Network Security vulnerabilities and Prevention Techniques of SQL Injection Detection

Abdalla Adel Abdalla Hadabi
Computer Science Department
Al–Neelain university
Khartoum, Sudan
Abdoadel30@gmail.com

Eltyeb Elsamani AbdElgabar
Computer Science Department
Al–Neelain university
Khartoum, Sudan
tayebsamani@gmail.com

Ali Abdallah Abakar Adam
Accounting Information
Systems Department
Al–Neelain university
Khartoum, Sudan
Misterali88@gmail.com

Mohamed Eltayeb Ahmed
Human Resource Department
Ministry of labour & Administrative Reform
Mhmdtayeb.cn@gmail.com

Alsammany Alnour Abdalrhman Mohammed
Center of Information
Al–Neelain university
Khartoum, Sudan
smmany@gmail.com

المستخلص

يعد حقن الاستعلام الهيكلى(SQL Injection Attack) هجومًا شائعًا على الويب و يمثل تحديًا لأمن تطبيقات الويب ؛ تتسبب هذا النوع من الهجمات في خسائر مالية في جميع أنحاء العالم بالإضافة إلى كونها تنتهك خصوصية بيانات المستخدم. أصبح اكتشاف حقن الاستعلام الهيكلى موضوعًا يشغل الباحثين مؤخرًا. جذبت كيفية الدفاع عن هجمات حقن الاستعلام الهيكلى بشكل فعال انتباه المتخصصين والباحثين في أمن الويب. الهدف من هذه الورقة العلمية هو مراجعة أسباب هجوم حقن الاستعلام الهيكلى

واستراتيجيات الوقاية الحديثة. قمنا بمسح كل الاعمال المنشورة في السنوات الأخيرة التى تتعلق بهذا الهجوم. تم استخدام أشجار القرار(Decision trees) و (SVM) على نطاق واسع في السنوات الماضية. أيضًا ، يمكننا أن نلاحظ أن خوارزميات التعلم العميق(Deep learning algorithms) قد تم استخدامها بشكل متزايد في حل هجمات حقن الاستعلام الهيكلى .يعد استخدام تقنيات التعلم الآلي لاكتشاف هجوم حقن الاستعلام الهيكلى مفيدًا للغاية وواعدًا لتأمين التطبيقات.

الكلمات المفتاحية: حقن الاستعلام الهيكلى، تعلم الالة، مطابقة الانماط، امن الشبكات.

Abstract

SQL injection is a popular web attack and has been a challenging matter for network security; SQL causes financial losses worldwide as well as user data offensive. SQL injection detection has become a hot topic recently. How to defend SQL injection attacks effectively has drawn the attention of web security professionals and researchers. The objective of this paper was to review SQL injection attack causes and modern prevention strategies. We surveyed the literature published in recent years. Decision trees and SVM have been widely used in the last years. Also, we can notice that deep learning algorithms have been increasingly used in solving SQL attacks. Using machine learning techniques for SQL attack detection is very useful and promising to secure applications.

Keywords: SQL injection attack, Machine Learning, pattern matching, Network security.

1. Introduction

Machine learning is broadly studied in many security problems and applications due to its technical advances in recent years. Machine learning models proved many achievements, in dealing with different and complex tasks, and demonstrates abilities close to humans or even exceeds humans [1]. Machine learning models, has transformed from laboratory innovation to real world applications in numerous domains. Such as fraud detection, industrial protocol

security analysis[2], ML for internet of things security[3] [4] [5]. Cyber security [6] [7]. SQL injection [8].

However, these current methods still suffer from security difficulties in next-generation devices and systems due to the shifting behaviors of security threats and zero-day attacks. Machine learning techniques can provide a solution to address the changing nature of security attacks by the analysis of data to recognize unseen patterns of data [9]. Machine learning (ML) methods found to be effective for identify or better prevent SQL attack This paper is organized as follows: section one introduction, section two discusses causes and prevention of SQL attack, SQL attack is illustrated in section 3, machine learning models for network security in section four, followed by section five results and discussion, last is the conclusion.

2. Causes and Prevention

Due to the heterogeneity and the diversity of attack methods and the variability of attack modes methods, attack method changes frequently with the development of website technology, SQL injection detection is still a stimulating problem. Developers usually create SQL statements by concatenating string which is submitted by users from web page. Due to the wide variety of SQL in different languages, there are too many encoding methods for constructing SQL statements, so there is a threat to be attacked anywhere via constructing SQL statements [10]. SQL is the placement of malicious code in SQL statements via web page input, is considered one of the most dangerous type of vulnerabilities that web applications are suffer from [11]. The SQL injection attack is one of the most critical for cloud (SaaS) vulnerabilities that allows attackers to disrupt the availability, confidentiality and integrity of user data [12]. Web application often has bugs which results as the security threat to the institution [13]. SQL attack comes through the back-end of the database [14].

