Design and Development of Web-based Expert System to Detect Network Intrusion

Asghaiyer Mohamed Asghaiyer, Purnomo Budi Santoso, Harry Soekotjo Dachlan
(Dept. Communication and Information System, Brawijaya University, Indonesia)

Abstract: It is very important that the security mechanisms of a system are designed so as to prevent unauthorized access to system resources and data. However, completely preventing breaches of security appear, at present, unrealistic. We can, however, try to detect these intrusion attempts so that action may be taken to repair the damage later. Expert system (ES) has been proposed to be the better solution of reasoning on the existing problems on the basis of expert knowledge.

Experts' brain are stored in a database called a knowledge base and then made the rules to give decisions in accordance with the knowledge base. In this research, knowledge base taken from the expert and journal. Expert system will be used to determine or predict the occurrence of intrusion detection systems based on the feature above. By simple means, any data packet is going to be examined by ES, which initially stores knowledge about intrusion characteristics from database as variables. The result of this research is anomaly packet already detected by the application according to the rule based on the expert and the journal.

Keywords: anomaly packet, expert system, intrusion, knowledge base, network packet.

I. INTRODUCTION

In the last three years, networking revolution has finally come of age. More than ever before, we see that the Internet is changing computing as we know it. More over, today, most people are connected to internet in nearly 24 hours a day. The possibilities and opportunities are limitless; unfortunately, so too are the risks and chances of malicious intrusions. A computer system should provide confidentiality, integrity and assurance against denial of service. However, due to increased connectivity (especially on the Internet), and the vast spectrum of financial possibilities that are opening up, more and more systems are subject to attack by intruders. These subversion attempts try to exploit flaws in the operating system as well as in application programs and have resulted in spectacular incidents like the Internet Worm incident of 1988.

Intrusion is something that may harm computers which connected to the network if it is being abandoned. Many kinds of attack comes from the vulnerability of network, of which without proper intrusion detection system. If a network has a good intrusion detection system, it has the capability to prevent any future damage or theft. Each data packets is grabbed by ES to compare its characteristics with the variables (intrusion characteristics), and the result is a decision taken by ES whether to hold the data packet or pass it to client. The scheduling method to grab-compare-decide by expert system should be able to make sure that all data packets are checked without any significant time consumption.

II. LITERATURE REVIEW

2.1 Definition of Expert System

Abraham (2005), stated that the basic components of an expert system are illustrated in Figure 2.1. The knowledge base stores all relevant information, data, rules, cases, and relationships used by the expert system. A knowledge base can combine the knowledge of multiple human experts. A rule is a conditional statement that links given conditions to actions or outcomes. A frame is another approach used to capture and store knowledge in a knowledge base. It relates an object or item to various facts or values. A frame-based representation is ideally suited for object-oriented programming techniques. Expert systems making use of frames to store knowledge are also called frame-based expert systems.

2.2 Knowledge Base

The highly specialized knowledge of the problem area is located in the knowledge base. This module contains the problem facts, rules, concepts, and relationships. The first step to build the knowledge base is to gather the knowledge from human experts, which should be stored in the knowledge base. After this step the knowledge must be coded in a form which is useful for automatic processing, which is subject of the knowledge representation. There are several techniques as to how to store the knowledge. As an example we could look at a rule-based representation. A rule is a if/then structure that logically relates information contained in the ‘if’ part of the rule to information in the ‘then’ part. An example above shows a sample rule-based knowledge representation.
2.3 Working Memory

The working memory contains the facts about a problem that are collected during one consultation of the expert system. When a new problem has to be solved, the user enters information about the problem into the working memory. The expert system then uses this information together with the facts in the knowledge base to infer new facts. These new facts are also stored in the working memory. The content of the working memory can also consist of facts that have been collected from external storage like databases, spreadsheets, or sensors, beside the information that is taken from the knowledge base.

2.4 Inference Engine

In logic, a rule of inference, inference rule, or transformation rule is the act of drawing a conclusion based on the form of premises interpreted as a function which takes premises, analyzes their syntax, and returns a conclusion (or conclusions). For example, the rule of inference modus ponens takes two premises, one in the form of "If p then q" and another in the form of "p" and returns the conclusion "q". The rule is valid with respect to the semantics of classical logic (as well as the semantics of many other non-classical logics), in the sense that if the premises are true (under an interpretation) then so is the conclusion. Typically, a rule of inference preserves truth, a semantic property. In many-valued logic, it preserves a general designation. But a rule of inference's action is purely syntactic, and does not need to preserve any semantic property: any function from sets of formulae to formulae counts as a rule of inference. Usually only rules that are recursive are important; i.e. rules such that there is an effective procedure for determining whether any given formula is the conclusion of a given set of formulae according to the rule.

