Blocking of SQL Injection Attacks by Comparing Static and Dynamic Queries

Jaskanwal Minhas¹ and Raman Kumar²

¹,² Department of Computer Science and Engineering,
¹ Sant Baba Bhag Singh Institute of Engineering and Technology, Jalandhar, Punjab, India.
² D A V Institute Engineering and Technology, Jalandhar, Punjab, India.
¹,² {er.ramankumar}@aol.in

Abstract — Due to internet expansion web applications have now become a part of everyday life. As a result a number of incidents which exploit web application vulnerabilities are increasing. A large number of these incidents are SQL Injection attacks which are a serious security threat to databases which contain sensitive information, the leakage of which cause a large amount of loss. SQL Injection Attacks occur when an intruder changes the query structure by inserting any malicious input. There are a number of methods available to detect and prevent SQL Injection Attacks. But these are too complex to use. This paper proposes a very simple, effective and time saving technique to detect SQLIAs which uses combined static and dynamic analysis and also defines an attack other than existing classification of SQLIAs.

Index Terms — Dynamic and Static query, SQL query, SQLIAs

I. INTRODUCTION

Due to the easy access to the internet nowadays most of services consist of web services. But the increasing use of internet for important applications poses a greater risk of attacks on the web. SQL-Injection Attack (SQLIA) is not affecting the systems like other attacks, but the ability of SQLIAs to obtain sensitive information, such as military systems, e-business, and banks, etc results into great security risk. The SQLIA occurs when an intruder changes the structure of the query by inserting any SQL keywords. Any secret information can be easily retrieved from database by using these attacks. Variety of techniques are available to detect SQL injection Attacks. Most preferred are Web Framework, Static Analysis, Dynamic Analysis, Combined Static and Dynamic Analysis and Machine Learning Technique. Web Framework[1] provides filters to filter special characters but other attacks are not detected. Static Analysis[2],[3] method checks the input parameter type, but fails to detect attacks with correct input type. Dynamic Analysis[4],[5] technique is capable of scanning vulnerabilities of Web application but is not able to detect all types of SQLIAs. Combined Static and Dynamic Analysis[6]-[9] includes the benefits of both but this method is very complex. Machine learning method[10] can detect all types of attacks but results in number of false positives and negatives. This paper proposes a very simple and effective method to detect SQL Injection Attacks which uses Combined Static and Dynamic Analysis technique. The method is a combination of SQL Query attribute values removal and combined Static and Dynamic Analysis technique. The rest of the paper is organized in the form of different sections. Section 2 describes the categories of SQLIAs. Section 3 discusses the related work. Section 4 explains the proposed method to detect SQLIAs. Section 5 describes the experimental results. Section 6 concludes this paper.

II. CATEGORIES OF SQL INJECTION ATTACKS

There are a number of different methods of SQL Injection Attacks which are serious threat for any database. For a successful SQLIA the attacker should append a syntactically correct command to the original SQL query. Now the following classification of SQLIAs in accordance to the Halfond, Viegas and Orso[11] researches is presented.

A. Tautologies

This type of attack injects SQL tokens to the conditional query statement which always evaluates true. This type of attack is used to bypass authentication control and access to data by exploiting vulnerable input field which use WHERE clause.

"Select * from loan where loan_no = 'L1' and branch = 'aaa' OR '1'=1" 

As the tautology statement (1 = 1) has been added to the query statement so it is always true.

B. Illegal/Logically Incorrect Queries

When any query is rejected, an error message is returned by SQL Oracle engine. This error messages help attacker to find vulnerable parameters in the Database.

C. UNION Query

By using this technique, attackers join SQLIA to safe query by using word UNION and then can get data about other tables from the application. Consider the following example

Select loan_no,branch from loan where loan_no=$loan_no

By injecting the following loan_no value:

$loan_no='L11 UNION select cust_name from customers

We will have the following query:
Select loan_no.branch from loan where loan_no=11
UNION select cust_name from customers which will join the result of the original query with all the customer names.

