FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 6 No. 3, June, 2022, pp 248 - 253
DOI: https://doi.org/10.33003/fjs-2022-0603-998

PERFORMANCE ANALYSIS OF CAPTCHA BASED BLOCKING AND RESPONSE OF AN INTRUSION DETECTION MODEL USING SIGNATURE

[*1]Abbas, Umar Farouk; [2]Abdulkarim, Abdulrahman

[1]Department of Mathematics and Statistics, Federal Polytechnic, Bauchi.
[2]Department of Information and Communication Technology, National Open University of Nigeria, Bauchi Study Center

*Corresponding authors' email: **ufabbas@fptb.edu.ng**

ABSTRACT

Intrusion Detection System is the process of intelligently monitoring the events occurring in a computer system or network, analyzing them for signs of violations of a security policy. Its primary aim is to protect the availability, confidentiality and integrity of critical networked information systems. This paper considered and reviewed a CAPTCHA based intrusion detection model. A method of incorporating signature was used along with the CAPTCHA in the intrusion detection model to clear the controversy identified in the existing model. The signature provides a means of identifying intruders that are able to by-pass the system as legitimate users. The model was implemented using a website hosted online. Dataset obtained from the site was analyzed based on Naïve Bayes classification model using confusion matrices. Implementation of the data analysis was carried out using RStudio software package. Analyzed results shows a better Detection Rate (DR), Accuracy (CR) and False Positive Rate (FPR). This shows that the developed system has significant capability of identifying intelligent spywares targeted at breaking CAPTCHA.

Keywords: Intrusion Detection System, Signature, CAPTCHA, Intrusion Prevention System

INTRODUCTION

With the advancement of the Internet and its capabilities, an increasing number of individuals are getting connected on a daily basis to take advantage of its services. On one hand, the Internet offers great commercial opportunities in terms of reaching end consumers. On the contrary, it introduces a lot of security threats to a business over the network. Various approaches are employed to assist the security of a business against threats and attacks. However, attackers are discovering new ways and techniques to break these approaches.

Intrusion could take a form of a process where software accesses web content that is protected with username and password, it uses various infiltration procedures to break the username and password. At a broader sense, intrusion may include both human and machine access to an account having web content that is secured with username and password. To overcome this problem, network security provides many techniques and one of the most important techniques is Intrusion Detection System (IDS) as suggested by (Souley & Abubakar, 2018). To achieve more security, the features of IDS, Intrusion Prevention System (IPS) and Honeypot can be combined as proposed by Yesugade et al (2016).

Intrusion Prevention System (IPS) is a tool that prevents spyware from entering a system. CAPTCHA is one of the methods employed in IPS. It's used in IPS to prevent unauthorized access to accounts by distinguishing between humans and spyware. With this fact, human being can solve CAPTCHA and gain authorization, spyware cannot.

With advancement in technology, Spyware creation is evolving at a rapid pace, and increasingly sophisticated spywares are designed to bypass CAPTCHA under IPS. Nachar, et. al., (2015) attained 57.3% - 76.7% success rate in bypassing CAPTCHA on Yahoo using edge and fuzzy logic segmentation and recognition method. Also, a success rate of 31.75% and 58.75% in solving reCAPTCHA was achieved by Sano, et. al., (2015) on Google's reCAPTCHA with continuous visual and audio symbols using sequence recognition method based on hidden markov model (HMM).

Based on the work by Abubakar, et al., (2020), CAPTCHA based IDS was designed and implemented to curb the security failure identified when CAPTCHA was used in IPS. The work integrated CAPTCHA in IDS to detect and identify spywares that bypass and break into the system. However, their method in making the software intruders to believe it is IPS may not be very efficient as some software intruders may attempt login ignoring the CAPTCHA and the textbox provided for it. It was observed in the model that, a genuine user can mistakenly fill in the CAPTCHA box, hence will be regarded as bot. Furthermore, bots may tend to neglect the CAPTCHA and can be considered been genuine, when this happened the entire system will be penetrated and compromised.

