International Journal of Core Engineering & Management
Volume-6, Issue-10, 2020 ISSN No: 2348-9510

MULTILINGUAL SUSPICIOUS TEXT DETECTION IN SOCIAL MEDIA

Juveria Fatima
Research Scholar, Dept. of CSE, Osmania University,
Hyderabad, Telangana state, India
fatima0912juveria@gmail.com

Mohammed Mahmood Ali
Dept. of CSE, Osmania University,
Hyderabad, Telangana state, India
mahmoodedu@gmail.com

Abstract

Social media is a platform for people to communicate freely with the people of their interest in their comfortable language leading to multilingual suspicious messages go untraceable due to the language barrier issue for detecting suspicious activities. Users of social media express their thoughts, activities, interests, location etc., with people of their interest which makes them an easy target for unintended users/ activities resulting in physical or psychological damage. Despite the fact that many approaches are developed/ proposed to overcome this issue they are mostly worked on major languages like English, Chinese, etc., or with small corpora of few languages. The proposed multilingual suspicious text detection system (MSTD) performs two steps to achieve multilingual suspicious messages detection in IMs/ SNS. The first step is to translate the user language to a particular language by identifying and then translating the language. The second step is to detect suspicious words in that translated text message and reporting the user with details. The proposed system is expected to perform better as it involves a language translator which eliminates the language barrier problem as compared with existing work which makes use of bag of words, machine learning, etc., in multilingual scenario. Language translator keeps the semantics of the message intact. Suspicious text detection is performed using WordNet which maps to the root word and set of pre-defined rules for domain mapping and specification of suspicious domain. The user with suspicious domain activity is reported.

Keywords: Social Media, Language Translation, Information Extraction Through WordNet, Predefined Rules for Domain Matching

I. INTRODUCTION

Social media allows people to use their comfortable languages to communicate with each other. Various networking and instant messaging platforms like Face book, Twitter, WhatsApp etc., are widely used today. People use these platforms for connecting with friends, making new friends and contact either for personal or professional reasons where people get exposure of different personalities/behaviours. This is pleasant most of the time but sometimes innocent people fall prey by unwanted activity towards the user of the account like bullying, blackmailing, threatening,

etc., there are many approaches developed for detecting user's suspicious behavior and reporting. Here we represent an approach for detecting suspicious text in multiple languages using Ontology based Information Extraction for extracting information and set of pre-defined Knowledge-based rules, for decision making process that are learned from domain experts and past learning experiences.

Social media has helped people to communicate freely with people of interest in their comfortable language however leads to multilingual suspicious messages go untraceable. Despite the fact that many approaches are developed/ proposed to overcome this issue but they are mostly worked on major languages like English, Chinese, etc., or with small corpora of few languages. The proposed system performs two steps to detect multilingual suspicious messages in IMs/ SNS. The first step is to convert the user language to a particular language. The second step is to detect suspicious words in that converted text message.

Here we use a language translator as a first step in multilingual suspicious text detection. The language translator will identify the language used by the user for intentional harm and then generates a translation based on its predefined rules for language translation. There are many language translators available today such as Google Translate, Bing translator, Yandex, etc., each have their own specialties and limitations. The second step is to detect the suspicious words in the translated text message/post. We use ontology based information extraction technique for extracting information and set of predefined knowledge based rules for domain prediction. We propose a model that discover and predict such messages that are sent using IM or SNS like Facebook, Twitter, LinkedIn, and others. Information Extraction expects to recover particular sorts of information from characteristic language text by handling them automatically. For instance, an information extraction framework may recover information about international indicators of nations from a lot of site pages while overlooking different sorts of information. Ontology-based information extraction has as of late rose as a subfield of information extraction. Here, ontologies - which give formal and express determinations of conceptualizations - assume a critical job in the information extraction process.

