uc3m | Universidad Carlos III de Madrid

Master Degree in Cybersecurity
2019-2020

Master Thesis

“An Analysis of Fake Social Media Engagement Services”

David Nevado Catalán

Juan Manuel Estévez Tapiador
Sergio Pastrana Portillo
Madrid, 2020

An Analysis of Fake Social Media Engagement Services

David Nevado Catalán
Universidad Carlos III Madrid

Abstract

Fake engagement services allow users in Online Social Network (OSN) to illegitimately increase their online reach and boost their perceived popularity. The demand for these services by actors with social, economic and political motivations has driven the development of a vast fake engagement underground market. Prior research 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 this work we study the social media fake engagement market, with a particular focus in Social Media Management (SMM) panels. These panels are essentially reselling platforms where a large variety of these services are offered. We crawl SMM panels daily for 4 months, generating a dataset with 2.8M entries grouped in 61k different services. This dataset allows us to elaborate a catalog with all the services for sale in this market and the platforms they target. We then carry out an economic analysis of fake engagement services. This analysis complements previous research efforts and can help shed some light on the market's operation and gain a better understanding of the ecosystem as a whole.

1 Introduction

Social media is gaining an ever increasing relevance in our day to day life. In the last decade 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 presence and influence in them. The pursue of these goals has fueled the development of a vast underground social media fake engagement market. In this market we can find a wide variety of services such as Instagram followers and Facebook likes [4, 9]. These services simulate interactions that manipulate OSN's recommendation algorithms in order to increase a user reach and visibility as well as boosting their perceived online popularity.

Previous research on the topic has revealed this to be a lucrative business involving multiple actors along its supply chain. Botnets appear to be the main supplier of fake engagement services [13]. However before these services reach their final customer they are often traded and resold in several platforms [12]. This phenomenon is constatated by the proliferation of SMM panels, which are essentially platforms that act as an intermediary between suppliers and intermediate or end users. Botnets have been a cybersecurity concern for a long time and they have undergone extensive research [3, 7]. Fake engagement and inorganic interactions in OSN have also been studied from the platforms' perspective in an effort to identify and eliminate fraudulent behavior [6, 8, 10, 15, 17]. However, research on the economics and operative of the fake engagement market has been quite limited.

In this work we carry out an extensive analysis of the SMM panels market as a key element of the fake engagement ecosystem. We start by describing the dataset elaboration process: the panel crawling and the data processing pipeline. In Section 3 we provide an overview of the fake engagement ecosystem by presenting a catalog of the offered services and analyzing their variations as well as the platforms towards which they are directed. In Section 4 we analyze the prices of these services and their variations across markets. We review the related work in Section 5 and conclude with a discussion and final remarks from the obtained results in Section 6.

2 Datasets

SMM Panels. In order to gather a dataset fit for our study we first searched for popular sites were these fake engagement services are traded. These sites are commonly referred as Social Media Panels or SMM Panels. We compiled a list of these panels using two methods: (i) Manually searching in Google terms *"buy Instagram likes"*, *"buy Facebook likes"*, *"buy followers"* and getting the most popular results; (ii) browsing 2 underground forums–Hackforums and BlackHatWorld–where these services are commonly advertised. In addition, we also used an already compiled lists of 343 such panels provided in a previous study [12]. However this list was put together in 2018 and most of its panels were no longer up. Some of the panels we initially collected also went down shortly after we started crawling data, and therefore we discarded them. This indicates that SMM panels are highly volatile, and thus motivates the crawling and offline storage of the content for the analyses. Our final list is composed by 58 panels, from which we elaborate our dataset.

Crawling and data normalization. We implemented a custom web crawler to gather the services offered in these panels. The crawler visited the panels 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 on which to carry out our analysis. This step involved some manual analysis to customize the parser to each panel's structure. First, we classify each panel entry according to the target platform (Instagram, YouTube, Facebook, etc.) and the provided service (likes, followers, comments, etc.). We then check in the services' name and description the presence of a set of keywords that indicate different variations of its provision (see 3.2). Finally, all prices are converted to USD 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, which price is not expressed in the common '$ per 1000' format. 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 dataset of 2.8M records.