This paper would therefore consider to make an improvement on the work of Abubakar, et al., (2020) by proposing and implementing a model that would eliminate the weakness identified in their work. In this paper, CAPTCHA would be used as a detection tool by integrating it with a signature. CAPTCHA an acronym for **C**ompletely **A**utomated **P**ublic **T**uring tests to tell **C**omputers and **H**umans **A**part is widely used by websites to distinguish abusive programs from real human users. A CAPTCHA is a question that demands an answer: a test used to determine the user's awareness, and ascertain their humanness. This test is a means of verifying a user's ability to identify information such as symbols, words, equations, understand the context in which it is shown, and label it accordingly.

Related Work

This section consists of an overview of other researchers' work on related topics. The review of related work is aimed at contributing towards clearer understanding of the nature and meaning of the problem selected for this study.

Abubakar, et al., (2020) proposed a study to design IDS with intelligent redirector using CAPTCHA, a trap for detecting Intrusion by intelligent spywares. The work was set based on a design of an email website which was integrated with the new CAPTCHA based IDS system and hosted online. A dummy honeypot system was designed in which the IP

addresses of the software intruders that intruded into the system is being captured. The work was implemented and evaluated based on the performance metrics which includes Detection Rate (DR), Precision (PR), False Positive Rate (FPR), correctly and incorrectly classified instances of the system. Researchers used captured IP addresses as the datasets for the study. The work employed Waikato Environment for Knowledge Analysis (WEKA) and Python to analyze the experimental data. Their findings indicated a high percentage of Detection Rate and a very low False Positive rate as compared to the existing system. However, this system faking software intruders to believe it is IPS may not be very efficient as some software intruders may attempt login ignoring the CAPTCHA and the textbox provided for it, this will re-direct the intruder to the login authentication and somehow this may have been an intrusion bypassing the fake IPS and the CAPTCHA-trap IDS.

Malav, et al. (2016) proposed a system which combines the concept of Intrusion Detection System (IDS), Intrusion Prevention System (IPS) and Honeypot. Because various exploits are being used to compromise the network. These exploits are capable of breaking into any secured networks. In order to increase efficiency of network security, they introduce Honeypot. Honeypot detect attacks with the help of IDS; trap and deflect those packets sent by attackers. The result of their work indicates that the system handles multiple clients using the concept of Honeypot. They also pointed out that, Intrusion detection system (IDS) monitor whole network and looks for intrusion. When any intrusion occurs Honeypot will be activated. This activated Honeypot will divert the traffic to dummy/virtual servers & back track the source (IP address) or origin of that attack. The drawback of the system is that since it supports multiple clients including an attacker, the system can easily be compromised.

Khudadad & Huang, (2018) carried out a study to compare and explore the latest Anomaly detection methods for WSNs to enhance the workings of IDS in wireless environment. Series of experiments were directed to evaluation and simulate every approach in order to elaborate the effective detecting method. The methods explored were Cluster Based Method, Support Vector Machine (SVM), Naïve Bayes (NB) Method, and the Random Forest (RF) method. Many critical assessments metrics such as Confusion Matrix were used in time to Construct Model, General Classification Ratio and Memory Usage. KDD Cup 99 Intrusion Detection was used as a dataset. Result showed a suggestion to include Data Mining methods to efficiently notice the attacking threads or intrusions into WSN.

The work of El-Mourabit, et al. (2015) has compared and evaluated the newest anomaly detection intrusion techniques used in wireless sensor network, to improve the efficient technique for IDS in Wireless Sensor Network (WSN). According to their findings, the decision of choosing efficient IDS is a compromise between technique employed and performance metrics. Some critical evaluation metrics were used such as confusion matrix, general classification rate, time to build model, and memory consumption. For implementation, KDDCup'99 intrusion detection dataset on WEKA tool was used. According to their results, it was highly recommended to use the data mining techniques to detect effectively the intrusions and attacks in WSN. Issues such as hierarchical clustering patterns, using machine learning in resource management problem of WSNs, selecting and preprocessing an appropriate dataset are open and needed further research on their work.

