International Journal For Technological Research In Engineering
Volume 3, Issue 9, May-2016 *ISSN (Online): 2347 - 4718*

A SURVEY ON RECOGNIZING MALEFICENT FACEBOOK APPLICATIONS

Gurudatta HR[1], Shilpa KA[2]
[1]M.Tech, [2]Assistant Professor, School of Computing and Information Technology
Reva University, Bengaluru, India

Abstract: ***Online social media administrations like Facebook witness an exponential increment in client movement when an occasion happens in this present reality. This movement is a blend of good quality content like information, personal views, opinions, comments, and also low quality content like rumors, spam, and other malicious content. In spite of the fact that, the great quality content makes online social media a rich wellspring of data, utilization of low quality content can corrupt client encounter, and have unseemly effect in this present reality. Malicious Web destinations are a foundation of Internet criminal exercises. Accordingly, there has been wide enthusiasm for creating frameworks to keep the end client from going by such destinations. Social systems administration has turned into a mainstream route for users to meet and collaborate online. Users invest a lot of energy in well known social system stages, (for example, Facebook, MySpace, or Twitter), storing and sharing an abundance of individual data. Two of the main reasons for the lack of studies on Facebook are the strict privacy settings, and limited amount of data available from Facebook, as compared to Twitter. In this paper, we break down to which degree spam has entered social systems. All the more definitely, we investigate how spammers who target social systems administration locales work. Facebook is about times greater than its next greatest partner Twitter, and is at present, the biggest online social system on the planet. In this writing overview, we audit the current exploration work done on Facebook, and study the methods used to recognize and break down low quality content on Facebook. We likewise endeavor to comprehend the impediments postured by Facebook regarding accessibility of information for accumulation, and examination, and attempt to comprehend if existing systems can be utilized to distinguish and think about low quality content on Facebook***

Index terms: Online Social Media (OSM), Malicious content, Internet criminal exercise, Spammers

I. INTRODUCTION

In the course of the most recent couple of years, social networking sites have become one of the primary routes for users to keep track and communicate with their companions online. Sites, for example, Facebook, MySpace, and Twitter are reliably among the main 20 most-seen sites of the Internet. In addition, insights demonstrate that, by and large, users invest more energy in famous social networking sites than on any other site [2]. Shockingly, this numerous users likewise pulled in light of a legitimate concern for malicious gatherings. Specifically, spammers are continually searching for approaches to achieve new casualties with their spontaneous messages. From a security perspective, social systems have special attributes. Initially, information access and communication depends on trust. Users commonly share a significant measure of individual information with their companions. This information might be open or not. On the off chance that it is not open, access to it is directed by a system of trust [4]. Facebook is currently the largest social network on the Internet. Here Users connect by mutual consent and is used to keep in touch with friends and family, share what people are up to, and consume information about real world events. Usually, user profiles are not public, and the right to view a user's page is granted only after having established a relationship of trust (paraphrasing the Facebook terminology, becoming friends) with the user. When a user A wants to become friend with another user B, the platform first sends a request to B, who has to acknowledge that she knows A. When B confirms the request, a friendship connection with A is established. However, the users' perception of Facebook friendship is different from their perception of a relationship in real life. Most of the time, Facebook users accept friendship requests from persons they barely know, while in real life, the person asking to be friend would undergo more scrutiny. In the past, most Facebook users were grouped in networks, where people coming from a certain country, town, or school could find their neighbors or peers. The default privacy setting for Facebook was to allow all people in the same network to view each other's profiles. Thus, a malicious user could join a large network to crawl data from the users on that network. This data allows an adversary to carry out targeted attacks. In Facebook there are approximately 1.32 billion monthly active users, 4.75 billion posts were made per day and over 300 petabytes of data are stored. Here Spammers exploit context of event to lure victims into scams. Facebook Spammers make $200 million just by posting links. From a security point of view, social networks have unique characteristics. First, information access and interaction is based on trust. Users typically share a substantial amount of personal information with their friends. This information may be public or not. If it is not public, access to it is regulated by a network of trust. In this case, a user allows only friends to view the information. Unfortunately, social networking sites do not provide strong authentication mechanisms, and it is easy to impersonate a user and sneak into a person's network of trust. Moreover, it often happens that users, to gain

