

Álvaro Herrero  
Emilio Corchado  
Carlos Redondo  
Ángel Alonso (Eds.)

# Computational Intelligence in Security for Information Systems 2010

Proceedings of the 3rd International Conference  
on Computational Intelligence in Security  
for Information Systems (CISIS'10)



The series "Advances in Intelligent and Soft Computing" contains publications on various areas within the so-called soft computing, which include fuzzy sets, rough sets, neural networks, evolutionary computations, probabilistic and evidential reasoning, multi-valued logic, and related fields. The publications within "Advances in Intelligent and Soft Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both theoretical and applied. An important feature of the series is the short publication time and world-wide distribution, allowing a rapid and broad dissemination of research results.

Álvaro Herrero · Emilio Corchado · Carlos Redondo · Ángel Alonso (Eds.)  
*Computational Intelligence in Security for Information Systems 2010*

This volume of *Advances in Intelligent and Soft Computing* contains accepted papers presented at the 3rd International Conference on Computational Intelligence in Security for Information Systems (CISIS'10), which was held in the beautiful and historic city of León, Spain, in November, 2010.

Since its origin, the main aim of CISIS conferences has been to establish a broad and interdisciplinary forum for Computational Intelligence and Associated Learning Paradigms, which are playing increasingly important roles in a large number of application areas related to Security in Information Systems.

CISIS'10 provided a broad and interdisciplinary forum to present the most recent developments in several very active scientific areas such as Machine Learning, Infrastructure Protection, Intelligent Methods in Energy and Transportation, Network Security, Biometry, Cryptography, High-performance and Grid Computing, and Industrial Perspective among others. The global purpose of the CISIS series of conferences has been to form a broad and interdisciplinary meeting ground offering the opportunity to interact with the leading research teams and industries actively involved in the critical area of security, and have a picture of the current solutions adopted in practical domains.

ISSN 1867-5662

ISBN 978-3-642-16625-9



*Available  
online*  
*springerlink.com*









<b>Chapter 3: Image, Video and Speech Processing</b>	
Performances of Speech Signal Biometric Systems Based on Signal to Noise Ratio Degradation.....	73
<i>Dzati Athiar Rami, Salima Abdul Samad, Aini Hussain</i>	
Lipreading Using $n$ -Gram Feature Vector.....	81
<i>Preetly Singh, Vijay Larmi, Deepika Gupta, M.S. Gaur</i>	
Face Processing for Security: A Short Review .....	89
<i>Ion Marqués, Manuel Graña</i>	
<b>Chapter 4: Network Security</b>	
Ontologies-Based Automated Intrusion Response System .....	99
<i>Verónica Mateos Lanchas, Víctor A. Villagrà González, Francisco Romero Bueno</i>	
Semi-supervised Fingerprinting of Protocol Messages.....	107
<i>Jérôme François, Humberto Abdehur, Radu State, Olivier Fester</i>	
Monitoring of Spatial-Aggregated IP-Flow Records.....	117
<i>Cynthia Wagner, Gerard Wagener, Radu State, Thomas Engel</i>	
Improving Network Security through Traffic Log Anomaly Detection Using Time Series Analysis .....	125
<i>Aitor Corchero Rodríguez, Mario Reyes de los Mozos</i>	
A Threat Model Approach to Threats and Vulnerabilities in On-line Social Networks .....	135
<i>Carlos Laorden, Borya Sanz, Gonzalo Alvarez, Pablo G. Bringas</i>	
An SLA-Based Approach for Network Anomaly Detection ...	143
<i>Yasser Yasami</i>	
Understanding HoneyPot Data by an Unsupervised Neural Visualization .....	151
<i>Álvaro Alonso, Santiago Porras, Enaitz Ezpeleta, Ekihiotz Vergara, Ignacio Arenaza, Roberto Uribeetxeberria, Urko Zurutuza, Álvaro Herrero, Emilio Corchado</i>	
<b>Chapter 5: Watermarking</b>	
Permuted Image DCT Watermarking .....	163
<i>Reena Gunjan, Saurabh Maheshwari, M.S. Gaur, Vijay Larmi</i>	

