INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 3 ISSUE 1 (2017) PAGES 1558 - 1562
RECEIVED : 5/01/2017. PUBLISHED : 19.01.2017

November 19, 2016

REAL-TIME DETECTION OF TRAFFIC FROM TWITTER STREAM ANALYSIS

[1]A. Kavipriya M.Phil.,[2]Mrs. R. Vasugi M.Sc., M,Phil.,

[1]Research Scholar, Department of Computer Science, Bharathiyar Arts and Science College for Women's, Deviyakurichi, Salem.

[2]Assistant Professor, Department of Computer Science, Bharathiyar Arts and Science College for Women's, Deviyakurichi, Salem.

Abstract

Social networks have been recently employed as a source of information for event detection, with particular reference to road traffic congestion and car accidents. In this paper, we present a real-time monitoring system for traffic event detection from Twitter stream analysis. The system fetches tweets from Twitter according to several search criteria; processes tweets, by applying text mining techniques; and finally performs the classification of tweets. The aim is to assign the appropriate class label to each tweet, as related to a traffic event or not. 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. We employed the support vector machine as a classification model, and we achieved an accuracy value of 95.75% by solving a binary classification problem (traffic versus non-traffic tweets). We were also able to discriminate if traffic is caused by an external event or not, by solving a multi class classification problem and obtaining accuracy value of 88.89%.

Keyword: Traffic, Network, Network-based anonymization and processing (NAP)

1. INTRODUCTION

Network security consists of the policies and practices adopted to prevent and monitor unauthorized access, misuse, modification, or denial of a computer network and network-accessible resources. Network security involves the authorization of access to data in a network, which is controlled by the network administrator. Users choose or are assigned an ID and password or other authenticating information that allows them access to information and programs within their authority. Network security covers a variety of computer networks both public and private that are used in everyday jobs; conducting transactions and communications among businesses, government agencies and individuals. Networks can be private, such as within a company, and others which might be open to public access. Network security is involved in organizations, enterprises, and other types of institutions. It does as its title explains: It secures the network, as well as protecting and overseeing operations being done. The most common and simple way of protecting a network resource is by assigning it a unique name and a corresponding password.

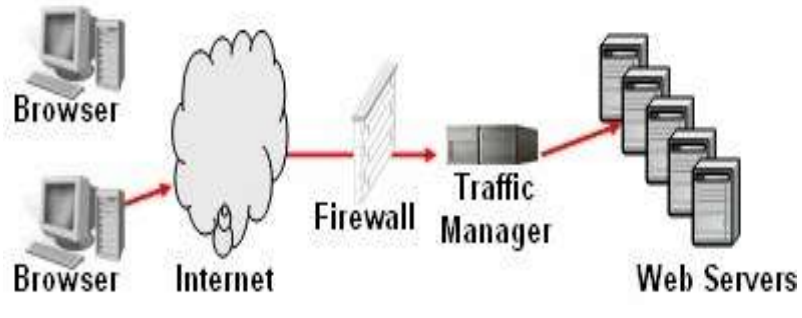

Fig 1. Network Traffic Manager

2. RELATED WORK

In this paper "A Survey on the Application of Genetic Programming to Classification" Classification is one of the most researched questions in machine learning and data mining. A wide range of real problems have been stated as classification problems, for example credit scoring, bankruptcy prediction, medical diagnosis, pattern recognition, text categorization, software quality assessment, and many more.The use of evolutionary algorithms for training classifiers has been studied in the past few decades. Genetic programming (GP) is a flexible and powerful evolutionary technique with some features that can be very valuable and suitable for the evolution of classifiers. This paper surveys existing literature about the application of genetic programming to classification, to show the different ways in which this evolutionary algorithm can help in the construction of accurate and reliable classifiers.

In this paper "Tweet Analysis for Real-Time Event Detection and Earthquake Reporting System Development" Twitter has received much attention recently. An important characteristic of Twitter is its real-time nature. We investigate the real-time interaction of events such as earthquakes in Twitter and propose an algorithm to monitor tweets and to detect a target event. To detect a target event, we devise a classifier of tweets based on features such as the keywords in a tweet, the number of words, and their context. Subsequently, we produce a probabilistic spatio temporal model for the target event that can find the center of the event location. We regard each Twitter user as a sensor and apply particle filtering, which are widely used for location estimation. The particle filter works better than other comparable methods for estimating the locations of taget events. As an application, we develop an earthquake reporting system for use in Japan. Because of the numerous earthquakes and the large number of Twitter users throughout the country, we can detect an earthquake with high probability (93 percent of earthquakes of Japan Meteorological Agency (JMA) seismic intensity scale 3 or more are detected) merely by monitoring tweets. Our system detects earthquakes promptly and notification is delivered much faster than JMA broadcast announcements.

