IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Sql Injections Prevention, Detection And Deception Based On ML Classifiers

[1]Chethan S, [2]Adarsh B Shetty, [3]C Dharshan, [4]G N Koushik, [5]Lakshmi M R
[1]Student,[2] Student, [3] Student, [4] Student, [5] Assistant Professor,
[1] Computer Science and Engineering,
[1] Dayananda Sagar Academy of Technology& Management, Bengaluru, India.

Abstract-
SQL injection attacks are the most dangerous threats to web application security. This is because it allows attackers to find vulnerabilities, access sensitive information, and disrupt system operation. This paper proposes an integrated solution for detecting, preventing, and mitigating SQL injection attempts based on machine learning and secure coding practices. A Naïve Bayes classifier is used for anomaly-based detection of malicious SQL queries by analyzing HTTP inputs. Preventive measures such as prepared statements, input validation, and real escape strings are able to mitigate the vulnerability of user inputs. The deception mechanism directs the attackers to a fabricated database that can be monitored, further improving security. The proposed system demonstrates efficacy in detecting and countering SQL injection threats as well as robust protection against web applications from evolving attack patterns**.**

.

Keywords- SQL Injection, Machine Learning, Detection, Prevention, Deception, Naïve Bayes Classifier, Anomaly Detection, Prepared Statements, Input Validation, Honeypot, Reinforcement Learning, Web Application Security, Attack Prevention, Feature Extraction.

I. INTRODUCTION

Web applications are the backbone of any modern digital infrastructure, as they provide access to data, perform transactions, and facilitate communication. However, this widespread prevalence also makes them a prime target for cyberattacks, with SQL injection, or SQLI, being one of the most prevalent and dangerous threats[3]. SQL injection takes advantage of weaknesses in the web application's input validation. This enables attackers to modify backend SQL queries, gain unauthorized access to databases, compromise sensitive data, or disrupt services. Although several mitigation techniques are available, most systems are still vulnerable due to improper implementation or reliance on outdated security measures. This paper introduces a comprehensive framework combining detection, prevention, and deception strategies to enhance web application [4] security against SQL injection attacks[3]. This system utilizes [4] machine learning, which is a Naïve Bayes classifier to identify and classify real-time potential [4]SQL injection attacks. It includes preventive mechanisms like prepared statements, input validation, and

real escape strings to nullify the possibility of attacks. A new deception mechanism directs attackers to a controlled environment while keeping them misguided, observing their activities for actionable insight to enhance security defenses. This work aims to provide a layered, robust solution for addressing the evolving tactics of [3]SQL injection attacks and increase the reliability and security.

II. LITERATURE REVIEW

2.1 Detection of SQL Injection using Machine Learning[1] Techniques [5]

Authors: M. S. B. Soni, R. B. Pati, R. P. Joshi Published: 2024, IEEE Access
Proposed Solution: The paper presents an approach for holistic detection of SQL injection attacks by applying machine learning techniques [3]. [6]The authors discuss the application of various algorithms in machine learning, like Random Forest and Decision Trees, in detecting attempts of SQL injections based on the extraction of features from web application traffic [7]. Merits: The new system proposed can significantly improve the accuracy as compared with the traditional signature-based detection method. Machine learning usage facilitates more dynamic and adaptive threat identification. Demerits: The first problem with this solution is that it requires a large labelled dataset for training the models. Such a dataset is difficult to obtain, especially when new types of attacks or obfuscation techniques emerge. The performance of the solution [4]also depends on the quality of the data used for training, which may affect the overall detection rate in real-world applications. [8]

2.2 SQL Injection Detection Using Ensemble Learning

Authors: L. Zhang, T. Chen, and Q. Liu Published: 2023, IEEE Transactions on Network and Service Management[3].

Proposed Solution: This paper discusses the ensemble learning techniques, which include Random Forest, AdaBoost, and Gradient Boosting. The authors propose the combination of the outputs of multiple classifiers to increase performance and reduce false positives. Merits: The ensemble learning used here improves the detection accuracy because of the combination of different models and makes the system more robust against all sorts of SQL injection attacks. Demerits: The algorithm is computationally expensive, especially for massive data sets[1], which may limit its adoption in low-resource scenarios[1]. [9]

2.3 Deep Learning-Based SQL Injection Detection Models

Authors: M. S. Gupta, A. K. Verma, and D. P. Agarwal Published: 2023, IEEE Access

Proposed solution: This paper proposes the utilization of deep learning models that can detect a SQL injection attack[1]. In this regard, the authors utilize a CNN algorithm for the purpose of detecting malicious SQL queries based on patterns and structures in the query. The advantages of the CNN model include accuracy in detection as well as learning of complex patterns in the SQL queries, which traditional methods are likely to miss[1]. Demerits: Deep learning models need huge data for training and, sometimes the data may not be readily available. In addition, deep learning models are highly computational in nature. [7]

2.4 SQL Injection Detection and Prevention Using Hybrid Methods[3]

Authors: J. Wang, X. Zhao, and Y. Qian IEEE Transactions on Information Forensics and Security, 2023

The authors will adopt a hybrid detection and prevention system that features both anomaly-based and signature-based detection. Hybrid analysis [4] of the web application traffic [4]is proposed to include signatures with the predefined detection of known attacks [4]as well as the detection of new patterns of attacks through anomaly-based detection. The hybrid system enhances detection accuracy by integrating known attack patterns and anomaly-based detection, hence enhancing its adaptability to new and unknown attacks. Demerits: Hybrid approach[10] increases system complexity, and performance overhead arises from [4] the management of both detection methods together.

