ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.5

Design the Framework for Detecting Malicious Mobile Webpages in Real Time

Sneha .J.Tripathi[1], Prof. V.S Gangwani[2]
M.E.Student[1], Assistant Professor[2]
Department of Computer Science & Engineering
H.V.P.V College of Engineering, Amravati, India

Abstract:
Mobile specific webpages differ significantly from their desktop counterparts in content, layout and functionality. Accordingly, existing techniques to detect malicious websites are unlikely to work for such webpages. The disclosed technology includes techniques for identifying malicious mobile electronic documents, e.g. webpages or emails, based on static document features. The static features may include mobile-specific features, such as mobile web API calls, hosted mobile-specific binaries, no script content, or misleading URL tokens visible on a mobile specific inter face. The static features may instead or also include various JavaScript JS features, HTML features, and URL features detected in numbers outside ranges expected for desktop electronic documents. These features may be used With machine learning techniques to classify benign and malicious documents in real time. Recent years have witnessed the increasing threat of phishing attacks on mobile computing platforms. Finally, design a browser extension to protect users from malicious mobile websites in real-time. In doing so, this application provide the first static analysis technique to detect malicious mobile webpages.

Key Words: Mobile Apps, ranking fraud detection, historical ranking records, evidence aggregation, review, ranking and rating.

I. INRODUCTION

Internet connected mobile devices are going to outnumber humans [2]. Moreover, global mobile data traffic is expected to increase 13-fold between 2012 and 2017. Both platform-specific applications (“native apps”) and browser-based applications (“web apps”) enable mobile device users to perform security sensitive operations such as online purchases, bank transactions and accessing social networks. The distinction between native apps and web apps on mobile devices is increasingly being blurred. HTML5 becomes universally deployed and mobile web apps directly take advantage of device features such as the camera, microphone and geolocation, the difference between native and web apps will vanish almost entirely. A recent study of smartphone usage shows that more people browse the Web than use native apps on their phone. The trend and the increasing use of web browsers on modern mobile phones warrant characterizing existing and emerging threats to mobile web browsing. Although a range of studies have focused on the security of native apps on mobile devices, efforts in characterizing the security of web transactions originating at mobile browsers are limited. Mobile web browsers have long underperformed their desktop counterparts. However, recent improvements in processing power and bandwidth have spurred significant changes in the ways users experience the mobile web. Modern mobile browsers provide rich functionality equivalent to their desktop counterparts using web technologies such as HTML, JavaScript, and CSS. Furthermore, browsers on mobile platforms now build on the same or similarly capable rendering engines used by many desktop browsers. Mobile users are three times more likely to access phishing websites than desktop users [3]. Mobile phishing is particularly dangerous due to the hardware limitations of mobile devices and mobile user habits. We did a comprehensive study on the security vulnerabilities caused by mobile phishing attacks, including the web page phishing attacks, the application phishing attacks, and the account registry phishing attacks. Existing schemes designed for web phishing attacks on PCs cannot effectively address the various phishing attacks on mobile devices. Mobile devices are increasingly being used to access the web [1]. However, in spite of significant advances in processor power and bandwidth, the browsing experience on mobile devices is considerably different. These differences can largely be attributed to the dramatic reduction of screen size, which impacts the content, functionality and layout of mobile web pages. Identify the malicious URLs based on dynamically extracted lexical patterns from URLs. They developed a new method to mine their URL patterns, which are not assembled using any pre-defined items and thus cannot be mined using any existing frequent pattern mining methods. It can provide new flexibility and capability malicious URLs algorithmically generated by malicious programs. Content, functionality and layout have regularly been used to perform static analysis to determine maliciousness in the desktop space. Features such as the frequency of iframes and the number of redirections have traditionally served as strong indicators of malicious intent. Due to the significant changes made to accommodate mobile devices, such assertions may no longer be true. For example, whereas such behavior would be flagged as suspicious in the desktop setting, many popular benign mobile web pages require multiple redirections before users gain access to content. Previous techniques also fail to consider mobile specific webpage elements such as calls to mobile APIs. For instance, links that spawn the phone’s dialer can provide strong evidence of the intent of the page. New tools are therefore necessary to identify malicious pages in the mobile web. The coming and the rising fame of systems, Internet, intranets and conveyed frameworks, security is getting to be one of the central purposes of exploration. Web substance is experiencing a critical change. Static features of mobile webpages derived from their HTML and JavaScript content, URL and advanced mobile specific capabilities. Our design detects a number of malicious mobile webpages not precisely detected by existing techniques such as Virus Total and Google Safe Browsing. Finally, we discuss the existing tools to detect mobile malicious webpages and phishing attack and build a browser extension.

