République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
École Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbès

THESIS

To obtain the diploma of **State Engineer**
Field: **Computer Science**
Specialty: **Ingénierie des Systèmes Informatiques (ISI)**

Theme

Enhancing Web Application Security through Advanced Techniques and Deep Learning-Based Web Application Firewall

---

Presented by:
Lachemat Mohamed Fouad
Slamat Mohamed Souhaib

Submission Date: **29 June 2023**
In front of the jury composed of:

Dr. Alaa Eddine Belfedhal Supervisor

Academic Year : 2022/2023

Abstract:

Web applications play a vital role in today's digital landscape, serving as platforms for various online services. However, with the increasing sophistication of cyber threats and the ever-evolving nature of web vulnerabilities, ensuring the security of web applications has become a paramount concern. This thesis addresses this challenge by proposing an advanced approach to enhance web application security through the implementation of a Deep Learning-based Web application firewall.

The objective of this research is to develop a robust and intelligent WAF capable of effectively detecting and mitigating web application attacks. The proposed WAF leverages state-of-the-art Deep Learning techniques, specifically the DistilBERT model, for payload content analysis and classification. By training the model on a diverse dataset comprising normal and malicious payloads, the WAF learns to identify patterns and distinguish between legitimate and malicious requests.

To evaluate the performance of the implemented WAF, comprehensive testing is conducted using various attack scenarios and real-world web application traffic. The results demonstrate the effectiveness of the WAF in accurately detecting and mitigating web application attacks while maintaining a low false positive rate. The WAF exhibits high accuracy and efficiency, with real-time response times, making it suitable for deployment in production environments.

In addition to the WAF implementation, this thesis also explores advanced techniques such as WordPiece tokenization and training on specific datasets to further enhance the model's accuracy and understanding of payload content. These techniques contribute to the overall effectiveness of the WAF in identifying and mitigating both known and emerging web application threats.

Overall, this research contributes to the field of web application security by providing an advanced and intelligent solution for detecting and mitigating web application attacks. The proposed Deep Learning-based Web Application Firewall, along with its advanced techniques, strengthens the security infrastructure of web applications, safeguarding them against a wide range of potential threats and ensuring the protection of sensitive data and user privacy.

Keywords: WAF,web application firewall , firewall ,cybersecurity, web vulnerabilities

LIST OF ACRONYMS

AI Artificial Intelligence. 25–27, 38

BERT Bidirectional Encoder Representations from Transformers. 36, 37

CDN content delivery network. 23

CNN Convolutional Neural Network. 5, 34, 35

DL Deep Learning. 25

DNS Domain Name System. 23

FNN Feed Forward Neural Network. 32

HTTP Hypertext Transfer Protocol. 21, 22

IBM International Business Machines Corporation. 26

IDS intrusion detection system. 21, 22

IoT Internet of Things. 22

IPS intrusion prevention system. 21, 22

ISO the International Organization for Standardization. 12

LSTM Long Short Term Memory. 34

ML Machine Learning. 25

MLP Multi-Layer Perceptron. 32

NGFW Next-generation firewalls. 22

NLP Natural language processing. 35–37

OSI The open systems interconnection. 12

OWASP Open Web Application Security Project. 21

PCI DSS Payment Card Industry Data Security Standard. 22

RNN Recurrent Neural Network. 5, 32–34

SQL Structured Query Language. 43

WAF Web application firewall. 1, 3, 12, 21–24, 40, 43

XSS Cross-site scripting. 40