com average 620, 236, and 234 views, respectively. Looking at the number of replies to each configuration file thread, we see that nulled.to leads with 56 replies per thread with nethingoez.com, cracking.org, and webcracking.com taking the second, third, and fourth places with 21, 18, and 10 replies per thread, respectively. More details can be seen in Table 1.

Table 1. Cracking forum threads and posts

	cracking.org	nethingoez.com	nulled.to	webcracking.com
Number of config file threads	3,197	708	833	698
Average number of config file thread views	620.15	235.71	1,044.39	233.90
Average number of config file thread replies	18.24	20.58	55.79	9.50
Number of config file subforum posts per user	8.21	8.86	3.24	2.69
Num. of unique users in config file subforums	7,500	1,719	14,446	2,720

Exploring the number of unique threads created by each topic starter – including website administrators – in the config file subforums, we observe the following: webcracking.com leads with 11.6 threads per user, second place is occupied by cracking.org with 7.2 threads per user, and nethingoez.com and nulled.to are last with 6.6 and 2.1 threads per user, respectively. On the other hand, when it comes to the number of thread creators in the configuration file subforums, cracking.org leads with 446 unique users and is followed by nulled.to, nethingoez.com, and webcracking.com with 390, 107, and 60 unique thread creators, respectively.

Further, we observe a small but extremely active set of users, most of whom are website administrators, on all four web forums. Combined, the config file threads created by those users are as numerous as all config file threads created by 98% of users across all four underground forums. In other words, the vast majority of thread creators in the configuration file subforums tend to create very few threads – between one and 19 – while a few select users are responsible for the creation of dozens and even hundreds of different threads.

Interestingly, the thread posting activity is somewhat similar across all forums in that there are short periods of high activity, such as 10–20 new threads posted in a 24–48 hour period, followed by several weeks of moderate to low activity with only a few config file threads posted per day.

2.5 Posts

When looking at the posting activity on the config file subforums, we see that a typical nethingoez.com user[2] has 8.9 posts/messages under their belt, followed by cracking.org, nulled.to, and webcracking.com users with 8.2, 3.2, and 2.7 posts, respectively (Table 1). In other words, it is safe to say that a typical user downloads between 2.7 and 8.9 config files since one has to post a reply before being able to access the thread attachments such as configuration files, and there is very little incentive for the posters to keep posting in the same config file thread once they have gained access to the attachments except to report an error, which we observed very rarely. If we expand the search to include all messages posted by the config file subforum posters on the four underground forums, we observe that nethingoez.com leads with 272 posts per user, followed by cracking.org (94 posts), nulled.to (71), and webcracking.com (67).

If we look at the individual users who post in the config file subforums, we see 7,500 unique users on cracking.org, 80 of which have more than 100 posts each, and nine have more than 200 posts each. The statistics are even grimmer for the other three forums: out of 1,719 nethingoez.com users, only five have made more than 100 posts, none have made more than 200 posts. None of the 14,446 nulled.to users have more than 100 posts under their belts, and only two out of 2,720 webcracking.com users have made more than 100 posts. Also, if we include all messages posted by the same users and not only those in the config file subforums, we see that only 22 nethingoez.com users, 20 cracking.org users, three webcracking.com users, and two nulled.to users have posted more than 2,000 messages each, with the vast majority of all users having posted fewer than 200 messages. Essentially, this continues to highlight how although a small core are very active, the vast majority of users are generally content to interact infrequently on each website. Our observations coincide with previous studies on the subject [36].

[2] In this Section we are looking at the users who posted at least one message in the config file subforums since we are unable to get the data on those who do not post any messages.

2.6 Post Content

When looking at the messages posted by users, we observe that the vast majority of the posts are non-informative and appear to have been made to satisfy web forums' requirements for accessing the content attached to the original posts (OP). Examples of such messages include: "thanks man," "thank for sharing," "thanks bro," "thxxxxxxxxxxxxx," and "thank for share." We also saw several instances of posters asking for help or reporting a config file that is no longer working due to the recent changes made by Facebook/eBay/etc. However, more often than not, such posts were left un-addressed. This further solidifies our view that most users on these websites are in the "script kiddie" mold of miscreant rather than a more nefarious and skilled hacker.

3 Cracking Tools

All cracking tools that we tested work in a similar manner. First, the user must configure the tool, which typically includes loading the config file (or specifying the parameters manually), selecting the word list to use (which is a text file containing username/password combinations), specifying the keywords for success and failure, loading the proxy list, and selecting the number of threads to use. The tool then sets up the connection by completing the three-way TCP handshake and starts to send HTTP or HTTPS packets to the target website (usually to the login page) with the credentials from the word list. After that, the cracking tool parses the HTML response it receives from the target website and determines whether the credentials are valid or not by looking for success and failure keywords specified earlier.

3.1 Cracking Tool Popularity

Looking at the number of threads dedicated to each cracking tool, we see that Sentry MBA is the most popular one across all forums. This makes sense based on the fact that it is free, relatively stable, has an intuitive graphical user interface, and is one of the oldest crackers in our test, with the first beta version of the original Sentry, the predecessor of Sentry MBA, dating back to April 25, 2003 [45]. Vertex, Account Hitman, and Apex occupy the second, third, and fourth places, interchangeably. Interestingly, all paid cracking tools that we tested – AIOHNB, vCrack, and Multi-Hacker – are orders of magnitude less popular than their free counterparts. One reason for this could be that both AIOHNB and Multi-Hacker do not support config files and, compared to Sentry MBA and Account Hitman, it is considerably more difficult to create a config file for vCrack. In addition, although we were not able to identify cracking tools' names in most of the nulled.to threads, a manual analysis of a 50-thread sample suggests that 98% or more of them are Sentry MBA. More details can be seen in Table 2. Also, due to the underground forums' structure, we had to group several cracking tools, namely EZLeecher, Forum Leecher, ZLeecher, and Fj Leecher, into one supergroup called "Leechers".

Table 2. Cracking tool threads

Cracking tool	cracking.org	nethingoez.com	nulled.to	webcracking.com
Sentry MBA	2,500	579	157	374
Vertex	196	–	–	72
Account Hitman	113	60	–	60
Apex	72	–	–	46
AIO Checker	40	–	–	66
AIOHNB	26	–	–	25
Leechers	16	–	–	27
E.F.R Checker	37	–	–	–
Sparta	31	–	–	–
Other	52	1	1	26
Unknown	114	68	675	2

3.2 Websites Targeted by the Config Files

An analysis of thread titles and attachments across all forums reveals that file sharing and downloading services, such as uploaded.net, 1fichier.com, and real-debrid.com are the most popular targets for the configuration files. Gaming websites and distribution platforms such as leagueoflegends.com, store.steampowered.com, and origin.com take a distant second place. Third place is occupied by adult-oriented websites. More details can be seen on Fig. 1.