A Developed Watermark Technique for Distributed Database Security .....	173
<i>Hazem M. El-Bakry, Mohamed Hamada</i>	
<b>Chapter 6: Cryptography</b>	
Trident, a New Pseudo Random Number Generator Based on Coupled Chaotic Maps .....	183
<i>Amalia Beatriz Orúe López, Gonzalo Alvarez Marañón, Alberto Guerra Estévez, Gerardo Pastor Dégano, Miguel Romera García, Fausto Montoya Vitini</i>	
The Impact of the SHA-3 Casting Cryptography Competition on the Spanish IT Market .....	191
<i>Manuel J. Martínez, Roberto Uribeetxeberria, Urko Zurutuza, Miguel Fernández</i>	
<b>Chapter 7: Industrial and Commercial Applications of Intelligent Methods for Security</b>	
A New Task Engineering Approach for Workflow Access Control.....	203
<i>Hanan El Bakkali, Hamid Hatim, Ilham Berrada</i>	
OPBUS: Fault Tolerance Against Integrity Attacks in Business Processes.....	213
<i>Angel Jesus Varela Vaca, Rafael Martínez Gasca</i>	
A Key Distribution Scheme for Live Streaming Multi-tree Overlays .....	223
<i>Juan Álvaro Muñoz Narango, Juan Antonio López Ramos, Leocadio González Casado</i>	
Intelligent Methods for Scheduling in Transportation .....	231
<i>M<sup>re</sup> Belén Vaguerizo García</i>	
Author Index .....	239



- Vol. 71. Y. Demazeau, F. Dignum, J.M. Corchado, J. Bujó, R. Corchuelo, E. Corchado, F. Fernández-Riverola, V.J. Julián, P. Pawłowski, A. Campbell (Eds.)  
*Trends in Practical Applications of Agents and Multiagent Systems, 2010*  
ISBN 978-3-642-12432-7
- Vol. 72. J.C. Augusto, J.M. Corchado, P. Novais, C. Analide (Eds.)  
*Ambient Intelligence and Future Trends, 2010*  
ISBN 978-3-642-13267-4
- Vol. 73. J.M. Corchado, P. Novais, C. Analide, J. Sedano (Eds.)  
*Soft Computing Models in Industrial and Environmental Applications, 5th International Workshop (SOCO 2010), 2010*  
ISBN 978-3-642-13160-8
- Vol. 74. M.P. Rocha, F.F. Riverola, H. Shatay, J.M. Corchado (Eds.)  
*Advances in Bioinformatics, 2010*  
ISBN 978-3-642-13213-1
- Vol. 75. X.Z. Gao, A. Gaspar-Cunha, M. Köppen, G. Schaefer, and J. Wang (Eds.)  
*Soft Computing in Industrial Applications, 2010*  
ISBN 978-3-642-11281-2
- Vol. 76. T. Bastiaens, U. Baumöl, and B.J. Krämer (Eds.)  
*On Collective Intelligence, 2010*  
ISBN 978-3-642-14480-6
- Vol. 77. C. Borgelt, G. González-Rodríguez, W. Trutschnig, M.A. Lubiano, M.A. Gil, P. Grzegorzewski, and O. Hryniewicz (Eds.)  
*Combining Soft Computing and Statistical Methods in Data Analysis, 2010*  
ISBN 978-3-642-14745-6
- Vol. 78. B.-Y. Cao, G.-J. Wang, S.-Z. Guo, and S.-L. Chen (Eds.)  
*Fuzzy Information and Engineering 2010*  
ISBN 978-3-642-14879-8
- Vol. 79. A.P. de Leon F. de Carvalho, S. Rodríguez-González, J.F. De Paz Santana, and J.M. Corchado Rodríguez (Eds.)  
*Distributed Computing and Artificial Intelligence, 2010*  
ISBN 978-3-642-14882-8
- Vol. 80. N.T. Nguyen, A. Zegrzwa, and A. Czyżewski (Eds.)  
*Advances in Multimedia and Network Information System Technologies, 2010*  
ISBN 978-3-642-14988-7
- Vol. 81. J. Düh, H. Hüfagel, E. Juritsch, R. Pfitzger, H.-K. Schimany, and Hans Schönegger (Eds.)  
*Data and Mobility, 2010*  
ISBN 978-3-642-15502-4
- Vol. 82. B.-Y. Cao, G.-J. Wang, S.-L. Chen, and S.-Z. Guo (Eds.)  
*Quantitative Logic and Soft Computing 2010*  
ISBN 978-3-642-15659-5
- Vol. 83. J. Angeles, B. Boulter, J.J. Clark, J. Koveses, and K. Siddiqi (Eds.)  
*Brain, Body and Machine, 2010*  
ISBN 978-3-642-16258-9
- Vol. 84. Ryszard S. Choraś (Ed.)  
*Image Processing and Communications Challenges 2*  
ISBN 978-3-642-16294-7
- Vol. 85. Álvaro Herrero, Emilio Corchado, Carlos Redondo, and Ángel Alonso (Eds.)  
*Computational Intelligence in Security for Information Systems 2010*  
ISBN 978-3-642-16625-9

