DETECTING AND CHARACTERIZING SOCIAL SPAM CAMPAIGNS

Mythili.J[1]
Assistant Professor
Computer Science and Engineering
Sengunthar College of Engineering, Tiruchengode, India

Naveena.C[2]
Computer Science and Engineering
Sengunthar College of Engineering
Tiruchengode, India

Abstract - With 20 million installs a day, third-party apps are a major reason for the popularity and addictiveness of Face book. Unfortunately, hackers have realized the potential of using apps for spreading malware and spam. The problem is already significant, as we find that at least 13% of apps in our dataset are malicious. So far, the research community has focused on detecting malicious posts and campaigns. In this paper, we ask the question: Given a Face book application, can we determine if it is malicious? Our key contribution is in developing FR AppE— Face book's Rigorous Application Evaluator—arguably the first tool focused on detecting malicious apps on Face book. To develop FRAppE, we use information gathered by observing the posting behaviour of 111K Face book apps seen across 2.2 million users on Face book. First, we identify a set of features that help us distinguish malicious apps from benign ones. For example, we find that malicious apps often share names with other apps, and they typically request less permission than benign apps. Second, leveraging these distinguishing features, we show that FRAppE can detect malicious apps with 99.5% accuracy, with no false positives and a high true positive rate (95.9%). Finally, we explore the ecosystem of malicious Face book apps and identify mechanisms that these apps use to propagate. Interestingly, we find that many apps collude and support each other; in our dataset, we find 1584 apps enabling the viral propagation of 3723 other apps through their posts. Long term, we see FRAppE as a step toward creating an independent watchdog for app assessment and ranking, so as to warn Face book users before installing

Keywords — Face book apps, malicious, online social networks, spam.

I. INTRODUCTION

A social networking service is an online service, platform, or site that focuses on facilitating the building of social networks or social relations among people who, for example, share interests, activities, backgrounds, or real-life connections. A social network service consists of a representation of each user (often a profile), his/her social links, and a variety of additional services. Most social network services are web-based and provide means for users to interact over the Internet, such as e-mail and instant messaging. Online community services are sometimes considered as a social network service, though in a broader sense, social network service usually means an individual-centered service whereas online community services are group-centered. Social networking sites allow users to share ideas, activities, events, and interests within their individual networks.

1.2 SOCIAL NETWORK CATEGORIES

1.2.1 Social Connections
Keeping in touch with friends and family members is one of the greatest benefits of social networking. Here is a list of the most widely-used websites for building social connections online.

- *Face book: Arguably the most popular social media utility, Face book provides a way for users to build connections and share information with people and organizations they choose to interact with online.*
- *Twitter: Share your thoughts and keep up with others via this real-time information network.*
- *Google+: This relatively new entrant to the social connection marketplace is designed to allow users to build circles of contacts that they are able to interact with and that is integrated with other Google products*

- *MySpace: Though it initially began as a general social media site, MySpace has evolved to focus on social entertainment, providing a venue for social connections related to movies, music games and more.*

1.2.2. Multimedia Sharing

Social networking makes it easy to share video and photography content online. Here are some of the most popular sites for multimedia sharing.

- *YouTube: Social media platform that allows users to share and view video content*
- *Flickr: This site provides a powerful option for managing digital photographs online, as well as for sharing them with others.*

1.2.3. Professional

Professional social networks are designed to provide opportunities for career-related growth. Some of these types of networks provide a general forum for professionals to connect, while others are focused on specific occupations or interests. A few examples of professional social networks are listed below.

- *LinkedIn: As of November of 2011, LinkedIn had more than 135 million members, making it the largest online professional network. Participants have an opportunity to build relationships by making connections and joining relevant groups.*
- *Classroom 2.0: Social network specifically designed to help teachers connect, share and help each other with profession-specific matters.*

1.2.4. Informational

Informational communities are made up of people seeking answers to everyday problems. For example, when you are thinking about starting a home improvement project or want to learn how to go green at home, you may perform a web search and discover countless blogs, websites, and forums filled with people who are looking for the same kind of information. A few examples include:

- *Super Green Me: Online community where individuals interested in adopting green living practices can interact*
- *Do-It-Yourself Community: Social media resource to allow do-it-yourself enthusiasts to interact with each other*

1.2.5. Educational

Educational networks are where many students go in order to collaborate with other students on academic projects, to conduct research for school, or to interact with professors and teachers via blogs and classroom forums. Educational social networks are becoming extremely popular within the educational system today. Some examples of such educational social networks are listed below.

- *The Student Room: UK-based student community featuring a moderated message board and useful resources related to school*
- *The Math Forum: A large educational network designed to connect students with an interest in math, this site provides interaction opportunities for students by age group.*
- *ePALS School Blog: This international social network for K-12 students is designed to build international connections to promote world peace.*

1.2.6. Hobbies

One of the most popular reasons many people use the Internet is to conduct research on their favorite projects or topics of interest related to personal hobbies. When people find a website based on their favorite hobby, they discover a whole community of people from around the world who share the same passion for those interests. This is what lies at the heart of what makes social networks work, and this is why social networks that are focused on hobbies are some of the most popular. A few examples of hobby-focused social networking sites include:

- *Oh My Bloom: Social media site specifically for gardening enthusiasts. It features groups, forums, blogs, video content and more.*
- *My Place at Scrapbook.com: Designed specifically for scrapbooking enthusiasts, users can create profiles, share information, post updates and more.*

