An Ontology based Enhanced Framework for Instant Messages Filtering for Detection of Cyber Crimes

Ms Ankita M. Shendurkar
ME, CSE (Scholar)
GHRCEM, Dept of Computer Science,
SGBAU Amravati University,
Amravati (MH), India
ashendurkar@gmail.com

Prof. Nitin R. chopde
Asst Professor, HOD
GHRCEM, Dept of Computer Science,
SGBAU Amravati University,
Amravati (MH), India
nitin.chopde@raisoni.net

***Abstract*—** Instant messaging is very appealing and relatively new class of social interaction. Instant Messengers (IMs) and Social Networking Sites (SNS) may contain messages which are capable of causing harm, which are untraced, leading to obstruction for network communication and cyber security. User ignorance towards the use of communication services like Instant Messengers, emails, websites, social networks etc, is creating favorable conditions for cyber threat activity. It is required to create technical awareness in users by educating them to create a suspicious detection application which would generate alerts for the user so that suspicious messages are not ignored. Very limited research contributions were available in for detection of suspicious cyber threat activity in IM. A context based, dynamic and intelligent suspicious detection methodology in IMs is proposed, to analyze and detect cyber threat activity in Instant Messages with relevance to domain ontology (OBIE) and utilizes the Association rule mining for generating rules and alerting the victims, also analyses results with high ratio of precision and recall. The results have proved improvisation over the existing methods by showing the increased percentage of precision and recall.

Keywords- : *Instant Messengers; Social Networking Site; Ontology based Information Extraction; Association Rule Mining*.

*****

I. INTRODUCTION

One of the developments in Web technology is instant messaging, the process of instant communication between people. Instant Messaging (IM) has been stepping into the workplace as well as people's daily life at remarkable speed. More and more people, of all ages, have started using instant messaging software and the numbers are continuing to grow Instant messengers (IMs) have become an integral part of today's state of the art communication system. The instant messaging system generally follows the client-server model. Communication between two or more people based on typed text messages, over the internet or Local Area Network (LAN).

IM differs from several other web applications because of its real-time nature of user interactions, e.g. on-line presence notification and instant messages. Nowadays, it's difficult to survive without IMS as users are enthusiastic about it. Trillions of messages are sent daily through Instant Messenger and Social Networking Site. Instant Messengers (IMs) and Social Networking Sites (SNS) may contain harmful and suspicious messages, which are untraced, leading to obstruction of network communication and cyber security.

Organized crimes have adopted on-line chatting technique to send these suspicious messages. An answer to this current back draw is to discover suspicious messages from the written messages. Our Contribution includes upgrading the existing IMS using data mining technique of Associative rules, Ontology based information retrieval technique (probabilistic models), that is guided with pre-defined Knowledge based rules and ARM. Early detection of suspicious messages from instant messaging systems (Mobile Phone, IM and SNS) is fessible with our proposed Framework to spot and predict the sort of cyber threat activity and trace the criminal details. Data mining approaches were used to find the frequent patterns and extract predefined knowledge base rules on a larger scale in IMs but however, the detected suspicious words had larger percentage of false positives and false negatives. This was obvious because the entire detection process was content based and it did not to detect the threat activity pertaining to that domain. Although in some IM systems, Association rule mining was used beside domain ontologies, which were applied right from information extraction and mapping them with pre-defined knowledge base rules to identify the domains, still it could not detect the suspicious words intelligently, as the main intention or *context* behind its usage could not be identified. However, its performance was slightly better than those not using domain ontologies [4]. Avoiding ambiguity over the relevance of the identified suspicious words is the main motive behind this work in order to enhance the instant messaging system performance. This could be done by identifying the *context* behind chatting messages in order to reduce the percentage of both false positive and true negatives. The extracted domain and its context if mined together can generate more interesting rules by using Association rules, in order to generate instant alerts dynamically for the victim, with the best possible performance i.e. maximum true positives percentage [4].

