Oligopoly of Big-Tech companies despite global mistrust

Mikołaj Artur Dębicki
Computer Science, NYUAD
mikolaj.debicki@nyu.edu

Advised by: Yasir Zaki, Talal Rahwan

ABSTRACT

Have you ever discussed a random topic with you friends only to see it in your Facebook ads on the next day? With the oligopoly of the Big-Tech companies the privacy of regular users has been continually breached, yet the companies like Google or Apple are still thriving. In our work we try to challenge the two phenomenons which we believe contribute the most significantly to that situation.

First, we conducted a worldwide survey and verified what people actually think about the Big-Tech companies and how trustful they are towards them. For narrowing down the scope of the research we decided on Apple, Facebook, Google and Huawei as the companies which we investigated. We performed global surveys, asking respondents from 12 countries all around the globe.

Second, we run a practical study in order to verify whether the Facebook application is indeed recording the audio of their users without social awareness. In order to answer the research question of "Does the oligopoly of Big-Tech companies influence the privacy of their users in any negative way" we went through robust experiments, testing various smartphone parameters such as the battery consumption.

KEYWORDS

Facebook, Apple, Google, Huawei, privacy, terms of use

Reference Format:
Mikołaj Artur Dębicki. 2022. Oligopoly of Big-Tech companies despite global mistrust. In *NYUAD Capstone Project 2 Reports, Spring 2022, Abu Dhabi, UAE.* 6 pages.

1 INTRODUCTION

Smartphones are everywhere nowadays. Whether we discuss work related scenarios or leisure activities their use became an inevitable part of our lives. With such powerful devices at the palms of our hands is there anything which they cannot do? That kind of thinking leads to suspicion when it comes to technology. Especially, when the ads begin to appear not so randomly on one's news feed. Countless news articles mentioned the possibility of Apple, Facebook, Google or Huawei secretly abusing their users' privacy but actually there have been not so many proper studies conducted on the matter [3]. We hear a lot about the general suspicion of the public but the actual data on that is also not readily available. If we extend our study's reach over the United States' population the data even ceases to exist.

Therefore, in this study we first conducted a worldwide survey, spanning over 12 countries and then also performed robust physical experiments in order to investigate whether Facebook really is listening to our conversations. This allowed us to verify whether the worldwide public really believes that Big-Tech companies are recording them and whether that is a genuine fact in itself.

2 RELATED WORK

As mentioned before, not much research has been done in the field yet. In fact, there have only been two formal studies conducted on the issue. Both provide valuable insights, yet miss on some critical components and scientific methods. Also, none of them takes into account the actual social perception of the issue.

2.1 Wandera study

The most notable study so far was conducted by a United Kingdom based security company Wandera [5]. Their goal was to create a report on whether the popular Big-Tech companies actually record our conversations. During the experiment they have played a loop of pet food ads for 30 minutes over a time span of 3 days, where an iPhone and a

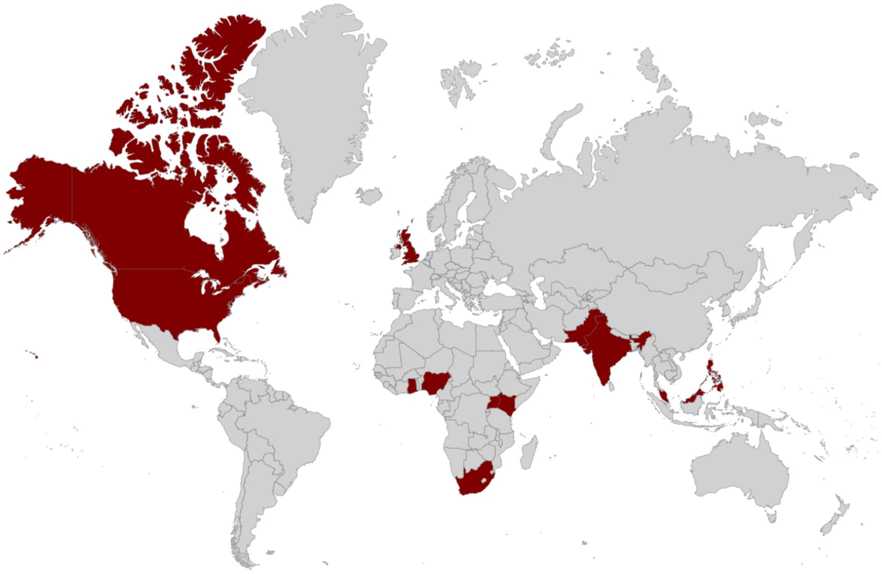


Figure 1: World map with highlighted countries which were chosen for the global survey.