II. LITERATURE SURVEY

Significant study has been done in the past few years on the security of mobile [1]. Discrepancies between mobile and desktop webpages demand investigation. Existing static analysis techniques and tools for detecting malicious webpages are focused on desktop web pages. Therefore, they are unable to detect mobile specific threats with high accuracy Study found that many existing tool work with desktop browsers. Author manually visited mobile specific known malicious webpages collected from Phish-Tank, from the Google Chrome mobile browser and observed that these webpages are flagged as malicious on the Chrome desktop browser, but not on the Chrome mobile browser whose users are the real targets of the mobile malicious webpages. kAYO [1] is an existing technique that detects mobile malicious webpages. kAYO makes these detections by measuring 44 mobile relevant features from webpages, out of which 11 are newly identified mobile specific features. kAYO provides 90% accuracy in classification, and detects a number of malicious mobile webpages in the wild that are not detected by existing techniques such as Google Safe Browsing and Virus Total. Finally, a browser extension was build using kAYO that provides real-time feedback to users. kAYO detects new mobile specific threats such as websites hosting known fraud numbers and takes the first step towards identifying new security challenges in the modern mobile web. Y. Zhang studied the problem that the biggest threat on the mobile web at present is believed to be phishing. The best known non-proprietary content-based approach to detect phishing webpages is Cantina. Cantina suffers from performance problems due to the time lag involved in querying the Google search engine. Moreover, Cantina does not work well on webpages written in languages other than English. Finally, existing techniques do not account for new mobile threats such as known fraud phone numbers that attempt to trigger the dialer on the phone. Dr. Jitendra Agrawal [9] studies malicious web page detection through classification Technique. Detection of malicious web has become a necessary and hot topic of research as numbers of internet users are increasing at a high pace. There are lots of challenges regarding this detection process. First the number of online URL is very large. Second web environment uses diverse platform and difficult to find security solution for them. Third now threats are become more and more complex and used various obfuscation techniques to bypass detection techniques. The existing detection techniques are focused only on single type of attacks only. New generated malicious web pages exploit multiple types of attacks for targeting the client. Cloaking type of attacks is difficult to detect because these web respond differently to browser and crawler. Size of web is a big challenge in the process. Paul C. van Oorschot [11] studies measuring of SSL indicators on mobile browsers. Although these browsers aim for equivalent functionality to traditional desktops, their smaller screen size has resulted in significant changes to the presentation and availability of SSL indicators. Their study presents the first large scale, cross-sectional measurement of this class of applications and compares the security indicators used in the overwhelming majority of mobile browsers to their traditional desktop counterparts.

Leyla Bilge and Engin Kirda [6] studied the way to find malicious domains using passive DNS analysis. The domain service (DNS) is a crucial component of the Internet. DNS provides a two-way mapping between domain names and their IP addresses. Just as DNS is a critical service for the functioning of benign Internet services, it has also started to play an important role for malicious activities. For example, bots resolve DNS names to locate their command and control servers, and spam mails contain URLs that link to domains that resolve to scam servers. Their study introduced EXPOSURE, a system that employs passive DNS analysis techniques to detect malicious domains. Manos Antonakakis [4] presented Notos, a dynamic reputation system for DNS. Davide Canali [5] proposed a filter for the large-scale detection of malicious web pages. A system whose aim is To provide a filter that can reduce the number of web pages that Need to be analyzed dynamically to identify malicious web pages. As malware on the Internet spreads and becomes more sophisticated, Anti-malware techniques need to be improved in order to Be able to identify new threats in an efficient, and, most important, automatic way. Adrienne Porter Felt [7] studied and examine the threat of phishing on mobile devices. A successful phishing attack has two parts: the user must be conditioned to enter her credentials in a certain setting, and the attacker must be able to imitate that setting. He studies real mobile applications and web sites to understand the scenarios in which users enter passwords on mobile phones, and then we propose attacks that subvert these scenarios. Many applications and web sites link to each other for the purpose of social sharing and payment, both of which require the user to enter her authentication credentials in contexts where the user has no way to identify who is receiving those credentials. Users are therefore likely accustomed to switching from one application to another and then entering their passwords into the second application, without any way to verify the authenticity of the second application. A malicious application can link the user to a social networking or payment web site, and then present the user with a fake login screen. Alternately, an attacker can intercept the interaction and substitute a fake login screen for the intended target. His research will motivate us in defenses against mobile phishing. Anh Le and Athina Markopoulou propose PhishDef, a system that performs proactive, on-the-fly classification of phishing URLs using only lexical features and the AROW algorithm. By using only lexical features, PhishDef reduces the page loading latency and avoids reliance on remote servers. By implementing the AROW algorithm, PhishDef achieves high classification accuracy, even with noisy data, while at the same time having low computation and memory requirements. Guang Xiang, Jason Hong, studied CANTINA+, a layered solution for phishing web page . They proposed CANTINA+, the most comprehensive feature-based approach in the literature including eight novel features, which exploits the HTML Document Object Model, search engines and third party services with machine learning techniques to detect phish. Moreover, they designed two filters to help reduce FP and achieve runtime speedup. The first is a near-duplicate phish detector that uses hashing to catch highly similar phish. The second is a login form filter, which directly classifies web pages with no identified login form as legitimate. CANTINA+ employs a hash-based filtering module to inspect web pages in the beginning of the layered pipeline to identify near-duplicate attacks in a fast and highly reliable fashion.

