VTIS: A Volunteered Travelers Information System *

Roland Varriale
Department of Computer Science
University of Illinois at Chicago
Chicago, IL, USA
rvarri2@uic.edu

Shuo Ma
Department of Computer Science
University of Illinois at Chicago
Chicago, IL, USA
sma21@uic.edu

Ouri Wolfson
Department of Computer Science
University of Illinois at Chicago
Chicago, IL, USA
wolfson@cs.uic.edu

ABSTRACT

VTIS is a dynamic notification system that takes in a user's route and calculates the time-delay imposed by disruptions to the normal traversal. The disruptions are calculated by using crowdsourced notifications. This is accomplished by the creation of a client side application for notification display and a server infrastructure that will process and store the event information. We have devised a system that will generate personalized notifications for users based on a provided path, temporal range, and set of transportation modes. At a high level, the functionality of this system is to identify events that affect the user's route and notify the user of these events. The VTIS will provide a multimodal notification system based on information mined from Twitter data and volunteered information from VTIS users. This information will be stored to create a repository of transportation events., This repository will be queried to notify affected users of events that may affect their route.. Although this outlined problem has been solved previously, our approach is novel in several ways: (1)accounting for multiple modes, (2)combining user input with mined data, and the (3)modeling method used to calculate effects on the user's route. Some of these methods have been implemented separately; however, a comprehensive system has not been constructed that includes all of these items.

General Terms

Human Factors

Keywords

crowdsource, traveler information system, ITS, twitter, data mining

1. INTRODUCTION

The transportation problem of calculating disruptions to a user's route is difficult. Many services attempt to alleviate this difficulty; however, there are deficiencies that are evident. Commonly used services fall short in one of several regards: limitation to single transportation mode, requiring human intervention when processing data, or a lackluster feature set. To create a useful system it would be beneficial to alleviate these deficiencies.

The scope of our system encompasses the entire city of Chicago and surrounding suburbs. Many other systems that have been implemented such as ILDOT Traffic Alerts or CTA Bus Tracker force you to constrict your queries to certain predefined routes or segments along selected roadways. By utilizing a transportation network that includes many different modes we can alleviate these shortcomings.

The strength of this system lies in the ability to accurately model and calculate the effects of transportation events on a user's trajectory. One naïve approach is to check which events happen within some specified radius of the user's route; however, this approach does not accommodate the temporal aspects of transportation. To account for the temporal aspect, we employ an intensity modeling that takes into account several factors such as transportation mode, severity, distance, and time. Specifically, given an event, a location and a time elapsed since the event happens, the modeling outputs an intensity value which represents the impact that the event may have on the given location. Using the event's original location coupled with this calculated intensity value we can determine how impactful the event will be to a user's route.

The rest of the paper will be structured in the following way. Section 2 will outline previous work that has been performed within the field of traveler information systems. Section 3 will give a high level overview and highlight the advantages the VTIS provides over pre-existing systems. Section 4 will outline the procedure for entering a route and interacting with the system. Finally, Section 5 will demonstrate the methodology used to extract useful information and apply it to users' routes.

2. RELATED WORK

IBM and The University of Calgary have performed some work in creating transportation information infrastructures through Smarter City and TrafficPulse, respectively. The work that was done primarily utilized human processing power. Although the work accomplished is seminal in the field, it is highly desirable to process the data present within the tweet independent of human intervention. The work that has been done builds an excellent foundation for future work and development in this area; however, it seems insufficient

---

*This research was supported in part by the U.S. Department of Transportation National University Rail Center (NURAIL), Illinois Department of Transportation (METSI), National Science Foundation grants IIS-1213013, CCF-1216096, DGE-0549489

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.Request permissions from Permissions@acm.org.
ACM SIGSPATIAL IWCTS'13, Nov 05-08 2013, Orlando, FL, USA
Copyright 2013 ACM ISBN 978-1-4503-2527-1/12/11 ...$15.00.

insofar as it should be a reliable replacement for a real-time traffic observatory[1].