How to defense SQL injection attack effectively becomes the focus and frontier of web security nowadays. It is the programmer's duty to write intelligent code that prevents such attacks. [15] Presented a hybrid approach based on the Adaptive Intelligent Intrusion Detector Agent. [16] proposed classifier uses

combination of Naïve Bayes machine learning algorithm and Role Based Access Control mechanism for detection.

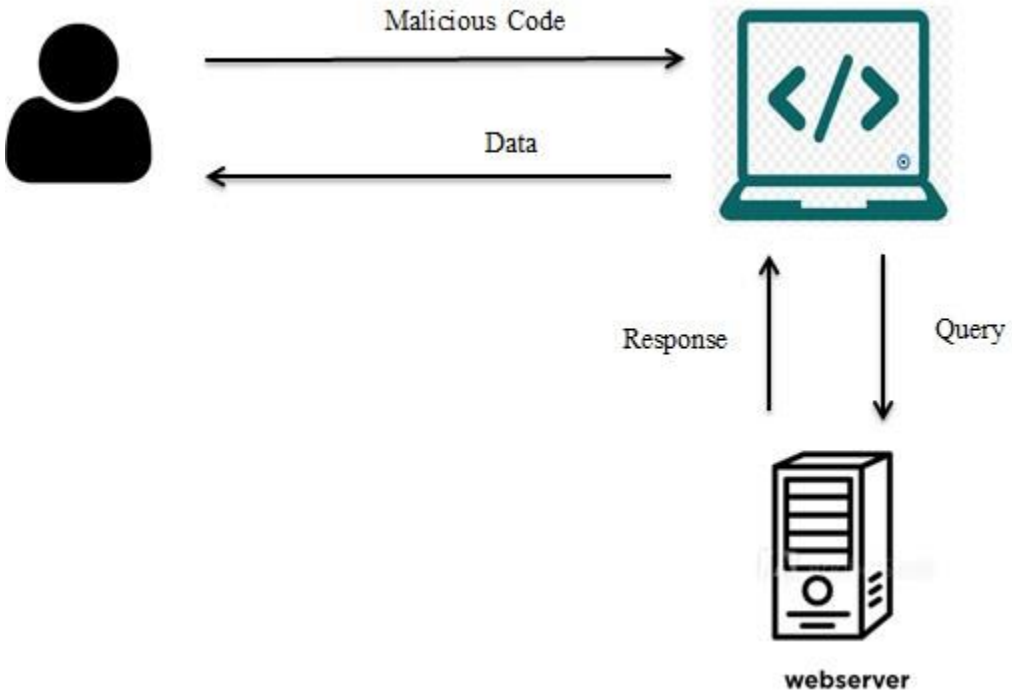

Figure 1. SQL Injection attack.

3. SQL Injection attack

SQL Injection attack is the first of the top 10 security threats announced by the OWASP. Meanwhile, SQL injection considered to be one of the most critical threats to the Web application, an attacker could get access to classified database that contain sensitive data. Primarily, SQL injection attack uses the embedded SQL query along with the user input in such a way that tricked the database into considering it is an SQL code [8]. Several approaches been introduced for identification and prevention of structured query language injection attack (SQLIA). Such as: pattern matching, parsing, and machine learning. Machine learning (ML) approach was found to be effective for SQLIA mitigation [17].

The authors in [10] introduced an adaptive deep forest-based method to detect the complex SQL injection attacks. They concluded that the method, which they

proposed, had a better performance compared to classical machine learning methods and deep learning methods. ML-Driven approach based on machine learning and an evolutionary algorithm to detect holes automatically in WAFs (Web Application Firewalls). ML-Driven picks attacks that exhibit patterns and incrementally learns attack patterns from previously generated attacks according to their testing results [18]. Manjunatha and Kempanna proposed system with a machine-learning algorithm to provide accurate evaluation that discovers vulnerabilities in web applications. Their analysis for vulnerabilities that directly deal with weak or absent of input validation [19]. Authors in [20] introduced machine learning model to detect malicious strings by classifying the input strings given to the web applications, they also concluded that classical methods for preventing SQL attack are not effective. Li et al. Proposed an SQL injection attack detection method based on memory. It can automatically learn the representation of data, and has a strong advantage to deal with high-dimensional data [21]. Raul et al. offered a solution of classifying SQL queries based on the features of the initial query string. The output calculation used a similarity metric. The probabilistic nature of the learned model gives us the advantage to adapt to new threats even in operation [22]. Peng et al in [23] proposed SQL injection detection method based on neural network. They first acquired the history of users URL access log data from the Internet Service Provider (ISP).