popularity, accept any friendship request they receive, exposing their personal information to unknown people. Networks of trust are important from a security point of view, because they are often the only mechanism that protects users from being contacted by unwanted entities. Another important characteristic of social networks is the different levels of user awareness with respect to threats. While most users have become aware of the common threats that affect the Internet, such as e-mail spam and phishing, they usually do not show an adequate understanding of the threats hidden in social networks. In particular, we look at three distinct areas, viz., a) attack and detection techniques with respect to malicious content on Facebook, and b) analysis of events using online social media data. At that point, we take a gander at the different limitations that Facebook posture, which makes event analysis and detection of malicious content on this network a difficult issue. Towards the end, we talk about the implications and research holes in identifying and analyzing malicious client generated content on Facebook amid events. abundance of information, and additionally the simplicity with which one can achieve.

II. RELATED WORKS

J. Ma, L. K. Saul, S. Savage, and G. M. Voelker [2] proposed a "Beyond blacklists: Learning to detect malicious Web sites from suspicious URLs" Malicious Web sites are a cornerstone of Internet criminal activities. As a result, there has been broad interest in developing systems to prevent the end user from visiting such sites. This paper, describes an approach to this problem based on automated URL classification, using statistical methods to discover the tell-tale lexical and host-based properties of malicious Web site URLs. These methods are able to learn highly predictive models by extracting and automatically analyzing tens of thousands of features potentially indicative of suspicious URLs. The resulting classifiers obtain 95-99% accuracy, detecting large numbers of malicious Web sites from their URLs, with only modest false positives. Issue based on robotized URL order, utilizing factual strategies to find the obvious lexical and host-based properties of malicious Web site URLs. These techniques can learn profoundly prescient models by separating and naturally analyzing a huge number of elements possibly characteristic of suspicious URLs. The subsequent classifiers acquire 95-99% precision, recognizing extensive quantities of malicious Web sites from their URLs, with just unobtrusive false positives [9]. A. Makridakis, E. Athanasopoulos, S. Antonatos, D. Antoniades, S. Ioannidis, and E. P. Markatos [3] proposed a "Understanding the behavior of malicious applications in social network", MyPageKeeper is a Facebook app designed for detecting malicious posts on Facebook. Once a user installs app, it periodically crawls posts from the user's wall and news feed. It then applies URL blacklists as well as custom classification techniques to identify malicious posts. My pagekeeper identifies social malware granularity of an individual post, without grouping together posts made by the given application. My pagekeeper determination of whether to flag that post does not take into account the application responsible for the post. The large fraction of posts monitored by my pagekeeper is not posted by any applications. Even among malicious posts identified by my pagekeeper it was not having any associated application. It relies on SVM (support vector machine). SVM evaluates every URL by combining information obtained from all posts containing that URL. Malicious post receives few like and comments. It identifies certain spam keywords such as FREE; DEAL etc and some of the posts tend to have similar text messages. Once URL is identified as malicious my pagekeeper marks all post containing the URL as malicious.

III. MALICIOUS CONTENT ON FACEBOOK

The popularity and reach of Facebook has additionally pulled in a great deal of spam, phishing, malware, and different sorts of malicious movement. Aggressors draw casualties into tapping on malicious connections indicating outer sources, and in proficient their network. These connections can be spread either through individual messages (chats), or through divider posts. To accomplish most extreme perceivability, aggressors want to post interfaces freely. Ordinarily, an assailant starts the assault by posting images with attention snatching reviews, which incite users to like, share, or remark on them to view them. The activities of preferring, remarking or sharing spread these pictures into the casualty's network. Once the image is spread, the casualty is diverted to a malicious website, which can advance taint her PC, or companions network through phishing, malware, or spyware.