1.2.7. Academic

Academic researchers who want to share their research and review results achieved by colleagues may find academic-specific social networking to be quite valuable. A few of the most popular online communities for academics are:

- *Academia.edu: Users of this academic social network can share their own research, as well as follow research submitted by others.*
- *Connotea Collaborative Research: Online resource for scientists, researchers and clinical practitioners to find, organize and share useful information.*

_____

TABLE 1- SUMMARY OF THE DATASET COLLECTED BY MYPAGEKEEPER

Data set Name	# of apps	
	Benign	Malicious
D-Total	6273	6273
D-Sample	6067	2528
D- Summary	2257	491
D- Inst	6063	3227
D- Complete	2255	487

II. MALICIOUS APPS ECOSYSTEM

Our analysis in Section III shows that malicious apps are rampanton Face book and indicates that they do not operate in isolation. Indeed, we find that malicious apps collude at scale—many malicious apps share the same name, several of them redirect to the same domain upon installation, etc. These observed behaviors indicate well-organized crime, with a few prolific hacker groups controlling many malicious apps.A common way in which malicious apps collude is by having one app post links to the installation page of another malicious app. In this section, we conduct a forensics investigation on the malicious app ecosystem to identify and quantify the techniques used in this cross promotion of malicious apps.

III. PERFORMANCE METRICS

The quality of political leaning scores is measured under two criteria. Classification: One should be able to directly infer the liberal/conservative stance of a source i from her sign of x_i, i.e., it is liberal if $x_i < 0$, or conservative if $x_i > 0$. Taking $\{x_i\}$ as ground truth, we say source i is correctly classified if the signs of x_i and x_i agree.[15] Classification performance is measured using the standard metrics of accuracy, precision, recall and F1 score.

Rank correlation: The set of scores $\{x_i\}$ induce a ranking of the sources by their political leaning. This ranking should be close to that induced by the ground truth scores $\{x_i\}$. We measure this aspect of performance using Kendall's τ, which varies from -1 (perfect disagreement) to 1 (perfect agreement). Given the modularity maximization method used here is analogous to the above eigenvector baseline; we treat its output as political leaning scores and report its performance. We also experimented with synchronous soft label propagation [54], similar to the algorithm in [19],but it did not produce better results. Table 2 reports the evaluation results. Our algorithm, in combining information from A, S and y, performs significantly better than all other algorithms in terms of Kendall's τ, F1 score and accuracy. We also observe that if no matrix scaling is applied in constructing W, the algorithm tends to assign all $\{x_i\}$ (except those of anchors) to the same sign, resulting in poor classification performance. In the remaining of this paper, we focus on the political leaning scores computed using cosine similarity.

TABLE 2- TOP MALICIOUS APPS IN D-SAMPLE DATASET

App ID	App Name	Post Count
236589745256987	What Does Your Name Mean?	1006
159874556255865	Free Phone Calls	793
235558955699555	The App	546
298975554123255	WhoSStalking?	434
142265865895611	FarmVile	210

IV. QUANTITATIVE STUDY

Worldwide Connectivity

No matter if you are searching for a former college roommate, your first grade teacher, or an international friend, no easier or faster way to make a connection exists than social media. Although Face book, Twitter, LinkedIn and Pinterest are probably the most well-known social networking communities, new websites are popping up regularly that let people connect and interact over the Web. With each of these sites, individuals can make new friends, build business connections or simply extend their personal base by connecting and interacting with friends of friends - which can have a multiplying effect.

V.CONCLUSION

Applications present convenient means for hackers to spread malicious content on Face book. However, little is understood about the characteristics of malicious apps and how they operate. In this paper, using a large corpus of malicious Face book apps observed over a 9-month period, we showed that malicious apps differ significantly from benign apps with respect to several features.

_____

For example, malicious apps are much more likely to share names with other apps, and they typically request less permission than benign apps. Leveraging our observations, we developed FRAppE, an accurate classifier for detecting malicious Face book applications. Most interestingly, we highlighted the emergence of app-nets—large groups of tightly connected applications that promote each other. We will continue to dig deeper into this ecosystem of malicious apps on Face book, and we hope that Face book will benefit from our recommendations for reducing the menace of hackers on their platform.

VI. REFERENCES

[1].C. Pring, "100 social media statistics for 2012," 2012 [Online]. Available: http://thesocialskinny.com/100-social-media-statistics-for-2012/

[2].Face book, Palo Alto, CA, USA, "Face book Opengraph API," [Online]. Available: http://developers.Face book.com/docs/reference/api/] "Wiki: Face book platform," 2014 [Online]. Available: http://en.wikipedia.org/wiki/Face book_Platform

[3]."Pr0file stalker: Rogue Face book application," 2012 [Online]. Available: https://apps.Face book.com/mypagekeeper/?status=scam_report-fb_survey_scam_pr0file_viewer_2012_4_4

[4]."Whiich cartoon character are you—Face book survey scam," 2012 [Online]. Available: https://apps.Face book.com/mypagekeeper/?status=scam_report_fb_survey_scam_whiich_cartoon_character_are_you_2012_03_30

[5].G. Cluley, "The Pink Face book rogue application and survey scam," 2012 [Online]. Available: http://nakedsecurity.sophos.com/2012/02/27/pink-Face book-survey-scam/

[6].D. Goldman, "Face book tops 900 million users," 2012 [Online]. Available: http://money.cnn.com/2012/04/23/technology/Face bookq1/index.htm

[7].R. Naraine, "Hackers selling $25 toolkit to create malicious Face book apps," 2011 [Online]. Available: http://zd.net/g28HxI