Measuring Cybercrime as a Service (CaaS) Offerings in a Cybercrime Forum

Ugur Akyazi
Delft University of Technology
u.akyazi@tudelft.nl

Michel van Eeten
Delft University of Technology
m.j.g.vanEeten@tudelft.nl

Carlos H. Gañán
Delft University of Technology
c.hernandezganan@tudelft.nl

ABSTRACT

The emergence of Cybercrime-as-a-Service (CaaS) is a critical evolution in the cybercrime landscape. A key area of research on CaaS is where and how the supply of CaaS is being matched with demand. Next to underground marketplaces and custom websites, cybercrime forums provide an important channel for CaaS suppliers to attract customers. Our study presents the first comprehensive and longitudinal analysis of types of CaaS supply and demand on a cybercrime forum. We develop a classifier to identify supply and demand for each type and measure their relative prevalence and apply this to a dataset spanning 11 years of posts on Hack Forums, one of the largest and oldest ongoing English-language cybercrime forum on the surface web. Of 28 known CaaS types, we only found evidence for only 9 of these in the forum. We saw no dramatic shifts in these offerings over time, not even after major underground marketplaces were being seized by law enforcement. Around 16% of first posts of the threads in the 'Market' section of the forum offers CaaS, whereas only 3% is focused on product-type criminal offerings. Within the types of CaaS, 'bot/botnet as a service', 'reputation escalation as a service' and 'traffic as a service' categories make up the majority (over 60%) for whole period in terms of both supply and demand. At least half of each CaaS offerings directs potential buyers to an instant messaging app or private message for transacting privately. In sum, we find that forums do in fact provide a channel for CaaS supply and demand to meet, but we see only a fraction of the CaaS landscape and there is no evidence in our data for the supposed growth of CaaS over time. We reflect on the implications of our findings for developing effective disruption strategies by law enforcement.

CCS CONCEPTS

• **Security and privacy → Economics of security and privacy**.

KEYWORDS

Cybercrime as a Service, CaaS, Cybercrime Forum, Underground Forums, Machine Learning, Natural Language Processing

1 INTRODUCTION

The rise of Cybercrime-as-a-Service (CaaS) is seen as a critical evolution in the cybercrime landscape. In Europe and elsewhere, its disruption has been marked the top priority for law enforcement [16]. Analogous to cloud services in legitimate markets, like platform-as-a-service, CaaS enables criminal entrepreneurs to develop and manage their business without the complexity of building and maintaining all required expertise, infrastructure and tools themselves. To illustrate, one study [4] compared services and prices from five leading CaaS providers on dark web and found that custom spyware could be purchased for around $200 and a month of SMS spoofing for only $20. In addition to technical tools, it is also possible to hire services, such as targeted account takeover [35]. The overall impact of CaaS is to make cybercrime more accessible to new criminals, as well as to support business models for advanced criminals via specialized business-to-business services [22].

A key area of research on CaaS is where and how the supply of these services is being matched with demand. This question is critical for developing effective disruption strategies by law enforcement. Simply put: how do CaaS suppliers find their customers? One of the promises of CaaS is that it is accessible for new entrants, so it cannot operate effectively within old and constrained model of closed, vetted and trust-based criminal networks. Various studies show that cybercriminals have been using open channels like online underground marketplaces, custom websites and forums to advertise and sell their services. Each type of channel has its pros and cons. Underground marketplaces were found to have only modest volumes of CaaS [58]. In general, their transaction mechanisms are more suited for products than for service-based models. Law enforcement seizures have further undermined the position of these marketplaces (e.g., Alphabay, Hansa, Dream [19]). Custom websites, such as those for booters [12, 25, 26, 39] work well with service subscriptions. These sites, however, operate outside of the reliable reputation mechanisms of marketplaces and forums; therefore are rife with fraudsters [35].

In light of the problems with underground marketplaces and custom sites, forums offer an alternative channel for CaaS offerings. Monitoring these forums thus offers us useful insight into the state of the CaaS ecosystem [7, 49]. Criminal activity is distributed over a variety of forums which might show different patterns. We distinguish between closed, vetted forums and open, freely accessible forums. The former are, by necessity, smaller and more difficult to enter. This limits their reach as a marketplace. Open forums are accessible to many more (potential) cybercriminals. If CaaS is indeed overtaking the criminal ecosystem, we should see its presence also increase in open forums. Hack Forums (HF) has a prominent position in the ecosystem of open forums. It is one of the most popular English-language hacking sites worldwide, according to Alexa, and constitutes one of the largest and oldest cybercrime forums on the internet. Large forums provide a critical alternative marketplace. If CaaS offerings are transacted via forums, they should be visible on HF. For this reason, HF provides a representative dataset for studying the presence of CaaS trade in forums.

Research on cybercrime forums has focused primarily on four topics: (i) social networking and organization of forum members [20, 29, 31, 37, 41, 43, 45, 55], (ii) identifying and profiling key actors [1, 5, 8, 27, 42, 43, 50], (iii) evolution over time [3, 21, 44, 59] and (iv) measurement of products, transactions and prices [6, 7, 15, 32, 43,

49, 51, 52]. Our work contributes to the third and fourth categories by presenting the first comprehensive and longitudinal analysis of 28 different types of CaaS supply and demand on a cybercrime forum. We develop a classifier and measure the relative prevalence of different CaaS types over time. We also track whether the offering references a private communication channel, as prior work found that after a match between supply and demand is reached, the actual transactions are often conducted via private channels like Telegram, Discord, Skype, Jabber, or IRC [10, 19]. We conduct our analysis on 11 years of data from Hack Forums, one of the largest and oldest ongoing English-language cybercrime forum on the surface web. In sum, we make these following contributions:

- We present the first comprehensive study on the prevalence of a large collection of Cybercrime-as-a-Service offerings on a cybercrime forum over time. We compare the prevalence to four product-type cybercrime offerings.
- We generate a labeled set of ground-truth data of 662 posts across different types of CaaS and 4 product-type offerings, with attributes to indicate whether the post is supply or demand, includes a private contact, or offers an external trading link. We make our labeled dataset available to other researchers upon request in order to facilitate research on CaaS on other forums and channels.
- We evaluate various machine learning models and get the best results with LinearSVC for the CaaS and product classification (0.76 F1-score and 0.95 AUC ROC), Logistic Regression for distinguishing supply and demand offerings (0.86 F1-score and 0.92 AUC ROC) and XGBoost for discovering the usage of private messages and other communication channels (0.98 F1-score and 0.94 AUC ROC). We supplement these models with decision function rule sets for the supply/demand tags and contact information.
- Of 28 known CaaS types, we only found evidence for only 9 of these in the forum. Around 16% of first posts in the `Market' section of our dataset involve CaaS where `bot/botnet as a service', `reputation escalation as a service' and `traffic as a service' categories make up the majority (over 60%) for the whole period in terms of both supply and demand. At least half of each CaaS offerings directs potential buyers to an instant messaging app or private message for transacting privately. We found no major changes in supply and demand in the period after the major law enforcement actions against the underground marketplaces, suggesting CaaS offerings on the latter did not migrate to clearweb cybercrime forums.

The rest of this paper is structured as follows. Section 2 describes the CaaS types and value chain model. Section 3 lays down our approach and explains our methodology of collecting, preparing and classifying the data. Section 4 presents our findings and Section 5 discusess these findings in terms of their implications for law enforcement. Section 6 discusses the limitations and ethics. Section 7 connects our work to earlier contributions and Section 8 concludes with future work recommendations.

2 CAAS TYPES AND VALUE CHAIN MODEL

It is important to understand the activities associated with a cyber attack to disrupt the business of cybercriminals who sell these attacks. These activities are, however, widely dispersed and inconsistently identified in the literature. Without a clear framework, it is also impossible to effectively grasp current cyber threats.

Michael Porter's value chain model [48] considers an organization as a system composed of subsystems, each with inputs, transformation processes, and outputs, in addition to support activities. As cybercrime has become a business, we will recognize activities that add value to cyber attack operations from the viewpoint of the value chain. This value-added processes include any activity in the business ecosystem of cybercrime that lets the attacker minimize the cost of cyber attacks and maximize the benefits. The support activities besides the primary activities, that are frequently ignored, are also important in promoting the operation of the cybercrime business, as they can enable the attacker to carry out an attack at a reduced cost and with greater profit. To perceive these processes, we make use of the cybercriminal value chain model (Figure 1) developed by K. Huang et al. [21] consisting of the primary activities of vulnerability discovery, exploitation development, exploitation delivery, and attack, as well as the supporting roles of cyber attack life-cycle operations, human resources, marketing and delivery, and technical support. They have validated the list of cybercrime services and this value chain framework with more than 30 senior executives, managers, and researchers focusing on cybersecurity from Fortune 500 companies and key cybersecurity solution providers.

