Facebook Applications' Installation and Removal: A Temporal Analysis

Dima Kagan, Michael Fire, Aviad Elyashar, and Yuval Elovici
Telekom Innovation Laboratories and Information Systems Engineering Department,
Ben-Gurion University of the Negev, Beer-Sheva, Israel
Email: {kagandi,mickyfi,aviade, elovici}@bgu.ac.il

Abstract—Facebook applications are one of the reasons for Facebook attractiveness. Unfortunately, numerous users are not aware of the fact that many malicious Facebook applications exist. To educate users, to raise users' awareness and to improve Facebook users' security and privacy, we developed a Firefox add-on that alerts users to the number of installed applications on their Facebook profiles. In this study, we present the temporal analysis of the Facebook applications' installation and removal dataset collected by our add-on. This dataset consists of information from 2,945 users, collected during a period of over a year. We used linear regression to analyze our dataset and discovered the linear connection between the average percentage change of newly installed Facebook applications and the number of days passed since the user initially installed our add-on. Additionally, we found out that users who used our Firefox add-on become more aware of their security and privacy installing on average fewer new applications. Finally, we discovered that on average 86.4% of Facebook users install an additional application every 4.2 days.

Keywords-Social Network Analysis, Social Network Privacy, Social Network Security, Facebook Application.

I. INTRODUCTION

In the last decade, online social networks have gained enormous popularity. Over a billion users worldwide are using these networks [18] to share information, communicate with friends, play games, etc. [6]. Recently, J.B. Duggan [12] discovered that 67% of adults in the United States use online social networks. These results demonstrate a 2% increase in comparison to a previous report [20]. The exponential growth of social networks created a reality where there are social networks for almost every usage. These vary from social networks for connecting with business colleagues like LinkedIn [7] and Xing [9] to social networks for animal lovers, such as Dogster [2], Catster [1] and YummyPets [10]. The biggest online social network is Facebook, [3] which has more than 1.11 billion monthly active users as of March 2013 [6]. The median Facebook user is 22 years old [8], and has 138 friends on average [6]. Facebook users have made 140.3 billion friend connections and used over 1.13 trillion likes [8]. Moreover, every 60 seconds Facebook users post 510,000 comments, update 293,000 statuses, and upload 136,000 photos [22]. One of the main reasons for Facebook popularity is its third-party applications [26]. The Facebook application platform popularity is growing rapidly; for example, more than 20 million Facebook applications are installed every day [24]. There are many different kinds of Facebook applications, such as utility, productivity, and even educational applications [5]. According to Nazir et al. [21], the most popular applications are games; approximately 230 million people play games on Facebook every month [13]. The most popular games such as "Candy Crush Saga" have more than 2.7 million daily active users [19].

In recent years, hackers and spammers have found Facebook applications to be an efficient platform for spreading malware and spam. Moreover, recent research has found that at least 13% of applications on Facebook are malicious [26]. Recently, Rahman et al. [26] described in their study that spammers use Facebook applications to lure their victims into clicking on specific malicious links. Additionally, hackers can take advantage of Facebook application platform properties to: (a) find their next potential targets across a large base of potential users [26], (b) use the trust between friends to infect more users [23], (c) exploit the application developers API [4] to collect information, such as personal information, photos, tags, posts, chat threads, etc., and (d) use the user's credentials to publish posts or to spread spam, advertising and phishing under the name of a legitimate user [17].

To deal with this problem, we developed an add-on that is part of the Social Privacy Protector (SPP) [15], [16]. The purpose of this add-on is to educate and to increase user awareness about the threat that lurks in social applications. Our add-on notifies users of their current number of applications. By notifying the users, we encourage removal of third-party applications that are not in use by the user.

In this study, we utilize a dataset that was collected by our dedicated Firefox add-on. It contains information about 44,541 different occasions collected from 2,945 users between May 2012 and June 2013. Each entry in our dataset consists of the user id, application number and date.

In our previous study [16], we presented preliminary statistics on this dataset. We analyzed this dataset and discovered that users who used the SPP add-on for application removal, removed more than 50% of all their installed applications one day after its installation. These results indicate that in many cases the installed applications are unwanted or unneeded applications. In this study, we perform a temporal analysis of the installed application data that was collected for a longer period of time red than in our previous study. We discovered that within the first week after the add-on's initial use, the user's number of applications decreased by 12.1% (see Table I) on average. Moreover, the application removal rate continued to grow up to 27.7% (see Table I) by an average of 63 days after the initial use. According to the results presented in this study, we can conclude that using our add-on made many users become more aware of the existence of unnecessary applications on their Facebook profiles.

