International Journal of Applied Sciences, Engineering and Management
ISSN 2320 – 3439, Vol. 04, No. 01, January 2015, pp. 25 – 29

Detecting Malicious Facebook Applications

P. Ramya Latha[1], P. Ramesh Babu[2]

[1]M. Tech Student, Dept of CSE, Amrita Sai Institute of Science and Technology, Paritala, Krishna-521180.
[2]Assistant Professor, Dept of CSE, Amrita Sai Institute of Science and Technology, Paritala, Krishna-521180.

Abstract: With 20 million installs a day, third-party apps are a main reason for the reputation and addictiveness of Facebook. Unluckily, hackers have realized the potential of using apps for scattering malware and spam. The problem is already major, as we find that at least 13% of apps in our dataset are malicious to date, the research community has focused on detecting malicious posts and campaigns. In this paper, we ask the question: given a Facebook application, can we decide if it is malicious? Our key contribution is in developing FRAppE Facebook's Rigorous Application Evaluator arguably the first tool focused on finding malicious apps on Facebook. To develop FRAppE, we use information gathered by observing the posting behavior of 111K Facebook apps seen across 2.2 million users on Facebook. First, we identify a set of features that aids us distinguish malicious apps from benign ones. For example, we discover that malicious apps often share names with other apps, and they typically request fewer permissions than benign apps. Second, leveraging these distinguishing features, we demonstrate that FRAppE can identify malicious apps with 99.5% accuracy, with no false positives and a low false negative rate (4.1%). Finally, we explore the ecosystem of malicious Facebook apps and recognize mechanisms that these apps use to spread interestingly, we find that many apps collude and support each other; in our dataset, we find 1,584 apps enabling the viral propagation of 3,723 other apps through their posts. Long-term, we see FRAppE as a step towards creating an independent watchdog for app assessment and ranking, so as to warn Facebook users before installing apps.

***Keywords**: Facebook Apps, Malicious Apps, Profiling Apps, Online Social Networks*

INTRODUCTION

Online social networks (OSN) make and encourage third party applications (apps) to improve the user experience on these platforms. Such enhancements consist of interesting or entertaining ways of communicating among online friends, and diverse activities such as playing games or listening to songs. For example, Facebook provides developers an API that facilitates app integration into the Facebook user-experience. There are 500K apps available on Facebook , and on average, 20M apps are installed every day. Furthermore, many apps have obtained and maintain a large user base. For instance, FarmVille and CityVille apps have 26.5M and 42.8M users to date. Recently, hackers have started taking advantage of the popularity of this third-party apps platform and deploying malicious applications. Malicious apps can provide a lucrative business for hackers, given the popularity of OSNs, with Facebook leading the way with 900M active users. There are many ways that hackers can benefit from a malicious app: (a) the app can reach large numbers of users and their friends to spread spam, (b) the app can obtain users' personal information such as email address, home town, and gender, and (c) the app can "re-produce" by making other malicious apps popular. To make matters worse, the deployment of malicious apps is simplified by ready-to-use toolkits starting at $25. In other words, there is motive and opportunity, and as a result, there are many malicious apps spreading on Facebook every day. Despite the above worrisome trends, today, a user has very limited information at the time of installing an app on Facebook. In other words, the problem is: given an app's identity number (the unique identifier assigned to the app by Facebook), can we detect if the app is malicious? Currently, there is no commercial service, publicly-available information, or research-based tool to advise a user about the risks of an app. As we show in Sec. 3, malicious apps are widespread and they easily spread, as an infected user jeopardizes the safety of all its friends. So far, the research community has paid little attention to OSN apps specifically. Most research related to spam and malware on Facebook has focused on detecting malicious posts and social spam campaigns. A recent work studies how app permissions and community ratings compare to privacy risks of Facebook apps. Finally, there are some community-based feedback driven efforts to grade applications, such as Whatapp though these could be very powerful in the future, up to now they have acknowledged little adoption.