Figure 1: Cybercriminal Value Chain Model

We can see each of these service's availability, pricing model and their estimated price (changes over time) in the upper part of Table 1, taken from the same study [21]. As shown in Figure 1,

these CaaS categories are mapped to the value chain model and then grouped into three di erent categories based on their availability: (i) existing, (ii) evolving and (iii) emerging. Although this extensive list of CaaS categories helped us to establish the structure of our study, we added more CaaS types because it did not cover all CaaS models we encountered in the literature. CAPTCHA solvers [36], Phone/SMS veri cation [55, 56], E-whoring [24, 43], Proxies [55] and Remote Desktop Protocol (RDP) [47] services are the later added CaaS categories whose properties are depicted in the lower part of Table 1. As a result, we had a de nitive set of 28 CaaS types that we wanted to identify in the posts of our forum dataset.

3 MEASUREMENT METHODOLOGY

Our measurement methodology consists of: (i) compiling and pre-processing a dataset of underground forums, (ii) creating and annotating `ground truth' listings manually, (iii) deploying a classi er to map the cybercrime products/services, supply/demand, contact and external links; and (iv) analyzing the dynamics of CaaS in the forums.

3.1 Data collection

We leverage the data collected by Cambridge Cybercrime Centre[1] on underground forums, namely CrimeBB. CrimeBB comprises posts from di erent forums as listed in Table 2. Hack Forums is one of the largest and oldest ongoing hacking communities having a separate marketing section. Accessible from the surface web, it has been the most popular hacking forum according to Alexa ranking for long years. HF has been connected to several high pro le events, including the release of the Mirai botnet source code [44]. Some of the contents of these surface web forums, unlike other types of underground markets, are legitimate, such as discussions pertaining to current a airs, gaming, and topics related to technology. These forums, however, are also used for the sharing of information on aberrant behaviour and trade in products and services of unlawful origin or use [42]. Previous study has found that these sites will o er a jump ahead to more serious online illegal activities [23]. We also think that selection of a public forum for evaluations will help to enable reproducibility of our ndings.

HF[2] has 10 di erent sections including *Market*, *Hacking*, *Life* , *Tech*, *Coding*, *Game*, *Groups*, *Web*, *GFX* and *Money*. As stated in the forum's code of conduct, all trading should take place in the *Market* section. Considering that our research is about investigating the trade-related posts, we naturally limited our scope only with this section. Since the sub-forums in the HF dataset [44] were not classi ed into these 10 sections, we classi ed them ourselves. Thus, we followed the links of the scrapes of each sub-forums in the dataset after registering in the HF website as a regular member and classi ed the forums into the di erent sections manually.

3.2 Data preparation

There are 29 sub-forums, 1,107,372 threads and 9,795,204 posts ranging from 17/11/2007 to 04/12/2018 in the *Market* section of Hackforums. We are interested in the rst-posts of these threads since they are more useful to provide information about the subject product and service types [7, 8, 40, 42]. Hence, we extracted rst-posts of each thread in this section making total of 1,104,046 posts.

Before starting the annotation of our ground truth we set the rules to di erentiate a CaaS type crimeware from a classic product type crimeware. We name an o ering as CaaS type if the owner of the post is either:

- renting an infrastructure or/and platform (e.g. bulletproof hosting, proxies),
- selling a service for committing a crime (e.g. hacker for hire, money laundering),
- selling a product but continuing to provide some required services remotely after sale (e.g. dashboard service for Ransomware as a Service, control board of a exploit kit),
- selling a product but giving lifetime customer support (i.e., giving support for only installing/setting up is not considered as CaaS).

3.3 Ground Truth

We created ground truth data by annotating each item of a sample population of posts to 5 groups of labels (Figure 2): (i) category of the service or product, (ii) supply/demand/exchange/other classi - cation of the o er, (iii) existence of private message, (iv) existence of contact information and (v) existence of external trading links.

Figure 2: Ground Truth labels

We randomly selected 12,000 posts (approximately 1% of 1,104,046) from all rst-posts of the *Market* section. In a rst phase, we annotated 500 posts. Out of these 500, only 19% of the posts belonged to the CaaS categories. To avoid over tting due to highly represented categories, the ground truth data needed to include at least 15 posts per category which is observed as the breaking point number in our labeling. For this purpose we boosted our ground truth by annotating more posts of the CaaS and product-type categories whose numbers were below 15. To nd these lacking categories we conducted a manual keyword search on the sample of 12,000 posts which led to 662 posts in total where 29% of the posts then belonged to the 14 di erent CaaS categories. Around 29% belonged to 'account' category, 21% belonged to the 'other' category, 9% belonged to 'cash-out' category and 12% was labeled as product-type crimewares.

Later we excluded ve categories (17 posts) from the ground truth due to their scarcity (less than 5 posts): e-whoring as a service, tra c redirection as a service, tool pool as a service, CAPTCHA solving as a service and money laundering as a service. It is also worth remarking that some of the CaaS types (14 out of 28) did not occur even once in our enlarged ground truth containing 662 posts. These zero-existing CaaS categories are from the 'existing' group (Exploit as a Service, Deception as a Service, Security Checker as a Service, Marketplace as a Service, Money Mule Recruiting as a Service, Hacker Training as a Service); the 'evolving' group (Personal

[1] https://www.cambridgecybercrime.uk/datasets.html
[2] Hack Forums. https://hackforums.net

Table 1: CaaS categories and their properties

	CaaS name	Status	Pricing model	Estimated price
EaaS	Exploit as a Service	Existing	Licence Subscription	up to more than $250,000 $150,000 per month
PLaaS	Payload as a Service	Existing	Pay-per-install Commission	$0.02-$0.1 per install 40%
DaaS	Deception as a Service	Existing	Subscription Commission	$85-$115 per month 40%
OBaaS	Obfuscation as a Service	Existing	Subscription	$50-$150 per month
SCaaS	Security Checker as a Service	Existing	Subscription	$25 per month
TRaaS	Tra c Redirection as a Service	Existing	Pay-per-click	$7-$15 per 1000 visitors
BNaaS	Botnet as a Service	Existing	Subscription	$40 per month
BHaaS	Bulletproof Hosting as a Service	Existing	Subscription	$300 per month
TAaaS	Tra c (including DDoS) as a Service	Existing	Subscription	$999 per month
REaaS	Reputation Escalation as a Service	Existing	Pay-per-record	$0.42-$0.7 per record
MPaaS	Marketplace as a Service	Existing	Licence Commission	$4500 per licence 2%-10%
MRaaS	Money Mule Recruiting as a Service	Existing	Licence	$1700 per licence
MLaaS	Money Laundering as a Service	Existing	Commission	2%-30%
HTaaS	Hacker Training as a Service	Existing	Licence	$200-$800 per person
PPaaS	Personal Pro le as a Service	Evolving	Licence	$4-$20 per record
TPaaS	Tool Pool as a Service	Evolving	Subscription	$4000 per month
RaaS	Reputation as a Service	Evolving	Subscription	-
HRaaS	Hacker Recruiting as a Service	Evolving	Subscription	-
VDaaS	Vulnerability Discovery as a Service	Emerging	Subscription	$542.04-$1810.31 per vulnerability
TSaaS	Target Selection as a Service	Emerging	Subscription	-
EPaaS RPaaS	Exploit Package as a Service Repackage as a Service	Emerging	Subscription	$4000 per month
DMaaS	Domain Knowledge as a Service	Emerging	Subscription	-
VEaaS	Value Evaluation as a Service	Emerging	Subscription	-
CPaaS	CAPTCHA solving as a Service	Existing	Pay-per-solution	$0.5-$20 per 1000 CAPTCHAs solved
PSVaaS	Phone/SMS Veri cation as a Service	Existing	Pay-per-challenge	$0.20 per challenge
RPSaaS	RDP/Proxy/Seedbox as a Service	Existing	Licence Subscription	$8-$15 per server $25-$250 per month
EWaaS	E-Whoring as a Service	Emerging	Subscription	-

Table 2: CrimeBB dataset

Forum	Language	Members	Threads	Posts	Oldest
Hackforums	EN	573,925	3,856,143	40,196,641	01/2007
Kernelmode	EN	1,441	3,144	25,024	03/2010
O ensive Community	EN	10,593	18,436	58,779	06/2012
Multiplayer Game Hacking	EN	452,186	739,527	8,907,938	12/2005
Stresserforums	EN	764	708	7,069	04/2017
Greysee	EN	440	1,239	6,969	06/2015
Garage4Hackers	EN	872	2,096	7,697	07/2010
SafeSkyHacks	EN	7,378	12,892	26,842	03/2013
Antichat	RU	77,865	242,408	2,449,221	05/2002
RaidForums	EN	43,278	33,100	124,776	03/2015

Pro le as a Service, Reputation as a Service) and the 'emerging' group (Vulnerability Discovery as a Service, Target Selection as a Service, Exploit Package as a Service, Repackage as a Service, Domain Knowledge as a Service, Value Evaluation as a Service) of the ones in Table 1. In conclusion, we found zero occurrences of 14 and inadequate occurrences of 5 of the CaaS categories while building the ground truth set before classi cation, so we reported 645 posts on the remaining 9 CaaS types and 4 product-types besides 'account' and 'cash-out' categories.