The remainder of this paper is organized as follows. In Section II, we give a brief overview on different related solutions that help users in social networks protect themselves from malicious applications. In Section III, we describe our methodology for analyzing the information and how we used linear regression to predict the application number. In Section IV, we present our study's initial results. Finally, in Section V, we present our conclusions regarding the change in awareness that resulted from our add-on notifications and offer future research directions.

II. RELATED WORK

Recent reports [27], [28] have indicated that due to the growth of social network popularity, there also has been a massive rise in malicious activity and security threats to online social network users. In recent years, social network users, social network operators, security companies, and academic researchers have proposed solutions to increase the security and privacy of social networks users. In the remainder of this section, we describe notable solutions in the area.

A. Detecting Malicious Applications

Detection is the most standard way to deal with security and privacy problems. There are many works on this topic and many different ways to detect malware. For example, Rahman et al. [25] presented MyPageKeeper, a Facebook application that protects Facebook users from socware. MyPageKeeper is based on a Support Vector Machine (SVM) classifier that uses a main feature specific keyword occurrence in a post made by an application. MyPageKeeper was able to identify socware posts and alert the user with 97% accuracy, but was unable to detect malicious applications. Websense Defensio [11] is a Facebook application from Websense that monitors posts in a user's profile and determines whether they are legitimate, spam, or malicious. Defensio also uses SVM to detect malicious posts and in addition they could delete them. Abu-Nimeh et al. [11] used Defensio as a platform to study malicious links. They found that about 9% of the studied posts were spam or malicious. In 2012, Rahman, et al. [26] improved his previously mentioned work. Rahman, et al. developed the FRAppE: A tool that can identify malicious applications by using the application information as features. Some examples include the number of permissions required, the domain reputation of redirect URI, and others. FRAppE can detect malicious applications with 99.5% accuracy and a low false negative rate 4.1%.

B. Increasing User Awareness of the Threat

Another approach is informing the user about possible threats that can jeopardize his or her security and privacy. By using this approach, it is possible to stimulate users to react and to protect themselves. This tactic is about teaching users what kind of threats exist and how they should protect themselves.

In 2012, Xu et al. [29] chose to redesign the Facebook application authentication dialogue to increase user awareness

of the permissions that are required by applications. Recently, Fire et al. [15] presented the SPP software, a Firefox add-on and a Facebook application. This software contains three protection layers, which improve user privacy by implementing different methods. (1) *The friends layer* - suggests friends who might pose a threat and then restricts these friends' exposure to the user's personal information. (2) *The privacy settings layer* is based upon different types of social network usage profiles. (3) *The application layer* - alerts users to the number of installed applications on their Facebook profiles.

In our previous study [16], we presented initial results, which were based on a dataset from 1,676 users collected between the 27th of June, 2012, and the 10th of November, 2012. In this study, we perform a temporal analysis on our complete dataset, which contains information from 2,945 users between May 2012 and June 2013. Unlike our previous study, this is the first study where we focused on analyzing user application installation and removal from different aspects for long periods of time.

III. METHODS AND EXPERIMENTS

Our study's main goal is to study how awareness affects installation and removal of social applications. To perform our research, we collected data from the Facebook online social network by using the SPP Firefox add-on (see Figure 1), which collected the following information: (1) *Hashed User Id*, (2) *Installed Application Number* - the number of installed Facebook applications on the user's Facebook account, and (3) *Date* - the date when the information was collected. To avoid duplicate entries, we collected only one entry for each user per day. Our final collected dataset included 2,947 Facebook users with unique hashed user ids across 351 different days between the months of May 2012 and June 2013. On average, we obtained information of 63 days and 15.1 entries per user. In addition, we discovered that Facebook users have a mean of 41.86 applications.

Fig. 1: Warning about installed applications.

