COUNTERING
THE MALICIOUS
USE OF
SOCIAL MEDIA

THE BLACK MARKET FOR SOCIAL MEDIA MANIPULATION

NATO StratCom COE
Singularex

ISBN 978-9934-564-31-4

ISBN: 978-9934-564-31-4
Authors: NATO StratCom COE
Research: Singularex
Project manager: Sebastian Bay
Text editor: Anna Reynolds
Design: Kārlis Ulmanis

Riga, November 2018

NATO STRATCOM COE
11b Kalciema iela
Riga LV1048, Latvia
www.stratcomcoe.org
Facebook/stratcomcoe
Twitter: @stratcomcoe

Singularex is a Social Media Intelligence and Analytics company based in Kharkiv, Ukraine.
Website: www.singularex.com
Email: hello@singularex.com

This publication does not represent the opinions or policies of NATO.

The views expressed here are solely those of the author in his private capacity and do not in any way represent the views of NATO StratCom COE. NATO StratCom COE does not take responsibility for the views of authors expressed in their articles.

> ” Social media manipulation is undermining democracy, but it is also slowly undermining the social media business model.

Introduction

Around the turn of the decade, when the popularity of social media sites was really beginning to take off, few people noticed a secretly burgeoning trend — some users were artificially inflating the number of followers they had on social media to reap financial benefits. Even fewer noticed that organisations such as the Internet Research Agency were exploiting these new techniques for political gain. Only when this innovation in information warfare was deployed against Ukraine in 2014 did the world finally become aware of a practice that has now exploded into federal indictments,[1] congressional hearings,[2] and a European Union Code of Practice on Disinformation.[3]

At the heart of this practice, weaponised by states and opportunists alike, is a flourishing black market where buyers and sellers meet to trade in clicks, likes, and shares. NATO Strategic Communications Centre of Excellence in Latvia and the Ukrainian Social Media analytics company Singularex have undertaken a joint venture to map the online market for social media manipulation tools and services. We have scanned the dark web and tracked down sellers, buyers, and victims in an attempt to understand what a potential customer, wishing to wage information warfare, can purchase online.

Understanding the market

The White Market

Anyone interested in promoting content on social media has a few basic **social media metrics** to work with — followers, views, likes, comments, and shares. Social media platforms have developed a range of paid advertising tools users can purchase to promote their posts beyond their own set of followers. More advanced users work with advertising agencies to generate reach, often through selected popular users referred to as influencers.

And just as newspapers with many subscribers can charge more for advertising than newspapers with fewer subscribers, popular influencers can charge more for their services. Reports indicate that while an influencer with 100,000 followers might earn $2,000 for a promotional tweet, an influencer with a million followers can earn as much as $20,000 per promotional tweet.[4]

This social media ecosystem has many vulnerabilities, both technical and cognitive, that can be exploited using influence techniques such as social proof, flaming, deceptive identities, and bots.[5] The most potent techniques are now packaged and sold on the black market of social media manipulation.

The Black Market

The black market of social media manipulation is divided into three overlapping categories — the easily accessible open market, the dark web, and the offline word-of-mouth market.

These are all unofficial channels that provide users the opportunity to buy likes, shares, comments, subscribers, and accounts. Services range from vending machines in Russia,[6] to sophisticated software and highly-managed information manipulation services offering extensive customer support.

This industry exists in contradiction to the platforms' Terms of Service, which don't allow likes, comments, or subscribers to be bought; yet this is still legal in most countries. A simple search on your favourite search engine will produce a long list of providers offering to sell tools and services for social media manipulation. In fact, our research team encountered a number of paid ads on both Google and Bing promoting social media manipulation tools.

These services operate in a grey zone and actively develop ways to manipulate social

> Our research team encountered a number of paid ads on both Google and Bing promoting social media manipulation tools.

media platforms. The black market services are unregulated, unreliable, and come with a number of disadvantages:

1. **Technical risks** – social networks monitor and block users who circumvent their terms of service.

