An Analysis of Fake Social Media Engagement Services

David Nevado-Catalán[1], Sergio Pastrana[1], Narseo Vallina-Rodriguez[2], Juan Tapiador[1,*]

Abstract

Fake engagement services allow users of online social media and other Internet platforms to illegitimately increase their online reach and boost their perceived popularity. Driven by social, economic, and political motivations, actors are increasingly demanding these services, which has incentivized the rise of a vast underground market tailored to the fake engagement ecosystem. Prior research in this area has been limited to the study of the infrastructure used to provide these services (e.g., botnets), and to the development of fake engagement detection algorithms in the targeted platforms. In contrast, the platforms in which these services are sold (known as *panels*), and the social media fake engagement market at large has not received much research attention. This paper studies Social Media Management (SMM) panels, a key piece of this complex ecosystem. These panels are reselling platforms in which a large variety of fake engagement services are offered, often advertised in underground forums. We crawl 86 representative SMM panels daily for 4 months, generating a dataset with 2.8M entries grouped into 61k different services. This dataset allows us to elaborate a catalog with all the services for sale and the platforms they target. Specifically we carry out an economic analysis of fake engagement services, including trading activities by analyzing 7k threads in underground forums. Our analysis shows an extensive variety of customizations that allow buyers to select features such as the quality of the service, the speed of delivery, the country of origin, as well as personal attributes of the fake account (e.g., gender). The price analysis reveals significant disparities between prices of the same product across different markets. This might be an indication that the market is still undeveloped and sellers do not know the worth of the services they offer, leading them to underprice or overprice. We also observe a drop in prices with respect to values reported by previous research. Overall, our results provide insights that complement previous research efforts, shedding some light on the market operation and helping to gain a better understanding of the ecosystem.

[*]Corresponding author

Email addresses: `100421545@alumnos.uc3m.es` (David Nevado-Catalán), `spastran@inf.uc3m.es` (Sergio Pastrana), `narseo.vallina@imdea.org` (Narseo Vallina-Rodriguez), `jestevez@inf.uc3m.es` (Juan Tapiador)

[1]Universidad Carlos III de Madrid.

[2]IMDEA Networks, ICSI, and AppCensus

Keywords: Fake engagement services, social networks, cybercrime, security economics

1. Introduction

Online social media play a critical role in our day to day life. In the last decade Online Social Networks (OSN) have proven to be an effective marketing tool for businesses as well as a platforms for sharing information and news. This makes them an attractive target for actors that realize the potential of increasing their online presence and influence. The pursue of these goals has fueled the development of a vast underground social media fake engagement market. This market offers a wide variety of services, such as Instagram followers or Facebook likes, potentially obtained through illicit means [8, 19, 24, 7]. These services provide interactions and are used to manipulate OSNs' recommendation algorithms in order to increase reach and visibility of online profiles, as well as boosting their perceived popularity.

Previous research on the topic has demonstrated that fake engagement services are a lucrative business involving multiple actors along its supply chain. As in the case of other cyberthreats [17, 6, 32, 11, 27, 24], botnets appear to be the main supplier of fake engagement services [24]. However, before these services reach their final customer, they are often traded and resold in underground online platforms [23]. This phenomenon is confirmed by the proliferation of Social Media Management (SMM) panels, which are essentially platforms that act as an intermediary between suppliers and intermediate or end users. Fake engagement and inorganic interactions in OSN have also been studied from the platforms' perspective in an effort to identify and eliminate fraudulent behavior [18, 33, 20, 16, 36]. However, the research community has overlooked the economics and operation of the fake engagement ecosystem has been quite limited (we review related work in Section 2).

In this work we carry out an extensive and empirical analysis of the underground economy of SMM trading. Our key research goals are to study the catalog of fake engagement services that are offered, their main features and prices, and the platforms that they target. To do so, we first identify a representative set of SMM panels collected both from general-purpose search engines and from two specialized underground forums—Hackforums and BlackHatWorld. We then crawl and analyze data from the resulting set panels daily for 4 months, a dataset which we complement with trading activities and the conversations related to fake engagement services occurring in *Hackforums* and *BlackHatWorld*, which are popular underground forums focused on the trading and discussion of illicit activities [25, 29] that contain dedicated boards for social media products and services. In total, we collect and compile a dataset from 58 SMM panels totaling 2.8M listings. We also leverage a large dataset of underground forum data to study more than 7,063 discussion threads offering SMM services.

The main findings and contributions of our analysis are:

- We conduct a quantitative study of the market providing fake engagement services for social media. To do so, we compile a dataset of offered services by crawling daily 86 SMM panels where they are advertised during a period of 4 months. This dataset consists of 2.8M entries grouped in 61k different service variations. Using this dataset, we have identified 294 different services targeting 59 platforms including the major Internet OSN, review services, video and music platforms.

- We analyze the catalog of available engagement services and observe that most of them are offered with an impressive variety of customizations, including the quality of the service, the speed of delivery, the country of origin, as well as personal attributes of the fake account such as the gender. Such a rich catalog indicates that the market counts on a substantial underlying infrastructure to deliver the services.

- We perform an economic analysis of the ecosystem. Specifically, we analyze their prices and their variations across markets, and how different customizations affect the market. Our results indicate lower prices than those reported in previous research, and also large disparities between the price for the same service across markets.

- We complement our analysis with a study of the presence of these services in underground forums. Our results confirm that they align with those being sold in dedicated panels. Also, we observe that actors re-selling SMM services tend to start providing some free products to gain reputation, and that they complement with other illicit activities.

Overall, our research sheds light on the vast underground ecosystem of fake SMM. We show that dedicated panels services are prevalent, though often volatile, and reachable through forum advertisements. We show that prices fluctuate across sites, which suggests that this market is at an early stage, and that the underlying supply chain supporting it is still unstable. We also observe how customized services (e.g., followers from specific gender or location) increase prices and provides specialization to re-sellers.
Dataset. We open source our dataset at `https://github.com/davidnevadoc/fake_SE/tree/master/Datasets`

2. Related work

Fake engagement in OSN is a problem that has gained relevance in the last decade. Many studies have analyzed how malicious actors abuse social media platforms for varying purposes, including marketing fraud such as fake product reviews [20, 22] and spamming campaigns [14], but also political and social manipulation [3] and other forms of misinformation [34]. More broadly, these services are used as a means to boost influence and increase reach.

The fake engagement problem has been approached from multiple perspectives. One line of study mainly focuses on the detection of the presence of fraudulent entities [5] or behavior (e.g., fake reviews [31] or spamming [13]) in OSN. In particular, fake engagement detection studies have been carried out for the most prominent platforms such as Facebook [14, 8], Twitter [35, 36], YouTube [18] and Instagram [38, 33]. These studies present a wide variety of methods leveraging Machine Learning (ML) techniques, specially behavioral clustering to study community structures and identify groups exhibiting common patterns of behavior [16, 37, 4].

Another line of study focuses on how these services are delivered, which actors are involved, and how it fits into the broader cybercriminal ecosystem. A 2016 work [24] studies the Linux/Moose botnet and how it is used for Social Media Fraud (SMF). This work not only deals with the technical aspects and operation of the botnets but also with the platforms that it targets, its clients, and potential motivations and an estimation of the revenue generated by its operators. This research was continued in 2018 with a heavier focus on the market aspects of the SMF supply chain [23]. In particular, it elaborates on the relationship between the botnet and the reseller panels and how the revenue is distributed among these actors. Our work extends the research carried out in these studies by making an in-depth analysis of the reselling panels, gathering an exhaustive catalog of the offered services, and estimating their prices attending their different variations.

3. Datasets

In this section we describe the two datasets used to carry out our analysis: data crawled from online SMM Panels and the CrimeBB dataset of underground forums.

3.1. SMM Panels dataset.

In order to gather a representative dataset for our study, we first searched for popular sites trading fake engagement services. These sites are commonly referred as Social Media (Marketing) Panels or SMM Panels. We compiled a list of these panels using two methods: (i) By manually doing Google searches using terms such as *"buy Instagram likes"*, *"buy Facebook likes"*, *"buy followers"*, etc.; and (ii) browsing 2 underground forums – Hackforums and BlackHatWorld – where these services are commonly advertised. In addition, we used a compiled list of 343 such panels provided in a previous study [23]. However this list was put together in 2018 and most of its panels are no longer up as of this writing. In fact, some of the panels we initially collected also went down shortly after we started crawling data. The panels that went down in the 2 fist weeks of crawling were discarded and not added to the dataset. Our final list is composed by 58 panels, from which we build our dataset. The size of these panels in terms of offered services is quite diverse, however most of them (75%) did not exceed 1000 services offered simultaneously. The full distribution of panel sizes is illustrated in Table 1.

Daily entries	Nº panels	%	cumulative %
0 - 200	5	9.62	9.62
200 - 400	7	13.46	23.08
400 - 600	11	21.15	44.23
600 - 800	8	15.38	59.62
800 - 1000	8	15.38	75.00
1000 - 1200	3	5.77	80.77
1200 - 1400	3	5.77	86.54
1400 - 1600	4	7.69	94.23
1600+	3	5.77	100.00

Table 1: Median number of service entries in SMM panels.

