Data Sharing and Exposure: Findings from Descriptive and Network Analysis of Instant Games on Facebook

Stacy Nicholson
snicholson@towson.edu

Joyram Chakraborty
jchakraborty@towson.edu

Aisha Ali-Gombe
aaligombe@towson.edu

Robert J. Hammell II
rhammell@towson.edu

Towson University
Towson, MD 21252, USA

Abstract

Do you know how your data are being shared? This is a privacy concern that continues to be unclear to users interacting with third-party applications (apps) on social networking sites. Although the new privacy laws and regulations in recent years have forced developers to present end users with better permission control and privacy policy statements, which roughly specify what types of data can be collected and share with third parties, the lack of clarity in these options makes it difficult for average users to understand the full specifications of how their data are being shared. Using a combination of descriptive and network traffic analysis, this study examines Instant Games apps on Facebook to understand and identify how the user's data is being shared and its potential security exposure. The network analysis found that 86% of user's identifiable data (name, photo, and location) are shared with third-party domains. This work supports the need for greater transparency regarding user data sharing with third-party apps/services on social networking sites.

Keywords: Data Sharing, Privacy, Third-party apps, Instant Games, Data Exposure, Descriptive Sharing Analysis.

1. INTRODUCTION

Data sharing with third-party apps on social networking sites (SNS) has grown exponentially. These third-party apps and services can collect a huge amount of user data with every interaction leaving users concerned about their privacy when using such apps. Such concerns are underscored by the Facebook Cambridge Analytica data privacy scandal, where up to 87 million pieces of personally identifiable user information were captured and misused by a third-party quiz app (TIME, 2018). SNS such as Facebook offer third-party apps a diverse user-centric and cross-platform environment where they can expand their services and products to a large pool of

online users. Facebook achieves this with the aid of APIs, policies, or terms of services, which provide third-party apps such as Instant Games with access to the user data. For example, Facebook APIs (Facebook, 2021a) allow Instant Games (IG) to access user's names, locale, profile photos, etc., which is privacy-centric. According to pCloud (2021), Facebook shares 57% of the user data. Once the user data leave Facebook servers, there is little insight into how it is used or shared with other associated third parties. Although general Facebook permission and privacy controls have seen improvements over the years, when it comes to its third-party IG, the notion of implied permission and the all-or-nothing opt-in request for user data access are areas that still need significant improvements. The need for delivering better and unambiguous control options to users to ensure that third-party IG apps have limited access to the user information cannot be overemphasized.

The user experience may differ when interacting with third-party apps, since users may not know to whom their user data is being shared, or where their data is transmitted to and stored. The user can attempt to understand this by first going through each IG's privacy policy or lengthy IG terms of service. However, the data sharing policy given by Facebook and the IG apps are presented in a broad category listing regarding third-party partners with which the user's data are shared. Additionally, it does not exactly list who these third-parties are by name and the full scope of what types of user data are shared. Greater transparency is needed about exactly who, where, and the types of data being shared with these third-party apps and services.

In this study, we examined twenty (20) IG featured on Facebook's gaming platform, to understand how the user's data are being shared and collected, and the potential security exposure. We looked at cross-origin data sharing employed by Facebook. Cross-origin data sharing in this study relates to data originating on Facebook being shared with third-party IG applications, servers, and services. The contribution of this paper is realized through two approaches. The first is the use of *Descriptive Sharing Analysis*, where we examine the IG data sharing requirements and control options. The second is *Dynamic network traffic analysis,* where we conduct a network traffic analysis to pinpoint the types of user data being shared with third-party servers or services, and potential security exposure of user data that may occur during transmission over the network to these third-party domains.

Cross-origin data sharing in the scope of IG on Facebook is an area of concern that needs addressing. Note that examining the extent to which an adversary can exploit vulnerabilities in IG apps on Facebook is not an aim of this paper, nor are there any suggestions offered to provide mitigation measures at this time. This study focuses on the specific objectives presented later in this section.

The types of user's data being shared with third-party IG apps/services are personally identifiable information (PII), and other types of user information that may put users at risk and invade their privacy. Risks include the fact that a user's name, location, profile photo, connections, etc. are not only shared with third-party IG apps but also with other services and stored on servers outside of Facebook. Thus, in addition to being concerned about how Facebook shares their data, users also need to consider how the third-party apps/services use and reshare their data to others. In addition, the amount of data sharing being done may open the door for the user to be constantly tracked while engaging in game play activities each time they use an IG app. For example, concerns regarding vulnerability to Man-in-the-Middle (MITM) network sniffing are real, as will be shown in this study's data analysis.

While there are numerous examples in the literature focused on privacy and security concerns related to Facebook (Li et al., 2015)(Gross et al., 2005)(Al-Shamaileh et al., 2017)(Dhami et al., 2013)(Kumar et al., 2019), no work was found that addressed the key IG third-party data sharing privacy and control aspects. There is a critical need for IG users to understand how their data are being shared, and to understand the potential for data exposure.

The overarching research goal supported by the work in this paper is the examination of IG on Facebook to bring an understanding and awareness of data sharing and privacy leakage relating to the flow of data across IG domains. The efforts reported herein begin this overall research thread by examining the following specific questions:

1) What types of user data are being shared with IG third-party applications? The study examined the types of user data and their device details that are being shared, which includes: name, photo, location, gender, and other associated data.

2) With whom is the user data shared? (IG developers / third-party domain / third-party analytics, etc.). In answering this question we

explored with whom and where the user data is being shared as it relates to the IG app's domain, third-party servers, and services.

3) What types of security exposure exist? The study investigated potential user data exposure during transmission that may expose and compromise the user's privacy. In addition, permission and privacy controls limitations were also examined as they relate to IG that may further impact the user's privacy.