Álvaro Herrero, Emilio Corchado,  
Carlos Redondo, and Ángel Alonso (Eds.)

# Computational Intelligence in Security for Information Systems 2010

Proceedings of the 3rd International  
Conference on Computational Intelligence in  
Security for Information Systems (CISIS 2010)



## Editors

Dr. Álvaro Herrero  
University of Burgos  
Civil Engineering Department  
Polytechnic School  
Francisco de Vittoria s/n  
09006 Burgos  
Spain  
E-mail: ahcosio@ubu.es

Carlos Redondo  
Fundación Centro de  
Supercomputación de  
Castilla y León  
24071 León  
Spain  
E-mail: carlos.redondo@fesc.es

Emilio Corchado  
University of Salamanca  
Departamento de Informática y Automática  
Facultad de Biología  
Plaza de la Merced s/n  
37008 Salamanca  
Spain  
E-mail: escorchado@usal.es

Ángel Alonso  
Fundación Centro de  
Supercomputación de  
Castilla y León  
24071 León  
Spain

ISBN 978-3-642-16625-9

e-ISBN 978-3-642-16626-6

DOI 10.1007/978-3-642-16626-6

Advances in Intelligent and Soft Computing

ISSN 1867-5662

Library of Congress Control Number: 2010937280

© 2010 Springer-Verlag Berlin Heidelberg

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typeset & Cover Design: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper  
5 4 3 2 1 0  
springer.com

## Preface

The 3<sup>rd</sup> International Conference on Computational Intelligence in Security for Information Systems (CISIS 2010) provided a broad and interdisciplinary forum to present the most recent developments in several very active scientific areas such as Machine Learning, Infrastructure Protection, Intelligent Methods in Energy and Transportation, Network Security, Biometry, Cryptography, High-performance and Grid Computing, and Industrial Perspective among others. The global purpose of CISIS series of conferences has been to form a broad and interdisciplinary meeting ground offering the opportunity to interact with the leading research team and industries actively involved in the critical area of security, and have a picture of the current solutions adopted in practical domains.

This volume of Advances in Intelligence and Soft Computing contains accepted papers presented at CISIS 2010, which was held in León, Spain, on November 11–12, 2010. CISIS 2010 received over 50 technical submissions. After a thorough peer-review process, the International Program Committee selected 25 papers which are published in this conference proceedings. This allowed the Scientific Committee to verify the vital and crucial nature of the topics involved in the event, and resulted in an acceptance rate close to 50% of the originally submitted manuscripts.

The selection of papers was extremely rigorous in order to maintain the high quality of the conference and we would like to thank the members of the Program Committee for their hard work in the reviewing process. This is a crucial process to the creation of a conference high standard and the CISIS conference would not exist without their help.