2.1 Smarter City

IBM has built a concept named Smarter City [1] that encompasses many sensors and information extraction techniques to build an infrastructure in which information is shared ubiquitously. This is accomplished by using social sensors such as crowd sensing[2] and traditional sensors, such as those found in cellular phones. Information is readily shared and accessed across platforms as to improve the overall experience of members within the Smarter City.

2.2 TrafficPulse

TrafficPulse[2] is a system created at the University of Calgary which harnesses users' cell phone sensors to identify traffic occurrences and situations, namely through the collection of spatiotemporal traffic data. The basic tenets of the system are (specified on their website):

- Enable participants to join the urban sensor web by volunteering their Smartphone's to perform opportunistic sensing tasks (e.g., collecting spatio-temporal traffic velocity data)
- Enable participants to join existing, or create new participatory sensing projects through the use of a Smartphone's touch-based GUI and an online social networking platform
- Enable the general public, through an open social networking sensor web portal, to browse, search for, access, and visualize TrafficPulse data.

Results are automatically tabulated; however, the findings are manually extracted.

2.3 Waze

Waze[3] is a navigation application that leverages users' input and cell phone sensors to update a geographic map in real-time. Hazards such as police, accidents, and traffic jams are identified by users and displayed on a map within the application so users may avoid these events. Users may also edit the map's geometry to correct any inaccuracies they may find. Neither event information nor map edits are verified in any way.This application is available on iOS and android across 20+ countries.

2.4 Illinois Department of Transportation (IDOT) Traffic Alerts[4]

This traffic alert system provides notifications based on several available criteria. Users may subscribe to certain segments and receive alerts twice per day by filling out the route information such as: segments of roadways , two times of day and days of the week you want to receive alerts, average speed threshold, and types of events to receive alerts for. Within the IDOT road infrastructure there are 72 road segments that are covered by this alert system.

2.5 CTA Bus Tracker

The CTA Bus Tracker[5] has an option for receiving alerts about specified bus routes. There are several options:

1. Receive alerts for all stops along a specific, fixed bus route
2. Specify a stop number, a start time, a time interval, an end time, and days of the week to repeat.

There are severe limitation to this service because of the modal restrictions. Although useful within its limited application, it is not extendible beyond the public bus mode. The CTA also provides a service for tracking trains[6] that is very similar in nature.

3. OVERVIEW OF THE VTIS

Previously, transportation data within Twitter was under utilized due to the inaccuracies of the geo-locational data associated with each tweet. Older methods were able to geo-locate tweets with an average error of 535.57 miles. [3] With new findings in the area of Twitter location mining, a tweet can be more precisely located by using landmarks and street names to further constrain the location [1]. Using this technology we can place information from verified Twitter sources along a route chosen by a user utilizing a subscription/notification system. Moreover, we can utilize these sources to create a machine learning classifier that leverages trusted data to identify useful tweets from unknown sources. Furthermore, Twitter serves as a desirable platform for informational transport due to the widespread accessibility and expeditious publication of real-time events [4, 5, 3]. Twitter's growth in recent years as a medium to broadcast news and events is increasing at an impressive rate[6], and makes it an enticing choice as a vehicle by which travel information can be conveyed.

In addition to retrieving relevant information from Twitter, users are able to submit events that are observed along their route. This functionality will allow for inclusion of events that are difficult to detect, or where data is very sparse. For example, an accident that occurs on an arterial street in the suburbs may not attract the attention of twitter accounts; however, if a user enters the event it will be available for notifications to other users.

We feel that trust of the users' input is a very important aspect to the functionality of this VTIS. Many different types of attacks to Intelligent Transportation Systems (ITS) have previously been outlined[7]. Currently no security features are present that would handle such attacks; however, these attacks will be taken into future consideration to preserve the integrity of the system.

