Pervasive and Mobile Computing 20 (2015) 1–12

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

A privacy mechanism for mobile-based urban traffic monitoring

Chi Wang [a], Hua Liu [b], Kwame-Lante Wright [b,*], Bhaskar Krishnamachari [b], Murali Annavaram [b]

[a] Microsoft Research, Redmond, WA, USA

[b] University of Southern California, Los Angeles, CA, USA

ARTICLE INFO

Article history:
Received 25 July 2013
Received in revised form 25 October 2014
Accepted 20 December 2014
Available online 27 December 2014

Keywords:
Mobile computing
Traffic monitoring
Privacy

ABSTRACT

In mobile-based traffic monitoring applications, each user provides real-time updates on their location and speed while driving. This data is collected by a centralized server and aggregated to provide participants with current traffic conditions. Successful participation in traffic monitoring applications utilizing participatory sensing depends on two factors: the information utility of the estimated traffic condition, and the amount of private information (position and speed) each participant reveals to the server. We assume each user prefers to reveal as little private information as possible, but if everyone withholds information, the quality of traffic estimation will deteriorate. In this paper, we model these opposing requirements by considering each user to have a utility function that combines the benefit of high quality traffic estimates and the cost of privacy loss. Using a novel Markovian model, we mathematically derive a policy that takes into account the mean, variance and correlation of traffic on a given stretch of road and yields the optimal granularity of information revelation to maximize user utility. We validate the effectiveness of this policy through real-world empirical traces collected during the Mobile Century experiment in Northern California. The validation shows that the derived policy yields utilities that are very close to what could be obtained by an oracle scheme with full knowledge of the ground truth.

1. Introduction

In existing sensor networks, power-constrained sensors are deployed in the targeted area and data is collected until the sensors run out of battery power or the collection time window expires. There are several disadvantages in such traditional sensor networks. First, the size of the sensor network is usually small. Second, most sensors are power constrained and hence may require replacing or recharging of their batteries; either of these tasks is intrusive to the sensing process and can sometimes be time consuming if the sensing environment is not easily accessible.

1.1. Motivation for participatory sensing

In order to overcome the shortcomings of traditional sensor networks, researchers have proposed projects such as *MetroSense* [1] and *Participatory Sensing* [2]. This new generation of sensing projects is based on the concept of "people-centric

* Corresponding author.
E-mail addresses: chiw@microsoft.com (C. Wang), hual@usc.edu (H. Liu), kwamelaw@usc.edu (K.-L. Wright), bkrishna@usc.edu (B. Krishnamachari), annavara@usc.edu (M. Annavaram).

http://dx.doi.org/10.1016/j.pmcj.2014.12.007
1574-1192/

sensing" at a large scale (e.g., campus, town, or metropolis). People are central to the sensing experience and represent the key architectural component in this new paradigm. In this category of sensing, human-carried sensors are brought into the environment that we are interested in sensing. The key element of such sensing is that people might be sensing their surroundings as they go about their daily activities without even making any explicit effort to sense. Mobile phones have become a key enabler for such *passive sensing*. Mobile phones are typically equipped with several integrated sensors, such as GPS, microphones, Bluetooth, and Wi-Fi. These features make the phones attractive for such participatory sensing projects.

In this paper, we focus on one particular participatory sensing application, namely urban traffic monitoring. In this traffic monitoring application, sensors, namely GPS, are integrated either into a mobile phone or into a user's vehicle. These sensing systems have the potential to radically improve the accuracy and timeliness of traffic information. In this application, several users driving on various road segments can use their GPS-enabled sensors to accurately determine their speed and position information. The measured information is then transmitted to a backend aggregation server. The aggregator collects segmented traffic reports from individual users and combines the reports to obtain complete traffic conditions of an entire road stretch. The global traffic information is in turn used by the aggregator to provide real-time traffic and travel time estimates to all the users in the system. Traffic sensing is an important application class where the accuracy of traffic estimation improves with the increasing number of participants.

1.2. Importance of privacy

While the motivation for traffic sensing using mobile phones is clear, the approach described above where the user reports speed and position information to the aggregator, potentially compromises the participant's privacy. In traditional sensor networks, since a sensor node is not associated with a particular individual the need for privacy is relatively low. However, in participatory sensing, when a mobile phone is being used as a sensor, the sensing device and the participant are closely tied together. A mobile phone identifies the sensor uniquely with a participant's identity. The data sensed is not only indicative of the participant's surroundings, but also reveals the participant's location and speed. Hence, we have to take the device holder's (application subscriber) privacy into account when designing the system. If the accurate location/speed information is intercepted by malicious attackers, the attackers can reveal the phone's identity by investigating the MAC layer packet headers. Once the identity of the device holder is revealed with *precise* location and speed information, the participant is exposed to the attacker. Imagine the day when an unwary traffic sensing participant gets a speeding ticket as an SMS message!

The goal of this paper is to study the privacy risks in traffic sensing. In order to protect users' privacy, we derived a utility based application method, which lets the users update the system with "just enough" information to the backend server, trading some data accuracy for improved user privacy. In this research, we consider the *location granularity* as a mechanism to obfuscate the users' precise location information. For instance, using a coarse location granularity the user can inform the aggregator that he/she is currently driving *somewhere* between two exits on a freeway without disclosing the precise location. Privacy is better protected but the system can still maintain reasonable service quality.

The paper is organized as follows. Section 2 describes the traffic monitoring application. Sections 3 and 4 depict our novel mathematical formulation of the problem, including the Markov-based road condition model and utility modeling. In Section 5, we propose a practical policy that suggests a near-optimal decision on maximizing a user's utility. Our experimental methodology and results are presented in Sections 6 and 7. Finally we present some related works in Section 8 and conclude our work in Section 9.

2. Application description

We believe that mobile based urban traffic monitoring systems will help relieve traffic conditions in the future and help application users estimate traffic conditions on the road with privacy taken into consideration.

In the simplest version of this application, we envision the use of virtual trip lines (VTLs) [3] to help coordinate data gathering. Virtual trip lines are GPS coordinates of a line that is *virtually* drawn on top of any road segment by the traffic application administrator. Mobile devices monitor their location using GPS and when they cross a VTL the device sends a raw update to a backend server with accurate position (VTL id) and speed information. The backend server aggregates the information obtained from multiple devices and uses it to estimate the current traffic conditions and provide accurate traffic and drive time estimates back to the mobile devices in real time. This information can then be used to alert the vehicle drivers about possible traffic congestions and even suggest alternate routes.

However, for the users on the road, the major privacy concerns are focused on users' exact location and speed. If the user's update information is overheard, or maliciously detected by eavesdroppers, the user's privacy is compromised by revealing the exact location and speed information. Even though the application may not need the user's identity when collecting the traffic condition updates, the MAC layer of the mobile devices implicitly reveals a user's identity by the MAC address. In this case, the simplest version for the traffic monitoring application does not preserve the user's privacy. We need to modify the application to provide better privacy protection. Therefore, we propose a utility based privacy preservation model for the traffic monitoring application. This modified application considers the tradeoff between the users' desire to protect privacy, and their requirement to have accuracy on traffic estimation error. It also provides a policy to optimize this tradeoff. That