2.5 Characteristics of Expert Systems

The most important characteristic of any expert system is that the knowledge is separated from control logic. This distinguishes an expert systems from a conventional program, where it is often not possible to add knowledge without changing the source code of the program at the same time. The tasks of modifying and maintaining the system are easier with this separation. Another important feature of an expert system is that it embodies the expertise of a human expert. The knowledge of such an expert is used as well as his problem-solving skills. Because expert systems focuses on knowledge of a clearly defined domain they poorly perform outside their area of expertise. At each time in an interaction between the system and the user it should be possible for the user to ask the system how it has reached the actual conclusion. This is important because an expert system does not have a well-defined structure like a conventional program that makes it easy to understand what the system is doing. Lets look at an example:

*Expert System: The battery is bad.*
*Person: How?*
*Expert System: Since your car won’t start,*
*I assumed there was a problem with the electrical system.*
*Once I found the battery voltage was below 10 volts,*
*I knew that the battery was bad.*

An expert system can also explain why it is asking a specific question. The explanation of the system makes the user feel more comfortable because he can reproduce why the system is asking this question. Lets look at an example:

*Expert System: Will the car not start?*
*Person: Why?*
*Expert System: If I know that the car won’t start,*
*then I usually assume the problem is in the electrical system.*

2.6 Building Expert Systems

The development of an expert system is a highly iterative process. Most of the time it is not possible to design first the whole system and then just implement the design. Instead it is needed to build a first prototype, then use this prototype to get more knowledge from the human experts and re-implement the prototype. This process is repeated until there is a final system that is able to do what the users are expecting it to do.

2.6.1 Knowledge Engineering

One major part is the Knowledge Engineering. Usually there are human experts who know much about their specific domain, and application developers who have a lot of knowledge in programming but don’t know much about the domain where the expert system is used in the future. In the process of Knowledge Engineering both groups of people are involved and try to find an efficient way as to how to encode the human experts knowledge in the expert system.
2.6.1.1 Assessment
During the assessment phase, studies are made to determine the feasibility and justification of the candidate problem. After this process the overall goals of the project have to be defined. These goals together with the information gathered by the process of Knowledge Engineering are then used to identify the sources of needed knowledge.

2.6.1.2 Knowledge Acquisition
Knowledge Acquisition is the process of acquiring, organizing, and studying the knowledge from human experts. In the early stage of expert system development the objective of this step is to uncover key concepts and general problem-solving methods used by the expert. In a later stage the results of the testing are used to explore for more detailed information. Knowledge Acquisition has long been recognized as the bottleneck in expert system development.

2.6.1.3 Design
After the knowledge acquisition, the design phase is used to define the overall structure and organization of the system’s knowledge. Methods are defined for processing the knowledge and the appropriate software tool is chosen. After this step an initial prototype system is built, which serves as the focal point for further interviews with the human experts. As mentioned above this task is repeated several times until a final design is reached.

2.6.1.4 Testing
The Testing phase is not a separate task. It is rather a continuous process throughout the whole project. After every step in building the expert system, it is tested and new knowledge is added to it. These tests should not only be done by the programmers, but also by the end-users because it is very important that the end-users are able to use the system and that the system is well adapted to the user’s needs.

2.6.1.5 Documentation
The Documentation addresses the need to compile all of the project’s information into a document. The Documentation must also support the knowledge engineer during the development of the system. It should contain a knowledge dictionary that provides a well organized representation of the system’s knowledge and the included problem solving procedures.

2.6.1.6 Maintenance
Because of the highly iterative process in developing an expert system, it is also important to periodically maintain the system. The system’s knowledge may need to be refined or updated to adapt the system to the actual circumstances. If the knowledge is separated from the control, this process is more flexible and easier as if the whole knowledge is included in the logic.

2.6.2 Knowledge Representation
Knowledge is represented in some symbolic form that can be manipulated by the expert system and by the users of the expert system. There exist many different forms of knowledge representation. The knowledge engineer must choose the knowledge representation technique best suited for the given application. Some of the most important types of knowledge are presented in the following subsection.

2.6.2.1 Types of Knowledge
Knowledge can be divide into the following types:

a. Procedural knowledge describes how a problem is solved, e.g. rules, strategies, and procedures.
b. Declarative knowledge describes what is known about a problem, e.g. concepts and objects.
c. Meta-knowledge describes knowledge about the knowledge.
d. Heuristic knowledge describes a rule-of-thumb.
e. Structural knowledge describes knowledge structures, e.g. rule sets and relationships of concept.

