A New Approach in Diagnosing and Preventing SQLIA with Large Language Models (LLMs)

Amin Rezanejad
Master's student in Computer Engineering, University of Guilan, Rasht, Iran.
aminrezanejad20@yahoo.com

Amir Seyed Danesh
Faculty of Technology and Engineering, East of Guilan, University of Guilan, Rudsar-Vajargah, Iran.
seyeddanesh@guilan.ac.ir

Farid Feyzi
Assistant Professor, Faculty of Engineering, University of Guilan, Rasht, Iran
feizi@guilan.ac.ir

ABSTRACT

SQL injection attack is considered to be one of the most important and common methods of intrusion into databases. The current research was conducted with the aim of improving the security of databases and web applications by relying on artificial intelligence (AI) and natural language processing (NLP). According to the research done in the past, to detect and prevent SQL injection attacks, we will present a new approach using natural language processing (NLP) approaches such as large language models (LLMs), which has the ability to reduce the vulnerability of the database. and ineffectiveness of SQL Injection hacker attacks.

KEYWORDS: SQL injection attack, large language models, Natural Language Processing, database security, software vulnerability.

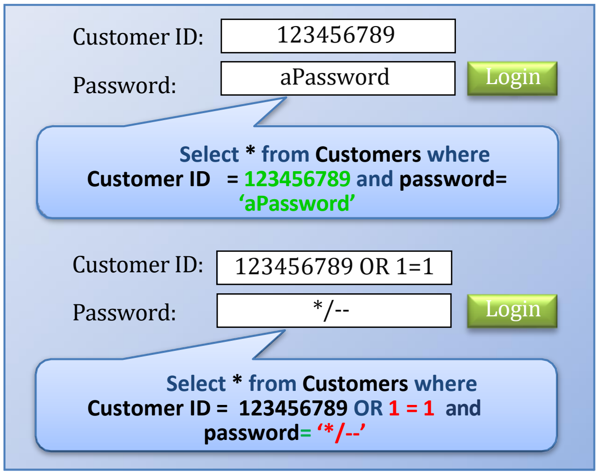
1 INTRODUCTION

Protecting important and confidential information in computer systems and computer networks has always been an important issue. Also, in this context, with the expansion of the use of the Internet and the increasing popularity of web applications, the importance of the topic of data and information security has doubled. Hackers are always looking to infiltrate websites and break into databases, and SQL injection attack is one of the most common and important methods of database hacking. After this introduction, in the second part of this research, we will introduce the SQL injection attack. In the third part, we will talk about the past valuable researches in the diagnosis and prevention of SQL injection attacks with machine learning techniques. Then, in the fourth part, we will discuss a new approach using natural language processing (NLP) approaches such as large language models (LLMs) to identify and prevent SQL injection attacks.

2 INTRODUCTION OF SQL INJECTION ATTACKS

From the point of view of the Open Web Application Security Project (OWASP) Foundation, which plays a special role in huge projects to improve software security and secure the web, SQL injection attack is placed in the list of top ten security threats in web applications.[1] According to research by the European Union Cyber Security Agency (ENISA), two-thirds of web application attacks include SQL injection attacks.[2] Most websites and online applications are susceptible to this type of attack [1]. In the results of the survey of the European information technology security company Balabit about the best hacking methods obtained from the participants of the black hat events, the injection attack has won the third place.[3] Also, according to Freepik, hackers were able to steal 8.3 million records (information registered by users on this site) through SQL injection.[4] Even the SQL injection attack for the database of RFID[5] systems is considered a serious threat [2]. Therefore, with the expansion of web applications, the emergence of intelligent networks, Internet of Things (IoT) and cyber physical systems (CPS), the number and intensity of SQL injection attacks will be added and increased, and as a result, the need to pay attention to the security of databases in the digital world. makes it very important.

