International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Review Article

A Review on Various Approaches for Detection of Cyber Crimes via Instant Messaging Filters

Ankita M.Shendurkar†* and Nitin R.Chopde†

†G.H.R.C.E.M Department of Computer Science & Amravati University, India

Accepted 27 Jan 2015, Available online 01 Feb 2015, **Vol.5, No.1 (Feb 2015)**

Abstract

Instant messaging is popular and relatively new form of social interaction. Instant Messengers (IMs) and Social Networking Sites (SNS) may contain harmful and suspicious messages, which are untraced, leading to hindrance for network communication and cyber security. In this paper, we provided an introduction to various frameworks and review the details of different systems developed so far. Here we attempted to analysis various frameworks and classify them based on different factors, which leads to a better understanding on their operation. We also discuss the implementation details of these systems including the tools used by various authors and the metrics used to measure their performance.

Keywords: *SNS, IMS, NLP, Information Extraction, Suspicious message.*

1. Introduction

A. Instant Messaging system

This section provides an overview of Instant Messaging System, with the rapid *growth* of the *internet*, security has become a major concern. In the present day world, people are so much habituated to Social Networks. Currently, with Instant Messenger (IM) and Social Networking Site (SNS) trillions of instant messages are sent through social media. Taking advantage of this case criminal adapted to send suspicious messages via mobile phones, Instant Messengers and Social Networking Sites. Tracing such criminal activity dynamically is a difficult task. In this paper, we surveyed various architectures of Mobile Phones, Instant messengers and Social Networking sites. These studies helped us to develop a new Framework for scanning and filtering the text messages. TDB (Text Database) Word Net, is a lexical database, which includes words that is useful for our study. It is used as features for classification of words from unstructured text. The ultimate objective is to improve existing IMS system using data mining technique of Associative rules, Ontology based information retrieval technique (probabilistic models), which 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 possible with such Framework which identifies and predicts the type of cyber threat activity and traces the criminal details.

B. Plan of Paper

Many researchers have developed techniques to detect suspicious and malicious web content, here, Section 2 presents literature concerning instant message framework and work done till date, and Section 3 explains the problem statement and related discussion. Section 4 shows the comparative analysis of important approaches and 5 concludes the paper.

2. Literature Survey

This section presents related literature concerning instant messaging frameworks,

In 2014, Mohammed Mahmood Ali, Khaja Moizuddin Mohammed, Lakshmi Rajamani proposed a paper on, "Framework for Surveillance of Instant Messages in Instant messengers and Social networking sites using Data Mining and Ontology", they have proposed a framework further these messages are put under surveillance along with detail of culprits.

Innumerable terror and suspicious messages are sent through Instant Messengers (IM) and Social Networking Sites (SNS) which are untraced, leading to hindrance for network communications and cyber security. Authors has proposed a Framework that discover and predict such messages that are sent using IM or SNS like Facebook, Twitter, LinkedIn, and others. Further, these instant messages are put under surveillance that identifies the type of suspected cyber

*Corresponding author **Ankita M.Shendurkar** is a M.E. (CSE) Scholar and **Nitin R.Chopde** is working as Assistant Professor

Table 1 Comparison

Sr. No.	Title of the paper	Authors	Approach	Result
1	**Framework For Surveillance Of Instant Messages In Instant Messengers And Social Networking Sites Using Data Mining And Ontology.**	Mahmood Ali, Khaja, Lakshmi Rajamani	Ontology Based Information Exaction Technique With Predefined Knowledge Based Rule Checked With ARM.	A Framework Which Predict Suspicious Message, Further These Messages Are Put Under Surveillance Along With Detail Of Culprits.
2	**Phishing Detection In Instant Messengers Using Data Mining Approach**	Mahmood Ali,Lakshmi Rajamani	Association Rule Mining Technique (Apriori Algorithm) To Detect Deceptive Phishing.	An Anti-Phishing System APD That Dynamically Traces Out Any Potential Phishing Attacks When Messages Exchanged Between Clients Of An IM System.
3	**A Social Approach To Security: Using Social Networks To Help Detect Malicious Web Content.**	Robertson, Yin Pan, And Bo Yuan	A Comprehensive Method Combining Traditional Security Heuristics With Social Networking Data To Aid In The Detection Of Malicious Web Content	Detects Malicious Web Content With The Help Of Heuristics Based On Social Networking Data.
4	**Ontology-Based Information Extraction: An Introduction And A Survey Of Current Approaches.**	Daya C. Wimalasuriya, Dejing Dou	An Introduction To Ontology-Based Information Extraction And Review The Details Of Different OBIE Systems	Aims To Retrieve Certain Type Of Information From NL Text By Processing Them.

threat activity by culprit along with their personnel details.