Crawling strategy. We implemented a custom web crawler to gather the services offered in the selected 58 panels. The crawler visited each panel daily from March 20th to August 17th 2020, recovering from each page the tables where the services are advertised. We then parse the tables to obtain structured data (i.e., product or service entries) that will be subsequently analyzed. This step involved some manual analysis to customize the parser to particularities of each panel. First, we classify each entry according to the target platform (e.g., Instagram, YouTube, Facebook, etc.) and the provided service (e.g., likes, followers, comments, etc.). We then check each service name and description for the presence of a set of keywords that indicate different variations of its provision, such as geographical or quality modifiers (see section 4.2).
Price normalization. Service prices are converted to USD [3] and reviewed to conform to the format '$ per 1000', since this is the format in which the majority of services are expressed. Some manual review was necessary as some services do not conform to the format and adjustments needed to be made after reading the service description. This was typically the case for expensive services, like Amazon or Google Business reviews. In these cases the price was *per unit* despite the name of the column. Even after manual review, these adjustments were difficult to make as it is common to find inconsistencies and contradictions within the service fields. In many cases the services' *Name* field specifies a maximum amount available to order (e.g., "Instagram likes [50k]") that does not correspond with the amount under the *Max. order* field. This issue is specially concerning when two different prices are given in the *Price* and *Description* fields. We later found out that this is a consequence of reselling: Some panels resell a service copying the name and description but changing the price, often creating such contradictions in the process. In these cases we choose the value specified in the *Price* field. After this process we obtained a curated

[3] Exchange rate on 16/10/2020: EUR 0.84, IDR 14.7k, IR 71.43

dataset of 2.8M records.

Service indexation. We processed the original dataset to generate a second dataset with aggregated service data. The second dataset contains each different service only once, with extra fields that indicate its duration in days, number of price changes during the crawling period, and basic metrics on its prices (i.e., mean, standard deviation, maximum and minimum values, and quartiles). To generate this dataset, we first needed to differentiate services, within and across panels, which is not straightforward because it is common to find the exact same service name and description in several panels. We even find services having the same spelling mistakes. However, even if this may indicate a common underlying service or a resell, we have no ground truth to draw any solid conclusion. In the same way, we cannot assume that services that use different wording in their names or description are different ones. One may be a resell of the other, or both of them may be resells of an underlying common service. Having acknowledged this uncertainty, we established the following criteria for differentiating services in our analysis: Two services are considered different if: (i) they come from different panels; or (ii) they have a different ID within the panel; or (iii) according to the preprocessing, the service has undergone a significant modification (typically, due to new features being removed or added). The result of this process is a dataset formed by 61k different services.

3.2. CrimeBB dataset

We also use the CrimeBB dataset [28] to study the fake engagement services ecosystem in underground forums. This dataset is freely available for researchers from the Cambridge Cybercrime Centre.[4] We got access to a dataset containing more than 91 million posts gathered from 34 different underground forums. Some of them are general-purpose forums, while others are specialized (e.g., in video-game hacks and cheats, malware or online accounts). The forums are subdivided in different categories and bulletin boards dedicated to particular activities, like hacking, graphics or marketplaces. Concretely, five of these forums contain special sections dedicated to SMM products and services. We selected these SMM related threads and collected a dataset by extracting the heading (title given to a conversation thread) and the content of the first post of each thread, its author, the number of replies, the post timestamp and the forum in which it was posted. After this process, we obtained a dataset of 27k threads.

Thread classification. A preliminary exploration of the data shows that the posts gathered, while having a common theme, have different purposes. We could find tutorials and guides (e.g., on how to grow popularity of Twitter [5] or YouTube accounts [6]), requests for advice or help, service offers, and service requests. As our focus is to study the market offer, we trained a classifier to identify offerings, i.e., threads that offer SMM products and services. First we

[4] `www.cambridgecybercrime.uk`

[5] https://hackforums.net/showthread.php?tid=5854028

[6] https://hackforums.net/showthread.php?tid=2664677

constructed a labeled dataset of 1.2k entries by leveraging common tags used in the thread headings, which are usually placed between square brackets and which indicate offers (e.g., `[wts]` stands for *want to sell*) or other purposes (e.g.,`[wtb]`, which stands for *want to buy*) We then built an NLP classifier using a Recursive Neural Network (RNN) [1] on top of a pre-trained encoder (BERT) [9] and trained it on 70% of the labeled posts. When tested on the remaining 30%, which is used as validation set, we obtain a F1 of 0.983 (accuracy 0.980, precision 0.986, recall 0.981). These are good performance metrics since the dataset is balanced (45% vs. 55%). Using this classifier on the unlabeled threads, we identified a total of 7k threads related to Social Media (SM) service offers, sales, or advertisements, which we use as a base for our analysis in section 6.2.

4. The fake engagement ecosystem

In this section we study the market ecosystem of the SMM panels. We first describe a catalog of all the services found in our dataset in order to illustrate the variety and scale of this market. Then, we classify these services and identify the social media platforms for which they offer services. To do so, we define several metrics to explore which online platforms are the most popular across this panels (more details are given in Section 4.2). We conclude this section by analyzing the size of the studied panels.

4.1. Prominent services offered and platforms targeted

Catalog. We identified a total of 294 different services (excluding variations or customizations) across 59 different platforms. The obtained catalog of services is summarized in Figure 1. Due to the dimension of the catalog, and to ease visualization, less popular services and platforms have been grouped together.

There is great disparity in the popularity of the offered services. On the one hand, a couple of dozens of services are present in nearly all the studied panels. In addition, each panel contains a substantial amount of entries for each of them, offering the services in different variations. On the other hand, we find many small services that are typically found only at the bottom of a few panels. Among the prominent services we find many fake engagement services for popular OSNs like Instagram and Facebook. Website traffic (visits) is also a common service offered, with listings being present in most panels and having many available customizations. We also observe a significant amount of services directed to music platforms, mostly consisting of fake plays. The popularity of fake plays is perhaps more surprising than that of the previous two. We speculate that their popularity may be partially caused by the simplicity in which the service can be provided. One account can provide many plays for the same track, playlist or album. In fact, in some of these platforms no account is required. In addition, there is no need for user interaction at all in order to simulate organic behavior, which is something crucial in other services in order to not get banned by the platform. In general the offer for SMM services in

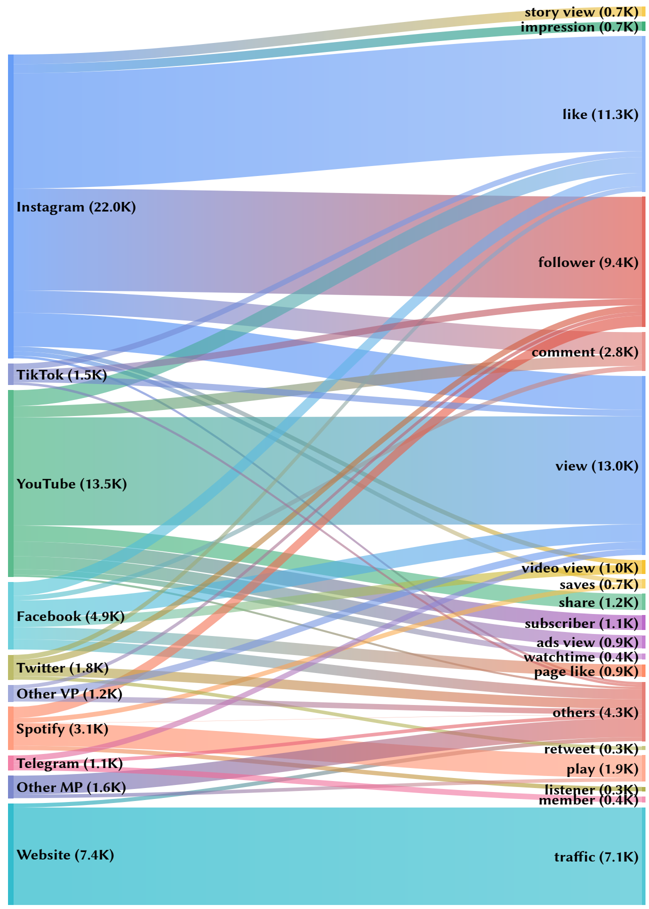

Figure 1: Catalog of the most popular services found in SMM panels. Small music/audio and video platforms have been grouped under the labels *Other MP* and *Other VP*.

platforms for independent music artists reflects the demand of cheap marketing strategies in a very crowded and competitive environment.

We find marginal services that present interesting characteristics. They fall into the following categories: (i) Premium accounts for video platforms (e.g., Netflix, Disney+, HBO), music platforms (e.g., Spotify, Amazon Prime Music) and adult platforms (e.g. Brazzers) ; (ii) real-looking accounts (with profile picture, followers, posts, etc.) for Instagram, Twitter and other OSN; (iii) organic mobile applications installs for the Apple Store and Google Play[12]; (iv) reviews and ratings for sites (e.g., Amazon, Google Business, LinkedIn, IMDb and TripAdvisor). We take a closer look at SMMs offering more expensive services such as review and rating services, and access to premium accounts in section 5.2.

Lastly, it is worth mentioning that among the offered services it is common to find the SMM panels themselves. These services offer packages that include hosting, a front-end website and an API for the easy deployment of a reselling panel. The presence of these listings illustrate the common practice of reselling in this market [23], which we confirm during our analysis of underground forum data in section 6.

Popularity. In order to quantify the popularity of services across platforms in these panels, we use the following 3 metrics: (i) The mean number of daily entries of each service across all panels; (ii) the number of different variations for each service identified during the study duration; and (iii) the percentage of panels where the service was present. The top 20 services according to these metrics are presented in Table 2.