A. Attack techniques

Keeping in mind the end goal to distinguish and contain malicious posts on Facebook, or any OSM, it is essential to investigate and comprehend the techniques that are, or can potentially be conveyed by attackers to spread such content. To this end, Patsakis et al. [Patsakis et al. 2009] portrayed how Facebook can be abused and changed over into an attack stage, with a specific end goal to increase some sensitive information, which can complete a flawless attacking against a client. Writers made a Facebook application for exhibit purposes that at first glance was a straightforward application, yet on the foundation it gathered helpful information. This application executed malicious code on the casualty's program, and gathered the IP location of the client casualty, the program form, the OS stage and whether some particular ports are open or shut. Fan et al. [Fan and Yeung 2010] proposed a virus model based on the application network of Facebook [3]. Writers also displayed the virus engendering with an email virus demonstrate and looked at the practices of virus spreading in Facebook and email network. Their discoveries revealed that while Facebook gives a stage to application designers, it also gives the same opportunity to virus spreading. Truth is told, the virus was found to spread quicker on the Facebook network if users invest more energy in it. The aftereffect of their recreation demonstrated that, despite the fact that a malicious Facebook application pulls in just a couple of users at the outset, it can even now spread quickly. That is on account of

users may trust their companions of Facebook and install the malicious application.

B. Detection techniques

Facebook has its own particular insusceptible framework to protect its users from undesirable, malicious content [Stein et al. 2011]. Researchers at Facebook constructed and sent a cognizant, scalable, and extensible real time framework to ensure their users and the social diagram. This framework performs real time checks and characterizations on each read and composes. Keeping in mind the end goal to recognize and contain malicious posts on Facebook, or any OSM, it is essential to investigate and comprehend the techniques that are, or can potentially be sent by attackers to spread such content. To this end, Patsakis et al. [Patsakis et al. 2009] depicted how Facebook can be misused and changed over into an attack stage, so as to increase some sensitive information, which can complete a flawless attacking against a client. Authors made a Facebook application for show purposes that at first glance was a straightforward application [6]; however on the foundation it gathered helpful information. This application executed malicious code on the casualty's program, and gathered the IP location of the client casualty, the program form, the OS stage and whether some particular ports are open or shut. This information was then transmitted to the authors over email. Huber et al. displayed a companion in-the-center attack through seizing session treats. Authors clarified how it was conceivable to imitate the casualty utilizing this system, and associate with the network without legitimate approval. Notwithstanding, this strategy was proposed in 2011, when utilizing HTTPS to interface with the website was optional. 13 Post 2013, all correspondence on Facebook utilizes encryption (HTTPS) as a matter of course, which implies that such attacks are not any more conceivable. Fan et al. [Fan and Yeung 2010] proposed a virus model based on the application network of Facebook. Actually, the virus was found to spread quicker on the Facebook network if users invest more energy in it.

C. Characteristics of Malicious content

The key characteristics of Malicious content on Facebook are of three aspects

*1)Textual contents and URLs***:** we found that the most common type of malicious posts in our dataset were the ones with URLs pointing to adult content and incidental nudity, and marked unsafe for children by WOT. The second most common type of malicious posts comprised of negative and questionable category URLs. These categories comprised of malware, phishing, scam, misleading claims or unethical, spam, hate, discrimination, potentially unwanted programs. Posts containing untrustworthy sources of information were the third most common type of malicious posts.

2) Entities posting malicious contents: Content on Facebook is generated by two types of entities – users and pages. Pages are public profiles specifically created for businesses, brands, celebrities, causes, and other organizations. Unlike users, pages gain "fans," people who choose to like a page. In our dataset, we identified pages by the presence of category field in the response returned by Graph API search, during the initial data collection process. The category field is specific to pages we used this field to differentiate between pages and user profiles. We found that pages were more active in posting malicious URLs as compared to legitimate URLs.