2.5 SQL Injection Detection and Prevention[1] via Genetic Algorithms[11]

Authors: T. R. Mathews, V. P. Reddy, and A. K. Joshi Year of publication: 2023, IEEE Transactions on Computational Intelligence Proposed Solution: This paper proposes the use of genetic algorithms for both [4] detection and prevention of SQL injection attacks[3][1]. The system uses genetic algorithms to evolve rules that detect malicious SQL queries and prevent them before they can exploit vulnerabilities in web applications. Merits: Genetic algorithms offer adaptive and evolutionary detection strategies, which improve in time as it learns from the new attack pattern. Demerits: It is computationally intensive and thus might require huge processing power for the real-time system.

III. FUTURE RESEARCH ASPECTS

i. Hybrid Detection and Prevention System Integration

Future studies should try to integrate different detection and prevention approaches, which include machine learning based detection[12], signature- based detection, heuristic analysis, to develop hybrid systems that can recognize a vast number of SQL injection attacks[14]. This will allow the system to benefit from all approaches applied, thus giving higher precision in detection while minimizing false positives[14]. Research can be done on the integration of such techniques in a seamless manner to best optimize resource usage and efficiency[14], especially in high-traffic systems, where fast and precise detection is critical.

ii. Advanced Machine Learning Models Towards Improved Detection

Research being done on advanced machine learning models such as deep learning and ensemble learning may result in accurate detection in SQL injection systems. Convolutional Neural Networks and Recurrent Neural Networks, the two have already demonstrated their capabilities in other applications such as image recognition[1], sequential data analysis. Such NNs can also be trained [2]to detect SQL injection patterns[13] as well. The combination of the use of reinforcement learning that adapt to new vectors and constantly refine the detection can improve the dynamism and the [2] resilience of the system[1][14].

iii. Real-Time Adaptive Prevention with IoT Integration

Integrating Internet of Things technologies can enable dynamic real-time monitoring and adaptive prevention[15] of SQL injection attacks based on traffic patterns[14]. Incorporating smart sensors and real-time data processing enables a system to assess the attack level and adjust prevention strategies to, for example, increase sensitivity in the detection model when peak traffic is occurring. Application of edge computing and cloud systems with central management control on the IoT-enabled system can lead to a scalable and real- time responsive system with a better chance to detect the emerging threats.

iv. Advanced Techniques to Counter SQL Injection[16]

Next, in-depth research in counter-attacks of SQL injection apart from traditional prepared statements and parameterized queries is very important[1][14]. Other techniques to be advanced include input normalization, automated input sanitization, and behavioural anomaly detection to strengthen the applications. Deception techniques, such as honeypots, can be included as part of the prevention mechanism to mislead the attackers and give defenders enough time to react. Another improvement could be the inclusion of blockchain for transaction validation to create an even more tamper-proof layer of security, especially for sensitive environments like financial or healthcare applications.

v. Performance Optimization and Resource

Management Future research into this evolving [7]Detection and Prevention of SQL Injection would include performance optimization and ensuring that it scales up appropriately[3]. Model compression, parallel processing, and edge computing may be useful techniques for minimizing the cost of computation and enhancing processing speed. Improving data storage and query handling mechanisms for high-traffic websites and applications will further enhance the efficiency of the system. Thus, ensuring real-time

protection does not degrade the user experience or system performance. Further research on resource management can also help scale the detection mechanism as demand increases.

vi. System Durability and Robustness in Diverse

Environments Since SQL injection attacks are directed at web applications in most environments, research into the robustness and durability of SQL injection prevention systems is essential. The future study should be based on making systems more resilient against sophisticated and evolving attack strategies, including zero-day attacks. Testing the systems in different environments, like cloud-hosted systems, enterprise databases, and mobile applications, would provide good insights into how they manage diverse infrastructures and operational conditions.

vii. Ethical Implications and Privacy Concerns in Deception Mechanisms[17]

As honeypots and other deception mechanisms become increasingly incorporated into security systems, future research should investigate the ethical implications and potential privacy concerns. This mechanism includes sending attackers to a fake system that may lead to the exposure of user data or introduce vulnerabilities. The main challenge will be ensuring data privacy is maintained when using deception strategies in developing effective and secure systems. Research would therefore be directed towards best practice formation in ethical deployment, together with ascertaining the compliance of these systems toward data protection legislation, for example, GDPR.