Our warmest and special thanks go to the Keynote Speakers: Prof. Ajith Abraham from MIR-Labs, EUROPE and Dr. Jorge Ramón Aguirre - Universidad Politécnica de Madrid (Spain).

Particular thanks go as well to the conference main sponsors, namely Junta de Castilla y León, Supercomputing Center of Castilla y León, University of León, and Technical Co-Sponsors: IEEB - SECCION ESPAÑA, IEEB Systems, Man and Cybernetics-Spanish Chapter, MIR-Labs, and the International Federation for Computational Logic who jointly contributed in an active and constructive manner to the success of this initiative.



Table 1 summarizes the performance of the classifier obtained with the test set.

Table 1 Results of the classification on the test set

Class	Precision	Recall	FP rate
Without defects in HDGL	30%	60%	35%
With defects in HDGL	87%	65%	40%

Following this approach, the agent devoted to evaluate whether the coil can enter the hot dip galvanizing line forecasts the likelihood of the blister to be visible, which allows the operator of the line to decide whether recycle that particular coil.

## 5 Conclusions

This paper has shown an approach to the control of a intricate production environment based on a decentralized configuration of different agents which provide the means to gain deeper knowledge on the actual status of the underlying processes. The multi-agent system relies on data mining techniques and algorithms to improve the performance of the processes and to reduce the environmental footprint. The decentralized configuration resulted apposite to the actual configuration of the different manufacturing facilities, which reinforces the belief in this approach.

## References

1. Wiemker, R., Ekin, A., Ofter, R., Bilow, T., Rogalla, T.: Unsupervised extraction and quantification of the bronchial tree on ultra-low-dose vs. standard dose CT. In: Proceedings of the SPIE, San Diego CA, vol. 6143 (2006)
2. Ou-Yang, C., Juan, Y.C.: Applying process mining approach to support the verification of a multi-agent system. *Journal of Systems Science and Systems Engineering* 19(2), 131–149 (2010)
3. Mao, Y.X., Shi, B.L.: An Efficient Method for Mining Generalized Frequent Itemsets. In: 2nd International Conference on Intelligent Information and Database Systems, Hue City Vietnam, pp. 82–92 (2010)
4. Pei, J., Han, J., Laks, V.S.: Mining frequent itemsets with convertible constraints. In: Proceedings of the 17th International Conference on Data Engineering, Heidelberg, Germany, pp. 433–442 (2001)
5. Pei, Z., Liu, X., Zou, L.: Extracting association rules based on intuitionistic fuzzy sets. *International Journal of Innovative Computing Information and Control* 6(6), 2567–2580 (2010)
6. Zaki, M.J.: Scalable algorithms for association mining. *IEEE Transactions on Knowledge and Data Engineering* 12(3), 372–390 (2000)

## 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 case-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\_ivanzp, fcofds, escorchado, jbaajo}@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



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 CBRI4SQL (a Case-Based Reasoning Intrusion Detector), capable of detecting attacks based on SQL code injection. CBRI4SQL 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. CBRI4SQL 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

way [11]. These features make the agents suitable to face the SQL injection attack problem. A distributed hierarchical multi-agent system presents a great capacity for the distribution of tasks and responsibilities, error recovering, adaptation to new changes and high level of learning. These factors are key to achieve a robust and efficient solution. One main innovation of the proposed architecture is the use of a CBR agent [5], which presents a great capacity of learning and adaptation. This CBR mechanism additionally incorporates a mixture of a neural network [10] and support vector machine (SVM) [12] in order to identify SQL injections. The types of agents within the architecture are described as follows:

- **Sensor agents:** Located in each of the devices accessing the database. They have 3 specific functions: a) the capture of datagrams launched by the devices, b) Order TCP fragments to extract the request's SQL string, c) Syntactic analysis of the request's SQL string. The duties of the agent Sensor end when the results (the SQL string transformed by the analysis, the result of the analysis of the SQL string and the user data) are sent to the next agent at the hierarchy of the classification process.
- **FingerPrint agents:** The numbers of agents FingerPrint depend on the workload at a given time. An agent FingerPrint receives the information of a Sensor agent and executes a pattern matching known attacks stored at a previously built database. The FingerPrint agent finishes its task when it sends its results to the Anomaly agent. The results of the FingerPrint agent consist of the SQL string transformed by the analysis, the result of the analysis of the SQL string, the user data and the results achieved by pattern matching.
- **CBRI4SQL agents:** These agents are based on the CBR model. They are the key component of the classification process. Their strategy is based on a case-based reasoning mechanism that incorporates a mixture of neural networks. These agents retrieve those past cases that are the most similar to the new case to be classified, train the neural networks with the recovered cases and generate the final classification for the new case. The result of the classification is sent to the Manager agent for the evaluation.
- **Manager agent:** This is the agent responsible for decision-making, evaluation and coordination of the overall operation of the architecture. It evaluates the final decisions for classifications, manages alerts of attacks and coordinates the necessary actions when an attack is detected.
- **Interface agent:** This agent allows the interaction of the user of the security system with the architecture. The interface agent communicates the details of an attack to the security personnel when an attack is detected. Its ability to work on mobile devices allows a ubiquitous communication to manage the alerts immediately.

Fig. 1 depicts the hierarchical multi-agent architecture showing different types of agents in charge of the classification of SQL queries.



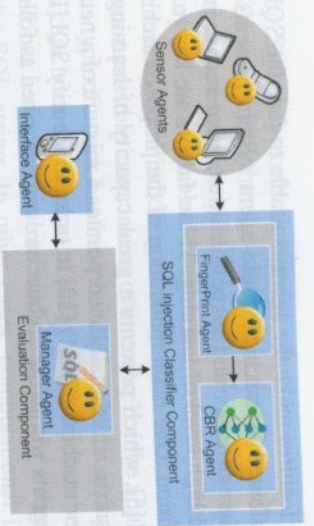


Fig. 1 Description of the distributed hierarchical multi-agent architecture

### 3 Classifier CBR Agent

In this section the CBRid4SQL agent is presented, with special attention paid to its internal structure and the classification mechanism of SQL attacks. This mechanism combines the advantages of CBR systems, such as learning and adaptation, with the predictive capabilities of a combination integrated by ANNs and SVMs. The use of this combination of techniques is based on the possibility of using two classifiers together to detect suspicious queries in the most reliable way possible.

**Table 1** Structure of the problem definition and solution for a case of SQL query classification

Problem Description fields		Solution fields	
IdCase	Integer	Idcase	Integer
Session	String	Classification_Query	Integer
User	String		
IP_Address	String		
Query_SQL	Query_SQL		
Affected table	Integer		
Affected field	Integer		
Command_type	Integer		
Word_GroupBy	Boolean		
Word_Having	Boolean		
Word_OrderBy	Boolean		
Numer_And	Integer		
Numer_Or	Integer		
Number_literals	Integer		
Length_SQL_String	Integer		
Cost_Time_CPU	Float		
Start_Time_Execution	Time		
End_Time_Execution	Time		
Query_Category	Integer		

In terms of CBR, the case is composed of the following elements of an SQL query: (a) Problem Description that describes the initial information available for generating a plan. The problem description consists of: case identification, user session and SQL query elements. (b) Solution that describes the action carried out in order to solve the problem description, in this case, prediction models. (c) Final State that describes the state achieved after that the solution has been applied.

A Multilayer Perceptron (MLP) and an SVM are applied simultaneously to carry out the prediction of the new query. Subsequently, a new inspection is performed which can be done automatically or by a human expert. In the case of the query resulting as suspicious, further inspection is carried out manually by a human expert. At this stage the most similar cases will be selected by means of a Growing Cell Structure (GCS) network [15], and then visualized by a dimensionality reduction technique which employs the neuronal model called Cooperative Maximum Likelihood Hebbian Learning (CMLHL) [16]. As a result of such visualization, the human expert will graphically see the relationship between the suspicious query and the recovered queries. During learning, memory information regarding the cases and models will be updated. The different stages of the CBR reasoning cycle associated with the system are comprehensively described in the following sections.