SYSTEM ANALYSIS

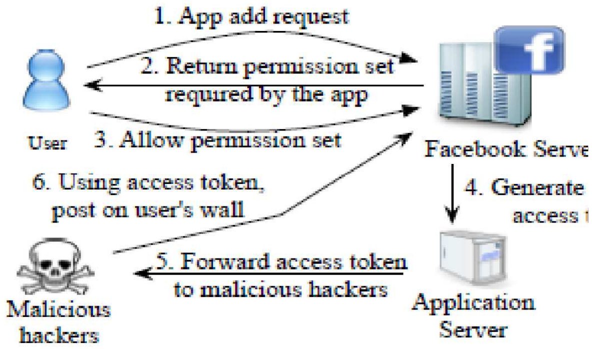
Existing System:

Hackers have begun taking advantage of the popularity of this third-party apps platform and deploying malicious applications. Malicious apps can provide a lucrative business for hackers, given the popularity of OSNs, with Facebook leading the way with 900M active users . There are many ways that hackers get benefit from a malicious app:(a) the app can attain large numbers of users and their friends to multiply spam, (b) the app can obtain users' personal information such as email address, home town, and gender, and (c) the app can "re-produce" by making other malicious apps popular.

Proposed System:

In this work, we develop FRAppE, a suite of efficient classification techniques for finding whether an app is malicious or not. To build FRAppE, we use data from My Page Keeper, a security app in Facebook that wathes the Facebook profiles of 2.2 million users. We examine 111K apps that made 91 million posts over nine months. This is arguably the first comprehensive study focusing on malicious Facebook apps that focuses on quantifying, profiling, and understanding malicious apps, and synthesizes this information into an effective detection approach.

SYSTEM ARCHITECTURE


1. App add request
2. Return permission set required by the app
User
3. Allow permission set
Facebook Serve
4. Generate access t
6. Using access token, post on user's wall
5. Forward access token to malicious hackers
Malicious hackers
Application Server

RELATED WORK

This is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, ten next steps are to determine which operating system and language can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. This support can be obtained from senior programmers, from book or from websites. Before building the system the above consideration r taken into account for developing the proposed system. This section provides background to the research through a review of some of the literature on privacy. The literature review is focused on those areas central to the scope of this research.

Detecting spam on OSNs. Gao et al. analyzed posts on the walls of 3.5 million Facebook users and showed that 10% of links posted on Facebook walls are spam. They also presented techniques to identify compromised accounts and spam campaigns. In other work, Gao et al. and Rahman et al. develop efficient techniques for online spam filtering on OSNs such as Facebook. While Gao et al. rely on having the whole social graph as input, and so, is usable only by the OSN provider, Rahman et al. develop a third-party application for spam detection on Facebook. Others present mechanisms for detection of spam URLs on Twitter. In contrast to all of these efforts, rather than classifying individual URLs or posts as spam, we focus on identifying malicious applications that are the main source of spam on Facebook. Detecting spam accounts. Yang et al and Benevenuto et al. developed techniques to identify accounts of spammers on Twitter. Others have proposed a honey-pot based approach to detect spam accounts on OSNs. Yardi et al. analyzed behavioral patterns among spam accounts in Twitter. Instead of focusing on accounts created by spammers, our work enables detection of malicious apps that propagate spam and malware by luring normal users to install them. App permission exploitation. Chia et al. investigated the privacy intrusiveness of Facebook apps and concluded that currently available signals such as community ratings, popularity, and external ratings such as Web of Trust (WOT) as well as signals from app developers are not reliable indicators of the privacy risks associated with an app. Also, in keeping with our observation, they found that popular Facebook apps tend to request more permissions. They also found that 'Lookalike' applications that have names similar to popular applications request more permissions than is typical. Based on a measurement study across 200 Facebook users, Liu et al. showed that privacy settings in Facebook rarely match users' expectations. To address the privacy risks associated with the use of Facebook apps, some studies propose a new application policy and authentication dialog. Makridakis et al. use a real application named 'Photo of the Day' to demonstrate how malicious apps on Facebook can launch DDoS attacks using the Facebook platform.King et al. conducted a survey to understand users' interaction with Facebook apps. Similarly, Gjoka et al. study the user reach of popular Facebook applications. On the contrary, we quantify the prevalence of malicious apps, and develop tools to identify malicious apps that use several features beyond the required permission set. App rating efforts. Stein et al. describe Facebook's Immune System (FIS), a scalable real-time adversarial learning system deployed in Facebook to protect users from malicious activities. However, Stein et al. provide only a high-level overview about threats to the Facebook graph and do not provide any analysis of the system. Furthermore, in an attempt to balance accuracy of detection with low false positives, it appears that facebook has recently softened their controls for handling spam apps . Other Facebook applications that defend users against spam and malware do not provide ratings for apps on Facebook. Whatapp collects community reviews about apps for security, privacy and openness. However, it has not attracted much reviews (47 reviews available) to date. To the best of our knowledge, we are the first to provide a classification of Facebook apps into malicious and benign categories.