Note that 'account' category is a combination of the accounts from games and social media, email and websites; while 'currency exchange' type posts are sitting under 'cash-out' category where the vendors ask for or propose gift cards, paypal or digital currency to be monetized. We put 'DDoS as a service' type o erings under the 'Tra c as a service' category to make it compatible with categorization of [22] as listed in Table 1. Lastly, the category 'other' contains the rest of the posts talking about cyber crime related subjects like 'e-books, guides, coding, scamming reports, chat invitations, webpage designing, signature spaces, domain names' and irrelevant subjects like 'fast-food coupons, wireless dog fence collar, mobile phones, movie tickets, shoes, laser pointer, computer parts'.

We sometimes had di culties to decide the labels since some posts are containing more than one category like NightOwl Botnet

Service| Bulletproof Hosting|Lifetime Support|Setup| Crypt| #1 On HF o ering both bulletproof hosting and a botnet. We annotated them regarding the contextual weight of the announced categories in that post or their inclusiveness status of one another. If several categories are stated evenly in the post, we chose the rst one according to the sentence order. Another interesting case to be expressed is the posts like:'Looking to Buy U2bviews account With 100K+ Credits.' This post is labeled as 'account' even though the stated account most probably would be used for reputation escalation. However we were right not to label it as 'Reputation Escalation as a service' since the demand is not for the service itself but for the account having that purpose.

While creating the supply/demand/exchange/other label, we labeled the posts including 'free' o ers (e.g.[Free to everyone!] FUD Java Drive By [FREE]) as 'supply' since they are also o ering services at no cost. During the contact info annotations it was frequently confusing to extract the contact information from the posts like:'Was wondering how much someone would o er for this email. Not sure OG but whatever. disclosed@live.com' it is indeed an o ered email account but not a real contact of the post owner.

For the external trading links, we especially looked in the supply type posts since it is not logical to encounter these links in demanding posts. Sometimes, external links could be only for showing the photos or properties of the real/similar product like a spoiler but not for shopping. After all, we could only nd four posts including any 'external trading link' in their contents. Most probably, the service/product sellers don't need to give external links to conduct their business. Otherwise this outcome signi es that either they share the trading links to the serious buyers in private contact channels or in the following posts of the same thread which we don't inspect.

3.4 Classi cation by Machine Learning and Decision Functions

The classi cation phase itself consists of four steps: (i) data cleaning, (ii) tokenizing, (iii) model selection and (iv) training and evaluation using the ground truth.

In the data cleaning phase, we removed all English stop-words, punctuation, numbers, URLs, accents of all unicode characters and non-textual content such as image, video frames, code, citing or attachment which were annotated earlier with delimiters by the dataset providers. We then lemmatized the words in order to group together the in ected forms of a word so they can be analyzed as a single item. We used a chi-squared feature extraction giving the most-frequent words used in discriminating the categories to take out the non-sense words and adding them to the stop-words set to be excluded in the analysis.

Next, we tokenized each listing which is the concatenation of heading and content of the rst-post of each thread and computed a tf-idf (term frequency - inverse document frequency) value for each of the resulting 4,587 unique tokens. Term Frequency (tf) indicates the number of occurrences of a particular term in document while Inverse Document Frequency (idf) of a term is the number of documents in the corpus divided by the document frequency of a term (the number of documents containing the term). To calculate the tf-idf, we used a max-df (maximum document frequency) equal to 0.7 this discards words appearing in more than 70% of the listings, and replaced tf with 1 + log(tf) by applying sublinear tf-scaling. We made experiments with setting the analyzer of tf-idf-vectorizer to 'char' instead of 'word' and setting the ngram-range more than one word; but observed no improvements in the classi cation results inducing us to relinquish these changes. We also extracted meta-data features of polarity and subjectivity using sentiment function of TextBlob library in Python from text of the posts and combined them with the textual features in a `pipeline' to improve the classi cation. Polarity is coded as a oat number which lies in the range of [-1,1] where 1 means positive statement and -1 means a negative statement. Subjectivity is also coded as a oat number which lies in the range of [0,1] where subjective sentences refer to personal opinion, emotion or judgment contrary to the objective ones referring to factual information. However we ended up utilizing only the textual features in our models due to ine ciency of these meta-data features in distinguishing our categories.

In the third step of our classi cation process, we selected the model with best evaluation results under a 10-fold strati ed (preserves the imbalanced class distribution in each fold) cross-validation with default parameters. From the depicted preliminary evaluation results in Figure 3, we can see that LinearSVC model has the highest F1-score (0,76) among other eight ML models (Stochastic Gradient Descent, Random Forest, Decision Tree, AdaBoost, Multinomial Naive Bayes, K-Nearest Neighbor, Logistic Regression and XGBoost). We also assessed the performance of these models by resampling with SMOTE (Synthetic Minority Over-sampling Technique), thereby increasing the cardinality of each category in the training phase to match the size of the largest category for the purpose of mitigating the negative impact of the imbalance between the categories. LinearSVC kept its rst place with a higher F1-score (0,77) although the order of the some other models have changed (Figure 12 in Appendix), hence we decided to pursue the LinearSVC model in our analysis for the crimeware classi cation.

We also implemented 'decision functions' together with the ML algorithm to reinforce the classi cation as conducted in [8, 42, 62]. Then the model consists of two parts: (i) regular expression (RegEx) calls used to search for certain keywords in the posts to classify them and (ii) an ML algorithm to classify the not-classi ed posts from the rst part. For instance, a rule-set consists of the union of 'anyone, someone' and 'hack, jack, decrypt, reverse engineer' keywords for the 'hacker as a service' category; and another rule-set 'bots, installs, updates, bot shop, slaves, DoS' keywords for the `bot/botnet as a service' category. However, we couldn't get better evaluation values after adding decision functions while classifying the service and product categories. Note that while the decision functions didn't improve the classi cation of the cybercrime categories, it did improve the accuracy of the classi cation of the other tags related to the supply/demand and contact information.

We performed hyperparameter tuning for our selected model LinearSVC applying gridsearch algorithm with 10-fold cross validation but ended up implementing LinearSVC (svm package of scikit-learn 0.24.2 library) with default parameters in `balanced' mode for class_weight parameter which automatically adjusts weights inversely proportional to class frequencies in the input data, since the latter settings had presented a relatively better performance.

Figure 3: Preliminary model evaluation

All correct predictions (recall) are depicted in the diagonal of our normalized confusion matrix (Figure 4) where each row represents the instances in an actual category and each column represents the instances in a predicted category. We can easily see that especially three of the categories (hacker as a service, malware and reputation escalation as a service) are more confused with the category 'other' resulting in low recall values. We elaborate on this misclassi ca-tion issue with illustrating examples in the following subsection. The precision, recall, f1-score values and occurrences of the each class items (imbalance) under name -support- can be viewed on the classi cation report displayed in Figure 13 in the Appendix.

Figure 4: Normalized confusion matrix heatmap

Figure 5 shows the receiver operating characteristics curve (AUC-ROC) of the model. The AUC-ROC is one of the most important evaluation metrics for checking any classi cation model's perfor-mance as it shows how good is the model in distinguishing classes. Our model can di erentiate the categories quite well with AUC-ROC values over 0,95 for every cross-validation splits.