Research Gap

Based on the survey and review on related work as discussed above, most of the research work identified security threats from intruders and attackers on the network. The security solutions proposed by most of the researchers were based on intrusion detection using a detection technique or combination of detection technique and a preventive technique. As presented from the work by Abubakar, et al., (2020), their work integrated CAPTCHA in Intrusion Detection System to detect and identify spywares that bypass and break into the system. Therefore, attackers can be detected, prevented and blocked, although it was evident that due to the weakness identified from their work, some software intruders may attempt login ignoring the CAPTCHA and the textbox provided for it, this would re-direct the intruder to the login authentication and somehow this could be an intrusion bypassing the fake CAPTCHA trap Intrusion Detection System.

METERIALS AND METHODS

This section describes the methodology and materials adopted in the development of the system.

Working Principle of the proposed System

The system model describes a developed website with login account, a text-based CAPTCHA and a signature generating module incorporated in the site. The CAPTCHA contains randomly generated text along with an instruction to click and insert a signature. The signature is meant to confuse non-human software intruders that can read CAPTCHA using intrusion technique to read the text but would not be aware of the instruction.

Any login with text from the generated CAPTCHA without considering the instruction to insert the generated signature along with the CAPTCHA would be granted access but be regarded as intrusion. This is the detection part of the system. Furthermore, login with the generated text CAPTCHA and without the signature would be regarded as intrusion and further attempt would be blocked by the admin using the captured IP Address of the intruding machine. This aspect is the prevention part of the system.

Below is the architecture of the proposed system shown in figure 1.

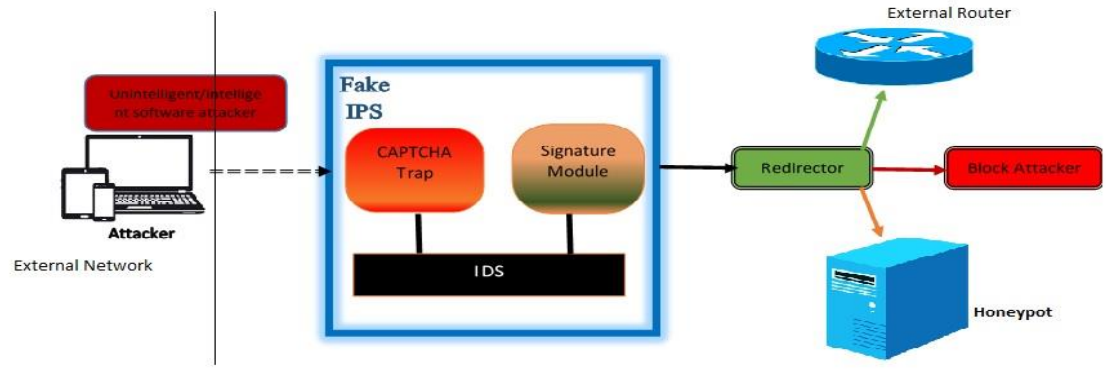

Figure 1. Architecture of the Proposed System

System Implementation Process

The implementation of these system is on a designed email service accout, which was designed and hosted online on a website as similar to that of the existing system.

Figure 2 shows the system interface, through which both genuine user and intruders can interact with the system. The interface demands Username/email, Password, CAPTCHA test and Signature. Any user that login with only username/email and password with CAPTCHA test only is identify as an intruder, even if the user pass the CAPTCHA test. A genuine user reads the attached message of signing in with signature. The system now distinguishes genuine users and intruders by the attached signature attempt. Therefore, login with a signature is considered genuine, otherwise is intruding.