To perform temporal analysis of our dataset we carried out the following steps. First of all, we divided all the users into two groups as follows: (1) *Regular Users* - users whose number of installed Facebook applications did not decrease between the initial and last use of the add-on (2) *Add-on Users* - users who decreased the number of installed Facebook application between their initial and last use of the add-on. Next, for each user, we calculated $AppChangeRatio(u, d)$: the ratio between the change in the number of applications user, u, had after, d, days and the number of applications that, u, had when the add-on was initially installed. $AppChangeRatio(u, d)$ was calculated only for the first 63 days after the add-on was initially installed, where 63 days is the average add-on use period. The formal definition of the $AppChangeRatio$

function is as follows:

$$AppChangeRatio(u, d) := ACR(u, d) :=$$

$$\frac{AppNum(u, 0)\text{-}AppNum(u, d)}{AppNum(u, 0)} \quad (1)$$

Where $AppNum(u, d)$ is a function that returns the number of applications user, u, had at a day, d, after the add-on installation. Then, for each day, d, we calculated the average $ACR(u, d)$ for each group as follows:

$$AvgAppChange(u, d) := AvgAC(u, d) :=$$

$$\frac{\sum_{u \in Users} ACR(u, d)}{|Users|} \quad (2)$$

We performed *t-test* to reject the null hypothesis, which states that the mean number of applications of the *Regular Users* and the *Addon Users* is identical and the observed differences are merely random. To find correlation between the application number and time, we conducted a linear regression experiment for each of the groups. The independent variable is the number of days passed since the initial install. The dependent variable is the percentage change in the applications number since the initial install.

Lastly, we tested if there are specific days of the week when users install or remove more applications as opposed to other days. To test this we calculated the average change in application number for every day of the week, w, for all the users, u. The change in application number is the calculation of difference in the number of applications between sequel days, we defined *AppNumDelta* as:

$$AppNumDelta(u, d) := ANDelta(u, d) :=$$

$$AppNum(u, d) - AppNum(u, d + 1) \quad (3)$$

The results were divided into two cases:

$$I(w) = \frac{\sum_{\{wday(d)=w | ANDelta(u,d)>0\}} ANDelta(u, d)}{|\{u \in Users | wday(d) = w \wedge ANDelta(u, d) > 0\}|} \quad (4)$$

$$R(w) = \frac{\sum_{\{wday(d)=w | ANDelta(u,d)\leq 0\}} ANDelta(u, d)}{|\{u \in Users | wday(d) = w \wedge ANDelta(u, d) \leq 0\}|} \quad (5)$$

where $wday(d)$ returns the day of the week for d. $I(w)$ is the average number of applications installed on that day of the week, w, and $R(w)$ is the average number of applications removed on that day of the week, w.

IV. RESULTS

At the end of our analysis, our *Regular Users* group consisted of 2,545 users and the *Add-on Users* group consisted of 400 users, out of which 52 (13%) also used our Friends Analyzer application, which is responsible for identifying a user's friends, who may pose a threat to the user's privacy. Therefore, we can classify these 52 users as privacy concerned users. Further, this suggests that less than 2% of all Facebook users are highly aware of their privacy.

TABLE I: The Percentage Change in the Applications Number

Group	After 1 Day	After 7 Days	After 63 Days
Regular Users	2.02%	11.9%	40.7%
Add-on Users	-5.4%	-12.2%	-27.7%

Afterward, we performed a Welch Two Sample t-test to prove the assumption of difference between the two groups. The t-test confirms that the *Add-on Users* ($\mu = 0.236$, $\sigma_{dev} = 0.12$) and the *Regular Users* ($\mu = -0.19$, $\sigma_{dev} = 0.05$) have a significant difference in their mean values $t = 25.936$, $p - value < 2.2e - 16$. Next, we used the linear regression method to find the correlation between the change ratio in number of applications and period of time in days.
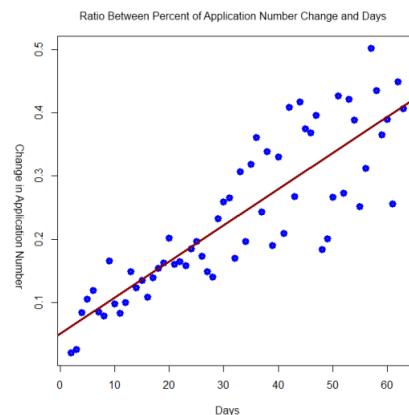

Fig. 2: The ratio between the percentage change in the applications number and days passed for the *Regular Users*. The solid line equation - $ApplicationChangePercent = 0.006 Days + 0.05$