2. BACKGROUND

Instant Games on Facebook

Since it first launched in 2016 as a cross-gaming platform, IG on Facebook offers social games that users can play with friends or other players on its platform. Over 1000 free IG are currently available on Facebook's IG platform (Facebook, 2020c), offering a wide variety of game categories. With millions of active players each month (Statista, 2020), the IG platform has grown in popularity over the years. Some popular games on the platform include Zynga's Words with Friends, Angry Birds, and 8 Ball Pool, all of which are accessed instantly from the Facebook gaming web portal. The style of each IG gameplay (Trivia & Word, Action, Cards, etc.) offers unique gaming options to users. IG are often played through the Facebook mobile app or its website. Users can also engage in group tournaments.

Instant Games apps are built using various technologies that allow developers to create games without device restrictions, enabling them to share code across devices with little to no changes (Facebook, 2021a). According to Facebook (2020b), games on Facebook are hosted as a portal. However, the actual game content can be hosted from the developers' web server or other third-party servers. Data APIs for player, context, locale, & entry point are accessed during the game loading (Facebook, 2021b). With Facebook's Terms of Service (TOS), data, and web service APIs, third-party IG can access user data available on the Facebook platform.

Data Sharing and Privacy Concerns

Prior research shows that users are concerned with their privacy as it relates to who can access their data, and privacy on Facebook, (Golbeck & Mauriello, 2016)(Wang et al., 2011)(Malik et al., 2016). In one study (Johnson et al., 2018) it was observed that 94.6% of Facebook users deny access to their data such as photos and content posts to persons outside of their friend network to protect their privacy. However, the study Johnson et al.(2018) is limited when it comes to addressing third-party insider threats.

The study by Golbeck & Mauriello (2016) states that users' overall concerns about privacy increased after seeing information about how apps access their data and the amount of personal data that can be obtained. The study also mentions that users did not fully understand what data Facebook apps could access about them after going through the "Facebook Data Policy" document. The study further shows that users were concerned about how Facebook handles their data with respect to selling or releasing data, identity theft, and legitimate apps being fraudulent. According to Facebook (2020a), they "don't sell any of your information to anyone"; however, prior reports TIME (2018) show that the user data is overly shared with third parties.

Impact on Users

The study conducted by Gross et al. (2005) shows that when the user shares their data with Facebook (full date of birth, gender, photo, etc.) it exposes them to various privacy implications such as demographics re-identification and face re-identification. For instance, if an adversary gains access to a user's friend or community network on Facebook, they could correlate a "comparatively large number of users to outside, de-identified data sources"(Gross et al., 2005). The authors also mention identity theft as an additional re-identification risk factor that may arise.

In various news reports, Facebook user's data were leaked even when the user had privacy settings configured globally to protect their data by limiting it to just friends, family, and close associates (TIME, 2018)(Dance et al., 2018). A recent security breach left 533 million Facebook users' data exposed online from 106 countries (Insider, 2021). Findings from Li et al. (2015) state that "privacy leakage could still happen even if a user correctly configures his privacy settings due to the exploits caused by inherent conflicts between privacy control and OSN functionalities". The analysis presented later demonstrates these problems are legitmiate issues with respect to IG.

Analysis Approach

As noted in this study's objectives, this work seeks to investigate third-party data sharing in a social networking environment related to the use of IG. Also, in this work we start to examine permission and privacy controls (objective question 3). Prior work such as Zang et al.(2015) and Jadhav Bhatt et al. (2018) shows one can

observe data sharing and data exposure instances through dynamic network analysis and monitoring. Regarding control options as they relate to permission and privacy, prior studies were mostly done from a survey-based (Chia et al., 2012; Kayes & Iamnitchi, 2017), user perception (Golbeck & Mauriello, 2016), or design standpoint (Van Kleek et al., 2017). Most studies found within the literature have not used the style of descriptive sharing analysis conducted in this research.

3. STUDY DESIGN

This study examines and evaluates privacy in 20 IG apps using two important methodologies - descriptive and network traffic analysis. Only apps that fall under the IG banner were chosen for examination.

The IG chosen are popular games that fall into the following categories: solo gameplay, social with friends, and social with random players. The app categories are directly related to the social structuring of IG, allowing us to examine further Facebook's privacy control and IG data access requirements.

For the basis of understanding data sharing from a descriptive sharing analysis, a small sample (20 IG) was used. Preliminary examination of IG revealed a reoccurring pattern with the same default user data requirement structure across many games. This influenced the decision to use a small sample size for the descriptive and network analysis to be conducted.

Network traffic analysis is employed to observe the user data generated and exfiltrated from the IG in real-time during gameplay. Additionally, we examine and uncover other third-party domains that may have received copies of the user's data during transmission.

Experimental Setup
Figure 1 illustrates the data monitoring, capturing, and recording process using the Charles proxy (Charles Proxy, 2020)[trial version]. Due to space limitations, further details including the packet examination methodology are given in Appendix A.

4. DESCRIPTIVE SHARING ANALYSIS

Based on the study objectives in Section 1, this section presents preliminary answers to question 1 and the permission and privacy limitations of question 3. We do this by exploring the design of

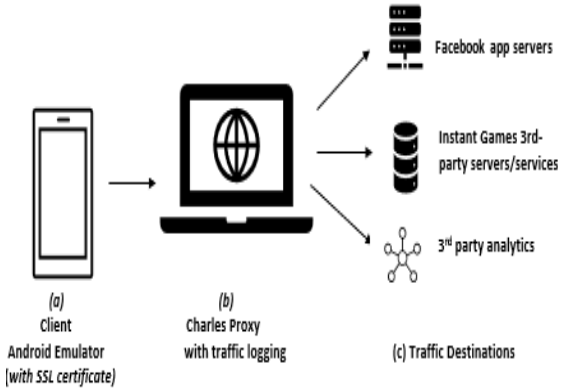

Figure 1. Network traffic capture workflow

the user permission model employed by IG, which Facebook allows. This analysis focuses on the types of IG apps that uses the same default data access requirement statement, as shown in

Figure 2. IG user data requirement dialog shown before gameplay (for each game tested in the experiment)

Figure 2, to inform the users of what types of data will be collected and shared.