The paper is organized into 6 sections. Section 2 discusses the related work of prevalent threat problems in Instant Messengers, the methodologies and approaches used to address the problem statement. Section 3 elaborates the architecture and design of the methodology and discusses the architecture of OBIE for domain and context extraction using the Triplet Algorithms. Section 4 represents the work flow of the system. Section 5 outlines the experimental setup to be established in the form of chat session transactions between chatters of IMs and addresses the implementation and testing of the suspicious words detection system in IMs and evaluates the performance in terms of precision and recall. Section 6 summarizes and highlights the contributions of this research work in precise. It also provides directions for future enhancements in this research area.

II. LITERATURE SURVEY

This section presents related literature concerning instant messaging frameworks,

Natural Language Processing (NLP) is that the processed approach to analyzing text that is based on both a set of theories and collections of technologies, and being an awfully active space of study and development, there is not a single agreed-upon definition that which may satisfy everyone, however there are some aspects, which would be part of any knowledgeable person's definition. Natural Language Processing (NLP) techniques play a crucial role in overcoming the shortcomings mentioned above, by giving the advantage of utilizing the semantics of the messages exchanged between chatters in instant messaging. The Semantic Web uses the ontology and could incorporate the NLP to extract triplets i.e. subject, predicate and the object of the message under process, as a part of Ontology Based Information Extraction (OBIE). Throughout the analysis study it has been observed that NLP has not been utilized to its full potential and at times it has only been used as an alternative to information extraction [4]**.**

Database mining has recently attracted tremendous quantity of attention within the database research because of its presence in many areas, including call support, market strategy and financial forecast. For discovering database mining an incremental updating technique is developed for efficient maintenance of the association rules when new transaction data are added to a transaction database. However, it's nontrivial to take care of such discovered rules in large databases because a database might enable frequent or occasional updates and such updates might not solely invalidate some existing strong association rules but also turn some weak rules into Strong ones [3].

In 2012, Mohd Mahmood Ali, and Lakshmi Rajamani has presented a paper, “APD: ARM Deceptive Phishing Detector System Phishing Detection in Instant Messengers Using Data Mining Approach” In this paper, authors had proposed, an Anti-phishing system APD system that dynamically traces out any potential phishing attacks when messages exchanged between clients of an IM System. In this paper, authors had presented an association rule mining technique (Apriori algorithm) to detect Deceptive Phishing. Suspicious messages sent using Instant Messenger between two or more or different clients/chatters stored in Transaction Database(TDB) using Information retrieval system technique(stemming, ignore words(IGWDB)) the frequent reoccurring words are extracted from the TDB dynamically using Association rule mining technique and stored in transaction pattern database(TPDB). The framework in this paper, try to detect deceptive phishing for messages in text format, which also includes encrypted patterns (ASCII codes) (Mohd Mahmood Ali, And Lakshmi Rajamani, 2012).

In 2010, Daya C. Wimalasuriya and Dejing Dou presented a paper, “Ontology-based information extraction: An introduction and a survey of current approaches”, the authors have explained an introduction to ontology-based information extraction and review the details of different OBIE systems developed so far. In this paper, they reviewed the field of OBIE and a number of systems that are categorized under it. Among other things, they have provided a definition for the field, identified a common architecture for the systems and classified the existing systems along different dimensions. Information extraction (IE) aims to retrieve certain types of information from natural language text by processing them automatically. Ontology-based information extraction (OBIE) has recently emerged as a subfield of information extraction. Here, ontology's which provide formal and explicit specifications of conceptualizations - play a crucial role in the IE process. Because of the use of ontology's, this field is related to knowledge representation and has the potential to assist the development of the Semantic Web [2].

In 2010, Michael Robertson, Yin Pan, and Bo Yuan suggested a paper, “A Social Approach to Security: Using Social Networks to Help Detect Malicious Web Content”, This paper presents a comprehensive method combining traditional security heuristics with social networking data to aid in the detection of malicious web content as it propagates through the user's network. The approach in this paper try to detect malicious web content with the help of heuristics based on social networking data. These approaches are tested successfully by using Face book account. The experimental results have shown very promising results, which predicts the presence of malicious content in the URLs.