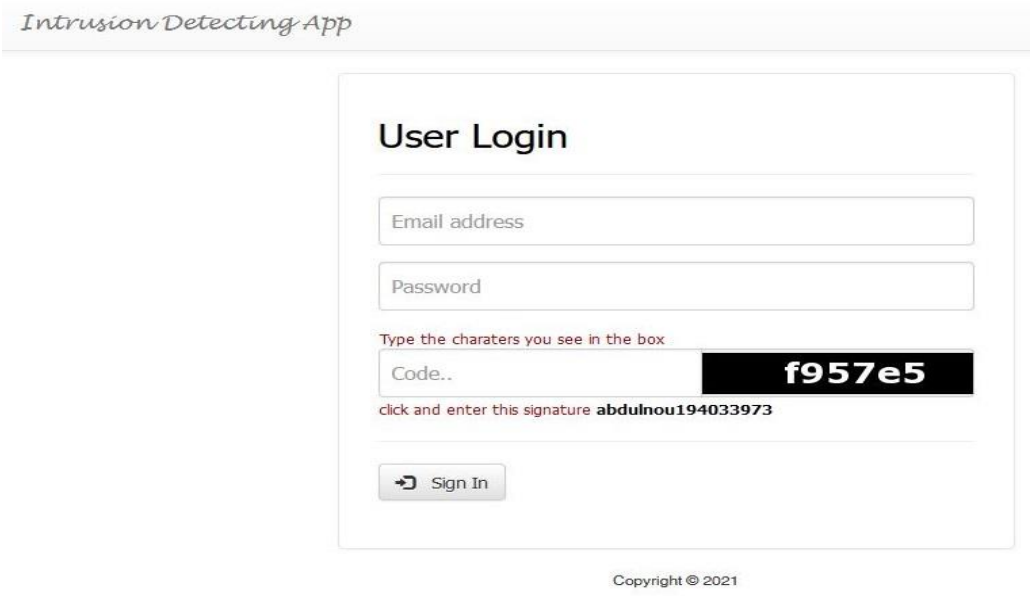

Figure 2. Login page of the system

All visitors identified without signature, can have their IP address blocked.

Experimental Data

Dataset to be used for the purpose of performance analysis would be a pool of data being captured from the site that is being hosted within a reasonable period of time. Among the data to be used as dataset for the analysis are:

1. IP address (Hostname)
2. Date of visit
3. Class (Human/bots)

Performance Metrics

For data analysis of this work, Naïve Bayes classification technique would be used to develop a model that would be used to classify our collected dataset.

Naïve Bayes classifier technique is based on the so-called Bayesian theorem particularly suited for situation where the dimensionality of the input is high. It is categorized as supervised learning as well as a statistical approach to classification (Stevens, 2016).

This method of analysis was also adopted to evaluate the performance of the existing model where the parameters obtained are as follows:

1. ***Detection rate (DR):*** The ratio between the number of correctly detected attacks and the number of the attacks. $DR = \frac{TP}{TP+FN}$
2. ***False positive rate (FPR):*** It is the ratio between the number of normal instances detected as attack and the total number of normal instances. $FPR = \frac{FP}{FP+TN}$

3. ***Classification rate/Accuracy (CR):*** It is defined as the ratio of correctly classified instances and the total number of the instances. $CR = \frac{TP + TN}{TP+TN+FP+FN}$
4. ***Precision rate (PR):*** Fraction of data instances predicted as positive, which are actually positive. $PR = \frac{TP}{TP+FP}$
5. ***Recall:*** It measures the missing part from precision rate, which is known as the percentage from the real attack covered by the classifier. This is the same as detection rate.
6. ***F-Measure (FM):*** Is the harmonic mean of the precision and recall at a given threshold. It is mostly preferred when only one accuracy metrics is needed as evaluation criterion. $FM = \frac{2}{\frac{1}{PR} + \frac{1}{Recall}}$

Where:

TP (True Positive): Intrusion successfully detected by the IDS.

FP (False Positive): Normal non-intrusive behavior that is wrongly classified as intrusive by the IDS.

TN (True Negative): Normal/non-intrusive behavior that is successfully labeled as normal/non-intrusive by the IDS.

FN (False Negative): Intrusion missed by the IDS, and classified as normal/non-intrusive.

Data analysis was carried out using R studio. R programming language was used for the analysis on the R studio. The "caret" and "e1071" library was used to implement the Naïve Bayes classifier and confusion matrix using the available dataset. These libraries are packages on the RStudio.

RESULT AND DISCUSSION

Presentation of Results

The performance of this system and that of the existing system were evaluated using the adopted metrics in Abubakar, et al. (2020). Both the robustness and the efficiency of the system were evaluated. Table 1 shows the analysis of the six-month IP addresses of all four-hundred and sixty-nine (469) visitors from October, 2021 to March, 2022, which were categorized into Genuine Users and intruders, according to the system.