Service indexation. From this original dataset we generated another aggregated and smaller one. The original dataset contains temporal information about the services, however, this information is not needed for all parts of our analysis. The second dataset we produced contains each different service only once, with extra fields that indicate its duration in days, number of price changes during this period and basic metrics on its prices (mean, std, max, min and quartiles). To generate this dataset we first needed to identify different services within and across panels. This is not straightforward as it is common to find the exact same service name and description in several panels, sometimes even with the same spelling mistakes. However, even if this may indicate a common underlying service or a resell we cannot safely make this assumption. In the same way, we cannot assume that services that use different wording in their names or description are in fact different. One may be a resell of the other, or both of them may be resells of an underlying common service. Having acknowledged this lack of certainty, 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 a feature that had been removed or added. The result of this process is a dataset formed by 61k different services.

3 The fake engagement ecosystem

In this section we study the characteristics of this market. We first describe a catalog of all the services found in the panels in order to illustrate the variety and the dimension of the market at hand. We then classify these services and identify the involved social media platforms. Concretely, we define several metrics to explore which platforms and services are the most popular across this panels. Next, in Section 3.2 we make an in depth analysis of the services, presenting the most common variations and customizations found. In this section we also study the topic of geo-targeting. We analyze from which countries these services are offered, and which platforms are targeted from each of them. We conclude the section by analyzing the size of the studied panels.

3.1 Prominent services and platforms

Catalog of services. The first step in our analysis was to elaborate a full catalog of all the services offered in these panels. We identified a total of 294 different services (without considering 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 or removed from this diagram.

Exploring this catalog we can observe a great disparity in the popularity of the offered services. One the one side, a couple of dozens of popular services are present in all or most of the studied panels and within this panels we find many entries devoted to each of them. On the other side, 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 is also very common, being present in most panels and having many available customizations. We have also observed 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

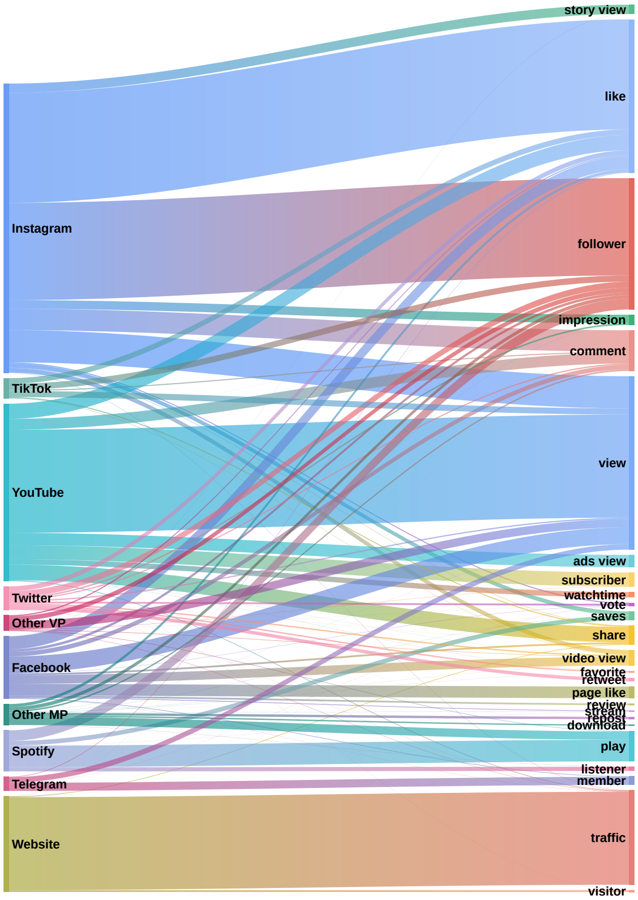

Figure 1: Catalog of the most popular services found in SMM panels. Small music/audio and video platforms have been grouped under the labels *Other MP* and *Other VP*.

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

Aside from these popular services we have found some others which are more marginal but are still interesting. These services are only present in a couple of panels and only occupy a few entries. They fall into the following categories: (i) Premium accounts for video platforms (Netflix, Disney+, HBO), music platforms (Spotify, Amazon Prime Music) and adult sites; (ii) real-looking accounts (with profile picture, followers, posts, etc.) for Instagram, Twitter and other OSN; (iii) mobile applications installs for the Apple Store and PlayStore. (iv) reviews and ratings for sites like Amazon, Google Business, LinkedIn, IMDb and TripAdvisor. In section 4.2 we take a closer look at review and rating services and premium accounts services, as they are among the most expensive services found in the studied panels.

