

SQL Injection Attack: A Systematic Literature Review on Detection, Prevention and Classification with Machine Learning Approach

¹Morufu Olalere ²Raji Abdullahi Egigogo, ³Umar Rukayya, ⁴Shafi'i Muhammad Abdulhamid ⁵Yisa Victor Legbo, ⁶Rasheed Gbenga Jimoh

^{1,2,3,4,5}Department of Cyber Security Science, School of Information and Communication Technology, Federal University of Technology, Minna, Nigeria.

⁵Department of Computer Science, Faculty of Information and Communication Technology, University of Ilorin, Nigeria.

¹lerejide@futminna.edu.ng, ²raji.pg610868@st.futminna.edu.ng, ³umarrukayya1@gmail.com, ⁴shafii.abdulhamid@futminna.edu.ng, ⁵yisavictor@futminna.edu.ng, ⁶jimoh_resheed@yahoo.com

Abstract -When it comes to web application, confidentiality, availability and integrity of individuals and organizations data are not assured. Open Web Application Security Project (OWASP) has identified SQL injection attacks as common threat to the web application. Consequently, many researchers have proposed different approaches for either detection, prevention or classification/categorization of SQL injection attack. Machine learning approach is one of the approaches existing in the literature, though not very much research outputs with this approach are available in the literature. This implies that, future researchers can still apply machine learning approach in addressing SQL injection attack problem. For this reason, this study presents a systematic literature review on SQL injection attack detection, prevention and classification based on machine learning approach. In order to obtain SQL injection attack related articles, various search engines and scholar databases were visited. The authors review analysis revealed that most of the proposed machine leaning approaches were proposed to only detect whether an application is vulnerable to SQL injection attack or not. Very few were proposed to prevent and classify the injection based on the attack type. It is our hope that this review will provide a theoretical background for future research and enable future researches to identify how and where machine learning approaches have been used to address SQL injection attack.

Index Terms - Categorization, Detection, Prevention, SQL injection Attack, Machine learning.

I. INTRODUCTION

Technology has burgeoned to the degree that individuals, groups and organizations keep save of vital and confidential information such as date of birth, password, username, credit card information, email address, mobile phone, student and staff identification number, last and first name, staff identification number and work address number that relates to individuals, groups organizations and partners (customers) on the World Wide Web (WWW). When a particular database is attacked, information in the database can be revealed to illicit users, even modified by the hackers or totally moped out from the database through various web

application vulnerabilities such Denial of Service (DoS), Cross Script (CSS) and Structural Query Language Injection Attacks (SQLI) [1].

The SQL Injection Attack (SQLIA) is a type of attack that injects malicious codes into the original query structure of a web application with the motive of modifying, deleting, retrieving/manipulating sensitive data that target databases connected web applications [2]. This vulnerability subsists when there is no proper input validation, standard error reporting and poor website administration [3]. Malicious code can be injected into a web application that is poorly designed in other to get access to the back end database. There are scores of location where users can input data in web applications such as URLs and login form, each leading to SQL injection attack opportunity resulting to loss of integrity, market value and confidentiality of an organization [4]. Various methods have been proposed to detect, prevent and even classify SQLI such as static, dynamic and machine learning based approach [2].

Machine learning is defined as a type of Artificial Intelligence (AI) that gives computers the ability to learn without being explicitly programmed. Machine learning focuses on the development of computer programs that can change when exposed to new data [5]. Addendum, machine learning has become one of the bases of information technology in which knowledge is discovered using different algorithms from a specified form of data over two decades [6]. The intrinsic ability to learn knowledge from data, technique of machine learning is believed to attract better attention in information retrieval, data mining and pattern recognition because data plays indispensable role in machine learning and learning algorithm that are used to learn knowledge and discover properties from the data [7]. There are various types of model in machine learning such as Neural Network (NN), Support Vector Machine (SMV) and Naive Bayes (NB) to mention but a few [8]. This study presents a systematic literature review on Detection, Prevention and Classification of SQL Injection Attacks using machine learning approach.

The seven (7) most popular attack types of SQL injection [10] were considered in the review. The Table 1 below illustrates the seven (7) most SQL injection attack types.