As a result of the difference we have shown between the two groups, we divided the linear regression into two cases: (1) *Regular Users* - users whose number of applications increased or did not change (see Figure 2), (2) *Add-on Users* - Users whose number of applications decreased (see Figure 3).
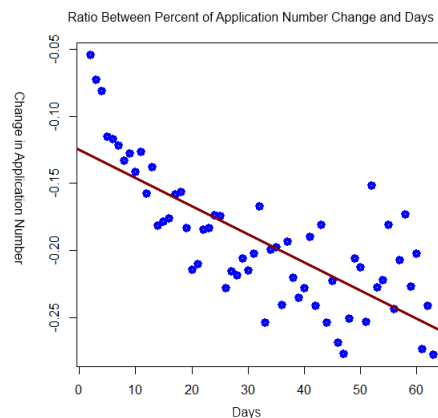

Fig. 3: The ratio between the percentage change in the number of applications and days passed of the *Add-on Users*. The solid line equation - $ApplicationChangePercent = -0.002 Days - 0.125$

Using linear regression, we received the following regres-

sion equations (see Figure 2 and 3):

$$RegularUsers : ApplicationChangePercent =$$
$$0.006 Days + 0.05 \quad (6)$$

where $R^2 = 0.736$, and $p - value = 2.2e - 16$.

$$AddonUsers : ApplicationChangePercent =$$
$$- 0.002 Days - 0.125 \quad (7)$$

where $R^2 = 0.57$, and $p - value = 1.351e - 12$

According to the linear regression results, the average Facebook user application number increases linearly over time, and the user installs about 7.15 applications each month.
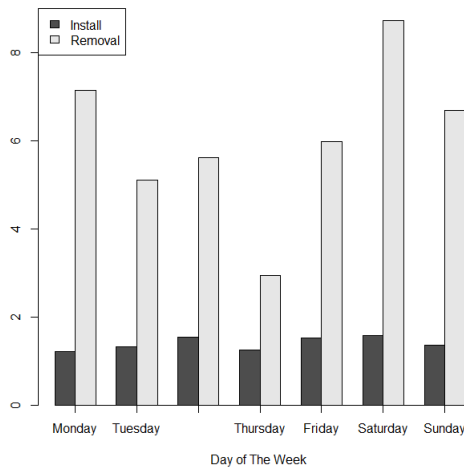

Fig. 4: Average application install and removal per day of the week.

In addition, we tested if there is a specific day of the week when users install or remove more applications than normal. We discovered that for both cases there is more activity on Saturdays and a significantly lower removal rate on Thursdays (see Figure 4).

V. CONCLUSIONS

In this study, we presented our initial methods and results in studying online social network applications with an aim of improving users safety and awareness. According to our results, it is possible to predict the number of applications a casual user has with high accuracy. Moreover, we discovered that users install many applications every month, and the average user installs approximately 7.15 applications each month. Currently, there are several malicious applications available. Rahman et al. states [26] that at least 13% of Facebook applications are malicious. By joining these two statistics together, we shockingly conclude that on average a Facebook user installs more than ten malicious applications each year. Our results show a possible solution to this problem. After the installation of our add-on, users become more aware of the number of installed applications they had. Users removed on average 12.1% of their applications after the first week, and they continued to remove even more applications afterward. Furthermore, by using the equations we discovered, it is possible to classify users who are not aware of the number

of applications they have.

In addition, we discovered that users install and remove more applications on Saturdays. We assume this is due to many users being at home instead of working on this day of the week. It is possible to use these results to notify users more often, in ways that are more noticeable and on specific dates regarding the danger of installing applications. For example, we can configure our add-on to give more alerts and add special alerts on Saturday.

The study presented in this paper is a work in progress with many available future directions. By gathering additional information about what kind of applications users tend to restrict, we can develop an algorithm for application removal recommendations. Moreover, when the same applications are restricted by many users, we can conclude with high likelihood that these applications are fake applications and recommend to Facebook and our users to remove these applications from the social network and their accounts. Another possible future direction is discovering the point in time when the *Add-on Users'* application numbers start increasing again, and at that point, to give the user a special warning regarding his or her number of applications.