D. Piggy-backed Queries
In this type of attack, intruders exploit database by using query delimiter, such as ";" by appending extra query to the original query. In this attack database receives and execute a multiple distinct queries. Normally the first query is legitimate query, whereas following queries could be illegitimate. So attacker can inject any command related to SQL to the database. In the following example, attacker inject " PO; drop table branch" into the designation input field instead of logical value. Then the application would produce the query:

Select salary from employee where emp_code='23467' AND designation='PO'; drop table branch

Because of ";" character, database accepts both queries and executes them. The second query is not legitimate and can drop branch table from the database.

E. Stored Procedure
Stored procedure is a part of database that programmer could set an extra abstraction layer on the database. By using stored procedure a user can store its own function according to the need. In stored procedure, a collection of SQL queries are included. As stored procedure could be coded by programmer, so, this is also one of the causes of SQLIA. Depending on specific stored procedure in the database there are number of different ways to attack. Create procedure dbo @emp_code varchar2, @designation varchar2 AS

EXEC("’’+@emp_code+’’ AND designation=’’+ @designation+’’);"
GO

If the intruder adds one more query after the legitimate query, then the normal query is converted into piggy backed query which is a type of SQLIA.

Select * from employee where emp_code='23451' and designation=’PO’; Shutdown;

After the execution of original query the second query which is illegitimate is executed and causes database shut down.

F. Alternate Encoding
In this type of attack the regular strings and characters are converted into hexadecimal, ASCII and Unicode. Because of this the input query is escaped from filter which scans the query for some bad characters which results in SQLIAs i.e. the converted SQLIA is considered as normal query.

G. Inference
By this type of attack, intruders change the behaviour of a database or application. There are two well known attack techniques that are based on inference: blind injection and timing attacks.

• Blind Injection: Sometimes developers hide the error details which help attackers to compromise the database. In this situation attacker face to a generic page provided by developer, instead of an error message. So the SQLIA would be more difficult but not impossible. An attacker can still steal data by asking a series of True/False questions through SQL statements. Consider two possible injections into the login field:

SELECT accounts FROM users WHERE user_name=’abc’ and 1 =0 -- AND password= AND pin=0

SELECT accounts FROM users WHERE user_name=’abc’ and 1 = 1 -- AND password = AND pin=0

If the application is secured, both queries would be unsuccessful, because of input validation. But if there is no input validation, the attacker can try the chance. First the attacker submit the first query and receives an error message because of "1=0". So the attacker does not understand the error is for input validation or for logical error in query. Then the attacker submits the second query which always true. If there is no login error message, then the attacker finds the user_name field vulnerable to injection.

• Timing Attacks: A timing attack lets attacker gather information from a database by observing timing delays in the database's responses. This technique by using if-then statement cause the SQL engine to execute a long running query or a time delay statement depending on the logic injected. This attack is similar to blind injection and attacker can then measure the time the page takes to load to determine if the injected statement is true. WAITFOR is a keyword along the branches, which causes the database to delay its response by a specified time. For example, in this attack the following query is inserted into user_name field:

"‘legalUser’ and ASCII (SUBSTRING ((select top 1 name from sysobjects),(1,1)) > X WAITFOR 5 -- ");"

Following query is produced from this:

SELECT accounts FROM users WHERE user_name="’legalUser’ and ASCII(SUBSTRING((select top 1 name from sysobjects),(1,1)) > X WAITFOR 5 -- ’ AND password=’’ AND pin=0

This query is used to extract a table name from the database. In this attack substring function is used to extract first character of first table’s name. If the ASCII value of character is greater than X, the attacker can get the character by using 5 seconds time delay in the response of database.

III. RELATED WORK
Order to detect and prevent SQLIAs a number of detection methods are available. This section explains the related work.

