International Journal on Advances in Engineering, Technology and Science

IJAETS, Volume: 1, Issue: 1, October 2015

Intelligent Web based Secure Browsing Implementation

Hardik Gohel

Assistant Professor
Atmiya Institute of Technology & Science, Rajkot, Gujarat, India
hagohel@gmail.com

Abstract: **The advent of the internet has changed the world with all significant ways since history of human life. The dominant growth of World Wide Web (WWW) and internet, we have entered in the world of intentional information. This paper is about to implementation of secure browsing as a part of web intelligence. The standard safe browsing and private browsing available in existing browsers are not useful to identify that queried URL is either malware, phishing or legitimate with no response body. This paper presents the implementation of secure browsing experimental API to allow applications to check URL against updated lists of suspected malware and phishing pages of any website.**

Keywords

Secure browsing, Web API, Malware, Phishing

I. INTRODUCTION

Browsing is the strategy for orientation to identify relevance things. In 1990, first browser invented by Tim Barners-Lee, WorldWideWeb (without space) was renamed Nexus [2]. In this two decades of web browsing there are various invention and tips and techniques specified for secure web browsing as well as transactions. Security in browsing is a feature of browsing which encrypt activity of any domain where possible and it makes it difficult for anyone else to access individual's information without permission. The application of security in web browsing is to protect data available on network and computer from contravenes of privacy. The browser may not be aware any kind of contravenes and may show user with safe connection mode. There are various contravenes includes operating system is contravenes and malware which cannot provide security to reading or modifying the browser memory of privilege mode [3]. There are so many misconceptions about security browsing. I may be thinking that I am safe but infected web pages exposed within seconds, so it is quite difficult to be updated with infected sites and awareness of risk will not work here. Let's start assessments by having some questions. As a user of web are we practicing secure web browsing? Are we avoiding risky sites? Are we specified limit time to spend online? Are we having solid policy for accessing internet? Are we using secure browser? Are we able to identify risky sites? If we can give answers of all above questions with "Yes" then we must get aware about our myths of secure web browsing [5].

II. EXISTING FICTIONS ABOUT SECURE WEB BROWSING

We are likely surfing from various perceptions about web security in which we are having some misinformation about protect ourselves from risk of browsing. As an individual we can go for elimination of internet access altogether which is not proper solution in competitive world. If we are going for strict walled perimeter which is also easily bypass by any user. [6].

2.1 Infected by malware

If I am saying that I am secure because I have not been infected or affected by malware. In such case we may not even know that we are affected. By stealing personal information and password many malware webs attacks. They may use my machine for spam distribution, malware or improper content without my knowledge. Let's say, I am using Quick Heal Total Security in which I have installed web appliance at its network gateway and immediately bannered more than 25 machines on its network for mistrustful behaviour – calling home to malware network [7].

2.2 Improper content surfing by users

Web filtering is the best way to prevent unnecessary access of internet. If we have not applied web filtering we really do not have idea about user's activities. There is one fact that more than 40% of corporate internet usage is improper and unobserved as well. The average of such activities is 1-2 hours per day per user. The inappropriate content includes gambling, pornography, social networking, travel planning and many more. [8]

2.3 Controlling of web usage

There are various web proxies available in the world of internet in which avoid web controlling and allow visit the site as per users requirement. Presently, more than hundreds of proxies are available, verification of which you can get by searching from Google in which you may search by "free proxy server" and 34 millions ways are there. [9]

2.4 Perceptions about dangerous web

If any individual is thinking that only gambling or porn sites are dangerous then according to softpedia [10] news there is greater than 83% hosting of malware sites are hosting. The website that you trust would be affects to your system and content of system. Malware sites might be popular with high-traffic venues and distribute malware to credulous browsers [13].

2.5 Experienced vs. Inexperienced users

Malware can be happen without any user's action or site visit. It does not make any sense with computer experience or expertise of individual. If any individual is using internet and browsing sites there are multiple chances of users at risk.