Samsung Galaxy smartphones were in a test room. The control sample was the same two devices held in a silent room for the same duration. During both tests they have given necessary user permissions to all the tested applications, locked the phones and left them in the rooms. In both cases, no pet food ads appeared on any applications. Furthermore, the data consumption over the tested 30 minutes was measured each time and the difference in amount used was minimal between the test and control samples. Contrary to that, Wandera has estimated that a native voice assistant would use up to 2 orders of magnitude more data, while recording and sending audio data to cloud over the test time span.

However, their methods had significant drawbacks and thus could be deemed unreliable. First and foremost, they are basing their assumption on a fact that audio data would be recorded and sent to the cloud straight away, which does not necessarily has to be the case. Second, their experiment was conducted only on locked phones, so it does not provide any information on the possible audio recording actions while using the suspected applications. At last, they have conducted their tests over regular time intervals which might not detect anything if the suspected recordings are taken at random times. All of the above mentioned drawbacks lead to the conclusion that much could be improved about the Wandera study and that their results are not fully reliable.

2.2 Northeastern experiment

An academic study was also conducted on the issue. A group of researchers at the Northeastern University in Boston performed a much more rigorous experiment [4]. They have tested over 17,000 Android applications from various marketplaces for both audio and video leaks. They have collected data on network consumption, permissions required and the use of third party libraries. Their study concluded that no application was recording audio without the users awareness.

Even though the Northeastern study was much more vigorous than the one conducted by Wandera, we believe that it was too broad to really verify whether the tech giants are eavesdropping on our conversations. Testing across 17,000 Android applications really limits the depth of analysis which could be conducted. Another key element which is missing is the complete absence of iOS devices in the experiment. Whereas, the Northeastern study did focus deeply on the Android devices, it also did its research in 2018, when the internal indicator capabilities of the devices' operating systems did not support access level indicators.

3 METHODOLOGY

Our research took span over 2 phases — a global survey and a physical experiment. Both were conducted in order to assure the validity of our data and to provide a global perspective on the possible privacy violations investigated in this paper while basing the research on the physical data.

3.1 Global survey

Before committing to the physical experiment we conducted a never before seen worldwide survey. We decided to run the survey prior to the experiment as that would remove many possible bias factors which could arise in the survey. Our goal was not to suggest anything to the respondents and keep the survey as unbiased as possible. The privacy issues are commonly mentioned on all continents, yet it is hard to find any reasonable data on the matter. This is why we decided on using Google Surveys in order to collect data from 12 countries spanning worldwide as presented in **Figure 1**. The countries we explored were (in alphabetical order): Canada, Ghana, India, Kenya, Malaysia, Nigeria, Pakistan, Philippines, South Africa, Uganda, United Kingdom, United States. We decided to select appropriately sized countries all around the world in order to keep our sample representative of the world population. We collected responses from 200 people in each of the countries. All the surveys were collected in English, as we have selected countries that offered that capability through the Surveys platform.

That data gives clear perspective on the global perception on the issue and might also provides valuable insight into different viewpoints towards privacy policies across the world depending on the potential factors such as the democracy index, etc., which could be explored in some future research. The questions covered a variety issues connected with privacy and Big-Tech companies.

(1) IRB Approved Consent Form
(2) Do you suspect any of these companies are secretly listening to your conversations when you are using your mobile device?
(3) Do you suspect any of these companies are secretly listening to your conversations when you are NOT using your mobile device?
(4) Have you ever switched off your phone or put it away while discussing a sensitive topic because you were worried about being recorded?
(5) To which degree are you worried that Google might use your private data in a way you consider to be inappropriate?
(6) To which degree are you worried that Facebook might use your private data in a way you consider to be inappropriate?
(7) To which degree are you worried that Huawei might use your private data in a way you consider to be inappropriate?
(8) To which degree are you worried that Apple might use your private data in a way you consider to be inappropriate?
(9) Have you ever used any products from these companies? (smartphones, mobile apps, web browsers, etc.)
(10) If you've discussed a topic in a conversation and it later appeared on your mobile as a suggestion, recommendation or an ad, describe that situation. (otherwise write N/A).