It is important to note that although these metrics combined give an idea of the popularity of the services offered, they may not reflect their demand, nor indicate the amount of fake engagement present in the target platforms. It is also relevant to point out that each metric is biased towards different kind of services. For example, web traffic is offered in many panels from different locations and with different referrers, as a consequence, it is overrepresented in terms of *entries/day*. Similarly, platforms that offer many different forms of interactions such as Instagram (that has likes, followers, impressions, views, story views, IGTV views, saves, reactions, etc.) are overrepresented compared to simpler ones in the *different variations* metric.

The top services show that likes and its variations are the most popular services for OSN (e.g., Instagram, Facebook). For video platforms (e.g., YouTube, Twitch) and music platforms (e.g., Spotify) the most popular services are, unsurprisingly, views and plays. However, YouTube's second most popular service is again like, possibly as a result of two factors: (i) likes are distinctive indicators used by the platforms to measure the quality of the content (e.g,. posts, videos, songs, etc.) so their manipulation can impact the recommendations algorithms; (ii) as opposed to other features like comments, likes are easier to manipulate automatically (e.g., botnets). This, as we will see in Section 5 makes them much cheaper, and thus boosts their popularity on the market.

We have used the same popularity metrics defined above to rank the platforms. The top 8 platforms are represented in Figure 2 indicating both their daily entries and number of service variations. We observe that Instagram and YouTube accumulate most of the offered services by a substantial margin, followed by Facebook, Spotify and Twitter. When we consider these results in relation to the number of active users of each platform we observe that Instagram (1,082 M) is by far the most targeted for its size compared to Facebook (2,603 M), YouTube (2,000 M) and Twitter (326 M)[7]. Figure 2 also showcases the bias of the used metrics we previously mentioned, which is particularly clear for Website.

[7]Source: Most popular social networks worldwide as of July 2020 (Statista.com)

Site	Product	Entries/day	different variations	% panels
Website	traffic	4,695± 493	7,066	72.4
Instagram	like	2,677± 235	8,362	100.0
YouTube	view	2,524± 436	7,836	98.3
Instagram	follower	1,995± 236	7,390	100.0
Instagram	view	1,084± 70	2,446	100.0
Spotify	play	971± 89	1,639	87.9
Instagram	comment	700± 52	1,622	94.8
YouTube	like	453± 35	1,151	96.6
Spotify	follower	440± 52	811	82.8
Facebook	like	401± 20	983	94.8
Facebook	page like	386± 44	883	93.1
YouTube	share	441± 145	1,165	84.5
YouTube	comment	390± 35	822	82.8
YouTube	ads view	351± 88	940	63.8
Instagram	impression	326± 20	654	91.4
Instagram	story view	311± 19	704	91.4
Twitch	view	290± 21	527	60.3
Facebook	video view	316± 49	657	94.8
Facebook	view	267± 46	1,272	70.7
YouTube	subscriber	267± 47	1,066	96.6

Table 2: Top 20 services in SMM panels.

4.2. Service customization

The listings advertised in the SMM panels typically provide a description of the offered service , such as its quality, the form and speed of delivery, or the refund policy. For many services offered, we can differentiate between low quality (or standard versions) from improved (or premium) versions. The latter come at a much higher price but promise more quality and reliable services. An example of this are the *comments* offered for multiple platforms. It is common to find very cheap services that provide *random* comments, whereas other services offer *customized* and real-looking ones for the same platform at a higher prize. We observe the same phenomenon for followers in various social networks. In this case, the cheapest services provide follows through bot-controlled low quality accounts, i.e., accounts with no publications and no followers. At increasing prices we see how they offer different forms of customization: more real-looking accounts or even the possibility of choosing the features such as the country or the gender of the account.

In order to analyze these customizations and their impact on the services' prices, we first identify recurring keywords and tags that characterize them. Then, we classify the services according to the presence of these keywords in the services names and descriptions. Since we do not actually buy the service,

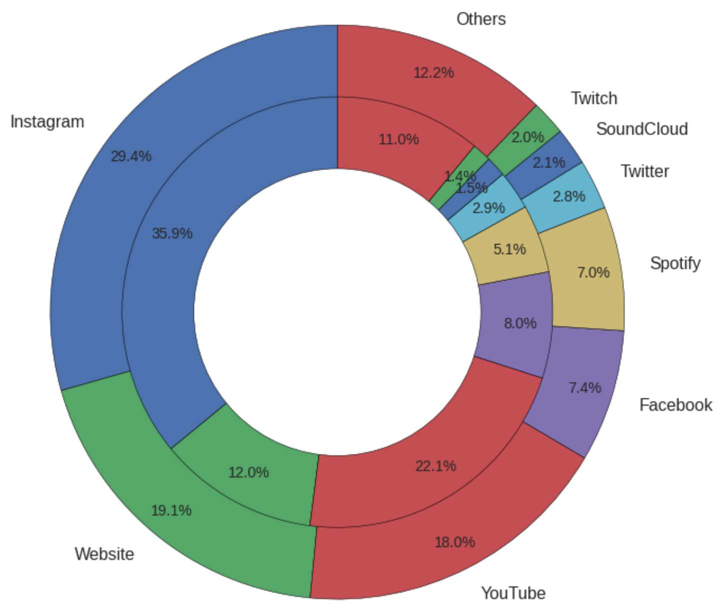

Figure 2: Top 8 targeted platforms by services in SMM panels. The inner ring represents the percentage of different services (variations) identified. The outer ring represents the percentage of mean daily entries.

we are not able to determine up to which point these keywords reflect a real difference in the service provision or are just a means to make it look more appealing. We next present the more relevant keywords identified and their presence in the panels. Later in Section 5 we will discuss their impact on the services' prices.

- **Bot / Real / Active**. These keywords are common across all services, especially `Real`. The use of these keywords indicates characteristics regarding the accounts used to provide services such as likes or followers. However, the concept of `Real` actions is quite broad and it is mostly used in other services like views and plays. In these cases it is not clear what information this keyword conveys about the service and is likely used as a marketing mechanism. The `Bot` keyword is much less common, as usual with negative keywords. Nonetheless it is fairly common for YouTube shares and App installs.

- **HQ / LQ**. These are common keywords found in the listings. However, they do not provide much information on their own, and need to be analyzed in the context of a particular service. For example, in a service like

Instagram followers they could refer to the quality of the account (real-looking or not), and in a *Spotify play* service they could refer to aspects such as drop rate or delivery time. Also, in many cases they may not provide actual information about the service, and their presence could be just a marketing strategy.

- **Drip / No Drip**. In contrast to the previous, these keywords are quite specific and are common only in a few services: *Spotify plays* and *followers*, *Twitch views* and *Website traffic*. The `Drip` keyword refers to a gradual delivery during an established period of time. This may be a desirable feature as it gives the appearance of a more realistic growth. It may also help avoid the detection mechanisms of the platforms, although this is unclear.

- **Drop / No Drop**. Many platforms and OSN try to detect and eliminate inorganic engagement. For this reason many low-quality services are expected to have substantial drop rates shortly after their delivery. Some services advertise the expected drop rate or state that there is no drop in their service. These keywords are common across most studied services.

- **Refill / No Refill**. As a result of platforms eliminating inorganic content and banning fake accounts some services such as *followers*, *likes*, or *plays* can suffer significant drop rates over time. Some services offer to compensate for these drop rates by refilling with the required service until reaching the agreed amount. This is usually not offered indefinitely but for a certain period of time: a few weeks or months, normally. We observe that refilling is common across most studied services.

- **Custom / Random**. These variations mostly appear for comments, specially in Instagram and YouTube. It is also one of the most influential customizations price-wise. Custom comments are several times more expensive that random or generic ones (see §5). We can attribute this price difference to the fact that custom comments may need some degree of human intervention or advance linguistic tools, which would substantially increase the cost of an otherwise completely automated process. This necessity of human intervention is probably the main reason behind reviews in *Amazon*, *LinkedIn* and *Google Business* ranking as the most expensive services. It will be interesting to study how the advancements in text-generating AI [30] will potentially reduce the need of human intervention and how this may impact these services.

- **Guarantee / No Guarantee / Refund / No Refund**. Many services are advertised as `Guaranteed`. This guaranty may be a refund of the payment or a replacement of the service in case it is not delivered. In general each panel has a guarantee and refund policy specified in its terms of service.

- **Slow / Fast / Instant**. It is common to find the speed and start time of the services advertised in their description. In some panels there is a field that reports the estimated delivery time of the service based on previous deliveries.
- **Male / Female**. This is the main demographic targeting we found aside from language and country of origin. However the possibility of choosing male or female accounts was only found in *Instagram* services, *YouTube comments* and some other review services.

4.2.1. Geo-targeting

In addition to the discussed variations, we analyze location-based targeting in the offered services (i.e., geo-targeted services). The majority of them offer the possibility to select a specific country from which the service will be delivered (e.g., source IP of web traffic, or the country of a registered account posting a comment in YouTube). For text-based services – e.g., comments and reviews –, most providers allow customers to select the language.

Table 3 presents the top 10 services ranked by the number of their location-based variations that allow for geo-targeting. We observe that more than 70% of *Website traffic* and *YouTube comments* services are geo-targeted, and more than 50% for *Spotify plays* and *YouTube ads views*. In general, we observe that services oriented to YouTube and Spotify have higher percentage of geo-targeted services than those of *Instagram*, where only 19% of likes or 19.55% of followers, allow for geo-targeting.

Regarding the location from which the services are offered we identified more than 60 different countries and regions. The most prominent ones by number of services are: USA, Brazil, India, UK and Russia. Note that we focus our study to English panels, and extending our dataset with Chinese, Russian and Spanish panels might have an impact in these results. Figure 3 depicts the platforms targeting for the services located in each of these 5 countries. These distributions percentages are presented with more detail in Table 4.

