International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 26 Issue 4 – JUNE 2019.

# Instant Messenger Surveillance Framework

Mary K G[#1] and Chandan M. N[*2]

[#] *Master of Computer Application, MCA, PESCE, Mandya, India*

[*] *Master of Computer Application, MCA, PESCE, Mandya, India*

***Abstract*—Through Social Networking Sites (SNS) and Instant Messengers Incalculable dreadful and suspicious messages are sent. These untracked messages lead to the obstruction of network communication and cyber security.**

**To uncover and anticipate the messages that are sent using IM or SNS like Facebook, Twitter, LinkedIn and others we propose a framework. Under this surveillance the instant messages are put to identify the type of uncertain cyber threat activity by offender along with their personal details. The development of framework is using the techniques like the Ontology based Information Extraction technique (OBIE), Association rule mining (ARM) a data mining technique with set of pre-defined Knowledge-based rules (logical),that are used for the process of decision making lettered from domain specialists and the learning experiences of the past suspicious datasets like GTD(Global Terrorist Database). To eliminate such cyber-crimes the exploratory results obtained will uphold.**

***Keywords-Instant Messengers (IM); Social Networking Sites (SNS); Ontology based Information Extraction; Association Rule Mining (ARM); Knowledge based rules.***

## I. INTRODUCTION

The evolvement of internet has led to the progression of innumerable cybercrimes. The suspicious messages sent through cell phones, Instant Messengers and Social Networking Sites were regulated by Hoodlums, which is hard to follow their violation powerfully. The E-crime division must be renovated with the furtherance of revolution to recognize the culprit. The cutout for sending messages, videos and audios were compacted by a considerable number of Instant Messaging Systems (IMS). They are not all prepared to perceive online suspicious messages.

The cybercrime activities are increasing grade by grade. The household and remote insight data are constructively gathered by the FBI, CIA and other government offices in order to avoid future digital assaults. In 2012 the Internet Crime Complaint Center (IC3) let out the report of cybercrimes with
the latest statistics and patterns of on line crime. We reviewed extraordinary models of Mobile Phones, Instant couriers and Social Networking locales. Another framework is built by all these examinations. WordNet, is a lexical database, the instantaneous messages positioned away in TDB (Text Database) checks and separates the words that are precious for our exam. The WordNet carries a huge measure of facts comprising of words and is applied as highlights for arrangement of phrases from unstructured content. Our Contribution consists of improving the contemporary IMS utilizing facts mining strategy of Associative standards, Ontology primarily based information restoration system (probabilistic fashions), that is guided with pre-characterized Knowledge based guidelines and ARM. Early recognition of suspicious messages from texting frameworks (Mobile Phone, IM and SNS) is conceivable with our proposed Framework to differentiate and expect the sort of virtual threat motion and observe the criminal subtleties.

## II. PROBLEM STATEMENT AND RELATED WORK

Nowadays the Instant Messaging Service (IMS) users are addicted such that thy are not able to live without it. Through IMS and emails trillions of messages are being sent. The way of communication with friends, companions and business colleagues have changed due to popular IMS like AOL, MSN, ICQ, Yahoo, Google etc. The investigator identifies and suspects the violation activities by understanding the gesture behind the relationships between the culprits. The famous Instant Messaging Systems are finding their way onto the handheld devices and cell phones which were limited to desktops permitting the users chat literally from anywhere.

The Social approach to detect malicious web content material for Facebook, with safety heuristics is limited to become aware of malicious URL links. To identify the criminal's behavior recently the Facebook static messages are scanned. The suspicious emails that are detected from static messages using decision tree induction are dependent on the highest entropy which identifies the messages are ambiguous or non-ambiguous.

## III. OVERVIEW OF PROPOSED SYSTEM

In this section we explore the operational phase of proposed Framework as shown in Fig. 1. The Suspicious Pattern Detection (SPD) algorithm initiates the steps to capture the instant messages that are communicated between the users and then, stores them into database for identifying suspicious messages.