III. IMPLEMENTATION OF SECURE BROWSING EXPERIMENTAL API

After having depth study of existing fictions of browsing of web we are moving to secure browsing experimental API which allows any application or user to check against continuously updated lists of alleged malware as well as phishing WebPages[11][12]. It is designed for providing minor interface for applications which queried as URLs. Generally, Google is maintaining database for malware and phishing sites and same database can be applied for any organization or corporate body. By using lookup application program interface, user has to query URL(s) by HTTP with GET request as well as POST request and will get position of URL from server itself [15].

3.1 Request of API Key

Initially, it requires generating key for authentication of user. For key generation you must register yourself into server. The purpose of API key is to make sure request has been sent to server without forged. Now the question is that whether the key is made of cryptography or there is logic behind of it? A little research acquiesces following for generating API key:

- ✓ Keys are long as 86 characters always.
- ✓ ABQIAAAA are first 8 characters always.
- ✓ 9 to 30 characters relate to account information
- ✓ 31 to 58 characters of URL
- ✓ 59 to 86 characters are for cryptographic hash function for validate a key which is hidden.

Key1: - ABQIAAAAnSuFwNN9J1XIntZ-MQNgzRSqXurX4ktkq4z_JrMDm3dtRxGSqg

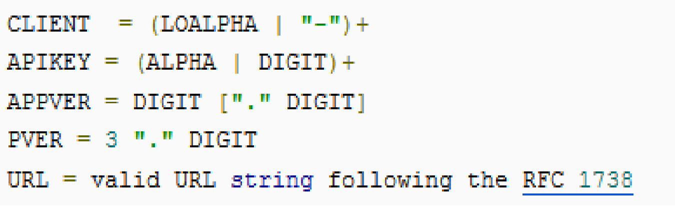
Key2:- ABQIAAAA29GuZNx6ZP4PAgfq_MaywRSttK5qEEQH70pg-5slDPB-tvGGvA

IV. OPERATIONS TO PERFORM API

The choice of operation is mandatory as well as important task. There are two operations, first one is GET method and another is PUT method.

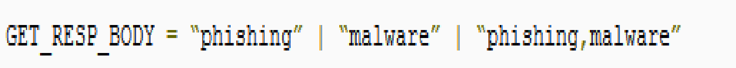
4.1 *HTTP Get Request*

In GET request client can go for one URL at the time. The Routing Backus-Naur Form (RBNF) of GET request is given bellow. It is syntax to form encoding rules in multiple routing protocol specifications.



```
CLIENT  = (LOALPHA | "-")+
APIKEY = (ALPHA | DIGIT)+
APPVER = DIGIT ["." DIGIT]
PVER = 3 "." DIGIT
URL = valid URL string following the RFC 1738
```

The response codes of GET request is given bellow which is very common for HTTP response. If URL is matches either Phishing, malware or both the response code will be 200.

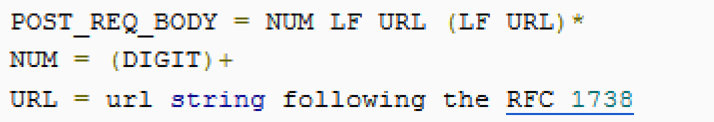


```
GET_RESP_BODY = "phishing" | "malware" | "phishing,malware"
```

Where, Phishing is the URL listed in Phishing list, malware is the URL listed in malware list and third one is matching both.

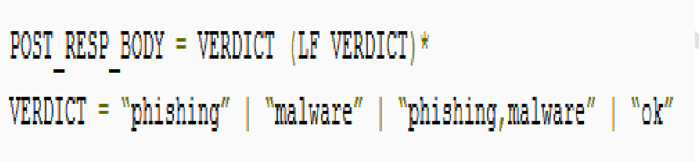
4.2 HTTP Post Request

In POST request client can go for set of URLs, nearby 500, at the time. The Routing Backus-Naur Form (RBNF) of POST request is given bellow. It is syntax to form encoding rules in multiple routing protocol specifications.



```
POST_REQ_BODY = NUM LF URL (LF URL)*
NUM = (DIGIT)+
URL = url string following the RFC 1738
```

In above code, various lines separated by LF. NUM is indicating the number of URLs in body. Other lines of URLs are several lines given in body. It is necessary to put URL as in valid form.



```
POST_RESP_BODY = VERDICT (LF VERDICT)*
VERDICT = "phishing" | "malware" | "phishing,malware" | "ok"
```

The response body same as to GET response body but server will return exact result as requested in original order.

