FIT5190 Introduction to IT Research Methods

Assignment 2

Critical Review of Published Research

SQL injection attack detection

HU Ying (2919****)

SOUTHEAST UNIVERSITY - MONASH UNIVERSITY JOINT

GRADUATE SCHOOL

Submission date: 12/5/2017

Abstract

Along with the increasing trend in the developments of web applications recently, the security issues in web applications also deserve more attention. SQL injection attack is a type of code injection attacks, which is one of the most serious threats to web security. This review mainly presents some existing detection techniques in terms of different analysis methods. To better clarify and evaluate the detection methods, the types of SQL injection attacks are also introduced briefly. For each detection method or model, the performance is evaluated from several aspects. In conclusion section, the strengths for these analysis methods are described from an overall review.

Keywords: SQL injection, SQL injection detection, web security

Introduction

As web applications become more widely used, web security also faces more threats. Structured Query Language (SQL) injection attack is one of the most frequent attack techniques used by attackers. In the condition that more web applications are developed based on Browser/Server model, users access web service mainly through WWW browsers. But the validation of user input is often insufficient during the development process of web application [1]. Given that condition, attackers utilize predefined SQL statements and append a series of string concatenation of static strings and variables in an undesirable way to access databases without authorization, leading to an unintended effect on databases **Error! Reference source not found.**.

Since SQL injection just exploits the vulnerability of databases in web applications to request queries, hardly distinguished from legal user requests, firewalls are unable to prevent SQL injection attacks. Apart from that, tradition pattern matching methods are only able to recognize a few types of SQL injection attacks, due to the diversity of SQL queries. Furthermore, there are a large number of SQL injection tools widely spread on the Internet, attackers can quickly intrude target websites with the aid of these tools, causing huge harm. One of the serious consequences caused by SQL injection are exposing the user privacy stored in databases. Website contents can also be tempered so that illegal information would be published. Even worse, once attackers successfully change or add administrator accounts so that database servers are completely controlled by attackers, Trojans are likely to be embedded into the server to conduct further attacks.

Because of the serious threats brought by SQL injection, effective methods for detecting and preventing SQL injection attacks are desired to be proposed. This review aims to introduce some existing detection methods from different analysis techniques. The rest of this review is organized as follows. In section 2, the scope and method of this literature review are described. In section 3, the SQL injection types are specified, and different SQL injection attack detection methods are classified and introduced separately. In section 4, the interpretation and conclusion of this review are conducted.

Scope and method

In this review, all papers are gained from Monash Library and google scholar by using searching keywords related to SQL injection detection. These papers specify their techniques and propose their detection methods or both detection and prevention methods. For the background knowledge, SQL injection attack is defined, and its classification are introduced briefly. We mainly focus on the detection techniques of SQL injection attacks using different analysis methods including static analysis, dynamic analysis, combined static and dynamic analysis, and machine learning methods. For each analysis method, a brief description of its research field is stated, with at least one related paper introduced. A brief evaluation is also stated for each detection methods.

Body of the review

SQL injection attacks is a kind of code injection attacks where SQL queries are likely to be changed by un-sanitized user input leading to undesired consequences on the database [3]. In this section, several attack types are presented. Then different analysis methods are introduced. Some methods have capability to detect most types of SQL attacks, while some just indicate their effectiveness in part of types.

SQL injection types

a) Tautologies

A tautology attack refers to injecting additional conditional statements into SQL queries so that the queries are always evaluated as true [3]. Typical attack intents include bypassing authentication and identifying injectable parameters.

b) Illegal/logically incorrect queries

This attack allows attackers reveal the vulnerable parameters existing in a web application by analyzing the error responses through inputting a series of incorrect queries [3].

c) Union queries

In union queries, the queries are designed to treat the application in order to get query results from an unintended database. Attackers inserts a UNION query into the query sentences to achieve this goal.

d) Other attack types

There are some other common SQL injection attacks including Stored Procedure, PiggyBacked Queries, Alternate Encodings and Inference (Blind Injection & Timing Attacks).

Analysis methods

Static analysis

Static analysis refers to discovering the existence of security vulnerabilities in web applications to detect and prevent SQL injection attacks. Static analysis methods mainly focus on analyzing SQL query sentences to validate the user input effectively so as to avoid illegal SQL injection. According to the analysis results, the source code of the web application needs to be modified [4].

Jang and Choi **Error! Reference source not found.** develop a technique to detect SQL injection attacks by comparing the malicious query results against the query result of the normal query. This technique substitutes vulnerable input variables into valid variables, and estimates query result size in two conditions by using an Equivalence and Largest Selectivity algorithm. The SQL injection attack will be detected when the query result size of its original query sentence is not equal to that of its substitute. The implementation of this technique also utilizes some features of Java including type-safe. But it lacks of considering multiple table queries in the web application during query parsing analysis.

Dynamic analysis

Dynamic analysis refers to analyzing responses from the web application after runtime scanning. Dynamic analysis methods are required to generate different kinds of test case input so as to collect enough responses covering cases as much as possible. It mainly focuses on locating vulnerability on SQL injection attacks [4]. Thus, the input cases have a decisive influence on the accuracy of the method.