REFERENCES

- [1] Kumar, P. & Pateriya, R. k. (2012). A survey on SQL injection attacks, detection and prevention techniques. Proceedings of the 3rd International conference of computing communication and networking technology. July, 26- 28.
- [2] Sayyed, M., Sadegh, S. & Bahare. T. P. (2014). Study of SQL Injection Attacks and Countermeasures *International Journal of Computer and Communication Engineering*, 2(5). 2013. Retrieved from <http://www.ijcce.org/papers/244-E091.pdf> and accessed on 20th march 2017
- [3] Jignesh, D. & Bhushan, T. (2017). Assessment of SQL Injection Solution Approaches. *International Journal of Advanced Research in Computer Science and Software Engineering*. Retrieved from www.ijarcsse.com and accessed on 20th April 2017
- [4] Lawal, M. A., Abubakar, M. D. & Ayanloye, O. S. (2016). Systematic literature review on SQL injection. *International Journal of Software Computing*, 11(1), 26-32. 2016 Retrieved from <http://docsdrive.com/pdfs/medwelljournals/ijscmp/26-35.pdf> and accessed on 20th April, 2017
- [5] Sumit., D., et al (2015). Applications of Artificial Intelligence in Machine Learning: Review and Prospect. *International Journal of Computer Applications (0975 – 8887)* 115 (9). Retrieved from <http://research.ijcaonline.org/volume115/number9/pxc3902402.pdf> and accessed on 20th April, 2017
- [6] Alex, S., & Vishwanathan, S.V.N. (2008). Introduction to machine learning. University of Cambridge: Press Syndicate published by the press syndicate. Retrieved <http://alex.smola.org/drafts/thebook.pdf> and accessed on 10th January 2017
- [7] Wei-Lun Machine Learning Tutorial 2011. Retrieved from <http://disp.ee.ntu.edu.tw/~wpjols/Machine%20Learning%20Tutorial.pdf> and accessed on 25th January 2017
- [8] Ian, H. W., & Eibe F.(2005). Data mining: A Practical Machine learning tools and techniques. 2nd edition Morgan kaufmann Publishers.
- [9] Romil R., & Shailendra, K. S. (2012) SQL injection attack Detection using SVM. *International Journal of Computer Applications*, 42(13). 2012. Retrieved from <http://research.ijcaonline.org/volume42/number13/pxc3877043.pdf> and accessed on 15th March 2017
- [10] Naghmeh, M. (2015). A Pattern Recognition Neural Network Model for Detection and Classification of SQL Injection Attacks. Retrieved from <https://www.researchgate.net/publication/271072307> and accessed on 20th May 2017
- [11] Fredrik, V. Darren, M. & Giovanni, V. (2005). A Learning-Based Approach to the Detection of SQL Attacks 2005. Retrived from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.94.9201&rep=rep1&type=pdf> and accessed on 20th November, 2017.
- [12] Aniruddh, L., & Phalke, D. A. (2016). SQL Injection Attack and User Behavior Detection by Using Query Tree, Fisher Score and SVM Classification. *International research journal of engineering and technology (IRJET)* 3 (6).
- [13] Naghmeh, M. S. (2014). Employing Neural Networks for the Detection of SQL Injection Attack. Retrieved from <https://www.researchgate.net/publication/264799665> and accessed on 4th April 2017
- [14] Christine, B., Ahmed, E., & Saad D. (2016). Detection of SQL Injection Using a Genetic Fuzzy Classifier System. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 7(6). Retrieved from <https://thesai.org> and accessed on 25th January 2017
- [15] Abdelhamid, M., Youcef, B. & Ahmed, S. (2014). Improving Web Application Firewalls to detect advanced SQL injection attacks. Information Assurance and Security (IAS). 10th International Conference. Retrieved from <http://ieeexplore.ieee.org/document/7064617/> and accessed on 20th April 2017
- [16] Amit, B and Tushar, V. (2016) SQL Injection detection using Baye's Classification. 3rd International Conference on Resent Innovation of Science Engineering and Management.
- [17] Anamika J. & Geetha, V. (2014). SQL Injection Detection using Machine Learning. International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT). IEEE. Retrieved from <http://ieeexplore.ieee.org/document/6993127/> accessed on 10th march 2017