SQL injection attack, also known as SQLIA[6], is one of the most dangerous and risky cyber attacks where attackers can perform unauthorized actions on the victim's database. For example, an attacker could inject SQL database code into a form that expects a simple username. If the form input is not secure, it will lead to the execution of SQL code, which will be possible depending on the security level of the website, the level of vulnerability of the database and the depth of penetration of the attacker from data leakage to information theft and other risks. Therefore, a hacker or an unauthorized user can impersonate a privileged user, and take control of the database. For a detailed understanding, let's do a simple comparison between normal and malicious SQL commands (Figure 1).



Customer ID: 123456789
Password: aPassword

```
Select * from Customers where
Customer ID   = 123456789 and password=
'aPassword'
```

Customer ID: 123456789 OR 1=1
Password: */--

```
Select * from Customers where
Customer ID =  123456789 OR 1 = 1  and
password= '*/--'
```

Figure 1. Comparison between normal and malicious commands

[1] https://owasp.org/www-project-top-ten/

[2] https://patchstack.com/articles/website-hacking-statistics/

[3] https://www.vanillaplus.com/2016/02/15/15415-top-10-hacking-methods-revealed-in-survey-by-balabit/

[4] https://www.bleepingcomputer.com/news/security/freepik-data-breach-hackers-stole-83m-records-via-sql-injection/

[5] Radio Frequency Identification

[6] SQL Injection Attack

For example, in a normal SQL command, the goal is to check a customer that matches the entered Customer Id through the list of bank customers, and after finding, return the records of that customer. Now, if the user enters the customer ID, say 123456789, in the web page form.

SELECT * FROM customers WHERE Customer Id= 123456789 and Password= 'a password';

As a result, the resulting SQL query outputs the record for the customer with its own Customer Id, exactly what the developer who wrote the API expected to happen. But malicious SQL Injection commands are designed with the aim of penetrating the system and obtaining sensitive information or destroying the database. These types of commands can be used by exploiting security vulnerabilities in applications that communicate with the database. In the same example, an attacker can enter a conditional logic in the input field next to the customer ID:

SELECT * FROM customers WHERE Customer Id= 123456789 OR 1=1 and Password= ' */--' ;

Now this command looks for the customer id or equality test 1 equals 1 and since the logical expression is true for all, as a result the database returns all the data in the customers table to the attacker who executes the query, then By using the `--` clause, the `password' section is completely disabled and ignored. And just as easily SQLIA works by targeting a vulnerable API[1]. API in this case is the software interface through which the server receives and responds to requests. It is an imperative for software development engineers to thoroughly test their programs for vulnerabilities and to continuously and intelligently take countermeasures to prevent hacker attacks and cybercrimes.

3 PAST EFFORTS IN SQL INJECTION ATTACK DETECTION AND PREVENTION

Machine Learning (ML) refers to the ability of a computer system to learn based on imitation of human learning methods with training and testing on data, and Deep Learning (DL) is a sub-branch of machine learning based on Artificial Neural Networks(ANN). Numerous researches have been conducted to counter this threat and various artificial intelligence techniques have been proposed to detect SQLIA using machine learning [3]. SQL injection attack can be detected using machine learning[4]. The machine learning approach has been proven to be suitable not only for preventing existing known attacks, but also for preventing future unknown attacks. An injected SQL statement can be easily detected, provided that an appropriate classifier is used and up-to-date data is used for training[5]. Past data is useful for identifying attack patterns, understanding detected traffic, and even predicting future attacks before they occur [6]. Another advantage of machine learning algorithms is that they cover a wider range of SQL queries. Also, the detection accuracy increases and the false positive rate decreases [7]. In the following, we will examine the efforts of respected researchers to deal with SQL injection attacks.