The proposed decision tree based classification method is an incremental and user-feedback based extension of a decision tree induction algorithm named Ad Infinitum which show that Ad Infinitum algorithm is a good choice for threatening e-mail detection as it runs fast on large and high dimensional databases, is easy to tune and is highly accurate, outperforming popular algorithms such as Decision Trees, Support Vector Machines and Naive Bayes. [5]

All of these techniques are inefficient to predict such types of activity. Pre-defined knowledge base rules have been considered which are to be mapped with the ontology generator through NLP, to obtain the domain of the suspected word. The user is alerted the e-crime department is notified if any suspicious activity is detected. The overall detection is content based and thus in spite of using the NLP, the *context* part is not explored, the use of ontology is to information extraction and domain identification only. The overall performance does not reduce the false positives and false negatives.

III. PROPOSED ALGORITHM

As our main area of discussion is to detect the suspicious words from Instant Messenger. For that I have developed an ontology based enhanced framework for detection of cyber crime by using data mining and Information extraction technique. In this system framework will include instant messengers as the web service, TDB for suspicious words, a pre-processing data module for extraction of action word using Natural Language Processing. Then, Ontology Based Information Extraction Technique which will be used for tracking the behavior of user based on current, previous and further messages. Such framework will help to trace criminal activity dynamically.

A. **Implementation of IMs as web service**

In this module, a website is developed which will contain an instant messages that can send and receive messages from other users. Instant messaging, a social-based communication medium provides us with extensive social network information concerning its users. Instant Messaging (IM) has been placed into workplace as well as people's daily life at remarkable speed. Instant messaging is a method for real-time communication over the Internet. The instant messaging system generally follows the client-server model. Communication between clients occurs either via a server, or a server brokers it

B. **Collection of database of suspicious words**

In this module, we are collecting a data set of suspicious word which will be used for further processing. The experimental results are shown when tested with dataset collected from Global Terrorism Database (GTD).

C. **Preprocessing of data using Natural Language Processing (NLP)**

In this module, following steps will be performed for NLP. In the first steps Part -Of –Speech (POS) tagging. In this module, the input messages will be tagged according to the POS like Noun, Pronoun, and Verbs etc. English is often a very frustrating language to study when it is not your first language. Even identifying the "part of speech" for a word is complicated. For example, most verbs in English can also be nouns; you can "run" to the store, but you can also go out for a run, or ride a bobsled down a run. For NLP to understand a word like "run" it must be able to understand and interpret the surrounding terms and understand the context in which the word is used. In the second step Chunking Technique is applied. Where only the action words will be extracted and words which do not constitute any action or meaning to the message i.e stop words will be removing from the message.

D. **Ontology Based Extraction Technique (OBIE)**

In this module, the users messaging behavior will be tracked and studied so that we can conclude the behavior of the user based on his current, previous and further messages.

E. **Rule based generation**

In this module, we will be developing a top K-rule algorithm for developing the association rules so that we get the rules for particular threat activity. The Association Classification rules for the three input components. First we apply rules to the user's IP address; we check if the user is malicious, based on the following criteria,

1. If the ontology suggests.
2. If the words are not proper.
3. If the user is reported malicious previously.

These are the rules for the user. They are calculated on per user basis and are based on the entries done by the user in the system.

IV. SUSPICIOUS MESSAGE DETECTION FRAMEWORK

In this Section we explore the operational phases of Framework, shown in Fig. The Suspicious Pattern Detection (SPD) algorithm initiate the steps to capture the instant messages that are sent between the clients/users and stores them into database for identifying suspicious messages using Ontology based Information Extraction (OBIE) technique. The following steps are performed by the system:

Step I: Sign up and Login

Firstly User has to sign up to the Instant Messaging Application by his username and password. User has to log in into application by inserting user name and password.

Step II: Analyze Input Message

System read the message from the input string.

Step III: Filtering of Message

Now the filtering of message is done by the following process: firstly it checks for the suspicious words from the dataset, and then it checks the messages by NLP, and removes the non action words and precedes the message.

Step IV: Ontology

After this, also check the type of words found in ontology, here words are comparing from safe ontology to suspicious ontology. If the safe counts are less than suspicious counts then it will consider being suspicious ontology.