Method
In this analysis, the goal is to examine data sharing requirements and the control checks built into the platform to regulate the excessiveness of the apps exploiting user data. We will analyze the design of the user permission model and privacy controls in both pre and post gameplay.

Permission Requirements before Gameplay
The following short permission requirements statement is presented to the user before gameplay in the IG examined as seen in Figure 2:

Friends, connections, and other people that you play with will be able to see your game activity. The game will have access to your name, profile photo, locale/language, and the people you know who also play this game.

The above statement has ambiguities and lacks clarity regarding permission. For example, the collection purpose and the actual organization name are not made upfront to the user. In addition to this, the user may not be aware that the collected data are being shared outside of Facebook with third-party partners. Furthermore, the word "locale" is not a familiar term; users may not understand that their location (country, city, etc.) at the time of use will be collected and potentially tracked. Additionally, the term "connections" in the permission statement does not state what this involves. Prior research by Van Kleek et al.(2017) indicates that a user can make more informed decisions when presented with the app permission requirements and purposes for collecting their data.

Before users begin playing an IG, they must confirm that they agree to grant access to their user data before they can proceed. As shown in Figure 2, the user's permission control abilities are limited; no option is given to edit what gets shared with the app. There is a sense of forced user permission in that the user has only two choices: 1) Click play to accept all the data-sharing requirements, or 2) Exit out and decide not to play. Unlike standalone mobile apps on platforms such as Android and Apple, Facebook does not give users the choice to selectively disable what data (such as location, contacts, photo, etc.) will be shared with the IG.

Apart from the permission control limitation that the user experiences, these IG apps may be viewed as privacy-invasive apps, based on the amount of user data being accessed, shared, and collected. Along with the permissions ambiguities that exist, there is a lack of clarity regarding how the app will interact with the user's device.

Another concern is that users may not read or go beyond the short data requirement statement in Figure 2 (or read it at all) before engaging in gameplay. Prior research by Sigmund (2021) and Obar & Oeldorf-Hirsch (2018) has shown that users tend to disregard reading lengthy privacy policy (PP) and TOS documents due to information overload. In addition to this, users may not understand the PP, since according to Fabian et al. (2017) PP requires a level of education such as high school or some college to fully comprehend. Furthermore, another study (Obar & Oeldorf-Hirsch, 2018) indicated that 74% of participants skipped the PP or TOS document altogether.

Privacy Control after Gameplay
Figure 3 shows that the user cannot modify any part of the information requested by the IG apps. The only option the user has is to remove the IG app altogether. However, according to Facebook IG settings, removed apps may still have access to information previously shared with them. Once an IG collects the user data, there are no procedural restrictions on how third-party IG use, store, and share user's data.

Summary of Descriptive Sharing Analysis
The findings from the descriptive analysis contributes to an understanding of data sharing and privacy concerns that arise with IG.

In regard to study objective Question #1, the descriptive analysis showed from the pre and post gameplay assessment that a subset of the user data (name, profile photo, locale/language, and the people you know who also play this game) is explicit in nature. However, there is a notion of implied data access where other user data found in post gameplay (time zone, gender, messenger connection, etc.) were not made known to the user beforehand. This highlights the concern that excessive data sharing can be done without the user's prior knowledge.

With respect to study objective Question #3, we also observed how inadequate the existing Facebook privacy and permissions settings are to protect the user data from third-party insider threats. The all-or-nothing access approach creates little room for the users to make modifications. A more optionable permission model would allow for users to make better decisions and have more control over what user information gets accessed and shared with these IG pre and post gameplay. The short permission requirements statement given before gameplay (Figure 2) shows that the user is not told that their gender, time zone, city, and messenger connection is being collected/accessed/shared beforehand; it is after gameplay that this extra user data is being mentioned (Figure 3). Hence, the user's ability to judge whether to accept the terms of the game is diminished.

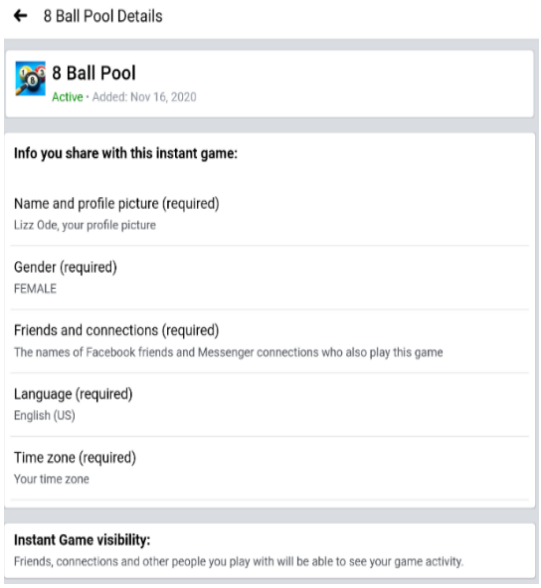

Figure 3: IG user data requirement dialog shown after gameplay (for each game tested in the experiment)

With respect to the privacy portion of objective Question #3, the pre/post comparison gives a benchmark for exploring potential data leakage over the network. This examination can start by checking for excessive data collection being done that was not revealed to the user either before or after gameplay.