Sania [6] is capable of intercepting SQL queries between the web application and the database during the development and debugging phases to reveal SQL injection vulnerabilities automatically. In the beginning, it analyzes the SQL queries generated from the HTTP request to discover the existing vulnerabilities. Then it exploits these vulnerabilities to generate targeted SQL queries to carry out SQL injection attacks elaborately. After sending these targeted request, Sania collects the parse tree of the malicious queries. If the parse trees of the innocent queries need to be compared with these of malicious queries to determine whether the attack succeeds or not. Through these steps, Sania could detect the vulnerabilities in the web application effectively. Since the vulnerabilities always appears in leaf nodes of parse trees, the accuracy of Sania is higher than that of analyzing HTTP responses. In addition, Sania is more efficiently with one SQL query injected by more than two vulnerabilities. Though a number of security vulnerabilities would be discovered through the dynamic analysis method, further repair measures are supposed to be taken by developers to prevent these attacks. The advantage is that there is no modifications in web applications unlike static analysis methods.

Combined Static and Dynamic analysis

Combining static and dynamic analysis method usually has the complementary advantages from both analysis method. Recently, there are more studies using both static and dynamic analysis in SQL injection attack detection.

Prabakar et al. [7] propose a method based on Aho–Corasick pattern matching algorithm for detecting and preventing SQL injection attacks. This method is composed of two phase: static phase and dynamic phase. In the static phase, each pattern in user input queries are compared with anomaly patterns in static pattern list where a number of known anomaly patterns is preset, by using Aho-Corasick pattern matching algorithm. If one pattern is matched, the user query will be rejected directly. Otherwise, anomaly score value will be calculated for the user query for actions in dynamic phase. In this phase, if the anomaly score is higher than the threshold assumed, the administrator needs to judge manually whether these user queries are anomaly or not, then updates the static pattern list by adding new-found anomaly pattern. In this method, the static pattern list is the judgment criterion for the anomaly user input queries. The updating process in dynamic phase guarantees the adaptability to new types of SQL injection attacks. But with the anomaly patterns explosively increasing nowadays, not only more space and memory may be needed, but also the matching speed may slow down. The other flaw is that there is no concrete SQL injection types clarified and targeted in that paper.

Lee et al. [4] propose a detection method for SQL injection attacks combining static and dynamic analysis. Fixed SQL queries in the web application are analyzed previously by applying a function removing the attribute values in the SQL queries. At runtime, the user-generated queries also apply the same function to remove the attribute values. Then two strings obtained from the processed fixed SQL query and user input query operate exclusive OR operator. If the result equals zero, the user query is normal. Otherwise it is abnormal. Different from the method [7] mentioned before, this detection method is completely automatically without manual operations. Besides, this paper applies this method in different kinds of SQL injection attacks including tautologies, illegal/logically incorrect queries, union queries, Piggy-Backed queries and stored procedures. The most obvious advantages in this method is its low computational complexity and constant time cost.

Machine learning

With the rapid increasing amount and variety of malicious codes, it is more and more difficult for traditional detection techniques to detect SQL injection attacks effectively. The application of machine learning becomes a new trend in SQL injection attack detection. Because of its advantages in pattern extraction and classification.

Joshi and Geetha [8] use Naïve Bays machining learning algorithm to classify the SQL queries. The first step is to extract features through blank separating. Then queries are broken into several meaning full elements. This method combines Role Based Access Control, where the role of the user is regarded as an evaluation factor. The classification of malicious queries and normal queries depends on the calculation of prior probability and likelihood following Bayes' rule. The proposed model tests three types of SQL injection attacks: tautology, union and comments, reaching a high accuracy.

Interpretation and conclusion

In this review, SQL injection attacks detection methods are introduced from different analysis methods: static analysis, dynamic analysis, combined static and dynamic analysis, and machine learning methods. There are many types of SQL injection attacks with different injection features, so it is one of key factors whether proposed methods are able to detect the majority of attacks. From the methods mentioned above, static analysis seems play a crucial role for an effective detection, where the extraction of patterns or features of SQL queries is conducted. More logical and considerate extraction contributes a higher accuracy. Given varieties of malicious injection queries, where there may still be a cognition gap, a dynamic phase is also necessary to implement further evaluation or even give feedbacks to modify the analysis results in previous static phase. Therefore, combined static and dynamic methods may be superior to these applying either static or dynamic analysis to some extent. With regard to machine learning methods, the most obvious character is its advantages in feature extraction. The model could be optimized through plenty of training to achieve better accuracy.

References

[1] Halfond, W. G. J., & Orso, A. (2005). AMNESIA:analysis and monitoring for NEutralizing SQL-injection attacks. Ieee/acm International Conference on Automated Software Engineering, 174-183.

[2] Halfond, W. G., Viegas, J., & Orso, A. (2006). A classification of SQL-injection attacks and countermeasures. In Proceedings of the IEEE International Symposium on Secure Software Engineering, 1, 13-15.

[3] Valeur, F., Mutz, D., & Vigna, G. (2005). A learning-based approach to the detection of sql attacks, 3548, 533-546.

[4] Lee, Inyong, Jeong, Soonki, Yeo, Sangsoo, & Moon, Jongsub. (2012). A novel method for SQL injection attack detection based on removing SQL query attribute values. Mathematical and Computer Modelling, 55(1 2), 58-68.

[5] Jang, Y. S., & Choi, J. Y. (2014). Detecting SQL injection attacks using query result size. Computers & Security, 44, 104-118.

[6] Kosuga, Hanaoka, Hishiyama, & Takahama. (2007). Sania: Syntactic and Semantic Analysis for Automated Testing against SQL Injection. Computer Security Applications Conference, 2007. ACSAC 2007. Twenty-Third Annual, 107-117.

[7] Prabakar, M., Karthikeyan, M., & Marimuthu, K. (2013). An efficient technique for preventing SQL injection attack using pattern matching algorithm. Emerging Trends in Computing, Communication and Nanotechnology (ICE-CCN), 2013 International Conference on, 503-506.

[8] Joshi, A., & Geetha, V. (2014). SQL Injection detection using machine learning. 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 1111-1115.