Computers & Security 124 (2023) 103013

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

An analysis of fake social media engagement services

David Nevado-Catalán[a], Sergio Pastrana[a], Narseo Vallina-Rodriguez[b], Juan Tapiador[a,*]

[a] Universidad Carlos III de Madrid, Spain
[b] IMDEA Networks, Spain

ARTICLE INFO

Article history:
Received 10 February 2022
Revised 20 October 2022
Accepted 13 November 2022
Available online 17 November 2022

Keywords:
Fake engagement services
Social networks
Fraud
Cybercrime
Security economics

ABSTRACT

Fake engagement services allow users of online social media and other web platforms to illegitimately increase their online reach and boost their perceived popularity. Driven by socio-economic and even political motivations, the demand for fake engagement services has increased in the last years, which has incentivized the rise of a vast underground market and support infrastructure. 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 algorithms to detect and remove fake activity in online targeted platforms. Yet, the platforms in which these services are sold (known as *panels*) and the underground markets offering these services have not received much research attention. To fill this knowledge gap, this paper studies Social Media Management (SMM) panels, i.e., reselling platforms—often found in underground forums—in which a large variety of fake engagement services are offered. By daily crawling 86 representative SMM panels for 4 months, we harvest a dataset with 2.8 M forum entries grouped into 61k different services. This dataset allows us to build a detailed catalog of the services for sale, the platforms they target, and to derive new insights on fake social engagement services and its market. We then perform an economic analysis of fake engagement services and their trading activities by automatically analyzing 7k threads in underground forums. Our analysis reveals a broad range of offered services and levels of customization, where buyers can acquire fake engagement services by selecting features such as the quality of the service, the speed of delivery, the country of origin, and even personal attributes of the fake account (e.g., gender). The price analysis also yields interesting empirical results, showing significant disparities between prices of the same product across different markets. These observations suggest that the market is still undeveloped and sellers do not know the real market value of the services that they offer, leading them to underprice or overprice their services.

1. Introduction

Online Social Networks (OSN) have become an effective marketing tool for businesses of all sizes and kinds as well as a popular platform for sharing information and news, either legitimate or fake. Many businesses, individuals and organizations have realized the potential of using online social media for increasing their online presence, economic benefits or influence. This has fueled the development and consolidation of a vast market that offers social media engagements such as Instagram followers, TikTok likes or Spotify plays. These services, which in many cases are delivered through illicit means (Cascavilla et al., 2021; De Cristofaro et al., 2014; Lieber, 2014; Paquet-Clouston et al., 2016), can be abused to manipulate the platform's recommendation algorithms or to boost a user's perceived popularity. They also enable cybercrime activities (Bhalerao et al., 2019).

Previous research on the topic has demonstrated that fake engagement services are a lucrative business involving multiple actors along its supply chain (Bhalerao et al., 2019; Stringhini et al., 2012). As in the case of other cyberthreats (Boshmaf et al., 2013; Farooq and Zhu, 2019; Kambourakis et al., 2017; Paquet-Clouston et al., 2016; Pastrana and Suarez-Tangil, 2019; Salamatian et al., 2019), botnets are leveraged as the main supplier of fake engagement services (Paquet-Clouston et al., 2016). However, before these services reach their final customer, they are often traded and resold in underground online platforms (Paquet-Clouston and Bilodeau, 2018). This phenomenon became evident with the pro-

* Corresponding author
E-mail addresses: 100421545@alumnos.uc3m.es (D. Nevado-Catalán), spastran@inf.uc3m.es (S. Pastrana), narseo.vallina@imdea.org (N. Vallina-Rodriguez), jestevez@inf.uc3m.es (J. Tapiador).

https://doi.org/10.1016/j.cose.2022.103013

liferation of SMM panels, which are essentially reselling 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. Work in this area has focused on the detection of fake activities (e.g., reviews and likes) and accounts (Caruccio et al., 2018; Jiang et al., 2014; Li et al., 2016; Lim et al., 2010; Sen et al., 2018; Wang, 2010), or in the measurement of fake engagement in specific platforms (e.g., Twitter Stringhini et al., 2012 or Facebook De Cristofaro et al., 2014). Other works have shed light on how the cybercrime underground offers the necessary infrastructure (Paquet-Clouston and Bilodeau, 2018) to, and is fueled by Bhalerao et al. (2019) fake engagement product and services. While the detection and analysis of the supporting infrastructure has been widely studied, the research community has overlooked the economics and operation of the fake engagement ecosystem, and work in this area has been quite limited (see Section 2). This is relevant because cybercrime activities are driven by economic factors, so a better understanding of the economics of the fake engagement ecosystem is of help to assess the maturity of this underground economy and to assist for market intervention (Collier et al., 2021).

To fill this gap, we carry out an extensive empirical analysis of the underground economy of SMM trading. Our key research goals are studying (*i*) the catalog of fake engagement services that are offered, (*ii*) their main features and prices; and (*iii*) 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 online communities: Hackforums and BlackHatWorld —two popular underground forums focused on the trading and discussion of illicit activities (Pastrana et al., 2018a; Portnoff et al., 2017)—. We then crawl and analyze data from the resulting set panels daily for 4 months. Using this methodology, we collect and compile a dataset formed by 2.8 M listings from 58 SMM panels. We augment this dataset with 7063 discussion threads obtained from the forums, about the trading activities and the conversations related to fake engagement. Our analysis combines automated and manual techniques, including text analysis (e.g., use of regular expressions to mine prices from the data), Machine Learning (e.g., to extract relevant threads from forums) and Social Network Analysis (e.g., to analyze the interactions of key actors).

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 and share with the research community a new dataset of offered services by crawling daily 58 SMM panels where they are advertised during a period of 4 months. This dataset consists of 2.8 M entries grouped in 61k different service variations. This dataset allows us to identify and further analyze 294 different services targeting 59 platforms, including the major Internet OSN, review services, video and music platforms. We describe the dataset compilation in Section 3.
- Using a combination of manual and automated text analysis techniques, we break down 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 (see Section 4). 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 in Section 5. Specifically, we analyze their prices and their variations across markets, and how different types of customization affect the market. Our results indicate lower prices than those reported in previous research (Paquet-Clouston and Bilodeau, 2018), and also large disparities between the price for the same service across markets. We observe that prices fluctuate across sites, which suggests that this market is still in an early maturity stage, possibly because the underlying supply chain supporting it is unstable.
- We complement our analysis with a study of the presence of these services in two relevant underground forums (Section 6). Our results confirm that they align with those being sold in dedicated panels. Also, we observe that actors reselling SMM services tend to start providing some free products to gain reputation, and that they complement their focus on SMM with other illicit activities.

Overall, our research sheds light on the vast underground ecosystem and marketplace of fake SMM. We show that dedicated panels services are prevalent, though often volatile, and reachable through forum advertisements. We also observe how customized services (e.g., followers from specific gender or location) increase prices and provides specialization to re-sellers. We believe our work will help future researchers and assist content providers to better understand the nature of fake engagement in the underground economy.

Dataset To foster reproducibility and independent verification of our results, we open source our dataset at: https://github.com/0xjet/smmpanels.

2. Related work

Fake engagement services gained research attention in the last decade. Most studies in this area can be grouped into two main categories that we describe next.

Detection and usage of fake engagement services Fake engagement has been widely used for market fraud, mostly by means of fake product reviews in popular online marketplaces such as Amazon (Lim et al., 2010; Mukherjee et al., 2012) or user-generated content sites such as TripAdvisor (Ott et al., 2011) or Yelp (Yao et al., 2017). Several works have analyzed the use and detection of fake engagement in OSN. One line of study mainly focuses on the detection of the presence of fraudulent entities (Boshmaf et al., 2016) or behavior (e.g., fake reviews (Ruan et al., 2020) or spamming (Fu et al., 2018)) in OSN. In particular, fake engagement detection studies have been carried out for the most prominent platforms such as Facebook (De Cristofaro et al., 2014; Gao et al., 2010), Twitter (Stringhini et al., 2012; Wang, 2010), YouTube (Li et al., 2016) and Instagram (Sen et al., 2018; Zarei et al., 2020). Lim et al. (2010) studied the detection of fake reviews as observed directly in the targeted platform by analyzing anomalous deviations from other reviews. Mukherjee et al. (2012) focus on groups of fake reviewers working in a collaborative setting. Gao et al. showed that fake engagement services for OSN might be also used in spamming campaigns, e.g., to advertise Phishing sites in Facebook (Gao et al., 2010). Bessi et al. describe how social bots were used in political manipulation during the 2016 US elections (Bessi and Ferrara, 2016). Indeed, these services, together with other other forms of misinformation (Sharevski et al., 2022), are used as a means to boost influence and increase reach. These studies present a wide variety of methods leveraging ML techniques, specially behavioral clustering, to study community structures and identify groups exhibiting common patterns of behavior (Beutel et al., 2013; Jiang et al., 2014; Wang et al., 2013), or the use of DeepLearning, relying on anomalies and deviations from regular ratings (Aghakhani et al., 2018; Lim et al., 2010). Most of these works rely on data collected directly from the targeted OSN, e.g.,

Table 1
Related work on social marketing fraud (SMF).