Interestingly, there are significant differences in the distribution of targeted platforms across countries. By observing the percentage of services under the *Others* label we realize that services USA and UK are much more evenly spread in terms of platforms when compared to the other 3. Brazil shows the complete opposite effect having more than 60% of all its services targeting *Instagram*.

5. Price analysis

This section analyzes the prices of the services advertised in SMM panels. We first provide a catalog containing the range of prices found for the most popular services. Then, we analyze the prices variability during the period of the study and contrast it to their variability within and across panels. This analysis is extended further with a review of the cheapest and most expensive services we encountered. We conclude the section by studying the higher cost of the providers offering customized services (See Section 4.2).

Site	Product	Total	Geo-targeted	%
Website	traffic	7,066	5,447	**77.09**
YouTube	view	7,836	3,449	44.01
Instagram	like	8,362	1,589	19.00
Instagram	follower	7,390	1,445	19.55
Spotify	play	1,639	950	**57.96**
Instagram	comment	1,622	622	38.35
YouTube	comment	822	597	**72.63**
YouTube	ads view	940	522	**55.53**
YouTube	share	1,165	421	36.14
Spotify	follower	811	377	46.49

Table 3: Presence of geo-targeted services.

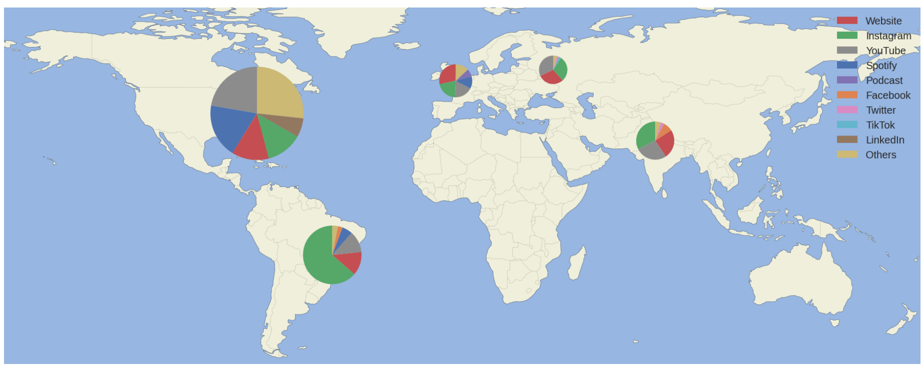

Figure 3: Target platform distribution for the top 5 countries of origin of geo-targeted services: USA, Brazil, India, Russia and UK. The pie charts are scaled by the number of services offered in each country.

The first step in order to get useful metrics on the services' price was to carefully filter the data to remove potential errors and outliers. However, it is difficult to distinguish whether the high price is due to exorbitant sellers or errors, or if the service is just extremely expensive (e.g., those between $\$10^5$ and $\$10^7$, which corresponds to \$100 and \$10k per unit). We filter services that match any of the following conditions: (i) Its price is higher than $\$10^7$. This is the equivalent of a \$10k price per unit; (ii) it has been active less than 10 days; (iii) its minimum order is above than 5 and its maximum order is above 100; and (iv) it is classified as a package or bundle. The last two conditions serve the purpose of eliminating services that may be valid but are not relevant for studying the price of bulk services. Therefore, we try to eliminate highly specialized services or packages such as *Instagram Followers* from verified accounts, or a service consisting of daily comments and views for an specific YouTube channel during 3 months. Aggregating these services with regular *Instagram followers* or *YouTube views* would provide a distorted view of the market.

	USA	Brazil	India	UK	Russia
Nº Targeted Services	2,843	1,785	1,160	1,006	865
YouTube	22.30	11.93	26.98	17.69	32.14
Spotify	18.85	5.83	-	12.13	-
Website	12.94	12.94	24.74	28.23	29.60
Instagram	12.63	63.64	32.50	21.37	29.13
LinkedIn	6.58	-	-	-	-
Facebook	-	2.58	8.10	-	-
Twitter	-	-	2.84	-	2.31
TikTok	-	-	-	-	2.43
Podcast	-	-	-	7.46	-
Others	26.70	3.08	4.83	13.12	4.39

Table 4: Platform targeting distribution for the top 5 countries of origin of geo-targeted services.

Table 5 reports the prices for the top 20 most popular services. We note that we use the median and quartiles instead of mean and standard deviation as, despite our filtering efforts, there are outliers that demand the use of robust statistical metrics. This choice may explain why the results presented in this paper differ substantially from mean values reported in previous research studies. However, it could also be the case that these services have gotten cheaper in the last years as the market and the diversity of services grew.

5.1. Price stability

This section studies the fluctuation in the different services' prices to understand the overall market volatility. Unfortunately, measuring these aspects is not trivial due to the fact that, instead of modifying the price in a service, most services are typically taken down or just re-branded.

Yet, out of the 61k observed services, 88.6% of them remained stable and never changed the price of their offerings. We test if these stable services had a shorter lifespan than those that did change. We used a Welch's t-test on the duration in days of fixed prices services and non-fixed price services. We obtained t-value$= -25.58$ and confirmed the difference in duration was relevant. Having the mean duration of fixed price services, 45.3 days, being significantly shorter than that of the non-fixed services, 57.9 days. Although there are some differences between the percentage of fixed-price services across panels, they are not significant so we conclude that this is a general practice in the market.

With this limitation we study the price evolution of aggregated services such as *Instagram followers* and *YouTube views* instead of individual variations. As in Table 5, we first filter out outliers according to the criteria previously described. Then, we evaluate the price of the most popular services during the first 10 days of April, May, June and July, so we measure local variances over four one-month

Site	Product	count	min	Q1	median	Q3
Website	traffic	5,340	0.09	0.36	0.39	0.60
Instagram	like	4,732	0.06	0.80	1.43	2.88
YouTube	view	3,729	0.25	1.35	2.10	3.00
Instagram	follower	4,057	0.08	2.20	4.62	8.50
Instagram	view	1,085	0.00	0.03	0.08	0.25
Spotify	play	1,373	0.33	1.20	2.10	3.77
Instagram	comment	1,058	0.24	8.76	25.00	60.00
YouTube	like	769	0.60	5.00	9.00	13.56
YouTube	share	766	0.40	1.35	1.89	2.10
Spotify	follower	692	0.24	1.50	2.55	4.50
Facebook	like	682	0.40	2.23	4.32	7.70
YouTube	comment	669	0.78	29.50	42.00	62.26
Facebook	page like	592	1.20	5.03	9.79	17.00
YouTube	ads view	588	0.72	2.20	2.50	3.20
Instagram	impression	284	0.02	0.10	0.17	0.45
Facebook	video view	377	0.05	0.13	0.22	0.70
Instagram	story view	550	0.01	0.07	0.18	0.39
Twitch	view	419	0.24	1.05	1.60	1.93
YouTube	subscriber	515	1.30	10.80	16.52	22.80
Facebook	view	653	0.05	28.80	65.00	120.00

Table 5: Prices of top 20 most advertised services.

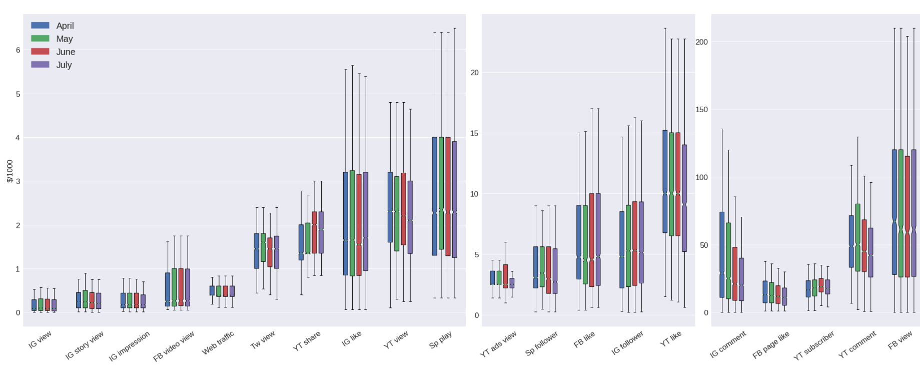

Figure 4: Boxplot of the most of the top-20 most popular services during 4 months. The whiskers extend to the higher/lower observation within one Interquartile range (IQR) of the quartiles. *We note that, due to the variability in the amount of services, the graph has been divided to use different scales and enhance the visibility of the graph.*

periods. The results are presented in the boxplots rendered in Figure 4, being the x-axis a service, with colors representing each month studied.

In this figure we can observe that the price distribution did not change significantly during these 4 months. *Instagram comments* is the service that exhibited the greatest change, starting at \$29 in April and steadily decreasing down to \$20 by July. However, even in this service the variation across time is minimal in comparison to the price deviation within and across panels in the market. In fact the most important observation is the wide range of prices. This is particularly clear for the services in the right side of the figure. For example, we can see that for *YouTube comments* in May we had simultaneously prices below \$5 and above \$130. *Facebook views* present the same effect, having prices below a dollar and above \$200. A potential reason for such a variance is due to the presence of service customization, since it is expected that geo-targeted custom comments for YouTube are more expensive than random non geo-targeted ones. However this is not the only reason for the high variance. These panels are mainly reselling platforms and they often take services from other panels and resell them at higher prices. As a result we often find the exact same service offered at very different prices across several panels. In the next section we illustrate this issue with a particular service: *Google Business reviews*.

