Secured web application based on CapsuleNet and OWASP in the cloud

Rohith Vallabhaneni, Sanjaikanth E. Vadakkethil Somanathan Pillai, Srinivas A. Vaddadi, Santosh Reddy Addula, Bhuvanesh Ananthan

1Department of Information Technology, University of the Cumberlands, Williamsburg, USA
2School of Electrical Engineering and Computer Science, University of North Dakota, Grand Forks, USA
3Department of Electrical and Electronics Engineering, PSN College of Engineering and Technology, Tirunelveli, India

ABSTRACT

The tremendous use of sensitive and consequential information in the advanced web application confronts the security issues. To defend the web application while it processing the information must requires the security system. The detection of attacks of web is made by the payload or HTTP request-based detection in association with the scholars. Some of the scholars provide secured attack model detection; however, it fails to achieve the optimal detection accuracy. In concern with these issues, we propose an innovative technique for the attack detection the web applications. The proposed attack detection is based on the novel deep CapsuleNet based technique and the process begins with pre-processing steps known as decoding, generalization, tokenization/standardization and vectorization. After the pre-processing steps the information are passed to deep CapsuleNet for extracting the features for attaining the temporal dependencies from the sequential data. The subtle patterns in the information also detected using the proposed work. Simulation is effectuated to demonstrate the effectiveness of the proposed work and compared with other existing works. Our proposed system provides better accuracy in detecting the attacks than the state-of-art works.

1. INTRODUCTION

Technology that operates within the browser window of a website is known as an internet app, companies must communicate with one another and provide products distantly, organizations privately and efficiently communicate with clients through web applications [1]. Regular quantity and utilization of web-based applications, like robots to communicate, intelligent assistants, and engines of suggestions in online shopping, on platforms like Instagram, Twitter, message boards, and numerous other services are growing substantially due to web technology [2]. Our everyday and personal lives now revolve around gadgets, and at the same duration, online apps have emerged as the main focus of computer hackers. The use of cloud services [3] is a relatively new idea that allows organizations and people to contract out technical support. People were reluctant to embrace cyberspace because it was sending away sensitive data from somewhere else. As more alluring online resources become available, opposition to online adoption has begun to wane as a substitute for safety issues. Using online applications and network authorities, the exchange of information is rapidly extended globally, which boosts the profitable growth of electronic commerce. This growth
includes both the examination of electronic hazards resulting from web use and the result of e-commerce. It thus affects the three safety targets of anonymity, reliability, and accessibility. To reduce the shortcomings, several measures have been implemented, including the use of security technologies such as internet-based firewalls and intrusion detection systems. Still, developing apps that are devoid of strength requires code-level encryption.

Developers must educate users on the risks associated with knowledge and the effects of assaults on pipelines as they evaluate and enhance these systems. Therefore, identifying the attempts and releasing the stretch is the final objective. Thus, creating a safer and more advanced online program is required. It continuously acquires program and privacy upgrades, so it is constantly up to current and less in danger of vulnerability intrusions. The primary drawback of web pages is that some of them present erroneous data. Some people attempt to obtain sensitive data to defraud customers of funding or provide subpar products or amenities. The primary drawback of web pages is that some of them present erroneous data. Some people attempt to obtain sensitive data to defraud customers of funding or provide subpar products or amenities. The attack detection in the web applications are challenging one with the following reasons:

- Attackers can easily exploit the comments, encoding payloads, and usage of puzzlement methods. This might affect the input validation process.
- Most of the web-based applications are outdated in the base of code and it can easily be attacked by attackers. Finding and resolving process in these cases are arduous and consumes high time.
- While considering the XSS attack, multifaceted protection solutions are necessary for defending the web page from this type of attack. Meanwhile, understanding the context is essential, since it might have change for the same application and make it arduous to detect the system.

The roadmap of the article are enclosed as, the review of the work is stated in section 2. The proposed work is elaborated in section 3. The simulation results are brought in the section 4. The work is concluded in section 5.

2. LITERATURE SURVEY

Al-Mohannadi et al. [4] have described a threat intelligence approach analyzing assault information obtained through wireless web services to assist with the risky knowledge that is now operating. The framework operates solely when there is a substantial amount of data available for evaluation linked to electronic attacks. Stretchy heap, an extremely versatile and scalable solution for information analysis and presentation, is used to analyze the log data. A more affordable method of obtaining assessment and utilizing the manufacturing process is examining the inspection of honeypot information to acquire security information. Hence, implementing it with current records is not feasible.

