Exploiting Mobile Kinetic Data for Transportation Apps

Dongyao Chen and Kang G. Shin
chendy@umich.edu,kgshin@umich.edu
University of Michigan, Ann Arbor

ABSTRACT

Ubiquitous sensing devices, such as smartphones and wearables, are fast-growing in various applications of the transportation ecosystem. Existing work often focuses on the use of crowdsourced geolocation (e.g., longitude and latitude pair) data and spatio-temporal data mining techniques to derive meaningful/useful insights. Recently, researchers have shown that more safety-critical information can be derived by harvesting and analyzing mobile kinetic (movement-related) signals, such as gyroscope and accelerometer data collected via the driver's smartphones. This approach has the potential for an in-depth traffic pattern analysis which is efficient and scalable, thus enhancing the transportation ecosystem significantly. We first survey the related work in this area, then present a framework for collecting and analyzing mobile kinetic data, and finally discuss related research issues.

1 INTRODUCTION

Ubiquitous devices, such as smartphones and wearables, are fast emerging in many transportation-related applications. For example, ride-sharing services like Didi, Uber, and Lyft use smartphones for flexible coordination of drivers and passengers. Auto insurance companies also use the insuree's smartphone for monitoring his/her driving behavior to adjust his/her premium and deductible, which is proven to be an effective way to incentivize safe driving.

Existing work has been focusing on analysis of geolocation (i.e., longitude and latitude pair) data and spatio-temporal data mining [1, 2] to gain meaningful insights. For example, researchers have shown the feasibility of mining the crowdsourced GPS trajectories of moving vehicles and/or pedestrains. To analyze GPS trajectories, various spatio-temporal mining algorithms ([3]) have been proposed for clustering and deriving insights from trajectories. Representative insights that can be derived from GPS traces include finding the region-of-interest [4], segmenting city districts [5], and the estimated time of arrival (ETA) [6].

In this position paper, we focus on the mobile kinetic (movement-related) sensor (i.e., the gyroscope and accelerometer) data that can be collected from various mobile devices, including smartphones and wearables. Compared to the geolocation data, the mobile kinetic data has two unique advantages. First, the mobile kinetic data is not affected by the signal distortion induced by the changing environment (e.g., urban-canyon effect [7]) — a key limitation for the satellite-based positioning systems. Second, the kinetic data can capture the *dynamics* of moving objects (e.g., vehicles) at a high sampling rate, thus enabling an in-depth analysis of vehicle dynamics (as we will elaborate in Sec. 2). The advantages of the mobile kinetic data are illustrated with an example in Fig. 1. Specifically, Fig. 1 (a) shows the GPS trajectory of a left-turning vehicle. Fig. 1 (b) presents the gyroscope data trace of a smooth left turn, whereas

Published in CIKM 2019 Workshop on AI in Transportation.

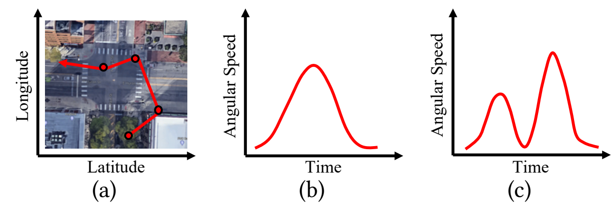

Figure 1: Different data traces of a left turn at an intersection. (a) Geolocation data trace; (b) gyroscope trace of a smooth turn; (c) gyroscope trace of an interrupted turn.

Fig. 1 (c) shows the data traces indicating that the vehicle turn is , for example, by crossing traffic, jaywalkers, etc. This insight can be used to enhance the safety at an intersection. For example, an intersection that shows a high rate of interruptions is likely to be accident-prone and should, therefore, be noticed by both navigation apps (e.g., for path planning) and local transportation departments.