Step V: User's messaging Behavior

Now check if the user has already sent suspicious message to the user. If this user has sent suspicious messages to more than 1 users previously. Then user will be in the list of suspicious user.

Step VI. Checking for the number of malicious activities

If the numbers of suspicious activities are more than 2 then mark it as a malicious user and block that user. After that he unable to send the message to the proper user.

Step VII: Detected Suspicious Message

Then the proper user gets the message by malicious user that the suspicious message is removed by the system. He will alert for next malicious users message.

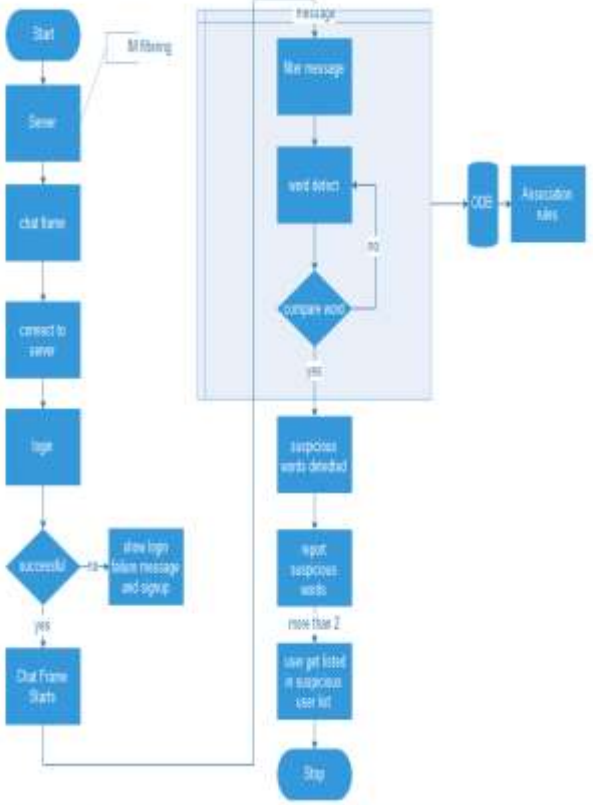

FIGURE 1: FLOW CHART OF THE SYSTEM

V. SIMULATION RESULTS

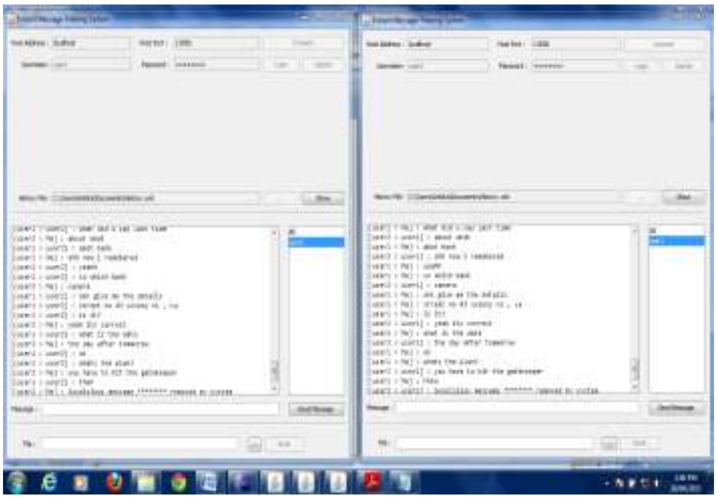

Figure 2: Sample screenshot 1

In figure 1 illustrates the sample screenshot in which two users are communicating with others. From 2300 transactions showing some transaction which is detected by the instant messaging filter system.

In figure 2 illustrates the sample screenshot among 2300 transactions which detect the suspicious message and removed from the system. The proper user gets the message by malicious user that the suspicious message is removed by the system. He will alert for next malicious users message.