The concept of privacy

What is privacy? It is an almost customary feature of any analysis of privacy to begin with a disclaimer about the inherent difficulty of defining exactly what 'privacy' is and disaggregating its various dimensions. It is something that is taken for granted and most people would have a sense of what privacy is but have difficulty putting it into words. The concept and meaning of privacy has long been debated by philosophers, social scientists, academic lawyers and other scholars. All definitions, to some extent, are based on assumptions about individualism and about the distinction between the realms of civil society and the state. However, many gloss over essential cultural, class-related and gender differences. Literature on privacy tends to give readers an overwhelming sense that privacy is a deeply contested concept, which often varies according to context and environment. (Bennett & Grant, 1999)

According to Bennett and Raab (2003), in Western culture, the modern claim to privacy and the contemporary justification for information privacy as a public policy goal was derived from a notion of a boundary between the individual and other individuals, and between the individual and the state. This concept of privacy rests on a construct of society as comprising relatively autonomous individuals and on notions of differences between the privacy claims and interests of different individuals. According to John Stuart Mill (as cited in Bennett & Raab, 2003), there should be certain 'self-regarding' activities of private concern, contrasted with 'other-regarding' activities to community interest and regulation. Shils (as cited in Bennett & Raab, 2003) argued that privacy is essential for the strength of American pluralistic democracy because it bolsters the boundaries between competing and countervailing centres of power. Dr Alan Westin, a leading academic (whose book *Privacy and Freedom* has shaped virtually all current thinking about privacy as a public issue), reinforced the importance of privacy for liberal democratic societies – in contrast to totalitarian regimes:

A balance that ensures strong citadels of individual and group privacy and limits both disclosure and surveillance is a prerequisite for liberal democratic societies. The democratic society relies on publicity as a control over government, and on privacy as a shield for group and individual life. Westin also addresses the specific functions that privacy plays. It promotes freedom of association. It shields scholarship and science from unnecessary interference by government. It permits the use of a secret ballot and protects the voting process by forbidding government surveillance of a citizen's past voting record. It restrains improper police conduct such as unreasonable search and seizure. It also serves to shield those institutions, such as the press, that operate to keep government accountable.

In a seminal law review article Samuel Warren and Louis Brandeis (1890) defined privacy simply as "the right to be let alone" – to go about life free from unreasonable interference by external forces.

Privacy has also been defined comprehensively:

Privacy is a concept related to solitude, secrecy, and autonomy, but it is not synonymous with these terms; for beyond the purely descriptive aspects of privacy as isolation from the company, the curiosity, and the influence of others, privacy implies a normative element: the right to exclusive control of access to private realms… the right to privacy asserts the sacredness of the person;… any invasion of privacy constitutes an offence against the rights of the personality – against individuality, dignity, and freedom. Arnold Simmel .Privacy can be divided into the following facets Territorial privacy – concerning the setting of limits on intrusion into the domestic and other environments such as the workplace or public space.

- Privacy of the person – this is concerned with protecting a person against undue interferences such as physical searches and drug testing, and information that violates his or her moral sense;
- Privacy of communications, covering the security and privacy of mail, telephones, email and other forms of communication;
- Privacy in the information context – this deals with the gathering, compilation and selective dissemination of personal information such as credit data and medical records.