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. These services are also mentioned in Paquet-Clouston et al. [12] and they illustrate how common reselling practices are in this market.

Popularity. In order to quantify the popularity of services and platforms in these panels, we measured 3 parameters: (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 1.

It is important to note that although these metrics combined give an idea of the popularity of the services, 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 many different locations and with many 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 (Instagram, Facebook). For video (YouTube, Twitch) and music platforms (Spotify) the most popular services are, unsurprisingly, views and plays. However, YouTube's second most popular service is again like. The popularity of these services is probably a result of two factors. On the one hand they are distinctive indicators used by the platforms to measure the quality of the content (posts, videos, songs, etc.). Therefore their manipulation can effectively be used to impact the recommendations algorithms. On the other hand, they are easier to provide by automated means (e.g. botnets) than other services such as comments or subscribers. This, as we will see in Section 4 makes them much cheaper, and thus boosts their popularity on the market.

We have used the same metrics to classify the platforms. The top 8 platforms are represented in Figure 2 ranked by daily entries and by 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 (1082 M) is by far the most targeted for its size compared to Facebook (2603 M), YouTube (2000 M) and

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

Table 1: Top 20 services in SMM panels.

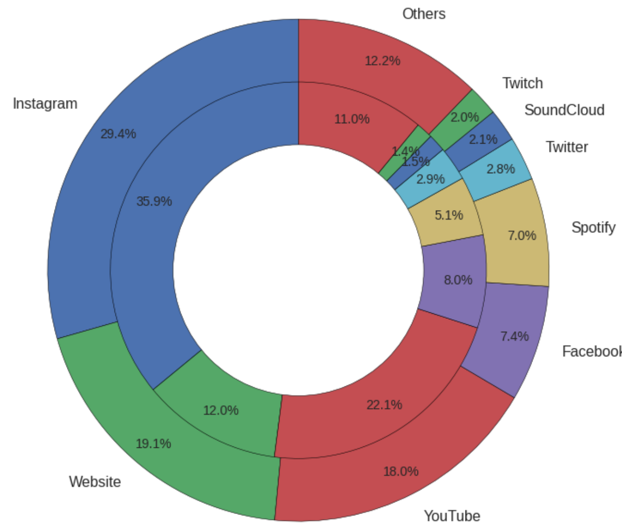

Figure 2: Top 8 targeted platforms by services in SMM panels. The inner ring represents the percentage of different services (variations) identified. The outer ring represents the percentage of mean daily entries.

Twitter (326 M)[1]. This figure also showcases the bias of the used metrics we previously mentioned, which is particularly clear for Website.

3.2 Service customization

The services advertised in the SMM panels typically have a description that elaborates on details about the service, such as its quality, the form and speed of delivery, the refund policy, and other service-specific features. For many services, we can differentiate between low quality or standard versions from improved or premium versions that come at a much higher price. 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 comments 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 find different forms of customization: more real-looking accounts and 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 identified recurring keywords. We then classified the services according to the presence of these keywords in the services names and descriptions. Unfortunately we are not able to determine up to which point these keywords reflect a real difference in the service or are just a means to make the service look more appealing. Nonetheless, we next present the most general and relevant keywords identified and later in Section 4 we will discuss their impact on the services' prices.

- **Bot / Real / Active**. These keywords are common across all services, specially Real. The use of these keywords indicate characteristics regarding the accounts used to provide services such as likes or followers. However, Real is quite broad and is often 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 keywords are also quite common. However, they do not provide much information on their own and they need to be analyzed in the context of a particular service. 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. In many cases they may not provide any real information about the service and their presence may 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 refer to a gradual de-

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

livery 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 the 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 couple of weeks or months normally. As with the Drop keyword these are very 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 Section 4). 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 [14] 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 review 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 2: Presence of geo-targeted services.

3.2.1 Geo-targeting

In addition to the discussed variations it is important to analyze location-based targeting in these services. The majority of them offer the possibility to select a specific country from which the service will be delivered. For the services that involve text–comments and reviews mainly–the language can also be chosen. However, despite this customization being available for most services it is particularly prevalent in a few of them.