III. PROBLEM OF ANALYSIS

Existing static analysis techniques and tools for detecting malicious webpages are focused on desktop webpages.

Therefore, they are unable to detect mobile specific threats with high accuracy. Several webpages built specifically for mobile, return empty pages when rendered in a desktop browser. Thus, even existing dynamic analysis techniques that execute websites in desktop browsers on virtual machines, are ineffective on such mobile websites. Finally, signature based tools such as Google Safe Browsing currently only work with desktop browsers. Existing application kAYO is similar to those of existing malicious website detection tools using static analysis. Many comprehensive set of features makes it harder to evade kAYO, as seen from evaluation over a large dataset. Existing application statically crawled the top million websites of Alexa. But it did not collect webpages that use JavaScript to detect and redirect to the mobile webpage. It also missed the mobile webpages represented by ways other than the ones used by the top 1,000 websites. It doesn't gathering all mobile webpages from Alexa top one million. The potential of bad activity in the mobile web could increase yet further over time.
To address such existing problem we proposed a real-time detector. The design, structure and languages used to build native apps are usually more complex as compared to web apps. Here oue system is billed as web app. The goal of the proposed model is to ensure that average users make better security decisions of web apps. Nevertheless, we believe that maintaining significant overlap between the permission models for native and web apps will help average users in making informed security decisions. More importantly, security experts will be able to use the proposed model to design tools similar to the malware detection tools for native apps. Our proposal deals with restricting web apps access to the resources on a user's device. Our browser extension adds value for two reasons. First, the mobile specific design enables detection of new threats previously unseen by existing services. Second, building an extension allows immediate use of our technique. We developed a browser extension for mobile which informs users about the maliciousness and phising attacks of the webpages they intend to visit. Our goal was to build an extension that runs in real-time. Therefore, instead of running the feature extraction process in a mobile browser, we outsourced the processing intensive functions to a backend server. In our proposed work user enters the URL he wants to visit in the extension toolbar. The extension then sends the URL backend server over HTTPS. The server crawls the mobile URL and extracts static features from the webpage. This feature set is input to our model, which classifies the webpage as malicious or benign. Next it will check if there is any phishing attack on site. The output is then sent back to the user's browser in real-time. If the URL is not malicious and free from phishing attack according to our app, then it will open webpage in the browser automatically. Otherwise, a warning message is shown to the user recommending them not to visit the URL or visit on their own risk. Users of the extension will browse both mobile specific and desktop webpages since not all websites offer a mobile specific version.

IV. PROPOSED WORK

Proposed work includes the following –

- The proposed method focus on mobile specific threats. Proposed method work on the mobile specific specific webpages. Existing technique to detect malicious websites are unable to work on mobile. Here determination is based on the static as well as dynamic features.
- The proposed method is outlined in figure this system use URL to get malicious content. The proposed methodology also introduced the cross site scripting (XSS). Cross site scripting (XSS) is a type of computer security vulnerability typically found in web application. XSS enable attackers to inject client side scripts into web pages viewed by other users. XSS refers to client side code injection attack wherein an attacker can execute malicious script into a legitimate website or web application.
- Our application is use to check the malicious function and phishing site. Here SLD (second level domain) and OCR (optical character recognition) technique are also introduced. OCR is technique that convert image into text to detect valuable phishing attack.
- Here user first enters the URL he wants to visit then the system will match the URL with existing malicious URL. If URL match then it will show warning message otherwise go to the next step .
- Next it will read the HTML tags . If it found any iframe tags then it shows the warning message that the URL contain malicious function and add the URL in database else goto next step.
- Next our application will read for cross site scripting. If any unwanted script found then it will show a message that URL conation malicious function and add the URL in the database. Otherwise go to the next step.
- Finally we check for the phishing attack . Here SLD is use with the help of OCR technique to find the phishing site. Now if the SLD matches with the text extracted by the OCR then the site is safe otherwise a warning message will pop.