Table 1: The visitors categorised in Month

Month	Total visitors	Genuine	Intruders
OCT, 2021	**80**	50	30
NOV, 2021	**110**	16	94
DEC, 2021	**80**	23	57
JAN, 2022	**69**	14	55
FEB, 2022	**78**	19	59
MAR, 2022	**52**	11	41
TOTAL	**469**	**133**	**336**

Analyses of the two systems were conducted in RStudio using the caret and e1071 packages. R was considered due to its availability and richness of its classification and evaluation metrics (Zhuoheng et al., 2018). Figure 4 shows the snapshot of the analysis.

```
- library(e1071)
- library(caret)
- ps = read.csv("Exceldata1.csv")
- train =ps[1:280,]
- test = ps[281:469,]
- ps_mod = naiveBayes(States~.,data=train)
- pre_ps = predict(ps_mod,test)
- confusionMatrix(pre_ps, as.factor(test$States), mode="everything")
Confusion Matrix and Statistics

          Reference
Prediction   0   1
         0 142  42
         1   4   1

               Accuracy : 0.7566
                 95% CI : (0.689, 0.816)
    No Information Rate : 0.7725
    P-Value [Acc > NIR] : 0.7314

                  Kappa : -0.006

 Mcnemar's Test P-Value : 4.888e-08

            Sensitivity : 0.97260
            Specificity : 0.02326
         Pos Pred Value : 0.77174
         Neg Pred Value : 0.20000
              Precision : 0.77174
                 Recall : 0.97260
                     F1 : 0.86061
             Prevalence : 0.77249
         Detection Rate : 0.75132
   Detection Prevalence : 0.97354
      Balanced Accuracy : 0.49793

       'Positive' Class : 0
```

Figure 4. The Snapshot of the Analysis of the result of the proposed system using RStudio

The proposed system used the same metrices to measure the robustness in terms of Detection rate, False Positive and True Negative Rates, as seen in table 2. Both the proposed and existing systems performance are compared in the table. The Output of the proposed system will be used as captured in Figure 4.

Table 2. The Performance Differences of the Existing System with the proposed system using the first three evaluation metrices

SN	Evaluation Metrics	Existing System	Proposed System	Performance Difference
	Number of datasets	**96**	**469**	**373**
1.	Detection Rate	0.45	0.75	0.30
2.	False Positive Rate	0.36	0.20	0.16
3.	True Positive Rate	0.75	0.77	0.02

Table 2 shows the numerical representation of the performance metrices with the clear differences that indicates the proposed system supercedes the existing system base on the above table.

The performance of the two system can also be compared with respect to the precision, recall and F-Measure as shown in Table 3.

Table 3. Evaluate the robustness based on Precision, Recall, F-Measure and Accuracy of the two systems.

SN	Evaluation metrics	Existing System	Proposed System	Performance Difference
1.	Precision	0.75	0.77	0.02
2.	Recall	0.75	0.97	0.22
3.	F-Measure	0.75	0.86	0.11
4.	Accuracy	0.70	0.75	0.05

Table 3 describe the differences of the two systems.

Discussion of findings

The proposed system was measured and compared with the existing system using these metrics which includes; Detection Rate, False Positive Rate, True Positive Rate, Precision, Recall, F-measure to ascertain the robustness of the system. All the above listed metrics have values ranges from 1 to 0; and performed better when the values go from 0 to 1 with the exception of false positive rate, which takes the reverse. The proposed system has a Detection rate of 0.75 or 75% while that of the existing system had Detection rate of 0.45 or 45%, false positive rate of 0.20 or 20% against 0.36 or 36% of the existing system and True positive rate of 0.77 or 77% against 0.75 or 75% of the existing system. This means our system is 1.7 times better in detecting bots and has 20% less problems than the existing system with 36% issues in misclassifying human as bots.