#### 3.1 Retrieve

The retrieval phase consists of two phases; case retrieval and model retrieval. The case retrieval is performed by using the *Query\_Category* attribute which retrieves queries from the case memory which were used for a similar query in accordance with the attributes of the new case. Subsequently, the models for the MLP and associated with the recovered cases are retrieved. The recovery of these memory models allows the improvement of the system's performance so that the time necessary for the creation of such models is considerably reduced, mainly in the case of the ANN training.

#### 3.2 Reuse

The reuse phase initially considers the information of the retrieved cases and the recovered models of the MLP and the SVM. The combination of both techniques is fundamental in the reduction of the false negative rate. The inputs of the MLP an SVM are: *Query\_SQL*, *Affected\_table*, *Affected\_field*, *Command\_type*, *Word\_GroupBy*, *Word\_Having*, *Word\_OrderBy*, *Numer\_And*, *Numer\_Or*, *Number\_literals*, *Number\_LOL*, and *Length\_SQL\_String*. The number of neurons in the hidden layer of the MLP is  $2n+1$ , where  $n$  is the number of neurons in the input layer. Finally, there is only one neuron in the output layer. The activation function selected for the different layers has been the sigmoid.



As the neurons exiting from the hidden layer of the neural network contain sigmoidal neurons with values between  $[0, 1]$ , the incoming variables are redefined so that their range falls between  $[0.2-0.8]$ .

At the same time as the estimation through the use of neuronal networks is performed, estimation is also carried out by the SVM application, a supervised learning technique applied to the classification and regression of elements. The algorithm represents an extension of nonlinear model [12].

The selected kernel function in this problem was polynomial. The values used for the estimation are dominated by decision values and are related to the distance from the points to the hyperplane.

Once the output values for the ANN and the SVM are obtained, the mixture is performed by way of a weighted average in function of the error rate of each one of the techniques. Before carrying out the average, the values are normalized to the interval  $[0, 1]$ , as SVM provides positive and negative values and those of greater magnitude, so that it could affect the final value in greater measure if it is not redimensioned.

### 3.3 Revise

The revise phase can be manual or automatic depending on the output values. The automatic review is given for non-suspicious cases during the estimation obtained for the reuse phase. For cases detected as suspicious, with output values determined experimentally in the interval  $[0.35, 0.6]$ , a review by a human expert is performed. As CBR learns, the interval values are automatically adjusted to the smallest of the false negatives. The greater limit is constantly maintained throughout the iterations. The review consists of recovering those queries similar to the current one together with their previous classifications. To do so, a clustering technique (for the selection of similar requests) and a neuronal model (for the reduction of dimensionality) are combined to generate an informative visualisation in 2D or 3D.

The selection of similar cases is carried out through the use of a neuronal GCS network, the different cases are distributed in meshes and the mesh containing the new case is selected. To visualize the cases (those in the selected mesh), the dimensionality of data is reduced by means of the CMLHL neuronal model [16] which performs Exploratory Projection Pursuit by unsupervised learning.

### 3.4 Retain

The learning phase updates the information of the new classified case and reconstructs the classifiers offline to leave the system available for new classifications. The ANN classifier is reconstructed only when an erroneous classification is produced. In the case of a reference to inspection of suspicious queries, information and classifiers are updated when the expert updates the information.

## 4 Results and Conclusions

To check the performance of the proposed model, experiments were run over a dataset with 518 select, 89 update and 77 delete queries (both legal and illegal). A sample query is:

```
select * from request_client, client where id_client = id and id = 'test' AND
ORD(MID((CONCAT(CHAR(52), CHAR(52))), 1, 1)) > 63 AND '1'=1' OR
id_client=1
```

The accuracy reached by the CBRidSQL agent is 99% compared to 90.5% of Bayesian Network, 94.5% de Naive Bayes and 87.2% of linear regressions.