Supply and demand o erings. To classify the CaaS type posts into either supply or demand o erings, we implemented four di erent settings: (i) ML-only, (ii) ML with resampling, (iii) ML with decision

Figure 5: ROC curve of cross-validations

functions, and (iv) ML both with resampling and decision func-tions. We used the same parameters from the CaaS classi cation for the ML models, but only with di erent keywords in the decision function rule-sets; i.e.,'bid, refund, stock, shipping, buyers only, i can provide, i can give, my price, i will sell, serious o ers, payment is via' for supply;'WTB, seek, ready to pay, where can, is there, I'll pay, I want to buy, looking to buy, if you can do, willing to buy, is anyone, can anyone, grateful, need assistance, if anyone has, paying well, any suggestions' for demand. We observed in our ground truth that CaaS type posts contain only 'supply' or 'demand' o erings but not any 'exchange' or 'other' ones. Thus we preferred includ-ing only 'supply' or 'demand' labeled posts of the ground truth in this training phase since our goal is to classify only the CaaS type posts into their supply/demand classes. LinearSVC gave the best results with the rst three settings with 0.832, 0.838 and 0.849 F1-scores respectively, while Logistic Regression outperformed all the other algorithms with the last setting with 0.86 F1-score (and 0.92 AUC ROC). Consequently, we applied Logistic Regression with resampling and decision function to all of the ground truth data.

Contact information. We conducted 'contact information' classi-cation in two stages (and with four settings of ML-resampling-decision function combinations described earlier): 1) whether the post owner asks for 'private message (pm)' and 2) whether there is any contact information (pm and others) in the post or not. We used all of the ground truth items in the training of contact information classi ers, since they were incurred to binary labeling respecting the existence of `private message' and `other contact applications'. It is notable that we could not expand this classi cation deeply to other types of communication apps like 'telegram, discord, skype, jabber, whatsapp, wechat, facetime, viber, etc.' by reason of possible confusion by our ML models since these apps are at the same time o ered as products or services in the posts of 'phone/sms veri -cation as a service, hacker as a service, reputation escalation as a service and tra c as a service' categories.

In the rst stage, we leveraged the keywords 'pm, private message, direct message' in the regex rule-set to nd the posts having 'private

message' for contact. It is expected to make an exact classi cation by this decision function, nevertheless there are still cases that cause confusion. These are (i) the forum members sometimes prefer to request private messaging by only using the word 'message' which we can not add in our rule-set due to a confusion probability with other contact types like 'skype message', (ii) the rule-set keywords are written in some posts as part of a quoted story but not to mean a contact request (iii) the rule-set keywords are used for the intention of communication but in negative manner, telling that they don't want to contact via private messaging, (iv) 'pm' abbreviation is used to express 'perfect money'. Thus we again asked for the help of ML models in addition to decision function to enhance the results. After trials of several combinations stated earlier under 10-fold cross-validation, ML with decision function setting implementing XGBoost algorithm having 0.982 F1-score (and 0.96 AUC ROC) is chosen as the winner whilst all other models except LinearSVC (0.967) had almost same F1-scores in this setting. 0.953, 0.961, 0.969 are the F1-scores of decision function only model, ML (XGBoost) only model and ML (XGBoost) with resampling model, respectively.

In the second stage (searching for not only private messages but all kinds of contact information), we also executed the regex keyword search of the rst phase as a starting point for the ML model. XGBoost with decision function model is again winner of the 10-fold cross-validations with 0.981 F1-score (and 0.94 AUC ROC) whilst ML only model and ML with resampling model implementing XGBoost show performance of 0.933 and 0.939 F1-scores respectively.

External trading links. We couldn't run any model to search and classify the 'external trading links' stated in the posts since there was not enough labeled ground truth data (only four labelings as explained in the previous subsection) to run a supervised ML model.

3.4.1 Misclassifications. It is not surprising that even the best performing ML algorithm misclassi es some of the posts in a forum dataset, so did ours. To assess the impact on the nal results, we brie y exemplify the misclassi ed posts while classifying them into crimeware types. We can give some examples:

'e-whoring pack' misclassi ed as 'other': 'Im still working on the e-book and give it out to the customers when its nished.' is a sentence from the related post. Misclassi cation is most probably due to the word 'e-book' which is mostly used in the posts of 'other' category.
'hacker as a service' misclassi ed as 'account': Expressions of 'instagram account' in 'Need someone to Jack an Instagram. Hey i need someone to get me a password to an instagram account;' 'youtube account' in 'I was wondering if it's possible to pay someone out there to get a YouTube account banned.' 'yahoo email' in 'Willing to pay for yahoo email hack. I need someone to help me by getting an email pw for, i.e.' 'Combat Arms Game' in '[Coders] Combat Arms Game Hacking. I am searching for coders who are able to code hacks for combat arms.' possibly causes misclassi cation.
'malware/hacker tool' misclassi ed as 'other': Misclassi cation of the post including the sentence 'Looking for a 7-Layer Attack Script' is most probably due to the word 'script' which is mostly used in the posts of 'other' category.
'obfuscation as a service' misclassi ed as 'bot/botnet as a service': The word 'DDoSeR 4.2 bot' in this sentence 'Can anyone crypt my DDoSeR 4.2 bot without corrupting it?' the post is seemingly the reason of this misclassi cation.
'reputation escalation as a service' misclassi ed as 'other': Using the keywords 'ad y' and 'twitter' in 'I made an little and easy website for ad y and twitter clicking', 'social media' in '[NEW] Sterineb's Social Media Booster Service [NEW]' 'youtube' in 'I'm currently looking for someone who can deliver a lot (and consistent) YouTube views every week' presumably results with misclassi cation.

4 RESULTS

CaaS and product-type. We observe a substantial amount of CaaS type crimewares (15.6%) in our whole dataset as depicted in Figure 6 which includes all of the content types (supply, demand, exchange and other). However, `account' and `other' categories are dominant accounting for around 70% of the total number of posts (39.7% and 29% respectively). Posts related to 'cash-out' (12.6%) and 'product-type - composed of crypter, e-whoring pack, exploit, malware/hacker tool (3.1%) were lower in number than CaaS o erings. We can see sample posts from these nal categories in Figure 7 and read their real counts and percentages in Figure 14 of Appendix. To understand this dominance of 'account' and 'other' categories, it will be helpful to revisit Section 3.3 where the products comprised by these two categories are also described.

Figure 6: CaaS categories vs. the rest

When we delve into the diversity of only CaaS o erings in the market in Figure 8, it is seen that we could only nd 9 CaaS types in the forum. In these existing CaaS types, we can easily distinguish 'bot/botnet as a service (BNaaS)', 'reputation escalation as a service (REaaS)' and 'tra c as a service (TAaaS)' categories with 22%, 24%

Figure 7: Sample posts from the nal categories

and 18.5% of the total supply CaaS posts and with 26%, 19% and 16% of the total demand CaaS posts respectively. This result is not surprising given the multi-facet use of botnets for spreading malware, creating DDoS attacks, and fraudulently boosting social media accounts or web-shopping page rankings. The importance of social media and prevalence of trading on the internet justi es why there is this much supply/demand to reputation escalation services. And of course well-known usage of DDoS services (booters/stressers), spam email/SMS bombs and ad frauds explain the high volume of the 'tra c as a service' type o erings.

It is always informative to look into time series analysis of longitudinal data to better explore the changing dynamics over time. At rst glance, in the upper plot of Figure 9 where again all of the content types (supply, demand, exchange and other) are represented together, we see that number of posts in the 'CaaS categories' and 'all of the categories' have both a decreasing trend after 2012 which is also coherent with the decreasing numbers of all of the posts in whole Hackforums platform as illustrated in [4]. It seems that the period between 2011-2012 was the golden age for Hackforums community with the peak number of trading posts being over 20k per month. It is also remarkable that CaaS o erings keep their ratio

Figure 8: Diversity amongst CaaS categories

in the whole set almost steady during the full time period. To put it di erently, the prevalence of CaaS did not increase over time, in sharp contrast to the more popular understanding that cybercrime was increasingly dominated by service models. This understanding has even shaped law enforcement agendas across the world. In the

EU, for example, disruption of CaaS was mentioned as a top priority in all crime, not just cybercrime [16].

The various CaaS categories that were identi ed on the forum are depicted in the lower plot of Figure 9 so that we can probe into their popularity throughout the years. It is interesting to see that 'bot/botnet as a service' having the highest number of occurrences in total as shown in Figure 8 loses its popularity against 'reputation escalation as a service' after 2013 and also against 'hacker as a service' after 2015 over time. Furthermore, 'tra c as a service' and 'obfuscation as a service' after 2014, 'bulletproof hosting as a service' after 2015 also present a declining occurrence in the platform.

Figure 9: Time series plots

Supply and demand o erings. Supply/demand classi cation results for CaaS categories per se can be inspected in the Figure 10. We can see that supply and demand posts are in balance for most of the categories except demand posts looking for 'hacker as a service (HRaaS)' and 'phone/sms veri cation as a service (PSVaaS)', and supply o erings for 'bulletproof hosting as a service (BHaaS)' being higher (more than 70%) than the rest. The results indicate that the CaaS market in the forum is not one-sided, so that both buyers and sellers foster each other to keep the trade alive. It is also remarkable to see the high demand for hacker and phone/sms veri cation services for future market predictions.

