International Journal of Advanced Information Science and Technology (IJAIST) ISSN: 2319:2682
Vol.4,No.7, July 2015 DOI:10.15693/ijaist/2015.v4i7.102-106

Security analysis of Web Log Files against IP Spoofing and Brute Force Attack Using Genetic Algorithm and Neural Network

Neha
M.E. Student
NITTTR
Chandigarh, India

Maitreyee Dutta
Professor and Head
NITTTR
Chandigarh, India

***Abstract*— Nowadays it is very essential to sustain a high level security to ensure protected and trusted communication of information between various organizations. But secured data communication over the web and any other network is always under threat of intrusions and misuses. The main aim of this paper is to progress the security of web log files against a couple of attack named as IP Spoofing and Brute force using genetic algorithm and neural network. The timely detection of these particular types of attacks, although quite challenging, yet it is very much obligatory to safeguard the network assets as well as the end users. This research paper focused on two types of attacks named as Brute Force attack and IP Spoofing with the prevention technique like genetic algorithm. To avoid these threats and to progress the security of web log files, this research gives a healthier result. In this proposed work, we have analyzed security risk over web log files against brute force attacks and IP spoofing and also presented the proposed framework. The outcome evaluation has been performed by calculating parameters like FAR, FRR and accuracy with the help of neural classifier. The research would also compare the accuracy of IP spoofing and brute force attack. The whole simulation has taken place in the MATLAB 7.10 environment.**

***Index Terms* — GDI, PTL, Reversible logic, Microwind, Low power VLSI, CMOS design.**

I. INTRODUCTION

The World Wide Web has encountered wonderful development in organizations, individuals, governments and it found that web applications can create efficiencies as well as a logical solution to the complications of communicating and additionally directing commerce. Web applications similar to e-business, internet banking, enterprise coordinated effort and supply chain management suites, presumes that at the least 92% of Web applications are helpless against some type of attacks. So from this data we have seen that the private data of individuals must be kept secret and confidential and Integrity of them must be given by developer of web application, yet this is unrealistic, there is no any certification for safeguarding the fundamental databases from current attacks. Attacks take place when un-trusted information, for example, command, a query, or argument, is sent to interpreter. There are numerous kinds of online attacks which influence the system badly and infect the system in such a mode that the server finds itself to get better. In the same measures, attacks like brute force or IP spoofing has made its place in the swarm of online theft and attacks. The brute force attack is an attack in which the combination of passwords is throwing to the login system to crack the password. It has been seen time and again that the combination of password cracks the database. To thwart the system for such attacks optimization algorithms are designed to guarantee that the system remains safe and sound. Additionally, there is another attack named as IP spoofing attack which leads to send request to the server in an arbitrary manner, before the server deals with one IP another request hits the system and makes the server response slow as well as time-consuming. This scenario can be frequently seen in the academia sites where results day becomes a hectic task for the server to respond to the entire request at the same time and it leads to slowing down the server.

Specifically, this research paper focused on two types of attacks named as a brute force attack and IP spoofing with the prevention technique like Genetic Algorithm (GA). To avoid these threats and to progress the security of web log files, this research gives a healthier result. In this proposed work, we have analyzed security risk over web log files against brute force attacks and IP spoofing and also presented proposed framework. The base of our system is composed of intrusion detection systems (IDS) which utilize log file dataset to discover intrusion. The IDS scans all the log files being transmitted from the routers for malicious content and known virus signatures with the help of genetic algorithm (GA) and also utilizing neural classifier. The evaluation of our system, utilizing the log file testing dataset, shows a better ratio of detecting attacks and a low false positive ratio. It likewise supports easy alteration, usability and also it can be measured without problems.

II. PROPOSED WORK

This Section is portioned into two parts:

a. Brute force Attack implementation, encryption, prevention and classification.

b. IP spoofing Attack implementation, detection and classification.

2.1 Brute force attack

In this part, GUI system is designed for this proposed framework, in which log file is uploaded.