A. Web Framework
A Web Framework is a software framework that is designed to support the development of dynamic websites, web applications and web services. Some Web Frameworks offer SQL Injection Attack prevention methods, PHP provides magic quotes[1]. Magic quotes is a special feature of PHP language in which special characters ’”,/’,NULL are pre pended with a backslash
before being passed on to detect SQLIAs. But magic quotes support only four special symbols. SQLIAs with other symbols are not detected.

B. Static Analysis

Static Analysis method analyzes the SQL query sentences to detect SQLIAs. This method verifies user’s input type to reduce SQLIAs. JDBC checker[2] validates user input to prevent SQLIAs by using ISA. But if the malicious query has correct syntax and type, attack cannot be detected. Combined Automated Reasoning and Static Analysis method by Wasserman[3] uses FSA to detect SQLIAs. The use of FSA under approximation of the SQL grammar makes this technique too restrictive to remove some possible malicious queries from the represented set. The main problem with all static analysis techniques is that these require source code modification and most of the techniques are just used for web applications written in java.

C. Dynamic Analysis

Dynamic Analysis can locate vulnerabilities of SQLIAs without any source code modification. Paros[4] is a tool written in java used to locate vulnerabilities in web applications. Through its proxy nature all HTTP and HTTPS data between server and client, including cookies and form fields, can be intercepted and modified. Paros is not perfect because it uses predetermined attack codes to scan and uses HTTP response to the success-rate of the attack. Sania[5] finds and collects all the SQL Injection vulnerabilities between web application and database and automatically generate elaborate attacks according to the syntax and semantics. Then the Sania compares the parse tree of intended SQL query and those resulting after attack to assess the safety of vulnerable spots. Due to the use of parse tree Sania is more accurate to detect SQLIAs than any other dynamic analysis technique. Dynamic Analysis methods are useful because no source modification is required. But the vulnerabilities found in web application must be manually fixed by developer and not all of them can be found before predefined attack.

D. Combined Static and Dynamic Analysis

Combined Static and Dynamic Analysis includes the advantages of both Static and Dynamic Analysis techniques. SQLCheck[7] presents the definition of command injection attacks and gives a sound and complete algorithm for preventing SQLIAs based on context-free grammars and compiler parsing techniques. AMNESIA[6] is a model based approach to detect SQLIAs. This technique builds a model of all legitimate queries. Then each dynamic query is compared with this model to detect SQLIAs. Parse Tree[8] is used to compare static SQL query with dynamic SQL query to detect SQLIAs. Wei[9] proposed a technique to detect attacks in stored procedure using Control flow graph.

E. Instruction-Set Randomization

In Instruction set Randomization SQL query is encoded by inserting a random value in the query. SQLrand[12] uses this technique to protect the query from SQLIAs. But this method is not effective if the random key is predicted.

F. Machine Learning

Machine Learning is a technique in which training set i.e. a model is prepared which contain all the legitimate queries belonging to any web application. At run time all the requested queries are compared with queries in the training set to detect SQLIAs. Intrusion Detection System (IDS) by Valeur and colleagues[10] is based on this technique. But this technique results in number of false positives and negatives if the poor training set is used.

G. Prepare Statements

Prepare statements[13] are used in SQL to separate the values in query from the structure of query. In this the skeleton of SQL query is defined and then the values are filled at run time. SQLIA is detected if there is any change in the structure of query. But the main limitation of this method is that the whole web application needs to be modified in order to apply this method.
Blocking of SQL Injection Attacks by Comparing Static and Dynamic Queries

Fig.3. Overview of AMNESIA

H. Taint Based Technique

Java Dynamic Tainting[14] is a tool implemented in java. This tool chases string for taint information instead of character and then sanitizes query strings which are generated using tainted input. But the primary limitation of this tool is that this is not able to detect SQLIAs in numeric fields.

IV. PROPOSED METHOD

In this research work combined static and dynamic analysis technique is used to reduce false positives and false negatives. The static query structure is compared with dynamic query. In this database is maintained to store the valid query structure. These valid queries are also known as static queries. The attribute values of dynamic queries are removed during run time and compared with previously stored static queries having same number of tokens as in dynamic query. The symbols needed in the proposed algorithm are given in Table I.