Contact information. In Figure 11 we can observe both the distribution of 'only private message' and 'all types of contact info' amongst the CaaS supplies. The implied contact info aside 'pm' is the communication apps like 'telegram, discord, skype, jabber, whatsapp, wechat, facetime, viber, etc.'. Post authors of at least half of the each CaaS categories requested using an instant messaging app or private messages to continue the trade privately. It is also worth noting that these authors rely on private messages within the Hackforums platform for their private communication.

Figure 10: Supply/demand ratio in CaaS categories

Figure 11: PM/Contact info in supply posts of CaaS categories

5 DISCUSSION

We aimed to understand how cybercriminals are trading in new services via crime forums, as part of the emergence of CaaS. We observe some amount of CaaS o erings in the forum market across diverse categories. Out of 28 known CaaS models, only 9 were found on the HF forum and just 3 dominate supply and demand: 'bot/botnet as a service', 'reputation escalation as a service' and 'tra c as a service'. The ratio of the service o erings to other crimeware o erings remained stable over the last years. Our evidence does not support the dominant idea that CaaS is rising and the next evolution of cybercrime. Of course, this might be happening on other forums. Future research will have to bear this out, but we should note that Hack Forums constitutes one of the largest and oldest crime forums on the internet. If it the evolution is not visible here, it might not be such a widespread phenomenon as is often assumed. Yes, there might be forums that cater more speci cally to this niche, but we have to keep in mind that one of the promises of CaaS is that it lowers entry barriers for new criminals and facilitates their criminal activities. For this scaling e ect to occur, one would need new entrants to nd the places where one can purchase these services. That rules out more secretive and trust-based forums. Of

course, we should not over-generalize our analysis of a single, if large, forum. It does raise the question of where the feared growth in the matching of supply and demand for CaaS is happening? If not in forums, then in other channels? Two other channels are underground marketplaces and custom sites. Earlier on marketplaces, however, R.van Wegberg et al.[58] found no evidence for this evolution either. For custom sites, we know that there are successful examples e.g., certain booters but also that this channel is rife with fraud and generally seemed to provide less sophisticated o erings. In sum, our nding show that law enforcement might need to re-evaluate the priority that is currently being given to CaaS. Perhaps this trend is mostly relevant for speci c threats, rather across the whole cybercrime landscape, notwithstanding anecdotal evidence for 28 CaaS models.

Another nding of our study is to demonstrate that more than half of the cybercrime trade is dealt with privately via messaging apps and private messages on the forums. Based on these insights, LEA (Law Enforcement Agencies) and messaging app application vendors could leverage scraping of these contact points and take them down, raising the transaction costs for the miscreants abusing these platforms.

Next, the dominance of 'reputation escalation as a service' and 'phone/sms veri cation as a service' o erings suggests that service providers, website developers and application vendors need to design systems that can withstand these type of compromises.

Last but not least, models like those we developed for our analysis show that data on CaaS o erings can be gathered from forums for threat intelligence. For this reason, we also make our ground truth dataset available upon request to researchers in industry, government and academica. Data collection on a wider site of forums would not only bene t LEAs, but also threat intelligence vendors which provide customers with insights into the changing threat landscape.

6 LIMITATIONS AND ETHICS

Sometimes contact info is written not in the body but in the footer of the post. These footer contents are not scraped in our dataset. We could manually look into these footer contents and annotate the posts in the ground truth, but ML classi er would not be able to do the same. Hence, we neglected the contents of the footers and annotated the posts merely with the help of their body contents. There are also some posts whose contact info are embedded in an image in the content which our classi er could not mine through. We also see some contact info in the content for the purpose of storytelling but not for real-intention like this post including 'skype': 'This is an auction for 308k FIFA Ultimate team coins The reason im selling these is because i dont have FIFA anymore The auction will end when people stop posting on the thread sorry for the ugly bastard on skype he has a nose like ibra of the thread'. "My Personal Clash Of Clans Almost Maxed TH8 Account". Therefore, we did our best under these limitations to analyze the communication ways the forum members use while trading cyber crimewares.

We only analyzed the thread headings and rst posts of each thread in 'Market' section for a good reason, thus we may have missed the 'contact info' and 'external trading link' clues expressed in the following posts of those threads. Moreover our results are based on the observation of a single large forum. Thus, we do not analyse the posts on other forums. We also used a limited size of ground truth due to the long duration of labeling the posts into more than 30 cybercrime categories (at the beginning). Another limitation for our ground truth labeling was using only one annotator so lack of an inter-annotator agreement.

The research approach was designed with ethical considerations at the center. We complied with the Cambridge Cybercrime Centre's data sharing agreements. Furthermore, we note that all of the data used in our study was gathered from publicly available sources as anyone could register and access these digital forums on the surface web. Given the anonymous nature of these platforms, it is unlikely anyone used their real name so the data did not contain any type of Personal Identi able Information (PII).

7 RELATED WORK

7.1 CaaS business

We founded our research on the conceptual framework of K.Huang et al.[22] where they conducted an extensive survey of the CaaS services utilizing the value chain perspective in a systematic way which we talked more in Section 2. The authors described 24 di erent existing, evolving and emerging CaaS types of primary and support activities feeding the cybercriminal service ecosystem. They also demonstrated how the popular cyber crimewares like ransomware, social engineering attacks, fake reputation generation and privacy explosion follow a path of the framework presented in their work. To analyze the pro tability of cybercriminal businesses, the return on investment (ROI) of ransomware attack business is exempli ed in the Supplementary Material document for their survey. Furthermore, they didn't forget to suggest several strategy implementations for combating cybercrime.

Some of the CaaS crimewares and attacks like DDoS[12, 15, 26, 28, 39], reputation escalation[55, 61], exploit[18], obfuscation, value evaluation[53], personal pro le (impersonation)[9], deception[57], CAPTCHA solvers[36], phone/SMS veri cation[56], e-whoring [24, 43], tra c[13, 14], proxy[55], RDP[47], vulnerability discovery[50], password cracking (enclosed in 'hacker as a service' category in our study)[33], ransomware[2, 11, 34, 46, 54], bullet-proof hosting[38] as a service have been studied academically or analyzed technically in previous papers and security blogs.

7.2 Cybercrime analysis on Underground Forums/Markets

Various researchers have applied di erent information retrieval and text mining methods on online forums. Even though the methods resembles the ones we leveraged in this study, our study is the rst to classify the posts into large number of novel CaaS categories. The closest related work to ours is by R.van Wegberg et al.[58], which tracks the evolution of commoditization of cybercrime on online underground marketplaces and identi es the market supply over time. Our work di ers from this previous work in that we especially focus on CaaS o erings in the forums and not only the supply but also demand and contact aspects of these o erings over time.

In [44] S.Pastrana et al. describe CrimeBot, an online forum crawler, which is used to update and maintain the CrimeBB dataset

to be used for large-scale and longitudinal analysis. They provided a case study analysing currency exchanges in Hackforums community which shows their evolution over years and tracks the previous activity of the key actors by using SVM classi ers. The characteristics and pathways of 'key actors' in Hackforums of CrimeBB dataset who have been linked to criminal activities are analyzed in [42] by applying social network analysis (SNA), k-means clustering and natural language processing for the sake of deterring the involvement of young people away from a career in cybercrime. While analysing their pathways by looking for changes in their interests and the number of questions posted as they spend more time in the forum, they leveraged a binary classi er to identify the questions. A hybrid approach of classi cation combining statistical technique (Linear SVM) and heuristics (rule set) were implemented and accurate enough evaluation results (Precision=0.88, Recall=0.85 and F1=0.86) were obtained.

A. Caines and his colleagues [8] also conducted a classi cation study over the posts for three properties: post type, author intent and addressee from the HackForums of CrimeBB dataset. They found that a hybrid rule-based (logical) ML (statistical) model performs best for post type and author intent, whereas a purely statistical model (SVM, XGBoost and Linear model) is best for addressee. R. Portno et al. identi es the posts from eight underground forums of several languages related to transactions as determining the nature of the post whether it is an o er to buy, sell, currency exchange or none by the help of SVM model in their study [49]. Hackforums is one of the forums they scraped (only trade-related subforums) and analyzed for not only identifying the product types but also extracting product names and prices. One of the conclusions of their analysis is that the additional features like the length of the post or the reputation of the author did not improve the classi cation accuracy, that we also observed in our results. After assessing their classi er both within and across forums and getting promising results, they performed two case studies to uncover the popularity of original vs. bulk/hacked accounts, and detect the high-demanded currencies.