3)Metadata: There are various types of metadata associated with a post, for example, application used to post, time of post, type of post (picture / video / link), location etc. Metadata is a rich source of information that can be used to differentiate between malicious and legitimate users we also observed significant difference in the content types that constituted malicious and legitimate content. Over 50%of legitimate posts containing a URL were photos or videos whereas this percentage dropped to below 6% for malicious content. A large proportion of these photos and videos were uploaded on Facebook itself. This was one of the main reasons for facebook.com being the most common legitimate domain in our dataset. We used these, and some other features to train multiple machine learning algorithms for automatic detection of malicious content

IV. PROPOSED FRAMEWORK

In this paper, we develop FRAPPE, a suite of efficient classification techniques for identifying whether an application is malicious or not. To build FRAppE, we use data from MyPage- Keeper, a security app in Facebook. We find that malicious applications significantly differ from benign applications with respect to two classes of features: On-Demand Features and Aggregation-Based Features. We present two variants of our malicious app classifier—FRAppE Lite and FRAppE.

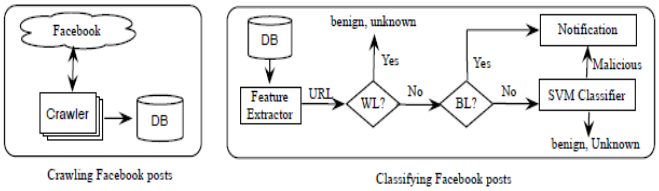

Fig 4.1.System Architecture of Proposed framework

FRAppE Lite is a lightweight version that makes use of only the application features available on demand. Given a specific app ID, FRAppE Lite crawls the on-demand features for that application and evaluates the application based on these features in real time. FRAppE—a malicious app detector that utilizes our aggregation-based features in addition to the on-demand features. On-deamand feature include app name, category, company and required permission set. Aggregation based feature include the similarity of app names and URL posted by application over time. FRAPPE focuses on quantifying, profiling and understanding malicious app and synthesizes information into an effective detection approach. Hence, FRAPPE redirects URLs, number of required permission set and use of different client ID in app installation URL. Some of the advantages of proposed framework are FRAPPE (Facebook

Rigorous Application Evaluator) is arguably is the tool to detect malicious apps and It provides security to users profiles from malicious apps.

V. CONCLUSION

In this survey, we explored various research attempts towards exploring the Facebook network, analyzing malicious content on it, and analyzing events on online social media in general. The aim of this survey was to look at relevant literature, which could aid in studying and combating malicious user generated content spread on Facebook during events. In this survey, we Investigated different research endeavors towards investigating the Facebook network, analyzing malicious content on it, and analyzing events on online social media in general. The aim of this survey was to look at relevant literature, which could aid in studying and combating malicious user generated content spread on Facebook during events. In order to keep this survey focused, we did not cover a variety of possibly relevant research areas including detection of compromised / fake accounts, and Sybil nodes in the Facebook network, detection of spam on other social networks like Twitter, credibility / trustworthiness of information of user generated content, and event detection in online social media. We also looked at the various challenges and limitations posed by Facebook. Apart from technical limitations, there exist various research gaps in existing literature, which are yet to be addressed and explored.

REFERENCES

[1] Lee and J. Kim. Warningbird: Detecting suspicious urls in twitter stream. In NDSS, 2012.
[2] 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.
[3] A. Makridakis, E. Athanasopoulos, S. Antonatos, D. Antoniades, S. Ioannidis, and E. P. Markatos. Understanding the behavior of malicious applications in social networks. Netwrk. Mag. of Global Internetwkg., 2010.
[4] 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..
[5] Lee, J. Caverlee, and S. Webb. Uncovering social spammers: social honeypots + machine learning. In SIGIR, 2010
[6] K. Thomas, C. Grier, J. Ma, V. Paxson, and D. Song. Design and Evaluation of a Real-Time URL Spam Filtering Service. In Proceedings of the IEEE Symposium on Security and Privacy, 2011.
[7] N. Wang, H. Xu, and J. Grossklags. Third-party apps on facebook: privacy and the illusion of control. In CHIMIT, 2011.
[8] H. Gao, Y. Chen, K. Lee, D. Palsetia, and A. Choudhary. Towards online spam filtering in social networks. In NDSS, 2012.
[9] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao. Detecting and characterizing social spam campaigns. In IMC, 2010.
[10] Y. Liu, K. P. Gummadi, B. Krishnamurthy, and A. Mislove. Analyzing facebook privacy settings: user expectations vs. reality. In IMC, 2011.