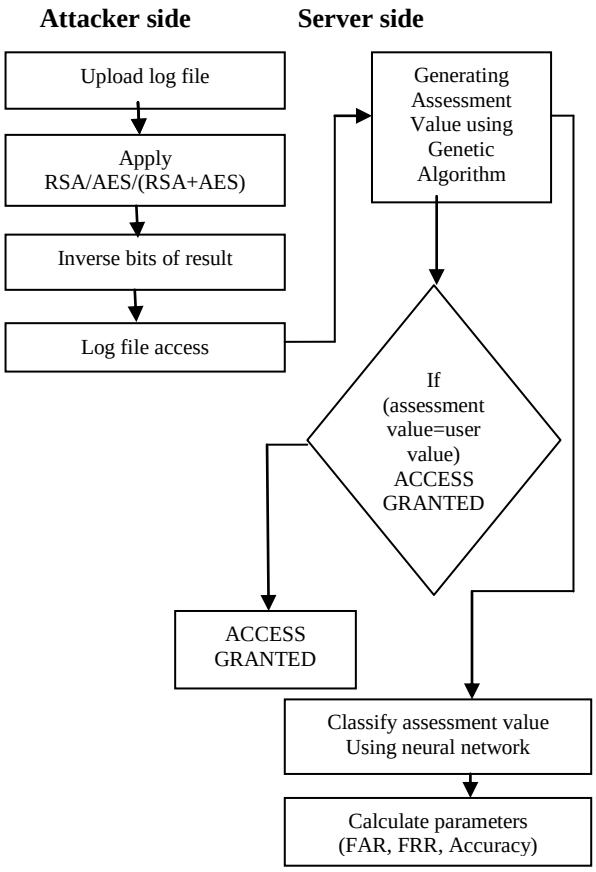

Figure 1. Flowchart of Brute force Attack implementation, encryption, prevention and classification.

Then RSA/AES/RSA+AES encryption algorithms are used to reduce the security risk over the log files. Then inverse the bits of the results and later access log file and send it to the server side for generating assessment value utilizing genetic algorithm. After generating assessment value first checks the assessment value with user value and if it is matched as well as equal with user value then Access is granted otherwise Access is denied. Hereafter classify the assessment value using neural network. The output of this attack is given on the basis of three parameters as FAR, FRR and accuracy which also tells about the system that how much efficient results it proved.

2.2 IP Spoofing

In this part, GUI system is designed, where server is firstly initialized and uploads the log file.

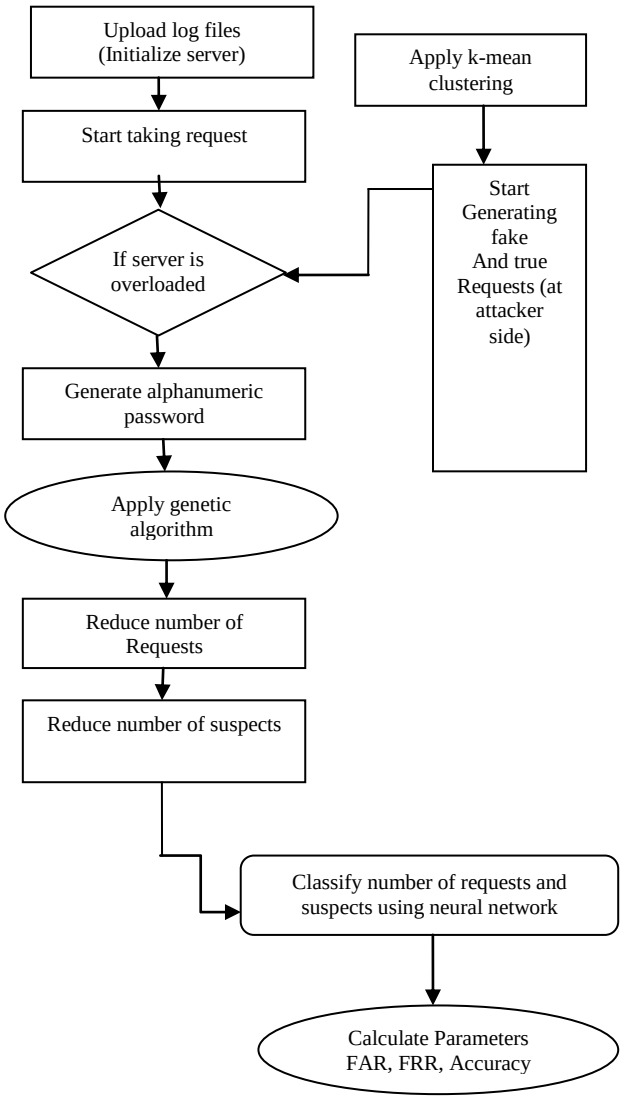

Figure 2. Flowchart of IP spoofing Attack implementation, detection and classification