II. PROBLEM STATEMENT AND RELATED WORK

Few approaches are proposed for suspicious text detection in multilingual context however they concentrate on majorly spoken languages like English, Arabic, Chinese etc., are worked with small corpus and so on. We try to develop a system which detects users involved in suspicious activities in social networking sites or instant messaging in multilingual context. Mostly users use their comfortable or mother tongue so for every language making a detection algorithm is a big task as language is big barrier in achieving this. So we use language translator to overcome this issue. Later suspicious detection technique can be applied. If suspicious words are traced, then the domain of suspicion is found. The suspected user can be reported to cybercrime department with user details like e-mail, message, phone number, etc., Lexicon based search was used to identify the language [9] or with the help of Word-Net suspicious pattern detection was achieved but worked only for widely used language English [1]. Some systems have been developed for small corpus of languages for detecting cyber activities [2][3]. Few have worked on known data to predict the outcome or pattern of cybercrime from selected criminal users [5]. The approaches turned out to be quite dynamic for crime detection [1] but rarely concentrated on multilingual dynamic suspicious text detection which this approach is trying to achieve.

International Journal of Core Engineering & Management
Volume-6, Issue-10, 2020 ISSN No: 2348-9510

III. PROPOSED FRAMEWORK FOR MULTILINGUAL SUSPICIOUS TEXT MESSAGE The architecture of the proposed system is being discussed in Fig 1.

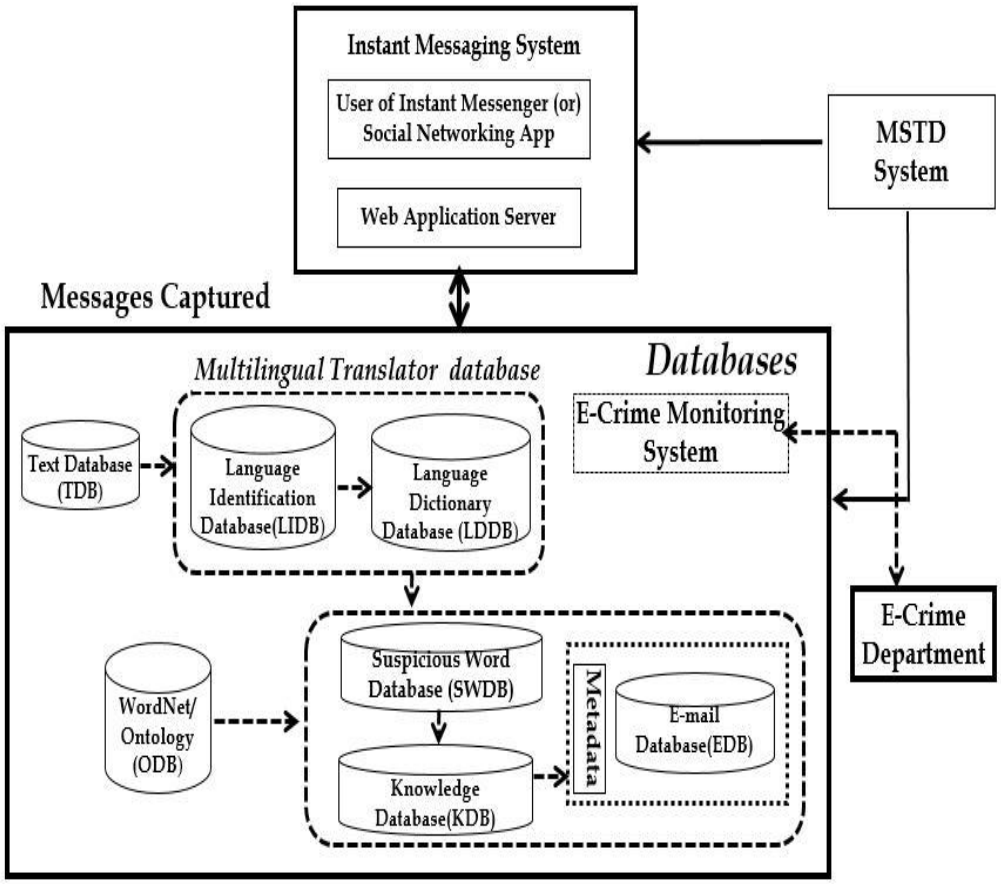

Fig1: System Architecture

Steps:

1. The messages are stored in text database (TDB)
2. Message is pre-processed by removing special characters, repeated words etc., and stored as list
3. The language translator detects the language of the message using Unicode which is unique for each character in a language test bed or language identification database (LIDB) of MTDB
4. After the language of the message is identified the translator translates the message to the required language
5. The message is now analyzed for suspicious word which are in suspicious word database (SWDB) using OBIE technique for information extraction
6. If suspicious words are found, then with the help of pre-defined rules domain is predicted
7. Reports the user with user's details from user Metadata