Table 1. The algorithms used to prevent SQL attack.

Author	Algorithm
[18],[12], [6], [21],[15],[24]	Decision Tree
[22],[12],[14],[15],[25],[24]	SVM
[10],[10]	Deep Forest
[20],[25]	Random Forest
[19],[16],[25]	Naive Baye
[12],[14]	Logistic Regression
[23],[11],[15],[24]	ANN
[12],[26]	KNN
[12]	Deep ANN

From the table above we can notice that decision tree, SVM have been widely used in the last three years. Also, we can notice that deep learning algorithms have been increasingly used in solving SQL attack. Machine learning is a promising technique to address SQL attack and network security problems.

4. Machine Learning Models for Network Security

Working with methods for reasoning under uncertainty is now one of the most interesting areas of machine learning. Machine learning has been used for several decades to tackle a broad range of problems in many fields of applications [27]. Concurrently, dramatic enhancements in machine learning models have enabled exceptional analytical capabilities. Machine learning mechanisms with the potential to improve prediction and existing threat patterns. As well as those that can learn about new attacks yet to be seen by the classifier [5]. Machine learning models can deal with the ever-increasing amount of data in addition to heterogeneous types of data [2].

5. Results and Discussion

The objective of this paper was to review SQL injection attack causes and modern prevention strategies. We surveyed the literature published by researchers in the recent years. We surveyed the literature available in the recent years. Decision tree and SVM have been widely used in the last years. Also, we can notice that deep learning algorithms have been increasingly used in solving SQL attack.

6. Conclusion

In conclusion, this paper employs machine-learning models against security attacks. The models which we built was evaluated using four performance metrics, using the four metrics allows us to see a bigger picture of our model and how it is expected to behave in different scenarios. This model detects SQL injection only; researchers should pay attention to other types of cyber-attacks so they could be addressed by machine learning algorithms.

References

[1] M. Xue, C. Yuan, and H. Wu, “Machine Learning Security : Threats , Countermeasures , and Evaluations,” vol. 8, 2020.

[2] J. Men, Z. Lv, X. Zhou, Z. Han, H. Xian, and Y. N. Song, “Machine learning methods for industrial protocol security analysis: Issues, taxonomy, and directions,” IEEE Access, vol. 8, pp. 83842–83857, 2020, doi:

10.1109/ACCESS.2020.2976745.

[3] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani,“A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security,” IEEE Commun. Surv. Tutorials, vol. 22, no. 3, pp. 1646–

1685, 2020, doi: 10.1109/COMST.2020.2988293.

[4] M. Bagaa, T. Taleb, J. B. Bernabe, and A. Skarmeta, “A Machine Learning

Security Framework for Iot Systems,” IEEE Access, vol. 8, pp. 114066–

114077, 2020, doi: 10.1109/ACCESS.2020.2996214.

[5] F. Liang, W. G. Hatcher, W. Liao, W. Gao, and W. Yu, “Machine Learning for Security and the Internet of Things: The Good, the Bad, and the Ugly,” IEEE Access, vol. 7, pp. 158126–158147, 2019, doi:

10.1109/ACCESS.2019.2948912.

[6] S. Intrusion and D. Model, “SS symmetry IntruDTree : A Machine Learning

Based Cyber,” no. May 2017, pp. 1–15, 2020, doi: 10.3390/sym12050754.

[7] R. A. Calix, S. B. Singh, T. Chen, D. Zhang, and M. Tu, “Cyber security tool kit (cybersectk): A python library for machine learning and cyber security,” Inf., vol. 11, no. 2, 2020, doi: 10.3390/info11020100.

[8] T. S. Gunawan, M. K. Lim, M. Kartiwi, N. A. Malik, and N. Ismail, “Penetration Testing using Kali Linux : SQL Injection , XSS , Wordpres , and WPA2 Attacks,” vol. 12, no. 2, pp. 729–737, 2018, doi:

10.11591/ijeecs.v12.i2.pp729-737.

[9] J. H. Park, “Symmetry-adapted machine learning for information security,”

Symmetry (Basel)., vol. 12, no. 6, pp. 1–4, 2020, doi: 10.3390/sym12061044.

[10] Q. I. Li, W. Li, and J. Wang, "A SQL Injection Detection Method Based on

Adaptive Deep Forest," pp. 145385–145394, 2019, doi:

10.1109/ACCESS.2019.2944951.

[11] K. Zhang, "A machine learning based approach to identify SQL injection vulnerabilities," Proc. - 2019 34th IEEE/ACM Int. Conf. Autom. Softw. Eng. ASE 2019, pp. 1286–1288, 2019, doi: 10.1109/ASE.2019.00164.