Mokbal et al. [5] have presented an artificial neural network (ANN) a method for detecting cross-site scripting (XSS) attacks that can be combined with the adaptive special retriever is suggested. The preliminary classifier is focused on stochastic scanning and initial information reliability. The subsequent form handles harvesting electronic information as characteristics of untreated knowledge and provides artificial neurons with these computational attributes. It serves as an additional level of protection for the local or the end-user side. Nonetheless, since storage is not accessed by multiple processes in online programs.

Sandikkaya et al. [6] have developed machine learning strategy an organizational structure to find harmful processes in cloud-based programs installed through platform-as-a-service (PaaS) services. To reduce the expense of launching an additional operating system for solitude, vendors might think about combining cloud apps. As a backup, firms can think about incorporating the technique into everyday distribution to boost protection even more. It is noted because the reliability of segmentation is exceptionally good. Thus, it is impossible to track client actions for each transaction.

Durai et al. [7] suggested SQL injection ontology (SQLIO) this puts the theory into practice flaws producing models and forecasting according to rules. The suggested structure effectively analyzed the risks and weaknesses that could lead to improper handling of internet-based intrusions. Ontology demonstrates the flaws, threats, and regulations that successfully manifest complex threats. The method of derivation ensures the list of possible assaults by taking into account the knowledge about website vulnerabilities. The information received is important for detectors and developers to manage threats indicate that the encryption procedure has been carefully thought out. It is insufficient, therefore, to offer the full safety answers required to create safe websites.

Chahal et al. [8] highlighted an orchestrated continuous vulnerability assessment (OCVA) focuses on the necessity of coordinating all computerized threat monitoring procedures’ ongoing risk assessments. It finds, tracks, displays, evaluates, reduces, and fixes system, resource, and website hazards. Offering the needed representations and statistics of the sensitive resources aids safety inspectors and producers in

Secured web application based on CapsuleNet and OWASP in the cloud (Rohith Vallabhaneni)
overcoming obstacles. From the start of the process of developing applications, employing freely available online scanning tools will not solely improve identification increases and reduce safety evaluation demands. Thus, websites consequently become less resistant to penetration.

Making the information accessible to unauthorized individuals is one way to mitigate the risk of cloud usage [9]. Sharing resources via the cloud creates new security issues, particularly when it comes to categorizing downloaded data [10], [11]. Conventional cloud formats protect data secrecy by encrypting it and isolating the hypervisor while allowing virtual machines to send visitors [12]. The best way to address the issue of cyber security on the cloud site is to encrypt information readiness; the recipient’s identifiable proof will undoubtedly verify the information and its accuracy and will guarantee the security and dependability of the application information [13]–[16]. These security protocols are employed to safeguard data, meet regulatory obligations, guarantee client safety, and create customized consumer inspection guidelines.

A thorough investigation has revealed that the average client always controls four cloud apps while running four. Furthermore, 41% of firms openly conceal the significant cost, according to the poll. We gradually monitor distributed computing security as our core workload moves to the cloud. This number is further supported by the Forbes 2017 research, which indicates that 49% of firms will postpone cloud operations and that 80% of IT budgetary plans will be spent within 15 months to fulfill the cloud’s aim [16]–[18]. Hammami et al. [19] proposed a comprehensive strategy to get rid of closed loops in the circulatory system as much as possible. Data mining encryption is what ensures the security of the information on the website. Nowadays, a lot of interdisciplinary training approaches have been put forth; sadly, not all of these approaches handle the issue of supplying data at particular times. A program is designed to ensure that different encrypted information signals in the cloud behave securely [20].

3. PROPOSED METHOD

The detailed explanations of proposed work in developing CapsuleNet based attack detection in the web application are provided in this section. The proposed work utilizes text classification methods and the information’s are pre-processed using the decoding, generalization, standardization/ tokenization techniques which are shown in Figure 1. The vectorization technique utilizes the word2vec which derives the required features. These are forwarded to the CapsuleNet and it extracts the features further in order to classify the information as attacks and normal. The details are explained below section.

![Figure 1. Proposed workflow for the detection of attacks in the web applications](image)

### 3.1. Pre-processing

The pre-processing step is the vital step in the attack detection system and it includes several process that are explained below in a wider context. The pre-processing step is crucial in ensuring the quality, relevance, and suitability of the data used for training and testing attack detection models. It directly impacts the performance, efficiency, and robustness of the detection system, making it a vital component in the overall cybersecurity infrastructure.