Predefined Rules:

Table 1: predefined table for domain prediction

Domain activity	Stem words to be detected for a given context
Robbery and Theft	Jewellery, bank, night, gun, knife, location, vehicle, break, night, keys, locker, plan
Match Fixation	Location, luxurious flat, cash, hotel, bet, gifts, virgin girl, bank, payment, loose game
Corruption Charges	Luxurious flat, money, bank, cheque, deposit, diamond, avoid tax, laptop, offshore account
Sexual Harassment	messages, beautiful, come, payment, spend night, location, jewellery, park hotel, car, gift, body parts, property, help, Job
Kidnap	Hijack, capture, seize, abduct, usurp, grab, gun, hostage, location, amount, kill, property, day, plan
Murder	kill, assault, assassinate, eliminate, gun, dagger, knife, stab, location, money, plan
Terrorist Attack	Bomb, vehicle, location, suicide attack, bag, holy place, laptop, demolish, payment , cash, day, time, damage, casualties, explosive
Drug Supply And Smuggling	Packet, brown sugar, cash, cocaine, hashish, M.tabs, Methaqualone, opium, charas, location, injection, Morphine, LSD STR/ECA, dibucaine

Rule1 - Predefined knowledge based rules for domain prediction which help in predicting the domain

Rule 2 - doubtful and unnoticed words are checked and corrected automatically based on nearness of stem words using ontology taxonomy constructed by rule1 in Table 1.

International Journal of Core Engineering & Management
Volume-6, Issue-10, 2020 ISSN No: 2348-9510

Flow Chart for Proposed System:

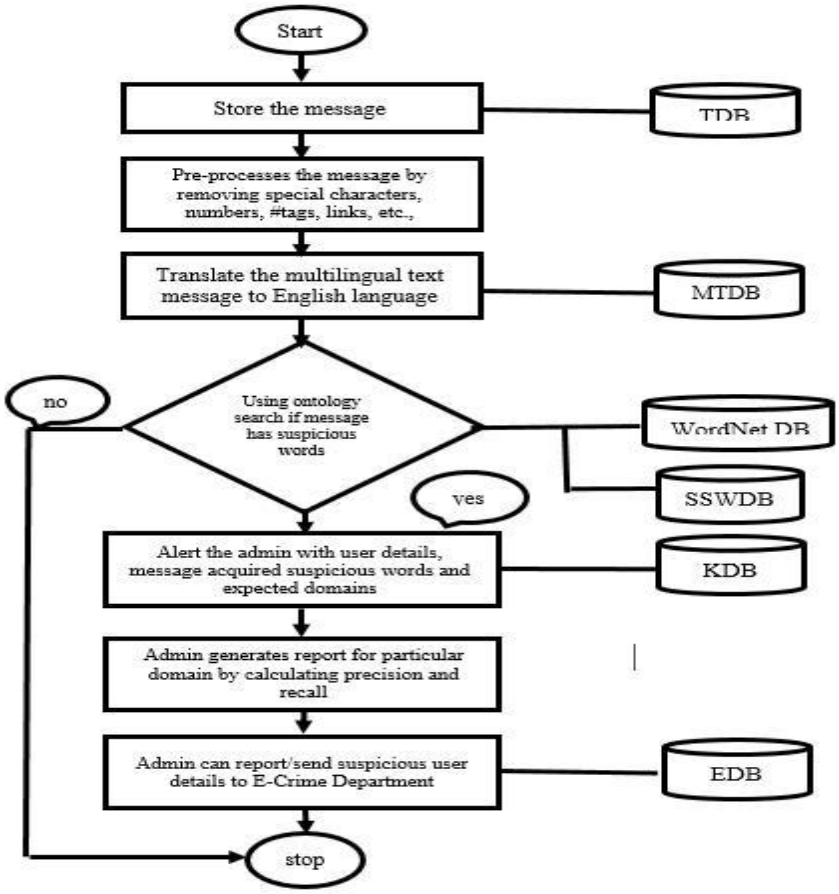

Fig 2: Proposed system flow chart

Proposed system flow chart is shown in fig 2. The steps below give the details of system flow.