When looking at each forum individually, we observe that, contrary to the other three forums, gaming website config files are much more popular on nulled.to than any other category. In contrast, file sharing and adult config file threads are the most pandered about on nethingoez.com and webcracking.com. The gaming website threads are few and far between. Another interesting finding was that fast food restaurants had more config file threads created for them than security software and financial services websites.

At first glance, one might be surprised that shopping and payment/financial services websites such as amazon.com, ebay.com, paypal.com, and wellsfargo.com are not very popular on these forums even though they [arguably] provide the highest return on investment. However, a brief look over several cracking tool discussion subforums would explain such low popularity of config files for payment and financial services websites – apparently, unlike most file sharing and adult sites, large banks and online shopping websites go after the miscreants who use cracking tools against their websites. In fact, a more in-depth search reveals a few posts by people who allegedly served time in jail for trying to brute-force online banking accounts.

Also, it must be noted that although we were able to categorize the majority of websites targeted by the configuration files, approximately 43% of thread titles could not be easily converted to one of the categories. Consequently, such thread

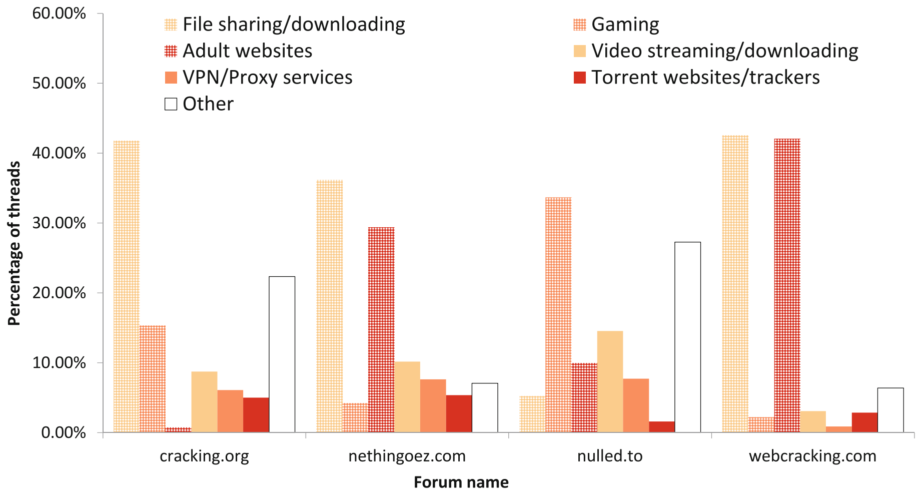


Fig. 1. Websites targeted by the config files

titles had to be omitted. Still, we believe the results reported in this section are representative of the population of websites that crackers target.

3.3 Overview of Cracking Tools' Functionality

To gain a better understanding of cracking software, we created fake accounts on several online social networks and used the most popular free and paid cracking tools to crack them. We chose not to test cracking tools on websites such as bankofamerica.com and ebay.com because, as discussed earlier, there have been several reports in underground forums of banks and large corporations pursuing individuals who tried to brute-force their customers' accounts. The free crackers studied include Sentry MBA, Account Hitman, and Vertex, which are the top three most popular cracking tools. The paid cracking programs were AIOHNB, vCrack, and Multi-Hacker. Similar to other cracking/hacking tools found in underground forums and marketplaces, the crackers we tested are for Microsoft Windows operating systems only. Also, at the beginning of our study, both AIOHNB and vCrack were paid tools. However, starting with version 2.7.0, the former appears to no longer require paid online activation [14] and the latter became open source on April 22, 2016 [16].

Furthermore, although most of the cracking tools we tested have a wide range of features, such as the ability to test proxy servers, check the validity of email accounts, and even optical character recognition (OCR) functionality to bypass CAPTCHAs, we concentrated our efforts on testing their abilities to check credentials.

Our first observation is that neither free nor paid cracking tools are particularly user friendly. One free and one paid cracking tool – Vertex and AIOHNB, respectively – refused to run unless additional files were downloaded (Fig. 2b, c and d). Interestingly enough, initially AIOHNB refused to work claiming that it required additional 'framework' files to operate. The cracker prompted us to install the said files (Fig. 2b), which implies there is a possibility that the miscreant could be downloading malicious files targeting themselves. Once all the required files for Vertex and AIOHNB were installed, we were able to start their graphical interfaces. This is a lot of trouble to go through when another cracking tool could be utilized instead.

Upon launching the cracking programs, we observed that half of them try to listen on a local port or issue HTTP GET requests. For example, Sentry MBA sends TCP packets to dyndns.com on port 80 to determine the external IP address of the machine. Account Hitman, on the other hand, does not send outgoing packets and only attempts to listen on TCP port 13121. Finally, AIOHNB tries to connect to cpc-prod3.canardpc.com on TCP port 80. Vertex, vCrack, and Multi-Hacker neither attempted to listen on a local port nor sent any outgoing packets.