Many previous systems such as those described in Section 2 focus on single modes of transportation. Although this is important we feel that encompassing multiple transportation modes allows for better functionality on multiple counts. We can encompass events occurring in more than one transportation mode. This is important because transportation modes often affect each other. For example, if a car crashes on a busy street it will delay multiple modes: auto, bus, and bicycle. In large cities it is very common for travelers to utilize many modes of transportation within

---

[1] http://www-03.ibm.com/innovation/us/thesmartercity
[2] http://sensorweb.geomatics.ucalgary.ca/projects/trafficpulse
[3] http://www.waze.com
[4] http://www.iltrafficalert.com/
[5] http://www.ctabustracker.com/bustime/eta/eta.jsp
[6] http://www.transitchicago.com/traintracker/

No	Filter(s)	Tweets Retrieved	Duration(mins)	Number of positive tweets
1	C and U	3808	120	61
2	C and K	29628	32	10
3	C, U and K	75945	22	0
4	Only U	67	37	15
5	Only K	24329	32	18
6	Only C	3578	60	0
C = Coordinates	U= UserIDs	K = Keywords		

Table 1: Statistics using various filters on tweets

the same day, sometimes within the same trip. Previously multiple applications or interfaces would need to be utilized to receive alerts for each mode separately and then compile them manually. Within the VTIS this is done automatically and the correlations between the modes are incorporated.

4. ROUTE CREATION AND SELECTION

We start by formally defining terms that will be used within the VTIS.

Definition 1. *(leg) A* ***leg*** *will be a directed path comprised of a single mode of transportation.*

Definition 2. *(route) A* ***route*** *will be an ordered set of legs, which will be specified by the user in the order they should be traversed.*

Provided a set of ordered legs, L = $\{L_1, L_2, .., L_n\}$, the respective Route, R, would be constructed by concatenating ($\bullet$) each leg to the previous leg such that R = $\{L_1 \bullet L_2 \bullet \ldots \bullet L_n\}$.

To initially interact with the VTIS a user must submit a spatial route with structure dictated by the Google Maps web application. Each leg should contain a start point, end point, and mode. If no start point is provided the previous end point will be used as the current start point. If these legs are followed, in order, the user should successfully depart the starting point and finish at the ending point. This will result in a number of ordered legs called a route, but also colloquially referred to as directions. A set of directions comprised of legs is depicted in Figure 1. This submission will be saved under the user's account.

Proper user input should be composed of a single, finite set of directions that can be easily decomposed into an ordered set of legs by using standard nomenclature set by Google Maps, see Figure 1.

Definition 3. *(subscription) A subscription S will consist of a tuple containing a user, a threshold(ψ), and a single selection from a set of routes* $(u, \psi, \{r_0 \ldots r_n\})$.

It is possible for a user to have multiple subscriptions; however, only one may be active at a time. This choice will be made from within the application. A user subscription will begin when the user chooses a trip from a list of previously created trips, selects a polling time length, and initiates the notification system. The dictated polling time to retrieve transportation events will allow users to control the amount of queries the application makes to minimize data usage and battery power. When specifying a route, a user must also specify some threshold he or she is willing to withstand in order to receive more accurate notifications. This threshold value will correspond to a value returned by an Intensity Function (IF). Upon detection of an event, the IF will be calculated for each mode, and affected routes with a threshold value that is less than or equal to the value returned by the IF will receive a notification. When these steps are completed the user's subscription will be said to have started. The subscription will not have a temporal range; however, it will complete when the user reaches the end point specified within the active trip. While the route is activated the user will receive notifications for all transportation events that may affect his or her trip. This will be outlined in detail within Section 5.5

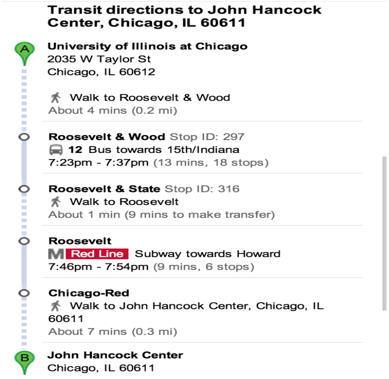
5. IMPLEMENTATION

5.1 Event Collection

The VTIS obtains data through two main methods: information extraction from Twitter and user input from the VTI application. Both sources are treated equally in terms of data validity and relevance. This is to say that we trust data coming from both sources in the same fashion.