The discourse on privacy as a policy issue has largely focused on information privacy and it is this facet of privacy that this research project will focus on. In this sense, privacy can be defined as "the claim of individuals, groups or institutions to determine for themselves when, how and to what extent information about them is communicated to others." (Westin, 1967, p7)

However, the rise to prominence of Internet communications and e-commerce has led to privacy of communications (and transmission) attracting more attention and concern. The increased concern with privacy of communications has caused some confusion between the meanings of information privacy and information security and the terms are often used interchangeably. As Clarke noted (as cited in Bennett & Raab, 2003), the term 'privacy' is used by some people to refer to the security of data or security of data during transmission as protection against various risks, such as data being accessed or modified by unauthorized persons. These aspects, however, are only a small fraction of the considerations within the field of 'information privacy'. That is, data security is a necessary but not sufficient condition for information privacy. An organization might keep the personal information it collects highly secure, but if it should not be collecting that information in the first place, the individual's information privacy rights are clearly violated.

IMPLEMENTATION MODULES

Malicious and benign app profiles significantly differ:

We systematically profile apps and show that malicious app profiles are significantly different than those of benign apps. A striking observation is the "laziness" of hackers; many malicious apps have the same name, as 8% of unique names of malicious apps are each used by more than 10 different apps (as defined by their app IDs). Overall, we profile apps based on two classes of features: (a) those that can be obtained on-demand given an application's identifier (e.g., the permissions required by the app and the posts in the application's profile page), and (b) others that require a cross-user view to aggregate information across time and across apps (e.g., the posting behavior of the app and the similarity of its name to other apps).

The emergence of AppNets: apps collude at massive scale:

We conduct a forensics investigation on the malicious app ecosystem to identify and quantify the techniques used to promote malicious apps. The most interesting result is that apps collude and collaborate at a massive scale. Apps promote other apps via posts that point to the "promoted" apps. If we describe the collusion relationship of promoting-promoted apps as a graph, we find

1,584 promoter apps that promote 3,723 other apps. Furthermore, these apps form large and highly-dense connected components, Furthermore, hackers use fast-changing indirection: applications posts have URLs that point to a website, and the website dynamically redirects to many different apps; we find 103 such URLs that point to 4,676 different malicious apps over the course of a month. These observed behaviors indicate well-organized crime: one hacker controls many malicious apps, which we will call an AppNet, since they seem a parallel concept to botnets.

Malicious hackers impersonate applications:
We were surprised to find popular good apps, such as 'FarmVille' and 'Facebook for iPhone', posting malicious posts. On further investigation, we found a lax authentication rule in Facebook that enabled hackers to make malicious posts appear as though they came from these apps.

FRAppE can detect malicious apps with 99% accuracy:
We develop FRAppE (Facebook's Rigorous Application Evaluator) to identify malicious apps either using only features that can be obtained on-demand or using both on-demand and aggregation based app information. FRAppE Lite, which only uses information available on-demand, can identify malicious apps with 99.0% accuracy, with low false positives (0.1%) and false negatives(4.4%). By adding aggregation-based information, FRAppE can detect malicious apps with 99.5% accuracy, with no false positives and lower false negatives (4.1%).

INPUT DESIGN AND OUTPUT DESIGN

INPUT DESIGN The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

- What data should be given as input?
- How the data should be arranged or coded?
- The dialog to guide the operating personnel in providing input.
- Methods for preparing input validations and steps to follow when error occur.

OBJECTIVES
1. Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.
2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.
3. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow

OUTPUT DESIGN
A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.
1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.
2. Select methods for presenting information.
3. Create document, report, or other formats that contain information produced by the system.
The output form of an information system should accomplish one or more of the following objectives.

- Convey information about past activities, current status or projections of the
- Future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action.