Note that this discussion does not aim to dispute that IG apps would need access to user data to render game services, improve features etc. However, a concern arises over whether the user understands that the game is a third-party app not owned by Facebook. Also, it is not clear that users grasp the extent to which their personally identifiable information and other user data is shared with IG servers and associated third-party services.

5. NETWORK TRAFFIC ANALYSIS

The descriptive analysis found partial gathering of the user data by the IG apps which provided a baseline for more detailed analysis. In this section, we use network traffic analysis to further explore the user data being accessed, and exfiltrated to IG domains and associated third-party servers and services. In addition, attempts are made to uncover potential data leakage.

Method

A dynamic network analysis was conducted on 20 IG hosted on Facebook's Instant Games web portal. The network analysis allows for real-time observation of data flow during gameplay and answers all three study objective questions outlined in Section 1 from a different point of view.

Figure 4. Captured third-party server connection.

In the dynamic network analysis, Charles proxy (see Figure 1) was used to monitor and record all network communication from the Android client emulator device on which the IG was played. Due to space limitations, further details are given in Appendix B.

Analysis of 20 Instant Games apps

In the findings, 16 out of the 20 IG examined had partially encrypted data and SSL vulnerabilities, where these apps were prone to Man-in-the-Middle (MITM) network sniffing. During the network analysis, we were able to trace the user data going to the IG domain third-party servers (AmazonAWS, Microsoft Azure, etc.), analytics services (such as google-analytics.com), and with other associated domains shown in Table 1 (Appendix C). Figure 4 shows a snapshot of an IG app third-party server connection where traces of user data were found during transmission.

While there is a need to access certain user data to render game services to devices and users, user comprehension about the types of game data being collected and shared is limited due to a lack of user understanding and clarity in the requirements. Users may not be aware of any potential security exposure that could impact their privacy until it is made known to the public.

Types of User Data Shared with Third-party

As each game is played, a list of domains with which the game interfaces is provided. After game play, every domain from the list was checked to see if user data appeared. Of the 20

IG examined, there were 59 third-party domains where the user's data were found as shown in Table 1 (see Appendix C; domains that did not include user data being shared/retrieved were excluded from the table). From the analysis, 86% of user's identifiable data (name, photo, and location) are shared with third-party domains.

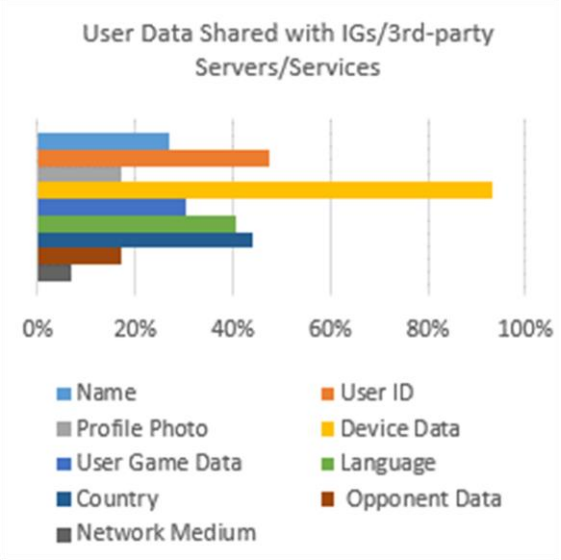

Figure 5. Percentage of domains in which each type of user data appeared.

Figure 5 depicts the percentage of the domains where types of user data were found. Note that device data was shared the most, as it was found in 93% of the domains. The type of network connection was shared the least at 7%. It is noteworthy that user "gender" data was not revealed in the network analysis even though it was collected by the IG apps during gameplay.

Instances of User Data being Shared

In the findings of the captured network packets, we found 191 instances of user data shared with third-party domains being exposed for 16 of 20 apps tested. Such a collection of user data could lead to the user's privacy being exploited on and off Facebook via re-identification, tracking, or other activities. Since the user data is being captured over the network before de-identification or obfuscation occurs, the user data is left exposed. A sample of the captured data is depicted in Figure 6 and Figure 7 for a user (user1).

Third-party Servers/Services Destinations

During our analysis, we found that the Google domain accounted for the most instances of where user data were collected (34%); AmazonAWS was second at 32%. Other frequently observed domains were Googleusercontent.com (14%) and Alicloud-Us (14%). Additionally, it was observed that most, if not all, IG apps store the user data on servers outside of Facebook.

For the 20 IG apps analyzed, we found that most of the user's data identified in Table 1 (Appendix C) were collected and shared with the IG's domain/servers. A smaller set of user data was shared with other associated third-party domains such as play.googleapis.com, collect15324sltrf .deltadna.net, googleapis.com, etc. as seen in Table1. In addition, third-party servers (AmazonAWS, Alicloud, Microsoft Azure Cloud, etc.) and third-party analytics (google-analytics.com, gameanalytics.com) were found. 16 of 20 apps (80%) were found to be transmitting user data to third-party domains and servers. Figure 8 shows a connection instance to a third-party analytics domain where the user's device data is being sent/shared.

Data Capture Details

The user-ID generated for 16 of 20 the IG apps was captured and exposed. User gameplay activities were also exposed, which allows the user's behavior pattern to be tracked and learned over time; This could put users at a disadvantage during competitive matches. Further encryption at all levels is needed.

Figure 7. User1 data captured location, etc.