Consider an example to explain the symbols used in Table I.

SQ: select * from users where user_id=? and password=?

DQ: select * from users where user_id='admin' and password='123456'

DQf: select * from users where user_id='' or '1'='1' and password='123456'

In algorithm 2 and algorithm 3, the function fn, calculates the total number of tokens in static and dynamic queries.

<table>
<thead>
<tr>
<th>Table 1 SQLIA's DETECTION SYMBOLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbole</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>I(sq)</td>
</tr>
<tr>
<td>f</td>
</tr>
<tr>
<td>fn</td>
</tr>
<tr>
<td>SQ</td>
</tr>
<tr>
<td>DQ(t,f)</td>
</tr>
<tr>
<td>DDQ(t,f)</td>
</tr>
<tr>
<td>SQT</td>
</tr>
<tr>
<td>DDQT</td>
</tr>
</tbody>
</table>

Algorithm $f(One\ SQL\ query)$

```
Enumerate\ \ Quotation\_Status=\{\ Quot\_Start,\ Quot\_End\}  
Input\ \ String=One\ SQL\ query;  
Output\_String=Null;  
Do\ \ while(\ \ not\ \ empty\ \ of\ \ Input\ \ String)  
| Char=Get\_Token(Input\_String);  
| If Char is a quotation character and if Char is in between Quot\_Start and Quot\_End  
| Replace Char with ‘?’  
| Else  
| Add Char to Output\_String;  
|}  
Return\ Output\_String;  
```

Algorithm 1: Algorithm which removes attributes values in a SQL Query

Algorithm $f(n(SQ,Total\_Token\_S))$  
//SQ is static query  
//Total\_Token\_S returns the total numbers of tokens in a static query  
Input\ String=SQ;  
Total\_Token\_S=0;  
Do\ while(\ \ not\ \ empty\ \ of\ \ Input\ \ String)\{  
| Char=Get\_Char(Input\_String);  
| If Char is a space character  
| Total\_Token\_S++;  
}

Algorithm 2: Proposed Algorithm for token calculation in static query

Algorithm $fn(SQ,Total\_Token\_S)$

```
//SQ is static query  
//Total\_Token\_S returns the total numbers of tokens in a static query  
Input\ String=SQ;  
Total\_Token\_S=0;  
Do\ while(\ \ not\ \ empty\ \ of\ \ Input\ \ String)\{  
| Char=Get\_Char(Input\_String);  
| If Char is a space character  
| Total\_Token\_S++;  
}
```

Algorithm 3: Proposed Algorithm for token calculation in static query
The following example shows the result of functions \( f \) and \( fn \). \( DQ_1 \) is a normal dynamic query and \( DQ_2 \) is an abnormal dynamic query. 

\[
SQ = \text{select * from users where user_id=? and password=?}
\]

\[
DQ_1 = \text{select * from users where user_id='admin' and password='123456'}
\]

\[
DDQ_1 = f(DQ_1) = \text{select * from users where user_id=? and password=?}
\]

Total number of tokens in dynamic query \( DDQT_1 \) = \( fn(DDQ_1) \) = 8

\[
DQ_2 = \text{select * from users where user_id='') or '1'='1'}\text{—and password='abc'}
\]

\[
DDQ_2 = f(DQ_2) = \text{select * from users where user_id=?--and designation=?}
\]

Total number of tokens in dynamic query \( DDQT_2 \) = \( fn(DDQ_2) \) = 7

Following formula is applied regardless of whether a query is normal or abnormal.

If \( (DDQT=SQT) \) then // if tokens are same

If \( (DDQ=SQ) \) // if queries are same

Then

Result = Normal

Else

Result = Abnormal

After removing attribute values this method locates for static queries having same number of tokens as in dynamic query. Then the dynamic query is compared character by character only with that static queries having same number of tokens. The technique of comparing dynamic queries only with that static queries having same number of tokens improves response time. If match is found requested dynamic query is valid query otherwise it is SQL Injection Attack.