CONCLUSION AND FUTURE WORK

Applications current a convenient means for hackers to spread malicious content on Facebook. However, little is understood about the characteristics of malicious apps and how they function In this work, using a large corpus of malicious Facebook apps observed over a nine 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 fewer permissions than benign apps. Leveraging our observations, we developed FRAppE, an accurate classifier for finding malicious Facebook applications. Most interestingly, we highlighted the emergence of AppNets large groups of tightly connected applications that promote each other. We will go on with to dig deeper into this ecosystem of malicious apps on Facebook, and we expect that Facebook will benefit from our recommendations for sinking the menace of hackers on their platform.

REFERENCES

[1] Besmer, H. R. Lipford, M. Shehab, and G. Cheek. Social applications: exploring a more secure framework. In SOUPS, 2009.

[2] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2, 2011.

[3] P. Chia, Y. Yamamoto, and N. Asokan. Is this app safe? A large scale study on application permissions and risk signals. In WWW, 2012.

[4] F. J. Damerau. A technique for computer detection and correction of spelling errors. Commun. ACM, 7(3), Mar. 1964.

[5] H. Gao, Y. Chen, K. Lee, D. Palsetia, and A. Choudhary. Towards online spam filtering in social networks. In NDSS, 2012.

[6] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao. Detecting and characterizing social spam campaigns. In IMC, 2010.

[7] M. Gjoka, M. Sirivianos, A. Markopoulou, and X. Yang. Poking facebook: characterization of osn applications. In Proceedings of the first workshop on Online social networks, WOSN, 2008.

[8] J. King, A. Lampinen, and A. Smolen. Privacy: Is there an app for that? In SOUPS, 2011.

[9] A. Le, A. Markopoulou, and M. Faloutsos. Phishdef: Url names say it all. In Infocom, 2010.

[10] K. Lee, J. Caverlee, and S. Webb. Uncovering social spammers: social honeypots + machine learning. In SIGIR, 2010.

[11] S. Lee and J. Kim. Warningbird: Detecting suspicious urls in twitter stream. In NDSS, 2012.

[12] Y. Liu, K. P. Gummadi, B. Krishnamurthy, and A. Mislove. Analyzing facebook privacy settings: user expectations vs. reality. In IMC, 2011.

[13] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker. Beyond blacklists: learning to detect malicious web sites from suspicious urls. In KDD, 2009.

[14] A. Makridakis, E. Athanasopoulos, S. Antonatos, D. Antoniades, S. Ioannidis, and E. P. Markatos.

[15] Understanding the behavior of malicious applicationsin social networks. Netwrk. Mag. of Global Internetwkg., 2010.

[16] M. S. Rahman, T.-K. Huang, H. V. Madhyastha, and M. Faloutsos. Efficient and Scalable Socware Detection in Online Social Networks. In USENIX Security, 2012.

[17] T. Stein, E. Chen, and K. Mangla. Facebook immune system 100 social media statistics for 2012. 11MillionBulkemailaddressesforsaleSaleprice$90http://www.allhomebased.com/BulkEmailAddresses.htm.

[18] Apppiggybackingexample.https://apps.facebook.com/mypagekeeper/?status=scam_report_fb_survey_scam_Converse_shoes_2012_05_17_boQ.

[19] Application authentication flow using oauth 2.0.http://developers.facebook.com/docs/authentication/.

[20] Bitdefender Safego. http://www.facebook.com/bitdefender.safego.

[21] bit.ly API. http://code.google.com/p/bitly-api/wiki/ApiDocumentation.

[22] Defensio Social Web Security. http://www.facebook.com/apps/application.php?id=177000755670.

[23] Facebook developers.https://developers.facebook.com/docs/appsonfacebook/tutorial/.

[24] Facebook kills App Directory, wants users to search for apps.http://zd.net/MkBY9k.

[25] Facebook Opengraph API. http://developers.facebook.com/docs/reference/api/.

[26] Facebook softens its app spam controls, introduces better tools for developers. http://bit.ly/LLmZpM.

[27] Facebook tops 900 million users. http://money.cnn.com/2012/04/23/technology/facebook-q1/index.htm.