Exploring Mobility Behaviours of Moving Agents from Trajectory traces in Cloud-Fog-Edge Collaborative Framework

Shreya Ghosh
Department of Computer Science and Engineering
Indian Institute of Technology Kharagpur
Email: shreya.cst@gmail.com

Soumya K Ghosh (Advisor)
Department of Computer Science and Engineering
Indian Institute of Technology Kharagpur
Email: skg@cse.iitkgp.ac.in

***Abstract*—Both analyzing mobility traces and understanding a user's movement semantics from mobile sensor data are challenging issues in ubiquitous computing systems. With the pervasiveness of sensor technologies, wireless networks and GPS-equipped devices, a huge volume of location information is being accumulated. Several techniques have been proposed to analyze the mobility traces and extract informative knowledge for varied location-aware applications. However, all of these applications necessitate an effective mobility-analysis framework to capture the movement behavior of individuals in minimum delay. This paper aims to develop a cloud-based mobility analytics framework to model peoples' mobility behaviour in varied granular scale and extract usable knowledge to provision location-aware services. The preliminary experimental results on real-life dataset depict the effectiveness of our proposed framework.**

***Keywords*-Cloud Computing; GPS Trajectory; Mobility Behaviour; Categorization; Transfer Learning.**

I. Introduction

The advances of location sensing technologies, wireless networks and mobile phones have generated huge volume of mobility traces. Analyzing and understanding the mobility traces are essential for better urban planning and management [1]. Moreover, the in-built sensors (GPS, Accelerometer, Sound etc.) of mobile-phones and other contextual information such as, activity-performed in a point-of-interest, app-usages or calender schedules of an individual enrich the movement history with more informative knowledge [2], [3].The multitude of mobility information fosters varied interesting and challenging problems namely, ride-sharing services (OLA/ UBER), efficient traffic management, personalized recommendation system, defense applications, mobile-user categorization [4], [5], [6] etc. Human activities generally follow temporal regularities and habitual sequences [3]. For example, people generally do not abruptly change their regular routines, such as travelling to work by a specific transportation mode at a particular time. They also do not easily change preferences such as listening to music or reading newspapers while travelling, or activity sequences such as having medicines after breakfast.

Location traces on a space-time continumm of any moving agents namely, people, vehicles are termed as *trajectory*. "Human moves with an intent" [7] and capturing the inherent knowledge behind any movement may help to facilitate several location based services or personalized recommendation systems. Therefore, one of the major objective of the work is to perceive the correlations among location, time and mobility behaviours of people. However, analysing huge volume of mobility data is time and compute-intensive. To handle huge amount of data, cloud, fog, edge are the most significant infrastructures to provide delay and energy efficient solution. This work aims to address following research questions:

1) How to analyse huge amount of mobility traces and extract usable knowledge to provision better urban-transportation planning and smart living? Whether cloud-fog-edge based paradigm provides effective solution? [1], [8], [9], [10]
2) How to model peoples' movement behaviour to facilitate several context-aware services effectively? Whether the behavioral differences in the movement patterns of the individuals can be captured and utilized to cluster/ categorize users? [2], [3], [11]
3) How to extract mobility-association rules (inter-sequence and inter-transactional) from trajectory log to summarize interesting movement behavior? [12], [6]
4) Can we map the mobility knowledge of one known region to another unknown (target) region of similar type (say, academic campus) for traffic, point-of-interest (POI)/land-use predictions etc.? [2]

The mobility analysis reveals the interconnections, correlations and differences among individuals' movement behaviors and their activities by exploring their mobility attributes. However, analyzing huge volume of accumulated trajectory traces is still a challenging task. The major objective of the work is to develop a cloud-based framework to extract informative knowledge from these mobile-trajectories and utilize the information in several applications.