5.2. High-end services

In Section 4 we analyse the most common services being offered (by prevalence of listings). In this section, we investigate what the most expensive services are. As explained above, the presence of outliers requires taking a semi-manual approach in this type of analysis. We rank the median price of the services to have an initial list of the most expensive services and then we manually review them. Fortunately such expensive services are uncommon, thus manual inspection is feasible.

We group the most expensive services in three categories: (i) review and rating services; (ii) accounts for subscription services and non-free interactions; and (iii) premium services, which are high end or enhanced versions of typically cheap services. We present the findings of each of this categories separately.

5.2.1. Review and rating services

These services normally consist of a written review plus some sort of rating. The services are advertised as custom reviews and most of them require the client to send the desired text to be submitted. We found this kind of services for Google Business, LinkedIn, TripAdvisor, IMDb and for the PlayStore and Apple Store. For these last two, the services offered app install plus review and rating.

Google Business reviews are found in 21 panels whose price range from \$1.4 and \$18.0 per review, and the available orders go from 1 to 100 reviews (although one service offered up to 1.5k reviews). When trying to identify which variables determine the difference in price, we discover that several services offer

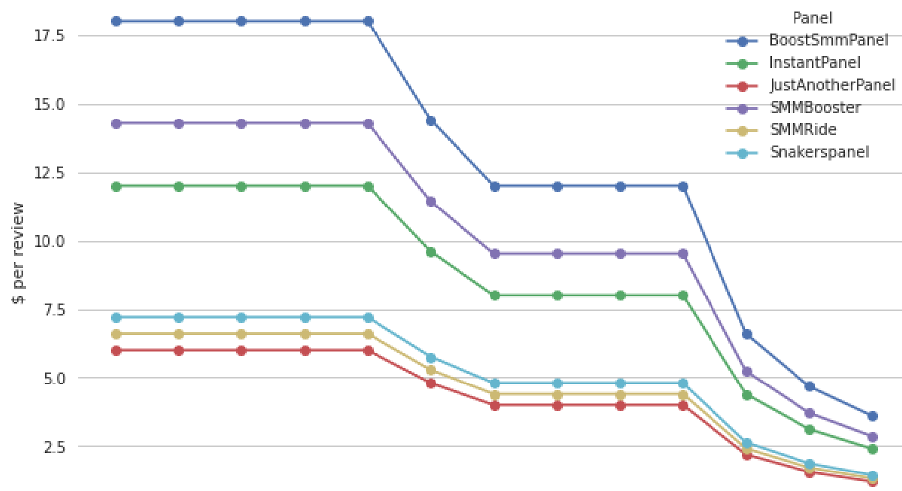

Figure 5: *Google Business review* prices for 13 exactly equal services (points) present in 6 different panels (lines). Y-axis represents the prices of the different services. Thus, points aligned vertically represent the prices of a same service across the 6 panels.

the exact same name and description in 6 different panels but at significantly different prices. Figure 5 shows the difference of price for a given service across panels. The shape represents that these services either have an underlying common services or are a reseller of one another. It is also worth noting that the difference in price across panels (vertically) is greater than the difference within a panel due to service variations (horizontally). This fact makes difficult the study of the services' customizations on the price because, as we see, aggregating prices from different panels may obscure the impact of these customizations.

TripAdvisor reviews are much less common, appearing only in 2 panels. In one of the panels the service description does not have much information besides stating it provided a custom review and its price oscillates between \$0.25 and \$0.72 per review. In the other panel, the service has a more detailed name and description that advertised a custom review plus a rating delivered from real accounts, their country of origin, the delivery rate, etc. It also grants a refill policy in case the reviews and rating are taken down. This service is priced at \$5 per review.

Interestingly, we find that **IMDb Votes** (which would be the equivalent of a rating in other platforms) are quite expensive. Prices range from \$15 to \$20 per 1000 in the low end, and from \$80 up to \$150 in the high end, where the country and gender of the voter could be selected.

5.2.2. Accounts for subscription services and non-free interactions

Services that are not free when acquired legitimately are unsurprisingly more expensive than free ones. Netflix accounts, Amazon Prime subscriptions and other services of this category are sold by a fraction of their legitimate price. These services typically offer the possibility of purchasing individual accounts but we also found services that offered bulk packages and were specifically advertised for resellers. Some of these services directly state that the accounts for sale are hacked accounts.

A particularly interesting case within this category is the service **Twitch Subscribers**. Twitch is a video streaming platform in which users can follow a streamer's channel for free or subscribe to it for a fee (typically around \$5 [8]) in order to support the streamer. Additionally users with an Amazon Prime account can subscribe to one channel for free. These are called Prime Subscriptions and despite being free for the user, the streamer still gets a portion of its value. We have found several services of such subscriptions with prices ranging from \$1.5 to \$3. A streamer that buys these services would not only boost his account, but could also potentially make money just from the service alone. The percentage of the subscription fee received by streamers is not fixed. For example, if he gets 40% of the fees, for \$5 worth subscription he would be getting \$2 per subscription. By purchasing subscription at a price below \$2 he/she would be on profit.

5.2.3. Premium services

It is common to find improved versions of popular services that retail at a much higher price. A very illustrative example for this are *Instagram likes* and *Instagram reach boost packages.* In Table 5 we can see that 1,000 *Instagram likes* usually go for around \$1.43. However we contrast this with a bundle that offers 1,000 *Power likes* and 30 posts for \$275. The service advertised next offers the same influencer *Power likes* with the possibility of choosing a quantity (from 500 to 5,000) at \$0.25 each, which is 17.5 times more expensive than the median price for a like. *Power likes* are a controversial topic. Basically they are likes provided from popular accounts– verified accounts, celebrities or influencer accounts, or accounts with a certain number of followers and reach–which allegedly have a big influence in Instagram's recommendation algorithm. Therefore they are supposed to increase an account's reach very effectively, which justify their price.

5.3. Low-end services

In contrast to the services described above, we now briefly review the cheapest services found in these panels. Indeed, we find various services are offered for free in order for clients to check the quality of the service and persuade them to make an investment by buying other services. These phenomena are also observed in trading from underground forums (see Section 6.2. Out of the 61k different services 75 where free services and they mostly consisted of *Instagram likes, comments and followers*; *TikTok views* and *YouTube views.* They offered a low quantity of views and a followers only for a limited amount of time.

[8]https://www.twitch.tv/p/partners/

The cheapest non-free services are video views and plays. For less than \$0.10 per 1000 we found: *TikTok views* between \$0.01 and \$0.03, *SoundCloud plays* for around \$0.05 and some *Instagram/IGTV views* between \$0.05 and \$0.08. Between \$0.10 and \$0.30 we found *Instagram story views* and *Facebook video views.* Above this price and up to \$0.60 we mainly find *web traffic.* The common factor across these services that could explain why they are so inexpensive is how easy they are to provide.

5.4. Customizations price impact

We conclude the prices analysis by studying the price variations of several services in function of the customizations identified in section 4.2. In order to do this we have chosen 3 customizations and we have analyzed them over 2 services where they were relevant. In addition, all 3 have been analyzed along side the geo-targeting variable. The reason for this is that geo-targeting is the most impactful variation price-wise and therefore it would cover up the impact of other variables in the price if it were not analyzed simultaneously.

In all 3 cases we have taken a similar approach. We have first applied some filtering to the data in order to remove outliers and erroneous data. The filtering used is the same as described at the beginning of this section but eliminating also the services with a price exceeding 5 times the third quartile of the prices (see Table 5). This extra condition was added in order to generate more useful visualizations. Then, we selected the locations where each service variation was most popular. Lastly, we generated graphs for each of this location depicting the distribution of prices for services with and without the customization. The results are presented in Figures 6, 7 and 8. In this figures each violin represents the prices for the service from different areas, with non geo-targeted services being placed in the leftmost side of the figure labeled as *Unspecified.* In each violin, the left side represents the price distribution of the services without the customization and the right side represents the price distribution of the services with the customization. In both sides the distribution has been cut at highest and lowest observations in the data.

YouTube and Instagram custom comments. The first case we studied was custom comments for YouTube and Instagram. The results obtained are presented in Figure 6. Here we observe that in both platforms there is a substantial difference in price between geo-targeted and non geo-targeted services. The most expensive locations are the US for YouTube and China for Instagram with a difference of \$41 (+256%) and \$45 (+300%) respectively compared to the non geo-targeted versions.

When comparing the prices for the *Custom* variation we don't see a clear difference in average prices. However, we see a different shape in the distributions, with custom comments tending to have a more spread out shape, with a heavier tail towards high prices. We would observe a more significant difference if we compared custom comments with those explicitly advertised as *Random.* However, as with many other negative keywords, we found much fewer of these services and therefore we did not have enough data to derive solid conclusions.