Work (year)	Research focus
Beutel et al. (2013); Fu et al. (2018); Gao et al. (2010); Lim et al. (2010); Mukherjee et al. (2012); Ott et al. (2011); Ruan et al. (2020); Yao et al. (2017)	Characterization and detection of fake reviews on different platforms.
Aghakhani et al. (2018); Boshmaf et al. (2016); Jiang et al. (2014); Lim et al. (2010); Wang et al. (2013)	Anomaly-based detection of fake accounts and automated bots on different platforms.
De Cristofaro et al. (2014)	Measurement of fake *likes* promotion using honeypot accounts in Facebook.
Stringhini et al. (2012)	Crawling and measurement of Twitter account markets.
Paquet-Clouston and Bilodeau (2018); Paquet-Clouston et al. (2016)	Analysis of the infrastructure supporting SMF by means of botnets.
Bhalerao et al. (2019)	Study on how SMF (mostly fake accounts) fuels other cybercrime activities in underground forums, such as romance scams and SIM swapping.
Our work	**Characterization and economic analysis of products and services offered in SMM panels and underground forums**

reviews or wall messages, to detect and describe its purpose. Few works, however, have studied where these services come from, how they are operated, and the economic factors of the trading ecosystem.

Operation and infrastructure of fake engagement systems Another line of study focuses on how these services are delivered, which actors are involved, and how they fit into the broader cyber-criminal ecosystem. A 2016 work by Paquet-Clouston et al. studies the Linux/Moose botnet and how it is used for Social Media Fraud (SMF) (Paquet-Clouston et al., 2016). This work not only explores 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 extended in 2018 with a heavier focus on the market aspects of the SMF supply chain (Paquet-Clouston and Bilodeau, 2018). In particular, it delves on the relationship between the botnet and the reseller panels and how the revenue is distributed among these actors. Bhalerao et al. presented a method to detect supply chains in underground forums (Bhalerao et al., 2019). They detected fake engagement services in at least 3% of the supply chains in Hackforums, and showed that they were fueling other cyber-criminal activities such as SIM swapping or romance scams. 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.

Table 1 provides a summary and comparison of key previous research work with our study.

3. Datasets and methodology

We gathered and studied two main datasets in this paper: (1) data crawled from online SMM panels; and (2) 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, which are commonly referred as Social Media (Marketing / Management) panels or SMM panels. To do so, we followed two complementary methods: (*i*) manually running Google searches using terms such as *"buy Instagram likes", "buy Facebook likes", "buy followers"*, etc.; and (*ii*) querying and manually browsing 2 underground forums – Hackforums and BlackHatWorld – where these services are commonly advertised. In addition, we leverage a list of 343 such panels compiled by a previous study conducted in 2018 (Paquet-Clouston and Bilodeau, 2018). Unfortunately, most of the panels indexed in this list either were not already active when we started our analysis or went down during the first few weeks of our crawling efforts. 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 shown in Table 2.

Table 2
Median number of service entries in the SMM panels of our dataset.

Daily entries	# 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

Crawling strategy We implement a custom web crawler to gather the services offered in the selected 58 panels. The crawler visits each panel daily from March 20th to August 17th 2020, recovering from each page the tables where the services are advertised. Then, we parse the tables to obtain structured data (i.e., product or service entries) that will be subsequently analyzed. This step involves some manual analysis to iteratively customize the parser so that we can successfully extract information regardless of the particularities implemented by 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.). Then, we 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 4.3).

Price normalization Service prices are shown using different currencies or unitary costs. Therefore, we convert them to USD[1] and normalize them to the format '$ per 1000', which is the most common format in the forums. Manual review is still necessary due to inconsistencies introduced by the service provider, including contradictions, typos, and other errors in the fields. For example, in the case of expensive services like Amazon or Google Business reviews, the price is often *per unit* despite the name of the column suggests per 1000s. In other cases, the service's *Name* field specifies a maximum amount available to order (e.g., "Instagram likes

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

[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. Further analysis revealed that these inconsistencies are a consequence of reselling: Some panels resell a service copying the name and description of the original provider, but changing the price. In these cases we choose the value specified in the *Price* field. After applying this process, we obtain a curated dataset of 2.8 M records.

Service discovery and indexation We process the original dataset to generate a second sanitized dataset with aggregated service data and without duplicates. It is common to find the exact same service name and description in several panels, sometimes even propagating spelling and grammar mistakes. 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. However, despite all the signals suggesting a common service provider, we have no ground truth to draw any solid conclusion about their uniqueness so we establish the following heuristics 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 with 61k different services. Then, for each unique entry, we extract information about the services such as the number of days that this service is advertised, the number of observed price changes, and basic price statistics (i.e., mean, standard deviation, maximum and minimum values, and quartiles).

3.2. CrimeBB dataset

We use the CrimeBB dataset (Pastrana et al., 2018b) to study the ecosystem and economics of fake engagement services in underground forums. This dataset is freely available from the Cambridge Cybercrime Centre.[2] Specifically, we study more than 91 million posts gathered from 27k SMM-related threads found in 34 different underground forums. While some of the forums cover general topics, others are specialized in video-game hacks and cheats, malware or online accounts. We identify and harvest SMM-related threads by extracting and analyzing the heading (title given to a conversation thread) and the content of the first post of each thread. Additionally, we collect metadata such as its author, timestamp, the number of replies and the forum in which it is posted.

Thread classification A preliminary exploration of the data shows that the posts gathered discuss different aspects of the SMM ecosystem. We find tutorials and guides (e.g., on how to grow popularity of Twitter[3]), service offerings (e.g., YouTube accounts[4] and requests for advice or help. As our goal is to automatically study service offerings, we trained a classifier by constructing a labeled dataset of 1.2k entries. To do so, we consider 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 build an NLP classifier using a RNN (Abadi et al., 2015) on top of the pre-trained encoder BERT (Devlin et al., 2019). In particular, we chose Bert-Mini, which has 4 transformer layers and 512 hidden embedding sizes, providing an adequate trade-off between performance and model complexity for our task. We fine-tune our model on 70% of the labeled posts and test it on the remaining 30%, which is used as validation set. As the dataset is balanced (45% vs. 55%), we obtain a F1 of 0.983 (accuracy 0.980, precision 0.986, recall 0.981). Using this classifier on the unlabeled threads, we automatically identify 7k threads related to SM service offers, sales, or advertisements, which we later analyze in Section 6.2.

[2] www.cambridgecybercrime.uk.
[3] https://hackforums.net/showthread.php?tid=5854028.
[4] https://hackforums.net/showthread.php?tid=2664677.

4. The fake engagement ecosystem

This section studies the market ecosystem of SMM panels. First, we create a catalog of all the services found in our dataset in order to analyze the variety and scale of this market. Then, we classify the services being offered and identify the associated social media platforms. Finally, we analyze service customization, i.e., different types and qualities for the same service.

4.1. Service catalog

Using the methodology described in Section 3.1, we identify a total of 294 different services (excluding variations or customization) across 59 different platforms. Fig. 1 shows the most common services that are offered for each platform. We can observe a great disparity in service popularity. On the one hand, a couple of dozens of services are present in nearly all the panels. Moreover, each panel contains a substantial amount of entries for each service, also offering variations of the services. On the other hand, we find many small services indexed at the bottom of a few panels.

Among the most prominent services, we find fake engagement services for popular OSNs like Instagram and Facebook, and website traffic (visits). We also observe services directed to generate fake activity in music platforms, mostly fake plays. We speculate that their popularity may be partially caused by the simplicity in which the service are provided, as just one account can be used to generate many plays for the same track, playlist or album. In fact, in some of these streaming platforms, no account is even required and there is no need for user interaction at all to simulate organic behavior. In general, the offer for SMM services in platforms for independent music artists reflects the demand of cheap marketing strategies in a very crowded and competitive environment.

We also find marginal services that present interesting characteristics. Manual classification allows us to categorize them as: (i) Premium accounts for video streaming 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 (Farooqi et al., 2020); and (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 5.3. Lastly, it is worth mentioning that the SMM panels themselves may be also indexed among the offered services as 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 (Paquet-Clouston and Bilodeau, 2018), which we confirm during our analysis of underground forum data in Section 6.