2. **Reputational risks** – reputable companies or influencers risk ruining their reputations if they are discovered using black market manipulation tools.

3. **High risk interactions with suppliers** – as the supply side operates in a grey zone, some sellers extort their customers and many of the advertised services are fraudulent.

4. **Unsustainable** – although there are a number of short-term benefits, an artificially inflated business model is unsustainable in the long run.

Although there are a number of disadvantages to using black-market tools, the advantages are clearly attractive given the availability of services.

1. **Easy to avoid detection** – although social media companies work to reduce inauthentic activity on their platforms, they are not keeping pace with the technological advancements of the black market suppliers, which means that it still easy to avoid detection by both social media companies and the public.[7]

2. **Simple product** – the products are easy to understand and market, and have a simple pricing structure.

3. **Multiple choices** – customers have a wide range of choices, and suppliers are constantly developing new products.

4. **Large potential reward** – successful social media manipulation can result in substantial short-term financial and political rewards.

The Dark Web

The 'dark web'[8] has become a safe haven for illegal, and often immoral activity. A Tor browser can give anyone with an Internet connection access to the dark web. But the dark web is difficult to navigate due to a lack of functional search engines and reliable services. As we began our investigation, we assumed that the most advanced services would be found there. However, after having searched extensively we concluded that although social media manipulation tools and services can be found on dark webforums and websites, prices are significantly higher than on the open Internet.

Although the dark web offers more advanced forms of social media manipulation than the open Internet, including tailored hacking of targeted accounts, we found no evidence of a market for advanced influence operations there.

The easiest way to explain the lack of services for social media manipulation on the dark web is that there is no reason to hide if you can advertise your products in the open. An abundance of suppliers use traceable domains to advertise their services, they process payments openly, and even place ads on popular search engines to attract customers.

The Black Market: Tools

This section provides information regarding the range, pricing, and types of black market social media manipulation services available online.

Fake accounts

Fake or compromised accounts serve as the basic tool for gaining access to a social media platform, a requirement for being able to manipulate the platforms with fake likes, views, comments etc. The price of purchasing such accounts largely depends on the level of network security for a given platform — the harder it is to register and maintain an account, the higher the price. The following characteristics affect the price of a fake account:

Account type
As in any other market, the cost of creating/maintaining a service influences its price. Accounts that can be registered

automatically are the cheapest but also the least reliable, while accounts that are registered manually are more reliable but also more expensive, as are genuine accounts that have been hacked.

- **Automatically registered accounts** — these accounts are registered using programs designed for the purpose and are usually not very sophisticated.

- **Manually registered accounts** — these accounts are registered by human operators and can often be padded with custom content.

- **Hacked accounts** — these are accounts that belong to real people whose credentials have been hacked or stolen. Because of their genuine historical record hacked accounts are better at circumventing platform safeguards. Hacked accounts are sometimes seen as consumables for collecting information about users, or as part of a short-term targeting scheme, because the legitimate owner can retake control of the account. However, since the real owners sometimes lack the will and ability to regain access, such accounts sometimes remain under the control of malicious users.

Account verification

Social media accounts can be verified by phone and by e-mail. Verified accounts are more expensive, but less likely to be blocked. The quality of the verification process is also a price factor — manual verification is more reliable, while automatically verified accounts are at greater risk of being blocked and the work done through these accounts (likes, comments etc.) deleted.

Content

The more content provided for a maliciously created account the more realistic it will look. Depending on the task an account is required to perform, a higher level of perceived authenticity can be more or less desirable. A basic account is usually sold in four set categories:

a. accounts with no content
b. accounts with a profile picture
c. accounts with a profile picture and a few photos
d. accounts with a profile picture, photos, and a range of posts

In addition to these general bulk categories more advanced services offer custom-made, highly-developed accounts that are nearly impossible to differentiate from genuine accounts.