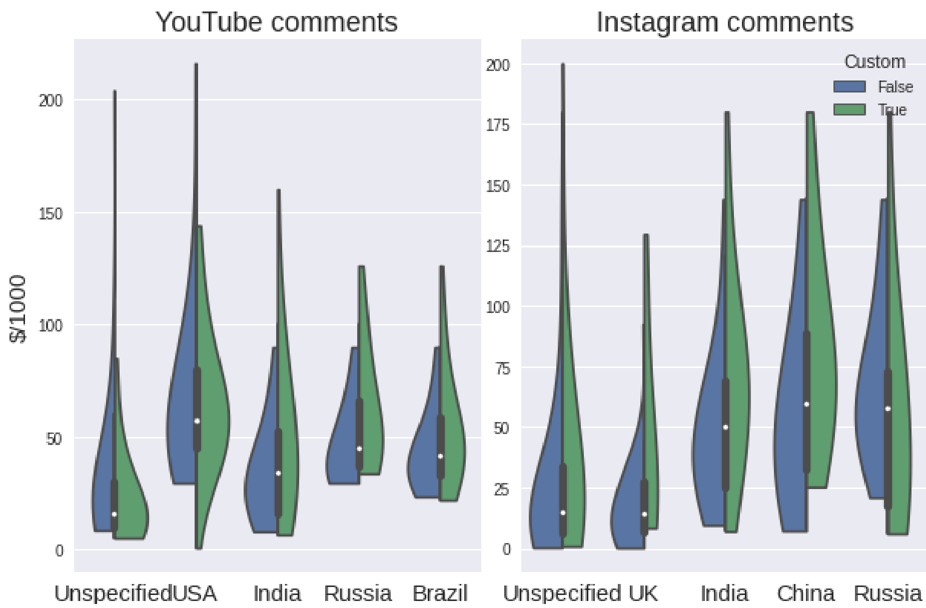

Figure 6: Prices distribution of YouTube comments and Instagram comments attending to geo-targeting and the *Custom* variation. Each individual violin plot represents the prices for a given region (horizontal axis). The left, blue side represents the distribution for non-custom comments, while the green, right side of the violin represents the distribution for the custom comments.

Instagram gender targeted followers and comments. In this case our objective was to study if gender targeting is a significant phenomenon and if so, estimate how relevant it is. We found that the platform where it is most common is Instagram, and particularly, in the followers and comments services. This customization however was not usually offered with geo-targeting with the only exception being Brazil. The results are presented in Figure 7. We can observe, as in the previous case, a significant price difference due to geo-targeting. The differences are of \$8 (+228%) and \$31 (+207%) for followers and customers respectively. In regards to gender targeting we did not focus on the specific gender and we grouped together the services that offered specifically male or female followers/comments. In this case we can clearly see a shift in the prices distribution of gender-targeted services. The price range for these services starts at a higher point and we see the median of the distribution also being notably higher.

Spotify and SoundCloud plays with refill policy. For the last case we selected plays in the 2 most popular audio and music platforms: Spotify and SoundCloud. SoundCloud despite being the second most popular has much fewer services than Spotify and the only geo-targeted SoundCloud plays we found were offered from the US. In contrast, Spotify plays are offered from 30 different countries from which he have selected the top 6. The variation studied

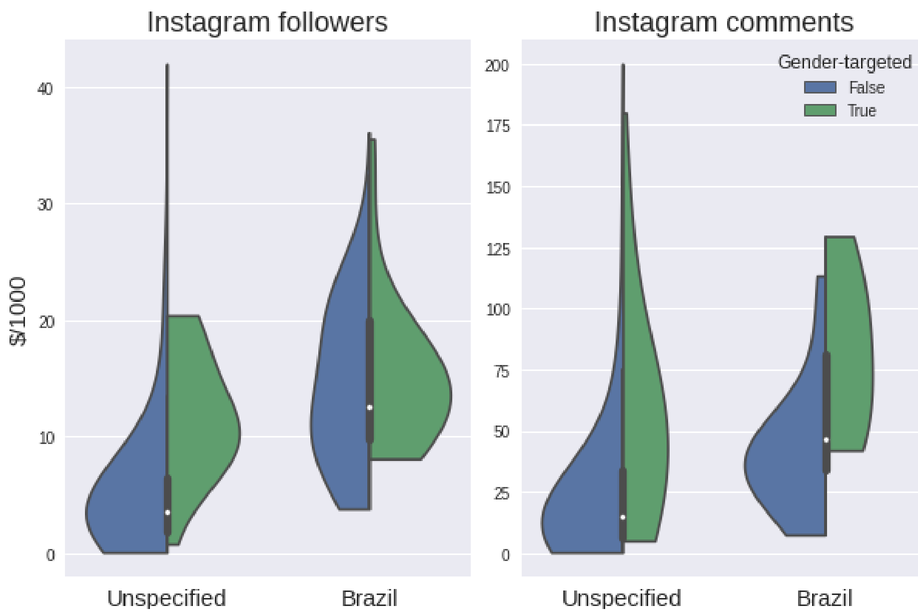

Figure 7: Prices distribution of Instagram followers and comments attending to geo-targeting and gender-targeting. Each individual violin plot represents the prices for a given region (horizontal axis). The left, blue side represents the distribution for non gender-targeted services, while the green, right side of the violin represents the distribution for specifically male or female services.

in this case was the *Refill* option, which is very common in many services. The results are presented in Figure 8 and just as in the two previous cases we observe a significant difference in prices due to geo-targeting. The most expensive location for Spotify plays is Germany, although it has very similar prices to France, UK, Canada and Brazil. The difference in prices between the most expensive locations and non geo-targeted services are \$2.6 (+195%) and \$0.05 (+100%) for Spotify and SoundCloud respectively. In relation to the *Refill* customization we can observe mixed results. We see that refillable services start at a higher price for SoundCloud plays and geo-targeted Spotify plays. We also observe that refillable services prices reach higher prices for Spotify plays from France, UK, Canada and Brazil but not the rest of locations nor SoundCloud plays. Therefore we see an upwards shift in the prices distribution for Spotify plays from these 5 countries but we observe the opposite trend in SoundCloud plays from the US, where all refillable plays prices are below the median. In general these results suggest slightly higher prices for refillable services although not in all cases. In order to draw more solid conclusions it would be necessary to analyze other variables that are closely related such as drop rates, speed of delivery and refill periods.

The results obtained in these 3 cases illustrate the impact of geo-targeting in the prices. It is clear to see how location or language customization is de-

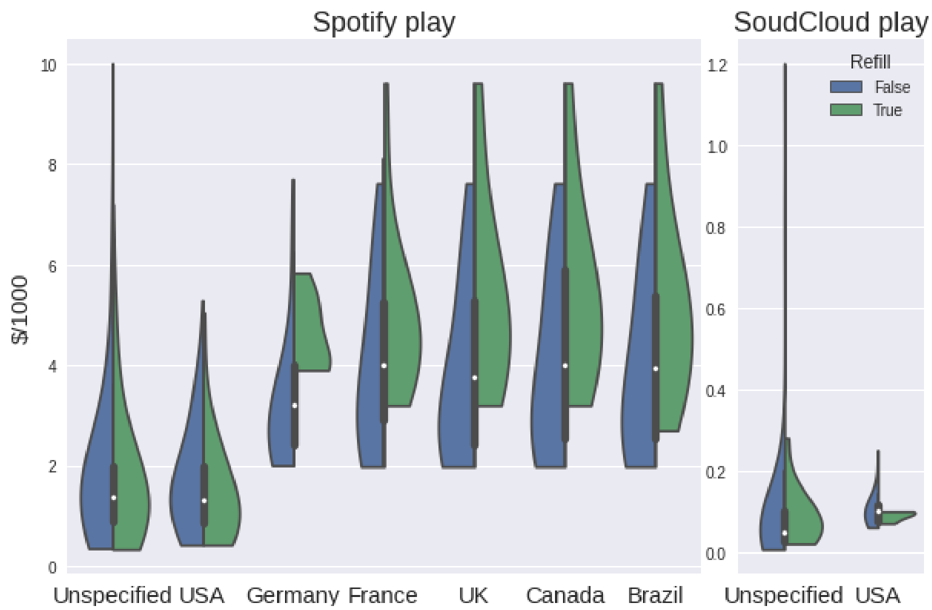

Figure 8: Prices distribution of Spotify and SoundCloud plays attending to geo-targeting and the *Refill* variation. Each individual violin plot represents the prices for a given region (horizontal axis). The left, blue side represents the distribution for services that are not advertised with *Refill*, while the green, right side of the violin represents the distribution for those which are.

sirable or even necessary in services such as product reviews. In others, such as YouTube views or Website traffic where the buyers objective may be to obtain benefits from advertisement fraud, selecting an adequate location may be very beneficial as advertisers often pay different rates for each country. Gender targeting also seems to have a substantial effect in price, although it is not as relevant as it is a much less widespread customization. This kind of targeting is probably useful in platforms like Instagram if the aim of the buyer is to influence the recommendation algorithm and direct its content reach towards certain communities. In general the effectiveness of these methods for achieving the customers goals make these very desirable services. This, together with the added difficulty for supplying targeted services, are probably the factors that drive their prices up.

6. Marketing and operation

In previous sections, we have discussed the market and economics of fake SMM panels as the actual storefronts where the services are being offered. In this section, we complement our study by analyzing two aspects related to the supply chain needed to promote and operate SMM product and services. First, we study the domain names used by the panels. Our goal is to understand how the security industry classify these sites and how are they ranked in terms of popularity (i.e., by traffic received). Second, we analyze the trading and discussions on this topic found in online underground forums, which gives us as unique opportunity to understand the ecosystem of the actors involved in this business.

6.1. Domain analysis

Domain classification. We now explore questions related to the domains registered to offer the SMM panels. We first run all domains through four domain classification services: Symantec, McAfee, Fortiguard, and OpenDNS. Our results suggest that these panels are largely seen as regular IT, business or marketing service providers. For example, McAffe classifies 45% of the domains as "Internet Services," 26% "Marketing/Merchandising," and 20% as "Media Sharing." Similarly, Fortiguard classifies 57% of the domains as "Information Technology" and 27% as "Business."