After initialization step, apply k–mean clustering for the partitions of number of requests into fake and true one. Then start taking requests and send to server. Then from attacker side, start generating fake requests which is sent to the server. If the server is overloaded with requests, then generate an alphanumeric password. Then, Genetic Algorithm is used to reduce the requests as well as a number of suspects. Eventually apply neural network for reduction to attain outputs. The output of this attack is given on the basis of three parameters as FAR, FRR and accuracy which also tells about the system that how much efficient results it proved.

III COMPUTATION PARAMETER

In the proposed framework of brute force attack and IP Spoofing, three parameters are calculated named as FAR, FRR and accuracy. It also tells about the system that how much efficient results it provides.

a. False Acceptance rate: The false acceptance rate or FAR, is also known as Type –II error. It is defined as the probability that the security system will mistakenly accept an access attempt by an illegal user. Mathematically, a system's FAR typically is calculated as the ratio of the number of false acceptances to the number of identification attempts.

FAR = Number of false acceptances / Number of Identification attempts (i)

b. False Rejection rate: The false rejection rate or FRR, is also known as Type –I error. It defined as the probability that the security system will mistakenly reject an access attempt by an authoritative user. Mathematically, a system FRR typically is calculated as the ratio of the number of false rejected to the number of identification attempts.

FRR = Number of false rejected / Number of Identification attempts (ii)

c. Accuracy: Accuracy evaluates overall correctly classified attempts.

Accuracy = TP+TN / (TP +TN+FA+FN) (iii)
In terms of percentage, it can be calculated as:

Accuracy = (100-(FAR+FRR)) % (iv)

IV EXPERIMENTS

4.1 Simulation work model

The whole implementation has been taken place in MATLAB 7.10 environment. The figure 3 shows the main graphical user panel of the proposed system (Brute force) having user interface controls in which a variety of buttons for various functions are designed such as choose a log data file, enter password text, approval of password, brute force encryption-AES, brute force encryption- RSA, apply hybrid algorithm (RSA+AES), brute force next phase (inverse the bits of result), generate combinations, prevention via like Genetic Algorithm and classify using neural network.

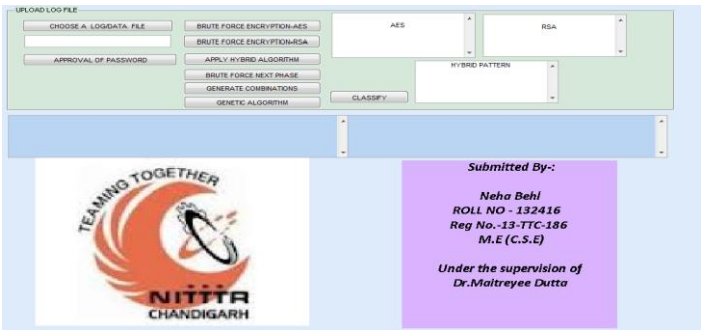

Figure 3. Main GUI of brute force attack using encryption algorithms.

The following steps present the different stages that need to minimize the security risk over log files against brute force attack with the help of popular encryption algorithm and optimization technique like Genetic Algorithm.

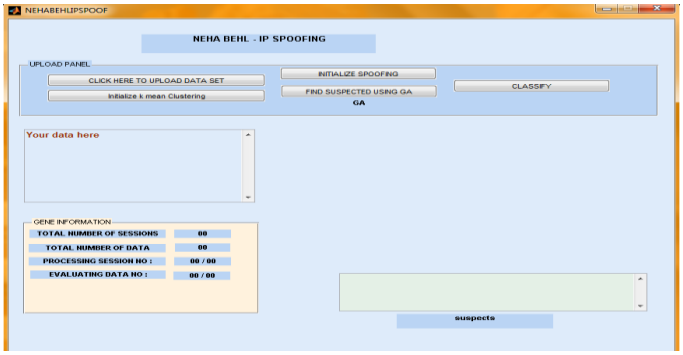

Figure 4. Main GUI of IP Spoofing

The figure 4 shows the second GUI of applying IP spoofing. In the above figure, there is shown various buttons like, upload data set, initialize k-mean clustering, initialize Spoofing, find suspected using Genetic Algorithm and classify using neural network. Then upload the dataset as input.