3. PROPOSED SYSTEM

Network-based anonymization and processing (NAP) framework, the first system for K-anonymous query processing in road networks. NAP relies on a global user ordering and bucketization that satisfies reciprocity and guarantees K-anonymity. We identify the ordering characteristics that affect subsequent processing, and qualitatively compare alternatives. Then, we propose query evaluation techniques that exploit these characteristics. In addition to user privacy, NAP achieves low computational and communication costs, and quick responses overall. It is readily deployable, requiring only basic network operations. We propose a framework for anonymous query processing in road networks. We design location obfuscation techniques that (i) provide anonymous LBS access to the users, and (ii) allow efficient query processing at the LBS side. Our techniques exploit existing network database infrastructure, requiring no specialized storage schemes or functionalities. We experimentally compare alternative designs in real road networks and demonstrate the effectiveness of our techniques.

4. SYSTEM MODEL

- Dataset preprocessing;
- Location Base Services:
- System model for ASR:
- Anonymous location-based queries:
- Anonymous query processing:

Dataset preprocessing;

- Server - Client computing or networking is a distributed application architecture that partitions tasks or workloads between service providers (servers) and service requesters, called clients. Often clients and servers operate over a computer network on separate hardware. A server machine is a high-performance host that is running one or more server programs which share its resources with clients. A client also shares any of its resources; Clients therefore initiate communication sessions with servers which await (listen to) incoming requests.

Location Base Services:

- Users are reluctant to use LBSs, since revealing their position may link to their identity. Even though a user may create a fake ID to access the service, her location alone may disclose her actual identity. Linking a position to an individual is possible by various means, such as publicly available information city maps.
- When a user u wishes to pose a query, she sends her location to a trusted server, the anonymizer through a secure connection (SSL). The latter obfuscates her location, replacing it with an anonymizing spatial region (ASR) that encloses u.
- The ASR is then forwarded to the LS. Ignoring where exactly u is, the LS retrieves (and reports to the AZ) a candidate set (CS) that is guaranteed to contain the query results for any possible user location inside the ASR. The AZ receives the CS and reports to u the subset of candidates that corresponds to her original query.

System model for ASR:

- The ASR construction at the anonymization process abides by the user's privacy requirements. Particularly, specified an anonymity degree K by u, the ASR satisfies two properties: (i) it contains u and at least another K * 1 users, and (ii) even if the LS knew the exact locations of all users in the system.
- We propose an edge ordering anonymization approach for users in road networks, which guarantees K-anonymity under the strict reciprocity requirement (described later).

Anonymous location-based queries:

- Considerable research interest has focused on preventing identity inference in location-based services. Proposing spatial cloaking techniques. In the following, we describe existing techniques for ASR computation (at the AZ) and query processing (at the LS). At the end, we cover alternative location privacy approaches and discuss why they are inappropriate to our problem setting. This offers privacy protection in the sense that the actual user position u cannot be distinguished from others in the ASR, even when malicious LS is equipped/advanced enough to possess all user locations. This spatial K-anonymity model is most widely used in location privacy research/applications, even though alternative models are emerging.

Anonymous query processing:

- Processing is based on implementation of the theorem uses (network-based) search operations as off the shelf building blocks. Thus, the NAP query evaluation methodology is readily deployable on existing systems, and can be easily adapted to different network storage schemes. In this case, the queries are evaluated in a batch. we propose the network-based anonymization and processing (NAP) framework, the first system for K- anonymous query processing in road networks. NAP relies on a global user ordering and bucketization that satisfies reciprocity and guarantees K-anonymity. We identify the ordering characteristics that affect subsequent processing, and qualitatively compare alternatives. Then, we propose query evaluation techniques that exploit these characteristics. In addition to user privacy, NAP achieves low computational and communication costs, and quick responses overall. It is readily deployable, requiring only basic network operations.

