# Security and Authentication of Web Browsers using FCM

Harish Singh Baghel
Computer Science Department
M.G.C.G.V
Chitrakoot, Satna, India

Bharat Mishra, PhD
Physical Science
M.G.C.G.V
Chitrakoot, Satna, India

## ABSTRACT
Web security is an important field area where security and storage is important during the flow of data from the client to server. Although there are various techniques implemented for the security of web data especially at the client side in web browsers. Here in this paper an efficient technique for the security of web data is implemented using combinatorial method of clustering and classification. The web log data collected at the server side is used for the training and testing of classification of various attacks and anomalies in the packets send through internet. The Fuzzy C-means clustering is first applied on the web log data to divide into a number of clusters of similar groups and then these clustered data are passed to the classification algorithm to generate decision tree to get the final classification of attacks and anomalies.

## Keywords
WWW, WML, Clustering, Classification, Fuzzy logics.

## 1. INTRODUCTION
Accessing of information is very easy in modern era via Internet. The World Wide Web provides its services through internet. WWW has a vast storage called as databases of information ready for users at the very moment. The user can have direct access to this information with the help of internet. Besides information it also provides various services through web application that is in the form of banking, shopping, music, video etc. The technique behind these services is continuously growing providing ease for the users. All they require is a medium to exploit or say use these services. That is in form of Web Browsers. A web browser is a web client which runs on users/clients computer requesting web server to operate requests generated by clientfor obtaining information or resources. The web server in return locates and sends the information to the web browser which thereby displays the results. It is almost similar to client-server model. Therefore web browser is a necessary requirement for a connection establishment and exchange of information between the user and World Wide Web.

Besides these services web browsers also provide various other small services through extensions, plug-ins, add-ons etc. These are the third party applications providing services like bookmarks, java script run, flash player etc. Due to such vast use of browser there are always security threats against them in multiple forms. Browser extensions introduce serious security vulnerabilities into the browsers or the websites to which extensions interaction takes place [1]. Extensions are able to read and manipulate content from websites, make network requests and can even access browser user data such as bookmarks and location history etc. This opens the attacker privileges of collecting users' private information and authentication details. Extensions are basically written in JavaScript and HTML in which JavaScript provides multiple methods for converting strings to code. Improper use of these methods may introduce code injection threats compromising the extension. Data can even execute if it is written into a page as HTML instead of as text [2]. Browser was designed and evolved from a simple client application that display static data into a complex networked operating system, to improve and manage multiple type of user's on-line experience. To fulfill these need browser extensions were introduced expanding the functionality of browsers by interacting with browser level events and so data. Some of the extensions are simple and make a little change in the appearance of web page or in the browser whereas other extensions may provide more sophisticated functionality.

There are multiple browsers present in the market fulfilling needs of users. Availability of multiple options has made the choice of web browser difficult and confusing. Mainstream browsers are getting tough competition from the upcoming browser alternatives with more unique features of its own. Web browsers also provide accessibility features for users with disabilities like blindness and low vision, hearing loss and motor impairments. Web browser is important as an application for conducting billions of dollars of Internet enabled commerce each year [3]. When a user visits a web page the contents of that page get stored in the browser's cache making it not to be re-requested and re-downloaded. This setting of the browser can be changed according to the user's requirement. Efficiently using the browser cache helps to improve end user response time and reduces utilization of bandwidth. If an item or web page is considered cacheable the browser will retrieve or load the item from cache on repeat or next visit to that page thereby increasing the efficiency behind the mechanism of web browser [4].

Web browsers in today's world are not subjected to computer systems only they run on multiple types of hardware like cell phones and tablet, PCs, desktop computers, laptops etc.. Hence a proper design becomes very important for the browser to be compatible with each and every device. Reference architecture for web browsers helps implementers to understand features and trade-offs while designing new systems and thereby assisting maintainers in understanding the code. Reference architecture serves as a template for creation of new systems by identifying areal features in which reuse can take place at the design level and also at the implementation level. A web browser retrieves documents from web servers and displays them on screen that can be the browser window itself or by passing the document to an external application. It also allows particular resources that can be requested explicitly by URL's or implicitly through attached hyperlinks. The conceptual architecture gives the