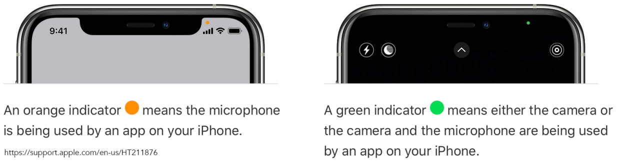
The survey also had an appropriate IRB approved consent form to mitigate the possible conflict of interest for people who were giving their opinion on Google via a Google owned platform.

3.2 Physical experiment

The physical experiment tried to overcome all the shortcomings of both Wandera and Northeastern studies. First of all, we conducted our experiment on both iOS and Android devices. We only used vanilla Android, as the possible provider additions would have potentially influenced the results of our research. We also conducted all the tests in 3 states of the application: foreground, background and locked. The application tested in the experiment was the main Facebook mobile application.

In our research we exposed the devices to audio advertisements in a similar manner to what was conducted by Wandera. We decided on fast food ads and created a playlist which was played in a loop throughout the whole experiment. However, building on top of their shortcomings we improved many of the previous experiment's features. We decided on the process of having a trial and control group, of one Android and one iOS device, which were Xiaomi Redmi Go and iPhone 7 Plus respectively. The experiment was scheduled to take 7 days and 6 test rounds were conducted on each day. The rounds lasted for 2 hours each, with the breaks in between to charge up the phones to the full battery capacity. That measure was taken to prevent different battery consumption level at different stages of power dissipation in mobile devices, which is known to commonly occur [6]. During the test rounds half of them were trial rounds and the other half were control samples. The trial rounds consisted of both phones being confined in a soundproofed place, where another device would stream the fast food ads. The control group repeated the same rounds with the only changed factor being the lack of the audio playback. There were 3 different types of trials, each dependent on 1 of the 3 states of the running app (foreground, background and locked).

The collected data was the battery percentage of the device before and after the trial as well as the screen capture of the devices needed for the system indicator analysis. System level indicators were introduced in Android 12 and iOS 14 respectively and present whether any application is using



An orange indicator ● means the microphone is being used by an app on your iPhone.

https://support.apple.com/en-us/HT211876

A green indicator ● means either the camera or the camera and the microphone are being used by an app on your iPhone.

Figure 2: Apple's access indicator feature.

the microphone or the camera from the operating system level (see **Figure 2**). Later, the recorded visual footage was analyzed in x30 speed to verify against the presence of the before mentioned system access indicators.

4 EVALUATION

Having concluded the international survey we discovered some interesting trends worldwide, while the research done on the physical experiment brought other rather unexpected results.

4.1 People consider switching off

Question number 4 of the survey asked whether the respondent has ever considered turning off their phone while discussing something important as they were scared of being recorded. **Figure 3** presents the worldwide responses distributed by country. Interestingly, African countries as well as western ones such as the United States or the United Kingdom, are significantly less worried about possibly being recorded, whereas Asian countries are clearly deeply concerned about the issue.

With as many as 48% of respondents admitting to having actually turned off their devices while discussing something secretive, the results seem to be evident — the global public is truly worried. Even in the least worried country, the United Kingdom, more than 40% of the respondents have admitted to at least considering switching off their mobile devices. The issue is clearly indicated by the 2 shades of red in **Figure 3**, where any red visualizes the value of “at least Considered”, presenting an astounding result. Clearly, almost all of the countries we surveyed present a significant tendency to worry about being recorded. We believe that by gathering a truly worldwide audience for our survey we were able to finally underline how prevalent the issue of assumed technological eavesdropping really is.

4.2 Trust ratings of the Big-Tech companies

Another shocking results we were able to verify were the trust ratings issued to the Big-Tech companies by our respondents. Our questions asked to which degree are they worried that a given company might use their private data in a way which they consider to be inappropriate. The answers varied from “Not worried at all”, to “Extremely worried” on a scale from 1 to 10.

The aggregated results in **Figure 4** present percentage numbers as vertical values indicating that the vast majority of our respondent is extremely worried about all of the companies misusing their data. Considering further analysis it is worth to point out that Facebook consistently received significantly higher numbers of negative ratings (7-10 points) worldwide, as compared to the other companies. That shows how Facebook is extremely mistrusted by the general public.