Account age

The history of an account, its age, is important for trust and credibility. The market offers accounts with a wide range of ages — from days to 7+ years old. The older the account, the more expensive it is, as an old account is less likely to be detected as malicious.

Manipulating Social Metrics

A large number of resellers offer social metrics manipulation (likes, comments, shares, views, followers, etc.) for all major platforms. Social metrics manipulation is accomplished using a range of tools:

a. fake automated accounts
b. special 'freelance' platforms (usually employing people from developing countries)
c. 'likes exchanges' where users are offered likes in return for likes
d. malicious software acting without a user's permission, e.g. through browser extensions or malware

Many suppliers also offer 'trending content' services on platforms such as YouTube. Some suppliers also offer manipulation of specified targets such as web panels, surveys, or recommendation sites in order to manipulate the outcome of a political survey or smearing a business competitor etc.

The Black Market: Segments

The price and quality of social media manipulation services can differ greatly depending on the time and efforts sellers need to produce the manipulation. We mapped a large number of social media manipulation services and were able to distinguish between three different price/ quality segments.

3rd category: The low-end segment provides the cheapest products with the lowest quality, such as automatically registered, but unverified accounts. These accounts have no content or profile picture, and all or most subscribers are bots. Low-end likes and comments are also created automatically by bots. There is a high probability that these accounts are identified and blocked, but as research by the NATO StratCom COE has shown, many simple bots on Twitter often survive long enough to enable them to produce upwards of 5000 posts, which still allows them to serve specific purposes.[9] These low-end segment products are easy for potential buyers to find using a standard search engine.

„ The social media manipulation market is largely dysfunctional in terms of cost and quality, and supply and demand.

2nd category: The mid-range segment consists of accounts that are mostly registered automatically. However, these accounts cost more because they are verified and contain basic user information reducing the risk that they are quickly identified and blocked. Tasks executed from these accounts are often performed automatically, but are also sometimes managed in bulk by human operators.The mid-range segment is harder for the potential buyer to find, and buyers may struggle to distinguish trustworthy service providers from those selling low-end services at a higher price. Navigating this market requires experience.

1st category: The high-end segment consists of accounts that have real friends and subscribers, content, and unique profile pictures. They are verified by phone/email and have been registered for at least one year. Such accounts are often genuine, but hacked, or manually registered and populated over time. Likes are added by real users and comments are written manually. These accounts can be tailored to specific demographics for a target audience with variables such as country, gender, and age. Subscribers to high-end accounts are often real users. Suppliers use all the tricks of their trade to create the appearance of real user activity so as not to arouse the suspicions of security systems and reduce the probability of blocking. The high-end segment is difficult to find without a personal referral from someone with experience in the market.

Comparative analysis of prices

Our analysis of the comparative cost of services in selected segments for various social networks (Fig. 1) reveals a number of patterns. For YouTube there is a clear relationship between price and quality — the higher the price the higher the quality. For Facebook likes are cheaper than subscribers, but for Instagram the opposite is true — likes tend to be more expensive than subscribers.

The overall pattern, however, is that the expected linear relationship between quality and price can't be seen. Normally the price of a service is related to the cost of the production. For the social media manipulation market this often isn't the case — the price of a low-quality service can be as high as the price for a high quality service, while different providers offer equivalent services at hugely different prices. This indicates that the market is largely dysfunctional in terms of cost and quality, and supply and demand.

