International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, No. 8, Pages 1-6
Published Online June – July 2016 in IJEAST (http://www.ijeast.com)

An intelligent anti-phishing approach for Fraudulent URL detection using ML

Mrunali Naik
Department of Computer Engineering,
MCERC, Nasik, Maharashtra, India
mrunalinaik303@gmail.com

Pranit Kolhe
Department of Computer Engineering,
MCERC, Nasik, Maharashtra, India
pranitkolhe99@gmail.com

Bharti Nikale
Department of Computer Engineering,
MCERC, Nasik, Maharashtra, India
bhartnikale09@gmail.com

Saurabh Sanap
Department of Computer Engineering,
MCERC, Nasik, Maharashtra, India
saurabhsanap90@gmail.com

Abstract— Trying to gather personal information through deceptive ways is becoming more common nowadays. In order to assist the user to be aware of the access to such websites, the implemented system notifies the user through email and also pop-up, when trying to access a phishing site. This paper proposes an approach of phishing detection system to detect blacklisted URL also known as phishing websites, so that individual can be alerted while browsing or accessing a particular website. Therefore, it can be utilized for identification and authentication and become a legitimate tool to prevent an individual from getting tricked.

Keywords: Blacklisted, phishing, Agile Unified Process (AUP), alert, pop-up notification, Email notification, Machine Learning

I. INTRODUCTION

Phishing can be defined as impersonating a valid site to trick users by stealing their personal data comprising usernames, passwords, accounts numbers, national insurance numbers, etc. Phishing frauds might be the most widespread cybercrime used today. There are countless domains where phishing attack can occur like online payment sector, webmail, and financial institution, file hosting or cloud storage and many others. The webmail and online payment sector was embattled by phishing more than in any other industry sector. Phishing can be done through email phishing scams and spear phishing hence user should be aware of the consequences and should not give their 100 percent trust on common security application. Machine Learning is one of the efficient techniques to detect phishing as it removes drawback of existing approach. The objectives which is the most vital thing in proposed project is to verity the validity of the website by capturing blacklisted URLs. To notify the user on blacklisted website through pop-up while they are trying to access and to notify the user on blacklisted website through email while they are trying to access. This proposed project will allow administrator to add blacklisted URL's in order to alert user during their inquiry. The two scope of project, which is well known as user scope and system scope. User has some responsibility towards the system. The system includes a few standards and policies that requires to be obliged in order to comply the system. The user can be notified if blacklisted website is being accessed. The admin can capture the blacklisted URL's to alert user. The system involves features like capturing blacklisted website, viewing blacklisted website, displaying pop-up notification and also displaying email notification.

II. PROPOSED METHOD

In the effort of developing the proposed system, a project methodology has to be selected and defined, as to generate a

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, No. 8, Pages 1-6
Published Online June – July 2016 in IJEAST (http://www.ijeast.com)

practicable development environment and realistic schedule. Thus, this project will be done using the Agile Unified Process (AUP) Lifecycle for its abridged development period and flexible process as referred in Figure 1. A baseline of hardware and software requirements are set. This is to ensure the operation system platform is capable to handle and perform the development of the system. The software that is used to develop system is using Mircrosoft Visual Studio 2010 Ultimate. Mircrosoft Visual Studio 2010 Ultimate generates the system C# was the most appropriate language to run the program. mySQL stocks up data and to implement database in this system. mySQL builds in database in the Mircrosoft Visual Studio 2010 Ultimate. In hardware, atleast 4GB RAM is required in laptop/PC to build the system .This ensures a smooth process during the development.