In Figs. 2 and 3 it can be seen the visualization of a suspicious query applying the CMLHL model. The suspicious query (C2) is shown in red, the legal queries (C0) are depicted in blue and the illegal queries (C1) in red. In Fig. 2, the axes represent the variables in low dimensionality. In Fig. 3 it is represented the same information with a surface map. In both images, it can be seen that the suspicious

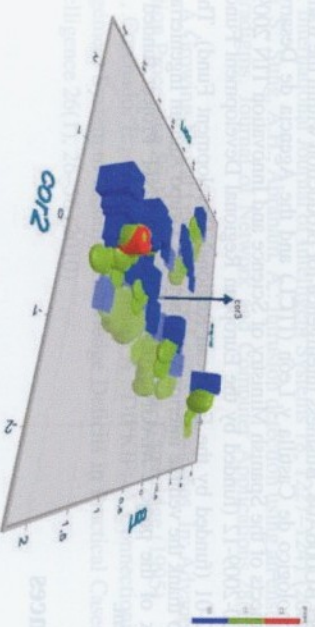


Fig. 2 Revision phase for a suspicious query

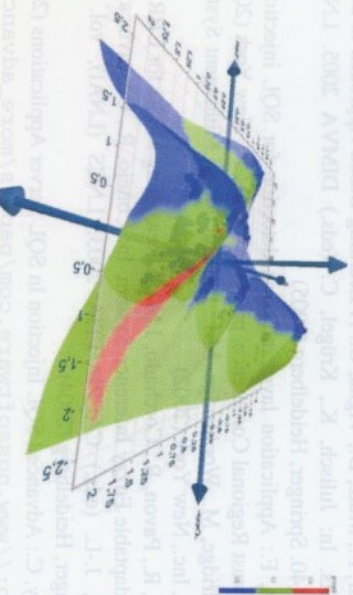


Fig. 3 Revision phase for a suspicious query with a surface map



query is represented with the illegal queries, and therefore the query would be classified as an attack.

SQL injection attacks on databases are a serious threat against information systems. As conclusions of this work, we can state that this paper presents a distributed hierarchical multi-agent architecture incorporating a novel type of agent based on the CBR model specially designed for detecting and blocking such attacks. This CBR agent handles a great adaptation and learning capacities using a CBR mechanism. In addition, it incorporates the prediction capabilities that characterize neural networks and SVM. As a result, an innovative and robust solution is presented allowing a significant reduction of the error rate during the classification of attacks and a different way to tackle SQL injection attacks using a distributed and hierarchical approach.

The empirical results show that the best methods are those that involve the use of neural network and SVM and, if it is considered a mixture, the predictions capabilities are remarkably improved. Besides, the revision phase through a neural visualization eases the classification of suspicious queries.

**Acknowledgments.** This research is partially supported through the Junta de Castilla and León project BU006A08, Business intelligence for production within the framework of the Instituto Tecnológico de Castilla y León (ITCL) and the Agencia de Desarrollo Empresarial (ADE), projects of the Spanish Ministry of Science and Innovation TIN 2009-13839-C03-03, CIT-020000-2009-12 (funded by the European Regional Development Fund) and TIN2010-21272-CO2-01 (funded by the European Regional Development Fund). The authors would also like to thank the vehicle interior manufacturer, Grupo Aniolin Ingeniería S.A., within the framework of the project MAGNO2008 - 1028.- CENIT Project funded by the Spanish Government.