V. Limitation of Secure Browsing API

- ✓ Internal structure of API is requiring to be known by users which include storage of hashed URLs in list of phishing and malware.
- ✓ Users are also requires to update their cache of hashed URLs.
- ✓ Users need to compile downloaded hashed value of URLs.
- ✓ Cannibalize of URLs by user themselves.
- ✓ API URLs are not hashed by this method which is vulnerable.
- ✓ There is no guarantee on Response time of requested API URLs.
- ✓ It is not efficient in the terms of bandwidth usage.

VI. Conclusion

The above implementation study is really effective for secure browsing in the terms of identifying Phishing and malware. If individual is not much concerned about privacy as well as latency of request in the term of response time, the above given application program interface is very much useful since it is simpler for implementation but lengthy to understand usage of it.

REFERENCES

[1] Y.Y.Yao, Ning Zhong, “Web Intelligence - Research Challenges and Trends in the New Information Age”, 1 ed. , Canada: Springer, 2011.

[2] Tim Berners-Lee, : “WorldWideWeb, the first Web client”, W3.org,

[3] Goodin, Dan, : “MySQL.com breach leaves visitors exposed to malware”. Retrieved 26 September 2011

[4] Smith, Dave, : “The Yontoo Trojan: New Mac OS X Malware Infects Google Chrome, Firefox And Safari Browsers Via Adware. IBT Media Inc.” Retrieved March 21, 2013

[5] Gaurav Aggarwal, Elie Bursztein, Collin Jackson, Dan Boneh, “An Analysis of Private Browsing Modes in Modern Browsers”, Stanford University ed. , 2012.

[6] “The 10 myths of safe web browsing”, 1 ed., http://sophos.williams.edu/media/myths-for-safe-web-browsing.pdf: Sophon, May 2010.

[7] Lincoln Spector, “When malware strikes: How to clean an infected PC”, http://sophos.williams.edu/media/myths-for-safe-web-browsing.pdf: PCWorlds, May 30, 2013 3:04 AM.

[8] Robert P. Lipschutz , “Web Content Filtering: Don't Go There”, http://www.pcmag.com/article2/0,2817,1532605,00.asp : PCMag, Accessed on May 1, 2014.

[9] Trefor Davies, “How to bypass the Virgin Media web filter to access Pirate Bay”, trefor.net, 6 May, 2012.

[10] Eduard Kovacs, “Canada Sees 83% Increase in Malware C&C Hosting, Study Finds”, http://news.softpedia.com/news/Canada-Sees-83-Increase-in-Malware-C-C-Hosting-Study-Finds-360787.shtml: Softpedia, June 13th, 2013, 17:11 GMT.

[11] Gohel, Hardik. "Looking Back at the Evolution of the Internet." CSI Communications - Knowledge Digest for IT Community 38.6 (2014): 23-26.

[12] Hardik, Gohel. "Design and Development of Combined Algorithm computing Technique to enhance Web Security." International Journal of Innovative and Emerging Research in Engineering (IJIERE) 2.1 (2015): 76-79.

[13] Gohel, Hardik. "Deliberation of Specialized Model of Knowledge Management Approach with Multi Agent System." National Conference on Emerging Trends in Information & Communication Technology. MEFGI, Rajkot, 2013

[14] Hardik, Gohel. "Design of Intelligent web based Social Media for Data Personalization." International Journal of Innovative and Emerging Research in Engineering (IJIERE) 2.1 (2015): 42-45.

[15] Gohel, Hardik, and Priyanka Sharma. "Study of Quantum Computing with Significance of Machine Learning." CSI Communications - Knowledge Digest for IT Community 38.11 (2015): 21-23.