[1] Application Programming Interface

Due to the fact that the injection attack is included in the Jirga of historical crimes, many methods have been proposed and used by researchers and engineers from the past to the present day. AMNESIA is one of the tools that uses pattern matching mechanism [8]. Dynamic Taint is a similar method based on pattern matching [9]. Also, SQLrand is a parsing method to prevent SQL injection[10]. And SQL Guard is another parsing method[11]. Anamika Joshi and Geetha V have proposed a method to detect SQL injection attack based on Naïve Bayes machine learning algorithm along with role-based access control mechanism, the maximum accuracy of the algorithm reaches 93.3%[12]. Nekkalapudi and Polinati et al.; presented a SQL injection detection model trained using the dataset and logistic regression technique. It allows users to run queries, then a file program applies it to a regression model to predict whether the query is normal or abnormal. The classification accuracy of this algorithm is 96.667% [13]. Krishnan et al.; They use five naive classifiers, passive aggressive classifiers, SVM, CNN and logistic regression to classify traffic into normal order or malicious order. Passive aggressive accuracy is 79%, SVM accuracy is 79%, and logistic regression accuracy is 92%. The accuracy percentage of the Naive Bayes classifier model is also 95%. The CNN algorithm tests and corrects many parameters at the same time, and with 97% accuracy, it is a better option for dealing with the issue of SQL Injection classification[14]. In the research of A. Alam and M. Tahreen and colleagues; A strategy for identifying SQL injection threats using machine learning algorithms such as the basic Naive Bayes model with an accuracy of 97.8 has been presented, which prevents malicious database queries[15]. In the studies of T. Pattewar and H. Patil et al.; Using the Naive Bayes classification algorithm, it gives findings that are 92.8% accurate, and it is thought that the classification of SQL injection threats using machine learning will result in high accuracy and low error rate results [16]. In the research of N. Gandhi and J. Patel, et al.; A combined CNN-BiLSTM technology is presented to detect SQL injection attacks. The combined CNN-BiLSTM machine learning model reduces the number of SQL injection attacks and is introduced as the best classifier for detecting malicious commands with an accuracy of 98% [17]. In K. Zhang's research, the main goal is to create a machine learning classifier for finding SQLI errors in PHP code that can identify vulnerable SQLI files. CNN model with 95.4% accuracy is considered a suitable method [18]. We reviewed some of the important contributions of countering SQL injection attacks by machine learning and deep learning techniques. Although there are many valuable works in this direction, they are devoted to limited aspects of the problem. For example, in past studies, the proposed solutions have not been successful in covering all the different types of SQLIA. Or sometimes the accuracy of detection has been addressed, but the speed of detection or in other words, real-time and preventive detection has not been investigated, while attackers usually try to reach their target in the shortest possible time. It is also worth noting that in order to bypass security, professional hackers and smart attackers sometimes execute multi-vector attacks against a target website. DDoS attacks[1], DNS hijacking[2], social engineering and other methods are used as a distraction to execute massive SQL injection attacks. This requires careful studies of all current and future security threats in databases and web applications.

4 PROVIDING A NEW APPROACH BASED ON NLP AND LARGE LANGUAGE MODELS (LLMS)

Large language models (LLMs) are advanced artificial intelligence models that use deep learning techniques such as Transformer and are trained on large sets of textual data. Transformer architecture, which was introduced in 2017 [19], has been one of the most successful models in natural language processing. For example, BERT[3] is one of the applications of Transformer, which is used in Many language processing tasks have improved the accuracies of previous models. Research shows that Transformer-based natural language processing (NLP) models have achieved advanced performance for name entity recognition, relation extraction, sentence similarity, natural language inference, and question answering

[1] Distributed Denial of Service

[2] Domain Name System Hijacking

[3] Bidirectional Encoder Representations from Transformers

[20]. Also, in 2020, a new model named Reformer has been introduced by Google, which is much more efficient than Transformer in terms of speed and memory.[1] Large language models (LLMs), such as GPT-3, PALM, LLaMA, and GPT-4, and products built on them, such as ChatGPT, have recently attracted a lot of attention [21, 22]. Many attentions from journalists [23, 24, 25], policy makers [26, 27, 28] and researchers in many fields [29, 30, 31] are expanding. Large language models (LLMs) consistently show significant performance on various tasks [32, 33, 34]. The capacity of artificial intelligence to process huge amounts of data in real time, learn from it and predict threats can play a transformative role in proactively dealing with cyber threats [35].