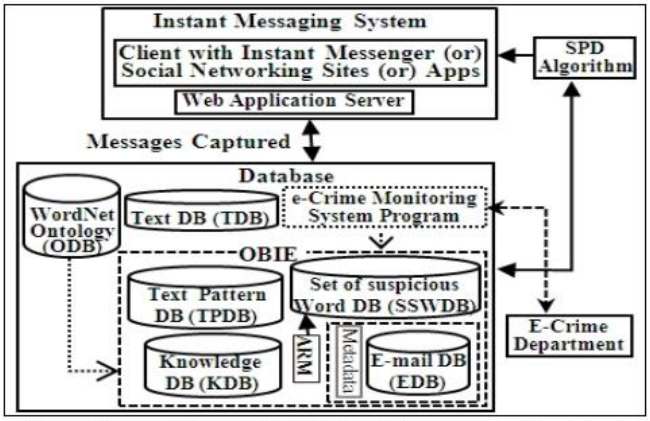


Fig. 1. Proposed Framework to detect suspicious messages from Instant Messaging Systems (IMS).

To store dynamic messages this framework uses database and Ontology Based Information Extraction techniques in order to obtain suspicious words from messages that are guided with the Pre-defined Knowledge based rules that are checked with ARM. The major tasks that are performed are: 1) Extraction of the word from the unstructured text. 2) E-crime monitoring system. 3) SPD Algorithm. The working flow of the framework is depicted in Fig.2 Steps of algorithm are illustrated as follows:

1. Firstly the messages are filtered to extract the unwanted words; the suspicious words are recognized using algorithm in this process.
2. The words that are found to be suspicious are marked as suspicious based on the pre-defined knowledge-based rules that are monitored with WordNet (ODB). Then it is checked with Association Rule Mining (ARM). The KDB maintains the detected stem words along with the domain to detect the undetected words. The Email_Id from and to which the suspicious words belong is identified by checking the metadata. And the relevant information is gathered.

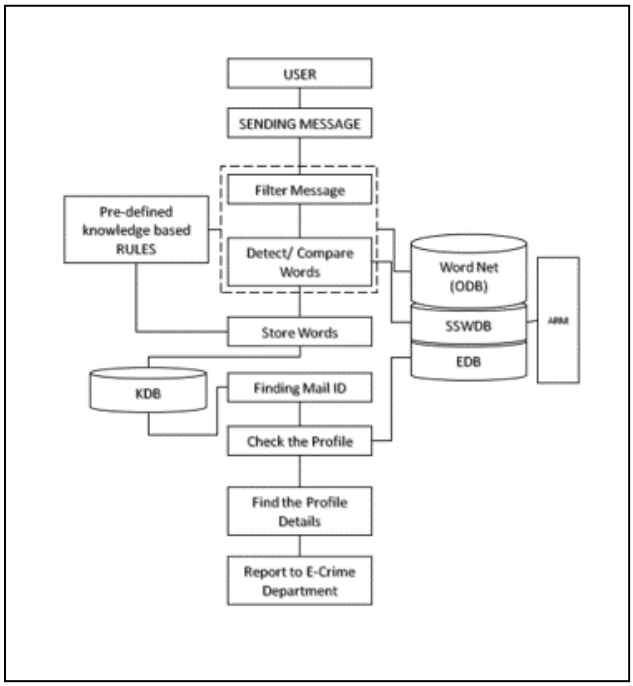


Fig. 2. Schematic cum algorithmic representation for proposed Framework for tracing criminals from instant text messages named as SPD algorithm.

1.Browsing the Email-id Database the profile details like Phone Number, Contact details, Company details, Age and other relevant information are extracted which are provided during the creation of email account, with the aid of Relational Wrapper Algorithm.
2. The e-crime monitoring system tracks the computer (IP -address), ISP and location details of the email-id account who sent the suspicious messages and the report is generated.
3.The generated report consists of type of cyber threat activity performed by the offenders in cyberspace, email profile details and other relevant information.
4. The action to enforce an inquiry on report will be taken by the E-Crime department under the E-Crime act. The domain to which the suspicious words belong is been predicted and mapped by the OBIE.
For this, we are using the database like (TDB, TPDB, ODB, SSWDB, KDB and Metadata).