#### 3.1.1. Decoding

The decoding process is done to restore the input information to the original source form, since most of the time the attackers use encoding method such as HTML entity encoding and URL encoding. This might have prevents the conventional validation technique [21], meanwhile, the hyperlinks are converted to the numerical values in this step. For example, the decoded data are generalized to mitigate the disturbances of unnecessary data.

---

3.1.2. Tokenization or standardization

This is the process of providing tokens to the input information depending up on the various scripting language applied. This process identifies the initial and last labels along with the windows event and function names. The unique tokens are allocated further. From the vocabulary list the tokens are verified and if it is available then it is taken or else, constant delimiter is applied to it. Standardization is mainly applied when there is SQL queries.

3.1.3. Vectorization

This process is for word embedding and we utilizes word2vec [22] technique and is effectuated after the completion of tokenization or standardization approach. The vectorization begins with the generation of vocabulary with the data that are obtained from the tokenization and it also consists of common words. Using the neural network the probability of appearance of word along with the adjacent words. Henceforth, the mapping of vocabulary is done with the vector embedding process.

3.2. Deep CapsuleNet for the classification of attacks

The CNN is the most widely used deep learning approach for the detection, classification and prediction of attacks; however, it also possesses some demerits. The feature extraction ability of the CNN is the highlight of the technique and for transforming the upper scalar determination primary capsule layer is utilized. The architecture of the proposed CapsuleNet is shown in Figure 2. It employs batch normalization technique after the activation function and for that it uses Funnel activation [23]-[25]. The non linear transformation is effectuated with the FReLU and thus forms the spatial dependencies and lead to the improvement of visual layouts capturing capacity along with regular convolutions. Meanwhile, the dynamic routing algorithm is used to upgrade the impact of averting the pooling layer. The test is carried out few samples of dataset and the final outcome is taken from the length of the Eigen vector. The neuron of fully connected CapsuleNet is scalar and the weight can be denoted as $W^g_{jk}$. The value of the weight, in numeric form also in. Meanwhile, the neuron of Capsule is named as CapsNet. The capsule neuron weights are upgraded using the back-propagation.

$$Q_k = \sum_j x_{jk} u_{k/j}^\wedge, u_{k/j}^\wedge = W^g_{jk} u_j$$ (1)

$$x_{jk} = Q(h_{jk}) \frac{\exp(h_{jk})}{\sum_i \exp(h_{jik})}_{\text{max}}$$ (2)

Figure 2. Proposed architecture for the classification of attacks in the web applications

The upper CcapsuleNet outcomes and the learnable weights are $W^g_{jk}$ and $u_j$ respectively. The matrix weights are multiplied for each capsule output as jth and kth. With the association of linear sum the auxiliary coefficients are evaluated.

Secured web application based on CapsuleNet and OWASP in the cloud (Rohith Vallabhaneni)
The next layer is shifted using the forward propagation value of $Q$ and with the sigmoid activation function. The weight up-gradation is done with the evaluation of loss function for the entire network. After completing the first module, the feature maps are forwarded to the convolutional layer and mitigate the channel numbers with the less spatial relationship and hence FReLU is replaced with the leaky-ReLU.

$$m_x = \sum_{i \in \text{num}} t_i \max(0, n^+ - \|V_i\|^2) + \rho (1 - t_i) \max(0, \|V_i\| - n^-)^2 \quad (3)$$

The attack detection in web pages are evaluated using $n^+$and $n^-$. The accurate label with the vector probability length $\|V_i\|$ is $t_i = 1$. Here the category number is $i$. The proposed system effectively classifies the information as normal or attacks using the CapsuleNet technique.

4. EXPERIMENTAL RESULT AND DISCUSSION

An investigational analysis is discussed and the evaluation criterion is listed in this section. POSIX operating system with Intel Xenon Gold 6145 CPU, 64 bit ELF platform structure, 96 GB memory, 16 TB ROM and Matplotlib Version 3.2.0 is utilized. An evaluation measures evaluates the performance model of this proposed work. To determine the metric, the specific application objectives solves the types of problem. To prevent and detect web attacks, the proposed model utilized and its evaluation is analyzed using the SQLI–XSS payload dataset [1]. The attack and normal SQL queries with various payloads present in it. The real positives proportions are recognized accurately with the prediction true positive rate is named as precision and recall measures the correctly expected positive value proportion. The precision and recall harmonic mean defines the F-score value an accurate prediction percentages are determined with the indicator of accuracy.