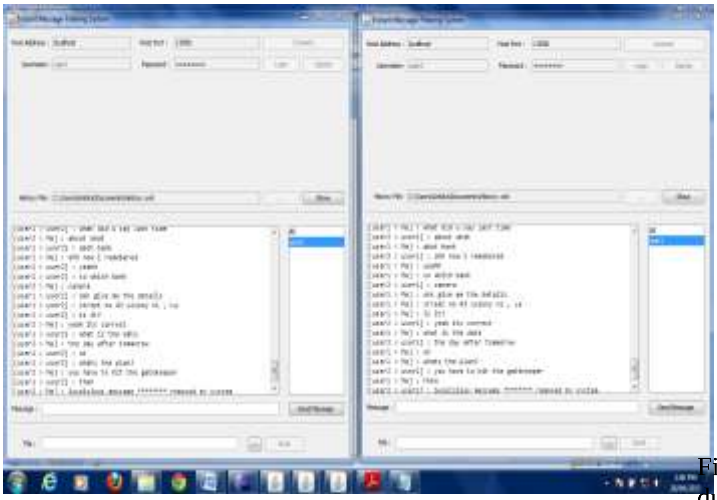

Figure 3: Sample screenshot 2

The System has been tested on 2300 transactions between the Chatters using domain ontology and generating improved rules through association rules to help suspicious words and generating ALERTS to the victim chatter.

- Check for how many times a proper person is detected as malicious, mark it as TP.
- Check for how many times an improper person is detected as improper and mark it as FN.
- Check for how many times a improper person is detected as proper, mark it as TN.
- Check for how many times a proper person is detected as improper, mark it as FP.

Terms	Framework outputs
Total True Positive	1903
Total True Negative	82
Total False negative	31
Precision	95.86
Recall	98.39

Table 1 Outputs obtained through TP, TN, FP, and FN

The total number of true positives obtained out of 2300 transactions was found to be 1903. The total numbers of true Negatives obtained out of 2300 transactions were found to be 82. The total number of False Negative obtained out of 2300 transactions was found to be 31. These precision and recall measure are calculated on the basis of True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN).

The generated values are mentioning above in the table no 1.

Precision = TP/(TP+TN)
=1903/(1903+82)=95.86%

Recall = TP/(TP/FN)
=1903/(1903+31)=98.39%

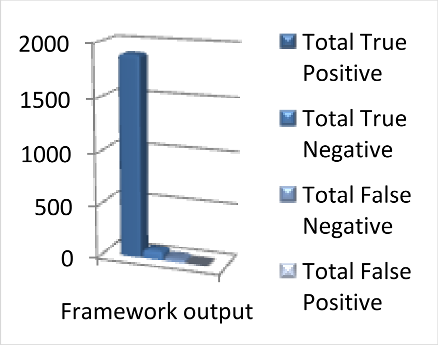

Figure 4 Graphical Representation of Table no 1

Figure 4 this graph presents the number of word detected during chat sessions, using our proposed framework. These detected words are categories as TP, TN, FN and FP values with the respective values 1903, 82, 31 and 0. With the help of these values, we are able to generate the recall and precision measure.

Figure 5 this graph presents the real scenario of the chat sessions which we performed during result analysis phase, which is based on the true detection rate of the system i.e. accuracy. These precision and recall values help to us to determine efficacy of our frameworks. These values are calculated using TP, TN, FN and FP values.

VI. COMPARISON OF PROPOSED FRAMEWORK WITH EXISTING FRAMEWORK

The comparative study based on the domain ontology has a much better performance in terms of both precision and recall with respect to the existing systems which is Framework for surveillance of instant messages in IMs and social networking sites using data mining and ontology. Also the key feature being the context which provides the dynamism and intelligence element which provides an upper edge in the comparison. The comparative analysis is based on the measures precision and recall. Here the main parameter is accuracy in which true detection rate of the system shows the higher the better.

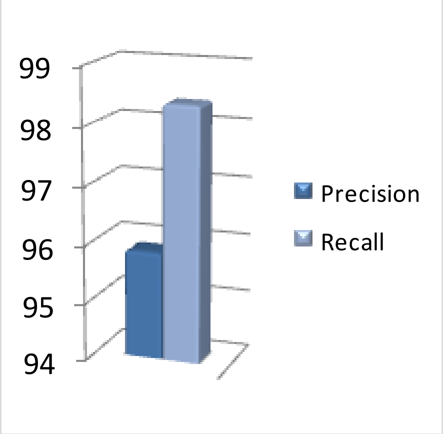

Figure 5 Graphical Representation of Table no 1

Table 2 shows the existing framework outputs in which Correctly Extracted, Total extracted correctly and Total Possible words Extracted are retrieved for calculating the measure Precision and recall.