In Table 2 we present the top 10 services ranked by the number of their geo-targeted variations. We observe that more than 70% of Website traffic, YouTube comments services are geo-targeted and more than 50% for Spotify plays and YouTube ads views. The other services oriented to these platforms have also a high percentage of geo-targeted services compared to those of other platforms such as Instagram.

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 would have an impact in these results. In Figure 3 we show graphs that represent the platform targeting for the services located in each of these 5 countries. These distributions percentages are presented with more detail in Table 3.

Interestingly, there are significant differences in the targeted platforms distribution between each of these 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.

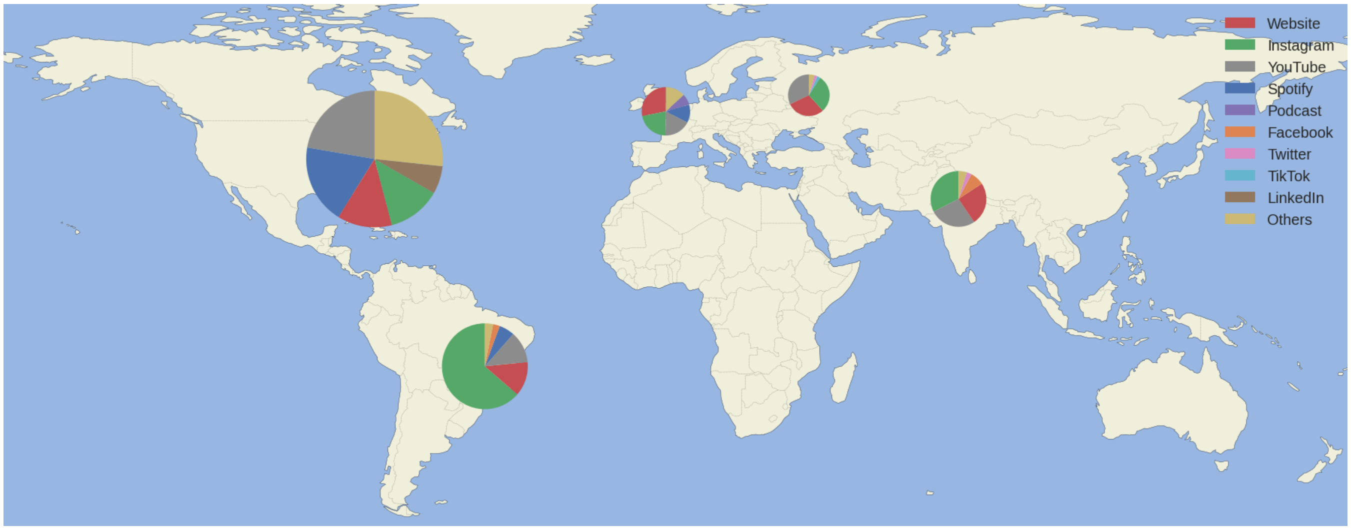

Figure 3: Target platform distribution for the top 5 countries of origin of geo-targeted services: USA, Brazil, India, Russia and UK. The pie charts are scaled by the number of services offered in each country.

	USA	Brazil	India	UK	Russia
N^{o} targeted services	2843	1785	1160	1006	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 3: Platform targeting distribution for the top 5 countries of origin of geo-targeted services.

3.3 Panels

To conclude this section we briefly analyze the panels we have used in our dataset. We have measured the size of the panels using the same metrics we used for identifying the most popular services: number of daily entries and number of different variations identified during the study period. Figure 4 represents the histogram and CDF of the panels' sizes using these two metrics. These measurements are also presented in Table 4. We observe that the most common size for a panel is found between 400 and 600 entries per day and that 75% of the panels do not exceed 1000 daily entries. In regards to the number of different services the 0-500 range is the most common and 84.5% of the panels got less than 1500.

Analyzing these two metrics side by side, we see a relatively low number of different services in relation to the size of the panel in daily entries. This indicates that the panels are not frequently updated with new services.

4 Price analysis

In this section we analyze in detail 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 impact of the customizations discussed in Section 3.2 on several services' prices.

