On using wavelets for detecting attacks to web-based applications

Cristian Cappo, Christian Schaerer
Polytechnic Department- National University of Asuncion
Campus UNA, San Lorenzo, Paraguay
email : (ccappo,cschaerer)@pol.una.py

Raul Ceretta Nunes
Electronic and Computer Science Department, Federal University of Santa Maria,
Technologic Center, Campus UFSM, Santa Maria, RS, Brazil
E-mail: ceretta@inf.ufsm.br

ABSTRACT

Detection attacks to web-based applications have recently received considerable attention, specially intrusion detection system (IDS) for use with HTTP. This work present the use of signal techniques to detect abnormal behavior in query send to web servers by application users. More specifically, we use wavelet transform to detecting anomalous length attribute in queries and then to the applications installed on the servers.

Web-based application use the Internet or intranet infrastructure to interact with their clients. The communications protocol used for this is the Hipertext Transport Protocol (HTTP).

Problem

The usual scenario of attack to web-based application is shown in the figure 1. The anomalous query that pretends violate the security of system is usually more long of normal query request.

The data model

Our detection analyze the HTTP request logged by most common web server. In our experiments we use only GET request. The data model used is proposed in [5] and is shown in the figure 2.

The wavelet transform

We apply the Haar wavelet transform in order to capture the scale variations in the length of the value of each parameter. The remarkable difference in the length of value indicates one possible attack (figure 3).

The process

The detection process is divided into two phases: the training phase, that characterizes the benign data, and the detection phase, that determine if a query is valid or an attack.

The following algorithm summarizes the procedure for detecting outliers in the query through Haar wavelets coefficients:

1. For each web-app server and different parameter of query obtain the length of value
   \[ X = (X_1, X_2, \ldots, X_n) \]
2. Apply discrete Haar wavelet transform to \( X \) in 2 levels
3. For two levels find the max coefficient \( d_{i,j} \) and save as the threshold
   \[ TL = \max \{ \text{coefficients of level} \} \]

Detection phase

1. For each web-app server and different parameter of query obtain the length of value
   \[ X = (X_1, X_2, \ldots, X_n) \]
2. Apply discrete Haar wavelet transform to \( X \) in 2 levels
3. For two levels compare the coefficient \( d_{i,j} \) with a threshold
   \[ |d_{i,j}| > 1.1 \times \max (TL) \]

Experiments

We apply the above detection procedure to query sent to several web-based application installed on the server of Polytechnic Department. The results show (see the table) that our approach is able to detect anomalous queries.

<table>
<thead>
<tr>
<th>Application</th>
<th>Number of queries</th>
<th>Number of attack</th>
<th>Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.pol.una.py/page.php">www.pol.una.py/page.php</a></td>
<td>760709</td>
<td>428</td>
<td>255</td>
</tr>
<tr>
<td><a href="http://www.pol.una.py/galeria.php">www.pol.una.py/galeria.php</a></td>
<td>20387</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td><a href="http://www.pol.una.py/educa/help.php">www.pol.una.py/educa/help.php</a></td>
<td>2554</td>
<td>23</td>
<td>18</td>
</tr>
<tr>
<td><a href="http://www.pol.una.py/educa/calendar/view.php">www.pol.una.py/educa/calendar/view.php</a></td>
<td>6315</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

Conclusion and future work

The proposed model has shown that the application of wavelet is a valid approach for the detection of attacks on web applications. These results are initials and we are currently working to incorporate the attribute character distribution on the model.

References