4.2. Service popularity

We use the following 3 metrics to measure service popularity: (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 shown in Table 3. It is important to note that these metrics give an idea of the service popularity but they may

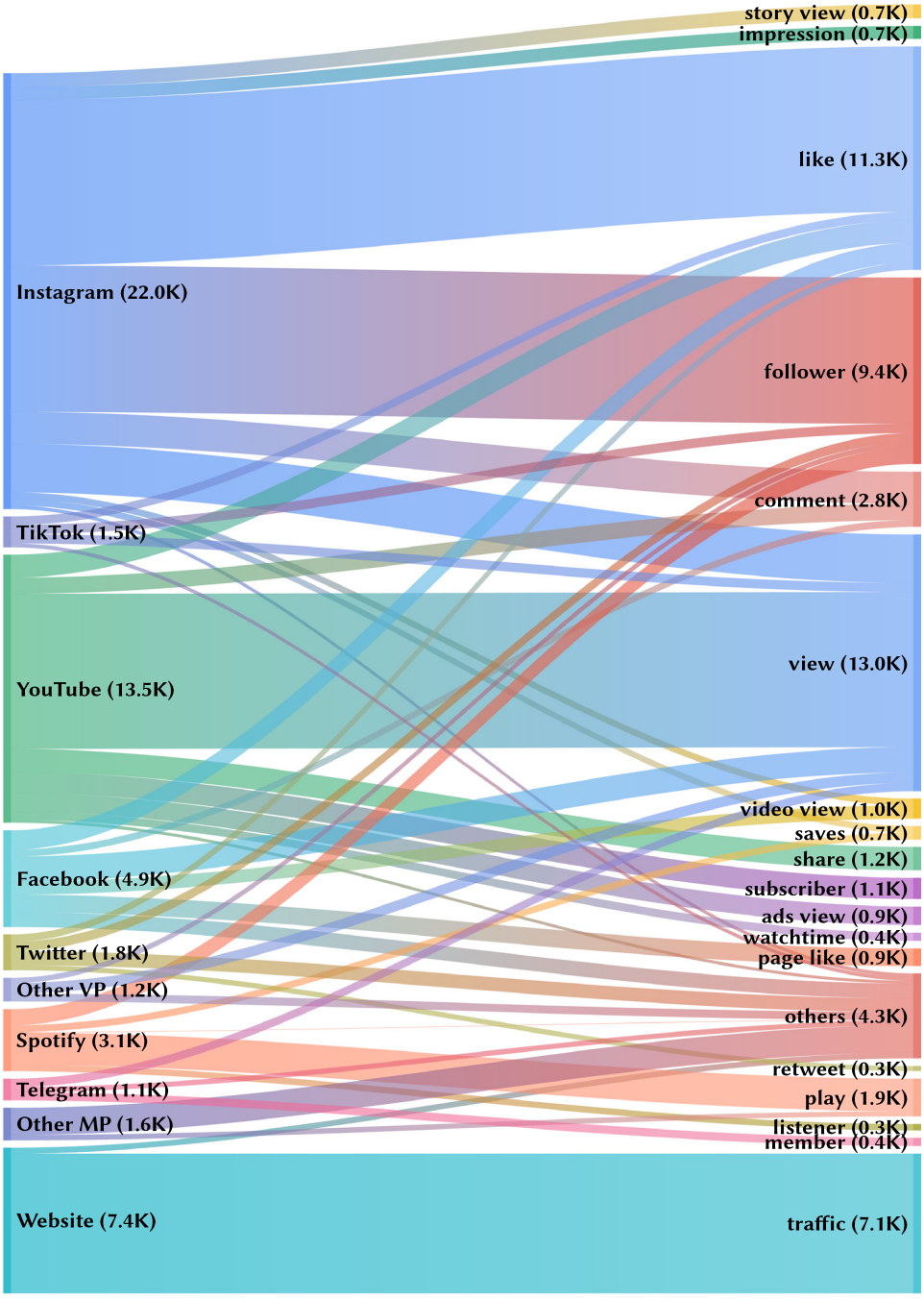

Fig. 1. Catalog of the most popular services found in SMM panels. Due to the scale of the catalog, and to ease visualization, less popular services and platforms are grouped together. Small music/audio and video platforms have been grouped under the labels *Other MP* and *Other VP*.

not reflect their actual demand nor their impact as fake engagement traffic in the target platforms. Another relevant observation is that each metric is biased towards different kinds of services. For example, web traffic is offered in many panels from different locations and with different referrers. As a consequence, it is over-represented 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 over-represented compared to simpler ones in the *different variations* metric.

The ranking of top services shows 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 likes, possibly because of two factors: (*i*) likes are platform-defined indicators of content quality and popularity (e.g,. posts, videos, songs, etc.), so their manipulation can impact the recommendations algorithms to attract organic users; and (*ii*) likes are easier to manipulate automatically (e.g., via botnets) as opposed to other features like comments. These properties, as we will see in Section 5, makes them an effective and cheap mechanism to boost content popularity.

We used the aforementioned popularity metrics to rank the targeted platforms. The top 8 platforms are shown in Fig. 2, indicating both their daily entries and number of service variations. We ob-

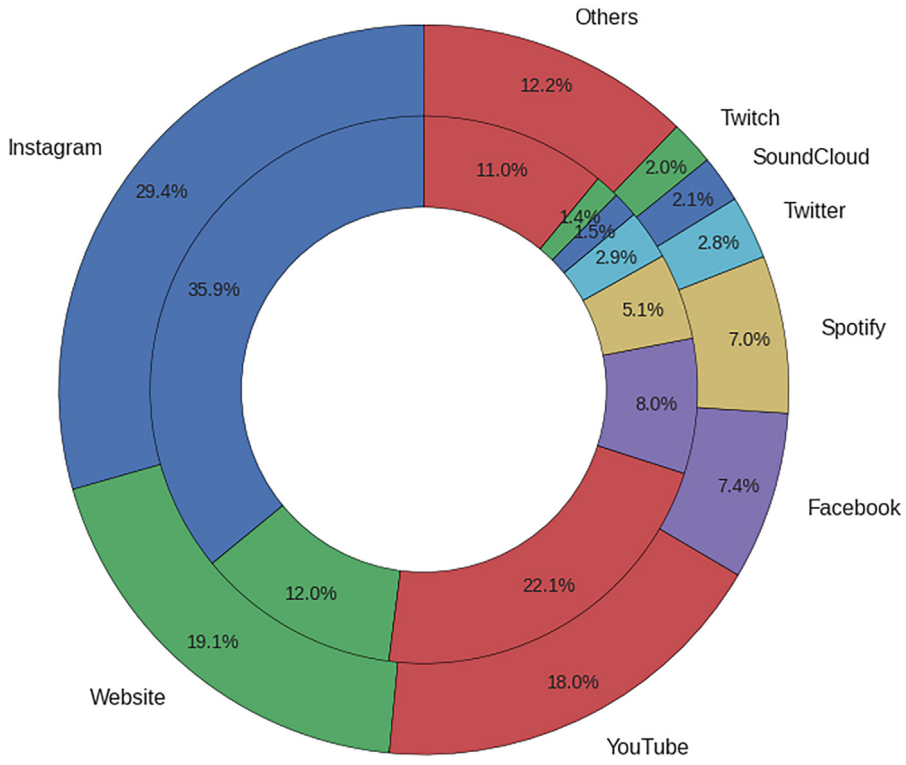

Fig. 2. Top 8 targeted platforms by services in SMM panels. The inner ring represents the percentage of different services (variations) identified. The outer ring represents the percentage of mean daily entries.

Table 3
Top 20 services in SMM panels.

Site	Product	Entries/day		Different variations	% panels
Website	traffic	4695	493	7066	72.4
Instagram	like	2677	235	8362	100.0
YouTube	view	2524	436	7836	98.3
Instagram	follower	1995	236	7390	100.0
Instagram	view	1084	70	2446	100.0
Spotify	play	971	89	1639	87.9
Instagram	comment	700	52	1622	94.8
YouTube	like	453	35	1151	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	1165	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	1272	70.7
YouTube	subscriber	267	47	1066	96.6

serve 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 (at the time the dataset was collected), we observe that Instagram (1082 M) is by far the most targeted for its size compared to Facebook (2,603 M), YouTube (2,000 M) and Twitter (326 M).[5] Fig. 2 also reveals the bias of the used metrics that we previously mentioned, which is particularly clear for the Website category.

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

4.3. Service customization

One interesting feature of SMM panels is the multiple customizations offered for each available service. 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, we can also find low quality (or standard versions) and more expensive, reliable and improved (or premium) offerings. An example of this are the *comments* offered for multiple platforms. While it is common to find very cheap services that provide *random* comments, other premium services offer *customized* and real-looking ones for the same platform at a higher prize. We note 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) whereas premium services offer customized and 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 further analyze customization, we first extract and analyze recurring keywords and tags that may characterize them. Then, we classify the services by the presence of these keywords in their names and descriptions, which we list below.[6] Afterwards, 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 is mostly used in other services like

[6] We note that as we do not actually buy the service, 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.

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 is 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 modifiers, 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 be sold as a way to avoid the detection mechanisms implemented by the targeted platforms, although this is speculative.
- **Drop / No Drop**. Many platforms actively try to detect and eliminate inorganic engagement to mitigate their potential harmful impact. 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 customization price-wise. Custom comments are several times more expensive that random or generic ones (see Section 5). We can attribute this price difference to the fact that custom comments may need some degree of human intervention or advanced linguistic tools, which would substantially increase the cost of an otherwise completely automated process. This need of human intervention is probably the main reason behind reviews in *Amazon, LinkedIn* and *Google Business* ranking as the most expensive services. In turn, this raises the question of how advancements in automated AI-generated texts (Radford et al., 2019) 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 guarantee 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 that 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.