A random forest method is used in [41] to predict which public posts are likely to trigger private messages by automatically labeling them. They evaluated the performance of their methods using data from three real forum leaks in di erent languages and had AUC results ranging from 0.65-0.77 when training and testing is performed on the same forum. R. Bhalerao et al. [7] propose a method that leverages machine learning and graph-based analysis to extract supply chains from cybercrime forums. Similar to our approach they classify the posts from two popular forums one of which is the HackForums to 14 product categories (not focused on as a service types) and get averagely 0.71 F1 from the Hackforums evaluations using strati ed k-fold cross-validation. They later identify the replies indicating that a user bought or sold the product with the validation result of 0.85 F1 averagely for the Hackforums. This step somewhat di ers from our supply/demand classi cation since we implement our method over rst posts of each thread but they implement over the replies to these rst posts. Their using character n-grams rather than word n-grams and using a weighted average of the precision scores across all the categories except 'other' to select the classi er rather than F1 score of all are other dissimilarities with our work.

The authors of RIPEx paper [17] identify and label IP addresses in security forums by utilizing a cross-forum learning method that they use a classi er from their source forums to identify seed information for training a classi er on the target forum. Thus it does not require training data for each new forum and achieves better performance than solely using the classi er of the current forum on the new forum. They found out that Logistic Regression method outperforms SVM, Bayesian networks and k-nearest neighbors methods using 10-fold cross-validation on their ground truth including both text information of the post and the contextual information like frequency of posting, average post length which captures the behaviour of the author. In another work [40] obtaining threat intelligence information from online forums and markets, a combination of semi-supervised and supervised methods including Naive Bayes, random forest, support vector machine and logistic regression are used to classify data. Logistic regression performed the best with 80% precision and 68% recall while evaluating on a small dataset of two English forums, and leveraging unlabeled data in a semi-supervised technique improved the recall about 10% with same precision.

While SVM outperformed the k-nearest neighbor, naive bayes and decision tree algorithms in terms of precision, recall and F-measure in [51] where the authors apply classi cation and topic modeling to ve hacker forums in order to explore the hacker assets like attachments, source code and tutorials; the Maximum Entropy classi er indicated that it is able to nd more actual malware selling and carding promotion threads in a Russian carding forum than SVM, Naïve Bayes, and kNN [30]. In a comparison of forum post classi cation into 9 categories, I. Deliu et al. [15] found that a conventional SVM produced results (approximately 98% accuracy and F1 score) that were on par with more modern Convolutional Neural Networks (CNN). Therefore they concluded that SVMs are superior for the purposes of practical, real-time cyber threat intelligence applications given the computational complexity of CNN architectures. There are also studies like [62] in which only rule-based classi ers are leveraged to identify selling and buying posts. It would be informative if they had the chance to validate their results since we found that purely rule-based classi cation is not enough to achieve an accurate supply/demand classi cation because of same keywords (sell, buy, o er, look for, seek, purchase, etc.) being used in both post categories.

8 CONCLUSIONS AND FUTURE WORK

We presented the rst study to longitudinally measure the prevalence of a large set of Cybercrime-as-a-Service (CaaS) o erings in a large cybercrime forum. In addition, we also compared this to product-type o erings and tracked the usage of contact info in the posts.

We observed that 15.6% of the rst posts in the 'Market' section o ers CaaS. Within this set, 'bot/botnet as a service', 'reputation escalation as a service' and 'tra c as a service' categories constitute the bulk amount (over 60%) for whole period in terms of both supply and demand. When we look into the time series analysis of longitudinal data it is seen that 'bot/botnet as a service' loses its popularity against 'reputation escalation as a service' after 2013 and also against 'hacker as a service' after 2015 over time. The

peak number of trading posts being over 20k per month during 2011-2012 period and continuous decline with small up and downs after the peak is also noticeable in Figure 9. Even though these small increases coincide with times when some big underground markets seized by LEA or closed themselves in the ending months of 2013, 2014 and 2017 [19], we can not tell for sure that these market closures and related loss of consumer trust are the obvious reason for these sudden interest in criminal forums.

An obvious line for future work is to apply our model to different English-language forums and seek for evidence of CaaS there. 'Safeskyhacks' and 'Offensive Community' forums of the crimeBB dataset [14] are two potential forums for these future researches since they also include a 'Market' section. In line with prior work, we only looked into the first posts of the threads in 'market' section. Future work could investigate the following posts of the CaaS related threads in order to understand the feedback for both supply and demand offerings. Examining the vendors offering CaaS would also be interesting for a future study. Finally, to verify the real transactions in the forums, law enforcement can leverage the private messages of seized forums that are related to CaaS offerings.

ACKNOWLEDGMENTS

We thank Cambridge Cybercrime Centre for providing us the CrimeBB dataset.

REFERENCES

[1] A. Abbasi, W. Li, V. Benjamin, S. Hu, and H. Chen. 2014. Descriptive Analytics: Examining Expert Hackers in Web Forums. In *Proceedings of the IEEE Joint Intelligence and Security Informatics Conference (JISIC '14)*. The Hague, The Netherlands, 56 63. https://doi.org/10.1109/JISIC.2014.18

[2] Lawrence Abrams. 2019. Jokeroo Ransomware-as-a-Service Offers Multiple Membership Packages. BleepingComputer. Retrieved August 21, 2020 from https://www.bleepingcomputer.com/news/security/jokeroo-ransomware-as-a-service-offers-multiple-membership-packages/

[3] Luca Allodi, Marco Corradin, and Fabio Massacci. 2016. Then and Now: On the Maturity of the Cybercrime Markets The Lesson That Black-Hat Marketeers Learned. *IEEE Transactions on Emerging Topics in Computing* 4, 1 (2016), 35 46. https://doi.org/10.1109/TETC.2015.2397395

[4] Bradley Barth. 2016. Snack attack: A crimeware-as-a-service menu for wannabe hackers. SC Media. Retrieved July 10, 2020 from https://www.scmagazine.com/snack-attack-a-crimewareas-a-service-menu-for-wannabe-hackers/article/527865

[5] V. Benjamin and H. Chen. 2012. Securing cyberspace: Identifying key actors in hacker communities. In *Proceedings of the IEEE Conference on Intelligence and Security Informatics (ISI '12)*. Arlington, VA, USA, 24 29. https://doi.org/10.1109/ISI.2012.6283296

[6] V. Benjamin, W. Li, T. Holt, and H. Chen. 2015. Exploring threats and vulnerabilities in hacker web: Forums, IRC and carding shops. In *Proceedings of the IEEE International Conference on Intelligence and Security Informatics (ISI '15)*. Baltimore, MD, USA, 85 90. https://doi.org/10.1109/ISI.2015.7165944

[7] Rasika Bhalerao, Maxwell Aliapoulios, Ilia Shumailov, Sadia Afroz, and Damon McCoy. 2019. Mapping the Underground: Supervised Discovery of Cybercrime Supply Chains. In *Proceedings of the APWG Symposium on Electronic Crime Research (eCrime)*. IEEE, Pittsburgh, PA, USA, 1 16. https://doi.org/10.1109/eCrime47957.2019.9037582

[8] Andrew Caines, Sergio Pastrana, Alice Hutchings, and Paula J. Buttery. 2018. Automatically identifying the function and intent of posts in underground forums. *Crime Science* 7, 19 (Nov 2018). https://doi.org/10.1186/s40163-018-0094-4

[9] Michele Campobasso and Luca Allodi. 2020. Impersonation-as-a-Service: Characterizing the Emerging Criminal Infrastructure for User Impersonation at Scale. In *Proceedings of the ACM Conference on Computer and Communications Security (CCS '20)*. Virtual Conference, USA. https://arxiv.org/abs/2009.04344

[10] Chainalysis. 2019. Crypto Crime Report: Decoding increasingly sophisticated hacks, darknet markets, and scams. Technical Report. Chainalysis. https://blog.chainalysis.com/2019-cryptocrime-review

[11] Catalin Cimpanu. 2017. Ranion Ransomware-as-a-Service Available on the Dark Web for 'Educational Purposes'. BleepingComputer. Retrieved August 21, 2020 from https://www.bleepingcomputer.com/news/security/ranion-ransomware-as-a-service-available-on-the-dark-web-for-educational-purposes/

[12] Ben Collier, Daniel R. Thomas, Richard Clayton, and Alice Hutchings. 2019. Booting the Booters: Evaluating the Effects of Police Interventions in the Market for Denial-of-Service Attacks. In *Proceedings of the Internet Measurement Conference (IMC '19)*. ACM, Amsterdam, Netherlands, 50 64. https://doi.org/10.1145/3355369.3355592

