A Distributed Hierarchical Multi-agent Architecture for Detecting Injections in SQL Queries

Cristian Pinzón, Juan F. De Paz, Álvaro Herrero,
Emilio Corchado, and Javier Bajo

Abstract. SQL injections consist in inserting keywords and special symbols in the parameters of SQL queries to gain illegitimate access to a database. They are usually identified by analyzing the input parameters and removing the special symbols. In the case of websites, due to the great amount of queries and parameters, it is very common to find parameters without checking that allow bad-intentioned users to introduce keywords and special symbols. This work proposes a distributed architecture based on multi-agent systems that is able to detect SQL injection attacks. The multi-agent architecture incorporates cased-based reasoning, neural networks and support vector machines in order to classify and visualize the queries, allowing the detection and identification of SQL injections. The approach has been tested and the experimental results are presented in this paper.

Keywords: SQL injection, Database Security, Intrusion Detection Systems, Multi-agent Systems, Case-based Reasoning, Unsupervised Projection Models.

1 Introduction

A potential security problem of databases is the SQL injection attack. This attack takes place when a hacker changes the semantic or syntactic logic of an SQL text

Cristian Pinzón · Juan F. De Paz · Emilio Corchado · Javier Bajo
Departamento Informática y Automática, Universidad de Salamanca,
Plaza de la Merced s/n, 37008, Salamanca, Spain
e-mail:{cristian_ivanp,fcofds,escorchado,jbajope}@usal.es

Álvaro Herrero
Department of Civil Engineering. University of Burgos,
C/ Francisco de Vitoria S/N, 09006, Burgos, Spain
e-mail: ahcosio@ubu.es

Á. Herrero et al. (Eds.): CISIS 2010, AISC 85, pp. 51–59.
springerlink.com

string by inserting SQL keywords or special symbols within the original SQL command. The SQL query will then be executed at the database layer of an application [1], [6], being extremely dangerous in the case of online applications as the answer to the query will be available through a web browser. The results of this attack can produce unauthorized handling of data, retrieval of confidential information, and in the worst possible case, taking over control of the application server.

Nowadays, this type of attack has been handled from distinct perspectives. The string analysis [7] has been the support of many others approaches such as [1] and [8], which carried out a more complete analysis applying a dynamic and hybrid treatment over the SQL string. In other cases, computational intelligence techniques have been applied to face the SQL injection attack, such as [9], [2], [3] with WAVES (Web Application Vulnerability and Error Scanner). These approaches apply machine learning techniques based on a dataset of legal transactions and artificial neural networks. Usually, many approaches present a poor performance, with high error rates (both false positive and false negative rates). The performance of misuse-based intrusion detection systems depend on the database, which requires a continue update in order to detect new attacks.

The proposal presented in this work tackles the SQL injection attack problem through a distributed hierarchical multi-agent architecture to detect SQL attacks in queries. The key component is the intelligent agent CBRid4SQL (a Case-Based Reasoning Intrusion Detector), capable of detecting attacks based on SQL code injection. CBRid4SQL is an agent that addresses the SQL injection problem from the Intrusion Detection standpoint by combining different Computational Intelligence techniques. This is the principal component of a distributed hierarchical multi-agent system aimed at detecting a wide range of attacks in dynamic and distributed environments. CBRid4SQL is a CBR agent [13] characterized by the integration of several techniques within the CBR mechanism. This mechanism provides the agents with a great level of adaptation and learning capability, since CBR systems make use of past experiences to solve new problems [13]. This is very effective for blocking SQL injection attacks as the mechanism uses a strategy based on anomaly detection [14]. The multi-agent system incorporates classification and visualization techniques in the different phases of the reasoning cycle.

The rest of the paper is structured as follows: section 2 focuses on the details of the proposed multiagent architecture while section 3 comprehensively explains the integrated classification model. Finally, section 4 describes how the proposed agent has been tested in the frame of a multi-agent system and presents the obtained results.

2 A Multi-agent Architecture for the Detection of SQL Injection

The agents are characterized through their capacities such as autonomy, reactivity, pro-activity, social abilities, reasoning, learning and mobility [4]. One of the main features of agents is their ability to carry out cooperative and collaborative work, when they are grouped into multi-agent systems to solve problems in a distributed