idea of developers thinking about the system and contains only relationships between subsystems meaningful to developers itself whereas the concrete architecture of a system is a high-level description of major subsystems that are implemented [5].Reference architecture for a browser captures fundamental subsystems common to systems of the browser and the relationships between the browsers subsystems. Evolution in telecommunication technology has led the reference architecture developers to work upon browsers over handheld devices. Today people access the Web with Internet connections using HTML and WML (Wireless Markup Language) browsers and present portable devices (e.g. Tablets, Cell-phones etc.) offer features like large memories, friendly graphical user interfaces, communication interfaces and possibility to install Internet browsers making them suitable for hosting applications that are normally performed by standard personal computers. They are limited to the lesser input/output capabilities like lack of alphanumeric keyboard and small displays etc. The development of modern browsers with voice input/output capabilities, graphic pointing, touch screens, small numeric keyboards etc. satisfies various user requirements. Modern web browsers built up on reference architectures are highly sophisticated. Browsers are software suites that can execute videoconferencing, play audio/video end even lets you allow to create and publish HTML pages. Browsers have blurred the line between computer and the Internet making computer and the Internet function as a single computer system [6, 7]. Browsers with greater usage are equally vulnerable to attacks and threats. But continuous increase in technology, security features reference architecture etc. the mechanism of web browsers will increase rapidly making them more efficient for users.

## 2. LITERATURE REVIEW

Dr. B. Mishra et.al. [8] Suggested various protocols working in different layers can enhance the performance of the web browsers. They gave the idea that modern web browsers contain more features as compared with the existing one. They are designed in such a way that the feature comparison is lot higher. Modern web browsers are implemented with multiple protocols making the browser efficient for transferring of information through browsers. With increasing the efficiency of protocols browser efficiency can directly be increased. The presented a comparative analysis of multiple protocols used in web browsers which are used for establishing different applications. The study shows that protocols are can provide more effective way of communication, security and the transfer of data in efficient manner in web browsers. As a future work they suggested that browser efficiency can be increased by sophisticated implementation of more protocols [8].

G. Wagner et.al.[9] gave the idea that modern web browsers do not sequentially navigate static web sites they manage multiple simultaneous tabs that display dynamic web content which may be running client-side JavaScript code. This parallelism was not fully taken by JavaScript virtual machine architecture. They proposed abstraction for multiple disjoint JavaScript heaps and named them as compartments. The methodology they used as by clustering objects into separate compartments through document origin where objects can only reference each other through wrappers and objects within same compartment can reference each other directly. With the help of this methodology garbage collection pause time can be reduced by permitting collection of compartments and cross origin object access policy can be enforced through wrappers. They implemented per-compartment GC and added a layer of abstraction to JavaScript heap and separated JavaScript data which was based on its origin [10].

C. Reis et.al. [10] Provided that websites today contains client side code which are similar to programs than being the documents. For this purpose the browsers should provide high performance, robust and responsive platform by identifying the boundaries of program and isolate them. They suggested three techniques in their research as giving abstractions of web programs and program instances and how these abstractions clarify browser components interaction identifying the program boundaries. They identified backward compatibility tradeoffs which defines how web content can be divided into simple programs without disturbing existing websites. They also suggested mechanism to improve performance, fault tolerance, resource management by isolating web program instances from each other through a multi-process browser architecture [10].

M. Ter Louw et.al. [11] they examined security issues of functionality extension (plug-ins) mechanisms that web browsers support. The extensions/plug-ins have unrestrained access in web browsers and thus are vulnerable to malware. Due to the lack of security mechanisms for browser extensions they implemented a malware application for Firefox named as browser Spy. The application they implemented was capable of observing activity of browser, taking control of it and cannot be detected. The results obtained were analyzed for performing security mechanisms by code integrity checking technique to control the extension installation and its process of loading. They described their implementation mechanism as a drop-in solution which can employ JavaScript and a faster in-browser solution that uses the browser's native cryptography implementation. They discussed techniques for runtime monitoring of plug-in behavior for defending threats generated by installed extensions. They ensured that browser only allow plug-ins installed by the user to be loaded and can even detect unauthorized changes made to installed plug-ins resulting in sealing of outside installation of malicious extensions [11].