[13] Joseph Cox. 2017. This Dark Web Service Makes Spamming Hackers Ridiculously Easy. Vice Media. Retrieved August 21, 2020 from https://www.vice.com/en_us/article/pg5d38/this-dark-web-service-makes-spamming-hackers-ridiculously-easy

[14] Dancho Danchev. 2018. Spamming vendor launches managed spamming service. ZDNet. Retrieved August 21, 2020 from https://www.zdnet.com/article/spamming-vendor-launches-managed-spamming-service

[15] Isuf Deliu, Carl Leichter, and Katrin Franke. 2017. Extracting cyber threat intelligence from hacker forums: Support vector machines versus convolutional neural networks. In *Proceedings of the IEEE International Conference on Big Data (IEEE BigData)*. IEEE, Boston, MA, USA, 3648 3656. https://doi.org/10.1109/BigData.2017.8258359

[16] Europol. 2017. EU Policy Cycle - EMPACT. Europol. Retrieved October 15, 2020 from https://www.europol.europa.eu/crime-areas-and-trends/eu-policy-cycle-empact

[17] Joobin Gharibshah, Evangelos E. Papalexakis, and Michalis Faloutsos. 2018. RIPEx: Extracting Malicious IP Addresses from Security Forums Using Cross-Forum Learning. In *Proceedings of the The Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) (Lecture Notes in Computer Science, Vol. 10939)*. Springer, Melbourne, Australia. https://doi.org/10.1007/978-3-319-93040-4_41

[18] Chris Grier, Lucas Ballard, Juan Caballero, Neha Chachra, Christian J. Dietrich, Kirill Levchenko, Panayiotis Mavrommatis, Damon McCoy, Antonio Nappa, Andreas Pitsillidis, Niels Provos, M. Zubair Rafique, Moheeb Abu Rajab, Christian Rossow, Kurt Thomas, Vern Paxson, Stefan Savage, and Geoffrey M. Voelker. 2012. Manufacturing Compromise: The Emergence of Exploit-as-a-Service. In *Proceedings of the Conference on Computer and Communications Security (CCS '12)*. ACM, Raleigh, North Carolina, USA, 821 832. https://doi.org/10.1145/2382196.2382283

[19] Rick Holland, Rafael Amado, and Michael Marriott. 2018. Seize and Desist? The State of Cybercrime in the Post-AlphaBay and Hansa Age. Technical Report. Digital Shadows. https://info.digitalshadows.com/SeizeandDesistReport-Press_Reg.html

[20] Thomas J. Holt, Deborah Strumsky, Olga Smirnova, and Max Kilger. 2012. Examining the social networks of malware writers and hackers. *International Journal of Cyber Criminology* 6, 1 (1 Jan. 2012), 891 903. http://cybercrimejournal.com/holtetal2012janijcc.pdf

[21] Thomas J. Holt. 2013. Examining the Forces Shaping Cybercrime Markets Online. *Social Science Computer Review* 31, 2 (2013), 165 177. https://doi.org/10.1177/0894439312452998

[22] Keman Huang, Michael Siegel, and Stuart Madnick. 2018. Systematically Understanding the Cyber Attack Business: A Survey. *ACM Comput. Surv.* 51, 4, Article 70 (July 2018), 36 pages. https://doi.org/10.1145/3199674

[23] Alice Hutchings and Richard Clayton. 2016. Exploring the Provision of Online Booter Services. *Deviant Behavior* 37, 10 (2016), 1163 1178. https://doi.org/10.1080/01639625.2016.1169829

[24] Alice Hutchings and Sergio Pastrana. 2019. Understanding eWhoring. In *Proceedings of the IEEE European Symposium on Security and Privacy (EuroS&P '19)*. Stockholm, Sweden, 201 214. https://doi.org/10.1109/EuroSP.2019.00024

[25] Thomas S. Hyslip and Thomas J. Holt. 2019. Assessing the Capacity of DRDoS-For-Hire Services in Cybercrime Markets. *Deviant Behavior* 40, 12 (2019), 1609 1625. https://doi.org/10.1080/01639625.2019.1616489

[26] Mohammad Karami, Youngsam Park, and Damon McCoy. 2016. Stress Testing the Booters: Understanding and Undermining the Business of DDoS Services. In *Proceedings of the 25th International Conference on World Wide Web (WWW '16)*. ACM, Montréal, Québec, Canada, 1033 1043. https://doi.org/10.1145/2872427.2883004

[27] Emily Kimpton and Helen Thackray. 2020. Knowledge is Power: An Analysis of Discussions on Hacking Forums. In *Proceedings of WACCO 2020: 2nd Workshop on Attackers and Cyber-Crime Operations, IEEE EuroS&PW 2020*. Conference, 477 482. https://conferences.computer.org/eurosp/pdfs/EuroSPW2020-7k9FlVRX4z43j4uE2SeXU0/859700a477/859700a477.pdf

[28] Daniel Kopp, Matthias Wichtlhuber, Ingmar Poese, Jair Santanna, Oliver Hohlfeld, and Christoph Dietzel. 2019. DDoS Hide & Seek: On the Effectiveness of a Booter Services Takedown. In *Proceedings of the Internet Measurement Conference (IMC '19)*. ACM, Amsterdam, Netherlands, 65 72. https://doi.org/10.1145/3355369.3355590

[29] E. Rutger Leukfeldt, Edward R. Kleemans, and Wouter P. Stol. 2016. Cybercriminal Networks, Social Ties and Online Forums: Social Ties Versus Digital Ties Within Phishing and Malware Networks. *The British Journal of Criminology* 57, 3 (Feb. 2016), 704 722. https://doi.org/10.1093/bjc/azw009

[30] Weifeng Li and Hsinchun Chen. 2014. Identifying Top Sellers In Underground Economy Using Deep Learning-Based Sentiment Analysis. In *Proceedings of the Joint Intelligence and Security Informatics Conference (JISIC '14)*. IEEE, The Hague, Netherlands, 64–67. https://doi.org/10.1109/JISIC.2014.19

[31] Jonathan Lusthaus. 2013. How organised is organised cybercrime? *Global Crime* 14, 1 (2013), 52–60. https://doi.org/10.1080/17440572.2012.759508

[32] M. Macdonald, R. Frank, J. Mei, and B. Monk. 2015. Identifying digital threats in a hacker web forum. In *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM '15)*. Paris, France, 926–933. https://doi.org/10.1145/2808797.2808878

[33] Derek Manky. 2013. Cybercrime as a service: a very modern business. *Computer Fraud & Security* 2013, 6 (2013), 9 – 13. https://doi.org/10.1016/S1361-3723(13)70053-8

[34] Per Håkon Meland, Yara Fareed Fahmy Bayoumy, and Guttorm Sindre. 2020. The Ransomware-as-a-Service economy within the darknet. *Computers & Security* 92 (2020), 101762. https://doi.org/10.1016/j.cose.2020.101762

[35] Ariana Mirian, Joe DeBlasio, Stefan Savage, Geoffrey M. Voelker, and Kurt Thomas. 2019. Hack for Hire: Exploring the Emerging Market for Account Hijacking. In *Proceedings of The World Wide Web Conference (WWW '19)*. San Francisco, CA, USA, 1279–1289. https://doi.org/10.1145/3308558.3313489

[36] Marti Motoyama, Kirill Levchenko, Chris Kanich, Damon McCoy, Geoffrey M. Voelker, and Stefan Savage. 2010. Re: CAPTCHAs: Understanding CAPTCHA-Solving Services in an Economic Context. In *Proceedings of the 19th USENIX Conference on Security (USENIX Security'10)*. Washington, DC, 28.

[37] Marti Motoyama, Damon McCoy, Kirill Levchenko, Stefan Savage, and Geoffrey M. Voelker. 2011. An Analysis of Underground Forums. In *Proceedings of the ACM SIGCOMM Conference on Internet Measurement Conference (IMC '11)*. Berlin, Germany, 71–80. https://doi.org/10.1145/2068816.2068824

[38] Arman Noroozian, Jan Koenders, Eelco Van Veldhuizen, Carlos H. Ganan, Sumayah Alrwais, Damon McCoy, and Michel Van Eeten. 2019. Platforms in Everything: Analyzing Ground-Truth Data on the Anatomy and Economics of Bullet-Proof Hosting. In *Proceedings of the 28th USENIX Conference on Security Symposium* (Santa Clara, CA, USA) *(SEC'19)*. USENIX Association, USA, 1341–1356.