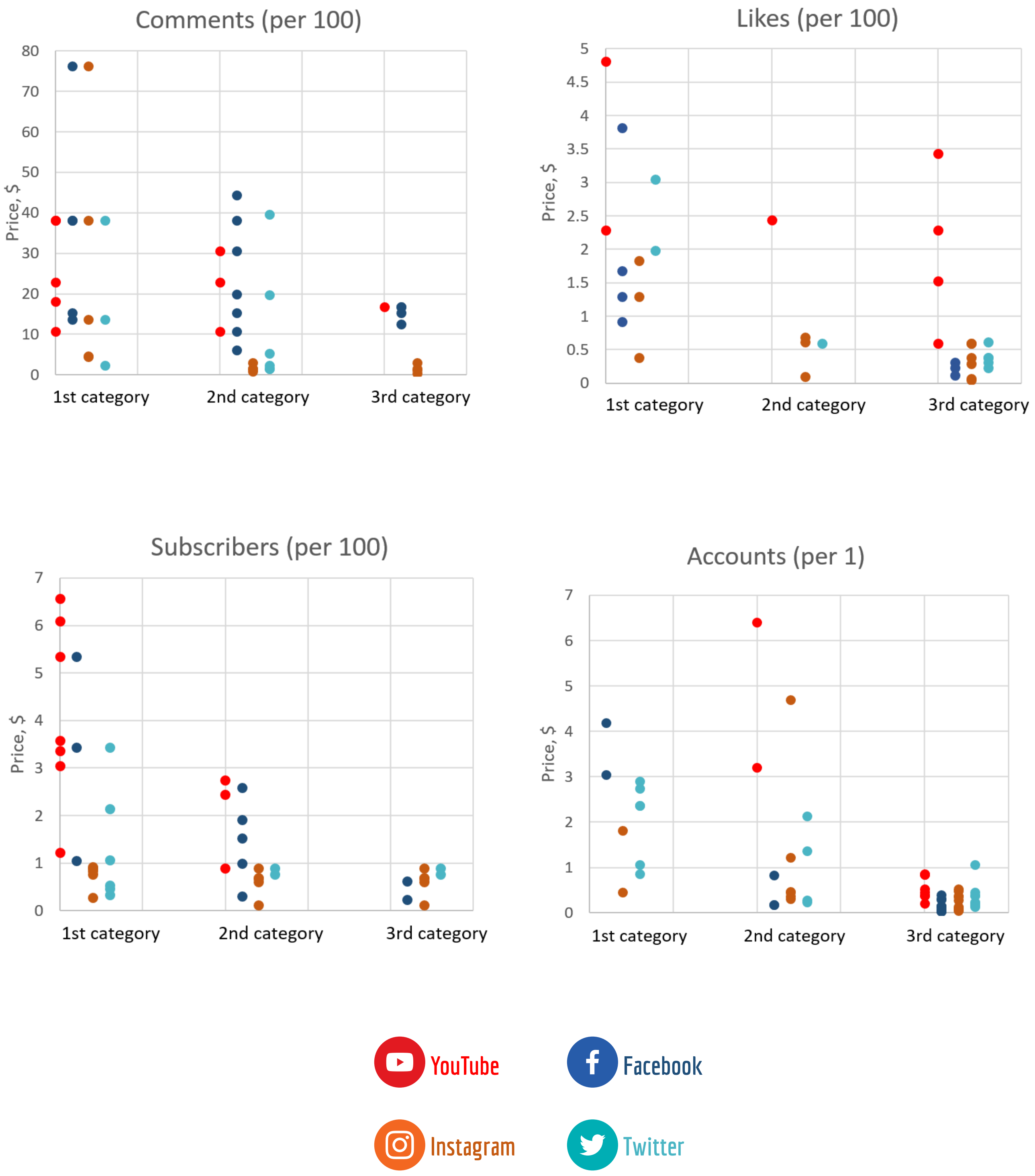


Figure 1. Comparative analysis of the price of subscribers, comments, likes, and accounts as offered by different service providers. Each dot represents the price of a specific service from a specific service provider.

Automation software

Suppliers selling mid- to low-range services use automation software. To further our understanding of the tools available we conducted an in-depth assessment of one of the most popular software packages for social media manipulation. This software package is marketed from a public web page that offers a comprehensive knowledge-base, video tutorials, and 24/7 support to help their customers use their software as effectively as possible. The development team consists of about 20 people and the business has an official account at a Russian bank. Most of their marketing activities are aimed at the private market.

