



Traffic Detection from Real Time Twitter Stream Analysis and Navigation System

Mrs. Kavita Sawant¹, Miss. Shital Pawar², Miss. Poonam Jadhav³, Mrs. Sayali Vidhate⁴, Mrs. Nirasha Bule⁵, Mrs. Snehal Patil⁶
Professor^{1,2}, BE Student^{3,4,5,6}

Department of Computer Engineering
Bharati Vidyapeeth College of Engineering for Women, Pune, India

Abstract:

Traffic congestion is a big problem the world is facing nowadays. People suffer very bad in terms of money and time. In this paper we present a system which detect the traffic and analyze reason of traffic using twitter stream analysis. Social networks are very useful source of information for event Detection with particular reference to road traffic congestion and car accident. Twitter is a social networking site which allows people to share and read tweets. The system fetches the tweets from twitter; applies natural language processing technique on them; categorizes the tweets related to traffic; notifies the registered users about it. Natural language processing (NLP) focuses on developing efficient algorithms to process text and convert it into machine understandable language. Here, we apply NLP on the tweets to detect the traffic.

Keywords: Text mining; Classification; Social network; Event detection.

I. INTRODUCTION

People these days are getting more and more active regarding their everyday activity. It is been addiction which can result in more productive application. It is regarding sedan drivers posting every minute detail of traffic and route. In this project user gets public traffic tweets. Apply tokenization, remove stop words, and apply stemming to tweets and Detect the cause based on the word which is stored into database. And then it provides the path using navigation system.

II. LITERATURE REVIEW

Pietro Ducange ,Eleonora D'Andrea, Beatrice Lazzerini, Member, IEEE, and Francesco Marcelloni, Member, IEEE[1] in this paper proposed a system which is build on Service Oriented Architecture (SOA). the SOA fetches the tweets from the twitter then classified it .also the system gives cause of the traffic. Techniques are used text analysis and pattern classification. Text were not well formatted.

Anusha Jalaparathi and A Suraj Kumar[2] Introduced a feature software packages and state-of the art techniques for text analysis and pattern classification these techniques and technologies are analyzed, and integrated.

Kunigiri eswar kumari, H. Ateeqahmed [3] In this paper data mining is used it means extracting the important data and used of the natural language processing(NLP), text mining and equipment learning algorithms are used for traffic detection.

III. PROPOSED SYSTEM

Our proposed system consists a new idea of android application where user need to login to the android application and then User can search path and they can see traffic on that way which

will be displayed on the map. Also user can select the alternate path.

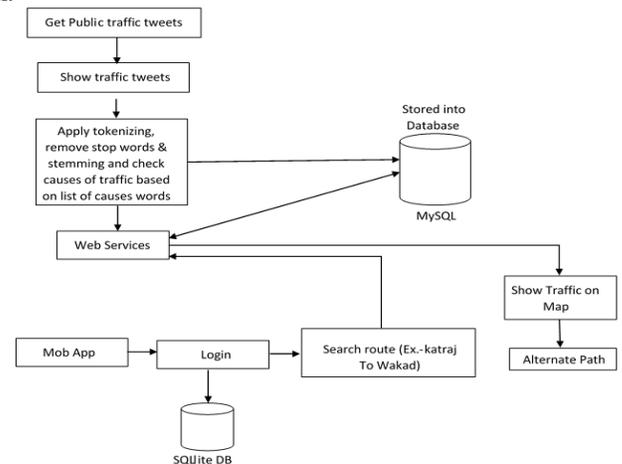


Figure.Error! No sequence specified. **Proposed System Architecture**

1. Twitter traffic system: In this module, we are fetching public traffic twists from twitter on basis of different traffic related keyword, which is stored in our database and if that keyword is present in the text of tweet, then that twit will be fetched. Most of the times while making twit, many people’s uses additional information in the twit such as links, symbols, hash tags but our system we are considering only text part of each twit, and the information like user id, hash tags, links are discarded. We are using NLP (Natural Language Processing) algorithm in which tokenization, remove stop words and stemming is applied to a particular tweet. We have maintained lists of causes (e.g. Accidents, Traffic, Jams, Vehicle breakdowns, etc.) and we check these causes in that particular tweet. Tweets and there causes are stored into database.