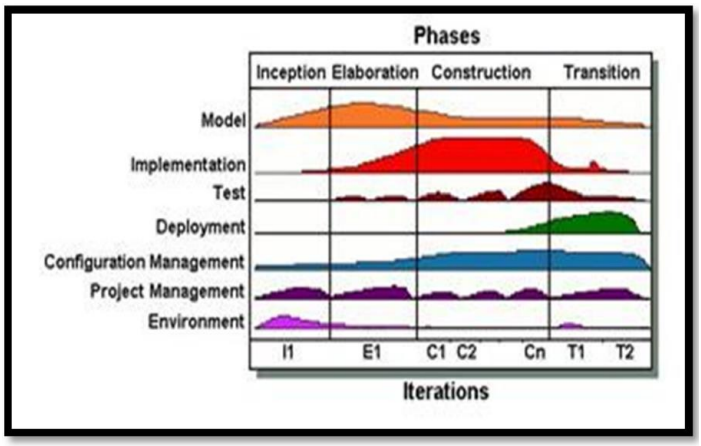

FIGURE 1: AGILE UNIFIED PROCESS MODEL

Based on Figure 2, admin can filter which URLs are blacklisted and which are not blacklisted by copy and pasting the URLs at "Site" row. Admin can classify blacklisted URLs as 1 and not blacklisted URL as 0. Admin can also edit, update and delete the sites once the URLs have been added.

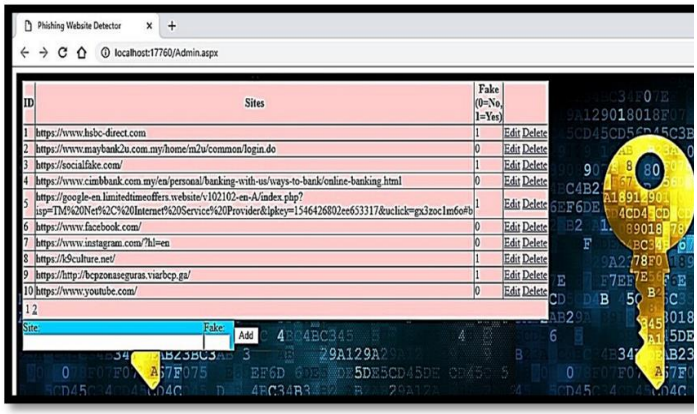

FIGURE 2: ADMIN PAGE

Based on Figure 3, this is a main page where user can classify which is blacklisted URLs and which is not blacklisted URLs by the color. Blacklisted URLs are in red colored row and non- blacklisted URLs are in white colored row. User can also inspect them by clicking on the URLs.

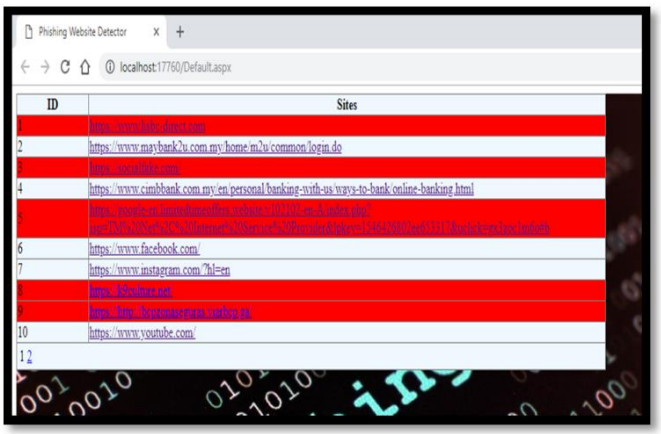

FIGURE 3: MAIN PAGE

URLs which are not blacklisted will redirect to the actual website once user clicks on it. Figure 4 shows that a pop up notification a when user clicks on the blacklisted URL. The pop-up notification is an alert box to apprise the user by questioning whether they wish to continue knowing that it may be a phishing site.