Table 4
Presence of geo-targeted services.

Site	Product	Total	Geo-targeted	%
Website	traffic	7066	5447	**77.09**
YouTube	view	7836	3449	44.01
Instagram	like	8362	1589	19.00
Instagram	follower	7390	1445	19.55
Spotify	play	1639	950	**57.96**
Instagram	comment	1622	622	38.35
YouTube	comment	822	597	**72.63**
YouTube	ads view	940	522	**55.53**
YouTube	share	1165	421	36.14
Spotify	follower	811	377	46.49

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

N Targeted Services	USA 2843	Brazil 1785	India 1160	UK 1006	Russia 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

4.3.1. Geo-targeting

In addition to the discussed variations, services may offer location-based targeting and customization (i.e., geo-targeted services). The majority of them offer the possibility to select a specific language—mostly text-based services like comments and reviews—and country from which the service will be delivered, including source IP of web traffic or the registration country of the accounts. Table 4 presents the top-10 services ranked by the number of 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 YouTube and Spotify have higher percentages of geo-targeted services than those of Instagram, where only 19% of likes or followers allow for geo-targeting.

As for the location from which the services are offered, we find 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. Fig. 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 5. 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 located in the USA and UK are much more evenly spread in terms of platforms when compared to the other three. 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 study the price ranges for the most popular services. Then, we analyze their price variability during the period of the study and compare it within and across panels. We further

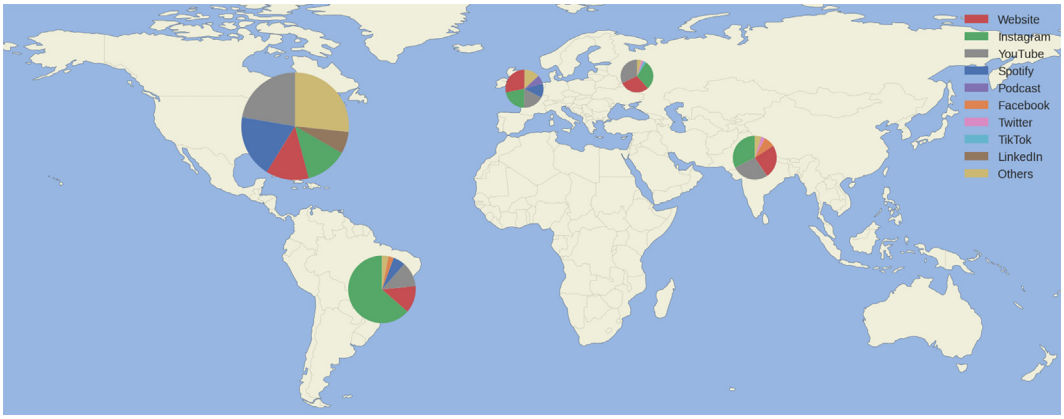

Fig. 3. Target platform distribution for the top 5 countries of origin of geo-targeted services: USA, Brazil, India, Russia and UK. The pie charts are scaled by the number of services offered in each country.

Table 6
Prices of top 20 most advertised services.

Site	Product	Count	Min	Q1	Median	Q3
Website	traffic	5340	0.09	0.36	0.39	0.60
Instagram	like	4732	0.06	0.80	1.43	2.88
YouTube	view	3729	0.25	1.35	2.10	3.00
Instagram	follower	4057	0.08	2.20	4.62	8.50
Instagram	view	1085	0.00	0.03	0.08	0.25
Spotify	play	1373	0.33	1.20	2.10	3.77
Instagram	comment	1058	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

extend this analysis with a review of the cheapest and most expensive services we encounter. We conclude with a study of the higher cost of services offering customized services.

5.1. Data sanitization

Before obtaining price metrics and statistics, it is critical to carefully filter the data to remove potential errors and outliers. Unfortunately, distinguishing whether a high price is due to sellers' high expectations, errors, or just because delivering the service is actually expensive (e.g., those between $\$10^5$ and $\$10^7$, which corresponds to \$100 and \$10k per unit) is challenging. We develop the following heuristics to filter out potential outlier services: (i) their 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 5 and its maximum order is above 100; and (iv) it is classified as a package or bundle. The last two conditions help eliminating services that may be valid but are not relevant for studying the price of bulk services. Therefore, we remove highly specialized services or packages from our analysis (e.g., Instagram followers from verified accounts) as they would provide a distorted view of the market.

Table 6 reports the prices for the top 20 most popular services. We report the median and quartiles instead of mean and standard deviation as, despite our filtering efforts, there are outliers that require robust statistical metrics. This choice may explain why the results presented in this paper differ substantially from mean values reported in previous research studies (De Cristofaro et al., 2014; Paquet-Clouston et al., 2016). However, it could also be the case that these services have gotten cheaper in the last years as the market developed.

5.2. Price stability

This section studies the fluctuation in the different services' prices over time to understand the overall market volatility. Unfortunately, measuring these factors is not trivial due to the fact that services are typically taken down or re-branded. Indeed, instead of modifying the price in a service, it is common to take it down and replace it with a new one under a different ID with a slightly different name. Yet, out of the 61k observed services, 88.6% of them had stable prices. We used a Welch's t-test to validate if services with stable prices had a shorter lifespan than those whose price fluctuated across time. With a t-value$= -25.58$ we confirmed that indeed, the difference in duration is significant. The mean duration of fixed price services is 45.3 days, 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 in mind, we study the price evolution of aggregated services such as Instagram followers and YouTube views instead of studying each individual service. We filter out outliers using the criteria defined to generate Table 6. Then, we evaluate the price of the most popular services during the first 10 days of April, May, June and July 2020, so we measure local variances for each one-month period. The boxplots rendered in Fig. 4 show the results of this analysis, the x-axis being a service and the colors representing each studied month. Fig. 4 suggests that prices do not change significantly during this 4-month period. Instagram comments is the service that exhibits the most significant variation, starting with a median of \$29 in April and steadily decreasing down to \$20 by July. However, this trend is minimal in comparison to the price range observable within and across panels in the whole market. In fact, a key observation is the wide range of prices that exist within services. This is particularly clear for the services plotted in the right side of the figure. For example, we can see that just for YouTube comments in the month of May, prices varied from \$5 to over \$130. Similarly, Facebook views range from less than a dollar to over \$200. A potential reason for such a variance is the presence of service customization, since it is expected that geo-targeted custom comments for YouTube are more

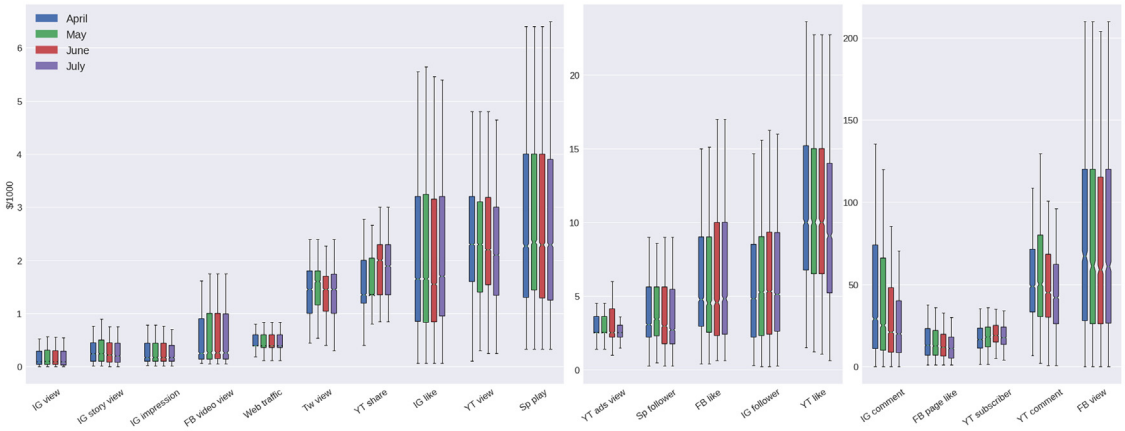

Fig. 4. Boxplot of the most of the top-20 most popular services during 4 months. The whiskers extend to the higher/lower observation within one IQR of the quartiles. *We note that, due to the variability in the amount of services, the graph has been divided to use different scales and enhance the visibility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

expensive than random non geo-targeted ones. Moreover, some of these panels may act as reselling platforms, thus offering them at a higher price than the original provider. In the next section we further illustrate this issue with a particular case study: Google Business reviews.

5.3. High-end services

In this section, we perform a deepr analysis of the most expensive services. The constant presence of outliers forces us to take a semi-manual approach to review each service. Fortunately, the number of expensive services is limited so manually inspecting each service 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 discuss each category separately in the following subsections.