In this section, we provide new insights into the applications of NLP to detect and prevent SQL injection attacks. Natural language processing (NLP) and large language models (LLMs) as a new and powerful approach to identify and prevent SQL injection attacks can be of interest to artificial intelligence researchers and software engineers. LLMs, which are a highly advanced machine learning model, have the ability to process large data, examine various patterns, and detect complex relationships. In the field of identifying SQL injection attacks, they can help identify SQL injection attacks by analyzing queries and identifying suspicious patterns. LLMs can analyze the structure of SQL queries using natural language processing techniques and help effectively prevent SQL injection attacks if suspicious patterns are identified. Therefore, using LLMs as a new method to detect SQL injection attacks can help software development engineers to effectively use SQL injection in their systems to monitor and fix cyber attacks. By using large language models, SQL injection attacks can be detected automatically and without the need for human intervention, and if the attack is discovered, it can be filtered. To monitor and remediate SQL injection threats in real time, LLMs can be trained on a dataset of known SQL injection attacks as well as on normal web traffic. This model can then be deployed in a real-time monitoring system, where it can analyze incoming web traffic and detect SQL injection attacks in real time. To explain the new approach, the practical way to implement LLMs to detect SQL injection in real time is to integrate it into a Web Application Firewall (WAF) (Figure 2). A web application firewall (WAF) is a security tool that can filter incoming web traffic to prevent attacks. By integrating LLM into a WAF, the model can analyze incoming SQL queries and determine whether they are normal or malicious and be used to detect SQL injection in real-time. By deploying LLMs in a real-time monitoring system, software development engineers can detect SQL injection attacks as they occur and take action to prevent dangerous damage to critical applications and data. One of the most important and significant potential benefits of our proposed approach can include improving the accuracy and speed of detection, facilitating threat monitoring and automatic attack prevention. In addition, our study sheds light on possible challenges and limitations with a brief hint at future research and future research directions. Things like variety and quality of training data, model size and computational efficiency, fine-tuning and transfer of learning, compression of large language models, transparency and interpretability, universality and generalizability, reliability and evaluation criteria, as well as ethical and legal considerations. They can be among the most important conditions to contribute to a safer cyber environment under scrutiny and research.

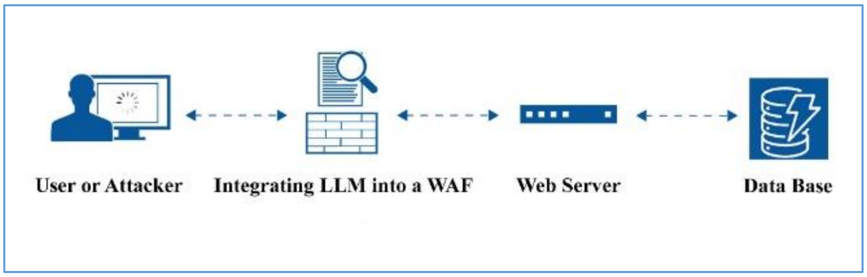


Figure 2. A new approach to integrating LLM into WAF

[1] https://ai.googleblog.com/2020/01/reformer-efficient-transformer.html

5 CONCLUSION

This research has revealed an opening and a pioneering move specifically in the realm of using large language models to improve web security and improve database security. The new approach of integrating LLM in WAF; It will be an advanced technique against complex threats and will provide a strong foundation for future research and practical applications that will contribute to the goal of realizing sustainable cyber security in the age of artificial intelligence. The insights obtained from this study can be useful in achieving this goal and thus provide a promising ground for future studies and discoveries. At the end, we end the article with the thought that risk reduction depends on increasing awareness and preparation, and software development engineers are required to review the past methods and use the approach in designing mechanisms related to data and system security. New technologies are aimed at monitoring threats and reducing risks.

6 REFERENCES

[1] K. Ross, SQL injection detection using machine learning techniques and multiple data sources, Department of Computer Science, Master’s Project, San Jos´e State University, 2018.

[2] Q. Zhang and X. Wang, ``SQL injections through back-end of RFID system," in *Proc. Int. Symp. Comput. Netw. Multimedia Technol.*, Jan. 2009, pp. 1_4.

[3] Yan, R.; Xiao, X.; Hu, G.; Peng, S.; Jiang, Y. New deep learning method to detect code injection attacks on hybrid applications. J. Syst. Softw. **2018**, 137, 67–77.