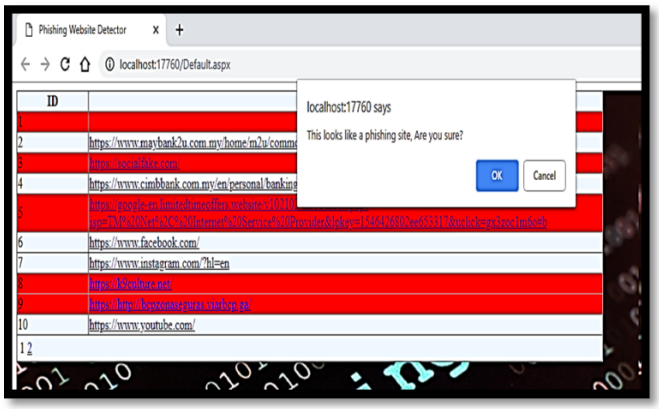

FIGURE 4: POP-UP NOTIFICATION

Based on Figure 5, a message from admin will be displayed once user clicks "OK" from the pop –up notification. This message will notifiy the user that the site accessed is confirmed a phishing site.

International Journal of Engineering Applied Sciences and Technology, 2016
Vol. 1, No. 8, Pages 1-6
Published Online June – July 2016 in IJEAST (http://www.ijeast.com)

FIGURE 5: MESSAGE FROM ADMIN

IV.CONCLUSION

After reviewing and researching for appropriate monitoring tools, proposed system has been identified and chosen to address the complexity of monitoring requirement for current situation. This software is designed to show awareness of the extensive level of its functionality, features that can be displayed in the monitoring era. The system fosters many features in comparison of other software. Its unique features such as capturing blacklisted URL's from the browser directly to verify the validity of the website, notifying user on blacklisted websites while they are trying to access through pop-up, and also notifying through email. This system will assist user to be alert when they are trying to access a blacklisted website. In conclusion, this system is designed for resources are used as intended, prevents from valuable information from leaks out, produce better control mechanism and alerts the user to keep their private information safe. Like any other programs, there are improvements which could be made into this system. Based on the capabilities which the current system processes, text message integration would a great recommendation that could be made to improve the program in the future. The future version of the application could also implement an option to directly notify the blacklisted website with a text message. The program could be made to access the list as an attachment. This text message integration function would further the usability of the application.

V. REFERENCE

[1] Matthew Dunlop, Stephen Groat, David Shelly (2010) " GoldPhish: Using Images for Content-Based Phishing Analysis"
[2] Rishikesh Mahajan (2018) "Phishing Website Detection using Machine Learning Algorithms"
[3] Purvi Pujara, M. B.Chaudhari (2018) "Phishing Website Detection using Machine Learning : A Review"
[4] David G. Dobolyi, Ahmed Abbasi (2016) "PhishMonger: A Free and Open Source Public Archive of Real-World Phishing Websites"
[5] Satish.S, Suresh Babu.K (2013) "Phishing Websites Detection Based On Web Source Code And Url In The Webpage"
[6] Purvi Pujara, M. B.Chaudhari (2018) "Phishing Website Detection using Machine Learning : A Review"
[7] Satish.S, Suresh Babu.K (2013) "Phishing Websites Detection Based On Web Source Code And Url In The Webpage"
[8] Tenzin Dakpa, Peter Augustine (2017) "Study of Phishing Attacks and Preventions"
[9] Ping Yi (2018) "Web Phishing Detection Using a Deep Learning Framework"
[10] Jalil Nourmohammadi Khiarak (2017) "What is Machine Learning"
[11] Sadia Afroz, Rachel Greenstadt (2018) "PhishZoo: An Automated Web Phishing Detection Approach Based on Profiling and Fuzzy Matching"
[12] Arun Kulkarni, Leonard L. Brown (2019) "Phishing Websites Detection using Machine Learning"
[13] Rohan Saraf , Mayur Khatri , Mona Mulchandani (2014) "Phish Tank-A Phishing Detection Tool"
[14] Sadia Afroz, Rachel Greenstadt (2017) "PhishZoo: Detecting Phishing Websites By Looking at Them"
[15] Matthew Dunlop, Stephen Groat, David Shelly (2010) " GoldPhish: Using Images for Content-Based Phishing Analysis"