Very few domains are spotted as related to malicious or suspicious activity. Symantec classifies 11% of the domains as "Suspicious," 2% as "Phishing," and 1% as "Malicious Sources/Malnets." The results for McAfee (2% as "Malicious Sites" and 1% as "PUPs (potentially unwanted programs)") and Fortiguard (1% as "Malicious Websites") are similar. We also run the domains through VirusTotal. The majority of the 75 detection engines (which mostly operate based on blocklists) spotted all domains as "harmless." In only 2 cases, 1 out of the 75 engines flags a domain as suspicious." These results confirm that the security industry do not consider these sites as incurring in any potentially harmful activity.

Domain popularity. We also analyzed the popularity of these domains using Alexa traffic ranks [2]. We compute the daily position for all the panels and then analyze the resulting time series. The median of the time series is 378,509, with a minimum of 21,062 and a maximum of 802,910. Sites ranked beyond 100,000 are generally deemed as statistically insignificant due to the scarcity of data available for them. Such low ranking positions suggest that these sites sustain a very reduced amount of traffic.

We tried to identify common actors by analyzing the domain registration information through the WHOIS service. However in many cases Third Party Service (TPS) are used to hide sensitive information (phones, emails, ...). In others, the registrar itself omits this information to protect users' privacy. We also analyzed the HTTPS certificates of the panels but we could not find any matches or information that could help with attribution.

6.2. Trade in underground forums

Products. Underground forums are a common place used for the trading of various illicit products and services [21, 29, 28, 25]. In order to understand the underground economy of SMM, we analyze a set of 7,063 forum threads providing SMM services (see Section 3). The titles of the threads are not as structured as in the case of SMM panels. This hinders a proper categorization

Instagram		Twitter		YouTube		Facebook		Snapchat	
Like	552	Follower	407	View	303	Like	231	Shoutout	81
Follower	436	Like	84	Like	230	Account	52	View	35
Shoutout	178	Account	45	Subscriber	80	Page like	43	Account	8
Account	133	Retweet	45	Comment	52	Follower	32	Follower	6
Unknown	803	Unknown	633	Unknown	554	Unknown	477	Unknown	97
Others (15)	88	Others (17)	89	Others (15)	82	Others (19)	91	Others (5)	9
Total	2190	Total	1303	Total	1301	Total	926	Total	236

Table 6: Top 6 platforms and their products being advertised in underground forums

of the services being offered. Also, often the same thread offers various services for various platforms. Yet, to get and overview of the type of service offered and the platform involved, we conduct a best-effort analysis looking for specific keywords in thread headings. In total, we observe 6,708 services from 37 platforms. Table 6 shows the most common services from the top 5 platforms traded in these forums. It can be observed that the top services align with the findings from SMM Panels (c.f. Table 2), with three notable differences. First, website traffic is not as popular in forums as it is in SMM panels. Second, Snapchat is the fifth platform traded in forums. Third, a common service being offered in forums is *shoutouts* (i.e., the promotion of a user account from another account, usually popular in terms of followers, in the form of a mention or a photo). We observe that these products are highly used for sourcing traffic intended for other activities, like eWhoring [15] or Cost-Per-Action (CPA) services.

Actors. In the analyzed threads, we identify a set of key actors[9] by means of number of SMM offerings, and conduct both a quantitative and qualitative analysis of them. Most of the threads in our dataset (82.7%) are from one particular forum: *HackForums*. We investigate the actors from this forum offering social media marketing services. From the total of 3,235 actors, we filter those that have not received any response, since they are probably low-impact services or even scams. This yields a total of 2,670 actors. From these, we focus on the top 751 actors that have at least 2 offerings, which account for 58.3% of the threads.

For the quantitative analysis, we follow the methodology from previous works [25, 26]. We extract the forum activity of the actors, including the number of posts, the reputation, the operating days (i.e., time between first and last post in the forum), the interests of the users and their social relations. Such social relations are obtained from forum interactions between the actors, i.e., responses in a thread or posts quoted. This allows us to build a social graph where each node is a forum actor and the edges represent their interactions. We compute popularity metrics such as the eigenvector of the graph or the H-index (a widely used metric used in academia to measure the popularity of researchers). For the interests of the actors, we analyze the number of posts

[9] In our work, we consider as an *actor* an account id posting messages in the forum. Thus, we do not consider cases where the same user operates various accounts.

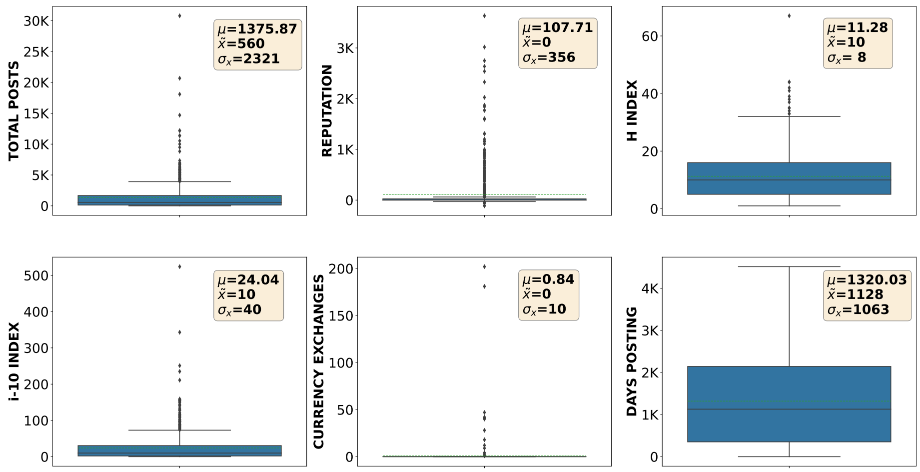

Figure 9: Forum activity for the top 50 users offering social media marketing services

and threads made in boards from the different categories, as provided by the forum, such as hacking, gaming, technology or market.

Figure 9 shows the aggregated statistics related to the forum activity for key actors. We found a wide diversity of actors according to their activity. For example, whereas the average number of posts is around 2.3k, some actors have posted more than 30k posts, while other only have a few dozens. We observe similar patterns in the reputation, with substantial differences between those highly reputed (with more than 1k positive points) and the majority which has no reputation at all (the median is 0). These authors have on average a H-index of 10, ranking similar to key actors initiating in cybercrime activities according to a previous study [25]. However, in general most SMM actors do not use the forum for currency exchange (average lower than 1), which is a board used to exchange and launder financial gains obtained from illicit activities [28].

Hackforums is divided in different categories, e.g. market, hacking or gaming. In order to characterize the interests of the authors, we analyze their activity in boards from these categories. We count the number of threads started and replies posted in these boards. Table 7 shows the most prevalent categories ranked in the first, second, and third position of the actors' interests. As expected, $\tilde{7}$5% of the actors have as primary interest the Market (with $\tilde{2}$0% having it in their second position), which includes boards specific for Social Media Marketing trading and also other marketplace related discussions. We observe that authors are also interested in the Common category, which includes general-purpose discussion about the forums, as well as miscellaneous topics (e.g., politics or religion). The most common category ranked in the third position is Hack, where actual discussion about computer and network hacking occurs. This suggests that, while the primary reason for the actors in the forum is to trade SMM, they show strong interests in other cybercrime-related topics.

1st category	2nd category	3rd category
Market (74.97)	Common (49.1)	Hack (24.35)
Common (15.85)	Market (19.31)	Gaming (18.41)
Web (2.4)	Hack (10.9)	Money (16.96)
Hack (2.26)	Money (7.31)	Common (15.65)
Money (2.13)	Gaming (5.66)	Web (8.84)
Gaming (1.6)	Web (4.41)	Graphics (4.2)
Graphics (0.4)	Coding (1.38)	Market (3.91)
Coding (0.27)	Graphics (1.38)	Tech (3.33)
Unknown (0.13)	Tech (0.28)	Unknown (2.32)
	Unknown (0.28)	Coding (2.03)

Table 7: Interests of the 751 actors offering SMM services in Hackforums

Top-10 actors. We conduct a qualitative analysis for the set of top 10 authors by number of offerings in our dataset. We analyze the threads started by these users in the forum, including those related with SMM but also other topics. By manual analysis, we observe a common pattern in most of the authors: they tend to start by providing a few amount of cheap or easy-to-get services (e.g., Instagram *shoutouts*, Youtube views or FB likes). This allows them to increase their reputation and to get known to the community, and it is a common practice in underground forums where trust is a valuable active that must be gained across time [10]. After a initial promoting period (typically a few weeks), they start trading the services. We also observe various activities potentially related to required components of the supply chain. These include SMM panel designs, Twitter bots, Instagram accounts with several followers (possibly used for selling shutouts) and automatic Youtube account makers. Interestingly, we found 2 cases where actors have built more than one SMM panels over time, only for selling them as quality products afterwards. This confirms that re-selling these products is a common practice. Finally, in parallel to their activity in SMM products, actors also operate other illicit businesses. For example, 2 of the 10 users analyzed provide and sell services related to eWhoring [15, 26], while other 3 actors provide accounts related to video games.

7. Concluding remarks

In this paper, we presented a study on the market providing fake engagement services for social media. We have compiled a dataset of offered services by crawling daily the SMM panels where they are advertised during a period of 4 months. This dataset consists of 2.8M entries grouped in 61k different service variations. Using this dataset, we have identified 294 different services targeting 59 platforms including OSN, review services, video and music platforms, etc.
Service customizations. We observe that most of these services are offered with an impressive variety of customizations that allow buyers to select features such as the quality of the service, the speed of delivery, the country of origin, as

well as personal attributes of the fake account (e.g., gender). The granularity of these customizations and the richness of the catalog hint at the existence of a substantial infrastructure underlying these services.