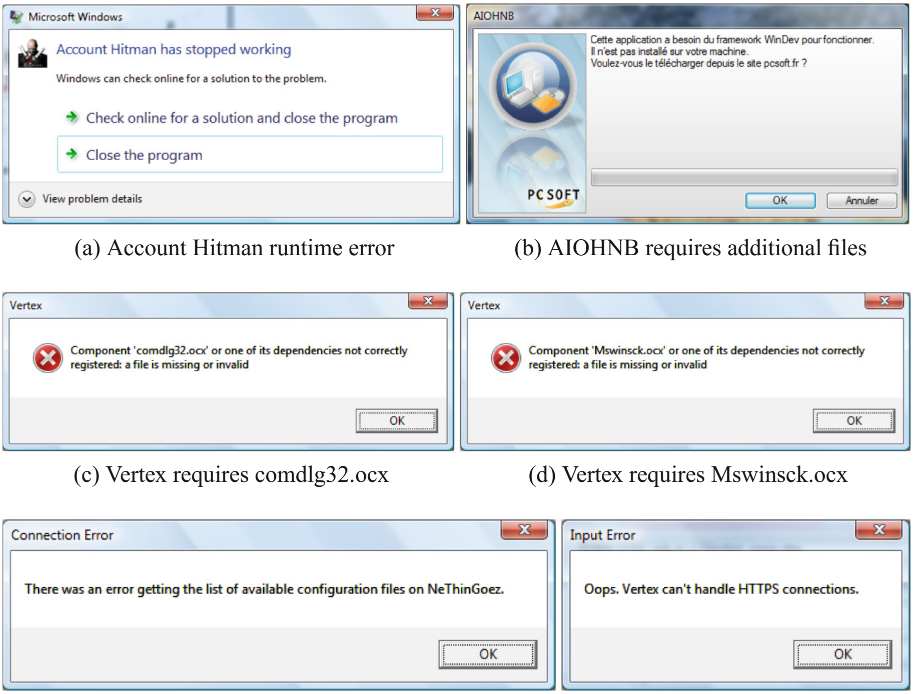
3.4 Issues Encountered

At a glance, both free and paid cracking tools appear to have nice, clean, easy-to-use interfaces. Additionally, they feature a wide array of settings and features ranging from the ability to use regular expressions to extract desired information from brute-forced accounts, such as street addresses and phone numbers, to automatic configuration file downloads directly from the graphical interface.

However, appearances can be deceiving. Upon closer examination, we found advertised features to be broken and others not operating as expected. During our testing, Account Hitman crashed on regular basis, an example of which can be seen on Fig. 2a. Vertex, on the other hand, refused to download updates or configuration files (Fig. 2e). A sleek UI cannot cover up the inability for these tools to function reliably.

In addition, despite being the most polished and widely used of the bunch, Sentry MBA had issues using custom HTTP headers, which require critical updates to circumvent server-based defense mechanisms. Fields such as *Referer*, *Accept-Language*, and *Cookie* could be easily changed via Sentry MBA to match those of any browser. However, using a custom *Accept-Encoding* header field breaks the TCP packet generated by Sentry MBA. Furthermore, we had to restart Sentry MBA several times during testing since it would sometimes refuse to use newly changed settings and would keep resetting itself to the old configuration. In all cases, a restart would solve such problems.

The paid cracking tools were not much better. vCrack for example, refused to work unless the number of threads numbered in the double digits. That is, it would not work with 1, 2, or 3 threads, but would run with 01, 02, and 10 threads. In addition, although vCrack would issue an HTTP GET request once we specified the correct thread number, it would not work as intended and would always claim that it verified the credentials for 0 user accounts even when supplied with valid username and password combinations. Further, although AIOHNB was the most feature-rich paid tool in our test with URL grabbing,



Account Hitman has stopped working

Windows can check online for a solution to the problem.

Cette application a besoin du framework WinDev pour fonctionner.
Il n'est pas installé sur votre machine.
Voulez-vous le télécharger depuis le site pcsoft.fr ?

Component 'comdlg32.ocx' or one of its dependencies not correctly registered: a file is missing or invalid

Component 'Mswinsck.ocx' or one of its dependencies not correctly registered: a file is missing or invalid

There was an error getting the list of available configuration files on NeThinGoez.

Oops. Vertex can't handle HTTPS connections.

(a) Account Hitman runtime error

(b) AIOHNB requires additional files

(c) Vertex requires comdlg32.ocx

(d) Vertex requires Mswinsck.ocx

(e) Vertex is unable to download config files

(f) Vertex cannot handle HTTPS

Fig. 2. Cracking tool crashes, errors, and notifications

proxy testing, and email checking modules, it had its fair share of issues – such as poor or missing translation, one example of which can be seen on Fig. 2b – even though it required additional files to "run" as described in Sect. 3.3.

Multi-Hacker was also not without its faults. When we tried to crack our Skype account, it claimed that the crack was successful despite the fact that we supplied it with invalid account names. We believe that an outdated Skype module of Multi-Hacker is to blame since Multi-Hacker actually worked on our Facebook and Instagram accounts.

3.5 Feature Comparison

Sentry MBA, Account Hitman, and Vertex all have very similar features, including multithreading support, the ability to use proxies, the ability to use and edit custom configuration files, and the ability to change the *User-Agent* HTTP header field. However, there are quite a few differences between these free cracking tools – some of which we believe to be responsible for that particular cracking tool's popularity (or lack thereof) – that warrants an explanation.

Despite being the most popular and stable of the bunch, Sentry MBA is the only free cracking tool in our test that does not check for updates, whether automatically or via a button click. Although Vertex has the least number of features compared to other free tools, it is the only cracker that supports direct download of configuration files from the underground forum nethingoez.com. Unfortunately, this feature was not working in our test (Fig. 2e), most likely due to the fact that nethingoez.com made several changes to its website during our data collection period.

When looking at the software's ability to create and edit configuration files, we found Sentry MBA's configuration file editor and creator more sophisticated than the one in Vertex, although we felt that it was not as intuitive to use as Account Hitman's.

While testing three crackers against our fake accounts, we noticed that there was virtually no difference in speed between them, and both Sentry MBA and Account Hitman correctly reported the results when supplied with both valid and invalid credentials. Vertex, on the other hand, reported all supplied username/password combinations as valid, despite the fact that half of them were invalid. Furthermore, Vertex refused to work with HTTPS websites (Fig. 2f), which, combined with the previously mentioned issues, make it the most buggy free cracking tool in our test.

Overall, out of free cracking tools, only Sentry MBA and Account Hitman were able to successfully verify credentials to our fake user accounts. They are also somewhat more polished and offer more features than the rest, which explains their popularity. However, it has to be mentioned that a large number of posters in underground forums have had success with Vertex, even though we did not. One possible explanation for this split could be that it needs the access to nethingoez.com in order to function properly. In addition, Vertex had not been updated recently, which, combined with the fact that all but one underground website in our data set had either changed domain names or modified their code during the data collection period, could have resulted in the cracker's failure that we encountered during testing as the program was looking for information that had either been moved or deleted.