REFERENCES

[1] Catster. http://www.catster.com/. [Retrieved: 9, 2013].
[2] Dogster. http://www.dogster.com/. [Retrieved: 9, 2013].
[3] Facebook. http://newsroom.fb.com/Key-Facts. [Retrieved: 9, 2013].
[4] Facebook apis. https://developers.facebook.com/docs/reference/apis. [Retrieved: 9, 2013].
[5] Facebook app center. https://www.facebook.com/appcenter/category/apps/?platform=web. [Retrieved: 9, 2013].
[6] Facebook newsroom. http://newsroom.fb.com/Key-Facts. [Retrieved: 9, 2013].
[7] Linkedin. http://www.linkedin.com/. [Retrieved: 9, 2013].
[8] One billion key metrics. http://newsroom.fb.com/imagelibrary/downloadmedia.ashx?MediaDetailsID=4227&SizeId=-1. [Retrieved: 9, 2013].
[9] xing. http://www.xing.com/. [Retrieved: 9, 2013].
[10] Yummypets. http://www.yummypets.com/. [Retrieved: 9, 2013].
[11] S. Abu-Nimeh, T. Chen, and O. Alzubi. Malicious and spam posts in online social networks. Computer, 44(9):23–28, 2011.
[12] M. Duggan and J. Brenner. The demographics of social media users2012. Pew Research Center, 2013.
[13] G. Finn. Facebooks mobile and web app hub, app center, officially launches. http://marketingland.com/facebooks-mobile-web-app-hub-app-center-officially-launches-13609. [Retrieved: 9, 2013].
[14] M. Fire, R. Goldschmidt, and Y. Elovici. Online social networks: Threats and solutions survey. arXiv preprint arXiv:1303.3764, 2013.
[15] M. Fire, D. Kagan, A. Elishar, and Y. Elovici. Social privacy protector-protecting users' privacy in social networks. In SOTICS 2012, The Second International Conference on Social Eco-Informatics, pages 46–50, 2012.
[16] M. Fire, D. Kagan, A. Elyashar, and Y. Elovici. Friend or foe? fake profile identification in online social networks. arXiv preprint arXiv:1303.3751, 2013.
[17] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao. Detecting and characterizing social spam campaigns. In Proceedings of the 10th ACM SIGCOMM conference on Internet measurement, pages 35–47. ACM, 2010.
[18] Itu. Trends in telecommunication reform 2012: Smart regulation for a broadband world. Technical report, 2012.
[19] I. Lunden. Facebook says it now has 235m monthly gamers, app center hits 150m monthly visitors. http://techcrunch.com/2012/08/14/facebook-says-it-now-has-235m-monthly-gamers-app-center-hits-150m-monthly-users/. [Retrieved: 9, 2013].

[20] M. Madden and K. Zickuhr. 65% of online adults use social networking sites. Technical report, 2011. [Retrieved: 9, 2013].

[21] A. Nazir, A. Waagen, V. S. Vijayaraghavan, C.-N. Chuah, R. M. D'Souza, and B. Krishnamurthy. Beyond friendship: modeling user activity graphs on social network-based gifting applications. In Proceedings of the 2012 ACM conference on Internet measurement conference, pages 467–480. ACM, 2012.

[22] D. Noyes. Top 20 valuable facebook statistics updated june 2013 . http://zephoria.com/social-media/top-15-valuable-facebook-statistics/. [Retrieved: 9, 2013].

[23] C. Patsakis, A. Asthenidis, and A. Chatzidimitriou. Social networks as an attack platform: Facebook case study. In Networks, 2009. ICN'09. Eighth International Conference on, pages 245–247. IEEE, 2009.

[24] C. Pring. social media statistics for 2012. http://thesocialskinny.com/100-social-media-statistics-for-2012/. [Retrieved: 9, 2013].

[25] M. S. Rahman, T.-K. Huang, H. V. Madhyastha, and M. Faloutsos. Efficient and scalable socware detection in online social networks. In Proceedings of the 21st USENIX conference on Security symposium, Security'12, pages 32–32, Berkeley, CA, USA, 2012. USENIX Association.

[26] M. S. Rahman, T.-K. Huang, H. V. Madhyastha, and M. Faloutsos. Frappe: detecting malicious facebook applications. In Proceedings of the 8th international conference on Emerging networking experiments and technologies, pages 313–324. ACM, 2012.

[27] Sophos. Security threat report 2013. Technical report, 2013.

[28] Websense. 2013 threat report. Technical report, 2013.

[29] H. Xu, N. Wang, and J. Grossklags. Privacy by redesign: Alleviating privacy concerns for third-party applications. In ICIA, 2012.