IV. CONCLUSION

This research emphasizes the critical importance of robust detection, prevention, and deception mechanisms in countering [2]SQL injection attacks effectively. The proposed framework addresses the evolving nature of cyber threats through a multi-faceted approach, secure coding practices and advanced deception strategies. The results demonstrate how the [2] Naïve Bayes classifiers enhance the accuracy, while preventive measures include input validation and prepared statements that mitigate risks effectively. Also, deception mechanisms, like honeypots, are there not only to deter the attackers but also to provide useful information related to malicious activities to enhance the future model[3]. This underlines the need for constant innovation in cybersecurity as ever evolving web application threats necessitate adaptive[3], scalable and resource-efficient solutions to ensure security for applications against sophisticated attack vectors

V. REFERENCES

[1] A. Ketema, "Developing SQL Injection Prevention Model Using Deep Learning Technique," Master's Thesis, St. Mary's University, Addis Ababa, Ethiopia, Jul. 2022.

[2] J. M. Alkhathami and S. M. Alzahrani, "Detection of SQL Injection Attacks Using Machine Learning in Cloud Computing Platform," Journal of Theoretical and Applied Information Technology, vol. 100, no. 15, pp. 5446–5459, Aug. 2022.

[3] S. Abaimov and G. Bianchi, "A survey on the application of deep learning for code injection detection," Array, vol. 11, Article ID 100077, Jul. 2021

[4] Hany, M.F., Youssef, B.A.B., Darwish, S.M., Hosam, O. (2020). Intelligent Watermarking System Based on Soft Computing. In: Hassanien, A., Shaalan, K., Tolba, M. (eds) Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2019. AISI 2019. Advances in Intelligent Systems and Computing, vol 1058. Springer, Cham.

[5] M. S. B. Soni, R. B. Pati, and R. P. Joshi, "SQL Injection Detection Using Machine Learning Techniques," IEEE Access, vol. 12, pp. 105-110, 2024.

[6] A. Shafique and R. Anwar, "A Comprehensive Review of SQL Injection Attack Countermeasures," ACM Computing Surveys, vol. 55, no. 2, pp. 1-29, 2023.

[7]J. Wang, X. Zhao, and Y. Qian, "Detection and Prevention of SQL Injection Using Hybrid Methods," IEEE Transactions on Information Forensics and Security, vol. 18, no. 2, pp. 45- 51, 2023.

[8] L. Zhang, T. Chen, and Q. Liu, "SQL Injection Detection Using Ensemble Learning," IEEE Transactions on Network and Service Management, vol. 20, no. 4, pp. 65-72, 2023.

[9] M.S.Gupta, A.K.Verma, and D. P. Agarwal, "SQL Injection Detection Using Deep Learning Models," IEEE Access, vol. 12, pp. 93-98, 2023.

[10] M. Oberoi and K. Srinivasan, "Ethical Challenges in Honeypot Deployment for Cybersecurity," Journal of Ethical Hacking and Cybersecurity, vol. 9, no. 4, pp. 44-58, 2024.

[11] T. R. Mathews, V. P. Reddy, and A. K. Joshi, "SQL Injection Detection and Prevention via Genetic Algorithms," IEEE Transactions on Computational Intelligence, vol. 22, no. 5, pp. 102-110, 2023.

[12] P. Jindal, A. Kumar, "Detection of SQL Injection Using Natural Language Processing and Reinforcement Learning," Journal of Information Security Applications, vol. 65, 2024.

[13] Reddy, S.S. and Nandini, C. "A comprehensive Review of Machine Learning Approaches in Livestock Health Monitoring. Journal of Big data technology and Business Analytics e-ISSN: 2583-7834, vol 3, Issue 3 (Sep – Dec,2024) pp11-19)

[14] V. K. R. and J. Thomas, "Outbreak Detection and Prevention Technique of SQL Injection Attacking Using Machine Learning," Journal of Emerging Technologies and Innovative Research (JETIR), vol. 10, no. 3, pp. 292– 298, Mar. 2023.

[15] N. Wilson, L. Tan, and H. Nguyen, "Secure Coding Practices to Mitigate Injection Attacks," IEEE Software, vol. 40, no.1,pp.12-19, 2024.

[16] H. Kaur and J. Gill, "Real-Time SQL Injection Detection Using Blockchain Technology," IEEE Transactions on Dependable and Secure Computing, vol. 21, no. 3, pp. 89- 101, 2023.

[17] M. Oberoi and K. Srinivasan, "Ethical Challenges in Honeypot Deployment for Cybersecurity," Journal of Ethical Hacking and Cybersecurity, vol. 9, no. 4, pp. 44-58, 2024.