## References

1. Halfond, W., Orso, A.: AMNESIA: Analysis and Monitoring for Neutralizing SQL-injection Attacks. In: 20th IEEE/ACM International Conference on Automated Software Engineering, pp. 174–183. ACM, New York (2005)
2. Valeur, F., Mutz, D., Vigna, G.: A Learning-Based Approach to the Detection of SQL Attacks. In: Julisch, K., Krügel, C. (eds.) DIMVA 2005. LNCS, vol. 3548, pp. 123–140. Springer, Heidelberg (2005)
3. Rietta, F.: Application layer intrusion detection for SQL injection. In: 44th Annual Southeast Regional Conference, pp. 531–536. ACM, New York (2006)
4. Woolridge, M., Woolridge, M.J.: Introduction to Multiagent Systems. John Wiley & Sons, Inc., New York (2002)
5. Laza, R., Pavon, R., Corchado, J.M.: A Reasoning Model for CBR. BDI Agents Using an Adaptable Fuzzy Inference System. In: Conejo, R., Utrera-Vicaya, M., Pérez-de-la-Cruz, J.-L. (eds.) CAEPLATIA 2003. LNCS (LNAI), vol. 3040, pp. 96–106. Springer, Heidelberg (2004)
6. Anley, C.: Advanced SQL Injection In SQL Server Applications (2002), [http://www.ngssoftware.com/papers/more\\_advanced\\_sql\\_injection.pdf](http://www.ngssoftware.com/papers/more_advanced_sql_injection.pdf)

7. Christensen, A.S., Moller, A., Schwartzbach, M.I.: Precise Analysis of String Expressions. In: 10th International Static Analysis Symposium, pp. 1–18. Springer, Heidelberg (2003)
8. Su, Z., Wassermann, G.: The essence of command injection attacks in web applications. In: 33rd Annual Symposium on Principles of Programming Languages, pp. 372–382. ACM Press, New York (2006)
9. Huang, Y., Huang, S., Lin, T., Tsai, C.: Web application security assessment by fault injection and behavior monitoring. In: 12th International Conference on World Wide Web, pp. 148–159. ACM, New York (2003)
10. Ramasubramanian, P., Kannan, A.: Quickprop Neural Network Ensemble Forecasting a Database Intrusion Prediction System. In: 7th International Conference Artificial Intelligence and Soft Computing, Neural Information Processing, vol. 5, pp. 847–852 (2004)
11. Corchado, J.M., Bajo, J., Abraham, A.: GerAnni: Improving Healthcare Delivery in Geriatric Residences. In: Intelligent Systems, vol. 23, pp. 19–25. IEEE, Los Alamitos (2008)
12. Vapnik, V.N.: An overview of statistical learning theory. IEEE Transactions on Neural Networks 10, 988–999 (1999)
13. Corchado, J.M., Laza, R.: Constructing deliberative agents with case-based reasoning technology. International Journal of Intelligent Systems 18, 1227–1241 (2003)
14. Mukkamala, S., Sung, A.H., Abraham, A.: Intrusion detection using an ensemble of intelligent paradigms. Journal of Network and Computer Applications 28(2), 167–182 (2005)
15. Fritzsche, B.: A Growing Neural Gas Network Learns Topologies. In: Advances in Neural Information Processing Systems, vol. 7, MIT Press, Cambridge (1995)
16. Herrero, Á., Corchado, E., Sáiz, L., Abraham, A.: DIPKIP: A Connectionist Knowledge Management System to Identify Knowledge Deficits in Practical Cases. Computational Intelligence 26(1), 26–56 (2010)

## 1 Introduction

Computational intelligence (CI) has been widely used to build Intrusion Detection Systems (IDSs) [1]. MOVICARIDS (Mobile Visualization Connection

Manuel Navarro, Vicente Illán

Departamento de Sistemas Informáticos y Computación,

Universidad Politécnica de Valencia,

Campus de Vera s/n, 46102, Valencia, Spain

e-mail: {navarro, illan}@dsic.upv.es

Enlilén Corchado

Departamento de Informática y Automática,

Universidad de Salamanca, Plaza de la Merced s/n

37008, Salamanca, Spain

e-mail: enlilen@usal.es

Alvaro Herrero

Civil Engineering Department, University of Burgos,

C/Prácticas de Valencia s/n, 09006 Burgos, Spain

e-mail: alh@ccia.uiburg.es