Interestingly, we found most paid cracking tools lacking in features compared to their free counterparts. For example, basic functionality such as *User-Agent* selection and pre- and post-login page actions were nowhere to be found in vCrack and Multi-Hacker. Furthermore, although both AIOHNB and Multi-Hacker feature a number of pre-built modules for popular websites like reddit.com and instagram.com, they neither support the external config files nor allow users to make any changes to the built-in modules, which will render the current tool versions useless once the target websites change their login pages. vCrack is the only paid tool in our test that supports the addition of external modules, although the process of creating a new module is much more involved compared to creating a config file for a free cracking tool. In our opinion, one of the very few advantages of the paid cracking tools over their free counterparts is the simplicity of use – one simply has to select the desired module, load the credential list, and click *Start*.

We were also surprised that the paid tools we tested had as many issues as the free cracking tools. Not only that, but vCrack was unable to verify any credentials in our tests, which might explain why the developer has chosen to make it open source.

In addition, it has to be mentioned that we tested all cracking tools on a small set of online social network websites to avoid unpleasant conversations with the authorities. As a result, it is possible that some cracking tools in our test would perform significantly better on websites like ebay.com, bankofamerica.com, and origin.com.

4 Detecting Cracking Tools

The best way to detect cracking tools attempting to access a website is to inspect the packets created by the crackers. Unfortunately, to an administrator or a security specialist who monitors the target website's traffic, the packets generated by the cracking tool would look almost identical to the packets generated by the popular browsers. Another method involves analysis of traffic and behavioral abnormalities, such as a large number of packets being sent from the same IP addresses over a short period of time and disproportionate number of login page requests from the same IP address compared to other web pages. Unfortunately, there are two disadvantages to such approaches. First, a miscreant can easily modify the timeout settings in the cracking tool and, instead of sending a packet every 30 seconds, the tool would wait for several hours in between the requests, which would make it very difficult to detect. Second, even if one could somehow find the pattern in the packet timestamps or user behavior, there would be no way for them to tell whether those requests were generated by a cracking tool or by a browser automator like Selenium.

In this study, we use a modification of the first approach – we capture the packets generated by the cracking tools and use the data from several protocol layers to differentiate between the cracking and legitimate packets. By being more detailed, our methodology focuses on identifying specific differences that can enable operators to flag cracking traffic over legitimate traffic.

4.1 Experimental Setup

A brief analysis of the packets generated by the cracking tools showed very little variation in terms of size and header values between each tool. Furthermore, there were not any noticeable differences when we compared them to the packets generated by several popular Internet browsers. Clearly, a more in-depth analysis was called for.

We started by creating a simple website with an HTML login form which would accept only one value for username and a password. If the supplied credentials are correct, the website would show fake user information, including name, address, and a phone number; otherwise, a short error message would be displayed on the HTML page. We then hosted this website on our own server and made sure that a Wireshark [22] instance was running in the background collecting packets.

We wanted to get both HTTP and HTTPS packet samples, which meant decrypting SSL/TLS data. To achieve this, several modifications had to be made. First, the Apache server had to be forced to use the weakest possible encryption by modifying the *SSLCipherSuite* parameter in the configuration file. Wireshark settings also had to be changed so that it would use the Apache's private RSA key to decrypt all HTTPS traffic.

4.2 Cracking Tool Packet Capture Methodology

All cracking tools were tested on a 64-bit version of Windows Vista SP2. Unfortunately, we were not able to test the three paid tools since all of them come with a pre-defined number of modules for popular websites like Facebook, Twitter, and Instagram, which makes it very difficult (impossible in the cases of Multi-Hacker and AIOHNB) to add a new website module. Using the existing modules would not work since all of them are for pre-defined HTTPS websites only. As a result, we were left with Sentry MBA, Account Hitman, and Vertex for packet generating purposes.

For each cracking tool, we changed the settings in such a way that we would get as many different packets as possible. For example, if a cracking tool worked with SSL, had several pre-defined *User-Agent* fields, and supported both GET and POST HTTP requests, then we would generate the packets for all possible combinations, such as HTTP POST request over an SSL connection with the first pre-defined *User-Agent*, HTTP GET request over a non-secure connection with the second pre-defined *User-Agent* field, and so on.

To get a wide range of packet samples from organic traffic, we used several versions of seven popular browsers and five different computers and virtual machines to simulate traffic of an average Internet user. The operating systems used ranged from Windows XP to Windows 10 to GNU/Linux, while the browsers included Firefox, Opera, Chrome/Chromium, Internet Explorer/Edge, SeaMonkey, K-Meleon, and Midori.

Overall, we captured 39 cracking packets generated by the cracking tools and 39 legitimate packets from the browsers, yielding 78 packets for subsequent analysis.

4.3 Packet Comparison

At first glance, the packets generated by the cracking tools look virtually identical to each other and to the packets created by the browsers. However, a closer examination reveals several differences between the legitimate and cracking packets.

We observe that, on average, the packets created by the cracking tools are 28% smaller than their legitimate counterparts. This difference is mostly due to the smaller HTTP payload in the cracking packets.

Moving down the layers, we see that both Ethernet and IP packet headers generated by the three cracking tools are virtually identical to each other as well as to the legitimate packets (with the exception of the IP length header field, which was explained above), which is what one would expect as developers generally let the networking libraries handle lower level packet creation.

Looking at the TCP header, we see that all packets are very similar with the exception of the source port numbers, options, and window size. Of these, only the last two are of interest to us. The difference in TCP options comes from the fact that, contrary to all cracking tools and browsers running on Windows, all GNU/Linux browsers in our test chose to set TCP option 8 (Timestamp). As for the TCP window size, most cracking tools in our test preferred values of 16,425, 65,040, and 65,700, while the browsers used a variety of different values, ranging from 229 to 65,568.

The most noticeable differences between the packets generated by the cracking tools and the Internet browsers are in the application layer, namely in the HTTP header. The first difference is that, in some instances, Sentry MBA uses HTTP version 1.0 while all browsers and the rest of the cracking tools use version 1.1. Furthermore, we observe that although all browsers in our test set the *Connection* field to *keep-alive*, both Account Hitman and Vertex set it to *close*. In addition, the *Accept-Language* header field varied significantly across the browsers and cracking tools, but it was also completely omitted in all packets generated by Sentry MBA. Also, the *User-Agent* field widely differed not only between legitimate and cracking packets but also between each browser instance. Finally, the HTTP *Pragma* header field was set in all packets generated by Sentry MBA while only two browser instances out of 39 used it.

When looking at the HTTP header fields which were exclusive to either cracking or legitimate packets, we observe that all browsers set the *Accept-Encoding* field while none of the cracking tools did. Further, *Accept-Charset*, *Upgrade-Insecure-Requests*, and *Cache-Control* header fields were set by six, 11, and five browser instances, respectively, while none of the cracking tools used them.