[4] AL-Maliki, M., Jasim, M. Comparison study for NLP using machine learning techniques to detecting SQL injection vulnerabilities. International Journal of Nonlinear Analysis and Applications, 2023; (): -. doi: 10.22075/ijnaa.2022.28365.4098

[5] C. Bockermann, M. Apel, and M. Meier, “Learning SQL for databases intrusion detection using context-sensitive modeling (extended abstract)" Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2009, vol. 5587 LNCS, pp. 196–205.

[6] Vähäkainu, P.; Lehto, M. Artificial intelligence in the cyber security environment. In Proceedings of the 14th International Conference on CyberWarfare and Security, ICCWS 2019, Stellenbosch, South Africa, 28 February–1 March 2019; pp. 431–440.

[7] A. Joshi and V. Geetha, "SQL Injection detection using machine learning," 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Kanyakumari, India, 2014, pp. 1111-1115, doi: 10.1109/ICCICCT.2014.6993127.

[8] W.G.Halfond and Aorso, “ AMNESIA Analysis and Monitoring for Neutralizing SQL-Injection Attacks,” Proc. IEEE and ACM International Conference on Automatic Software Engineering (ASE 2005), Long Beach, CA, USA, Nov 2005.

[9] V.Haldar, D.Chandra, and M.Franz, "Dynamic Taint Propagation for Java," Proc. 2 1s t Annual Computer Security Applications Conference, Dec 2005.

[10] S.W.Boyd and AD.Keromytis, "SQLrand: Preventing SQL Injection Attacks," Proc. the 2nd Applied Cryptography and Network Security (ACNS) Conference, pp. 292-302, Jun 2004.

[11] G.T.Buehrer, RW.Weide, and P.AG.Sivilotti, "Using Parse Tree Validation to Prevent SQL Injection Attacks," International Workshop on Software Engineering and Middleware (SEM), 2005.

[12] A. Joshi and V. Geetha, "SQL Injection detection using machine learning," 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Kanyakumari, India, 2014, pp. 1111-1115, doi: 10.1109/ICCICCT.2014.6993127.

[13] V.B. Polinati, S.C. Nekkalapudi, N.S. Sanjana and R.V. Bhupathiraju, SQL injection prediction web app using different machine learning algorithms Vinod, J. Eng. Sci. 13 (2022), no. 4.

[14] S.A. Krishnan, A.N. Sabu, P.P. Sajan and A.L. Sreedeep, SQL injection detection using machine learning, Rev. Gest˜ao Inova¸c˜ao e Tecnol. 11 (2021), no. 3, 300–310.

[15] A. Alam, M. Tahreen, M.M. Alam, S.A. Mohammad and S. Rana, SCAMM: detection and prevention of SQL injection attacks using a machine learning approach, PhD diss. Brac University, 2021.

[16] T. Pattewar, H. Patil, H. Patil, N. Patil, M. Taneja and T. Wadile, Detection of SQL injection using machine learning: a survey, Int. Res. J. Eng. Technol. 6 (2019), no. 11, 239–246.

[17] N. Gandhi, J. Patel, R. Sisodiya, N. Doshi and S. Mishra, A CNN-BiLSTM based approach for detection of SQL injection attacks, Proc. 2nd IEEE Int. Conf. Comput. Intell. Knowl. Econ. ICCIKE, 2021, pp. 378–383.

[18] K. Zhang, A machine learning based approach to identify SQL injection vulnerabilities, 34th IEEE/ACM Int. Conf. Automated Software Engin., 2019, pp. 1286—1288.

[19] A. Vaswani et al., "Attention is all you need," in Advances in Neural Information Processing Systems, 2017, pp. 5998--6008.

[20] Yang X, Chen A, PourNejatian N, Shin HC, Smith KE, Parisien C, Compas C, Martin C, Costa AB, Flores MG, Zhang Y, Magoc T, Harle CA, Lipori G, Mitchell DA, Hogan WR, Shenkman EA, Bian J, Wu Y. A large language model for electronic health records. NPJ Digit Med. 2022 Dec 26;5(1):194. doi: 10.1038/s41746-022-00742-2. PMID: 36572766; PMCID: PMC9792464.