A. Barth et.al. [12] they evaluated that websites rely on browser's security policy by using third party content in frames for security from harmful content. Applications are dependent on browser for isolation of frames from various security origins and provide an efficient and secure inter-frame communication. For this purpose they analyzed existing frame navigation policy and suggested a more sufficient and strict policy deployed in the open-source browsers because the browsers allow manipulation of frames through navigation. The policy also restricts communication between cooperating frames. For inter-frame communication through fragment identifier messaging they suggested fragment identifier messaging which provides confidentiality without authentication and post Message which provides authentication. The policy prevents attacks by allowing one frame to navigate another only if the frame is able to draw over the other frame's region of the screen. They presented that post Message communication suffer from vulnerability of converse security that can result in breach over confidentiality of channel which can be overcome by extending post Message to post Message API [12].

J. Kahng et.al. [13]they suggested that with constant increase in use of Internet and services provided by it the online security is gaining importance. Interaction through World Wide Web with internet for various services is fulfilled through web browsers and software's capable of accessing

web. However some security features are provided by WWW and web browsers also but they are required to be meaningful. Web browsers uses security policies to determine when to display security warnings due to lack of appropriateness the users sometimes gets interrupted and starts ignoring the warnings and messages when they can be useful to them. They proposed mechanism to regain user's attention towards such messages and warnings not by just changing the format or layout of the messages but by combining these designs with strict policies and making them aware of security settings, its importance and their computational environment [13].

M. J. Rees et.al. [14] stated that web browsers can overcome the use of web applications in accordance with the current trend. They proposed that UI designs, evaluation skills and implementation can be focused for such situation. The software architecture of the web requires HCI demands which will require training for the users designers in the architecture with the help of new web services which will bring new interface designs for the users. They demonstrated standard of future user interface design over the evolution of browser user interface. They proposed a browser independent application to resolve the international support constraints that can arise. The application is able to support DHTML DOM, XML, XMLDOM, XSLT, SOAP and UDDI. The user interface controls provided will be able to serve at both ends i.e. design time and run time with the help of distributed XML web services [14].

# 3. PROPOSED METHODOLOGY

**Fuzzy C-Means Clustering**

The Fuzzy C-means clustering uses the concept of categorizing the data in two or more clusters that belongs to the same group as generated in Fuzzy Logics. The objective function used in the Fuzzy Clustering provides the minimization of the function so that the clustering can be done efficiently. The Objective function can be given by:

$$O_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m} \|x_i - c_j\|2$$

Where m lies between $1 \leq m < \infty$

**Table 1.Annotations used in the Algorithm**

| | |
|---|---|
| $u_{ij}$ | It is defined as the degree of membership of $x_i$, in the present cluster j. |
| $x_i$ | It is denoted as the ith of the d-dimensional data to be measures in the dataset. |
| $c_j$ | It is defined as the d- dimension cluster center present in the dataset. |

The Algorithm consist of the following few steps along with the minimization of the objective function.

[1] First of all the Dataset whose clustering is done is initialized with the objective function as $S = [u_{ij}]$, it is a matrix which contains the values to be clustered of m rows and n columns.

[2] After each p-step compute the centroid of the matrix or vector matrix of the dataset denoted as C,

$$C^{(k)} = [c_j]$$

, and contains the vector Matrix S,

[3] Compute objective function and minimize the effect of the objective function using,

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} . xi}{\sum_{i=1}^{N} u_{ij}^{m}}$$

[4] Update each value of S[k] with next S [k+1]

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}}$$

[5] To check the computer value of S[k] and S[k+1], means ,

$$\|S^{k+1} - S^{k}\| < \in$$

Stop the process otherwise go to Step -2.

# 4. RESULT ANALYSIS

The table shown below is the analysis of correctly classified instances on the basis of number of cycles.

**Table 2. Analysis of Correctly classified instances**

| No. of Cycles | Correctly classified Attack (%) |
|---|---|
| 10 | 89 |
| 20 | 89.34 |
| 30 | 93.67 |
| 40 | 94.672 |
| 50 | 94.87 |
| 60 | 95.23 |
| 70 | 96.23 |
| 80 | 96.44 |
| 90 | 96.87 |
| 100 | 97.12 |

The table shown below is the analysis of precision, recall and F-score on the basis of number of cycles.

**Table 3. Analysis on Various Parameters**

| No. of Cycles | Precision | Recall | F-Score |
|---|---|---|---|
| 10 | 78.43 | 82.4 | 80.366 |
| 20 | 79.74 | 84.29 | 81.95189 |
| 30 | 82.47 | 85.39 | 83.9046 |
| 40 | 84.28 | 85.81 | 85.03812 |
| 50 | 86.48 | 86.23 | 86.35482 |
| 60 | 87.64 | 86.67 | 87.1523 |
| 70 | 88.81 | 86.87 | 87.82929 |
| 80 | 89.93 | 87.1 | 88.49238 |
| 90 | 91.38 | 88.92 | 90.13322 |
| 100 | 93.47 | 89.49 | 91.43671 |

The figure shown below is the analysis of precision, recall and F-score on the basis of number of cycles.