[39] Arman Noroozian, Maciej Korczyński, Carlos Hernandez Gañan, Daisuke Makita, Katsunari Yoshioka, and Michel van Eeten. 2016. Who Gets the Boot? Analyzing Victimization by DDoS-as-a-Service. In *Proceedings of the Research in Attacks, Intrusions and Defenses (RAID '16) (Lecture Notes in Computer Science, Vol. 9854)*. Springer, Evry, France, 1–20. https://doi.org/10.1007/978-3-319-45719-2_17

[40] Eric Nunes, Ahmad Diab, Andrew Gunn, Ericsson Marin, Vineet Mishra, Vivin Paliath, John Robertson, Jana Shakarian, Amanda Thart, and Paulo Shakarian. 2016. Darknet and deepnet mining for proactive cybersecurity threat intelligence. In *Proceedings of the Conference on Intelligence and Security Informatics (ISI)*. IEEE, Tucson, AZ, 7–12. https://doi.org/10.1109/ISI.2016.7745435

[41] Rebekah Overdorf, Carmela Troncoso, Rachel Greenstadt, and Damon McCoy. 2018. Under the Underground: Predicting Private Interactions in Underground Forums. arXiv:1805.04494 [cs.CR]

[42] Sergio Pastrana, Alice Hutchings, Andrew Caines, and Paula Buttery. 2018. Characterizing Eve: Analysing Cybercrime Actors in a Large Underground Forum. In *Proceedings of the 21st Research in Attacks, Intrusions Symposium (RAID '18) (Lecture Notes in Computer Science, Vol. 11050)*. Springer, Crete, Greece, 207–227. https://doi.org/10.1007/978-3-030-00470-5_10

[43] Sergio Pastrana, Alice Hutchings, Daniel Thomas, and Juan Tapiador. 2019. Measuring EWhoring. In *Proceedings of the Internet Measurement Conference (IMC '19)*. Amsterdam, Netherlands, 463–477. https://doi.org/10.1145/3355369.3355597

[44] Sergio Pastrana, Daniel R. Thomas, Alice Hutchings, and Richard Clayton. 2018. CrimeBB: Enabling Cybercrime Research on Underground Forums at Scale. In *Proceedings of the 2018 World Wide Web Conference (WWW '18)*. ACM, Lyon, France, 1845–1854. https://doi.org/10.1145/3178876.3186178

[45] Ildiko Pete, Jack Hughes, Yi Ting Chua, and Maria Bada. 2020. A Social Network Analysis and Comparison of Six Dark Web Forums. In *Proceedings of the IEEE European Symposium on Security and Privacy Workshops (EuroS&PW 2020)*. Virtual Conference, 483–492. https://conferences.computer.org/eurosp/pdfs/EuroSPW2020-7k9FlVRX4z43j4uE2SeXU0/859700a483/859700a483.pdf

[46] Michael Peters. 2019. *What Is Ransomware-as-a-Service? Understanding RaaS*. Security Boulevard. Retrieved August 21, 2020 from https://securityboulevard.com/2019/02/what-is-ransomware-as-a-service-understanding-raas

[47] Matías Porolli. 2019. *Cybercrime black markets: Dark web services and their prices*. ESET. Retrieved August 21, 2020 from https://www.welivesecurity.com/2019/01/31/cybercrime-black-markets-dark-web-services-and-prices

[48] Michael E. Porter. 1985. *Competitive advantage: Creating and sustaining superior performance*. Free Press.

[49] Rebecca S. Portnoff, Sadia Afroz, Greg Durrett, Jonathan K. Kummerfeld, Taylor Berg-Kirkpatrick, Damon McCoy, Kirill Levchenko, and Vern Paxson. 2017. Tools for Automated Analysis of Cybercriminal Markets. In *Proceedings of the 26th International Conference on World Wide Web (WWW '17)*. ACM, Perth, Australia, 657–666. https://doi.org/10.1145/3038912.3052600

[50] Sagar Samtani and Hsinchun Chen. 2016. Using social network analysis to identify key hackers for keylogging tools in hacker forums. In *Proceedings of the IEEE International Conference on Intelligence and Security Informatics: Cybersecurity and Big Data, ISI '16*. Tucson, Arizona, USA, 319–321. https://doi.org/10.1109/ISI.2016.7745500

[51] Sagar Samtani, Ryan Chinn, and Hsinchun Chen. 2015. Exploring hacker assets in underground forums. In *Proceedings of the International Conference on Intelligence and Security Informatics (ISI '15)*. IEEE, Baltimore, MD, USA, 31–36. https://doi.org/10.1109/ISI.2015.7165935

[52] Sagar Samtani, Ryan Chinn, Hsinchun Chen, and Jay F. Nunamaker Jr. 2017. Exploring Emerging Hacker Assets and Key Hackers for Proactive Cyber Threat Intelligence. *Journal of Management Information Systems* 34, 4 (2017), 1023–1053. https://doi.org/10.1080/07421222.2017.1394049

[53] Aditya K. Sood and Richard J. Enbody. 2013. Crimeware-as-a-service—A survey of commoditized crimeware in the underground market. *International Journal of Critical Infrastructure Protection* 6, 1 (2013), 28 – 38. https://doi.org/10.1016/j.ijcip.2013.01.002

[54] Pedro Tavares. 2018. *Mechanics Behind Ransomware-as-a-Service*. INFOSEC. Retrieved August 21, 2020 from https://resources.infosecinstitute.com/mechanics-behind-ransomware-as-a-service

[55] Kurt Thomas, Danny Yuxing Huang, David Y. Wang, Elie Bursztein, Chris Grier, Tom Holt, Christopher Kruegel, Damon McCoy, Stefan Savage, and Giovanni Vigna. 2015. Framing Dependencies Introduced by Underground Commoditization. In *Proceedings of the 14th Annual Workshop on the Economics of Information Security, WEIS '15*. Delft, The Netherlands. http://www.econinfosec.org/archive/weis2015/papers/WEIS_2015_thomas.pdf

[56] Kurt Thomas, Dmytro Iatskiv, Elie Bursztein, Tadek Pietraszek, Chris Grier, and Damon McCoy. 2014. Dialing Back Abuse on Phone Verified Accounts. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14)*. Scottsdale, Arizona, USA, 465–476. https://doi.org/10.1145/2660267.2660321

[57] Ashwin Vamshi. 2017. *Phishing as a Service - Phishing revamped*. Netskope. Retrieved August 21, 2020 from https://www.netskope.com/blog/phishing-service-phishing-revamped

[58] Rolf van Wegberg, Samaneh Tajalizadehkhoob, Kyle Soska, Ugur Akyazi, Carlos Hernandez Ganan, Bram Klievink, Nicolas Christin, and Michel van Eeten. 2018. Plug and Prey? Measuring the Commoditization of Cybercrime via Online Anonymous Markets. In *Proceedings of the 27th USENIX Security Symposium (USENIX Security '18)*. USENIX Association, Baltimore, MD, 1009–1026. https://www.usenix.org/conference/usenixsecurity18/presentation/van-wegberg

[59] Anh V. Vu, Jack Hughes, Ildiko Pete, Ben Collier, Yi Ting Chua, Ilia Shumailov, and Alice Hutchings. 2020. Turning Up the Dial: the Evolution of a Cybercrime Market Through Set-up, Stable, and COVID-19 Eras. In *Proceedings of the ACM Internet Measurement Conference (IMC '20)*. Virtual Conference, USA. https://www.researchgate.net/profile/Anh_Vu110/publication/344418946_Turning_Up_the_Dial_the_Evolution_of_a_Cybercrime_Market_Through_Set-up_Stable_and_Covid-19_Eras/links/5f73a395a6fdcc0086483051/Turning-Up-the-Dial-the-Evolution-of-a-Cybercrime-Market-Through-Set-up-Stable-and-Covid-19-Eras.pdf

[60] Marc Wilczek. 2018. *Cybercrime-as-a-Service: No End in Sight*. Link11. Retrieved August 21, 2020 from https://www.darkreading.com/endpoint/cybercrime-as-a-service-no-end-in-sight/a/d-id/1333033

[61] Haitao Xu, Daiping Liu, Haining Wang, and Angelos Stavrou. 2015. E-Commerce Reputation Manipulation: The Emergence of Reputation-Escalation-as-a-Service. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*. ACM, Florence, Italy, 1296–1306. https://doi.org/10.1145/2736277.2741650

[62] Ziming Zhao, Mukund Sankaran, Gail-Joon Ahn, Thomas J. Holt, Yiming Jing, and Hongxin Hu. 2016. Mules, Seals, and Attacking Tools: Analyzing 12 Online Marketplaces. *IEEE Security & Privacy* 14, 3 (2016), 32–43. https://doi.org/10.1109/MSP.2016.46