It is important to use an appropriate type of knowledge for a specific expert system application, because only a well adjusted structure is able to support effective problem solving. The next subsection describes different knowledge representation techniques.

2.7 Rules
In logic, a rule of inference, inference rule, or transformation rule is the act of drawing a conclusion based on the form of premises interpreted as a function which takes premises, analyzes their syntax, and returns a conclusion (or conclusions). For example, the rule of inference modus ponens takes two premises, one in the
form of "If $p$ then $q$" and another in the form of "$p" and returns the conclusion "$q$". The rule is valid with respect to the semantics of classical logic (as well as the semantics of many other non-classical logics), in the sense that if the premises are true (under an interpretation) then so is the conclusion.

### 2.8 Forward Chaining

Forward chaining starts with the available data and uses inference rules to extract more data (from an end user for example) until a goal is reached. An inference engine using forward chaining searches the inference rules until it finds one where the antecedent (If clause) is known to be true. When found it can conclude, or infer, the consequent (Then clause), resulting in the addition of new information to its data.

### III. METHODOLOGY

In this research, rule-based system is chosen because it has the advantages of using less time to perform the detection. The fast detection can be performed due to that the computer is not ‘thinking’ about another algorithm than just passing data packets to its rules, checking each data packet to sample of network intrusion, of which stored in memory, and lastly, decision is made after a data packet is done checked.

#### 3.1 Application Flowchart

Figure 3.1 is the flowchart for the network intrusion detection using expert system.

![Figure 3.1 NIDS using Expert System](image)

Figure 3.1 shows process that are performed in the server, when it started, the application will initially load all necessary information from database. The information including Knowledge Base, Rule sets, and other configurations related to run the Expert System to detect Network Intrusion. The knowledge base contains specific feature of different network instrusion behaviour, which taken from MySQL Database’s record about knowledge base and stored as variables of the web application, meanwhile the rule set is the rules that must be passed by realtime data packets. The rule sets are also taken from MySQL Database’s record about rule sets and stored to structure the application. These information are kept in Memory as variables or class’ objects with its properties, in order to minimize delay during runtime process. After initialization has completed, now server mode is in runtime where it continuously collects data packets, then extracts its features. The extracted features is then analyzed by using Expert system that resides in the memory during the initial process. And any result of data packet process is reported as threats or safe data. Reports are stored in the database for future analysis, display to user, and follow ups, either to block similar data packets or not.

When a threat can not be identified by the expert system, the application will automatically put the type of packet data into the label "warning" with the type of intrusion is “blank”.

www.iiosrjournals.org
3.2 User Design Interface

User interface need to be developed to ensure that system can be clearly understand the usage and appearance.

3.2.1 Login System

![Login System](image)

Figure 3.2 Login System

3.2.2 View Packet Data

![Packet Data Viewer](image)

Figure 3.3 Packet Data Viewer

3.3 Database Design

In this study, used 1 database and 2 tables for storing information associated with the log results in real-time network packet capture and store data about patterns of information from each feature IDS. Information stored in the database can be changed later if needed.

This study uses a MySQL database created by the organization opensource and free for both individual and commercial use. Figure 3.4 is a view of the MySQL database that already has a database with the name "ids".

![MySQL Database](image)

Figure 3.4 Display MySQL database containing ids database

From Figure 3.4 shows that the database "ids" has 2 tables, namely the log table and the table pattern. The following will explain in detail the contents of the log table and chart pattern.

<table>
<thead>
<tr>
<th>Table 3.1 The structure of table “log”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field Name</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>Id_packet</td>
</tr>
<tr>
<td>packet time</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Expert System in IDS, By : P. Budi, Harry, Ghazali
University of Brawijaya
Log table has several fields, namely id_packet which serves to store the unique identity of each row of data, field data from the packet containing the captured packets by application, and field time is the time when the packet is received in the database.