This software package comes with two modes — an 'assist mode' and a 'technical mode'. The 'assist mode' allows the purchaser to ensure that the account controlled by the software operates within credible parameters to keep the account from being banned. It is likely that developers have reverse engineered the social media companies' defense protocols to increase the survivability of the assisted accounts. The 'technical mode' is used to maintain and control accounts, as well as to gather information (usually from competitors' accounts).

The technical features of the software package include:

- Account registration
- Data collection from target social media accounts
- Assessment of collected data based on meta-data and search queries
- Execution of bulk subscriptions, likes, and follows — the software can predict the likelihood of a genuine account liking or following another account in return based on their historical behavior
- Posting comments en masse and mass messaging; a large number of automated accounts and long waiting periods are required if such operations are to avoid detection by social media platforms

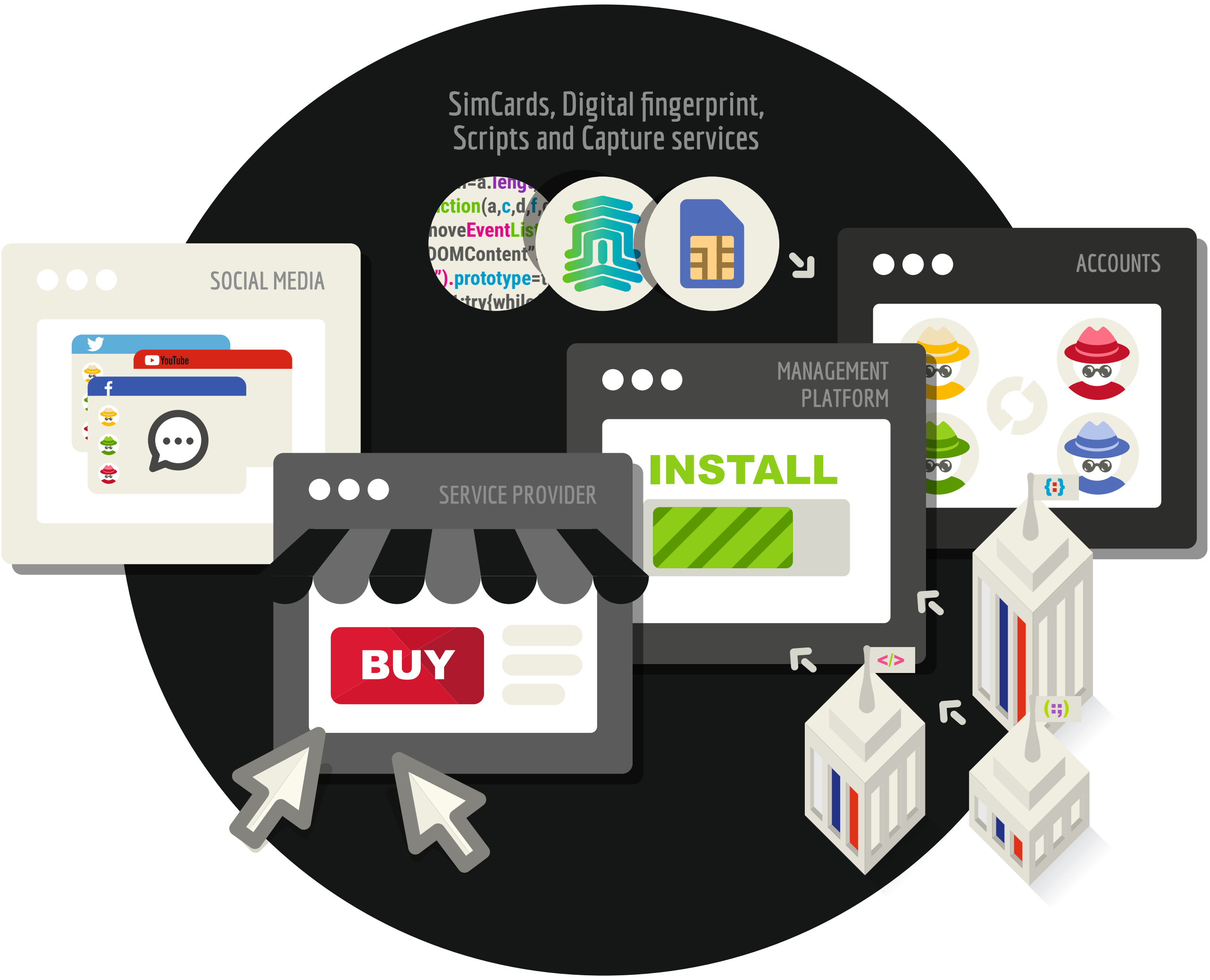
Such software packages rely on a number of subcontractors. Some of these subcontractors provide accounts, while others provide the infrastructure needed to register, verify, and maintain them, and to reduce the risk of detection, such as mobile proxies and virtual sim-cards, and cloud hosting.