The first step in order to get useful metrics on the services' price was to carefully filter the data to remove errors. The prices in these panels tend to have very noticeable outliers. Some of them are clearly erroneous (prices between $\$10^7$ and $\$10^{17}$) but with others it is difficult to tell (between $\$10^5$ and $\$10^7$, which corresponds to \$100 and \$10K per unit) if the high price is due to exorbitant sellers or errors, or the service is just extremely expensive.
The filtering applied consists in removing the 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. (iv) It is classified as a package or bundle. These last two conditions have the purpose of eliminating services that may be valid but are not useful for this analysis. Our objective is to analyze the price of bulk services, therefore we try to eliminate special services or packages. These special services are for example

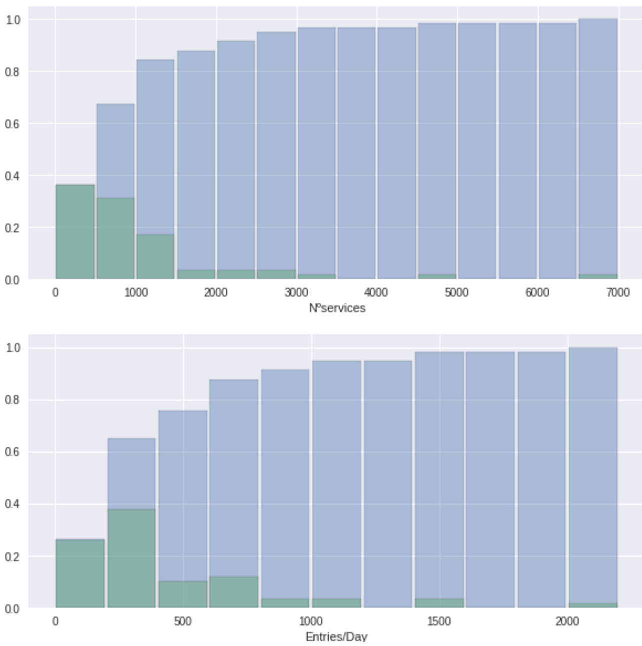

Figure 4: Panels size histograms (green) and CDF (blue) by number of total services observed in a 4 month period (top) and by number of daily services (bottom).

Total services	Nº panels	%	cumulative %
0 - 500	21	36.21	36.21
500 - 1000	18	31.03	67.24
1000 - 1500	10	17.24	84.48
1500 - 2000	2	3.45	87.93
2000 - 2500	2	3.45	91.38
2500 - 3000	2	3.45	94.83
3000+	3	5.17	100.00
Daily entries			
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 4: Distribution of the number total services of daily entries across SMM panels.

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

Table 5: Prices of top 20 most advertised services.

Instagram Followers from verified accounts, or a package (1 unit) consisting of daily comments and views for a YouTube channel during 3 months. Aggregating these services with regular Instagram followers or YouTube views would disrupt the metrics and give results that do not reflect accurately the reality of the market.

The prices for the top 20 most popular services are presented in Table 5. It is worth noting that we have used median and quartiles instead of mean and standard deviation. The reason behind this is that, despite of the filtering there are still outliers that demand the use of robust statistics. This choice is likely the reason why the results presented here differ substantially from those reported in previous research, where the mean price was used instead. However it could also be the case that this services have gotten cheaper in the last years, maybe as a result of the development of the market and the increase in the offer, but we cannot say for sure.

4.1 Prices stability

The next step in our study was to analyze the fluctuation in the different services' prices and the overall market volatility. Unfortunately it is not straightforward to measure the variability of a given service in these panels. This is due to the fact that normally, instead of modifying the price in a service it gets taken down and replaced with a new one under a different ID and a slightly different name. Out of the 61k observed services 88.6% of them never changed prices. Moreover, we checked if services that never changed their price had shorter lifespans that those that did. 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