To create a design table like Table 3.1, done at the command prompt windows system, as for the creation of application code table can be seen in Figure 3.5.

```
CREATE TABLE `log` (
    `id_packet` INT(3) NOT NULL AUTO_INCREMENT,
    `packet` LONGBLOB NOT NULL,
    `time` TIMESTAMP NOT NULL DEFAULT CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP,
    PRIMARY KEY (`id_packet`),
) ENGINE = MYISAM DEFAULT CHARSET = latin1;
```

Figure 3.5 Source code of Creating table “log”

<table>
<thead>
<tr>
<th>Keyname</th>
<th>Type</th>
<th>Unique</th>
<th>Packed</th>
<th>Column</th>
<th>Cardinality</th>
<th>Collation</th>
<th>Null</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRIMARY</td>
<td>BTREE</td>
<td>Yes</td>
<td>No</td>
<td>id_pattern</td>
<td>14</td>
<td>A</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

Space usage:
- Data: 2,536 B
- Index: 2,048 B
- Total: 4,584 B

Row Statistics:
- Format: dynamic
- Rows: 14
- Row length: 181
- Row size: 327 B
- Next autoindex: 19
- Creation: Dec 27, 2013 at 03:58 PM
- Last update: Dec 28, 2013 at 10:30 AM

As for storing the information about the list and the type of intrusion, made table pattern. This table contains information about the type and pattern of the intrusion, table pattern can be seen in the table below 3.3.

Table 3.3 The structure of table “pattern”

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Length/Values</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>id_pattern</td>
<td>Integer</td>
<td>11</td>
<td>Primary Key</td>
</tr>
<tr>
<td>category</td>
<td>Text</td>
<td></td>
<td>Category of intrusion Name of intrusion</td>
</tr>
<tr>
<td>name</td>
<td>Text</td>
<td></td>
<td>Information about pattern</td>
</tr>
<tr>
<td>pattern</td>
<td>Text</td>
<td></td>
<td>Filename of Application</td>
</tr>
<tr>
<td>status</td>
<td>Text</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To create a design table like Table 3.3, done at the command prompt windows system, as for the creation of application code table can be seen in Figure 3.6.

```
CREATE TABLE `pattern` (
    `id_pattern` INT(11) NOT NULL AUTO_INCREMENT,
    `category` TEXT NOT NULL,
    `name` TEXT NOT NULL,
    `pattern` TEXT NOT NULL,
    `status` VARCHAR(100) NOT NULL,
    PRIMARY KEY (`id_pattern`),
) ENGINE = MYISAM DEFAULT CHARSET = latin1;
```

Figure 3.6 Source code of Creating table “pattern”

Table pattern consists of several fields, namely: id_pattern that serves as a unique identifier of each row of data, category field which will store data about the categories of a intrusion which does include a probe or a denial of service or other. name field contains information about the name intrusion, filed pattern is complete information about the pattern of intrusion which according to its category, and the last field is the actual status of containing the file name of the application that corresponds to the existing categories.

Table 3.4 Index of table “pattern”

<table>
<thead>
<tr>
<th>Keyname</th>
<th>Type</th>
<th>Unique</th>
<th>Packed</th>
<th>Column</th>
<th>Cardinality</th>
<th>Collation</th>
<th>Null</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRIMARY</td>
<td>BTREE</td>
<td>Yes</td>
<td>No</td>
<td>id_pattern</td>
<td>14</td>
<td>A</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>
Design and Development of Web-based Expert System to Detect Network Intrusion

IV. IMPLEMENTATION AND RESULT

The tools used in this implementation are php designer and Xampp. Php designer is an IDE for php developer to arrange coding of programming. Xampp is a application bundling that consists of php, mysql, and mail server. MySql is a database engine that stored knowledge base and inference engine logic. These are implementation results of design interface:

4.1 Top menu / home menu

The main menu contains some information, including title of research, researcher name, and displays hyperlinks to other menus like the menu for editing the knowledge base, running applications and the menu contains information about the application.

4.2 Maintaining the Rule

This menu provides an easy way to manage rule base on the application. Admin can provide additional information network pattern behaviour, information and also the category that are used.

4.3 Inputting the Facts

4.3.1 Realtime Packet Data

In this section, application provides an easy way to the admin to import realtime packet data.
4.3.2 Pattern Management

In this section, application provides an easy way to the admin to make changes to the pattern, category of pattern, and information about related pattern.

Figure 4.3 Realtime Packet Captured

Figure 4.4 Pattern Management

Figure 4.5 List of Pattern
4.4 Running the Expert System

This menu is the center of this research, namely to make the system running and try the application of Expert System working on IDS.

Figure 4.5 Running the Expert System Home

V. CONCLUSION

The use of expert systems in the field of intrusion detection systems are still not widely used in the field of computer networks, but some researchers have tried but it still does not give good results. Design expert system has been created and validated by experts in the field of computer networks. After the design has been created and feasible for use, the next activity is to implement the design into a computer program and generate an expert system application in accordance with the purpose of research. Testing and validation conducted by researchers and accompanied by experts, experts from the field of computer networking is the center of information technology. The results obtained do not vary much with the results of calculations performed by the computer.

REFERENCES


