5.1 Introduction

This chapter discusses the implementation of the context based PDS Monitoring System in Instant messengers using Domain ontology and CBA. This is one of its kind application where in efforts have been made to mine the context of the Instant messages being exchanged for detecting the phishing words. As per the survey even though the earlier efforts were dynamic and possessed domain ontologies still they were content based only.

The PDS monitoring system was developed in Java Platform - Standard Edition 8 [41] and Oracle Database 11g Express edition [42]. The developed application is not a full-fledged Instant messaging system, but possesses the basic chatting facility of chat rooms, where in chatters login. The various chatting sessions discussed in the section 4.2 have been demonstrated here through the screen shots in section 5.2. Also the resulting phishing alerts for various phishing scenarios will also be demonstrated. Performance evaluation of the results will be discussed in the next section 5.3.
5.2 Screen shots

In the following figures depicting screen shots of chat sessions it is assumed that Chatter-1 i.e. phisher intends to extract deceitfully the confidential information from the Chatter-2. Here the implementation of the various scenarios discussed in the Section 4.2 are discussed.

![Figure 5.1 Screen shot Login Screen](image)

The Figure 5.1 shows the screen shot for the Login screens for both the chatters. When they login correctly the different types of Chat rooms are displayed. They are Web designing, ASP, Java, and JSP chat rooms, as shown in Figure 5.2. Chatters from the same chat rooms can chat, the assumption taken here is that both the chatters log into the same room for chatting. The two usernames of chatters are Syed and Admin. The chatter Syed is the phisher here trying to extract personal information from Admin.
Figure 5.2 Screen shot Chat Room selection by the Chatters

Figure 5.3 Screen shot Chat Initiation between the Chatters
After the chatting is initiated in Figure 5.3, based on prior acquaintance and trust gained the phisher tries to elicit personal information in the above Figure 5.4. The phishing words identified are favorite food, favorite past time, favorite teacher and lucky no. To these phishing queries the unsuspecting Chatter-2 has provided the requested details. Figure 5.5 shows the chat session with more clarity.
Figure 5.5 Screen shot Chat Session 1 detail with Phishing alerts

Figure 5.6 Screen shot Chat Session 2 with Phishing Alert
In the separate chat session shown in Figure 5.6 and 5.7, the phisher tries to extract confidential banking account details like the bank name, Internet banking, and requests for account creation tips. To these phishing queries the victim Chatter-2 gives the required information. But the colour indicators of the Phishing alerts have already highlighted the phishing words in RED colour dynamically at the same time.

In the next scenario shown in the Figure 5.8, the phisher tries to extract the password hints from the Chatter-2, which are promptly alerted to Chatter-2.

Figure 5.9 shows a normal chatting session between the chatters wherein the PDS system raises an alert for a harmless word accounting package (the reason
being the word accounting is stemmed to account, which is a strong phishing word belonging to the domain financial gain.)
Figure 5.10  Screen shot Chat Session 5 between chatters

Figure 5.11  Screen shot Chat Session 5 details
In the figures 5.10 and 5.11 a screen shots of normal chatting sessions are demonstrated, wherein, a chatter expresses his fondness for chocolates to such an extent that he uses the word “*kill*” to prove it. The PDS monitoring system evaluates the word *kill* according to its ontology domain and context using the modified OBIE module algorithms as a Harmless word, as chocolate falls under the *eatables* domain in the ontology and not under *Life threatening* domain, as discussed in Chapter 4. Therefore the colour indicator for the word *kill* is BLACK and it is a normal word.

The Figure 5.12 shows the screen shot of the Log file image which depicts the various logs generated during the chatting session. The log section belongs to the chat session shown in Figure 5.9 showing stemming operation of the word *accounting*.

![Screen shot of Log file depicting the background processes](image-url)
The Figure 5.13 shows the chatters logging out of their chat sessions.

5.3 Performance

The context oriented PDS Monitoring System has been tested on 110 transactions between the Chatters using domain ontology and generating improved rules through CBA to help detect Phishing words and generating ALERTS to the victim chatter.

5.3.1 Results and Observations:

Classifier performance is usually measured by accuracy, the percentage of correct predictions over the total number of predictions made. For this
Precision and Recall are used:

\[
\text{precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}
\]

i.e. the number of correctly predicted phish words.

\[
\text{recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}
\]

i.e. the total number of possible predicted phish words.

The total number of true positives (correctly identified phishing words) obtained out of 110 transactions were found to be 97.

The number of false positives (phishing words wrongly detected) out of 110 transactions were found to be 6.

Lastly the number of false negatives (phishing words which could not be detected) out of 110 transactions were 4.

There were 3 suspicious words detected which actually were phishing words.

\[
\text{Precision} = \frac{97}{97+6} = 94.17 \%
\]

\[
\text{Recall} = \frac{97}{97+4} = 96.03 \%
\]
5.4 Comparative Study

A comparative study of the existing three phishing detection systems in Instant messengers is shown in Table 5.1.

- Deceptive Phishing Detection System in Instant messengers using Data mining approach [28].
- Framework for surveillance of instant messages in instant messengers and social networking sites using data mining and ontology [33].
- Context oriented PDS for IM using domain ontology and CBA (the present work).

The comparative study shows that the CBA has a much better performance in terms of both precision and recall with respect to the existing systems mentioned above. Also the key feature being the context which provides the dynamism and intelligence element which provides an upper edge in the comparison.
Comparative study of the phishing detection systems:

<table>
<thead>
<tr>
<th>Phishing detection Method</th>
<th>Domain Ontology</th>
<th>Based on</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deceptive Phishing Detection System in IM based on ARM [28]</td>
<td>No</td>
<td>Content</td>
<td>80.72%</td>
<td>76.69%</td>
</tr>
<tr>
<td>Framework for surveillance of instant messages in Instant Messengers and social networking sites using data mining and ontology [33]</td>
<td>Yes</td>
<td>Content</td>
<td>85.67%</td>
<td>84.36%</td>
</tr>
<tr>
<td>Context oriented PDS monitoring system based on Domain ontology and CBA in IMS</td>
<td>Yes</td>
<td>Context</td>
<td>94.17%</td>
<td>96.03%</td>
</tr>
</tbody>
</table>

Table 5.1 Comparative Analysis