Steps:

1. Store the message in TDB text database
2. Pre-processes the message by removing special characters, numbers, #tags, links, etc.,
3. Translate the multilingual text message to English language using multilingual translation database
4. Using ontology search if message has suspicious words using WordNet database and suspicious words database SSWDB with set of pre-defined rules
5. Alert the admin with user details, message, acquired suspicious words and expected domains from knowledge database KDB
6. Admin generates report for particular domain by calculating precision and recall
7. Admin can report/send suspicious user details to E-Crime Department

International Journal of Core Engineering & Management
Volume-6, Issue-10, 2020 ISSN No: 2348-9510

Flow chart for language identification and translation:

Fig 3. Shows the steps involved in language identification process

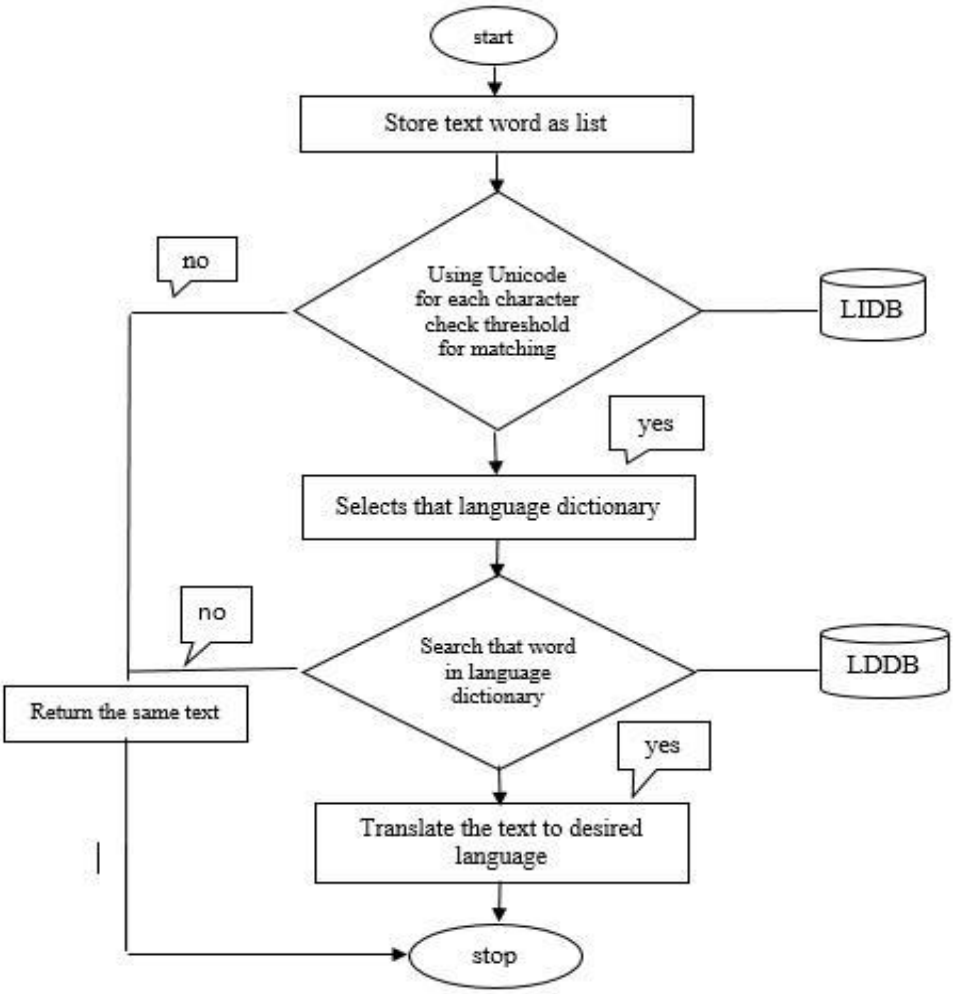

Fig 3: Language identification flow chart

Steps:

1. Store the word as list of characters
2. Using Unicode of each character check predefined threshold for language mapping from language identification test bed or database LIDB
3. If desired threshold is found, then select that language for that word or else return the same word
4. Now search that word in selected language dictionary LDDB
5. If a matching word is found, then translate that word or else return the same word

Unicode for Language Identification:

Unicode is a universal character encoding standard. It defines the way individual characters are represented in text files, web pages, and other types of documents. Unlike ASCII, which was designed to represent only Basic English characters, Unicode was designed to support characters from all languages around the world.