V RESULT AND COMPARISON

From the above experiment, we were able to perform security analysis of web log files to minimize the security risk of log files against IP Spoofing and brute force attack. As mentioned before, the scope of our experiment was focused on two types of attacks. Firstly, a framework of brute force attack is proposed with some popular encryption algorithm such as RSA, AES and hybridization of both (RSA+AES). In addition, prevention technique like genetic algorithm is applied for the better improvement of the result.
The security risk is also examined with the help of GA and these encryption algorithms such as GA-RSA, GA-AES, and GA-(RSA+ASES). The idea with GA is to use this power of evolution to solve optimization problems and when combined with these encryption algorithms gives the best security against these types of assaults.

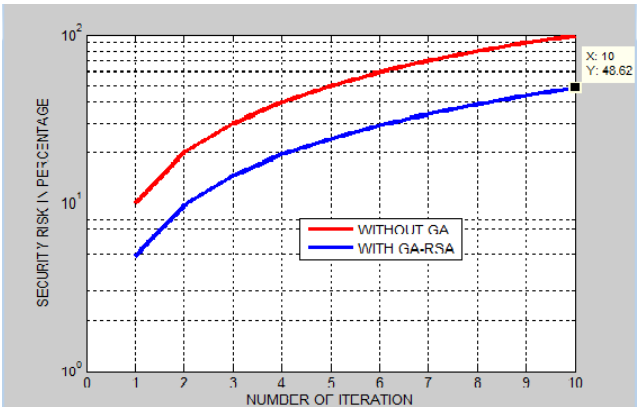

Figure 5. Graph represents value of security risk (%age) with GA –RSA

In above figure 5, it shows the graph without Genetic Algorithm and with Genetic Algorithm in combination with RSA. In this, firstly it optimizes and reduces the data by utilizing genetic algorithm and then it uses RSA for encryption purpose to save and protect the log file from any kind of security risk.

The idea with GA is to use the power of evolution to solve optimization problems and when combined with AES it saves and protects the log file from any kind of security risks.

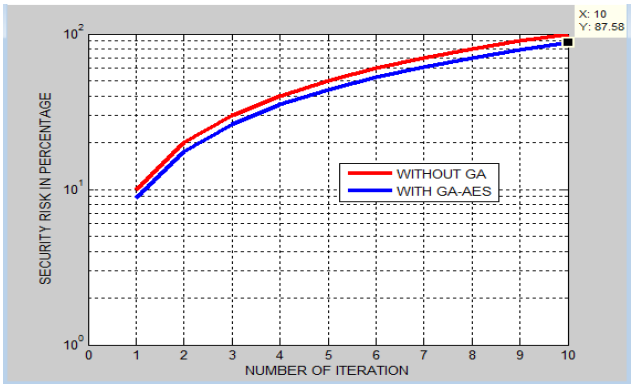

Figure 6. Graph represents the value of security risk (%age) with GA –AES

The figure 6 shows that the graph is created between security risk in percentage and number of iterations.

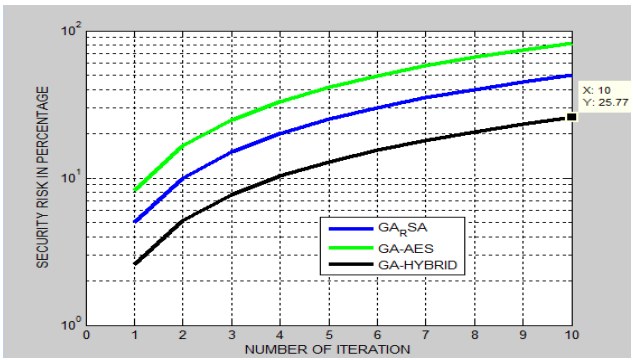

Figure 7. Represent the GA with RSA, GA with AES and GA with (AES+ RSA)