4.4 Classifier Training

Using either *Accept-Encoding* or *User-Agent* features for classifier training would give us a perfect accuracy in most machine learning algorithms since they are either unique to all approaches or provide a perfect split between browsers and cracking tools. However, we will not use them since *Accept-Encoding* and *User-Agent* header fields could be either patched by cracking tool authors or manually edited by advanced users.

Table 3 shows the features that were used to train the classifiers, which denotes that the top three most discriminating features according to both Chi-square and Information Gain tests are *Accept-Language* HTTP field, *Pragma* HTTP field, and packet size. *Accept-Charset* and *Cache-Control* HTTP header fields appear to be the least useful features according to both metrics.

Next, we used the RapidMiner data mining environment [43] to train several supervised machine-learning-based algorithms. For each classification exper-

Table 3. Classifier feature importance

Feature name	Chi-square	Information gain
HTTP Accept-Language	1.00	1.00
HTTP Pragma	0.61	0.62
Packet size	0.50	0.42
HTTP version	0.25	0.29
HTTP Upgrade-Insecure-Requests	0.13	0.15
TCP options	0.10	0.12
HTTP Connection	0.08	0.10
HTTP Accept-Charset	0.02	0.02
HTTP Cache-Control	0.0	0.00

Table 4. Classifier accuracy

Classifier	Accuracy	FP rate	FN rate
Random Forest	100.00%	0.00%	0.00%
J48	100.00%	0.00%	0.00%
PART	100.00%	0.00%	0.00%
CART	100.00%	0.00%	0.00%
Logistic Regression	98.72%	2.56%	0.00%
Neural Network	98.72%	2.56%	0.00%
Naive Bayes	75.64%	48.72%	0.00%

iment, we used a 20-fold cross-validation with stratified sampling. In a 20-fold cross-validation, the sample is divided into 20 parts: 19 parts are used as a training dataset, and the remaining part is used to test the classifier. This process is repeated 20 times, producing 20 results. The results reported subsequently are averages of the 20 runs.

The best performing algorithms were Random Forest, J48, PART, and CART, all of which had perfect accuracies. They are followed by Logistic Regression, Neural Network, and Naive Bayes, with the accuracies of 98.72%, 98.72%, and 75.64%, respectively. More details can be seen in Table 4. Also, although we do not know the exact reasons for such poor performance of the Naive Bayes classifier, one explanation could be that some of the features we used are not independent of each other given the class label, which could result in suboptimal probability estimates and wrong decisions [63].

When using AdaBoost to reduce the bias and improve the classifier accuracy even further, we observe that all classifiers' accuracies stay the same. Furthermore, in most cases the boosting was not possible due to the fact that only one classifier was used.

5 Related Works

There have been a number of studies on underground marketplaces and their economies. In what appears to be one of the first studies of modern cybercrime [35], Mann and Sutton analyzed Internet newsgroups, which are online, forum-like discussion groups where like-minded people can communicate with each other by posting messages. Mann and Sutton concentrated their efforts on two particular newsgroups: one with discussions on hacking encrypted satellite signals and another one on lock picking, safes, and other security devices. During the course of their study, the authors classify newsgroup members into categories, such as *hacker gurus*, *parasites*, *information providers*, and *money makers*. They also investigate the supply of and demand for illicit goods and services, and look into how newsgroup users with different levels of expertise interact with each other. This is in contrast to our study, where we target the subforums of four popular underground forums dedicated to cracking tools used to brute force user accounts and test stolen credentials.

A 2007 measurement study by Franklin et al. [23] focused on underground marketplaces and touched on some topics covered by our work. The authors used publicly posted IRC (Internet Relay Chat) messages to study malicious activities, such as spamming, online credential theft, and the sale of compromised hosts. They also proposed simple, low-cost countermeasures which could be used to disrupt the operations of such marketplaces. Similar studies shortly followed, with works by Cymru [15], Herley and Florêncio [26], and Fallmann et al. [20] concentrating their efforts on studying illegal IRC marketplaces. Unfortunately, not only have IRC chat rooms lost popularity among Internet users since that time, but underground black markets have also evolved from chaotic, difficult-to-control entities where there was little incentive for the miscreants not to scam each other to more orderly and better regulated marketplaces [5]. Further, the majority of these works looked at underground marketplaces as a whole. We focus on several smaller subforums, which allows for an in-depth analysis.

In [65], Zhuge et al. perform a measurement study on the underground economy within the Chinese Web. In the course of their study, the authors concentrate their efforts on underground marketplaces and their participants, which allows them to create a model describing the Chinese underground economy. Several similar and complementary studies followed, including the papers by Motoyama et al. [37], Christin [13], Yip et al. [60–62], Stone-Gross et al. [52], Garg et al. [24], Holt and Lampke [29], McCoy et al. [36], Radianti [42], Allodi et al. [6], Holt [27,28], and Sood and Enbody [51]. Our work is somewhat similar to those studies in that we also study underground marketplaces in the example of Web forums. However, unlike these works, which primarily focus on investigating the structure and organization of the underground forums as well as social interactions among their members, we look into the configuration files for cracking tools and user accounts used to share and download them.

Several studies propose various strategies for fighting cybercrime, ranging from making it more difficult and costly for the miscreants to operate to completely taking down underground communities. In [33], Leontiadis analyzes various types of online criminal networks, including underground forums and marketplaces, from both technical and economical perspectives. Leontiadis' study reveals that online criminal networks tend to have weak links, or *choke points*, which are critically-important components of online criminal networks. The author argues that targeting such components will increase criminal operational costs and reduce online crime. A somewhat similar strategy was proposed by Nadji et al. in [38] where the authors used two graph measures – graph density and eigenvector centrality – to investigate the structure of networks involved in criminal activities. The authors also analyzed different take-down strategies that could be used to shut down sophisticated criminal networks and determined that, in most cases, shutting down a few domain names would remove critical network links, thus, taking the whole criminal network down. Our work is similar to these and other studies [3,34,53,55] in that we also come up with ways to make it more difficult for the miscreants to engage in illegal activities. However, our work differs in that we are not really interested in taking down criminal networks; instead, we analyze the tools used by the criminals and develop machine learning classifiers that could be used by companies to make it more difficult and costly for the miscreants to attack them.

