

Navigation Using Twitter

Anmay Singh¹, Hrishikesh Rajpathak², Smit Bhagat³, Komal Chavan⁴

^{1,2,3,4}Dept of Computer Engineering

^{1,2,3,4}MAEER's MIT College of Engineering

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

Social networks have been recently employed as a source of information for event detection with particular reference to road traffic congestion and car accidents. Citizens these days are getting more and more active regarding their everyday activity. Somehow their addiction can result in more productive application. One such is regarding Sedan drivers posting every minute detail of traffic and route. In this project, user gets public traffic tweets from Twitter. Apply tokenize, remove stop words and apply stemming to tweet. Detect the causes on that tweet based on Words that are stored into database.

User logs in via android application, User can search path. Web portal gets array of latitude and longitude and sends return to traffic between that array with causes. Alternate path displayed with traffic. The traffic detection system was employed for real-time monitoring of several areas of the Italian road network, allowing for detection of traffic events almost in real time, often before online traffic news web sites. Twitter users to external attack servers. To cope with malicious tweets, several Twitter spam detection schemes have been proposed. These schemes can be classified into account feature-based, relation feature-based, and message feature based schemes. 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. PROPOSED METHODOLOGY

A. Motivation

We have proposed a system for real-time detection of traffic-related events from Twitter stream analysis. The system is also able to discriminate if a traffic event is due to an external cause, such as football match, procession and manifestation, or not. Web portal gets array of latitude and longitude and sends return to traffic between that array with causes. Alternate path displayed with traffic.

B. Objective

Our traffic detection system based on Twitter streams analysis is presented. And it detects the traffic events in real-time. Also Web part 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. Haversine method used to calculate distance between two latitude-longitude pairs, Triangulation for getting GPS Location.

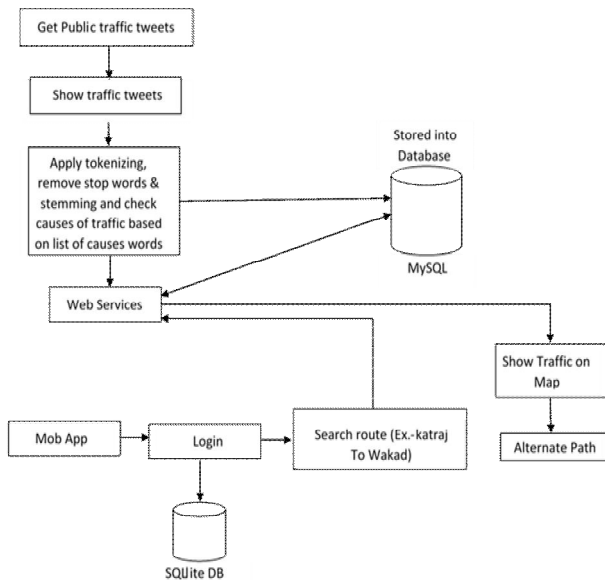
C. Literature survey

- 1) Pietro Ducange, Eleonora D'Andrea, Beatrice Lazzarini, Member, IEEE, and Francesco Marcelloni, Member, IEEE. 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.

- 2) Anusha Jalaparathi and A Suraj Kumar. 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.
- 3) Kunigiri eswar kumari, H. Ateeqahmed. 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. FUNDAMENTALS

A Architectural Diagram



B Technologies to be used

- JAVA:

Java has been tested, refined, extended, and proven by a dedicated community of Java developers, architects and enthusiasts. Java is designed to enable development of portable, high-performance applications for the widest range of computing platforms possible. By making applications available across heterogeneous environments, businesses can provide more services and boost end-user productivity, communication, and collaboration— and dramatically reduce the cost of ownership of both enterprise and consumer applications.

The original and reference implementation Java compilers, virtual machines, and class libraries were originally released by Sun under proprietary licenses. As of May 2007,

in compliance with the Specifications of the Java Community Process, Sun re-licensed most of its Java technologies under the GNU General Public License. Others have also developed alternative implementations of these Sun technologies, such as the GNU Compiler for Java (byte code compiler), GNU Class path (standard libraries), and Iced Tea-Web (browser plug in for applets).

- About Hibernate:

Hibernate an object-relational mapping framework for the Java language. It provides a framework for mapping an object-oriented domain model to a relational database. Hibernate solves object-relational impedance mismatch problems by replacing direct, persistent database accesses with high-level object handling functions. Hibernates primary feature is mapping from Java classes to database tables; and mapping from Java data types to SQL data types. Hibernate also provides data query and retrieval facilities. It generates SQL calls and relieves the developer from manual handling and object conversion of the result set.

- MySQL:

MySQL is the most popular Open Source Relational SQL Database Management System. MySQL is one of the best RDBMS being used for developing various web-based software applications.

MySQL is developed, marketed and supported by MySQL AB, which is a Swedish company. MySQL is the most popular Open Source Relational SQL Database Management System.

MYSQL Enterprise edition includes the most comprehensive set of advanced features & management tools for MYSQL.

MYSQL is the world's most popular open source database. Whether you are a fast-growing web property, technology ISV or large enterprise,

MYSQL can cost-effectively help you deliver high performance, scalable database applications

MYSQL is popular choice of database for used in web application & is a central component of widely used LAMP open source web application software stack.

MYSQL Query Analyzer: To optimize performance by visualizing query activity and fixing problem SQL code.

- **JDK:**

The Java Development Kit (JDK) is an implementation of either one of the Java Platform, Standard Edition, Java Platform, Enterprise Edition, or Java Platform, Micro Edition platforms released by Oracle Corporation in the form of a binary product aimed at Java developers on Solaris, Linux, macOS or Windows. JDK (Java SE Development Kit) Includes a complete JRE (Java Runtime Environment) plus tools for developing, debugging, and monitoring Java applications. JDK is required to develop and run Java applications and applets.