International Journal of Core Engineering & Management
Volume-6, Issue-10, 2020 ISSN No: 2348-9510

Ex. **चेक** के माध्यम से **भगतानु** न करें
091A 0915 0947(**चेक**), 0915 0947 (के), 092E 093E 0927 092F 092E (माध्यम), 0938 0947 (से), 092D 0941 0917 0924 093E 0928 (**भगतानु**), 0928 (न), 0915 0930 0947 (करें)

Language translation flow:

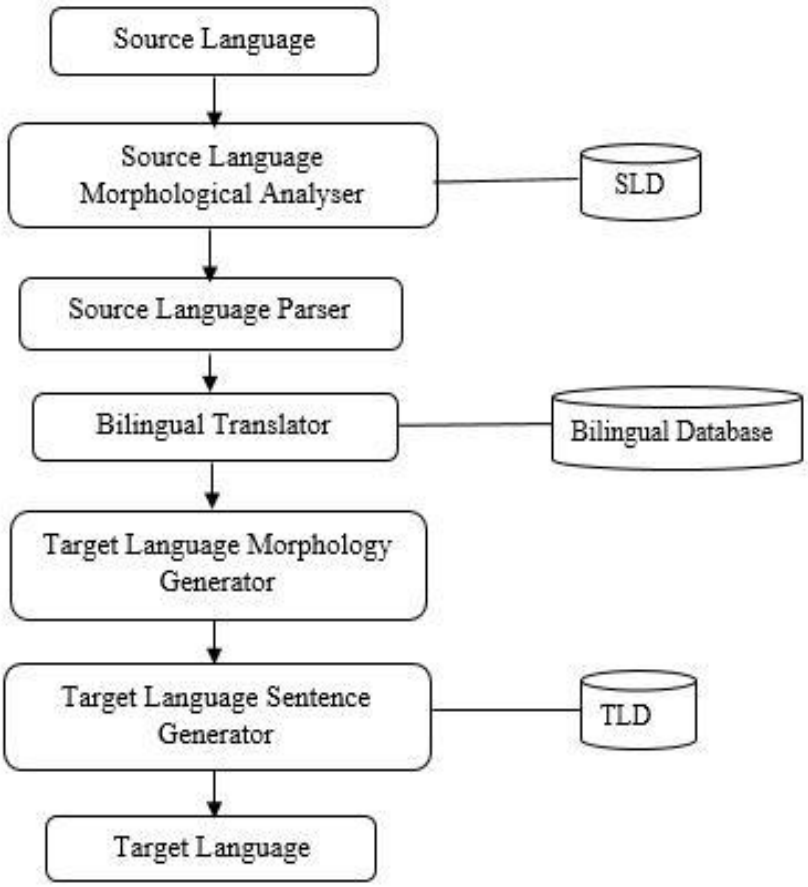

Fig 4: Language translation process

Morphological analyzer is a program for analyzing the morphology of an input word; the analyzer reads the inflected surface form of each word in a text and provides its lexical form, like for nouns it will provide gender, number, and case information, likewise for verbs it will provide tense, aspect and modularity. Whereas generation is the inverse process i.e., given a root and its grammatical features it will generate the word forms of the root word.

Parsing, syntax analysis, or syntactic analysis is the process of analyzing a string of symbols, either in natural language, computer languages or data structures, conforming to the rules of a formal grammar which describes how to form strings from a language's alphabet that are valid according to the language's syntax for instance, which groups of words go together (as "phrases") and which words are the subject or object of a verb. The fig 4 explains the language translation flow.

Steps:

1. Source language sentence in analyzed by Morphological analyzer which is the process of providing grammatical information about the word on the basis of properties of the morpheme it contains.
2. Source language parser is a program that works out the grammatical structure of sentences, for instance, which groups of words go together (as "phrases") and which words are the subject or object of a verb.
3. Bilingual translation is done with the help of bilingual database
4. Target language morphological generator is the inverse process i.e., given a root and its grammatical features it will generate the word forms of the root word.
5. Then target language parser will arrange the sentence according to its target language grammar.