4.1. Performance validation

The proposed evaluation accuracy with respect to the accuracies of training and testing is outlined in Figure 3. SQLI–XSS payload dataset based on its training as well as testing accuracies varies from 0 to 50 epochs. The legitimate is identified with the model ability that showed the astonishing result. While detecting harmful payloads, the accuracy increases from 20% to 99%. The maximum accuracy of training accuracy is 98% likewise, the testing accuracy increased up to 99% respectively. Compared to the accuracy of training, the proposed testing accuracy is higher. Figure 4 outlines the proposed evaluation loss with its testing and training loss results. The SQLI–XSS payload dataset depends upon the testing and training loss changes from 0 to 50 epochs. The identification of legitimate by the model capability that displayed the astonishing consequence. To determine the harmful payloads, the loss function minimizes from 0.3% to 0.06%. Where, 0.07% is for training loss and 0.06% to testing loss values. The loss of testing becomes minimal than training loss of proposed framework.

The comparative evaluation in terms of accuracy is plotted in Figure 5. The graphical representation shows that the comparison plot of accuracy based on the methods such as TI [5], ANN [6], ML [7], SQLIO [8], and proposed model. The level of accuracy get varied from 10 to 50 number of epochs. To perform on different data, an insightful information is offered SQL–XSS Payload dataset based on the proposed CapsuleNet structure. The low false positives with the highest accuracy of 96% is obtained based on the
secured web application based on CapsuleNet and OWASP in the cloud (Rohith Vallabhaneni)

The validation based on the precision outcomes with its comparative plot is outlined in Figure 6. The graphical plot reveals that the comparative plot of precision depending upon the methods namely TI [5], ANN [6], ML [7], SQLIO [8], and proposed model. The precision level become changed from the number of epochs such as 10 to 50. While describing on different data, an insightful information is obtainable dataset of SQL–XSS Payload according to the proposed CapsuleNet structure. The rates of low false positives by a superior precision rate of 95.40% is attained constructed on the proposed outline. In terms of precision, the proposed result is superior and worthy based on the dataset of SQL–XSS Payload. The model of proposed work undertakes 88.43%, 90.03%, 92.10%, 93.42%, and 95.40% in case of 10 to 50 number of epochs and its outcome develops higher than that of other existing TI, ANN, ML, and SQLIO models. The proposed demonstrates sophisticated capture rate of attack than previous methods.

The comparison validation in terms of recall is plotted in Figure 7. The graphical representation shows that the comparison plot of recall based on the methods such as TI [5], ANN [6], ML [7], SQL IO [8], and proposed model. The level of recall get varied from 10 to 50 number of epochs. To perform on different data, an insightful information is offered SQL–XSS Payload dataset based on the proposed CapsuleNet structure. The low false positives with the highest recall of 95.78% is obtained based on the proposed framework. In terms of recall, the proposed result is higher and good with respect to the dataset of SQL–XSS Payload. Our proposed model accomplishes 89%, 90.97%, 91.54%, 93.65%, and 95.78% recall in terms of 10 to 50 number of epochs that results becomes superior than other existing TI, ANN, ML, and SQLIO models. The proposed works proves sophisticated capture rate of attack than existing methods.
The evaluation based on the F-score results with its comparative plot is outlined in Figure 8. The graphical plot discloses that the comparative plot of F-score depending upon the methods namely TI [5], ANN [6], ML [7], SQLIO [8], and proposed model. The F-score level become changed from the number of epochs such as 10 to 50. While describing on different data, an insightful information is obtainable dataset of SQL–XSS Payload according to the proposed CapsuleNet structure. The rates of low false positives by a superior F-score rate of 95.58% is attained constructed on the proposed outline. In terms of F-score, the proposed result is superior and worthy based on the dataset of SQL–XSS Payload. The model of proposed work undertakes 88.70%, 90.48%, 91.81%, 93.53%, and 95.58% F-score in case of 10 to 50 number of epochs and its outcome develops higher F-score than that of other existing TI, ANN, ML, and SQLIO models. The proposed demonstrates sophisticated imprisonment rate of attack than existing works.

![Figure 8. The comparative evaluation in terms of F-score](image)

Table 1 depicts the comparative outcomes of time consumption. The tabulation representation of time consumption with its comparison result varying with respect to seconds. The comparative outputs of time consumption based on the methods like TI [5], ANN [6], ML [7], SQLIO [8], and proposed model. To describe various data, an insightful data is obtainable dataset of SQL–XSS Payload according to the proposed CapsuleNet structure. The time consumption output is 1.2s for proposed work that becomes minimal compared to previous TI [5], ANN [6], ML [7], and SQLIO [8] methods.