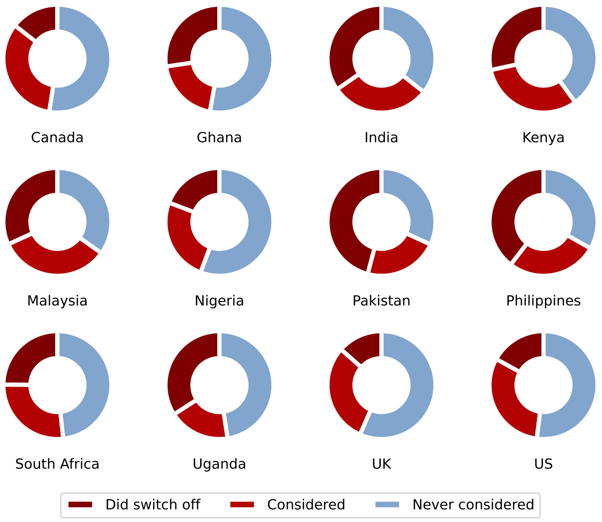


Figure 3: Answer distribution for the question "Have you ever considered switching off your phone while discussing something important?"

The other significant trend is the visibly decreased mistrust ratings among the African countries, especially noticeable in case of Huawei. The amount of “Extremely mistrust” (10 points) answers is almost halfway smaller in countries like Ghana and Nigeria, as compared to Canada and Philippines possibly indicating some cultural or economical factors influencing the mistrust ratings.

4.3 Physical Experiment Results

Having conducted the physical experiment the results were very contrasting to the previously gathered public opinion. The battery consumption levels were not higher when subjected to audio test rounds as compared to the control samples. They remained stable and comparable thought the whole experiment. Furthermore, having regularly scrolled though Facebook feeds of both devices, we did not notice any fast food ads in both test groups. At last, the close analysis of the recorded screen footage did not show any signs of appearing system access indicators, further defending the thesis that Facebook application is actually not recording any audio data of its users.

4.4 Other possible explanations

Having established the global mistrust towards Big-Tech companies and being unable to prove the merit of the worldwide accusations, we further researched the phenomenon. That brought us to exploring the premises of Google’s failed