Furthermore, some studies survey existing methods and suggest new strategies for detecting and preventing attacks on computer networks and Web applications. Papers by Sommer and Paxson [50], Lee and Stolfo [32], and others discuss and propose data mining and machine-learning-based approaches for network intrusion detection. Other studies, such as those by Douligeris and Mitrokotsa [18], Kumar and Selvakumar [31], and Bhuyan et al. [9], investigate defense mechanisms against distributed denial of service attacks. Further, some papers, such as the ones by Wang et al. [57] and Abreu [2], propose to use Web pages with dynamically changing content to make it more difficult for the miscreants to perform automated attacks on Web applications. Finally, there are studies that discuss the effectiveness of existing techniques for stopping automated attack tools [40]. Although our study is similar to all these papers in that we investigate automated attacks carried out with the help of computer networks, our work differs in that, in addition to the analysis of underground subforums, we concentrate our efforts on detecting network packets generated by the popular cracking tools, which, to the best of our knowledge, is the first work of its kind.

There are also articles and white papers that talk about credential stuffing attacks and cracking tools like Sentry MBA and suggest defense mechanisms, such as using complex passwords, avoiding password recycling, employing JavaScript anti-bot challenges, monitoring the traffic for specific HTTP *User-Agent* fields, and paying special attention to IP addresses responsible for a large number of failed logins [4,8,10,47,54,64]. Our study differs in that, in addition to config file subforum analysis, we go much deeper in our investigations of cracking tools as well as develop classifiers capable of detecting cracking packets.

Finally, there are also a number of short papers and articles, such as an article by Shulman [49] and a paper by Yip et al. [59], which provide a brief background on the operations of underground credential markets and give insights into their economies. In addition, a recent study found that cybercrimes are similar to violent crimes in that they both carry significant indirect and defense costs [7]. This is in contrast to traditional non-violent crimes, like car theft or tax fraud, which usually carry high direct costs, such as the price of a car, and relatively low indirect costs, such as psychological trauma and lost output. Further, Shin et al. [48] studied forum automators, which the miscreants use to spam legitimate forums with unrelated messages promoting their own websites. Shin et al. discovered that forum spam automators are fairly sophisticated and include a number of features – such as the ability to automatically solve CAPTCHAs and use anonymizing proxies – which help miscreants circumvent spam prevention mechanisms and avoid blacklisting. Although not directly related to our work, such articles and papers provide valuable insights into the underground cracking economy, some of which we indirectly use in our study.

6 Discussion

6.1 Data Collection Difficulties

One of the consistent traits encountered throughout this study is the large degree of paranoia that forum operators were operating under. In particular, one of our attempts at collecting data from webcracking.com was upended when our registered user was banned from the forums for "leeching," even though we were not downloading or posting cracking configuration files. We were only browsing. Such actions clearly impact our ability to collect data in a timely and complete manner but also point to a culture of distrust on these communities.

To try to avoid these arbitrary bans, we attempted to utilize a *VIP* membership, where premium content and laxer rule enforcement were supposedly benefits. In order to gain access to the VIP section, a user must send a monthly 'donation,' e.g. of $9.95, to the head administrator of webcracking.com. For this paper, we paid for one month worth of VIP access to determine if our data collection efforts could continue or if we would still be subjected to losing our accounts to bans.

We quickly discovered that the VIP membership was subjected to similar restrictions as the free membership, even though the advertisement promised the lifting of all restrictions. Furthermore, even though we strictly adhered to the specified restrictions, our account was temporarily banned for 10 days for downloading too many configuration files. Once the ban was lifted, we reduced the number of files downloaded to one file per 2–3 days. However, the administrators still permanently banned our account and the associated IP address for downloading too many files without uploading any in return, even though nowhere in the rules did it say that we had to upload any content in addition to paying for the VIP access.

6.2 Classifier Feature Selection

It could be argued that the features we used for classifier training – most of which come from the HTTP header – could be circumvented by the cracking tool authors, rendering our classifiers out-of-date. Although it is true that a developer could modify the packets created by their cracking software to make them virtually indistinguishable from those generated by a modern browser, we find it hard to believe that this thought had not crossed the minds of the cracking tool authors, especially considering that the free cracking tools that we tested were relatively mature with numerous versions released in the past several years. If the developers wanted the HTTP headers in their software's packets to resemble those of the popular browsers, they would have done so already.

6.3 Packet Samples

We also had to create our own config files, which we did without modifying the pre-defined HTTP header fields in any of the cracking tools. As a result, it is possible that the use of some config files for twitter.com, facebook.com, and other popular websites would result in mildly different packets than the ones we used in this study. Additionally, due to the difficultly of decrypting SSL traffic, we were unable to identify the encrypted payloads being sent to certain websites. Based on this shortcoming, we used our own website to get samples of legitimate and cracking packets, which we believe is representative of the packets that would be observed at encrypted websites, although we cannot know with absolutely certainty that this is the case.

6.4 Ethical Issues

In order to gain access to configuration file subforums, we had to post several messages from our underground forum accounts. In addition, for each downloaded configuration file, we were required to post at least one message and/or click on the *thank you* button. We strongly believe that none of these actions had a measurable effect on the underground forum economy.

On the other hand, paying $9.95 for one month worth of VIP access certainly did affect the underground forum economy – it made the cyber criminal(s) running the webcracking.com underground forum $9.95 richer. Furthermore, we violated the terms of use of several legitimate websites by creating fake accounts and carrying out credential stuffing attacks against them. Although these actions might be viewed as unethical, they were paramount to this study. Our actions could be compared to doctors and scientists running experiments on animals – although the lab animals suffer and often die painful deaths, the results of such experiments are used to save and improve human lives, which most consider a fair trade-off. Similarly, although it is unfortunate that we violated the terms of use of several websites and made the cyber criminals $9.95 richer, we feel that the benefits of our work far outweigh any moral or ethical concerns raised by it.

6.5 Future Work

Due to the difficulty and risks of collecting a large sample of cracking tools' packets, we were not able to test our classifiers on the real-world data. To rectify this, in the future we contemplate purchasing a dozen more cracking tools as well as downloading older versions of Sentry MBA, Account Hitman, and Vertex. For legitimate packets, we are considering including mobile browsers' packets as well as adding more flavors of GNU/Linux operating systems to our tests. Finally, we are planning on contacting Twitter, Instagram, and Facebook and asking for access to their decrypted traffic. This should give us a much larger sample of both cracking and legitimate packets, and allow us to test the performance of our machine learning algorithms in the wild, which appear to be very promising in preventing cracking tool based threats.