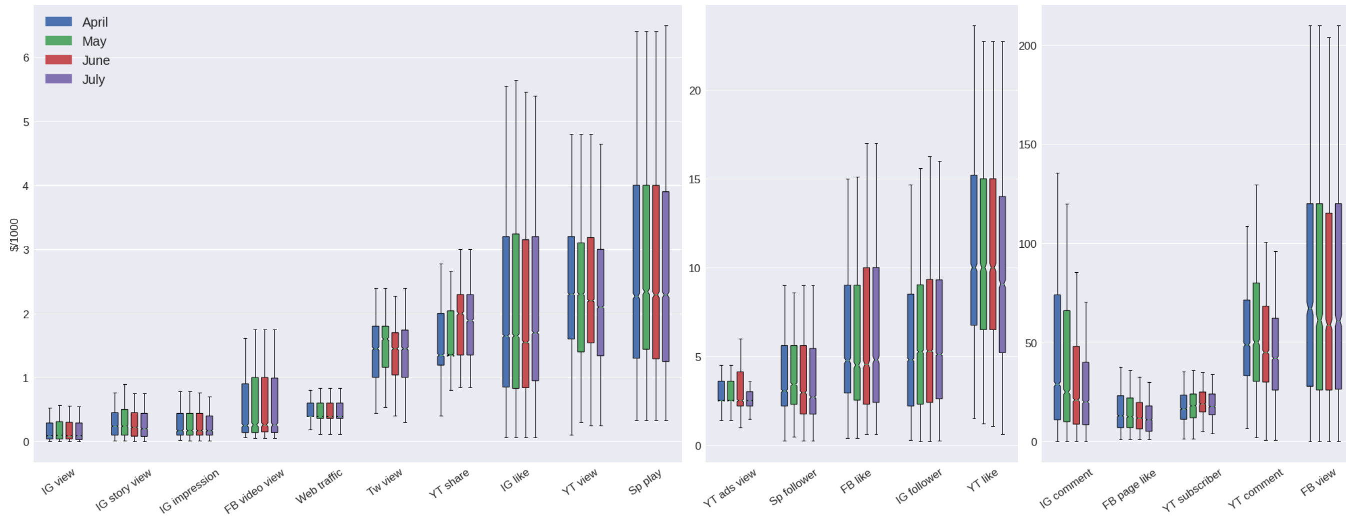

Figure 5: Prices of the top 20 most popular services during 4 months. Each group of boxes represents a different service. The boxes contain the prices between the first and third quartiles, while the whiskers extend to the higher/lower observation within one Interquartile range (IQR) of the quartiles.

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, 57.92 days. Although there are some differences between the percentage of fixed price services across the different panels they are not very significant and we concluded that this is a general practice in the market.

With this limitation we decided to study the prices evolution of aggregated services such as Instagram followers and YouTube views instead of individual variations of them. As we did with for the prices in Table 5 we first filtered out outliers and erroneous values using the conditions previously described. Then we evaluated the price of the most popular services during the first 10 days of April, May, June and July. The results are presented in Figure 5. Each group of boxes represents a service, and each color within the group represents one of the studied months. The boxes represent the prices between the first and third quartiles, the notch in the boxes represents the median price and the whiskers extend up to the highest and lower observation within one IQR of the quartiles. The graph has been divided in three and the services have been rearranged in order to use different scales that allow for a useful visualization of the prices.

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. Moreover, we cannot safely attribute this descending trend to a real lowering of the prices. We have to keep in mind that, due to the employed methodology the disappearance of high end variations or appearance of low end ones would also generate this effect on the results. In fact the most important observation we can derive from this visualization is the enormous range at which a service can be priced. 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. With Facebook views we observe the same effect, having prices below a dollar and above $200. One of the reasons for such a high variance is the fact that we are aggregating different variations of the services. It is to be expected that geo-targeted custom comments for YouTube are more expensive than random non geo-targeted ones. However this is not the only reason. 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.

4.2 High-end services

Besides the study of the most common services, it is interesting to investigate what are the most expensive and highest quality services offered in these panels. As we explained at the beginning of the section, the presence of high price outliers forced us to take a very manual approach for this task. We ranked the median price of the services to have an initial list of the most expensive services and then we manually reviewed them. Fortunately these expensive services are much less common that the ones we have been dealing with up until this point, and therefore manually reviewing them is an assumable task.