- [18] Garima, S., Dev, K., Unique, G. & Akhilesh P. S. (2015). SQL Injection Detection and Correction Using Machine Learning Techniques. Emerging ICT for Bridging the Future - Proceedings of the 49th Annual Convention of the Computer Society of India (CSI) 1(1) pp 435-442.. Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-13728-5_49 and accessed on 15th March 2017.
- [19] Yi Wang & Zhoujun. L. (2012). SQL Injection Detection with Composite Kernel in Support Vector Machine *International Journal of Security and Its Applications* 6(2).
- [20] Ritu, A., & Dharmendra, M. (2016). An Approach Based on SVM Classifier to Detect SQL Injection Attack *IJSRSET*, 2(3) 2395-1990.
- [21] Krit, K. & Chitsutha J. (2016). Machine Learning for SQL Injection Prevention on Server-Side Scripting Soomlek. Retrieved from <http://ieeexplore.ieee.org/document/7859950/?denied> and accessed on 15th March 2017.
- [22] Hanmanthu, B., Raghu, B. R., & Niranjana, P. (2015). SQL Injection Attack Prevention Based on Decision Tree Classification IEEE Sponsored 9th International Conference on Intelligent Systems and Control (ISCO) 2015.
- [23] Solomon, O. U. & J. William, B. (2017). Applied Machine Learning Predictive Analytics to SQL Injection Attack Detection and Prevention. IEEE. 2017. Retrieved from <http://www.napier.ac.uk/~media/worktribe/output-687590/applied-machine-learning-predictive-analytics.pdf> and accessed on 26th April 2017.
- [24] Cristian, et al. CBRid4SQL: A CBR Intrusion Detector for SQL Injection Attacks 2010. Retrieved from https://www.researchgate.net/publication/221053341_CBRid4SQL_A_CBR_Intrusion_Detector_for_SQL_Injection_Attacks and accessed on 25th January 2017
- [25] Kanika, E. & Prabhjot, E. k. (2015). A dynamic approach to detect & prevent sql injection attack to overcome website vulnerability *International Journal of Innovative Research In Science, Engineering And Technology* 4 (12).
- [26] Alexander, A. (2010). A Distributed System for Pattern Recognition and Machine Learning 2010. Retrieved from http://madm.dfki.de/_media/theses/alex-dispare-thesis.pdf and accessed on 7th April 2017
- [27] Lambert N., & Lin, K. S. (2010). Use of Query Tokenization to detect and prevent SQL Injection Attacks, in Proc. 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), pp. 438-440.
- [28] Atefeh, T., Suhaimi, I. & Maslin, M. (2011). SQL Injection Detection and Prevention Techniques. *International Journal of Advancements in Computing Technology* 3(7). 2011. Retrieved from <https://pdfs.semanticscholar.org/0ff9/2dd1d6e0878347be94dae6c06365247aa46b.pdf> accessed on 20th march 2017
- [29] Rubidha, D.D., Venkatesan, R. & Raghuraman. K. (2016). A Study on Sql Injection Techniques. *International Journal of Pharmacy & Technology*. Retrieved from <https://www.researchgate.net/publication/271072307> and accessed on 20th May 2017
- [30] Avinash, K.S.(2011). Detection and Prevention of SQL Injection Attack in Web Application. 2011. Retrieved from <http://dpSQLklk.org/> and accessed on 25th January 2017
- [31] W3resource. SQL injection September 2017. Retrieved from <https://www.w3resource.com/sql/sql-injection/sql-injection.php> and accessed on 23rd Desember, 2017.
- [32] Joseph, R., Manoj A.C., & Anto, M. D. (2014). An Approach to Detect and Prevent Tautology Type SQL Injection in Web Service Based on XSchema validation. *International Journal Of Engineering And Computer Science* 3(1), pg 3695-3699. 2014.
- [33] Kuldeep, R. (2011). Classification of SQL Injection Attacks and using Encryption as a Countermeasure. *International Journal of Advanced Research in Computer Science*. 2 (1), Jan. -Feb, 628-630

- [34] San-Tsai, S., Ting, H. W., Stephen, L. & Sheung L. (2007). Classification of SQL Injection Attacks and using Encryption as a Countermeasure.2007. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.62.0.9731&rep=rep1&type=pdf> and accessed on 30th October 017.