This may not be unconnected with the hidden signature that easily classified and distinguished normal user with the intruders. High scores for precision and recall show that the classifier is returning accurate results (Zhuoheng et al., 2018). From table 3, both precision and recall for the existing system is 0.75 or 75% where as that of the proposed system are 0.77 or 77% and 0.97 or 97% respectively. This means the proposed system is 1.03 and 1.3 times better than the existing system respectively. This of course is anticipated due to the complication, by the signature in the login. The accuracy of classification of both the existing and the proposed system was also captured in table 3, with 0.75 or 75% accurate in the proposed system and 0.70 or 70% accurate of the existing system. Here also our proposed system is 5% more accurate than that of the existing system. This is due to the high rate of the positive prediction of the proposed system over that of the existing system.

CONCLUSION AND FUTURE WORK

This work is an implementation of a CAPTCHA based blocking and response of an intrusion detection model. The work is aimed at improving the existing system developed using CAPTCHA based intrusion detection model with a redirector were it was observed that genuine user can be mistakenly regarded as an intruder and vice versa, which affects the systems performance on detection rate and false positive rate. The CAPTCHA based blocking and response of an intrusion detection developed using signature module, considering its efficiency on detection rate, precision and false positive rate. The analysis of the results signifies that the improved system is more effective than the existing system due to the signature that easily classify and distinguish normal user from intruders.

In a future work, there would be a need to make a research on the observed limitation made on the developed system were access is being granted to the system in the first successful attempt regardless of normal user or intruder. That is intruders get to be blocked only after gaining access to the system. Therefore, a further research would be needed in order for intruders to get blocked without getting access to the website.

REFERENCES

Abubakar, H., Souley, B., & Ya'u, A. G. (2020). An Improved CAPTCHA-Based Intrusion Detection System Based on Redirector Model. *Journal of Theoritical and Applied InformationTechnology, Volume 98, No. 03*, 429-440.

El Mourabit, Y., bouirden, A., Toumanari, A., & El Moussaid, N. (2015). Intrudion Detection Techniques in Wireless Sensor Network using Data Mining Algorithms: Comparative Evaluation Based on Attacks Detection. *International Journalof Advanced Computer Science and Application, Vol 6, No. 9*, 164-172.

Khudadad, M., & Huang, Z. (2018). Novel Intrusion Detection Methods for Security of Wieless Sensor Netwok. *Journal of Fundermental and Applied Sciences*, 173-189.

Malav, S., Avinash, M. S., Satish, N. S., & Sandeep, S. C. (2016). Network Security Using IDS, IPS, and Honeypot. *Interenational Journal of Recent Research in Mathematics Computer Science and Information Technology, Vol 2, Issue 2*, 27-30.

Milan, Sardana, H., & Singh, K. (2018). Reducing False Alarms in Intrusion Detection Systems – A Survey.

International Research Journal of Engineering and Technology, Volume 05, Issue 02, 9-12.

Nachar, R. A., Inaty, E., Bonnin, P. J., & Alayli, Y. (2015). Breaking Down CAPTCHA Using Edge Corners and Fuzzy Logic Segmentation/Recognition Technique. *Security and Communication Networks, Vol 8, No. 18*, 3995-4012.

Sano, S., Otusko, T., Itoyama, K., & Okuno, H. G. (2015). HMM- based Attacks on Google's ReCAPTCHA with Continous Visual and audio symbol. *International Journal of Information Processing, vol. 23, No. 6*, 814-826.

Souley, B., & Abubakar, H. (2018). A CAPTCHA – BASED INTRUSION DETECTION MODEL. *International Journal of Software Engineering & Applications, Vol.9, No.1*, 29-40.

Stevens, I., D. (2016). Using machine learning to detect bots in World of Warcraft. *Transactions on networking*19 (5).

Yesugade, K. D., Avinash, M. S., Satish, N. S., Sandeep, S. C., & Malav, S. (2016). Infracstructure Security Using IDS, IPS and Honeypot. *International Engineering Research Journal (IERJ), vol 2, issue3*, 851-855.

Zhuoheng, X., Zhenghao, Y., Simon, J., Micheal, R., Chris, R., Theerakorn, P., & Matthew, A. (2018). *Caret Versus Scikit-Leran AComparison of Data Science Tools.* Lanham Purdue University Krannert.