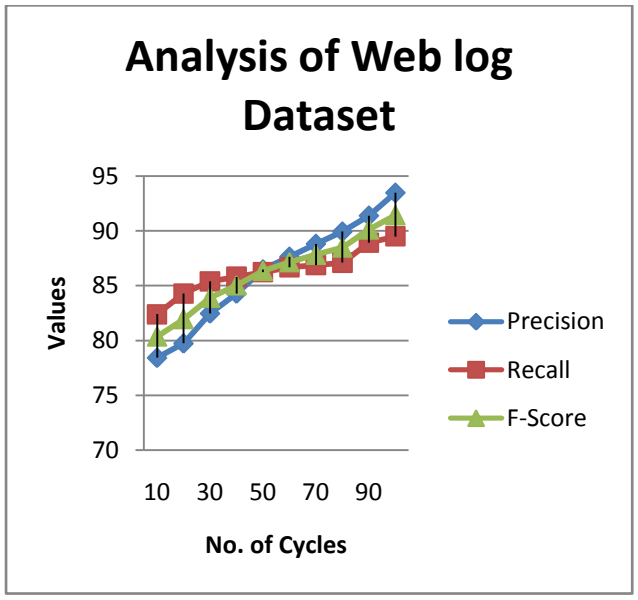


**Figure 4. Analysis of Web log Data**

## 5. CONCLUSION

During the transmission of data from the client to the server various security features are to be implemented such that the chances of eavesdropping are being reduced. Here in this paper the new concept of one time private key is used in the web browsers such that the browsers provides a special security and prevents from various security attacks. The technique implemented here in the web browsers shows the performance in terms of prevention of attacks as shown. The techniques also doesn't require additional storage cost and the chances of identity disclosure and replay attacks is removed since it only generates one key and destroy it every time.

## 6. REFERENCES

[1] Nicholas Carlini, Adrienne Porter Felt and David Wagner "An Evaluation of the Google Chrome Extension Security Architecture", 2012

[2] Sruthi Bandhakavi, Samuel T. King, P. Madhusudan and Marianne Winslett "VEX: Vetting Browser Extensions For Security Vulnerabilities", 2010

[3] Amogh Kulkarni, Jaison Salu John, Yohan John Thampi, Shravan Udaykumar, Gaurav Prasad and Vrinda Halarnkar "Web Browsers", 2009.

[4] Dawn Parzych "Caching Behavior of Web Browsers", Acceleration System Architect (ASA), F5 Networks, Nov- 2007.

[5] Alan Grosskurth and Michael W. Godfrey "Architecture and evolution of the modern web browser", Elsevier Science, 2006

[6] Cristiana Armaroli, Ivano Azzini, Lorenza Ferrari, Toni Giorgino, Luca Nardelli, Marco Orlandi and Carla Rognoni "An Architecture for a Multi-Modal Web Browser", 2001

[7] "How Web Browsers Work" Chapter 18, PART 4 HOW THE WORLD WIDE WEB WORKS.

[8] Dr. Bharat Mishra, Harish Singh Baghel,Manoj Patil and Pramod Singh "Study & Analysis of various Protocols in popular Web Browsers", International Journal of Advancements in Research & Technology, August-2012.

[9] Gregor Wagner, Andreas Gal, Christian Wimmer, Brendan Eich and Michael Franz "Compartmental Memory Management in a Modern Web Browser", ACM 978-1-4503-0263-0/11/06, 2011

[10] Charles Reis and Steven D. Gribble "Isolating Web Programs in Modern Browser Architectures",ACM 978-1-60558-482-9/09/04, 2009.

[11] Mike Ter Louw, Jin Soon Lim and V. N. Venkatakrishnan "Enhancing web browser security against malware extensions", Springer, 2008.

[12] Adam Barth, Collin Jackson and John C. Mitchell "Securing Frame Communication in Browsers", 2008

[13] Jennifer Kahng "Evaluating Web Browser Security Interfaces for a More Meaningful Design", 2001

[14] Michael J. Rees "Evolving the Browser towards a Standard User Interface Architecture", AUIC2002