V. Experiment and Evaluation

A. Experimental Result Analysis

This section shows the results of experiment. The main advantage of proposed method is its simplicity. The complexity of the algorithm is divided into two parts-first token calculation and second searching for dynamic query. In token calculation complexity depends upon the database we are using because each database has its own syntax but this token calculation technique is same for all databases. In searching for a query best case occurs when the algorithm finds the query in first comparison and the worst case occurs when the query is not in the list. The complexity for searching is \( O(n) \). But the best part of this method is that the whole list of static queries is divided according to the number of tokens. The dynamic query is
TABLE II Sample Input SQL Database

<table>
<thead>
<tr>
<th>S.No.</th>
<th>SQL Statements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Select * from branch</td>
</tr>
<tr>
<td>2</td>
<td>Select emp_code from employee union select emp_code from branch</td>
</tr>
<tr>
<td>3</td>
<td>Select branch_name from branch where emp_code='20496'</td>
</tr>
<tr>
<td>4</td>
<td>Select emp_code,branch_name,address from branch</td>
</tr>
<tr>
<td>5</td>
<td>Select * from branch where branch_name like 'S%'</td>
</tr>
<tr>
<td>6</td>
<td>Select * from employee where designation='PO'</td>
</tr>
<tr>
<td>7</td>
<td>Select designation,department,salary from employee where emp_code='19961'</td>
</tr>
<tr>
<td>8</td>
<td>Select address from branch where branch_name='STATE BANK OF PATIALA KAPURTHALA'</td>
</tr>
<tr>
<td>9</td>
<td>Select * from employee where salary between '20000' and '50000'</td>
</tr>
<tr>
<td>10</td>
<td>Delete from employee where designation='APRO'</td>
</tr>
<tr>
<td>11</td>
<td>Update employee set designation='PO' where designation='APRO'</td>
</tr>
<tr>
<td>12</td>
<td>Update branch set branch_name='STATE BANK OF PATIALA' where emp-code='20496'</td>
</tr>
<tr>
<td>13</td>
<td>Create or replace procedure abc(Vdesignation in char)is emp_no number;begin select emp_code into emp_no from employee where designation=Vdesignation;insert into branch values(emp_no,'OFFICE OF THE GENERAL MANAGER RAIL COACH FACTORY','RAIL COACH FACTORY KAPURTHALA');end;</td>
</tr>
<tr>
<td>14</td>
<td>Select emp_code,branch_name,address from branch where emp_code in(select emp_code from employee where designation='PO')</td>
</tr>
<tr>
<td>15</td>
<td>Select E.emp_code,designation,department,salary,B.emp_code,branch_name, address from employee E inner join branch B on E.emp_code=B.emp_code where designation='PO'</td>
</tr>
<tr>
<td>16</td>
<td>delete from branch where emp_code='17589'</td>
</tr>
<tr>
<td>17</td>
<td>delete from employee where department='DEPARTMENT OF POST'</td>
</tr>
<tr>
<td>18</td>
<td>Select branch_name from branch where emp_code='20496'</td>
</tr>
<tr>
<td>19</td>
<td>Select E.emp_code,designation,department,salary,B.emp_code,branch_name, address from employee E right join branch B on E.emp_code=B.emp_code where designation='PO'</td>
</tr>
<tr>
<td>20</td>
<td>select emp_code,salary from employee where designation like '_PO%'</td>
</tr>
<tr>
<td>21</td>
<td>select avg(salary) from employee group by department having department='THE MALL KAPURTHALA'</td>
</tr>
<tr>
<td>22</td>
<td>select designation,salary from employee union select designation,salary from employee</td>
</tr>
<tr>
<td>23</td>
<td>select max(salary) from employee group by designation having designation='PO'</td>
</tr>
<tr>
<td>24</td>
<td>Select designation from employee where salary='9200'</td>
</tr>
<tr>
<td>25</td>
<td>Select designation,count(designation) ‘no of designations’ from employee group by designation</td>
</tr>
</tbody>
</table>