Flow Diagram

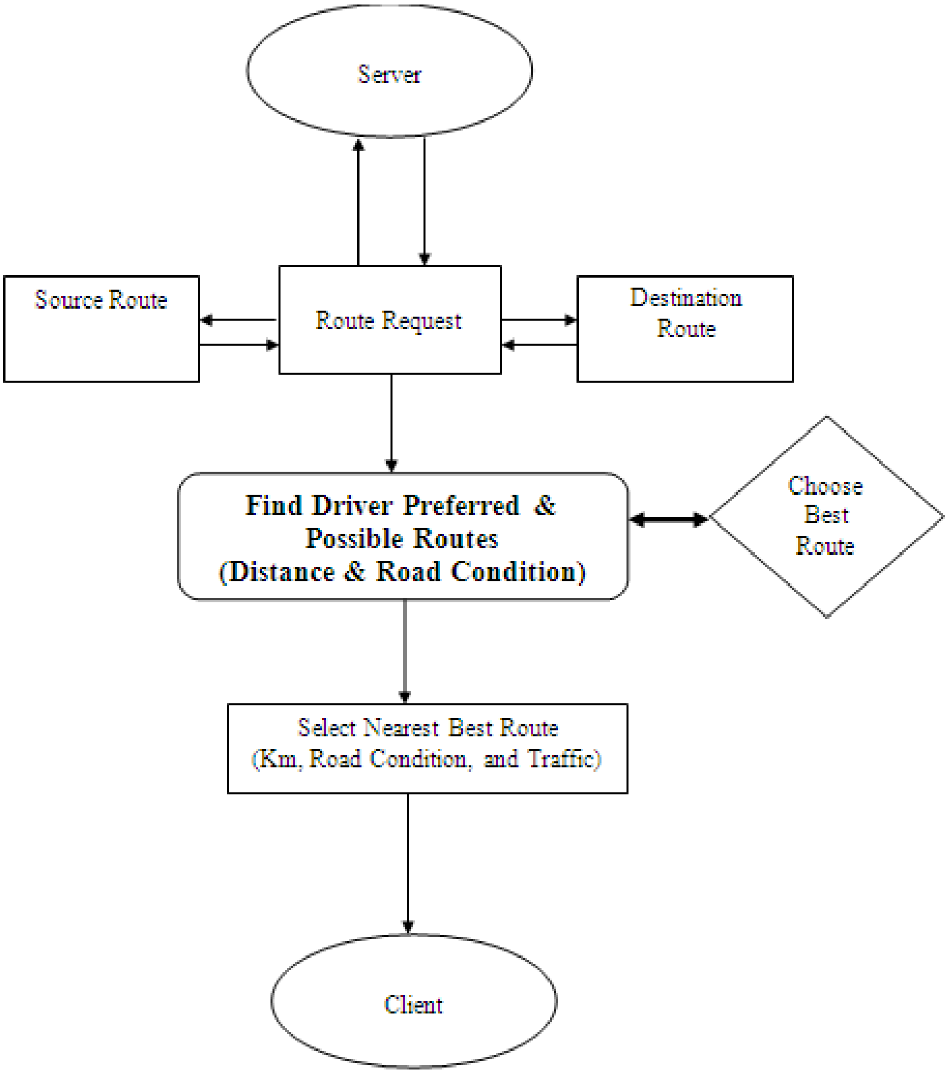
Server
Source Route
Route Request
Destination Route
Find Driver Preferred & Possible Routes (Distance & Road Condition)
Choose Best Route
Select Nearest Best Route (Km, Road Condition, and Traffic)
Client

Fig 2. Find Router Condition

CONCLUSION

Query evaluation techniques that exploit these characteristics. In addition to user privacy, NAP achieves low computational and communication costs, and quick responses overall. It is readily deployable, requiring only basic network operations. In the traditional spatial anonymity model, the data owner (e.g., a location-based service) makes its data available using a location server. It may, however, be the case that the owner is outsourcing its database to a third-party (and, thus, untrusted) location server. A challenge here is how to encrypt the owner's data so that they are hidden from the location server, while it can still process anonymous queries. Another interesting question is how (anonymous) users could verify that the location server did not tamper with the original owner data.

REFERENCE

[1] F. Atefeh and W. Khreich, "A survey of techniques for event detection in Twitter," *Comput. Intell.*, vol. 31, no. 1, pp. 132–164, 2015.

[2] P. Ruchi and K. Kamalakar, "ET: Events from tweets," in *Proc. 22nd Int. Conf. World Wide Web Comput.*, Rio de Janeiro, Brazil, 2013, pp. 613–620.

[3] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in *Proc. 7th ACM SIGCOMM Conf. Internet Meas.*, San Diego, CA, USA, 2007, pp. 29–42.

[4] G. Anastasi *et al.*, "Urban and social sensing for sustainable mobility in smart cities," in *Proc. IFIP/IEEE Int. Conf. Sustainable Internet ICT Sustainability*, Palermo, Italy, 2013, pp. 1–4.

[5] A. Rosi *et al.*, "Social sensors and pervasive services: Approaches and perspectives," in *Proc. IEEE Int. Conf. PERCOM Workshops*, Seattle, WA, USA, 2011, pp. 525–530.

[6] T. Sakaki, M. Okazaki, and Y.Matsuo, "Tweet analysis for real-time event detection and earthquake reporting system development," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 4, pp. 919–931, Apr. 2013.

[7] J. Allan, *Topic Detection and Tracking: Event-Based Information Organization*. Norwell, MA, USA: Kluwer, 2002.

[8] K. Perera and D. Dias, "An intelligent driver guidance tool using location based services," in *Proc. IEEE ICSDM*, Fuzhou, China, 2011, pp. 246–251.