In above figure 7, it shows the graph between security risk in percentage and number of iterations and represents the GA with RSA, GA with AES and GA with (AES+RSA) algorithms. Various combinations of GA are utilized here. RSA is a public-key encryption algorithm (asymmetric), while AES is a symmetric key algorithm. The two algorithms work very differently. So, hybridization of these algorithms (RSA+AES) improves the result and minimizes the security risk in a better way. Henceforth, we apply neural network on the above result for the classification purpose. In this model, it trained the number of inputs using neural network and classify the outputs. In the training section, neural classifier learns its own classification rules. A neural network can be trained to perform a particular function by adjusting the values of the weights between elements. In addition to, calculate parameters such as FAR, FRR and accuracy as shown below:

Table I. Parameter evaluation (Brute Force attack)

S. No.	FAR	FRR	ACCURACY
1	0.0169	2.409	97.5741
2	0.040071	2.4032	97.556
3	0.012957	2.4117	97.5753
4	0.11731	2.3839	97.4987
5	0.11349	2.383	97.50351
6	0.084458	2.3907	97.52484
7	0.047802	2.3978	97.55439
8	0.083819	2.3908	97.52538
9	0.11875	2.3819	97.49935
10	0.078192	2.3909	97.53908

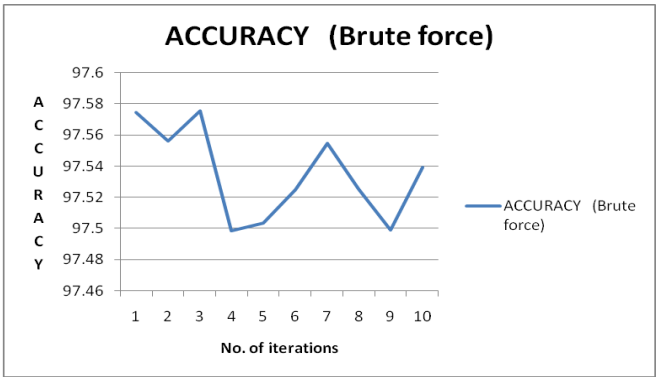

Figure 8. Graph of (Brute Force) Accuracy

According to Table I false acceptance rate (FAR), false rejection rate (FRR), and accuracy has been calculated. The value of FAR is seemed to be low and FRR is high relatively. In practical scenario, FAR should be low which shows the scenario that unauthorized user does not allowed to access an attempt. The value of FRR varies according to condition. The above graph in figure 8 shows the good amount of accuracy against brute force which tells the low access of an unauthorized user and high (as compared to FAR) rejection rate of an authorized user mistakenly. The second part of proposed work is to progress the security of web log files against IP Spoofing attack. The proposed framework is designed for the protection of log files by reducing number of fake requests as well as suspects with the help of prevention technique like GA. K-mean clustering is also used for generating fake and original requests .

Table II. Parameter evaluation (IP Spoofing)

S. No.	FAR	FRR	ACCURACY
1	0.002417	0.0095863	99.9888
2	0.0014866	0.010151	99.9884
3	0.0021483	0.0097493	99.881
4	0.0047706	0.0081585	99.871

5	0.0018217	0.012158	99.986
6	0.0027024	0.0094131	99.879
7	0.0017303	0.010003	99.9883
8	0.0017218	0.010008	99.9883
9	0.0029881	0.0092348	99.9878
10	0.0022505	0.0096813	99.9881

At last, neural network is used for the classification purpose and neural classifier computes parameters such as FAR, FRR and accuracy.

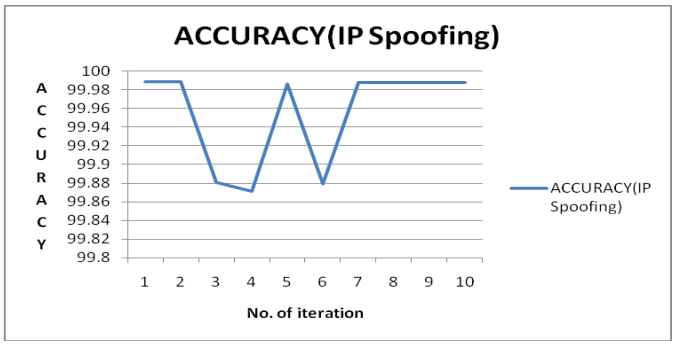

Figure 9. Graph of (IP Spoofing) Accuracy

Above Table II shows the values of FAR, FRR and accuracy. Figure 9 shows better accuracy as compared to previous framework. Eventually, we also compare the accuracy of brute force and IP spoofing attacks. The analysis shows that the value of FAR and FRR seems to be low in case of IP Spoofing as compared to Brute force.