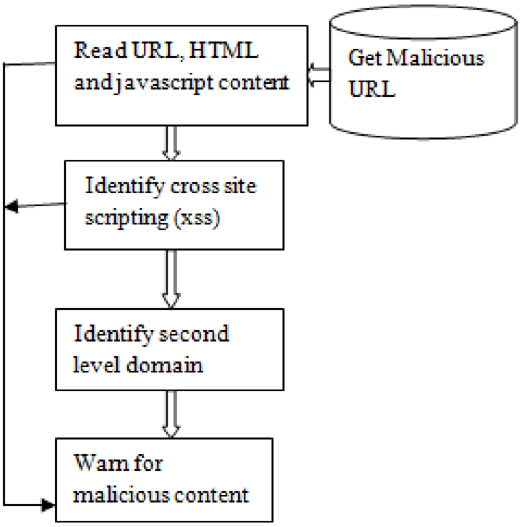

Figure.1. proposed methodology

- If application identifies that the pages are malicious then the proposed method will generate an output i.e. it detect a malicious webpages or phishing site.

V. CONCLUSION

In this way, we study the framework for detecting malicious mobile webpages in real time. Mobile webpages are significantly different than their desktop counterparts in content, functionality and layout. Therefore, existing techniques using static features of desktop webpages to detect malicious behavior for mobile specific pages. We designed and developed a fast and reliable static analysis technique that detects mobile malicious webpages and also detect phishing sites. Our application provides greater accuracy in classification, and detects a number of malicious mobile webpages in the wild that are not detected by existing techniques such as Cantina. Finally, we build a browser extension that provides real-time feedback to users. We proposed an application for mobile platforms. We identified the weaknesses of the heuristics-based anti-phishing schemes that highly rely on the HTML source code of web pages. Our application resolves this issue by using OCR, which can accurately extract text from the screenshot of the login interface so that the claimed identity of phishing attacker can be verified. We conclude that our application detects new mobile specific threats such as websites hosting and takes the first step towards identifying new security challenges in the modern mobile web.

VI. ACKNOWLEDGEMENT

I take his opportunity to express my hearty thanks to my guide Prof. V.S Gangwani for his guidance and sharing his findings for technical guidance and direction. Suggestions given by him were always helpful in this work to succeed. His leadership has been greatly valuable for me to work on this project and come with best out of it.

VII. REFERENCES

[1]. Chaitrali Amrutkar, Young Seuk Kim, and Patrick Traynor, Senior Member, IEEE “Detecting Mobile Malicious Webpages in Real Time” Chaitrali Amrutkar, Young Seuk Kim, and Patrick Traynor, Senior Member, IEEE

[2]. Charles Arthur, “Mobile internet devices ’will outnumber humans this year’.” http://www.theguardian. com/ technology/ 2013/feb/07/ mobile-internet-outnumber-people.

[3]. Chakradeo, S., Reaves, B., Traynor, P., and Enck, W., “MAST: Triage for Market-scale Mobile Malware Analysis,” Tech. Rep. GT-CS-12-01, College of Computing, Georgia Institute of Technology, 2012.

[4]. N. Provos, P. Mavrommatis, M. A. Rajab and F. Monrose, “All Your iFRAMEs Point to Us”, Proceedings of the 17th Conference on Security Symposium, SS, USENIX Association Berkeley, (2008); CA,USA.

[5]. D. Canali, M. Cova, G. Vigna, and C. Kruegel. Prophiler: a fast filter for the large-scale detection of malicious webpages. In Proceedings of the 20th International Conference on World Wide Web (WWW), 2011.

[6]. L. Bilge, E. Kirda, C. Kruegel, and M. Balduzzi. EXPOSURE: Finding malicious domains using passive DNS analysis. In Proceedings of the 18th Annual Network and Distributed System Security Symposium (NDSS), 2011.

[7]. A. P. Felt and D. Wagner. Phishing on mobile devices. In Web 2.0 Security and Privacy (W2SP), 2011.

[8]. “Cross-site Scripting (XSS) Attacks and Defense Mechanisms: classification and state-of-art” by Shashank Gupta and B.B Gupta ,14 September,2015, Springer.

[9]. Dr. Jitendra Agrawal, Dr. Shikha Agrawal, Anurag Awathe, Dr. Sanjeev Sharma. “Malicious Web Page Detection through Classification Technique: A Survey”. In Proceeding of the IJCST March 2017

[10]. C. Amrutkar, K. Singh, A.Verma, and P. Traynor. Vulnerable Me: Measuring systemic weaknesses in mobile browser security. In Proceedings of the International Conference on Information Systems Security (ICISS), 2012.

[11]. C. Amrutkar, P. Traynor, and P. C. van Oorschot. Measuring SSL indicators on mobile browsers: Extended life, or end of the road? In Proceedings of the Information Security Conference (ISC), 2012.

[12]. S. Garera, N. Provos, M. Chew, and A. D. Rubin. A framework for detection and measurement of phishing attacks. In Proceedings of the ACM workshop on recurring malcode, 2007.

[13]. A. P. Felt and D. Wagner. Phishing on mobile devices. In Web 2.0 Security and Privacy (W2SP), 2011.