The combination of CapsuleNet-based intrusion detection systems with OWASP-based secure web application practices in a cloud environment offers a comprehensive approach to cybersecurity, providing organizations with the tools and capabilities needed to detect, prevent, and respond to evolving cyber threats effectively. The combination of CapsuleNet-based intrusion detection systems with OWASP-based secure web application practices in a cloud environment offers comprehensive protection against cyber threats, providing organizations with the tools and capabilities needed to safeguard their web applications and data assets effectively.

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Time consumption (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI [5]</td>
<td>3.4</td>
</tr>
<tr>
<td>ANN [6]</td>
<td>6.5</td>
</tr>
<tr>
<td>ML [7]</td>
<td>5.5</td>
</tr>
<tr>
<td>SQLIO [8]</td>
<td>4.7</td>
</tr>
<tr>
<td>Proposed framework</td>
<td>1.2</td>
</tr>
</tbody>
</table>

5. CONCLUSION

The work in this article is about to secure the web applications in the cloud environment from the various attacks. To achieve this we proposed an approach known as deep learning based CapsuleNet technique which classifies the information as attack or normal using the extracted features. The information that is obtained from the dataset is pre-processed to remove the control words, blank words and decode the accessed to the original format. The vectorization is effectuated to extract the features and then the required features are extracted in the first layers of CapsuleNet. Further, simulation is made and analyzed the statistical parameters such as accuracy, precision, recall and time consumption to validate the performance of
the proposed work. Furthermore, the secured defended by the proposed work is higher with the attained accuracy, precision, recall and F-score with the values such as 96%, 95.40%, 95.78%, and 95.58%. Thus ensures the protection in cloud environment. In future, the authors will extent this research in CapsuleNet-based OWASP-based secured web applications in cloud systems, contributing to more effective, adaptive, and resilient cybersecurity solutions for protecting web applications against evolving threats and vulnerabilities.

ACKNOWLEDGEMENTS

The author would like to express his heartfelt gratitude to the supervisor for his guidance and unwavering support during this research for his guidance and support.

REFERENCES


Secured web application based on CapsuleNet and OWASP in the cloud (Rohith Vallabhaneni)
BIOGRAPHIES OF AUTHORS

Dr. Rohith Vallabhaneni received his Ph.D. degree in Information Technology from the University of the Cumberlands in the United States of America. As a Senior IEEE member, he has contributed significantly to various research journals, both as an author and co-author. His professional journey is marked by a strong work ethic and a profound ability to lead teams in addressing organizational challenges. His exemplary team leadership skills and dedication to his work underscore his contributions to the field of Information Technology. He can be contacted at email: rohit.vallabhaneni.2222@gmail.com.

Sanjaikanth E. Vadakkethil Somanathan Pillai (Senior Member, IEEE) holds an MS in Software Engineering from The University of Texas at Austin, Texas, USA, and a BE from the University of Calicut, Kerala, India. Currently pursuing a Ph.D. in Computer Science at the University of North Dakota, Grand Forks, North Dakota, USA, his research spans diverse areas such as mobile networks, network security, privacy, location-based services, and misinformation detection. He is a proud member of Sigma Xi, The Scientific Research Honor Society, underlining his commitment to advancing scientific knowledge and research excellence. He can be contacted at email: s.evadakkethil@und.edu.

Srinivas A. Vaddadi is a dynamic and forward-thinking professional in the field of Cloud and DevSecOps. With a solid educational foundation in computer science, Srinivas embarked on a journey of continuous learning and professional growth. Their relentless pursuit of knowledge and commitment to staying at the forefront of industry advancements has earned them recognition as a thought leader in the Cloud and DevSecOps space. He can be contacted at email: Vsad93@gmail.com.

Santosh Reddy Addula holds a master's degree in Information Technology from the University of the Cumberlands in Kentucky, United States of America. He has over five years of experience working in the IT industry, and he has showcased expertise across various domains in IT. He has 3+ patents, has contributed as an author and co-author of research articles, and has a role as a reviewer for esteemed journals such as IEEE, Springer, and Elsevier. He can be contacted at email: santoshaddulait@gmail.com.

Bhuvanesh Ananthan received the B.E. degree in Electrical and Electronics Engineering from Anna University in 2012, M.Tech. in Power System Engineering from Kalasalingam University in 2014 and Ph.D. degree from Faculty of Electrical Engineering of Anna University in 2019. He has published more than 65 papers in reputed international journals, 25 papers in international conferences and 10 books. He is a life time member of International Society for Research and Development, International Association of Engineers. He can be contacted at email: bhuvanesh.ananthan@gmail.com.