We identified two other developers who provide software of similar quality and scope (with 24/7 support and a wide range of functions) and a number of small-scale providers who offer cheaper and simpler services. Although the service providers we have been able to identify are based in a variety of countries, the three main software developers and the bulk of infrastructure maintenance services are Russian entities and many of them operate openly from Russia.

While it is possible for individual users to install and operate these software packages, it requires substantial know-how. To bridge the knowledge-gap many service providers and re-sellers offer services resembling a standard online shopping experience. This makes it easy for the general population to buy social media manipulation services.

The industry as a whole consists of several layers — software development, infrastructure maintenance, and the provision of technical services.

From a customer's point of view the automation tools used for the low-end and mid-range segments are sufficient for most commercial goals, but to solve more complex problems high-end, manual services are needed.



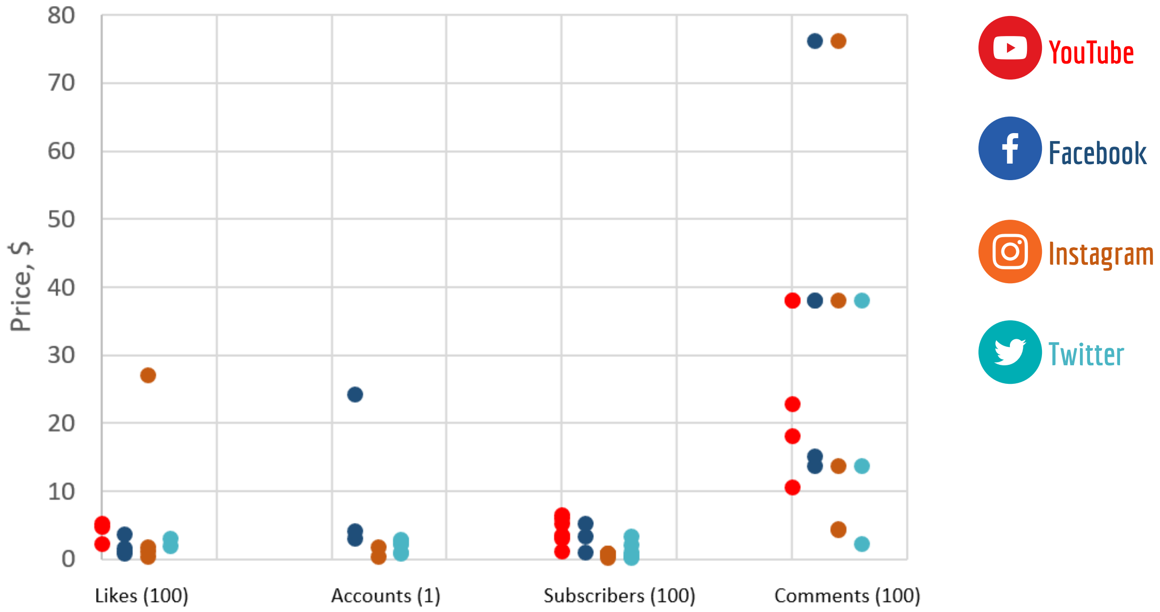


Fig. 3. High-end segment prices

High-end Services – Price range

An overall assessment of the high-end services segment (Fig. 3) shows that Instagram manipulation is the cheapest and YouTube manipulation is the most expensive. Boosting numbers of subscribers or followers is the cheapest service, while generating comments is the most expensive.

The price difference between platforms can likely be attributed to the difficulty in providing the services, however financial gain is also a possible reason as ad-platforms such as YouTube provide immediate financial rewards to users with artificially boosted performance ratings. The wide range of prices suggests that the market is ineffective due to limited sales channels, legal restrictions, and lack of knowledge among buyers.

Freelance platforms

While organisations with in-house staff provide some of the high-end services, such work may also be outsourced to freelancers who perform simple tasks such as liking, commenting, reposting, reporting, voting in polls, etc., with minimum risk to the service provider or the accounts supported. The simplest tasks are often performed by individuals willing to work for less than $1 per hour.

Top-tier freelance platforms, such as Upwork, are also involved in social media manipulation. Researchers have found requests posted in Upwork for people willing to write hyper-partisan news stories. Writers are paid anywhere from $1.50 to $50 dollars per story, with some employers expecting up to ten stories a day from each writer.[10]

DDoS