2. Web Service:

In this module, the communication between user and twitter traffic system is established over a network. Web service offers a various services to the user, via android application, where services are built by twitter traffic system and that driven through web services. Tweets and their causes are stored into database. Web service gets array of latitude and longitude of searched path and then the latitude and longitude of the traffic is compared with searched path with their causes. After comparing the longitude and latitude having traffic, it is displayed on the maps of Android device.

3. Mobile application:

In this module, the various services are provided to the user via android application. To get services user need to register after registration user will login to the application, after successful login User can search path (Ex-Karaj to Wakad) and they can see traffic on that way.).Also he can select the alternate path.

SOFTWARE REQUIREMENT AND TOOLS

- Eclipse mars, JDK 7,
- Android Studio 15
- Tomcat 7
- Database- MySQL 5.2, SQLite 3.8.6.
- Technology- Java, Hibernate.

IV. ALGORITHM

The system will include the following algorithm and methods

I. NLP (Natural Language Processing)

It is the process in which human language is automatically processed. it is the communication between computers and human language. NLP has relation with human computer interaction. it facilitates computers to extend meaning from human or natural language input. Natural language processing systems take strings of words (sentences) as their input and provide organized representations finding the meaning of those Strings as their output. The behavior of this output depend sheavily on the task athand.NLP performs key role in our system the sentence or word taken as input. Syntactic dissection gets at the grammar of the sentences. Semantic analysis deals accord with the meaning of words and sentences, the ways that words and sentences refer to elements in the world. In NLP there are further steps which we have used to process the tweets for making a real time result of traffic. These steps are mentioned below

- 1.Tokenization
- 2.remove stop word
- 3.steamming

1. Tokenization

In the specified sentence or any report or document, tokenization is the process of separating it into pieces, called *tokens* , perhaps at the same time eliminating certain letters, such as punctuation.

Here is an example of tokenization:

Input: Hello Welcome to world of technology;

Output:

Hello	Welcome	to	world	of	Technology
-------	---------	----	-------	----	------------

These tokens are individually known referred as terms or

words, but it is sometimes important to make a type/token distinction. A *token* is an object of a sequence of characters in some particular document that are grouped together as a productive semantic unit for processing. The big question of the tokenization phase is what are the correct tokens to use? We can separate whitespace and throw away punctuation characters. This is a starting point, but even for English there are a number of tricky cases, with each language there is always some issue in order to tokenize it.

2. Stemming

For grammatical reasons, documents or report are going to use different forms of a word, such as *complete, completing, and completed or organize, organizing*. Additionally, there are families of extendedly related words with similar meanings, such as *democracy, democratic, and democratization*. In many situations, it look like if it would be useful for a search for one of these words to return documents that contain another word in the set. The target of stemming is to deduct inflectional forms sometimes derivationally related forms of a word to a common base form. For instance:

car, cars, car's, cars' ⇒ car

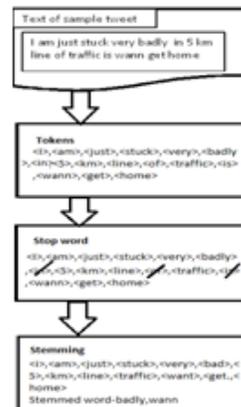
Dog, dogs, dog's=dog

However, the two words suspend in their flavor. *Stemming* usually refers to a crude heuristic process that cutoff the ends of words in the hope of achieving this target correctly most of the time, and often includes the removal derivational join.

3. Stop word

a an and are as at be by for from
has he in is it its of on that the
to was were will with

A stop word set of 25 words. Sometimes, extremely same words which would present to be of small value in helping specified documents matching a user need are excluded from the vocabulary entirely. These words are called *stop words*. The general technique for analyzing g a stop list is to sort the terms by *collection frequency* (the total number of times each term present in the document collection), and then to take the most of the time appeared terms, often hand-filtered for their semantic content relative to the domain of the documents being indexed, as a *stop list* , the members of which are then discarded during indexing. An example of a stop list is shown above . Using a stop list significantly deducts the number of postings that a system has to save. Following is the example of all above process which are applied on sample tweet.



I. Haversine method

They have sine formula is an equation important in navigation, giving great-circle distances between two points on a sphere from their longitudes and latitudes.

