

Using Social Networking Data as a Location based Warning System

Harsh Alkutkar, George Sam
Computer Department
MIT College of Engineering
Pune, India

Kailash Tambe, Bharati Ainapure
Computer Department
MIT College of Engineering
Pune, India

ABSTRACT

Twitter and Facebook are huge social networks that contain a lot of data that can be used for sentiment analysis. We often find out that a particular area we travel to is dangerous, after asking around. However what if we could use social networking data like ‘tweets’ to find out if a place is actually dangerous? This paper introduces how to use social networking data, analyze it and use it for alerting someone in a disaster prone or high crime rate prone area with the help of smartphones using natural language processing and sentiment analysis.

General Terms

Mobile Computing, Artificial Intelligence, Sentiment Analysis, Social Networks

Keywords

Social Networks, Data Mining, Warning System, Android

1. INTRODUCTION

The importance of social networks and their integration with in our lives is becoming more and more notable with the advent of Smartphones and Social Platforms such as Facebook and Twitter opening up their API's for further development. These API's allow one to find out various data such as sentiments about a particular event or the number of people

This information can be useful in finding out if any criminal activity or natural/artificial disasters are taking place in that particular area where the user is at the current moment.

A social networking service is a platform that focuses on the building of social networks or social relations among people who, for example, share interests, activities, backgrounds, or real-life connections. There are a 100 million active users on Twitter. About 55% of them access twitter from their mobile devices. [1] When you sign up with Twitter, you can use the service to post and receive messages to a network of contacts. Instead of sending a lot of e-mails or text messages, you send one message to your Twitter account, and the service distributes it to all your friends. Members use Twitter to organize gatherings, carry on a group conversation or just send a quick update to let people know what's going on. These tweets are geo tagged which means we can access the exact location of that tweet. Any criminal activity happening in that area, or sentiments about a particular area are geo tagged and can be found out easily with the API. By analyzing the words in the tweet we

can draw conclusions. For example if shootings frequently took place in a particular area of Los Angeles, they would originate from Compton, Los Angeles County ,CA and the analysis of words such as “robberies” , “shootings” , “drugs” can lead us to a general conclusion about the city Compton in Los Angeles. We can thus store this information in our database for further use as mentioned.

2. LITERATURE REVIEW AND RELATED WORK

Sentiment Analysis which automatically extracts and analyses the sentiments in written text has emerged as an active area of research. [2]

Algorithms for mining data from feedback such as the rating of movies [3] and hotels [4] from user reviews have previously been proposed. Comparatively however these algorithms relied on static data from websites and not live data from social networking websites. The flexibility of the Android operating system has driven innovation in this field resulting in various applications of the device. One application was an Android Phone Surveillance [5] System which proposed a monitoring scheme based on Android smart phone terminal, which makes it possible to monitor the target site anywhere and anytime under the coverage

of wireless network. Wireless location and surrounding search systems have been implemented and used widely for locating banks, shops etc and to provide navigation functions. [6] Visual Sentiment analysis on twitter data streams has also been done , indicating positive and negative tweets from all around the world using natural language processing.[7] A social emergency alert service [8] that makes use of the wide availability of smart phones and activates nearby social contacts in cases of emergency has also been researched. Also, a Dangerous Location Aware System for Assisting Kids Safety Care [9] provides insight as to how the system alerts a kid on the move of the possible dangers whenever he/she is near accident prone areas.

Tweet analysis websites such as Sentiment 140 [10] also exist. However, none of these solutions integrate social networking and do not employ natural language processing to find out dangerous locations dynamically.

3. UTILIZING AGGREGATED DATA ABOUT LOCALITIES

Smartphones are used everywhere in the world and are capable of various data aggregation and connectivity functions. Smartphones can be programmed to run specific software or in simple terms “apps”. Google Android is one such open source platform that allows development in a Java based environment. A simple application written on this device can easily keep pinging the central server where all the aggregated data is stored and ask for the rating of that particular area via simple

HTTP communication utilizing the GPS built into the mobile phones. The latitude and longitude sent by the phone will be first used for isolating all the tweets within a 5km radius of that latitude and longitude, and also any tweets containing any mention of the place with the associated keywords. A tweet depending on the associated keywords will be analyzed by the algorithm as a positive tweet or a negative tweet, along with a severity ranking. This severity ranking will be sent as a reply to the users phone. If the area is a very dangerous area, it will have a high severity ranking. This will be interpreted by the application and it may alert the user of any ongoing activities about that area or send a message to local authorities or loved ones letting them know that the user is in a dangerous area.