5.1.1 Processing Twitter Event Information

At the heart of the VTIS application lies the ability to identify transportation events on Twitter and extract the relevant information that is available. It has been shown that using specific Twitter accounts pertaining to travel and automotive transportation authorities, such as the CTA and CDOT, can yield very promising results when used in conjunction with a gazetteer or dictionary containing local streets and landmarks [2]. Identifying and extracting useful event information will be accomplished by selecting rele-



Transit directions to John Hancock Center, Chicago, IL 60611

A
University of Illinois at Chicago
2035 W Taylor St
Chicago, IL 60612

Walk to Roosevelt & Wood
About 4 mins (0.2 mi)

Roosevelt & Wood Stop ID: 297
12 Bus towards 15th/Indiana
7:23pm - 7:37pm (13 mins, 18 stops)

Roosevelt & State Stop ID: 316
Walk to Roosevelt
About 1 min (9 mins to make transfer)

Roosevelt
M Red Line Subway towards Howard
7:46pm - 7:54pm (9 mins, 6 stops)

Chicago-Red
Walk to John Hancock Center, Chicago, IL 60611
About 7 mins (0.3 mi)

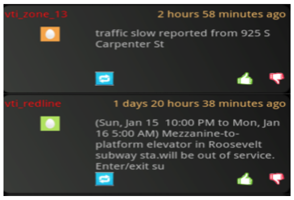
B
John Hancock Center
Chicago, IL 60611

Figure 1: A sample route on Google Maps

vant tweets through machine learning techniques. Using the Twitter API we can access these accounts using keywords and phrases obtain a corpus to train a classifier. This technique is preferable to string based or keyword based searching in both quantity and quality of data being retrieved as seen in Table 2.

5.1.2 Crowdsourcing from Travelers

Users of the VTIS are a valuable source of information, especially in areas where transportation sources are sparse. The VTIS allows users to input transportation events directly from within the application on their mobile device. Users are prompted for a location (exact or cross streets), classification of event (such as those in Table 6), and time observed. Once this data is submitted the event is stored within the VTIS for retrieval through notifications. In addition to inputting transportation events users will also be able to vote on the usefulness of events that have been detected or submitted. Upvote/downvote buttons will be present at the bottom of each notification within the VTIS as shown in figure 2. Once a vote is cast it should be stored in the event database and reflected in the user's notifications, once refreshed. This will assist travelers in discerning the validity of information that may have become stale or inaccurate.



vti_zone_13 2 hours 58 minutes ago

traffic slow reported from 925 S Carpenter St

vti_redline 1 days 20 hours 38 minutes ago

(Sun, Jan 15 10:00 PM to Mon, Jan 16 5:00 AM) Mezzanine-to-platform elevator in Roosevelt subway sta.will be out of service. Enter/exit su

Figure 2: Upvoting/Downvoting within the VTIS Applcation

5.2 Extracting a List of Transportation Authorities

We begin by compiling a list of transportation authorities of a chosen city, in our case Chicago. This list will assist us in identifying relevant and useful transportation tweets. Utilizing the Twitter API we can search by keywords such as: *reroute, earlier, closed, late, arrive, ramp, train, jam, accident, clear, traffic, detour, road*; and save the results [8, 9]. Since the Twitter API does not allow for multiple, simultaneous filters when performing this query we must separately query by location and then keyword. Once we perform these searches separately we can store the intersection of the two result sets. Once we cross reference the accounts we will have a list of active users tweeting about transportation events. This results in a list similar to Table 3. By using this technique we facilitate the portability of the VTIS. We will also take secondary sources that are trusted by our verified sources as truth. Such secondary sources may be people working for associations that interact with the aforementioned verified sources such as newscasters, local organizations, or government employees. By leveraging each of these categories we can build a more customized subscription/notification system tailored to meet each user's transportation needs.