Distributed Denial-of-Service (DdoS) attacks are used to take websites or servers offline at critical times. Such attacks are often used in ransom schemes and influence operations. The price of a DDoS attack is relatively small, starting from $5 for attacks on small websites lasting a few minutes to upwards of $200 dollars for large-scale attacks on small- to medium size websites. We were unable to find any service providers offering large-scale DDoS attacks on major websites, indicating that such services aren't sold in the open.

DDoS 2.0

Since the majority of online conversations have moved to social media platforms, the number of traditional DDoS attacks has decreased. It is difficult to mount a successful DDoS attack on Facebook or Google because of the sheer volume of ordinary traffic. As a result a new form of DDoS attacks has emerged. Coordinated attacks are now staged on pages inside social media platforms using automated or manual accounts that falsely report accounts or post triggering the platforms to take down content or ban accounts for moderation. Although the effect is usually temporary, this new form of directed attack achieves much the same purpose as traditional DDoS attacks once did. This type of attack is used to curtail the work of civil rights activists, independent journalists, etc. by attacking their social media accounts to preassure them to stop using the platforms.

The Black Market: Campaign Management

To further our understanding of high-end social media manipulation services, we found a service provider who agreed to an interview. This person described an industry that distinguishes itself from automated social media manipulation by providing start-to-finish campaign management services, rather than simple clicks and likes. According to our interviewee, clients are usually politicians or businesses that have no qualms about manipulating social media to achieve their desired end-state.

High-end service providers combine automated and personalised techniques and use a number of information distribution channels such as blogging, planted news stories, and attempts to influence journalists, to achieve significant effect. High-end campaigns also combine open social media advertising tools with malicious activity to reach a wider audience.

For a campaign to succeed it is important to manufacture impact on conventional media. Therefore, high-end service providers spend a great deal of effort on creating content that will get picked up by traditional media, but they also promote curated news stories online and suppress negative news coverage.

A large campaign would includes automated account management, the creation and promotion of specified content, as well as manipulation of real-life events to generate online and offline news coverage.

Conclusions

There is a large, vibrant online market for buyers and sellers of tools and services for social media manipulation. Some customers want more likes on their photos, some want to profit financially at the expense of the ad industry, and others want to influence the outcome of elections. Regardless of the end goal, the tools used are much the same.