Terms	Framework Outputs
Total Extracted Correctly	1779
Correctly Extracted	1703
Total possible words Extracted	1732
Precision	95.72
Recall	98.32

Table .2: Existing Framework Outputs

Precision = Correctly Extracted/Total extracted correctly*100

= 1703/1779*100

=95.72%

Recall = Correctly Extracted/ Total Possible words Extracted*100

=1703/1732*100

=98.35%

Terms	Framework Outputs
Total Extracted Correctly	1985
Correctly Extracted	1903
Total possible words Extracted	1934
Precision	95.86
Recall	98.39

Table .3: Proposed frame work outputs

Table 3 shows the proposed framework outputs in which Correctly Extracted, Total extracted correctly and Total Possible words Extracted are retrieved for calculating the measure Precision and recall. Here when we compare the result analysis of proposed framework outputs are higher than the existing framework outputs. As for true detection rate the output is higher the better.

Precision =Correctly Extracted/Total extracted correctly*100

= 1903/1985*100

=95.86

Recall = Correctly Extracted/ Total Possible words Extracted*100 =1903/1934*100 = 98.39

VII. CONCLUSION AND FUTURE WORK

Currently none of Instant Massagers and Social Networking Site and Mobile Phone (Apps) has ability to detect suspicious message during online chat. Framework aids the t to identify suspicious words from cyber messages and trace the suspected culprit. In existing system, they simply used POS Tagging. Due to which the speed of the system is not superior, in the proposed system we overcome this by applying the chunking technique which improves the speed of the system and shows the proper and accurate output. In future we would be developing this for Mobile System with Voice Input.

English is not the only medium of communication in a sub continent like India, it might be of multi lingual nature, the concept of translating and applying suspicious detection is a future challenge. Words from other languages might be written in English which cannot be identified by any ontology based tool and may be ignored which in turn may turn out to be a suspicious message. Images, sounds and videos are the formats which need a lot of work other than text formats where the vulnerability towards suspicious may always is present

The following issues and challenges are therefore identified where lot of research activities should be concentrated.

- Deceptive suspicious messages are sent in any format other than textual (Images, Audio, Video), then they are not detected.
- Rules lack multilingual support for detection of suspicious words.
- Issue with the interpretation of a message written in multiple languages.
- If the suspicious messages are encrypted, we need to detect using decryption techniques.

ACKNOWLEDGEMENT

I would like to thank to all the people those who have help me to give the knowledge about these research papers and I thankful to my guide with whose guidance I would have completed my research paper and make it to published, finally I like to thank to all the website and IEEE paper which I have gone through and have refer to create my review paper successful.

REFERENCES

[1] Mohammed Mahmood Ali, Khaja Moizuddin Mohammed, Lakshmi Rajamani, “Framework for Surveillance of Instant Messages in Instant messengers and Social networking sites using Data Mining and Ontology,” Proceeding of the IEEE Students' Technology Symposium, 2014.

[2] Daya C. Wimalasuriya, and Dejing Dou,“Ontology-Based Information Extraction”An Introduction and a Survey of Current Approaches,” Journal of Information Science”, Volume 36, No. 3, pp. 306-323, 2010.

[3] Wang Wei, Payam Barnaghi, and Andrzej Bargiela, “Probabilistic Topic Models for Learning terminological ontologies,” published by IEEE Tran., on Knowledge and data engineering, vol 22, no. 7 in july, 2010.

[4] Mohammad S. Qaseem and A. Govardhan, “Phishing Detection in Instant messanger Using Domain Ontology and classification based association rule An Innovative Rule Generation Approach”, International Journal of Information Sciences and Techniques (IJIST) Vol.4, November 2014.