5.2.1 Creating a Training Set and Classifier

No	User Name	Number of tweets
1	TotalTrafficCHI	2493
2	CMoleChicago	1966
3	thewatchmantwit	304
4	cta	104
5	metraBNSF	95
6	94_294_Tollway	66
7	TotalTrafficGRR	63
8	SarahJindra	56
9	metraUPNW	54
10	I_90_Tollway	48
11	ILI88thm	47
12	ILI55thm	46
13	ILI80thm	45
14	KyeCommuteNBC	44
15	Chicago_CP	43
16	ILI57thm	37
17	ChicagoDrives	37
18	MikePriesWBBM	34
19	ILI94thm	34
20	ILI90thm	33

Table 3: Chicago Twitter Transportation Authorities by Tweet volume

We create a training set from the tweets of these transportation authorities' accounts on Twitter and treat those as ground truth. Furthermore, we restrict our tweets to a bounded box, in this case within the City of Chicago and closely surrounding suburbs, to focus our efforts geographically. All tweets within that box are captured and scrubbed of extraneous information such as replies (@someUser), retweets (RT), and hashtags (#someWord). The Twittionary API is utilized to normalize words or phrases that are not common such as: lol, omg, etc. Once this is completed the tweets are saved for further inspection as test data. Events within these saved tweets are detected by using the Support Vector Machine (SVM) method.

$$w_i = tf_i \text{ X } idf_i$$

w_i is the weight of word i
tf_i is the term frequency of term i
idf_i is the log of total documents compared to the documents in which the term i appears

Certain CTA and transportation authority accounts (shown in Table 3 are used as ground truth in relation to the classifier. Tweets generated from these sources are used to train the classifier in order to detect other transportation related tweets. As we can see in Table 2 using machine learning/data mining techniques are superior in quantity to naïve string searching or keyword based searches alone. This conforms with similar work that was done in this area[10].

We have also devised a separate method to process tweets containing transportation events by using natural language processing. Once collected we parse each tweet and use a part of speech tagger (Stanford NLP POS tagger) to identify and mark each word with the appropriate part of speech. After each tweet has been marked we calculate a score based on the parts of speech that are present within the tweet. The results of using this method were favorable in identifying

Number of instances in test file	Using regular expression and thumb rules	Data Mining	Positive instances
9591	141	233	230
3782	62	149	147
1491	62	73	70
2085	44	76	76
1305	15	58	10
8055	156	359	357

Table 2: Transportation events by source

	Classified positive	Classified negative
Actual positive	2939	114
Actual negative	211	1145

Table 4: Naïve Bayes Confusion Matrix for Precision and Recall P=2939/(2939+211) = 0.93 r=2939/(2939+114) = 0.96

	Classified positive	Classified negative
Actual positive	3045	8
Actual negative	19	1337

Table 5: SVM Confusion Matrix for Precision and Recall p= 3045/(3045 + 19) = 0.99 r= 3045/(3045 + 8) = 0.99

relevant tweets and this method will be integrated with the previously described method for future use within the VTIS.

5.2.2 Mining the Twitter Data

Once the classifier is trained it is run on all tweets fetched within the bounded box mentioned in Section 5.2.1 and they are labeled as 0: non-transportation related, or 1: transportation related. In our example this will pertain to all tweets that are contained within the city of Chicago and nearby suburbs. Each tweet within this bounded box will be captured, inspected, and labeled. All tweets containing a ’1’ label are then saved within a relational database containing the following information: username, geographic location, mode, and time stamp. It may not be possible to obtain an accurate geographic location purely based on the information provided within the tweet metadata; however, it will be recorded to the most accurate extent possible[11]. If the transportation mode cannot be determined then all modes will be recorded for processing. The classifier trained using the Twitter accounts from Table 3 has higher recall and precision than the Naïve Bayes approach. This is shown in Tables 4 and 5.