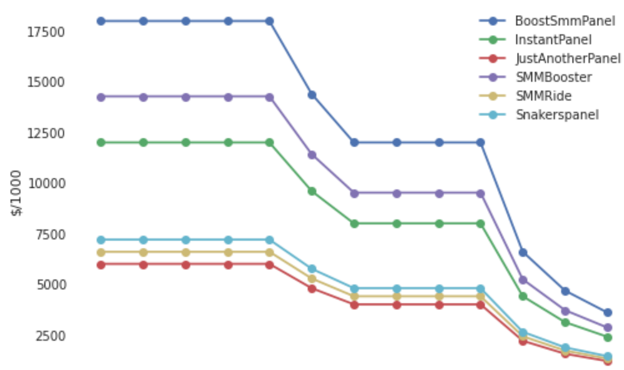

Figure 6: Google Business review prices for 13 exactly equal services present in 6 different panels. Each line represents the prices of the different services in one panel. The points in the same vertical axis represent prices for the same services in the 6 different panels.

The most expensive services found mainly consist of: (i) review and rating services, (ii) accounts for subscription services or paid services and (iii) the very high end or premium versions of typically cheap services. We present the findings of each of this categories separately:

- **Reviews and ratings**
  This 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 were found in 21 panels, their prices range from $1.4 and $18.0 per review and the available orders range from 1 to 100 reviews (although one service offered up to 1.5k reviews). When we tried to find out which variables determined the difference in price we discovered that several services were offered with the exact same name and description in 6 different panels but at significantly different prices. In Figure 6 we can appreciate the difference of price of the same services across the panels. More importantly, the matching shapes reflect that these services either have an underlying common services or are a resells of one another. It is also worth noting that the difference in prices 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 were much less common, appearing only in 2 panels. In one of the panels the service description did not have much information besides stating it provided a custom review and its price oscillated between $0.25 and $0.72 per review. In the other panel the service had 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 granted a refill policy in case the reviews and rating were taken down. This service was priced at $5 per review.

  With **IMDb** we found that in general all of its services were quite expensive. Votes, which would be the equivalent of a rating in other platforms went 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.

- **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 [2]) 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 he could potentially make money just from the service alone. The percentage of the subscription fee received by streamers is not fixed. For example, if he gest 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.

- **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

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

boost packages. In Table 5 we can see that 1000 Instagram likes usually go for around $1.43. However we contrast this with a bundle that offers 1000 *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. However, none of this is confirmed and there are those who claim they are a fraud.

4.3 Low-end services

In contrast to the services we just described, we now briefly review the cheapest services found in these panels. First among these services we find the free services. Many services offer a free service in order for clients to check the quality of the service and persuade them to make an investment by buying other services. 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.

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 video views for Facebook. 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.

4.4 Customizations price impact

We conclude the prices analysis by studying the price variations of several services in function of the customizations identified in section 3.2. 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 7, 8 and 9. 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 7. 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.

- **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 8. 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

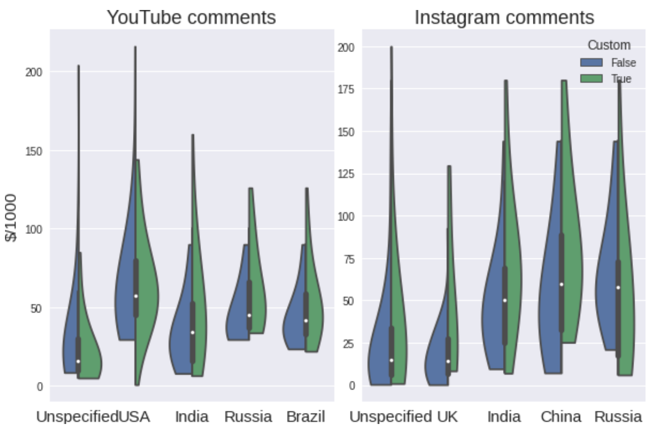

Figure 7: Prices distribution of YouTube comments and Instagram comments attending to geo-targeting and the *Custom* variation. Each individual violin plot represents the prices for a given region (horizontal axis). The left, blue side represents the distribution for non-custom comments, while the green, right side of the violin represents the distribution for the custom comments.