II. Related Works

The earlier works represent trajectory as episodes of *stop* and *move* [13]. However, recent research focus is on semantic analysis of trajectory traces to mitigate the semantic

978-1-7281-6095-5/20/$31.00 ©2020 IEEE
DOI 10.1109/CCGrid49817.2020.00016

gap between GPS log collected from mobile devices and the *intent* of users' movements. Parent *et al.* [14] discuss several semantic enrichment processes such as appending transportation modes, activity performed at different locations. The concept of *symbolic trajectory* is introduced in [15], where the authors propose a novel approach to capture a varied range of semantic descriptions from a geometric trajectory. There are several challenging applications of the semantic trajectory data analysis. Du *et al.* [16] study the personal traveling behaviors and collective mobility patterns to detect pick-pocket suspects. The land-use classification from taxi GPS traces is carried out in [17]. The work in [18] describes the formalization of a semantic-enriched knowledge discovery process for interpreting human movement behaviors. All these depict a broad range of literature on semantic trajectory data mining. The extraction and representation of usable knowledge are major challenges in effective information retrieval.

The unavailability of labelled dataset is a major challenge in supervised learning techniques. To overcome this issue and reduce the expensive data-labelling efforts, *Transfer Learning* [19] seems a way forward. There are several real-life applications of transfer learning namely, classification of webpages, email-spam detection [20] text document classifications [21] etc. The knowledge transfer approaches between multimodal data are presented in [22]. In [23], the transfer learning led to promising results in predicting air-quality of different cities where data are insufficient.

There are several existing works on different aspects of trajectory data mining, however, they are somewhat fragmented. This doctoral symposium paper provides a generic framework of mobility analysis where varied prospects of this domain is explored. Moreover, we have provided brief description of four trajectory data mining applications. We strongly believe that this paper will provide a clear guideline to the readers for designing an end-to-end trajectory data mining framework.

III. OVERVIEW OF *Mobility-Analysis* FRAMEWORK

Figure 1 depicts the overview of the proposed framework. GPS traces of moving agents (people, public and private vehicles etc.) and road-network information (real-time traffic information, road network structure etc.) are the inputs of the framework. In the initial pre-processing steps, after removal of erroneous entries, each GPS point is geo-tagged (closest point-of-interest) and stay-points (where user spent significant time duration) are extracted by trajectory-segmentation. The 'trajectory modelling' module processes the semantically enriched trajectories both from individual and aggregate levels. The clustering approach groups similar movement trips and outputs the signature movement trajectory for each such cluster and outlier trajectories are also extracted. Further, urban mobility dynamics (crowd-flow/ GPS footprint distributions etc.) among varied places may help to understand the travel demand of different functional regions of a city. As mentioned before, analysing voluminous amount of mobility data is a big challenge. Moreover, delay becomes a crucial parameter for several location based service. We develop a hierarchical cloud-fog-edge collaborative framework to provision any location-aware service in minimum delay [1]. The hierarchical architecture of the proposed framework is illustrated in Figure 2. On the topmost layer, the cloud with several homogeneous servers and databases are used for te core mobility analytics. It predicts the optimal path by analyzing present traffic condition. In the second layer, fog devices (Road-side unit or RSU) are present, which communicates with the users and provides service (or information). The services provided by the RSU may be time-critical applications [1] such as finding the nearest hospital and recommending the less congested route to an ambulance. At the bottom layer, the users can request for a particular service through devices like laptop, tab, mobile phones. The proposed approaches of the work and progress [1-12] so far are discussed briefly.

A. Mobility aware Cloud-service development:

Analysing large scale real-time time-series traces (GPS points) and heterogeneous data becomes challenging due to space and time complexity. Apparently distributed platform, like Cloud Computing paradigm seems to be efficient and effective solution but traditional cloud computing is not designed to process and analyse spatio-temporal traces like GPS. To this end, we develop an end-to-end cloud based trajectory-management system [10], [9] and [8] which is capable to i) index and store spatio-temporal traces along with additional information (social media, POI types etc.) ii) efficient spatio-temporal query retrieval iii) effective partitioning of spatial data (say, road-network of a country) and distribution of mining tasks (finding optimal paths, spatio-temporal clustering etc.) iv) computation offloading to reduce service-response time and providing real-time response. The algorithms or modules are developed using MapReduce paradigm to reduce the execution time of the process.