5.3.1. Review and rating services

This type of service 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 offerings for Google Business, LinkedIn, TripAdvisor, IMDb, Play Store and Apple Store. In the case of mobile app stores, the services offer app installs, reviews and ratings. We discuss three cases of interest next:

- **Google Business reviews** are offered in 21 panels and their price ranges from $1.4 to $18.0 per review as shown in the violin plots rendered in Fig. 5. The shape suggests that these services either have an underlying common services or act as resellers. In fact, we find that several services offer the exact same name and description in 6 different panels but at different prices. 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 difficults the study of the services' customization on the price because, as we see, aggregating prices from different panels may hide the impact of customization. The available orders go from 1 to 100 reviews, although one service offered up to 1.5k reviews.
- **Tripadvisor reviews** are much less common, appearing only in 2 panels. In one of the panels the service description does not have much information. The price of a custom review in one panel goes from $0.25 to $0.72. In the other panel, the service provides a more detailed name and description of the service, clearly stating that the service provides a custom review and a rating using real accounts. The vendor also advertises the origin country of the accounts and other features like the delivery rate, and also grants a refill policy in case the reviews and rating are taken down. This service is priced at $5 per review.
- **IMDb votes**, which would be the equivalent of a rating in other platforms, are quite expensive. They appear in 20 panels with prices that range from $15 to $20 per 1000 in the low end, to $80 and $150 per 1000 in the high end. The country and gender of the voter could be selected.

5.3.2. Accounts for subscription services and non-free interactions

Unsurprisingly, services that require paying a registration or subscription fee are more expensive than those that do not. However, Netflix accounts or Amazon Prime subscriptions are sold by a fraction of their legit price. These services offer the possibility of purchasing either individual accounts (most common model) or bulk packages, being specifically advertised for resellers. Some of these services also state that the accounts for sale are hacked or stolen ones.

A particularly interesting case within this category is the service **Twitch Subscribers**, a video streaming platform that allows users to follow a streamer's channel for free or subscribe to it for a fee (typically around $5[7]) in order to support the streamer. Additionally, users with an Amazon Prime account can subscribe to one Twitch channel for free. These are called Prime Subscriptions and despite being free for the user, the streamer gets revenue from Amazon. We find several services of such subscriptions with prices ranging from $1.5 to $3. Streamers purchasing these services would not only boost their accounts, but also potentially make money from Twitch: For example, if the streamer receives a 40% of the subscription fee value, then for a $5 worth subscription they receive $2. So by purchasing subscriptions at a price below $2 the streamer would be on profit. The percentage of the subscription fee received by streamers is not fixed but this is plausible scenario.

5.3.3. 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. Table 6 shows that 1000 Instagram likes usually cost around $1.43. However, 1000 Power likes and 30 posts cost $275. A given service 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. In

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

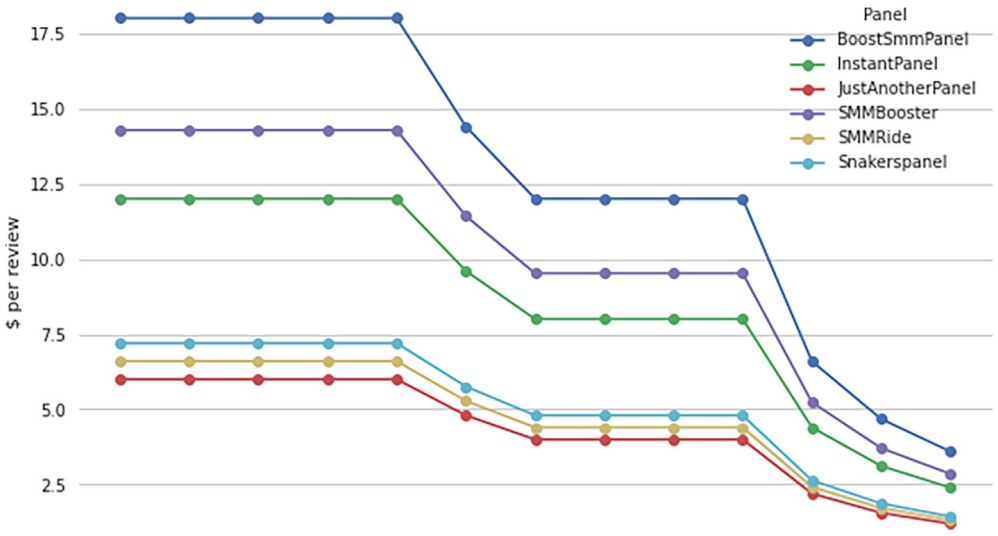

Fig. 5. *Google Business review* prices for 13 exactly equal services (points) present in 6 different panels (lines). Y-axis represents the prices of the different services. Thus, points aligned vertically represent the prices of a same service across the 6 panels.

fact, Power likes are a controversial topic as 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.

5.4. Low-end services

We now focus on the cheapest services indexed by these panels. Interestingly, we find that 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 and they offer Instagram likes, comments and followers; and TikTok or YouTube views. Typically, free services offer a low quantity of fake engagements for a limited amount of time. Yet, the cheapest non-free services are video views and plays. For less than $0.10 per 1000 actions we find TikTok views ($0.01–$0.03), SoundCloud plays ($0.05 aprox) and Instagram/IGTV views ($0.05–$0.08). In the range of $0.10 and $0.30, we find Instagram story views and Facebook video views. Up to $0.60 per action, we find web traffic. The fact that these services are easy to automate could explain why they are so inexpensive.

5.5. Impact of price customization

Finally, we study the price variations of services offering different levels of customization as presented in Section 4.3. To do so, we choose three types of customization and analyze them over two services where they are relevant. All of these types are analyzed along side the geo-targeting variable. Geo-targeted services are significantly more expensive and thus need to be studied separately to be able to observe the impact of the target customization.

In all 3 cases, we apply our filters to remove outliers and erroneous data. However, we also remove services with a price exceeding 5 times the third quartile of the prices (see Table 6) in order to generate more useful visualizations. Then, we select the locations where each service variation was most popular. Lastly, we generate graphs for each location depicting the distribution of prices for services with and without the customization. The results are presented in the violin plots rendered in Figs. 6–8. We note that non geo-targeted services are placed in the leftmost side of the figure labeled as *Unspecified*. In each violin plot, 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 results shown in Fig. 6 reveal a substantial difference in price between geo-targeted and non geo-targeted services in both platforms. 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 do not see a clear difference in average prices. However, there are significant differences in the shape of the distributions with custom comments tending to have a more spread out shape, with a heavier tail towards high prices. There is also a more noticeable difference if we compare custom comments with those explicitly advertised as *Random*. However, as with many other negative keywords, we find much fewer of these services and therefore we do not have enough data to drive solid conclusions.
- **Instagram gender targeted followers and comments.** We 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, 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 Fig. 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 select services offering fake plays in two of the most popular audio and music streaming platforms: Spotify and SoundCloud. We note that SoundCloud is less targeted than Spotify. The US is the only country for which geo-targeted ac-

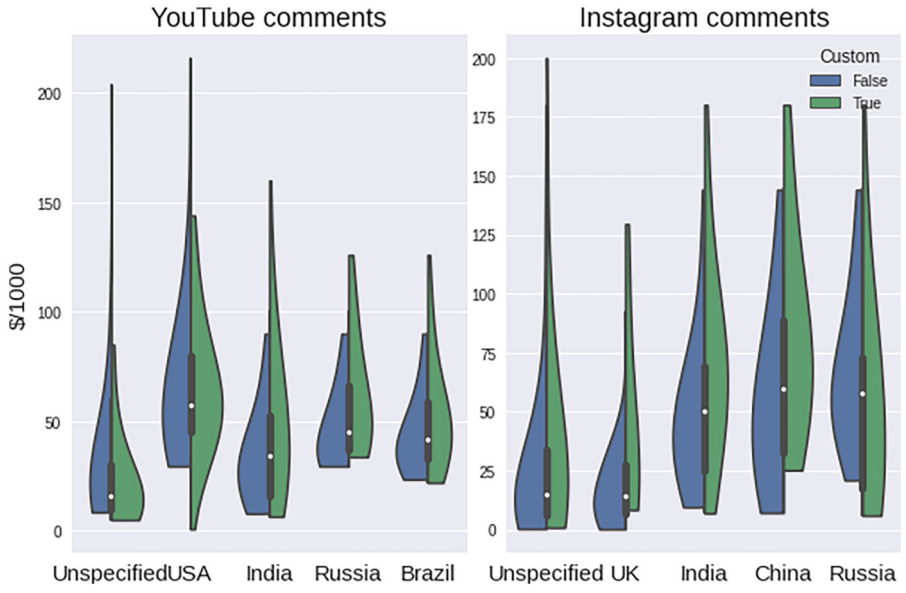

Fig. 6. Prices distribution of YouTube comments and Instagram comments attending to geo-targeting and the *Custom* variation. Each individual violin plot represents the prices for a given region (horizontal axis). The left, blue side represents the distribution for non-custom comments, while the green, right side of the violin represents the distribution for the custom comments. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