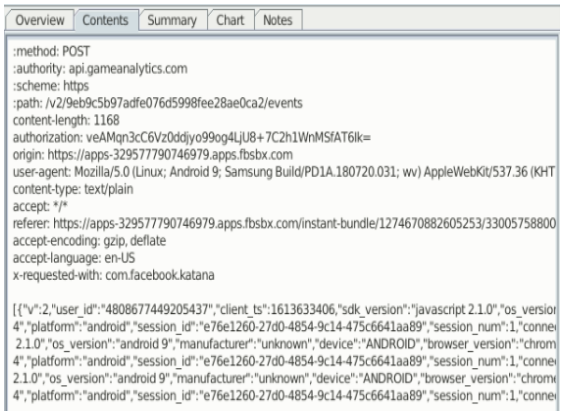

Figure 8. Captured third-party domain with user data being shared

Figure 9. Captured device data

A sample of the captured user device details is shown in Figures 8 and 9. These figures depict the user device details being shared, such as Android client (version, type, build number), machine OS type used to run the Android emulator (Linux), location, etc., to the third-party domain.

From the network packet analysis, user1's own Facebook account profile photo and that of friends and other players (opponents) were obtained. The images were found at Facebook domain URL *"platform-lookaside.fbsbx.com,"* which renders images to the IG apps. This means the users' profile photo was not obfuscated when stored on Facebook servers. Figure 10 shows a snapshot of a captured Facebook user's profile photo.

Summary of Network Traffic Analysis

It was observed that IG played as solo games had less instances of user data being collected and exposed than games played with opponents. Solo games include Brain Game, Candy Rain, Escape Now, Flip Bottle, Garden Tales, Rope Cut, Solitaire, and Sudoku.

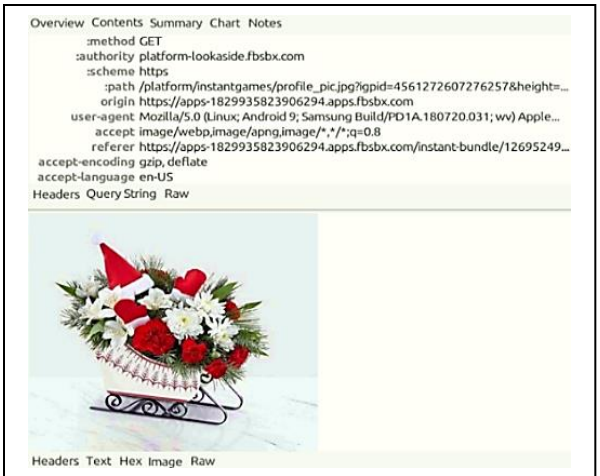

Figure 10. User1- test user profile photo on Facebook

IG played with opponents had the most instances of user data exposure through Facebook cross-origin data sharing with the IG. This held true regardless of whether the game was played with random opponents (8 Ball Pool, Angry birds, Dominoes, Ludo Club) or with friends (Chess, and Words with Friends). This illustrates that the security impact relating to data sharing in games played with an opponent opens the door for the user's privacy to be at risk.

Additionally, the analysis shows the lack of proper SSL certificate implementation and lack of consistent vulnerability checks in apps, putting user information at risk. A user's profile data on Facebook still needs protection against insider threats from third-party apps such as IG. The findings again show that the privacy setting employed by Facebook does not allow the user to have adequate control of what gets shared with IG. It also underscores the need for IG to state the purpose for collecting personally identifiable user data from Facebook.

Attacks on user privacy and data are possible from a social engineering standpoint (such as phishing) whether the user turns on messages from the IG during the gameplay or after. Users of IG need to be aware of phishing threats both on Facebook and during in-game play activity. IG that are played with random players, including game tournaments, may open the door for the user to receive phishing attempts.

Findings from the network analysis show that the types of user and game data being requested by the IG can pose a risk to the user's privacy when exposed to an unintended party. Results also highlight the excessive data gathering being done by non-transactional game apps.

6. STUDY LIMITATIONS AND ETHICAL CONSIDERATIONS

4 out of 20 IG played, namely Plants vs. Zombies, Playing Soccer, Supper Bowling, and Word Search Tournament (not listed in table), were not prone to the MITM sniffing. Hence, we were unable to trace and examine the types of user data that may have been sent. The user's emails, gender, age, phone number, and birthday were not obtained in the network packet evaluation. Therefore, we cannot conclude whether or not these data types were shared with the third-party IG apps from the network analysis even though "gender" and "time zone" were found in the descriptive sharing analysis.

The user data obtained in this study were not used to re-identify or further discover information about the user on or off Facebook. Additionally, the data obtained about players were limited to what is mentioned in the study and only used for research purposes.

7. DISCUSSION AND CONCLUSIONS

There is a need for greater transparency regarding how the users' data is being shared with respect to IG being used on the Facebook platform. Analysis shows that playing games on social networking sites such as Facebook can be privacy invasive. Users will not only have to be concerned with how Facebook shares their data, but also with how these IG apps do as well.

Both analyses illustrate the need for better privacy control and permission options to allow the user more control over their data when shared with IG third-party apps. The descriptive analysis highlighted the concern of excessive data sharing being done without the user's knowledge. It also demonstrated the inadequacy of Facebook privacy and permission settings to protect user data from third-party apps.

Findings from the network analysis gave a more detailed look into vulnerable areas of the IG that can lead to privacy leakage. Few IG apps are fully encrypted and secured against network sniffing, leaving user data exposed during network transmission. It is also clear that privacy settings are inadequate, and that Facebook global privacy settings do not necessarily translate to privacy protection with respect to third-party IG apps. The analysis further demonstrated the depth and breadth of user data that Facebook shares to third-party domains; the excessive data collection and sharing can leave users vulnerable to phishing and other attacks.

This research provides important knowledge into Facebook's data sharing with IG third-party applications. The three study objective questions posed in Section 1 were answered and discussed. In addition, the results also contribute to the limited literature that addresses the data sharing and exposure aspects with IG third-party gaming apps on Facebook. This work may also assist developers in recognizing potential vulnerabilities in their apps that may expose identifiable user's data. The contributions open the door for other approaches that will help users understand with whom and where their data are being shared.