Language Identification Algorithm:

Input: message Mi

Output: message language detected message Mt

1. Pre-process text message by removing special characters, numbers, etc.,
2. Tokenize each character of the words
3. store each character in list L[i]
4. while (L[i]!=0) search
5. For each test-set Ti
6. compare(L[i]==T[i][j])
7. match is found
8. Match = Match+1
9. If Match is >= threshold value // threshold common value defined by the system for all language, Language identified
10. Invoke language dictionary
11. check that word Mi in language dictionary
12. If match is found
13. return translate message Mt
14. else
15. return same word Mi //language not identified
16. End

IV. EXPERIMENTAL RESULT

a. Evaluation methods used:

The metrics used for result evaluation are precision and recall. The metrics are based on suspicious words retrieved after translation from user messages and with our suspicious words database

$$\text{Precision} = \frac{\text{correctly extracted words}}{\text{Total extracted words}}$$

$$\text{Recall} = \frac{\text{correctly extracted words}}{\text{Total no. of possible words}}$$

$$\text{Accuracy} = \frac{\text{correctly extracted words}}{\text{Total no. of words observed}}$$

International Journal of Core Engineering & Management
Volume-6, Issue-10, 2020 ISSN No: 2348-9510

b. Results obtained:

The results obtained are from suspicious database which has 8 domains and 95 words as stems words for defining the domain. We get approx. 1000 words from WordNet for root word extraction for all domains. The dataset is small and created by brainstorming as we could not get the real suspicious data due to authorization issues with social networking sites or instant messengers. But we have tried to work with real time scenarios which works better even in cross language communication scenarios. The results we conclude are increased by 10 times so as to get better performance metrics

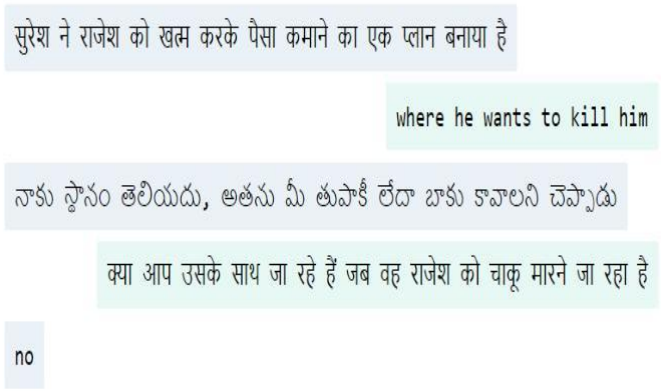

Fig 5. Murder domain scenario for message exchanged

The above fig.5 shows scenarios between two users w.r.t murder domain. The alert received at the admin with user details and suspicious words and expected domains shown in fig 6. The dominant domain is selected after report generation and report is stored in a excel file at the backend explained in fig 7.

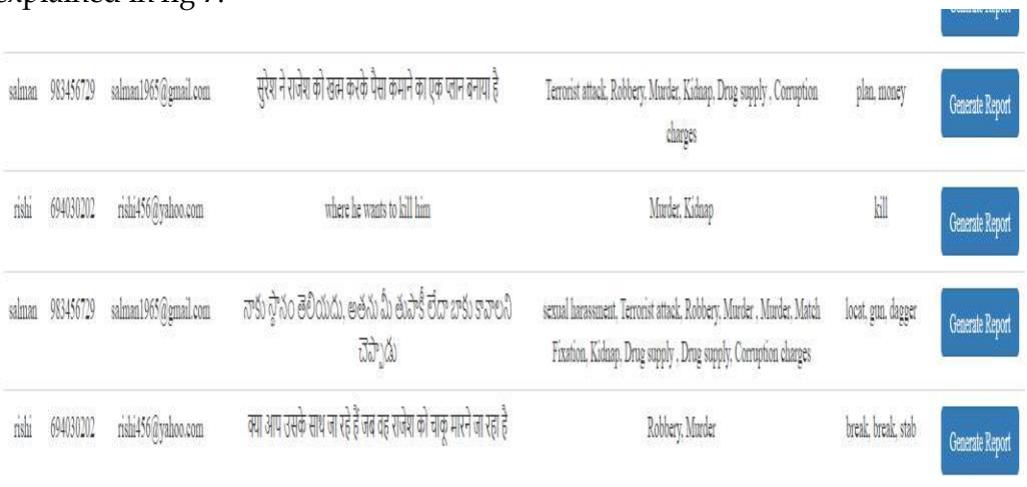

Fig 6. Alert at admin with suspicious words detected and user details

International Journal of Core Engineering & Management
Volume-6, Issue-10, 2020 ISSN No: 2348-9510

Table 2. The results shown here are for murder domain.