This framework is mainly based on Suspicious Pattern Detection (SPD) Algorithm. When suspicious messages are found the algorithm stores the text messages in TDB until the culprits are found by providing detailed report from KDB(Knowledge Database) and EDB(EmailId Database) to ECrime Department.

Algorithm SPD(TDB)
Input: Instant Messages from Instant Messenger Surveillance Framework stored in TDB.
Output: Detailed report of suspicious message sender and message to the ECrime Department.

1. Push the messages to TDB.
2. Scan TDB to match the suspicious pattern. Store these patterns in TPDB and map with SSWDB using the OBIE (Ontology Based Information Extraction).
3. Compare patterns of TPDB with patterns of SSWDB {
4. If TPDB==SSWDB. Then push patterns into KDB
   Else
   Do Nothing
5. EndIf
6. } Until TDB!=NULL
7. If TPDB==KDB THEN
8. Check for patterns in KDB.
9. If KDB=’TRUE’ then
10. Check for EDB
11. Report to ECrime Department.

## IV. FUTURE WORK

English is not the only medium of communication in a sub-continent like India, it might be of multilingual 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.
1.Deceptive suspicious messages are sent in any format other than textual (Images, audio, video), then they are not detected.

2.Rules lack multilingual support for detection of suspicious words.
3.Issue with the interpretation of a message written in multiple language.
4.If the suspicious messages are encrypted, we need to detect using decryption techniques.

## V. CONCLUSION

Framework aids the E-crime department to identify suspicious words from cyber messages and trace the suspected culprits. Currently existing Instant Messengers and social Networking Sites lack these features of capturing significant suspicious patterns of threat activity from dynamic messages and find relationships among people, places and things during online chat as criminals have adapted to it. The User Generated Content (UGC) testbed is proven to be useful, for monitoring terror and suspicious crimes in cyberspace which provides national and international security.
We used simple English terms like kill, murder, etc. But, in practical scenarios these words are in specific coding language, for example “picnic” is used instead of “kill”.

## REFERENCES

[1] (2012).[Online].Available:http://www.fbi.gov/sandiego/press-releases /2012/IC3-2011-Internet-crime-report-released.
[2] 3GPP2 partners, “Short Message Service Over IMS: third Generation Partnership project 2”, developed under 3GPP2, published in 2007.
[3] (2012). [Online]. Available: http://www.ontologyportal.org/.
[4] David W. Cheung, and et al., ”Maintenance of discovered association rules in large databases: An incremental updating technique, “ published by IEEE in 1996.
[5] M. Mahmood Ali, and L. Rajamani, “Framework for surveillance of Instant Messages, “ published by InderScience, IJITST, vol.5,2013.
[6] S. Chen, B. Mulgrew, and P. M. Grant, “A clustering technique for digital communications channel equalization using radial basis function networks,” *IEEE Trans. on Neural Networks*, vol. 4, pp. 570-578, July 1993.
[7] J. U. Duncombe, “Infrared navigation—Part I: An assessment of feasibility,” *IEEE Trans. Electron Devices*, vol. ED-11, pp. 34-39, Jan. 1959.
[8] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, and M. Miller, “Rotation, scale, and translation resilient public watermarking for images,” *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 767-782, May 2001.
[9] (2012). [Online). Available: http://www.digitaltrends.com/social-media/facebook-scans-chats-and -comments-looking-for-criminal-behavior/.
[10] Appavu, and et al., ”DataMining based Intelligent Analysis of threatening e-mail,” published by Elsevier in Knowledge-based systems in 2009.

**Mary K.G** received her Bachelor’s degree in Computer Applications from Mangalore University, India and she is currently pursuing MCA in VTU, India. Her current research areas include Data Mining and Ontology.

**Chandan M.N** obtained his MCA from VTU in 2017. He is working as an Assistant Professor in department of MCA, PES College Of Engineering, Mandya, India.