Social media manipulation is undermining democracy, but it is also slowly undermining the social media business model. During the past year, the social media giants have committed to better protect their platforms. Anecdotal evidence suggests this has led to an increase in the cost of manipulation tools and services. Despite the commitments of online platforms and leading social networks, codified in the EU Code of Practice on Disinformation, this study shows it is still surprisingly cheap and easy to manipulate social media.

Our research resulted in four surprising conclusions.

- First, we were impressed by **the scale of the black-market infrastructure** for developing and maintaining social metric manipulation software, generating fictitious accounts, and providing mobile proxies and solutions for SMS activation.

- Second, we were struck by the **openness of this industry.** This is no shadowy underworld, it's an open and accessible marketplace most users can find with little effort through any search engine.

- Third, we were surprised to find **service providers advertising on Google and Bing**. Providers trafficking in YouTube manipulation services buy ads from Google—the owner of YouTube—and fearlessly promote their services in public.

- Lastly, we were intrigued to learn that **Russian service providers seem to dominate the social media manipulation market.** Virtually all of the major software and infrastructure providers we were able to identify were of Russian origin.

The dark web—it turns out—is the least of our worries.

Endnotes

1 ”Internet Research Agency Indictment | Department of Justice”. Accessed 04 December 2018. https://www.justice.gov/file/1035477.

2 ”United States Senate Committee on the Judiciary”. Accessed 04 December 2018. https://www.judiciary.senate.gov/download/04-10-18-zuckerberg-testimony.

3 ”Code of Practice on Disinformation”. Digital Single Market. Accessed 04 December 2018. https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation.

4 Confessore, Nicholas, Gabriel J. X. Dance, Rich Harris, och Mark Hansen. ”The Follower Factory”. The New York Times, 27 January 2018, avs. Technology. ttps://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html

5 Pamment, James, Howard Nothhaft, and Alicia Fjällhed. “Countering Information Influence Activities: The State of the Art.” (2018).

6 Matsakis, Louise. ”This Russian Vending Machine Will Sell You Fake Instagram Likes”. *Motherboard* (blog), 06 June 2017. https://motherboard.vice.com/en_us/article/xw8yv3/russian-vending-machine-fake-instagram-likes.

7 ”Reducing Inauthentic Activity on Instagram”. *Instagram* (blog), 19 November 2018. https://instagram-press.com/blog/2018/11/19/reducing-inauthentic-activity-on-instagram.

8 The dark web is the World Wide Web content that exists on darknets, overlay networks that use the Internet but require specific software, configurations, or authorization to access. ”Dark Web”. *Wikipedia*, 04 December 2018. https://en.wikipedia.org/w/index.php?title=Dark_web&oldid=871937241.

9 Fredheim, Rolf. ”Robotrolling 4/2018”. NATO StratCom COE, November 2018. https://www.stratcomcoe.org/robotrolling-20184.

10 @DFRLab. ”INFLUENCE FOR SALE: Who Writes Your Hyperpartisan News”. *Medium* (blog), 05 July 2017. https://medium.com/@DFRLab/influence-for-sale-who-writes-your-hyperpartisan-news-8ba64ddafd58.

Prepared and published by the

NATO STRATEGIC COMMUNICATIONS
CENTRE OF EXCELLENCE

The NATO Strategic Communications Centre of Excellence (NATO StratCom COE) is a NATO accredited multi-national organisation that conducts research, publishes studies, and provides strategic communications training for government and military personnel. Our mission is to make a positive contribution to Alliance's understanding of strategic communications and to facilitate accurate, appropriate, and timely communication among its members as objectives and roles emerge and evolve in the rapidly changing information environment.

Operating since 2014, we have carried out significant research enhancing NATO nations' situational awareness of the information environment and have contributed to exercises and trainings with subject matter expertise.

www.stratcomcoe.org | @stratcomcoe | info@stratcomcoe.org