- **Eclipse:**

Eclipse is an integrated development environment (IDE) used in computer programming, and is the most widely used Java IDE. It contains a base workspace and an extensible plug-in system for customizing the environment. Eclipse is an integrated development environment (IDE) used in computer programming, and is the most widely used Java IDE. It contains a base workspace and an extensible plug-in system for customizing the environment.

IV. OVERALL DESCRIPTION

A. Product Perspective:

The developed system was installed and tested for the real-time monitoring of several areas of the road network, by means of the analysis of the Twitter stream coming from those areas. The aim is to perform a continuous monitoring of frequently busy roads and highways in order to detect possible traffic events in real-time or even in advance with respect to the traditional news media.

B. Requirement:

- **Software Requirements:**

- 1) Eclipse mars, JDK 7,
- 2) Android Studio 1.5
- 3) Tomcat 7
- 4) MySQL 5.2, SQLite 3.8.6

- **Hardware Requirements:**

- 1) 8 GB RAM
- 2) 500 GB HDD

C. Product Function:

- **Web Service:**

Web service is a method of communication between two electronic devices over a network. It is a software function provided at a network address over the Web with the service. A Web service generally as, a software system designed to support interoperable machine-to-machine interaction over a network.

Tweets and their causes are stored into database.

Web part 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. Bootstrap template GUI for web.

- **Android App:**

Login and registration of user. User can search path and they can see traffic on that way. Also user can select the alternate path.

V. ALGORITHM

- **NLP (Natural Language Processing) Algorithm:-**

It is the process in which human language is automatically processed. It is the communication between computers and human language. 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. 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:

1. Tokenization
2. Remove stop word
3. Stemming

Step 1: Tokenization

Tokenization is the process of separating it into pieces, called tokens. A token is an object of a sequence of characters in some particular document that are grouped together as a product semantic unit for procession. Here is an example of tokenization:

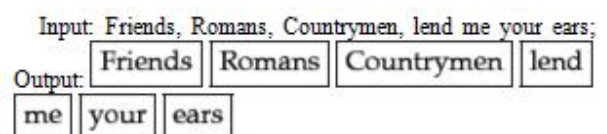


Figure 2.4: Tokenization

Step 2: Remove stop word

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.

Stopword list

a	been	get
about	before	getting
after	being	go
again	between	goes
age	but	going
all	by	gone
almost	came	got
also	can	gotte
am	cannot	had
an	come	has
and	could	ha

Figure 2.5: Remove stop word

Step 3: Stemming

The target of stemming is to deduct inflectional forms sometimes derivationally related forms of a word to a common base form. For instance:

waited, waiting, waits => wait
 closed, closing, closely => close

• Haversine method:-

The haversine formula is a very accurate way of computing distances between two points on the surface of a sphere using the latitude and longitude of the two points. Central angle Haversine can be computed, between two points with r as radius of earth, d as the distance between two points,

ϕ_1, ϕ_2 latitude of two points and
 λ_1, λ_2 longitude of two points respectively, as:

$$\text{hav} \left(\frac{d}{r} \right) = \text{hav}(\phi_2 - \phi_1) + \cos(\phi_1) \cos(\phi_2) \text{hav}(\lambda_2 - \lambda_1)$$

VI. CONCLUSION

We developed this system for Traffic detection of traffic-related events from Twitter stream analysis and we have also maintained lists of causes like for eg. Accidents, Traffic-Jams, Vehicle breakdowns, etc. Social networks have been recently employed as a source of information for event detection with particular reference to road traffic congestion and car accidents. We check these causes in that particular tweet: Showing traffic tweet with causes and Showing traffic between two points.

VII. DISCUSSION OF RESULT

Our traffic detection system based on Twitter streams analysis is presented. And it detects the traffic events in real-time. Also Web part gets an array of latitude and longitude of searched path and then the latitude and longitude of the traffic are compared with searched path with their causes. After comparing the longitude and latitude having traffic, it is displayed on the maps of Android device.

VIII. ACKNOWLEDGMENT

It gives us great pleasure in presenting the preliminary project report on NAVIGATION USING TWITTER. We are profoundly grateful to Prof. ANITA THENGADÉ for her expert guidance and continuous encouragement throughout to see that this project rights its target since its commencement to its completion. We would like to express deepest appreciation towards Prof. Dr. Anil. S. Hiwale, Principal, MIT COLLEGE OF ENGINEERING and Prof. BHARATI DIXIT, Head of Department of Computer Engineering whose invaluable guidance supported us in completing this project. At last we must express our sincere heartfelt gratitude to all the staff members of Computer Engineering Department who helped us directly or indirectly during this course of work.

REFERENCES

[1] TWITTER STREAM ANALYSIS FOR TRAFFIC DETECTION:
<http://www.ejournal.aessangli.in/ASEEJournals/CE145.pdf>

[2] Real-Time Detection of Traffic From Twitter Stream Analysis:
<http://ieeexplore.ieee.org/xpl/login.jsp?tp=arnumber=7057672url=http>

[3] Rucha Kulkarni, Sayali Dhanawade, Shraddha Raut, D. S. Lavhakare, Twitter Stream Analysis for Traffic Detection in Real Time, IJARIT, 2016

[4] Yisheng Lv, Yuanyuan Chen, Xiqiao Zhang, Yanjie Duan, and Naiqiang Li, Social Media Based

Transportation Research: the State of the Work and the Networking, IEEE, 2017

- [5] Shen Zhang, Using Twitter to Enhance Traffic Incident Awareness, IEEE 18th International Conference on Intelligent Transportation Systems, 2015