To effectively harvest the mobile kinetic data for various real-world applications, we need to address the following key challenges.
1. Sensor data alignment. Different postures (portrait vs. landscape) of the mobile device generates different sensor readings. So, coordinate alignment is necessary for calibrating the sensor data.
2. Complexity for analyzing the mobile kinetic data. Due to the high sampling rate of the mobile kinetic data, direct application of the legacy spacial-temporal analytics may incur prohibitive costs of time and computational resource.
3. Analysis of crowdsourced mobile kinetic data. The data analytics back-end needs to derive meaningful features from the mobile kinetic data traces.

To address these challenges, we need to investigate both front-end (e.g., smartphone apps for users) and back-end (data analytics engine in the cloud).

1.1 Framework Overview

Fig. 2 depicts a generic, mobile kinetic data collection and analysis system.
Data-collection application. A data-collection app needs data coordinate alignment to address the first challenge (as discussed in Sec. 1). To mitigate the impact of the varying posture, the coordinate alignment [8] transforms a mobile coordinate system to a fixed one, e.g., the Earth coordinate system. The developer can also implement other techniques to enhance the usability of the app. For example, dimension reduction [9] can be used to make the sensor data less bulky for its uploading to the cloud. Note that a light-weight task can also be processed locally.
Data analytics at the cloud. The heavy data analytics tasks on the collected mobile data are done in the cloud to address challenges 2 & 3. First, the data must be pre-processed to filter noisy and less meaningful data snippets. For example, a sharp spike in the gyroscope data is unlikely to be induced by driving maneuvers.

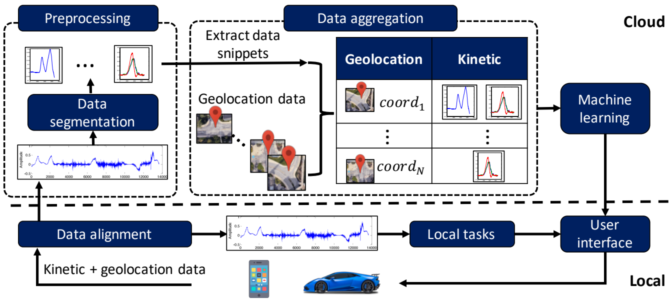

Figure 2: Ubiquitous sensor kinetics analysis for transportation apps that includes local and cloud-based analyses.

Next, kinetic data traces are aggregated based on the region or pattern of interests. For example, if the goal is to find the turning pattern of vehicles at an intersection, one can first narrow down the search space by performing geo-fencing on the collected gyroscope data. Then, one can extract the turning maneuver by recognizing the *bump*-shaped pattern of gyroscope traces. Finally, a machine learning technique can be implemented to process the aggregated data to derive insights based on the requirement of a given task.

2 RELATED WORKS

Compared to the geolocation data, mobile kinetic data collection and analysis for transportation apps is still in its infancy. VSense [8] proposed a vehicle steering maneuver detection algorithm that has linear time-complexity. The low overhead of VSense algorithm is essential for building real-time (or time-critical) apps and cyber-physical systems for assisted driving. Map++ [10] has shown the feasibility of using the pattern of mobile data to detect the occurrence of traffic signs, e.g., stop signs. TurnsMap [9] showed the feasibility of detecting risky intersections by analyzing he crowd-sourced mobile kinetic data. Based on the crowdsourced data from drivers' smartphones, TurnsMap used the sequential pattern of the gyroscope data to detect if an intersection has a left-turn protection — an enforcement that is proven to be critical for enhancing driving safety at intersections.

3 RESEARCH ISSUES

3.1 Adaptive Sensing and Analysis Scheme.

Traffic patterns may change dynamically due to the different time (day) of day (week) and/or environment. For example, the left-turn protection of an intersection may change with the time of day — the protection may be enforced during rush hours but turned off at night time. Moreover, a thus-derived pattern in a certain area may not be usable for other regions due to different driving habits and/or traffic regulations. Therefore, a scalable data collection and analysis framework should be adaptive to different environmental settings.