References

1. Sentry MBA (2016). https://sentry.mba/tool?id=1. Accessed 28 May 2017

2. Abreu, L.P.B.: Morphing Web Pages to Preclude Web Page Tampering Threats (2016)
3. Afroz, S., Garg, V., McCoy, D., Greenstadt, R.: Honor among thieves: a common's analysis of cybercrime economies. In: Proceedings of the eCrime Researchers Summit (eCRS). IEEE (2013)
4. Agarwal, S.: The Half-Day Attack: From Compromise to Cash with Sentry MBA (2016). https://goo.gl/Yb08S9. Accessed 01 June 2017
5. Allodi, L., Corradin, M., Massacci, F.: Then and now: on the maturity of the cybercrime markets (the lesson that black-hat marketeers learned). IEEE Trans. Emerg. Top. Comput. **1**, 1 (2015)
6. Allodi, L., Shim, W., Massacci, F.: Quantitative assessment of risk reduction with cybercrime black market monitoring. In: Proceedings of the Security and Privacy Workshops (SPW). IEEE (2013)
7. Anderson, R., et al.: Measuring the cost of cybercrime. In: Böhme, R. (ed.) The Economics of Information Security and Privacy, pp. 265–300. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39498-0_12
8. Ben-Meir, E.: Sentry MBA: A Tale of the Most Popular Credential Stuffing Attack Tool (2017). https://goo.gl/bFDn1b. Accessed 01 June 2017
9. Bhuyan, M.H., Kashyap, H.J., Bhattacharyya, D.K., Kalita, J.K.: Detecting distributed denial of service attacks: methods, tools and future directions. Comput. J. **57**(4), 537–556 (2014)
10. Bleau, H.: Credential Checking Services Soar in Popularity on Dark Web (2016). https://goo.gl/yq3Vxf. Accessed 01 June 2017
11. Buddah: Vertex 1.0.4 (2016). https://goo.gl/yORQUV. Accessed 28 May 2017
12. Burgess, M., Temperton, J.: The security flaws at the heart of the Panama Papers (2016). https://goo.gl/b49RaQ. Accessed 28 Oct 2016
13. Christin, N.: Traveling the Silk Road: a measurement analysis of a large anonymous online marketplace. In: Proceedings of the International Conference on World Wide Web (WWW). ACM (2013)
14. ConfigMasta: AIOHNB tool v 2.7.8 [Full version] (2016). https://goo.gl/PjYLl2. Accessed 28 May 2017
15. Cymru, T.: The underground economy: priceless. Technical report, Login: 31(6) (2006)
16. DavePS: voidproducts (2016). https://goo.gl/GaIKik. Accessed 28 May 2017
17. Davey, M.: Red Cross Blood Service data breach: personal details of 550,000 blood donors leaked (2016). https://goo.gl/ls3ZJM. Accessed 28 Oct 2016
18. Douligeris, C., Mitrokotsa, A.: DDoS attacks and defense mechanisms: classification and state-of-the-art. Comput, Netw. **44**(5), 643–666 (2004)
19. Drašar, M.: Behavioral detection of distributed dictionary attacks. Ph.D. thesis, Masaryk University, Brno, Czech Republic (2015)
20. Fallmann, H., Wondracek, G., Platzer, C.: Covertly probing underground economy marketplaces. In: Kreibich, C., Jahnke, M. (eds.) DIMVA 2010. LNCS, vol. 6201, pp. 101–110. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14215-4_6
21. Fiegerman, S.: Yahoo says 500 million accounts stolen (2016). https://goo.gl/EjJfTt. Accessed 28 Oct 2016
22. Foundation, W.: Wireshark - Go deep (2017). https://www.wireshark.org/. Accessed 20 May 2017
23. Franklin, J., Paxson, V., Perrig, A., Savage, S.: An inquiry into the nature and causes of the wealth of internet miscreants. In: Proceedings of the Conference on Computer and Communications Security (CCS). ACM (2007)

24. Garg, V., Afroz, S., Overdorf, R., Greenstadt, R.: Computer-supported cooperative crime. In: Böhme, R., Okamoto, T. (eds.) FC 2015. LNCS, vol. 8975, pp. 32–43. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-47854-7_3
25. H3God: >> MULTIHACKER << || HACK FB/IG/TWITTER/REDDIT/SKYPE + MORE ACCOUNTS - 8 HACKERS IN 1 (2016). https://goo.gl/WvRu1Z. Accessed 28 May 2017
26. Herley, C., Florêncio, D.: Nobody sells gold for the price of silver: dishonesty, uncertainty and the underground economy. In: Moore, T., Pym, D., Ioannidis, C. (eds.) Economics of Information Security and Privacy, pp. 33–53. Springer, Boston (2010). https://doi.org/10.1007/978-1-4419-6967-5_3
27. Holt, T.J.: Examining the forces shaping cybercrime markets online. Soc. Sci. Comput. Rev. **31**(2), 165–177 (2013)
28. Holt, T.J.: Exploring the social organisation and structure of stolen data markets. Global Crime **14**(2–3), 155–174 (2013)
29. Holt, T.J., Lampke, E.: Exploring stolen data markets online: products and market forces. Crim. Justice Stud. **23**(1), 33–50 (2010)
30. ImadTheMAD: Introduction: What is Account Hitman? (2011). https://goo.gl/i1dZhj. Accessed 28 May 2017
31. Kumar, P.A.R., Selvakumar, S.: Distributed denial of service attack detection using an ensemble of neural classifier. Comput. Commun. **34**(11), 1328–1341 (2011)
32. Lee, W., Stolfo, S.J.: Data mining approaches for intrusion detection. In: Proceedings of the USENIX Security Symposium (1998)
33. Leontiadis, N.: Structuring disincentives for online criminals. Ph.D. thesis, Carnegie Mellon University Pittsburgh, PA (2014)
34. Li, W., Chen, H.: Identifying top sellers in underground economy using deep learning-based sentiment analysis. In: Proceedings of the Intelligence and Security Informatics Conference (JISIC), pp. 64–67. IEEE (2014)
35. Mann, D., Sutton, M.: NETCRIME more change in the organization of thieving. Br. J. Criminol. **38**(2), 201–229 (1998)
36. McCoy, D., Pitsillidis, A., Jordan, G., Weaver, N., Kreibich, C., Krebs, B., Voelker, G.M., Savage, S., Levchenko, K.: PharmaLeaks: understanding the business of online pharmaceutical affiliate programs. In: Proceedings of the USENIX Security Symposium (2012)
37. Motoyama, M., McCoy, D., Levchenko, K., Savage, S., Voelker, G.M.: An analysis of underground forums. In: Proceedings of the SIGCOMM Internet Measurement Conference (IMC). ACM (2011)
38. Nadji, Y., Antonakakis, M., Perdisci, R., Lee, W.: Connected colors: unveiling the structure of criminal networks. In: Stolfo, S.J., Stavrou, A., Wright, C.V. (eds.) RAID 2013. LNCS, vol. 8145, pp. 390–410. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41284-4_20
39. Nakashima, E.: Russian government hackers penetrated DNC, stole opposition research on Trump (2016). https://goo.gl/IKkgjt. Accessed 28 Oct 2016
40. Ollmann, G.: Stopping automated attack tools. Whitepaper-NGS software insight security research (2005)
41. Pleasant, R.: Ubuntu Forums data breach exposes 2 million users (2016). https://goo.gl/IZJc0b. Accessed 28 Oct 2016
42. Radianti, J.: A study of a social behavior inside the online black markets. In: Proceedings of the International Conference on Emerging Security Information Systems and Technologies (SECURWARE). IEEE (2010)