TABLE IV Experiment Results

<table>
<thead>
<tr>
<th>Total no. of Static queries</th>
<th>No. of inserted SQL queries</th>
<th>Valid SQL queries</th>
<th>Detected SQL Injection attacks</th>
<th>Total time taken (milliseconds)</th>
<th>Average time per query (milliseconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>255</td>
<td>109</td>
<td>48</td>
<td>61</td>
<td>8580</td>
<td>78.7</td>
</tr>
</tbody>
</table>

TABLE V Accuracy Results

<table>
<thead>
<tr>
<th>Total no. of inserted queries</th>
<th>Total false positives</th>
<th>Total false negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>48(Valid queries)+61(SQLIAs)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

TABLE VI Performance Analysis

<table>
<thead>
<tr>
<th>Total no. of inserted queries</th>
<th>Total time taken by base paper method (Milliseconds)</th>
<th>Total time taken by proposed method (Milliseconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>109</td>
<td>25318</td>
<td>8580</td>
</tr>
</tbody>
</table>

TABLE VII DETECTION AND PREVENTION METHODS OF VARIOUS SQLIAs

<table>
<thead>
<tr>
<th>Symbols: ‘O’ defines that detection and prevention is partially possible</th>
<th>‘N/A’ Not Applicable</th>
<th>‘X’ defines that detection and prevention is impossible</th>
<th>‘ND’ defines that attack is not defined</th>
</tr>
</thead>
</table>
Compared character by character only with that static queries having same number of tokens. This results in less number of comparisons. Table II shows the sample database. Table III shows the sample injected SQL statements. Table IV, Table V and Table VI shows the experimental results.

### B. Comparison of various detection and prevention methods by attack types

Table VII shows the comparison of various detection and prevention methods according to various attack types. Halfond[11] classified SQL Injection attacks into seven major categories: Tautologies, Incorrect queries, Union queries, Piggy-Backed queries, Stored procedures, Inference and Alternate encodings. As shown in table VII one more category of SQLIAs is defined named White Space Manipulation Attack which is not defined in any other detection and prevention method. In this type of attack an attacker can manipulate white spaces to prevent an attack from detection.

### V. CONCLUSIONS & RECOMMENDATIONS

This paper proposes a very simple method to detect SQLIAs which compares static SQL queries with dynamic SQL queries after removing attribute values. To minimize the response time incoming queries are compared character by character only with that static queries having same number of tokens. In this one more attack known as white space manipulation attack other than existing classification of SQLIAs is defined and detected by proposed method. Removing of attribute values makes a SQL query independent of the database. So this method is used for any database. Future work will focus on to detect other types of attacks like cross site scripting attacks.

**ACKNOWLEDGMENT**

The authors also wish to thank many anonymous models.

**REFERENCES**


Ms. Jaskanwal Minhas is working as an Assistant Professor with the Department of Computer Science and Engineering, S B B S Institute of Engineering and Technology, Jalandhar. Before joining S B B S Institute of Engineering and Technology, Jalandhar she did her Bachelor of Technology with honours from R I E T, Phagwara.

Mr. Raman Kumar (er.ramankumar@aol.in) is working as an Assistant Professor with the Department of
Computer Science and Engineering, D A V Institute of Engineering and Technology, Jalandhar. Before joining D A V Institute of Engineering and Technology, Jalandhar, He did his Bachelor of Technology with honours in Computer Science and Engineering from Guru Nanak Dev University; Amritsar (A 5 Star NAAC University). He did his Master of Technology with honours Computer Science and Engineering from Guru Nanak Dev University; Amritsar (A 5 Star NAAC University). His major area of research is Cryptography, Security Engineering and Information security. He has published many papers in refereed journals and conference proceedings on his research areas.