Market analysis. The prices we observe for these services are significantly lower than those reported in previous studies. For example De Cristofaro et al. [8] report Facebook page likes for prices between $14.99 - $70 while we observed a range of $5.03 to $17. Similarly for Instagram likes we observe prices between $0.80 - $2.88 in contrast to the average of $19.54 reported by Paquet-Clouston et al. [24]. These differences might be a result of our methodology, in particular of our decision to filter out high price outliers. However it can also be indicative of a descending trend in the prices during the period between the studies.

The price analysis revealed very significant disparities between prices of the same product across different markets. This price differences is likely a consequence of the multiple resellers present in the supply chain [23]. As Paquet-Clouston et al. [24] point out, this can also indicate that the market is still undeveloped and sellers do not know the worth of the services they offer, leading them to underprice or overprice. There is also significant variance in prices within markets but this can be mostly attributed to the different available versions of the services. In particular, geo-targeting and gender targeting (i.e., followers of a specific gender) resulted in a substantial increase of the prices. However geo-targeting is much more common, being available for almost all services while gender targeting was present only in a few.

Trading in underground forums. Underground forums are nowadays a key component where panels are advertised. They allow newcomers to enter into the business by providing tutorial and guidance. Also, while not originally designed as markets, various threads are intended for the trading of fake engagement services. Our study confirms that the platforms and products being offered coincide with those offered in dedicated panels. Also, we analyze the ecosystem of the actors involved. We observe that some actors that initially provide free services to gain reputation before engaging in trading. Also, we note that actors are not always specialized in the trading of SMM, but this is combined with other lucrative illicit businesses.

References

[1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[2] Amazon.com. Alexa - top sites. [Online] Last Accessed 16 December 2021.

[3] Alessandro Bessi and Emilio Ferrara. Social bots distort the 2016 u.s. presidential election online discussion. *First Monday*, 21, 11 2016.

[4] Alex Beutel, Wanhong Xu, Venkatesan Guruswami, Christopher Palow, and Christos Faloutsos. Copycatch: Stopping group attacks by spotting lockstep behavior in social networks. pages 119–130, 05 2013.

[5] Yazan Boshmaf, Dionysios Logothetis, Georgos Siganos, Jorge Lería, Jose Lorenzo, Matei Ripeanu, Konstantin Beznosov, and Hassan Halawa. Íntegro: Leveraging victim prediction for robust fake account detection in large scale osns. *Computers & Security*, 61:142–168, 2016.

[6] Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov, and Matei Ripeanu. Design and analysis of a social botnet. *Computer Networks*, 57(2):556 – 578, 2013. Botnet Activity: Analysis, Detection and Shutdown.

[7] Giuseppe Cascavilla, Damian A. Tamburri, and Willem-Jan Van Den Heuvel. Cybercrime threat intelligence: A systematic multi-vocal literature review. *Computers & Security*, 105:102258, 2021.

[8] Emiliano De Cristofaro, Arik Friedman, Guillaume Jourjon, Mohamed Ali Kaafar, and MZ Shafiq. Paying for likes? understanding facebook like fraud using honeypots. 11 2014.

[9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2019.

[10] Benoît Dupont, Anne-Marie Côté, Claire Savine, and David Décary-Hétu. The ecology of trust among hackers. *Global Crime*, 17(2):129–151, 2016.

[11] Muhammad Junaid Farooq and Quanyan Zhu. Modeling, analysis, and mitigation of dynamic botnet formation in wireless iot networks. *IEEE Transactions on Information Forensics and Security*, 14(9):2412–2426, 2019.

[12] Shehroze Farooqi, Álvaro Feal, Tobias Lauinger, Damon McCoy, Zubair Shafiq, and Narseo Vallina-Rodriguez. Understanding incentivized mobile app installs on google play store. In *Proceedings of the ACM internet measurement conference*, pages 696–709, 2020.

[13] Qiang Fu, Bo Feng, Dong Guo, and Qiang Li. Combating the evolving spammers in online social networks. *Computers & Security*, 72:60–73, 2018.

[14] Hongyu Gao, Jun Hu, Christo Wilson, Zhichun Li, Yan Chen, and Ben Y. Zhao. Detecting and characterizing social spam campaigns. In *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, IMC '10, page 35–47, New York, NY, USA, 2010. Association for Computing Machinery.

[15] Alice Hutchings and Sergio Pastrana. Understanding ewhoring. In *2019 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 201–214. IEEE, 2019.

[16] Meng Jiang, Peng Cui, Alex Beutel, Christos Faloutsos, and Shiqiang Yang. Inferring strange behavior from connectivity pattern in social networks. volume 8443, 05 2014.

[17] G. Kambourakis, C. Kolias, and A. Stavrou. The mirai botnet and the iot zombie armies. In *MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*, pages 267–272, 2017.

[18] Yixuan Li, Oscar Martinez, Xing Chen, Yi Li, and John E. Hopcroft. In a world that counts: Clustering and detecting fake social engagement at scale. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, page 111–120, Republic and Canton of Geneva, CHE, 2016. International World Wide Web Conferences Steering Committee.

[19] Chavie Lieber. The dirty business of buying instagram followers. *Retrieved May*, 31:2019, 2014.

[20] Ee-Peng Lim, Viet-An Nguyen, Nitin Jindal, Bing Liu, and Hady Wirawan Lauw. Detecting product review spammers using rating behaviors. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, page 939–948, New York, NY, USA, 2010. Association for Computing Machinery.

[21] Marti Motoyama, Damon McCoy, Kirill Levchenko, Stefan Savage, and Geoffrey M Voelker. An analysis of underground forums. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, pages 71–80, 2011.

[22] Arjun Mukherjee, Bing Liu, and Natalie Glance. Spotting fake reviewer groups in consumer reviews. In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12, page 191–200, New York, NY, USA, 2012. Association for Computing Machinery.

[23] Masarah Paquet-Clouston and Olivier Bilodeau. Uncovering the wholesale industry of social media fraud: From botnets to bulk reseller panels. 2018.

[24] Masarah Paquet-Clouston, David Décary-Hétu, Olivier Bilodeau, and Thomas Dupuy. When greed for fame benefits large-scale botnets. 2016.

[25] Sergio Pastrana, Alice Hutchings, Andrew Caines, and Paula Buttery. Characterizing eve: Analysing cybercrime actors in a large underground forum. In *International symposium on Research in Attacks, Intrusions, and Defenses*, pages 207–227. Springer, 2018.

[26] Sergio Pastrana, Alice Hutchings, Daniel R. Thomas, and Juan Tapiador. Measuring ewhoring. *Proceedings of the Internet Measurement Conference*, 2019.

[27] Sergio Pastrana and Guillermo Suarez-Tangil. A first look at the crypto-mining malware ecosystem: A decade of unrestricted wealth. In *Proceedings of the Internet Measurement Conference*, IMC '19, page 73–86, New York, NY, USA, 2019. Association for Computing Machinery.

[28] Sergio Pastrana, Daniel R. Thomas, Alice Hutchings, and Richard Clayton. Crimebb: Enabling cybercrime research on underground forums at scale. In *Proceedings of the 2018 World Wide Web Conference*, page 1845–1854, Republic and Canton of Geneva, CHE, 2018. International World Wide Web Conferences Steering Committee.

[29] Rebecca S. Portnoff, Sadia Afroz, Greg Durrett, Jonathan K. Kummerfeld, Taylor Berg-Kirkpatrick, Damon McCoy, Kirill Levchenko, and Vern Paxson. Tools for automated analysis of cybercriminal markets. In *Proceedings of 26th International World Wide Web conference*, 2017.

[30] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. *OpenAI Blog*, 1(8):9, 2019.

[31] Na Ruan, Ruoyu Deng, and Chunhua Su. Gadm: Manual fake review detection for o2o commercial platforms. *Computers & Security*, 88:101657, 2020.

[32] Salman Salamatian, Wasim Huleihel, Ahmad Beirami, Asaf Cohen, and Muriel Médard. Why botnets work: Distributed brute-force attacks need no synchronization. *IEEE Transactions on Information Forensics and Security*, 14(9):2288–2299, 2019.

[33] Indira Sen, Anupama Aggarwal, Shiven Mian, Siddharth Singh, Ponnurangam Kumaraguru, and Anwitaman Datta. Worth its weight in likes: Towards detecting fake likes on instagram. In *Proceedings of the 10th ACM Conference on Web Science*, WebSci '18, page 205–209, New York, NY, USA, 2018. Association for Computing Machinery.

[34] Filipo Sharevski, Raniem Alsaadi, Peter Jachim, and Emma Pieroni. Misinformation warnings: Twitter's soft moderation effects on covid-19 vaccine belief echoes. *Computers & Security*, 114:102577, 2022.

[35] Gianluca Stringhini, Manuel Egele, Christopher Kruegel, and Giovanni Vigna. Poultry markets: On the underground economy of twitter followers. In *Proceedings of the 2012 ACM Workshop on Workshop on Online Social Networks*, WOSN '12, page 1–6, New York, NY, USA, 2012. Association for Computing Machinery.

[36] A. H. Wang. Don't follow me: Spam detection in twitter. In *2010 International Conference on Security and Cryptography (SECRYPT)*, pages 1–10, 2010.

[37] Gang Wang, Tristan Konolige, Christo Wilson, Xiao Wang, Haitao Zheng, and Ben Y. Zhao. You are how you click: Clickstream analysis for sybil detection. In *22nd USENIX Security Symposium (USENIX Security 13)*, pages 241–256, Washington, D.C., August 2013. USENIX Association.

[38] Koosha Zarei, Reza Farahbakhsh, and Noel Crespi. How impersonators exploit instagram to generate fake engagement? 01 2020.