Future work includes a comparative study replicating the same approach or other methods using a similar study for other SNS platforms like Instagram. In addition, this research provides a foundation for further research to be carried out to examine the PP for each IG in more detail to uncover their data sharing practices with data received from Facebook. Further work is certainly warranted to more deeply examine IG vulnerabilities beyond data sharing (such as exploits and attacks) and mitigation strategies to implement.

8. REFERENCES

Accuwebhosting. (2020). *Web Server information*.https://www.accuwebhosting.com/resources/show-web-server-detail

Al-Shamaileh, O., Aloudat, A., & Barikzai, S. (2017). User concerns about Facebook: Are they important? *ICIT 2017 - 8th International Conference on Information Technology, Proceedings*, *May 2017*, 291–296. https://doi.org/10.1109/ICITECH.2017.8080015

Charles Proxy. (2020). *Getting Started*. https://www.charlesproxy.com/overview/about-charles/

Chia, P. H., Yamamoto, Y., & Asokan, N. (2012). Is this app safe? A large scale study on application permissions and risk signals. *WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web*, *May 2014*, 311–320. https://doi.org/10.1145/2187836.2187879

Dance, G. J. X., LaForgia, M., & Confessore, N. (2018). As Facebook Raised a Privacy Wall, It Carved an Opening for Tech Giants. *The New York Times*, 1–17.

https://www.nytimes.com/2018/12/18/technology/facebook-privacy.html

Dhami, A., Agarwal, N., Chakraborty, T. K., Singh, B. P., & Minj, J. (2013). Impact of trust, security and privacy concerns in social networking: An exploratory study to understand the pattern of information revelation in Facebook. *Proceedings of the 2013 3rd IEEE International Advance Computing Conference, IACC 2013*, *February 2016*, 465–469. https://doi.org/10.1109/IAdCC.2013.6514270

Fabian, B., Ermakova, T., & Lentz, T. (2017). Large-scale readability analysis of privacy policies. *Proceedings - 2017 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2017*, *September*, 18–25. https://doi.org/10.1145/3106426.3106427

Facebook. (2020a). *Data Policy*. https://www.facebook.com/policy

Facebook. (2020b). *Games on Facebook*. https://developers.facebook.com/docs/games/gamesonfacebook

Facebook. (2020c). *Instant Games*. https://www.facebook.com/games/instantgames

Facebook. (2021a). *Facebook for Developers*. https://developers.facebook.com/docs/games/instant-games

Facebook. (2021b). *Instant Games SDK*. https://developers.facebook.com/docs/games/instant-games/sdk/fbinstant6.3

Golbeck, J., & Mauriello, M. L. (2016). User perception of Facebook app data access: A comparison of methods and privacy concerns. *Future Internet*, *8*(2). https://doi.org/10.3390/fi8020009

Gross, R., Acquisti, A., & Heinz, H. J. (2005). Information revelation and privacy in online social networks. *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society - WPES '05*, 71. https://doi.org/10.1145/1102199.1102214

Insider. (2021). *533 million Facebook users' phone numbers and personal data have been leaked online*. https://www.businessinsider.com/stolen-data-of-533-million-facebook-users-leaked-online-2021-4?op=1&scrolla=5eb6d68b7fedc32c19ef33b4&r=US&IR=T

Iplocation. (2020). *Who is hosting a website?* https://www.iplocation.net/who-is-hosting-website

Jadhav Bhatt, A., Gupta, C., & Mittal, S. (2018). Network Forensics Analysis of iOS Social Networking and Messaging Apps. *2018 11th International Conference on Contemporary Computing, IC3 2018*, *November 2019*. https://doi.org/10.1109/IC3.2018.8530576

Johnson, M. M. J., Egelman, S., & Bellovin, S. M. (2018). Facebook and Privacy: It's Complicated. *The SAGE Encyclopedia of Business Ethics and Society*, *Section 2*. https://doi.org/10.4135/9781483381503.n442

Kayes, I., & Iamnitchi, A. (2017). A Survey on Privacy and Security in Online Social Networks. *Online Social Networks and Media*, *3–4*(January 2015), 1–21. https://doi.org/10.1016/j.osnem.2017.09.001

Kumar, A., Jain, S., & Yadav, R. (2019). Flaws in Privacy and Security of Facebook. *International Journal of Computer Sciences and Engineering*, *7*(7), 326–331. https://doi.org/10.26438/ijcse/v7i7.326331

Li, Y., Li, Y., Yan, Q., & Deng, R. H. (2015). Privacy leakage analysis in online social networks. *Computers and Security*, *49*, 239–254. https://doi.org/10.1016/j.cose.2014.10.012

Malik, A., Hiekkanen, K., Dhir, A., & Nieminen, M. (2016). Impact of privacy, trust and user activity on intentions to share Facebook photos. *Journal of Information, Communication and Ethics in Society*, *14*(4), 364–382. https://doi.org/10.1108/JICES-06-2015-0022

Obar, J. A., & Oeldorf-Hirsch, A. (2018). The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, *23*(1), 128–147. https://doi.org/10.1080/1369118X.2018.1486870

pCloud. (2021). *Invasive apps*. https://blog.pcloud.com/invasive-apps/

Sigmund, T. (2021). Attention Paid to Privacy Policy Statements. *Information*, *12*(4), 144. https://doi.org/10.3390/info12040144

Statista. (2020). *Facebook: most popular game MAU 2020*. https://www.statista.com/statistics/278933/monthly-active-users-of-the-most-popular-facebook-games/.