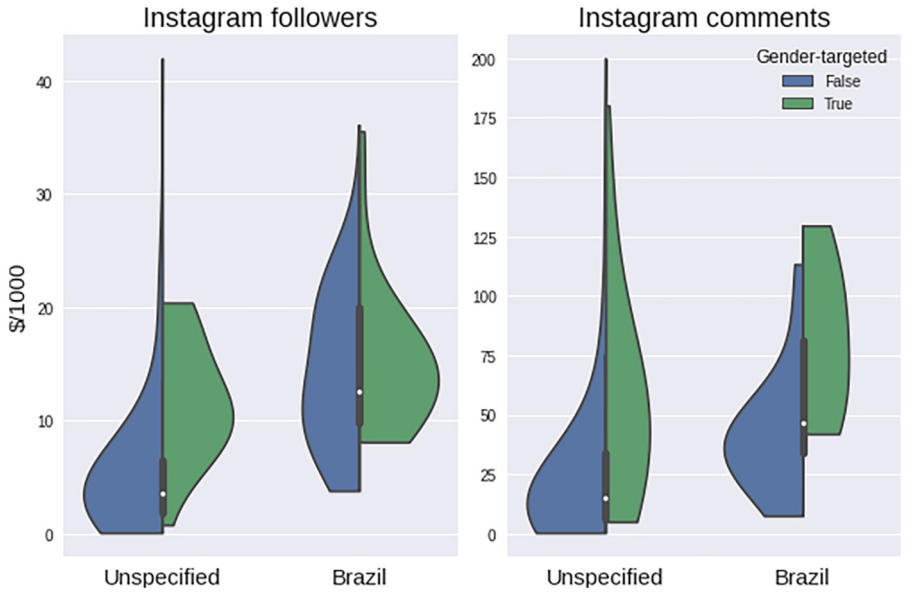

Fig. 7. Prices distribution of Instagram followers and comments attending to geo-targeting and gender-targeting. Each individual violin plot represents the prices for a given region (horizontal axis). The left, blue side represents the distribution for non gender-targeted services, while the green, right side of the violin represents the distribution for specifically male or female services. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tivity can be purchased for SoundCloud, whereas Spotify services are available from 30 different countries. 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. The most expensive location for Spotify plays is Germany, although the prices for France, UK, Canada and Brazil are similar. If we focus on the *Refill* option, we observe a significant difference in prices due to geo-targeting as shown in Fig. 8. We see that refillable services start at a higher price for SoundCloud plays and geo-targeted Spotify plays. We also observe that the prices for refillable services reach higher prices for Spotify plays from France, UK, Canada and Brazil but not for the rest of locations nor for SoundCloud plays. In short, while there is a tendency to increase the price for Spotify services in these five countries, we observe the opposite trend for SoundCloud plays from the US where all refillable plays prices are below the median. 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 three case studies illustrate the impact of geo-targeted price customization. In the case of services 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

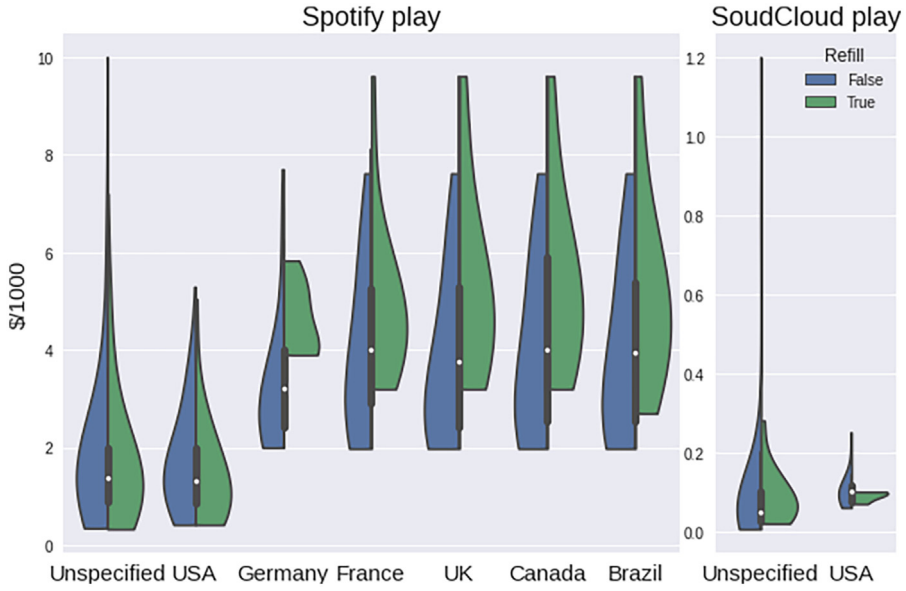

Fig. 8. Prices distribution of Spotify and SoundCloud plays attending to geo-targeting and the *Refill* variation. Each individual violin plot represents the prices for a given region (horizontal axis). The left, blue side represents the distribution for services that are not advertised with *Refill*, while the green, right side of the violin represents the distribution for those which are. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to have a substantial effect in price for services in certain platforms. Such targeting may render useful in platforms like Instagram if the aim of the buyer is to influence the recommendation algorithm and reach certain user demographics and 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

This section complements our previous analysis by looking at two aspects related to the supply chain and infrastructure needed to promote and operate SMM product and services. First, we study the domain names used by the panels to infer how these panels are perceived by the IT security industry and how are they ranked in terms of popularity (i.e., by traffic received). Second, we analyze the trading and discussions about SMM panels on online underground forums, offering us a unique perspective to map and understand the actors involved in this business.

6.1. Domain classification and popularity

For each panel, we extract the labels provided by four domain classification services: Symantec, McAfee, Fortiguard, and OpenDNS. However, it is known that domain classification services present fundamental limitations not only in their categorization and consistency, but also in terms of coverage (Vallina et al., 2020). In fact, our results suggest that these panels are largely seen as regular IT, business or marketing service providers. For example, McAfee 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 by these security firms. 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”. Just one engines flags two domains as “suspicious”. These results confirm that key actors in the security industry do not consider these sites as incurring in any potentially harmful activity.

We also analyzed the popularity of these domains using Alexa traffic ranks (Amazon.com, 2021). 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 the top-100k are generally deemed as statistically insignificant due to the scarcity of data available for them (Scheitle et al., 2018). Such low ranking positions suggest that these sites sustain a very reduced amount of traffic.

6.2. Trade in underground forums

Underground forums are known as common places for trading various types of illicit products and services (Motoyama et al., 2011; Pastrana et al., 2018a; 2018b; Portnoff et al., 2017). In order to understand the underground economy of SMM, we study 7063 forum threads providing SMM services (see Section 3). Unfortunately, a given thread may simultaneously offer services for multiple platforms and the titles of the threads are not as structured as SMM panels, which hinders a proper categorization of the services being offered. Therefore, we conduct a best-effort analysis by automatically looking for specific keywords in thread headings to get a birds’ view.

In total, we find threads offering 6708 services from 37 platforms. Table 7 shows the most common services from the platforms most traded in these forums. The top services align with the findings from SMM Panels (c.f. Table 3), with three notable differences: (*i*) website traffic is less popular in forums; (*ii*) Snapchat is the fifth platform most traded in forums; and (*iii*) *shoutouts* are more common on underground forums (i.e., the promotion of a user account from a popular account 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 eWhor-

Table 7
Top 6 platforms and their products being advertised in underground forums.

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

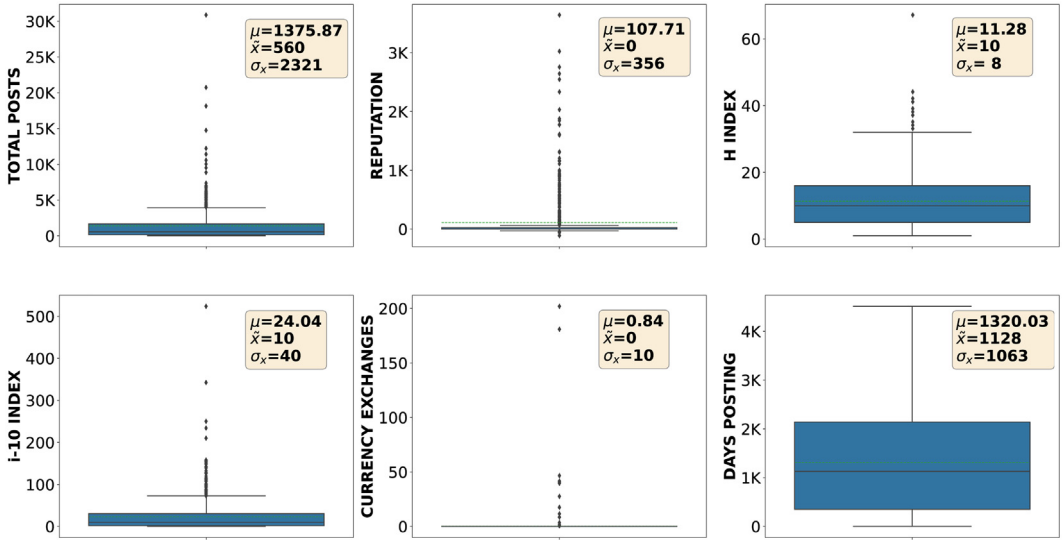

Fig. 9. Forum activity for the users offering social media marketing services in Hackforums.

ing (Hutchings and Pastrana, 2019) or Cost-Per-Action (CPA) services.