[5] David W. Cheung, and et al., “Maintenance of discovered association rules in large databases: an incremental updating technique,” published by IEEE in 1996.

[6] Appavu, and et al,“Data mining based intelligent analysis of threatening e-mail,” published by Elsevier in knowledge-based systems in 2009.

[7] Michael Robertson, Yin Pan, and Bo Yuan, “A Social Approach to Security: Using Social Networks to Help Detect Malicious Web Content,” published by IEEE in 2010.

[8] Sunitha Ramanujam, and et al., “A Relational Wrapper for RDF Reification,” E. Ferrari et al. (Eds.): TM 2009, IFIP AICT 300, pp. 196– 214, IFIP International Federation for Information Processing 2009.

[9] Tong Zhang, Fred Damerau , David Johnson, "Text Chunking based on a Generalization of Winnow," Journal of Machine Learning Research , 615-637,2002.

[10] M.W.Du, and S.C.Chang, "An Approach to Designing Very Fast Approximate String Matching algorithms," IEEE journal, 1994.

[11] Jer Lang Hong, "Data Extraction for Deep Web Using Word Net," published by IEEE Transactions on systems, man and cybernetics, 2011.

[12] H. Cunningham, Information Extraction, Automatic, "Encyclopedia of Language and Linguistics," second edition Elsevier Science, 2005.

[13] Saravanan Suba, Chistopher.T, PhD., "A Study on Milestones of Association Rule Mining Algorithms in Large Databases," published by International Journal of Computer Applications (0975 – 888) Volume 47– No.3, June 2012.

[14] Mohd. Mahmood Ali1, Mohd. S. Qaseem2, Lakshmi Rajamani3, A. Govardhan, "Extracting useful rules through improved decision tree induction using information entropy," International Journal of Information Sciences and Techniques (IJIST) Vol.3, No.1, January 2013.

[15] Zhijun Liu, Weili Lin, Na Li, and David Lee, "Detecting and Filtering Instant Messaging Spam A Global and Personalized Approach," published by IEEE , 2005.

[16] Fadi Thabtah, Peter Cowling, Y onghong Peng, "MCAR: Multi-class Classification," published by, IEEE 2005.

[17] Raghu Anantharangachar1, Srinivasan Ramani, S Rajagopalan, "Ontology Guided Information Extraction from Unstructured Text," International Journal of Web & Semantic Technology (IJWesT) Vol.4, No.1, January 2013.

[18] S Parija, S Acharya, "The Process of Information Extraction through Natural Language Processing," International Journal of Logic and Computation (IJLP), Volume (1): Issue (1) 2010.

[19] Sotiris Kotsiantis, Dimitris Kanellopoulos, "Association Rules Mining: A Recent Overview" Published by GESTS International Transactions on Computer Science and Engineering, Vol.32 (1), 2006.

[20] Tipawan Silwattananusarn1 and Assoc.Prof. Dr. KulthidaTuamsuk2, "Data Mining and Its Applications for Knowledge Management : A Literature Review from 2007 to 2012," International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.2, No.5, September 2012.

[21] Reza Soltanpoor, Mehran Mohsenzadeh, Morteza Mohaqeqi," A new approach for better document retrieval," First International Conference on Integrated Intelligent Computing, 2010.

[22] Siddharth Shah, N. C. Chauhan,S. D. Bhanderi "Incremental Mining of association rule: A survey," International Journal of Computer Science and Information Technologies, Vol. 3 (3) , 2012,4071-4074.

BIOGRAPHY

Miss Ankita M. Shendurkar is a Research Assistant in the Computer Science Department, G. H. Raisoni, College of Engineering and Management, Amravati University. She received Bachelor of Engineering degree in 2013 from PRMITR, Badnera, MS, India. Her research interests are Data Mining, Algorithms, etc.

Prof. Nitin R. Chopde is working as Head of Department at Computer Science Dept., G. H. Raisoni, College of Engineering and Management, Amravati University. He received Masters of Engineering degree from Sipna College of Engineering and Technology, Amravati, MS, India. His research interests are Networking, Cloud Computing, Data Mining, Algorithms, etc.