TIME. (2018, April 4). *Facebook's Cambridge Analytica Controversy Could Be Big Trouble for the Social Network.* https://time.com/5205314/facebook-cambridge-analytica-breach/

Van Kleek, M., Liccardi, I., Binns, R., Zhao, J., Weitzner, D. J., & Shadbolt, N. (2017). Better the devil you know: Exposing the data sharing practices of smartphone apps. *Conference on Human Factors in Computing Systems - Proceedings*, *2017-May*, 5208–5220. https://doi.org/10.1145/3025453.3025556

Wang, N., Xu, H., & Grossklags, J. (2011). Third-Party Apps on Facebook: Privacy and the Illusion of Control. *Proceedings of the 5th ACM Symposium on Computer Human Interaction for Management of Information Technology - CHIMIT '11*, *December 2011*, 1–10. https://doi.org/10.1145/2076444.2076448

Zang, J., Dummit, K., Graves, J., Lisker, P., & Sweeney, L. (2015). Who Knows What About Me? A Survey of Behind the Scenes Personal Data Sharing to Third Parties by Mobile Apps. *Technology Science*. https://doi.org/http://techscience.org/a/2015103001/

Appendices

Appendix A – Experimental Setup Details

Figure 1 illustrates the data monitoring, capturing, and recording process using the Charles proxy (Charles Proxy, 2020)[trial version]. A client (Samsung galaxy S9) Android emulator device at (a) is configured with the Charles proxy and setup to route traffic through the proxy server. At (b) using Charles HTTP/HTTPS proxy, network traffic generated is captured and saved on the system. The Traffic destination is shown at (c) where the first connection is to the Facebook server, then goes beyond facebook.com to included IG server/ third-party server/ 3rd party analytics. Once the client device and the user Facebook account are authenticated through the Facebook server then the IG gameplay can begin. The Facebook app was then launched on the client device, then followed by the selection of an IG to be played for about 10 to 20 minutes. One author was the main game player. This allowed for an adequate data sample and real time observation of the network packets during gameplay.

Packet Examination Modeling

The network packet analysis starts at the point where the game is first launched. For example, we pinpoint this at the following URL (apps-1825153594465039.apps.fbsbx.com) which is the entry point to the instant game bundle for "Angry Birds" IG. The (apps-xxx-apps.fbsbx.com) is used as the starting point for all the IG's played to better narrow down the actual local or third-party domains that the IG connect to during gameplay. This reduces the number of domains connection not involved during the gameplay activities. Each URL Domain and associated IP address is examined to further validate the third-party connections made during gameplay. Additional software was used to assist with the network packet tracing. these include *nslookup* net tool and web hosting search tool (Accuwebhosting, 2020)(Iplocation, 2020).

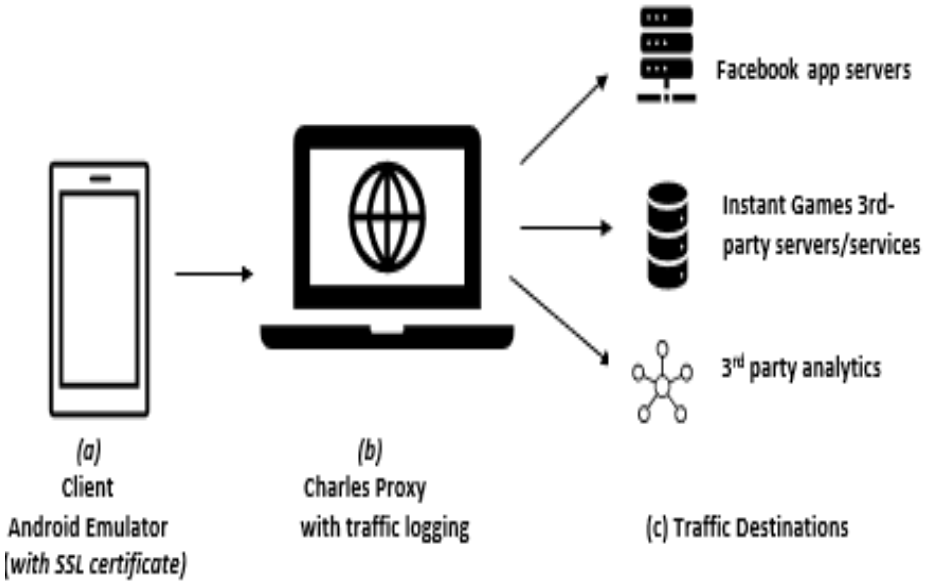

Figure 1. Network traffic capture workflow

Appendix B – Network Traffic Analysis Methodology Details

Charles was used to intercept the network traffic and record HTTP/ HTTPS traffic from the Android client emulator device. All the IG tested were run using Charles proxy which allowed for more accurate results regarding data sharing and exposure occurring with each application.

Once the network traffic was saved in Charles proxy (.chls file) the raw network packets logged were manually assessed. Charles proxy recorded and logged all the HTTP /HTTPS traffic sent and received during active gameplay on the client. Network traffic data was captured, which includes the full site address, remote URLs, GET/POST methods, request/response parameters. WebSocket's send/ received message were also recorded and analyzed. During our assessment, we checked for user data in plain text, URL domains, and device information (such as OS version, model, network connection medium, 3rd-party hosting domain, etc.) to answer the questions outlined in our study objective above