[12] D. Tripathy, R. Gohil, and T. Halabi, "Detecting SQL Injection Attacks in Cloud SaaS using Machine Learning," Proc. - 2020 IEEE 6th Intl Conf. Big Data Secur. Cloud, BigDataSecurity 2020, 2020 IEEE Intl Conf. High Perform. Smart Comput. HPSC 2020 2020 IEEE Intl Conf. Intell. Data Secur. IDS 2020, pp. 145–150, 2020, doi: 10.1109/BigDataSecurity-HPSC- IDS49724.2020.00035.

[13] R. Tommy, G. Sundeep, and H. Jose, "Automatic Detection and Correction of Vulnerabilities using Machine Learning," Int. Conf. Curr. Trends Comput. Electr. Electron. Commun. CTCEEC 2017, pp. 1062–1065, 2018, doi:

10.1109/CTCEEC.2017.8454995.

[14] S. O. Uwagbole, W. J. Buchanan, and L. Fan, "An applied pattern-driven corpus to predictive analytics in mitigating SQL injection attack," Proc. -2017 7th Int. Conf. Emerg. Secur. Technol. EST 2017, pp. 12–17, 2017, doi:10.1109/EST.2017.8090392.

[15] C. Pinzón, J. F. De Paz, J. Bajo, Á. Herrero, and E. Corchado, "AIIDA-SQL: An Adaptive Intelligent Intrusion Detector Agent for detecting SQL injection attacks," 2010 10th Int. Conf. Hybrid Intell. Syst. HIS 2010, pp. 73–78, 2010, doi: 10.1109/HIS.2010.5600026.

[16] A. Joshi and V. Geetha, "SQL Injection detection using machine learning,"

2014 Int. Conf. Control. Instrumentation, Commun. Comput. Technol. ICCICCT 2014, no. 2, pp. 1111–1115, 2014, doi:

10.1109/ICCICCT.2014.6993127.

[17] E. T. Jide and A. Sunday, "SQL Injection Attacks Predictive Analytics Using

Supervised Machine Learning Techniques," vol. 9, no. 4, pp. 139–149, 2020.

[18] D. Appelt, C. D. Nguyen, A. Panichella, and L. C. Briand, "A Machine- Learning-Driven Evolutionary Approach for Testing Web Application Firewalls," IEEE Trans. Reliab., vol. PP, pp. 1–25, 2018, doi:10.1109/TR.2018.2805763.

[19] M. Kempanna, "Web Security Aware by using Naive Baye ' s ML Technique," no. 4, pp. 3222–3230, 2020, doi: 10.35940/ijitee.D1325.029420.

[20] M. Venkata, S. Soma, and R. K. Megalingam, "Applying and Evaluating Supervised Learning Classification Techniques to Detect Attacks on Web Applications," no. 10, pp. 2222–2225, 2019, doi: 10.35940/ijitee.J9434.0881019.

[21] Q. Li, F. Wang, J. Wang, and W. Li, "LSTM-Based SQL Injection Detection Method for Intelligent Transportation System," vol. 68, no. 5, pp. 4182–4191, 2019, doi: 10.1109/TVT.2019.2893675.

[22] P. R. Mcwhirter, K. Kifayat, Q. Shi, and B. Askwith, "Journal of Information Security and Applications SQL Injection Attack classification through the feature extraction of SQL query strings using a Gap-Weighted String Subsequence Kernel," J. Inf. Secur. Appl., vol. 40, pp. 199–216, 2018, doi:10.1016/j.jisa.2018.04.001.

[23] P. Tang, W. Qiu, Z. Huang, H. Lian, and G. Liu, "Knowledge-Based SystemsDetection of SQL injection based on artificial neural network ☆," Knowledge-Based Syst., vol. 190, p. 105528, 2020, doi:10.1016/j.knosys.2020.105528.

[24] K. Kamtuo and C. Soomlek, "Server-Side Scripting," 2016.

[25] H. Gao, J. Zhu, L. Liu, J. Xu, Y. Wu, and A. Liu, "Detecting SQL injection attacks using grammar pattern recognition and access behavior mining," Proc. - IEEE Int. Conf. Energy Internet, ICEI 2019, pp. 493–498, 2019, doi:10.1109/ICEI.2019.00093.

[26] S. Ponmaniraj, R. Rashmi, and M. V. Anand, "IDS based network security architecture with TCP/IP parameters using machine learning," 2018 Int.Conf. Comput. Power Commun. Technol. GUCON 2018, no. 1, pp. 111–114,2019, doi: 10.1109/GUCON.2018.8674974.

[27] W. J. Murdoch, C. Singh, K. Kumbier, R. Abbasi-asl, and B. Yu, "Definitions , methods , and applications in interpretable machine learning," vol. 116, no. 44, 2019, doi: 10.1073/pnas.1900654116.