A natural solution to this problem is to construct a comprehensive data set that can cover diverse traffic patterns. However, this approach can incur a prohibitive cost. One way to mitigate this would be use of transfer learning [11] — i.e., reuse a pre-trained model for other regions. To build and/or validate the new system/algorithm, several simulation platforms, e.g., SUMO [12], Vissim [13], Carla [14], can be utilized.

3.2 Extending the Sensing Modality

As mobile devices are becoming the gateway for the connectivity of vehicles and the transportation ecosystem, other sensing modalities will likely be commonplace. Recent studies, e.g., CarLab [15], have demonstrated flexible automotive data collection with different sensing modalities, e.g., smartphone data and vehicle CAN-bus data. There can be several other issues on utilizing and coordinating various sensing modalities to facilitate smarter transportation.

3.3 Data Analytics Methods for Kinetic Data

Existing spatio-temporal data analytics may not be usable for analyzing the mobile kinetic data. For example, the Euclidean distance is the core of quantifying the spatial distance, and thus is essential for various geolocation data mining tasks, e.g., clustering. However, it is not usable for clustering kinetic data traces as there is no spatial connection between movements. An intuitive approach would be to explore the morphological similarity (e.g., dynamic time warping distance), which can induce a quadratic complexity of computing time. The remaining research question would then be: "how to cluster kinetic data traces efficiently?"

4 CONCLUSION

Mobile kinetic data has unique advantages in benefiting the transportation ecosystem. Compared to geolocation data, kinetic data allows a more in-depth analysis of traffic patterns. We presented a representative system for utilizing the mobile kinetic data and discussed some of the related research issues.

REFERENCES

[1] Jae-Gil Lee, Jiawei Han, and Kyu-Young Whang. Trajectory clustering: A partition-and-group framework. In *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data*, SIGMOD '07, pages 593–604, New York, NY, USA, 2007. ACM.

[2] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, pages 226–231. AAAI Press, 1996.

[3] Jing Yuan, Yu Zheng, Xing Xie, and Guangzhong Sun. Driving with knowledge from the physical world. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pages 316–324, New York, NY, USA, 2011. ACM.

[4] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. Mining interesting locations and travel sequences from gps trajectories. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, pages 791–800, New York, NY, USA, 2009. ACM.

[5] Jing Yuan, Yu Zheng, and Xing Xie. Discovering regions of different functions in a city using human mobility and pois. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, pages 186–194, New York, NY, USA, 2012. ACM.

[6] Google Maps Shows Estimated Time of Arrival. https://money.cnn.com/2017/03/22/technology/google-maps-eta/index.html, 2017.

[7] Hongmian Gong, Cynthia Chen, Evan Bialostozky, and Catherine T Lawson. A gps/gis method for travel mode detection in new york city. *Computers, Environment and Urban Systems*, 36(2):131–139, 2012.

[8] Dongyao Chen, Kyong-Tak Cho, Sihui Han, Zhizhuo Jin, and Kang G. Shin. Invisible sensing of vehicle steering with smartphones. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '15, pages 1–13, New York, NY, USA, 2015. ACM.

[9] Dongyao Chen and Kang G. Shin. Turnsmap: Enhancing driving safety at intersections with mobile crowdsensing and deep learning. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3(3), 2019.

[10] H. Aly, A. Basalamah, and M. Youssef. Map++: A crowd-sensing system for automatic map semantics identification. In *2014 Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pages 546–554, June 2014.

[11] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'14, pages 3320–3328, Cambridge, MA, USA, 2014. MIT Press.

[12] Simulation of Urban MObility. http://sumo.sourceforge.net/.

[13] Vissim simulator for transport planning, traffic engineering and traffic simulation. http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim/.

[14] Carla Simulator for autonomous driving research. http://carla.org/.

[15] Mert D. Pesé, Arun Ganesan, and Kang G. Shin. Carlab: Framework for vehicular data collection and processing. In *Proceedings of the 2Nd ACM International Workshop on Smart, Autonomous, and Connected Vehicular Systems and Services*, CarSys '17, pages 43–48, New York, NY, USA, 2017. ACM.