The lower value of FAR and FRR increases the good amount of accuracy which has been shown in comparison graph figure10. In practical scenario of biometric security system, the value of FRR varies according to circumstances. Herewith, unauthorized

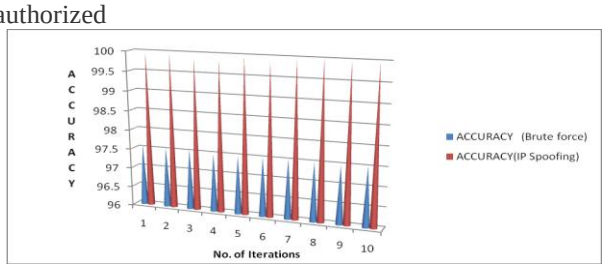

Figure 10. Comparison of accuracy

User will not be able to access an attempt as well as the system does not reject the authorized user mistakenly.

VI Conclusion

This research work analyses security risk over web log files against a couple of attack named as IP spoofing and brute force attack via genetic algorithm and some popular encryption algorithm. Then, a system is proposed, which makes use of neural network's learning ability and Genetic algorithm based optimization in web log files. And also neural network classifier is utilized for classification. This system uses the user defined dataset for measuring the performance. This implementation shows the healthier result by calculating FAR, FRR and the Accuracy values. In the future, this system could be improved by encrypting the data that is being transmitted on the network. As a result of this, in case of data leakage, the intruder would not be able to gain any important information.

REFERENCS

1. Pallavi Asrodia, Hemlata Patel, "Network Traffic Analysis Using Packet Sniffer" International Journal of Engineering Research and Applications, Volume 2, Issue3, pp.854-856, Jun 2012.
2. Wei Lu et. al, "Detecting New Forms of Network Intrusion Using Genetic Programming", IEEE Conference on Evolutionary Computation, pp. 2165 - 2172 , Dec. 2013.
3. ZU Wang, "Complement of an Extended Fuzzy Set", International Journal of Computer Applications, Volume 29, Issue 3, pp. 39-45, September 2011.
4.Tridivet. al, "A Real-time Intrusion Detection System Based on PSO-SVM", International Workshop on Information Security and Application, pp. 319-321, Nov. 2009.
5. R.natrajan, Dr. R. Sugumar, "A survey on attacks in web usage mining" International journal of innovative research in computer and communication engineering, Volume 2, Issue 5, May 2014.
6.Abdelhakim Herrouz, Chabane Khentout, Mahieddine Djoudi, "Overview of Web Mining Content Tool", The International Journal of Engineering and Science, Volume 2, Issue 6, pp. 1-6, Dec. 2014.
7. T. Xia, G. Qu, S. Hariri, M. Yousif, "An Efficient Network Intrusion Detection Method Based on Information Theory and Genetic Algorithm", IEEE International Conference on Performance Computing and Communications, pp. 11-17, April 2005.
8. Rimmy Chuchra, Bharti Mehta, Sumandeep Kaur, "Use of Web Mining in Network Security", International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 4, pp. 164-168, April 2013.
9. Neha, Dr. Maitreyee Dutta, "A Review paper on intrusion detection systems based on evolutionary algorithms", National Conference on Computing Technologies, pp. 37-42, April 2015.

Authors Profile

Neha is currently pursuing Regular M.E. from National Institute of Technical Teachers Training and Research, Chandigarh Sector 26. She has completed her B.Tech from Green Hills Engineering college, Kumarhatti, Himachal Pradesh University.

Dr. Maitreyee Dutta is currently working as Professor and Head in Electronics and Communication Engineering Department of National Institute of Technical Teachers Training & Research, Chandigarh, India. Former she was professor and Head in Computer Science and Department of National Institute of Technical Teachers and Training Research (NITTTR), India. She did her B.E from Gauhati University and ME from PEC University of technology. Her research area is Digital Image Processing, Digital Signal Processing, Data Warehousing and Data Mining etc.