5.3 VTIS Specific Notifications

Along with notifications that are generated by mining data on Twitter, we allow users to submit events that they observe along their route. A user must specify the location using an address or cross streets, an event classification, and the observation time. Once the information is submitted it will be recorded within the database for processing. This data treated in the same fashion as the Twitter data in accordance with the intensity calculations and notification procedures.

5.4 Calculating the influence of events on a Route

Definition 4. ***(Severity)*** *We will use the severity, S, of the event such that S ϵ {1... 10}.*

The severity of an event will depend strictly upon the type of event that is detected. Certain transportation events will exhibit a larger scale of influence upon travel time. For example, an accident occurring on a road would affect the overall travel time more than light traffic or a road closure. These event descriptions will conform to the standards set forth by ANSI D16.1-2007[7] Our intuition is that traffic events follow the propagation of shockwaves very closely, namely through kinematic waves. Kinematic waves have been used to model traffic events that perform in similar ways to waves on a body of water. Although originally demonstrated by Lighthill, Whitman, and Richards (LWR)[12, 13] there has been a significant amount of work done in this area, namely by Kerner[14, 15]. The premise behind this model is the preservation of the relationship between flow, density, and speed. We can observe a similar relationship between intensity, distance, and time with the VTIS. We adapt this LWR model to calculate the intensity of a transportation event upon a user’s route. This Intensity Function that will process a transportation event and calculate a value based on the transportation mode of the event, transportation mode of the route, distance, severity, and time.

Event	Severity
Accident	8
Delay	3
Road Closure	3
Light Traffic	5
Heavy Traffic	9
Fatality	10

Table 6: Sample Event Severity Measurement

By utilizing traffic data available to us by the Gateway Traveler System(GTS)[8] we can verify this intensity calculation. A myriad of information is available through the use of sensor data through the GTS site in the form of reports. By using the information within the Incident Report, which contains data such as: description of travel event, location of event, details, current status, start time, and approximate end time; we can verify events that are being reported or detected through the VTIS. Moreover, we can inspect the congestion caused by these events by utilizing the congestion, travel time, and detector reports.

---

[7] http://www-nrd.nhtsa.dot.gov/Pubs/07D16.pdf

[8] http://www.travelmidwest.com/lmiga/home.jsp

Mode	Rail	Car	Bus	Subway	Pedestrian
Rail	1.00	0.60	0.60	0.00	0.00
Car	0.60	1.00	1.00	0.00	0.80
Bus	0.60	1.00	1.00	0.00	0.80
Subway	0.00	0.00	0.00	1.00	0.00
Pedestrian	0.00	0.80	0.80	0.00	1.00
Bicycle	0.00	1.00	0.60	0.00	1.00

Table 7: Sample Modal Affect Matrix

We know that certain modes have greater interaction than others due to shared resources. For example, cars and buses share common roads as well as pedestrians and cyclists. These relationships directly affect the amount of influence an event with some mode has upon another mode. In Table 7 we outline these relationships insofar as they will be used to calculate influence. This table can be generate by comparing resources, such as roads, that are shared between modes. To create the actual table we would need to generate lists of available edges used by each transportation mode and identify common edges in each other mode of transportation. The number of common edges between any two modes divided by the total edges for each mode would yield the value shown in Table 7. For example, if the auto mode had 537 edges and 240 of those were common with bicycles; the bicycle mode is said to affect the auto mode with .45. By combining both an event's severity and the modal affect we can create a more complete model to represent the influence of some event on a specific route.