Terms	Result
Precision	93.8
Recall	75.0

name	mobile	e-mail	suspicious words	domain
rahul	59834032	rahul123@	night,locat,rob,come,plan,bank,key,place,gun	Robbery
veer	9.88E+08	veer001@	cash,amount,come,bet,money,break,payment,place,gift,bank	Match Fixation
raafe	45805804	raafe0987	help,locat,gift,plan,money,amount,captur,park,break,properti,gun	Kidnap
rahul	59834032	rahul123@	place,bomb,break,demolish,come,laptop,vehicl,cash	Terrorist attack
raafe	45805804	raafe0987	bank,break,chequ,money,day,properti,help,come,amount,deposit	Corruption charges
rahul	59834032	rahul123@	place,locat,money,cocain,packet,come,cash,amount	Corruption charges

Fig 7. Report at backend for E-crime department:

We have used GTD Global Terrorism Dataset for checking the accuracy of the system w.r.t terrorist activity domain with the help of pre-defined rules which help in predicting the domain. The GTD dataset has 135 columns and 191465 rows. We have used 2000 rows for extracting required suspicious data. Where if the suspicious word is found then it takes "1" and if it doesn't find then it's taken as "0". From the all retrieved 1's and 0's accuracy is calculated. Table 3 shows the accuracy results of terrorist activity domain whereas murder domain results are shown in table 2.

Table 3. GTD results

Terms	Result
Accuracy	97.392

c. Comparison of proposed system with current approaches: Table 4. Comparison feature analysis

Features	current social networks	MSTD System
Support for social interaction supported	Yes	Yes
Pre-defined	No	Yes

rules		
Language translation supported	No	Yes
Dynamic Suspicious text detection	No	Yes
Information extraction using ontology	No	Yes
Report generation	No	Yes
Precision and Recall calculated	No	Yes
Database & data mining supported	Yes	Yes

V. CONCLUSION

Language is not a barrier for detecting cybercrime in social media. Suspicious text messages can be dynamically detected in multilingual context. Cybercrimes can be avoided to an extent as the system gives details of suspicious user with predicted suspicious domain. Works on larger corpus as compared to the existing system. Concentrates on all sort of suspicious activities such as kidnap, robbery, murder, etc. This system generates an evidence for investigation

VI. FUTURE WORK

- There may be situations where the user might use some keywords such as kill, hit, etc., for fun. so detecting the sentiments behind a message is an issue but with the help of stance detection approach this can be avoided to some extent
- Can be upgraded for audio, video and picture reading messages.
- The authenticity of the user can be restored by taking a picture at the time of sending message.
- The system can be automated further by directly uploading reports to e-crime department which can speed up the tracing of culprit and avoid any mishap if acted on time.
- The study can be further pursued with other domains against other datasets for other domains specified in this paper

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", 2014
[2] M. Brinda, V. Vishnupriya, S. Rohini, M. Udhayamoorthi "Active chat monitoring and suspicious detection over internet", 2018
[3] Rohit pawar, Rajeev R. Raje "Multilingual cyber bullying detection system ", 2019
[4] Sameera khan, Pinki Vishwakarma "Cybercrime: to detect suspected user's chat using text mining", 2019
[5] Farkhund Iqbal, Benjamin C.M.Fung, Mourad Debbabi, Rabita Batool, Andrew Marrington "WordNet-based criminal networks mining for cybercrime investigation", 2019
[6] Madhura Mandar Phadke, Dr. Satish R.Devane "Multilingual machine translation: An analytical study, 2017
[7] Mausam Stephen Soderland Oren Etzioni Daniel S. Weld Michael Skinner Jeff Bilmes "Compiling a Massive, Multilingual Dictionary via Probabilistic Inference", 2009
[8] B.N.VNarsimha Raju, M.SV.S Bhadari Raju "Statistical machine translation system for Indian languages", 2016
[9] Ali Selamat, Nikolas Akosu "Word length algorithm for language identification of under resourced languages", 2016
[10] Alberto Jimenez-Feltström, "Text Language Detection," 2001
[11] Multilingual Compiler System and Method, Wael Abouel saadat, 2006
[12] M. Mahmood Ali, and L. Rajamani, "Framework for surveillance of instant messages, " published by inderscience in IJITST, vol. 5, 2013.
[13] www.wikepedia.com