Appendix C – Table 1

IG App	Key {1= yes}	WHERE	DATA TYPES / DATA SENT								
	Third-party Domains	(third-party servers/services)	Name	User ID	Profile Photo	Device Data	User Game Data	Language	Country	Opponent Data	Network Medium
8 Ball Pool	cm.miniclippt.com/18.221.123.145	AmazonAWS				1					
	1c4e92e5-e670-4add-8702-d1cb7ae90a41.goliath.atlas.bi.miniclippt.com/52.7.61.226	AmazonAWS		1		1	1	1	1	1	
	prod-pool-mci-os.mci.miniclippt.com/44.230.191.81	AmazonAWS	1	1	1	1	1			1	
Angry Birds	service.gamesparks.net/35.167.167.237	AmazonAWS				1					
	gsp-aeu007-se53.gamesparks.net/52.215.193.192	AmazonAWS	1	1	1	1	1	1	1	1	
	cloud.rovio.com/52.85.131.216	Cloudfront.net		1		1	1	1	1		1
	play.googleapis.com	Google				1		1	1		
Brain Game	api.hotbloodgame.com/47.88.29.29	Alicloud-Us		1		1	1				
	play.googleapis.com/172.217.15.106	Google				1		1	1		
Candy Rain	candy-rain-5.gb.sbs.softgames.de/3.248.75.44	AmazonAWS		1		1					
	www.google-analytics.com/172.217.15.110	Google				1					
	collect15042cndyr.deltadna.net/34.96.113.148	Googleusercontent.com				1					
	www.googleapis.com/142.250.73.234	Google				1		1	1		
	android.googleapis.com/142.250.73.234	Google				1					
Chess	o70863.ingest.sentry.io/35.188.42.15	Googleusercontent.com				1					
	cdn.gamevh.net/104.26.11.42	Cloudflare	1	1	1	1	1			1	
	play.googleapis.com/142.250.73.202	Google				1		1	1		
	www.googleapis.com/172.217.7.202	Google				1		1	1		
Dominoes	domino-battle-v2-onlineservice.jogatina.com/54.147.70.66	AmazonAWS	1	1	1	1	1		1	1	
	o148945.ingest.sentry.io/34.120.195.249	Googleusercontent.com		1		1					
	domino-battle-v2-match.jogatina.com/18.232.255.118	AmazonAWS	1	1	1	1	1		1	1	
	www.google-analytics.com/172.217.13.78	Google				1					
	play.googleapis.com/172.217.7.234	Google				1		1	1		
Escape Now	api.hotbloodgame.com/47.88.29.29	Alicloud-Us	1	1		1	1	1	1		
	hotblood.oss-us-west-1.aliyuncs.com/47.88.73.45	Alicloud-Us				1					
Flip Bottle	api.hotbloodgame.com/47.88.29.29	Alicloud-Us	1	1		1	1	1	1		
	api-new.hotbloodgame.com/47.254.89.48	Alicloud-Us		1							

	hotblood.oss-us-west-1.aliyuncs.com/47.88.73.45	Alicloud-Us				1					
	play.googleapis.com	Google				1		1	1		
Garden Tales	garden-tales.gb.sbs.softgames.de/3.248.75.44	AmazonAWS	1	1		1	1				
	collect15528grdnt.deltadna.net/34.96.113.148	Googleusercontent.com				1	1	1	1		
	www.google-analytics.com/172.217.15.110	Google				1					
	play.googleapis.com/172.217.7.170	Google				1		1	1		
GO4!	api.digitalmoka.com/52.232.87.186	Microsoft Azure Cloud		1		1					
	smfox3.digitalmoka.com/52.168.131.4	Microsoft Azure Cloud	1	1	1	1				1	
	ajax.googleapis.com/172.217.13.74	Google				1					
Ludo Club by Moonfrog Labs	igludostats.moonfroglabs.com/15.206.108.167	AmazonAWS		1		1	1	1	1		
	sentry.io/35.188.42.15	Googleusercontent.com		1		1					
	igludoprod.moonfroglabs.com/15.207.158.84	AmazonAWS		1		1					
	igl-game-50.moonfroglabs.in/15.207.64.55	AmazonAWS	1	1	1	1	1		1	1	1
	play.googleapis.com/172.217.7.170	Google				1		1	1		
Rope Cut	api.hotbloodgame.com/47.88.29.29	Alicloud-Us	1	1		1	1	1	1		
	android.clients.google.com/172.217.13.78	Google				1					
	android.googleapis.com/172.217.15.74	Google				1					
	play.googleapis.com/172.217.15.74	Google				1		1	1		
	hotblood.oss-us-west-1.aliyuncs.com/47.88.73.45	Alicloud-Us				1					
Solitaire	solitaire-farm-seasons.gb.sbs.softgames.de/54.72.112.17	AmazonAWS	1	1			1				
	collect15324sltrf.deltadna.net/34.96.113.148	Googleusercontent.com		1		1		1			
	www.google-analytics.com/172.217.15.110	Google				1					
Sudoku	service.gamesparks.net/52.27.70.199	AmazonAWS				1					
	gsp-aeu007-se28.gamesparks.net/52.215.193.93	AmazonAWS	1	1	1	1		1	1		1
	play.googleapis.com/172.217.5.234	Google									
	o178629.ingest.sentry.io/35.188.42.15	Googleusercontent.com	1	1							
	api.gameanalytics.com	AmazonAWS		1		1		1	1		1
Super Cricket	ig01.sagames.net/130.211.29.125	Googleusercontent.com	1	1	1	1	1	1	1	1	
Words With Friends by Zynga	3ljlfq9zlb.execute-api.us-west-2.AmazonAWS/13.32.202.32	AmazonAWS				1					
	api.zynga.com/44.237.165.167	AmazonAWS	1	1	1	1	1	1	1	1	
	dpyefmd1sj6u0.cloudfront.net/99.84.185.156	AmazonAWS				1					

	play.googleapis.com/172.217.7.202	Google				**1**		**1**	**1**		

Table 1. Indicates domains associated with IG apps that received instances of user data found during the network analysis. Empty cells illustrate that the user data were not obtained. The key value of [1] represents instances of where and who collected the user's data.