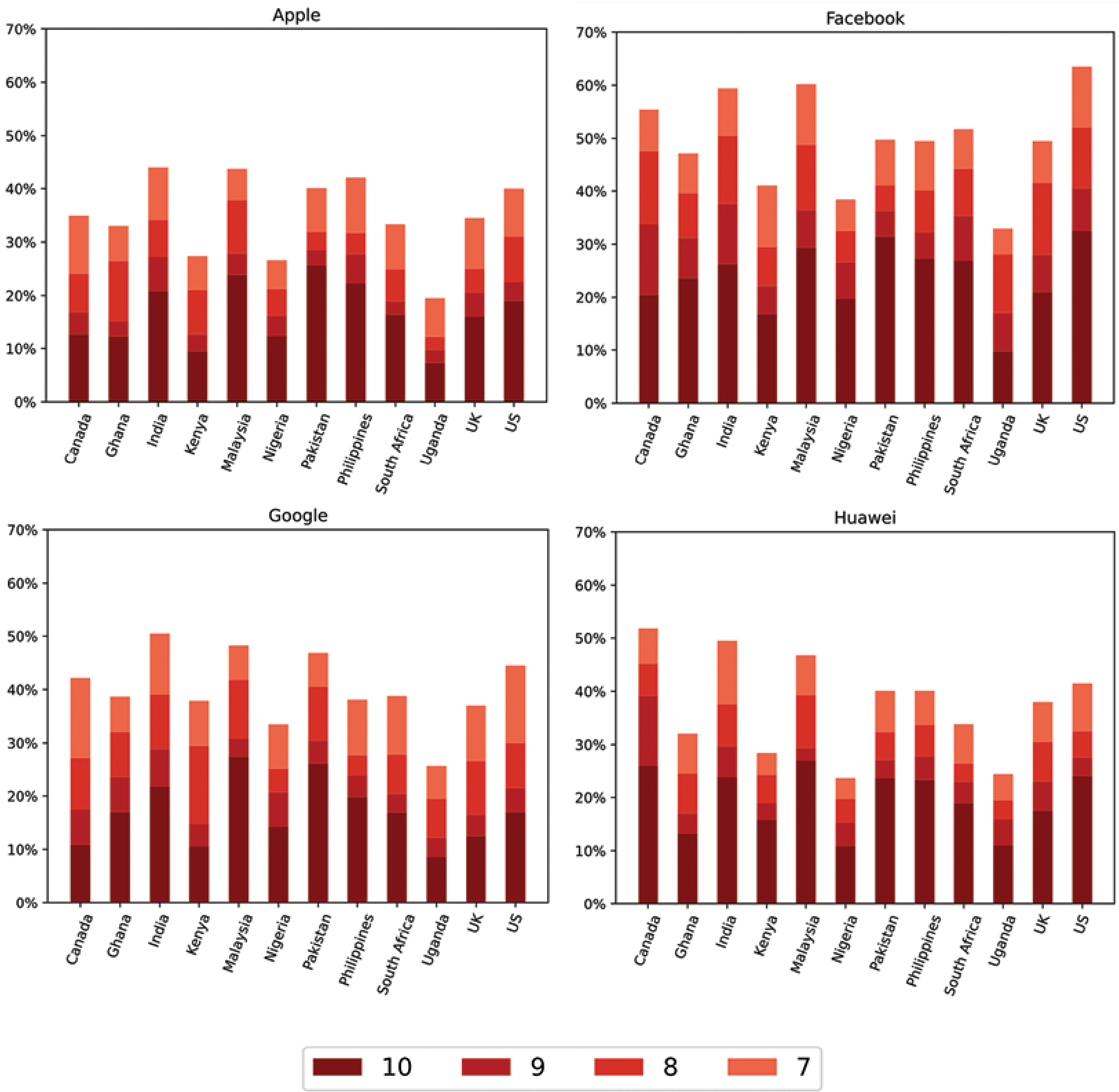


Figure 4: Selected answer distribution for the questions "To which degree are you worried that a given company might use your private data in a way which you consider to be inappropriate?" where 10 indicates "Extremely worried"

Federated Learning of Cohorts Project [1]. In it, each individual would be assigned to an informational bucket and the advertisers would choose which bucket do they advertise to. That network effect of similar people belonging to same buckets provides a reasonable substitute of an explanation for the assumed origins of “eavesdropped” data. If similar people are grouped into a same bucket, they would most probably share some key distinguishable characteristics, such as age, gender, hobby or location. Basing on that, the algorithms are capable of estimating suitable content for the given group with a high degree of certainty, without the need to record them.

Another, even more recent solution by Google named "Topics", takes an opposite approach and assigns each person with 15 out of 350 possible topics basing on the last 3 weeks of their browsing history [2]. That might as well explain the precision of the not random ads appearing in users' feeds.

5 CONCLUSION

The project was vigorously executed and did bring interesting results. The planning phase ensured that all the confounding factors of other experiments were not present in our study. The uniqueness of the worldwide scale of the experiment further enhanced the validity of the results, clearly determining that the world public is indeed mistrusting the Big-Tech companies. The vast majority of the respondents truly believed that their privacy is at risk even though we were not able to prove that they were actually being actively recorded. The psychological factor behind the phenomenon is highly shocking and might be a topic of further research. Both the global survey and the physical experiment clearly present, that the public perception of the approach of Big-Tech companies towards the privacy is strikingly negative and demands change.

REFERENCES

[1] Google [n.d.]. *Federated Learning of Cohorts (FLoC).* Google. https://privacysandbox.com/intl/en_us/proposals/floc/

[2] Google 2022. *Topics: The new Privacy Sandbox proposal for interest-based advertising.* Google. https://support.google.com/google-ads/answer/11899856

[3] Sean Keach. 2020. *Is Facebook listening to you? The truth and how to avoid snooping revealed.* https://www.thesun.co.uk/tech/7497249/facebook-listening-to-you-microphone-ads/

[4] Elleen Pan, J. Ren, Martina Lindorfer, Christo Wilson, and D. Choffnes. 2018. Panoptispy: Characterizing Audio and Video Exfiltration from Android Applications. *Proceedings on Privacy Enhancing Technologies* 2018 (2018), 33 – 50.

[5] Liarna La Porta. 2019. *Is your phone always listening to you?* Wandera. https://www.wandera.com/phone-listening/

[6] Jeff Shepard. 2021. *How to read battery discharge curves.* Battery Power Tips. https://www.batterypowertips.com/how-to-read-battery-discharge-curves-faq/