6.3. Actor analysis

Most of the threads in our dataset (82.7%) are from Hackforums. Therefore, we conduct and qualitative and quantitative analysis of the actors offering social media marketing services in this forum. We identify a set of key actors[8] by the number of their SMM offerings. From the total of 3235 actors operating, we study 751 actors that have at least 2 offerings, which account for 58.3% of the threads.[9]

For the quantitative analysis, we follow the methodology proposed in previous works (Pastrana et al., 2018a; 2019). Specifically, we extract and analyze the forum activity of the actors, including the number of posts, the reputation, activity intervals, user interests and their social relations, which are obtained from measuring forum interactions between a pair of 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 and threads made in boards from the different categories, as provided by the forum, such as hacking, gaming, technology or market.

Fig. 9 shows the aggregated statistics related to the forum activity for the 751 key actors. There is a wide diversity of actors according to their activity. For example, whereas the average number of posts is around 2.3k, some actors generated more than 30k posts, while other only have a few dozens. We also observe similar patterns in their reputation as in the number of positive points. There are substantial differences between a small set of 17 highly reputed actors (with more than 1k positive points) and the majority, which has no reputation at all (the median is 0). Actors offering SMM services in underground forums have on average a H-index of 10, ranking similar to key actors initiating in cybercrime activities according to a previous study (Pastrana et al., 2018a). 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 (Pastrana et al., 2018b).

Hackforums is divided in different categories, e.g., market, hacking or gaming. In order to characterize the broad interests of the authors, we analyze their activity in boards from these categories. Specifically, we count the number of threads started and replies posted in these boards. Table 8 shows the three categories capturing actors' interests. As expected, ≈75% of the actors have as primary interest the Market (with ≈20% 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 discussions 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 abuse and cybercrime-related topics.

[8] We define an *actor* as an account id posting messages in the forum. Thus, we do not consider cases where the same user operates various accounts.

[9] There are 2670 low-impact actors that have not received any response.

Table 8
Interests of the 751 actors offering SMM services in Hackforums.

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)

Top-10 actors We conduct a manual and qualitative analysis for top-10 authors (measured by the number of offerings that they offer). We analyze all the threads started by these users in the forum, including those related with SMM and other topics. We observe a common pattern in most of these popular authors: they tend to start by providing a small amount of cheap or easy-to-get services (e.g., Instagram shoutouts, Youtube views or FB likes) possibly to increase their reputation, attract users and to get known by the community. In fact, this is a common practice in underground forums where trust is a valuable active that must be gained across time (Dupont et al., 2016). After this initial self-promoting period that typically lasts a few weeks, they start trading the services. We notice various activities potentially related to required components of the supply chain, including SMM panel designs, Twitter bots, Instagram accounts with several followers (possibly used for selling shoutouts) and automatic Youtube account makers. Interestingly, we find two cases where actors have built more than one SMM panels over time, only for selling them as high-quality products afterwards. This confirms that panel re-selling is a common practice. Finally, in parallel to their activity in SMM products, these popular actors also operate other illicit businesses. For example, two of the ten top actors provide and sell services related to eWhoring (Hutchings and Pastrana, 2019; Pastrana et al., 2019), while other three provide accounts related to video games.

7. Discussion

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 2.8 M 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 customization We observe that most of these services are offered with an impressive variety of customization 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 types of customization 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. (2014) 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. (2016). 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 (Paquet-Clouston and Bilodeau, 2018). As Paquet-Clouston et al. (2016) 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.

8. Conclusion

This paper presents a measurement study of the fake engagement ecosystem for online platforms. We collect and open-source a new dataset of product listings from various SMM panels. Our work sheds light on the current status of the ecosystem, showing that this is a feature-rich and growing yet unstable market. We also analyze data collected from discussions arisen in underground forums. We believe that our work will inform researchers and security practitioners to better understand this underground economy. Future work includes crawling other sites, as well as conducting periodic re-crawls to offer a longitudinal view of the evolution and variation of the market. We also consider adding Chinese, Russian and Spanish panels to our dataset, as this would enrich the results drawn from them and would provide a global perspective on the market. To better understand re-selling activities, research on attribution could also be carried out by identifying common actors across the supply chain and across markets.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data is available in a public repository.

Acknowledgments

We thank the Cambridge Cybercrime Center for giving us access to the CrimeBB dataset. This work was supported by the EU Horizon 2020 Research and Innovation Program under Grant agreement no. 101021377 (TRUST aWARE); the Spanish grants ODIO (PID2019-111429RB-C21 and PID2019-111429RB-C22), and the Region of Madrid grant CYNAMON-CM (P2018/TCS-4566), co-financed

by European Structural Funds ESF and FEDER. The opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect those of any of the funders.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org. https://www.tensorflow.org/.

Aghakhani, H., Machiry, A., Nilizadeh, S., Kruegel, C., Vigna, G., 2018. Detecting deceptive reviews using generative adversarial networks. In: 2018 IEEE Security and Privacy Workshops (SPW). IEEE, pp. 89–95.

Amazon.com,. Alexa - top sites. [Online] Last Accessed 16 December 2021. https://www.alexa.com/topsites.2021

Bessi, A., Ferrara, E., 2016. Social bots distort the 2016 U.S. presidential election online discussion. First Monday 21. doi:10.5210/fm.v21i11.7090.

Beutel, A., Xu, W., Guruswami, V., Palow, C., Faloutsos, C., 2013. Copycatch: stopping group attacks by spotting lockstep behavior in social networks. pp. 119–130. doi:10.1145/2488388.2488400.

Bhalerao, R., Aliapoulios, M., Shumailov, I., Afroz, S., McCoy, D., 2019. Mapping the underground: supervised discovery of cybercrime supply chains. In: 2019 APWG Symposium on Electronic Crime Research (eCrime), pp. 1–16. doi:10.1109/eCrime47957.2019.9037582.

Boshmaf, Y., Logothetis, D., Siganos, G., Lera, J., Lorenzo, J., Ripeanu, M., Beznosov, K., Halawa, H., 2016. Ntegro: leveraging victim prediction for robust fake account detection in large scale OSNs. Comput. Secur. 61, 142–168. doi:10.1016/j.cose.2016.05.005.

Boshmaf, Y., Muslukhov, I., Beznosov, K., Ripeanu, M., 2013. Design and analysis of a social botnet. Comput. Netw. 57 (2), 556–578. doi:10.1016/j.comnet.2012.06.006. Botnet Activity: Analysis, Detection and Shutdown.

Caruccio, L., Desiato, D., Polese, G., 2018. Fake account identification in social networks. In: 2018 IEEE International Conference on Big Data (Big Data), pp. 5078–5085. doi:10.1109/BigData.2018.8622011.

Cascavilla, G., Tamburri, D.A., Van Den Heuvel, W.-J., 2021. Cybercrime threat intelligence: a systematic multi-vocal literature review. Comput. Secur. 105, 102258. doi:10.1016/j.cose.2021.102258.

Collier, B., Clayton, R., Hutchings, A., Thomas, D., 2021. Cybercrime is (often) boring: infrastructure and alienation in a deviant subculture. Br. J. Criminol. 61 (5), 1407–1423.

De Cristofaro, E., Friedman, A., Jourjon, G., Kaafar, M. A., Shafiq, M., 2014. Paying for likes? Understanding facebook like fraud using honeypots.

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2019. Bert: pre-training of deep bidirectional transformers for language understanding. arXiv:1810.04805.

Dupont, B., Côté, A.-M., Savine, C., Décary-Hétu, D., 2016. The ecology of trust among hackers. Global Crime 17 (2), 129–151.

Farooq, M.J., Zhu, Q., 2019. Modeling, analysis, and mitigation of dynamic botnet formation in wireless IoT networks. IEEE Trans. Inf. Forensics Secur. 14 (9), 2412–2426.

Farooqi, S., Feal, Á., Lauinger, T., McCoy, D., Shafiq, Z., Vallina-Rodriguez, N., 2020. Understanding incentivized mobile app installs on google play store. In: Proceedings of the ACM Internet Measurement Conference, pp. 696–709.

Fu, Q., Feng, B., Guo, D., Li, Q., 2018. Combating the evolving spammers in online social networks. Comput. Secur. 72, 60–73. doi:10.1016/j.cose.2017.08.014.

Gao, H., Hu, J., Wilson, C., Li, Z., Chen, Y., Zhao, B.Y., 2010. Detecting and characterizing social spam campaigns. In: Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement. Association for Computing Machinery, New York, NY, USA, pp. 35–47. doi:10.1145/1879141.1879147.