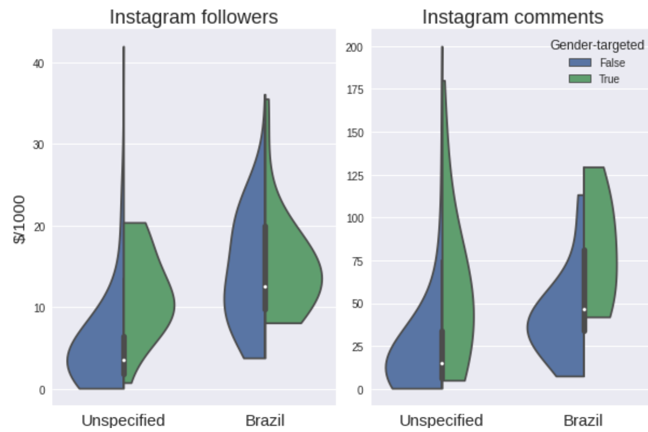

Figure 8: Prices distribution of Instagram followers and comments attending to geo-targeting and gender-targeting. Each individual violin plot represents the prices for a given region (horizontal axis). The left, blue side represents the distribution for non gender-targeted services, while the green, right side of the violin represents the distribution for specifically male or female services.

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 in this case was the *Refill* option, which is very common in many services.
  The results are presented in Figure 9 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 desirable 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.

5 Related work

Fake engagement in OSN is a problem that has gained relevance in the last decade. During this period many studies have been carried out on how malicious actors abuse social media platforms for diverse purposes. Among these purposes we

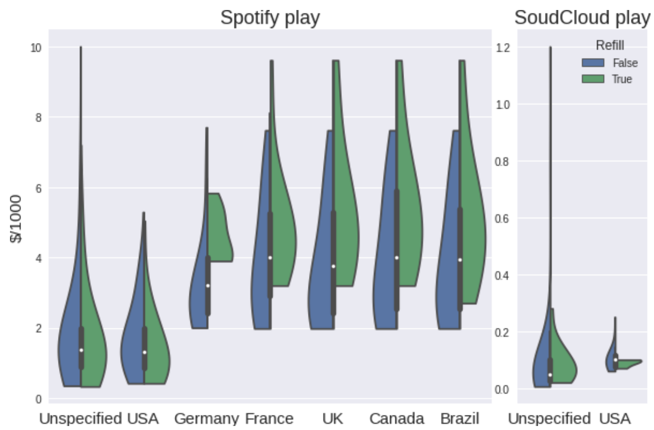

Figure 9: Prices distribution of Spotify and SoundCloud plays attending to geo-targeting and the *Refill* variation. Each individual violin plot represents the prices for a given region (horizontal axis). The left, blue side represents the distribution for services that are not advertised with *Refill*, while the green, right side of the violin represents the distribution for those which are.

find different forms of marketing fraud such as fake product reviews [10, 11] and spamming campaigns [5], but also political and social manipulation [1]. 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 or behavior in OSN. In particular fake engagement detection studies have been carried out for the most prominent platforms such as Facebook [4, 5], Twitter [16, 17], YouTube [8] and Instagram [15,19]. 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 [2, 6, 18].

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 [13] 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 [12] with a heavier focus on the market aspects of the SMF supply chain. 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.

6 Discussion and conclusions

To complement previous research on social media fake engagement we have carried out a study of the market that provides these services. 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. Within this dataset we have identified 294 different services targeting 59 platforms including OSN, review services, video and music platforms, etc. We observe that many of these services are offered with a great variety of customizations that allow buyers to select features such as the quality of the service, the speed of delivery and the country of origin. The granularity of these customizations and the richness of the catalog hint at the existence of substantial infrastructure underlying these services.

Market analysis. The first observation drawn from our analysis is that the prices we have found for these services are significantly lower than those reported in previous studies. For example De Cristofaro et al. [4] 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. [13]. These differences are undoubtedly 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 [12]. As Paquet-Clouston et al. [13] 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.

Future work We plan to extend our research by crawling further SMM panels during a period of time that allow us to study long-term trends and the evolution of the market. Along these lines 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.

Research on attribution could be carried out by identifying common actors across the supply chain and across markets. This can be attempted by using domain classifiers on the panels and by investigating underground communities such as

Hackformus and BlackHatWorld. In particular, researching Social Media threads on these communities could also be used to estimate the volume of this market as well as to gain insights on the operation of the services suppliers and the motivations of their customers. Moreover, it can help determine if or how this market is tied to other fraud schemes and illicit activities.

References

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

[2] 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.

[3] 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.

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

[5] 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.

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

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.

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

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

[14] 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.

[15] 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.

[16] 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.

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

[18] 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.

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