4. SYSTEM IMPLEMENTATION

4.1 Overview of the working

Twitter allows API integration , which means anyone can access its data after giving proper credentials. Once verified , one can write programs that take full advantage of twitter’s search services. For this we can use a server side script such as PHP which is supported by Twitter and parse the results from the Twitter web service which come in XML format. It is shown in Fig. 1. Sample Tweet

To implement the model, we create an application that will run on android phones. This application will keep pinging the server every X minutes OR when the user’s location changes. When the location changes the user’s latitude and longitude are sent to the server on which the PHP file resides – via the HTTP protocol. It is shown in Fig. 2 Working of the application After analysis of the tweets the user will receive a response from the server whether or not the area is safe enough or any observations have been made.

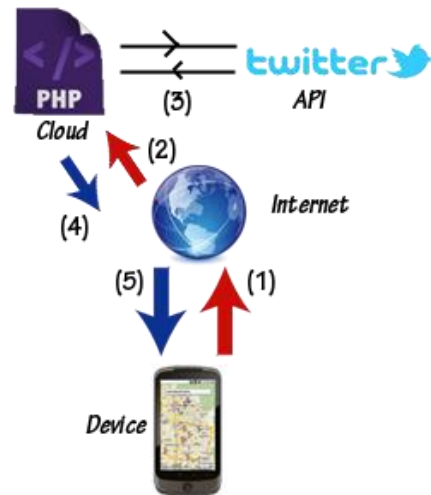


Fig. 2 Working of the System

4.2 Analysis of Tweets

Using natural language processing techniques, we determine topics or subjects, extract attributes from those topics employing attribute extraction and feature extraction and mapping attributes and opinions to measure sentiment value.[9]

There was a **shooting** at **century 16 aurora** about 30 minutes into the movie. Its **shocking!** . #batman

Century 16 Aurora	Shooting	Shocking
-2	-1	-1

Figure 3. NLP Analysis of tweets

The above example shows how a sample tweet can be analyzed for a danger severity rating. An attribute , that is the place in question (Century 16, Aurora) is associated with two negative words i.e – shooting and shocking. Each word has a predetermined severity index. These are added together to get a cumulative severity index for the place in question. Thousands of such tweets are analyzed and the average of all the cumulative severity ratings is considered for that particular place.

4.3 Algorithm

1. Application installed on the android phone
2. Application pings server at intervals or on location change events via HTTP GET or POST, sending the latitude and longitude co-ordinates of the current position
3. As soon as server gets location co-ordinates, the PHP script running on server accesses twitter API
4. Twitter API processes-
 - 4.1 Shoot a query to find all possible tweets in a 5km radius from the current location
 - 4.2 The results are analysed one by one and assessed depending on the key-words, or hash-tags, in the tweets.
 - 4.3 If the keywords correspond to negative sentiments or they specify events such as



Fig. 1 Sample Tweet

robberies, or other criminal activities, detrimental to social peace, then the locality is marked with a number corresponding to the danger severity along with all possible sightings and references to the tweets

- 4.4 If the criminal sightings or references have occurred in the last 1hr then the severity of the ranking is increased.
- 5 Based on the location details received from user, server sends ranking to mobile device
- 6 Mobile device alerts the user through ‘Visual Alerts’ about the location he is present in.

5. MOBILE APPLICATION FEATURES

Fig. 3 shows the Latitude and Longitude extracted from the device and shown on an Activity screen on the phone.



Fig. 4 Latitude and Longitude shown on the mobile device

The application will ask the user whether he wants to find out his current location using the GPS on his mobile. This will be done by the press of a button on the screen of his mobile. The button press event will trigger the application to send his Location specifics to the server and extract the Severity ranking from the server. This is shown in Fig. 4

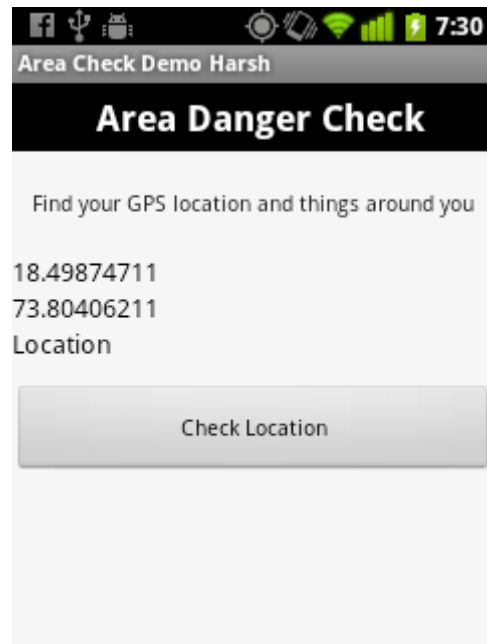


Fig. 5 The Screen where user is prompted whether he wants to find out his current location.