We can clearly see there are some transportation modes that have no influence on each other. This represents the disjoint nature of spatial resources between those two transportation modes. In contrast, a coefficient of 1.00 shows a 1:1 correlation between two transportation modes, which would be seen where all spatial resources are common. For example, in cities where trains are underground(subway) there would be no influence from any mode using above ground transportation, so the modal affect would be 0. Ideally the coefficients in Table 7 would represent the amount of shared spatial resources between any two modes.

5.5 Notifications

Notifications will occur when transportation events are detected and calculated to have influence upon the user's route. The user's location will be detected by leveraging the GPS sensor on his/her smartphone. As defined in Section 5.4, whenever an event is detected with intensity higher than the user's specified threshold, a notification will be generated and pushed to the user.

The polling time dictated by the user's preference will be used to query events recorded using the methods outlined in Section 5.1.1. This polling time will be a length of time which, upon passage of its specified value, the client side application will perform a query to the VTIS server using the user's route, location, and intensity threshold. A notification will be generated to the user if a transportation event is detected that intersects their path with an intensity higher than the user's supplied threshold. These notifications will be presented in the form of a list similar to aesthetics of the Twitter timeline. This list will be refreshed according to the user specified polling time when the VTIS is started. Notifications will be specific to the user's route and current location. No notifications will be generated for segments of a route that a user has already traversed.

Each notification should contain a normalized version of the detected tweet containing the event being described such as: location, severity, time, and user votes(described in Section 5.1.2. The VTIS will attempt to retrieve as much information as possible; however, it may be the case that the information is not presented and cannot be inferred. In that case, the publication will furnish as much information as possible.

6. REFERENCES

[1] Sílvio S Ribeiro Jr, Clodoveu A Davis Jr, Minas Gerais, Tatiana S Gonçalves, and Gisele L Pappa. Traffic Observatory : a System to Detect and Locate Traffic Events and Conditions using Twitter.

[2] Haggai Roitman et al. In *Harnessing the crowds for smart city sensing*, Maui, Hawaii, 2012. ACM Press.

[3] Zhiyuan Cheng, James Caverlee, and Kyumin Lee. You Are Where You Tweet : A Content-Based Approach to Geo-locating Twitter Users. pages 759–768.

[4] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake Shakes Twitter Users : Real-time Event Detection by Social Sensors. 2009.

[5] Graham Neubig. Safety Information Mining - What can NLP do in a disaster -. pages 965–973, 2011.

[6] Todd Wasserman. Facebook's u.s. growth slowing, but twitter's on a tear [study]. 2012.

[7] Shuo Ma, Ouri Wolfson, and Jie Lin. A survey on trust management for intelligent transportation system. In *Proceedings of the 4th ACM SIGSPATIAL International Workshop on Computational Transportation Science*, pages 18–23. ACM, 2011.

[8] Xiaohua Liu, Shaodian Zhang, Furu Wei, and Ming Zhou. Recognizing Named Entities in Tweets. (2008):359–367, 2011.

[9] David A Shamma and Elizabeth F Churchill. Tweetgeist: Can the twitter timeline reveal the structure of broadcast events? 2010.

[10] Alexander Pak and Patrick Paroubek. Twitter as a corpus for sentiment analysis and opinion mining. In *LREC*, 2010.

[11] Falko Schmid and Kai-Florian Richter. Extracting places from location data streams. International Workshop on Ubiquitous Geographical Information Services, 2006.

[12] Michael J Lighthill and Gerald Beresford Whitham. On kinematic waves. ii. a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 229(1178):317–345, 1955.

[13] Paul I Richards. Shock waves on the highway. *Operations research*, 4(1):42–51, 1956.

[14] Boris S Kerner and H Rehborn. Experimental features and characteristics of traffic jams. *Physical Review E*, 53(2):R1297, 1996.

[15] Boris S Kerner and P Konhäuser. Cluster effect in initially homogeneous traffic flow. *Physical Review E*, 48(4):R2335, 1993.