43. Ritthoff, O., Klinkenberg, R., Fischer, S., Mierswa, I., Felske, S.: Yale: yet another learning environment. In: Proceedings of the Tagungsband der GI-Workshop-Woche Lernen - Lehren - Wissen - Adaptivitat (LLWA) (2001)
44. Sang-Hun, C.: North Korea Stole Data of Millions of Online Consumers, South Says (2016). https://goo.gl/Ul7dmo. Accessed 28 Oct 2016
45. sentinel.deny.de: Sentry Readme (2003). http://sentinel.deny.de/ReadmeSentry.txt, https://goo.gl/eiTdBL. Accessed 01 June 2017
46. sentinel.deny.de: Sentry (2016). http://sentinel.deny.de/sentry.php, https://goo.gl/Dw2l3k. Accessed 01 June 2017
47. Shadows, D.: Protect Your Customer and Employee Accounts: 7 Ways To Mitigate the Growing Risks of Account Takeovers (2017). https://goo.gl/xrfhaO. Accessed 01 June 2017
48. Shin, Y., Gupta, M., Myers, S.: The nuts and bolts of a forum spam automator. In: Proceedings of the Conference on Large-scale Exploits and Emergent Threats (LEET). USENIX Association (2011)
49. Shulman, A.: The underground credentials market. Comput. Fraud Secur. **2010**(3), 5–8 (2010)
50. Sommer, R., Paxson, V.: Outside the closed world: on using machine learning for network intrusion detection. In: Proceedings of the Symposium on Security and Privacy (SP). IEEE (2010)
51. Sood, A.K., Enbody, R.J.: Crimeware-as-a-service a survey of commoditized crimeware in the underground market. Int. J. Crit. Infrastruct. Prot. **6**(1), 28–38 (2013)
52. Stone-Gross, B., Abman, R., Kemmerer, R.A., Kruegel, C., Steigerwald, D.G., Vigna, G.: The underground economy of fake antivirus software. In: Schneier, B. (ed.) Economics of Information Security and Privacy III, pp. 55–78. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-1981-5_4
53. Stringhini, G., Wang, G., Egele, M., Kruegel, C., Vigna, G., Zheng, H., Zhao, B.Y.: Follow the green: growth and dynamics in twitter follower markets. In: Proceedings of the Internet Measurement Conference (IMC). ACM (2013)
54. Thee, D.: Sentry MBA: A Tale of the Most Widely Used Credential Stuffing Attack Tool (2017). https://goo.gl/n8XY1U. Accessed 01 June 2017
55. Thomas, K., McCoy, D., Grier, C., Kolcz, A., Paxson, V.: Trafficking fraudulent accounts: the role of the underground market in Twitter spam and abuse. In: Proceedings of the Conference on Security (SEC). USENIX Association (2013)
56. Wagner, J.: Reset Those Passwords - Again: Over 6 Million ClixSense Users Compromised by Data Breach (2016). https://goo.gl/YBnkOL. Accessed 28 Oct 2016
57. Wang, X., Kohno, T., Blakley, B.: Polymorphism as a defense for automated attack of websites. In: Boureanu, I., Owesarski, P., Vaudenay, S. (eds.) ACNS 2014. LNCS, vol. 8479, pp. 513–530. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07536-5_30
58. Williamson, W.: What Happens to Stolen Data After a Breach? (2014). https://goo.gl/0ByDhi. Accessed 30 May 2017
59. Yip, M., Shadbolt, N., Tiropanis, T., Webber, C.: The digital underground economy: a social network approach to understanding cybercrime. In: Digital Futures 2012: The Third Annual Digital Economy All Hands Conference (2012)
60. Yip, M., Shadbolt, N., Webber, C.: Structural analysis of online criminal social networks. In: 2012 IEEE International Conference on Intelligence and Security Informatics (ISI), pp. 60–65. IEEE (2012)
61. Yip, M., Shadbolt, N., Webber, C.: Why forums?: an empirical analysis into the facilitating factors of carding forums. In: Proceedings of the Annual Web Science Conference (WebSci). ACM (2013)

62. Yip, M., Webber, C., Shadbolt, N.: Trust among cybercriminals? Carding forums, uncertainty and implications for policing. J. Polic. Soc. **23**(4), 516–539 (2013)
63. Zaidi, N.A., Cerquides, J., Carman, M.J., Webb, G.I.: Alleviating naive Bayes attribute independence assumption by attribute weighting. J. Mach. Learn. Res. **14**(1), 1947–1988 (2013)
64. Zavodchik, M.: Mitigating "Sentry MBA" - Credentials Stuffing Threat (2017). https://goo.gl/1JT0dQ. Accessed 01 June 2017
65. Zhuge, J., Holz, T., Song, C., Guo, J., Han, X., Zou, W.: Studying malicious websites and the underground economy on the chinese web. In: Johnson, M.E. (ed.) Managing Information Risk and the Economics of Security, pp. 225–244. Springer, Boston (2009). https://doi.org/10.1007/978-0-387-09762-6_11