Hutchings, A., Pastrana, S., 2019. Understanding ewhoring. In: 2019 IEEE European Symposium on Security and Privacy (EuroS&P). IEEE, pp. 201–214.

Jiang, M., Cui, P., Beutel, A., Faloutsos, C., Yang, S., 2014. Inferring strange behavior from connectivity pattern in social networks. vol. 8443. doi:10.1007/978-3-319-06608-0_11.

Kambourakis, G., Kolias, C., Stavrou, A., 2017. The Mirai botnet and the IoT Zombie Armies. In: MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM), pp. 267–272. doi:10.1109/MILCOM.2017.8170867.

Li, Y., Martinez, O., Chen, X., Li, Y., Hopcroft, J.E., 2016. In a world that counts: clustering and detecting fake social engagement at scale. In: Proceedings of the 25th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, pp. 111–120. doi:10.1145/2872427.2882972.

Lieber, C., 2014. The dirty business of buying Instagram followers. Retrieved May 31, 2019.

Lim, E.-P., Nguyen, V.-A., Jindal, N., Liu, B., Lauw, H.W., 2010. Detecting product review spammers using rating behaviors. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management. Association for Computing Machinery, New York, NY, USA, pp. 939–948. doi:10.1145/1871437.1871557.

Motoyama, M., McCoy, D., Levchenko, K., Savage, S., Voelker, G.M., 2011. An analysis of underground forums. In: Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference, pp. 71–80.

Mukherjee, A., Liu, B., Glance, N., 2012. Spotting fake reviewer groups in consumer reviews. In: Proceedings of the 21st International Conference on World Wide Web. Association for Computing Machinery, New York, NY, USA, pp. 191–200. doi:10.1145/2187836.2187863.

Ott, M., Choi, Y., Cardie, C., Hancock, J. T., 2011. Finding deceptive opinion spam by any stretch of the imagination. arXiv preprint arXiv:1107.4557.

Paquet-Clouston, M., Bilodeau, O., 2018. Uncovering the wholesale industry of social media fraud: from botnets to bulk reseller panels

Paquet-Clouston, M., Décary-Hétu, D., Bilodeau, O., Dupuy, T., 2016. When greed for fame benefits large-scale botnets.

Pastrana, S., Hutchings, A., Caines, A., Buttery, P., 2018. Characterizing eve: analysing cybercrime actors in a large underground forum. In: International symposium on Research in Attacks, Intrusions, and Defenses. Springer, pp. 207–227.

Pastrana, S., Hutchings, A., Thomas, D.R., Tapiador, J., 2019. Measuring ewhoring. In: Proceedings of the Internet Measurement Conference.

Pastrana, S., Suarez-Tangil, G., 2019. A first look at the crypto-mining malware ecosystem: a decade of unrestricted wealth. In: Proceedings of the Internet Measurement Conference. Association for Computing Machinery, New York, NY, USA, pp. 73–86. doi:10.1145/3355369.3355576.

Pastrana, S., Thomas, D.R., Hutchings, A., Clayton, R., 2018. Crimebb: enabling cybercrime research on underground forums at scale. In: Proceedings of the 2018 World Wide Web Conference. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, pp. 1845–1854. doi:10.1145/3178876.3186178.

Portnoff, R.S., Afroz, S., Durrett, G., Kummerfeld, J.K., Berg-Kirkpatrick, T., McCoy, D., Levchenko, K., Paxson, V., 2017. Tools for automated analysis of cybercriminal markets. In: Proceedings of 26th International World Wide Web conference.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., 2019. Language models are unsupervised multitask learners. OpenAI Blog 1 (8), 9.

Ruan, N., Deng, R., Su, C., 2020. GADM: manual fake review detection for O2O commercial platforms. Comput. Secur. 88, 101657. doi:10.1016/j.cose.2019.101657.

Salamatian, S., Huleihel, W., Beirami, A., Cohen, A., Médard, M., 2019. Why botnets work: distributed brute-force attacks need no synchronization. IEEE Trans. Inf. Forensics Secur. 14 (9), 2288–2299.

Scheitle, Q., Hohlfeld, O., Gamba, J., Jelten, J., Zimmermann, T., Strowes, S.D., Vallina-Rodriguez, N., 2018. A long way to the top: significance, structure, and stability of internet top lists. In: Proceedings of the Internet Measurement Conference 2018. Association for Computing Machinery, New York, NY, USA, pp. 78–493. doi:10.1145/3278532.3278574.

Sen, I., Aggarwal, A., Mian, S., Singh, S., Kumaraguru, P., Datta, A., 2018. Worth its weight in likes: towards detecting fake likes on Instagram. In: Proceedings of the 10th ACM Conference on Web Science. Association for Computing Machinery, New York, NY, USA, pp. 205–209. doi:10.1145/3201064.3201105.

Sharevski, F., Alsaadi, R., Jachim, P., Pieroni, E., 2022. Misinformation warnings: twitter's soft moderation effects on COVID-19 vaccine belief echoes. Comput. Secur. 114, 102577. doi:10.1016/j.cose.2021.102577.

Stringhini, G., Egele, M., Kruegel, C., Vigna, G., 2012. Poultry markets: on the underground economy of twitter followers. In: Proceedings of the 2012 ACM Workshop on Workshop on Online Social Networks. Association for Computing Machinery, New York, NY, USA, pp. 1–6. doi:10.1145/2342549.2342551.

Vallina, P., Le Pochat, V., Feal, A., Paraschiv, M., Gamba, J., Burke, T., Hohlfeld, O., Tapiador, J., Vallina-Rodriguez, N., 2020. Mis-shapes, mistakes, misfits: an analysis of domain classification services. In: Proceedings of the ACM Internet Measurement Conference. Association for Computing Machinery, New York, NY, USA, pp. 598–618. doi:10.1145/3419394.3423660.

Wang, A.H., 2010. Don't follow me: spam detection in twitter. In: 2010 International Conference on Security and Cryptography (SECRYPT), pp. 1–10.

Wang, G., Konolige, T., Wilson, C., Wang, X., Zheng, H., Zhao, B.Y., 2013. You are how you click: clickstream analysis for Sybil detection. In: 22nd USENIX Security Symposium (USENIX Security 13). USENIX Association, Washington, D.C., pp. 241–256.

Yao, Y., Viswanath, B., Cryan, J., Zheng, H., Zhao, B.Y., 2017. Automated crowdturfing attacks and defenses in online review systems. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, pp. 1143–1158.

Zarei, K., Farahbakhsh, R., Crespi, N., 2020. How impersonators exploit Instagram to generate fakeengagement? doi:10.1109/ICC40277.2020.9149431.

David Nevado-Catalán holds a M.Sc. in Cybersecurity by Universidad Carlos III de Madrid and B.Sc. degrees in Computer Science and Mathematics by the Universidad Autónoma de Madrid. His interests are in analytics, cybercrime and applied cryptography.

Sergio is Visiting Professor at Universidad Carlos III de Madrid, where he teaches various courses on cyberdefense and computer security. His research interests focus on different areas of security and privacy, including the measurement and analysis of the socio-technical factors and human aspects of cybercrime. He has published in top conferences such as WWW, IMC or RAID, and also in various international journals.

Narseo Vallina-Rodriguez is an Assistant Research Professor at IMDEA Networks and a Research Scientist at the Networking and Security team at the International Computer Science Institute (ICSI) at the University of Berkeley, USA. Narseo is also

a co-founder of AppCensus Inc. Narseo's research interests fall in the broad areas of network measurements, privacy, and mobile security. His research has been awarded with best paper awards at the 2020 IEEE Symposium on Security and Privacy (S&P), USENIX Security'19, ACM IMC'18, ACM HotMiddlebox'15, and ACM CoNEXT'14 as well as several industry grants and awards (e.g., Google Faculty Research Awards, DataTransparencyLab Grant, and Qualcomm Innovation Fellowship) and he has been the PI of several NSF, NSA, H2020, and Spanish projects. His work in the mobile security and privacy domain has influenced policy changes and security improvements in the Android platform while his research on the privacy and security risks of preinstalled Android applications has received the AEPD Emilio Aced Award and the CNIL-INRIA Privacy Protection Award, both in 2020. He is also the recipient of the IETF/IRTF Applied Networking Research Award in 2016 and the Caspar Bowden Award in 2020.

Juan Tapiador is Professor of Computer Science at Universidad Carlos III de Madrid, where he leads the Computer Security Lab. Prior to joining His research interests include binary analysis, systems security, privacy, surveillance, and cybercrime. He has served in the technical committee of conferences such as USENIX Security, ACSAC, DIMVA, ESORICS and AsiaCCS. He has been the recipient of the UC3M Early Career Award for Excellence in Research (2013), the Best Practical Paper Award at the 41st IEEE Symposium on Security and Privacy (Oakland), the CNIL-Inria 2019 Privacy Protection Prize, and the 2019 AEPD Emilio Aced Prize for Privacy Research. His work has been covered by international media, including The Times, Wired, Le Figaro, ZDNet, and The Register.