Once the location is received, the server will send back a severity ranking based on the algorithm back to the device. This information of ranking can be shown to the user on his device. This is shown in Fig. 5

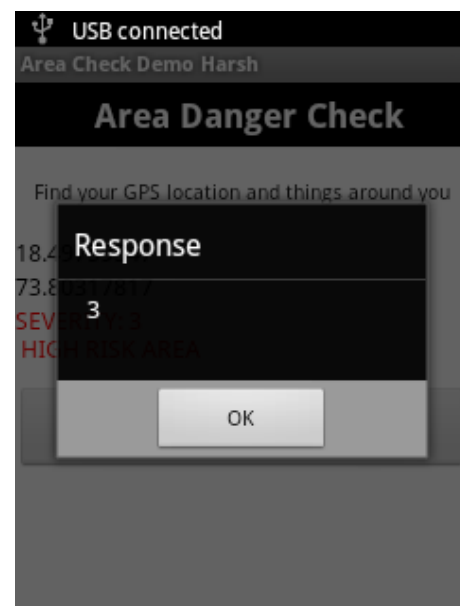


Fig. 6 Severity Ranking Shown on Screen

The device can then prompt the user about the problems in the location he is currently in. It can tell him based on the severity ranking, whether the area he is in is a high risk area or a low risk area for criminal activities. This is shown in Fig. 6 which shows the phone prompting the user of the risk factor of the location on the screen. The mobile device can also specify the type of activity whose occurrences are high in the area.

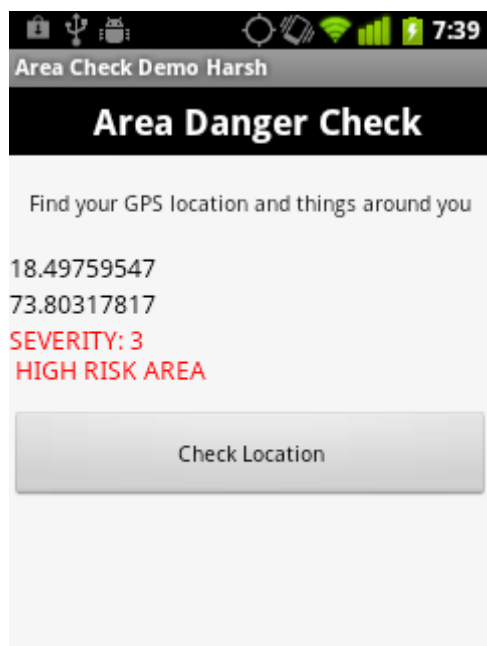


Fig. 7 The Mobile device prompt warning the user of Area beign High Risk Area

6. IMPROVEMENTS AND FUTURE WORK

Future improvements include gathering the crime rate statistics data from the local authorities and adding it into the database so the algorithm can also take the data into consideration. The phone can also go into a special mode in which if the user does not interact with it in a stipulated amount of time even after alerting him, the phone will wipe all the data on the phone and alert the local authorities with the latitude and longitude.

7. CONCLUSION

We can see from this paper that there is immense potential in harnessing the power of social networks and mobile technologies as a sophisticated warning system. Such a system will definitely benefit people all around the world.

8. REFERENCES

- [1] G. O. Young, "Infographic: Social Media Statistics For 2012," <http://www.digitalbuzzblog.com/social-media-statistics-stats-2012-infographic/>
- [2] Akshi Kumar, Teeja Mary Sebasitan 2012, *Sentiment Analysis: A Perspective on its Past, Present and Future*. International Journal of Intelligent Systems and Applications
- [3] Ion SMEUREANU ; Cristian BUCUR 2012, *Applying Supervised Opinion Mining Techniques on Online User Reviews*, Informatica Economica Journal
- [4] Han Xiao Shi. *A sentiment analysis model for hotel reviews based on supervised learning*, Machine Learning and Cybernetics (ICMLC), 2011 International Conference.
- [5] Pang Heming, Jiang Linying, Yang Liu, Yue Kun. *Design and Implementation of Android Phone Surveillance System* - IEEE 2010 International Forum on Information Technology and Applications
- [6] Li Xu Dong. *Android Based Wireless Location and Surrounding Search System* , Conference Publications
- [7] Ming Hao. *Visual sentiment analysis on twitter data streams* , Hewlett-Packard Labs., Palo Alto, CA, USA - Visual Analytics Science and Technology (VAST), 2011 IEEE Conference - Conference Publications
- [8] Ovelgönne. *Social Emergency Alert Service - A Location-Based Privacy-Aware Personal Safety Service*, M. ,Next Generation Mobile Applications, Services and Technologies (NGMAST), 2010 Fourth International Conference
- [9] Takata, K. *A Dangerous Location Aware System for Assisting Kids Safety Care* - - Advanced Information Networking and Applications, 2006. AINA 2006. 20th International Conference - Conference Publications
- [10] Sentiment140 – Twitter Analysis Website <http://www.sentiment140.com/>
- [11] Ming Hao- VAST, 2009 - Visual Opinion Analysis of Customer Feedback Data