[21] Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., Barham, P., Chung, H. W., Sutton, C., Gehrmann, S., et al. PaLM: Scaling language modeling with pathways. arXiv preprint 2204.02311, 2022.

[22] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Roziere,` B., Goyal, N., Hambro, E., Azhar, F., et al. LLaMA: Open and efficient foundation language models. arXiv preprint 2302.13971, 2023.

[23] Klein, E. This changes everything. New York Times, 2023. URL https://www.nytimes.com/2023/03/ 12/opinion/chatbots-artificial-intel ligence-future-weirdness.html.

[24] Perrigo, B. The new AI-powered Bing is threatening users. that's no laughing matter. Time, 2023. URL https: //time.com/6256529/bing-openai-chatg pt-danger-alignment/.

[25] Oliver, J. Last week tonight with John Oliver: Feb 26, 2023. URL https://www.hbo.com/last-week-to night-with-john-oliver/season-10/2-f ebruary-26-2022.

[26] J, P. and C, D. ChatGPT and large language models: what's the risk? National Cyber Security Center, 2023. URL August 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.acl-long.143. URL https: //aclanthology.org/2021.acl-long.143.

[27] Bartz, D. As ChatGPT's popularity explodes, U.S. law-makers take an interest. Reuters, 2023. URL https: //www.reuters.com/technology/chatgpt s-popularity-explodes-us-lawmakers-t ake-an-interest-2023-02-13/.

[28] Li, K., Hopkins, A. K., Bau, D., Viegas,´ F., Pfister, H., and Wattenberg, M. Emergent world representations: Exploring a sequence model trained on a synthetic task. In The Eleventh International Conference on Learning Representations, 2023. URL https://openreview.net/forum?id=DeG07 TcZvT.

[29] Chan, L., Garriga-Alonso, A., Goldowsky-Dill, N., Green-blatt, R., Nitishinskaya, J., Radhakrishnan, A., Shlegeris, B., and Thomas, N. Causal scrubbing: a method for rigorously testing interpretability hypotheses. Alignment Forum, 2022. URL https://www.alignmentfor um.org/posts/JvZhhzycHu2Yd57RN/causa l-scrubbing-a-method-for-rigorously-testing.

[30] Lund, B. D. and Wang, T. Chatting about ChatGPT: how may AI and GPT impact academia and libraries? Library Hi Tech News, 2023. doi: https://doi.org/10.1108/LHTN -01-2023-0009.

[31] Choi, J. H., Hickman, K. E., Monahan, A., and Schwarcz, D. ChatGPT goes to law school. Minnesota Legal Studies Research Paper, 23(03), 2023. doi: http://dx.doi.org/10. 2139/ssrn.4335905.

[32] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, Yifan Du, Chen Yang, Yushuo Chen, Zhipeng Chen, Jinhao Jiang, Ruiyang Ren, Yifan Li, Xinyu Tang, Zikang Liu, Peiyu Liu, Jian-Yun Nie, and Ji-Rong Wen. A survey of large language models. *CoRR*, abs/2303.18223, 2023.

[33] Jie Huang and Kevin Chen-Chuan Chang. Towards reasoning in large language models: A survey. In Anna Rogers, Jordan L. Boyd-Graber, and Naoaki Okazaki, editors, *Findings of the Association for Computational Linguistics: ACL 2023*, *Toronto, Canada, July 9-14, 2023*, pages 1049–1065. Association for Com-putational Linguistics, 2023.

[34] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Kaijie Zhu, Hao Chen, Linyi Yang, Xi- aoyuan Yi, Cunxiang Wang, Yidong Wang, Wei Ye, Yue Zhang, Yi Chang, Philip S. Yu, Qiang Yang, and Xing Xie. A survey on evaluation of large language models. *CoRR*, abs/2307.03109, 2023.

[35] KEREOPA-YORKE, Benjamin. Building Resilient SMEs: Harnessing Large Language Models for Cyber Security in Australia. *arXiv preprint arXiv:2306.02612*, 2023.