The Framework proposed in this paper will try to identify the type of cyber attack using Ontology based Information Extraction technique (OBIE), Association rule mining (ARM) a data mining technique with set of pre-defined Knowledge-based rules (logical), for decision making process that are learned from domain experts and past learning experiences of suspicious dataset like GTD (Global Terrorist Database) (Mohammed Mahmood Ali, Khaja Moizuddin Mohammed, Lakshmi Rajamani,2014).

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, 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 Facebook account. The experimental results has shown very promising results, which predicts the presence of malicious content in the URLs.

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.

3. Problem Statement and Discussion

The Internet has transformed our lives. It offers tremendous opportunities to share, connect, through Instant Massagers and Social Networking Site. Internet evolution led to growth of innumerable cybercrimes. Currently existing Instant Messengers and Social

Networking Sites lack these features of capturing significant suspicious patterns of threat activity from dynamic messages and find online chat, as criminals have adapted to it.

Currently existing Instant Messengers and Social Networking Sites lack these features of capturing significant Suspicious patterns of threat activity. To overcome this problem, we are proposing a better framework for instant messages filtering for detection of cyber crime. Organized crimes have adopted online chatting technique to send these suspicious messages as these systems have all the facilities and could serve as platform to spread across their information widely through socio-engineered and general text messages. A solution to this problem is to detect suspicious messages from the typed messages. Further these messages are put under surveillance then details of culprit are traced and reported to the E-Crime department with predictable type of threat activity. In this paper we try to figure out, a common better framework which ensures a proper filtering of instant messages.

The 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.

4. Comparisons

In the previous sections, we have define various technique to detect malicious web content and suspicious text messages which were predictable type of threat activity by using OBIE technique with pre defined knowledge based rule checked with ARM. A filtering of instant messages, which OBIE plays crucial role that predicts and maps the domain to which this suspicious word belongs. Therefore, among many methods we will comment on some important methods, see the comparative table (Table 1).

Conclusions

This survey has shown that it is in fact possible to detect malicious web content with the help a good Framework architecture. In this paper we explored and successfully reviewed various instant messaging frameworks. We studied various efficient frameworks and then discussed a system whose purpose was to detect and trace suspicious message dynamically.

References

Mohammed Mahmood Ali, Khaja Moizuddin Mohammed, Lakshmi Rajamani (2014), Framework For Surveillance Of Instant Messages In Instant Messengers And Social Neworking Sites Using Data Mining And Ontology, *Proceeding Of The IEEE Students' Technology Symposium.*

Mohd Mahmood Ali, And Lakshmi Rajamani (2012), APD: ARM Deceptive Phishing Detector SystemPhishing Detection In Instant Messengers Using Data Mining Approach, *Springer-Verlag Berlin Heidelberg*,

Jer Lang Hong (2011), Data Extraction For Deep Web Using Word Net, *Published By IEEE Transactions On Systems, Man And Cybernetics*.

Michael Robertson, Yin Pan, And Bo Yuan (2010) A Social Approach To Security: Using Social Networks To Help Detect Malicious Web Content, *Published By IEEE*.

Daya C. Wimalasuriya And Dejing Dou (2010),Ontology-Based Information Extraction: An Introduction And A Survey Of Current Approaches, *Published At Univ of Oregon.*

Wang Wei, Payam Barnaghi, And Andrzej Bargiela (2010), Probabilistic Topic Models For Learning Terminological Ontologies, *Published By IEEE Tran., On Knowledge And Data Engineering.*

Jerome R. Bellegarda, Fellow (2010), Part-Of-Speech Tagging By Latent Analogy*, IEEE Journal Of Selected Topics In Signal Processing*

Michael Robertson, Yin Pan, And Bo Yuan (2010), A Social Approach To Security: Using Social Networks To Help Detect Malicious Web Content, *Published By IEEE*.

Appavu, And Et Al (2009), Data Mining Based Intelligent Analysis Of Threatening E-Mail, *Published By Elsevier In Knowledge-Based Systems.*

Sunitha Ramanujam, And Et Al (2009), A Relational Wrapper For RDF Reification, *IFIP International Federation For Information Processing* IFIP AICT 300, Pp. 196– 214*.*

Tong Zhang, Fred Damerau , David Johnson (2002), Text Chunking Based On A Generalization of Winnow*, Journal Of Machine Learning Research.*

David W. Cheung, And Et Al (1996), Maintenance Of Discovered Association Rules In Large Databases: An Incremental Updating Technique*,. Journal of Information Science,* 36 (3) pp. 306–323

M.W.Du, And S.C.Chang (1994), An Approach To Designing Very Fast Approximate String Matching Algorithms, *IEEE Journal*.