Haversine formula:

$$a = \sin^2(\Delta\phi/2) + \cos \phi_1 \cdot \cos \phi_2 \cdot \sin^2(\Delta\lambda/2)$$

$$c = 2 \cdot \text{atan2}(\sqrt{a}, \sqrt{1-a})$$

$$d = R \cdot c$$

where, ϕ is latitude, λ is longitude, R is earth's radius (mean radius = 6,371km);

Example:-

Point 1: ,

Point 2: ,

Distance: 968.9 km (to 4 SF*)

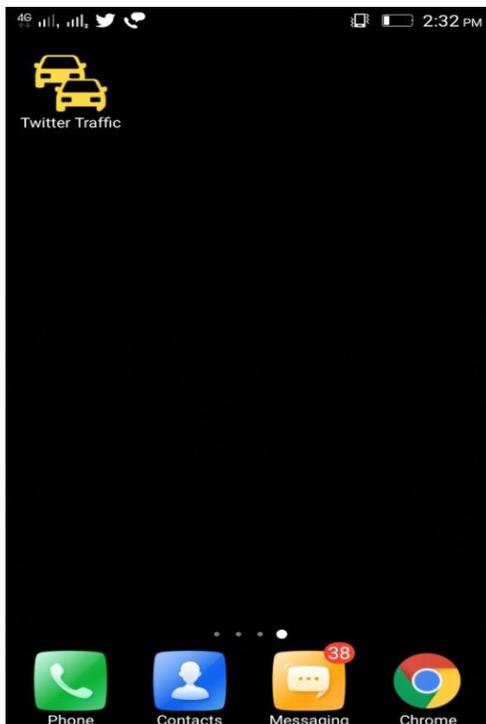
Initial bearing: 009° 07' 11"

Final bearing: 011° 16' 31"

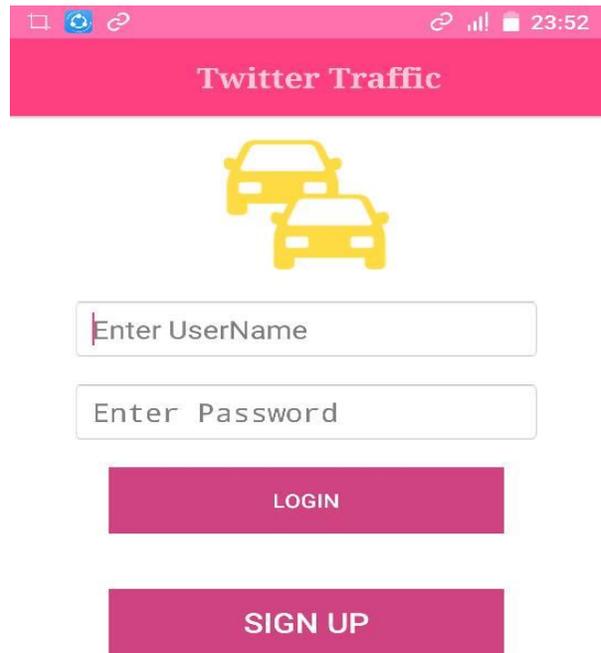
Midpoint: 54° 21' 44" N, 004° 31' 50" W

ADVANTAGES

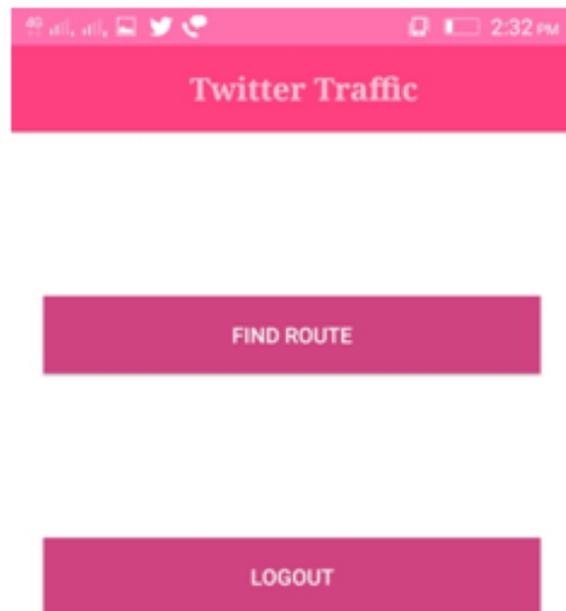
- It provides the path towards destination along with traffic details.
- System suggests the alternate path towards destination.
- It provides the cause of traffic (eg. Accident ,rally ,manifestation etc)
- System uses social media effectively to discover the traffic.



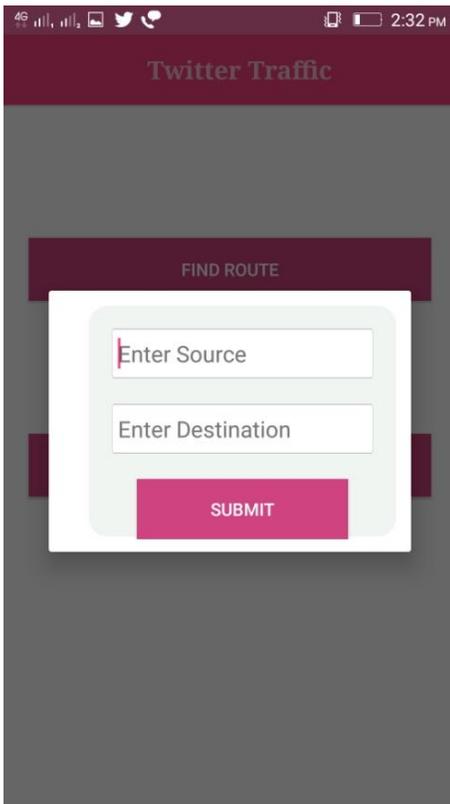
To make the system more friendly and interactive to the user we have developed an android application which will guide the user to find the minimum traffic path or route in very efficient way. The result of the analysis which is done at the web service level is provided to the user via android application



To interact with the android application or to use the application the user just need to do some authorization process such as he need to just register via application and after that he will get the authorized password and user name via mail and using this credentials he/she can login to the android application



Here user has to enter find route option for finding the route. And also user cans logout the application.



To find the path user need to login. Only after successful login, the user is able to find the path. to find the path he need to enter his source and destination location correctly. After entering the source and destination click on the submit button and user will get the map on his device which is first minimum traffic path towards the destination. Traffic is shown in blue darken line if it is present.

Eg Source Pune
Destination Mumbai



After getting first minimum traffic path, but user don't want to go via this first path for any reason he has the choice, to get next minimum path by clicking on find alternate route button after doing this, he need to do the same procedure like entering the source and destination and he will get the alternate path easily. After getting first minimum traffic path, but user don't want to go via this first path for any reason he has the choice, to get next minimum path by clicking on find alternate route button after doing this, he need to do the same procedure like entering the source and destination and he will get the alternate path easily.

V. CONCLUSION

In this paper we have proposed system which detect the traffic and analyze reason of traffic using twitter stream analysis and navigation system. It detects the cause based on the word which is stored into database. And then it provides the alternatives path using navigation system. Social media provide substantial details regarding traffic condition by using graphical areas. The data can be extracted and analyzed using streaming process and navigation system. By using this technique, we can identify the traffic status about particular area like traffic crash, traffic jam and it helps the drivers to notify the alternative path.

VI. REFERENCE

- [1]. "Real-Time Detection of Traffic From Twitter Stream Analysis" Pietro Ducange, Eleonora D'Andrea, Beatrice Lazzarini, Member, and Francesco Marcelloni, Member, IEEE, NOV 2015
- [2]. "Monitoring and Analysis of Real time detection of traffic from twitter stream analysis" Anusha Jalaparathi 1* and A Suraj Kumar 2 Sanketika Vidya Parishad Engineering College, Visakhapatnam – India, IJSC JUNE 2016
- [3]. "Estimation of Traffic with Accuracy through Twitter Stream Analysis" Assistant Professor, Dept of CSE, Dr. K. V. Subba Reddy Institute of Technology, India
E-mail: ateeqh25@gmail.com,,
- [4]. KUNIGIRI E SWAR KUMAR 1, H. A TEEQ A HMED PG Scholar, Dept of CSE, Dr. K. V. Subba Reddy Institute of Technology, India, E-mail: eswargutz@gmail.com. IJITECH JULY 2016