B. Mobility Pattern Modelling and User Categorization:

Movement pattern modelling is a challenging task. A variant of Bayesian network has been deployed to model spatio-temporal and semantic information of the movements. We define the summarized mobility trace as probabilistic directed graph $G = (V, E)$, where each node represents a random variable, consisting the visited place of the individual along with the timestamp and time spent at the point. Given the historical GPS log of users and pre-defined user-categories, user-classification/ categorization is carried out. The mobility features used for the categorization task are: i) *visit in types of places* (mf_1), ii) *Speed of movement*

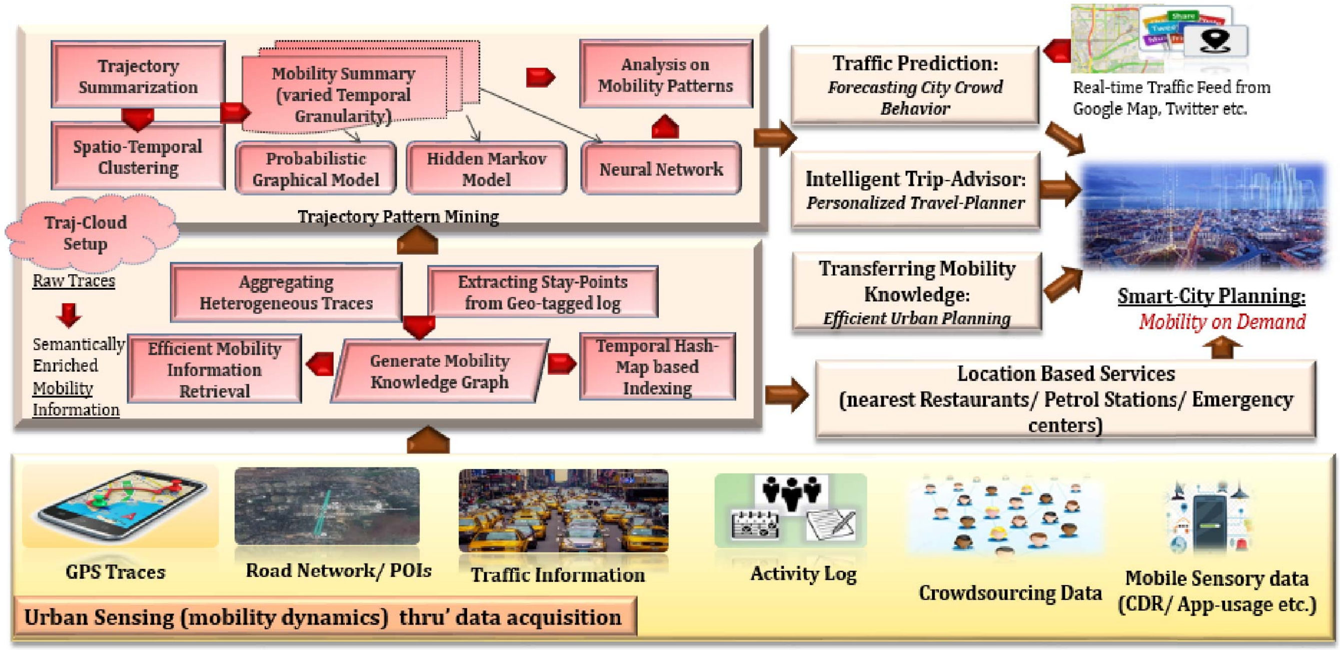

Figure 1: Flow Diagram of the *Mobility Analysis* work

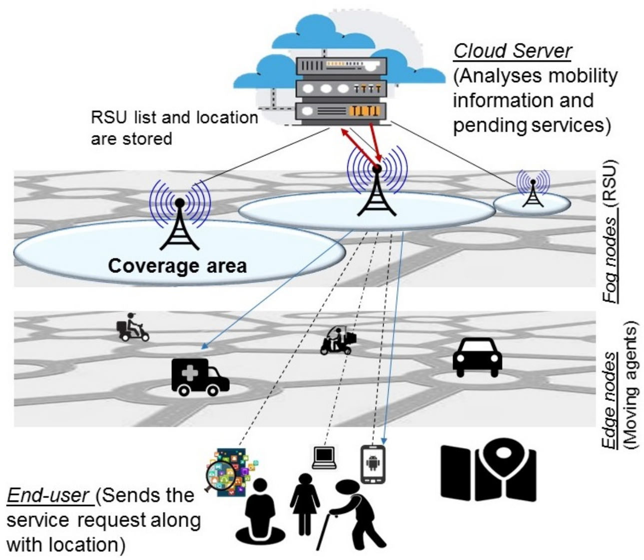

Figure 2: Hierarchical structure of Cloud-Fog-Edge nodes and end-user

or transportation mode (mf_2), iii) *User Movement patterns* (mf_3).

C. Mobility Association Rule Generation:

Mobility association rule mining finds out the rules of the form $A \Rightarrow B$, where $A \in \{T_s, T_d, E_d\}$ and $B \in \{P, T_d, E_d, D_{AB}\}$ and $A \cap B = \phi$ and $Support(A \Rightarrow B) > thresh$ from transactional database of GPS traces. Each GPS transaction consists of user-id, stay-point information (i.e, T_d:duration, T_s:Time-interval, P:place-information), edge-traversal information (i.e, E_d:distance and E_t:travel-duration to reach the stay-point). D_{AB} depicts derived attributes of the transactions in the mobility transactional database. [6]. These mobility rules are crucial to map travel demands in various places. For instance, the mobility events (road blockage/ traffic break down) effect other neighboring regions in a temporal sequence and subsequently it helps to predict travel demand and travel time efficiently. Our proposed framework may explore the rules like (R1): *When the region* A_1 *of a city experiences traffic congestion, then with 80% probability the traffic density of region* A_2 *will be higher after* δ *time-period and consequently high travel-demands in the spatial-neighborhood region of* A_2.

D. Transferring Knowledge-base from source to target region:

The objective is to extract knowledge from the GPS traces of source region (where sufficient labelled data is available) and transfer the knowledge, more specifically labelled data of source region (S) to target region (T) (where partial or no labelled data available) to perform classification task [4]. For example, each place exhibits some mobility features, such as spatio-temporal regularity in the density of GPS footprints. For instance, commercial places, like movie-hall or shopping mall are likely to have more GPS footprints in weekends. Road segments connecting to residential place and industrial area are likely to be congested in particular time-intervals. The places like, airports or railway stations have more or less similar GPS density throughout the day. If we can learn the interrelations, then With an appropriate knowledge transfer framework, it would be possible to identify the POIs (point-of-interests say, lecture hall/ cafeteria/ movie-complex or residential buildings) of the region T, even with no labelled GPS trace of individual people. The intuition is that the mobility dynamics of a given city can be characterized by the movement patterns of the people and may help in planning other newly-built cities

IV. Experimental Observations

The experimentation has been carried out using two real-life GPS datasets. Mobility traces of 65 students' are collected voluntarily from GPS-enabled smartphone devices/Google map timelines over two semester term (spring 2016 and autumn 2016). The participants installed an application (AndroSensor) and uploaded the GPS traces after each week in a web-portal which populated the back-end database. Another dataset of Dartmouth Campus, USA [24] consists of GPS traces of students and professors, employees. Table 1 presents the confusion matrix of user categorization task in two different regions using *precision* and *recall* measures. It may be noted that the user categorization of Dartmouth region has been carried out by transferring the mobility knowledge from Kharagpur region. The detailed process is described in [4]. The mobility association rule mining technique extracted several interesting mobility rules. Few rules are depicted as follows [S:Support, C:Confidence, J: Jmeasure (J), M: Mutual Information]:

- R1: User (High, HealthCare Center) ⇒ User (Low, Hospital, MedicineShop), S = 0.51, C = 0.87, J = 0.83, M = 0.87
- R2: User (High, Cafe, Restaurant, Transit) ⇒ User (Low, AcademicBuilding, ClassRoomComplex, Department), S = 0.78, C = 0.88, J = 0.78, M = 0.84.
- R3: TimeStamp (1, 3) ⇒ FootprintFrequency (High, AcademicComplex, Transit,Department), S = 0.51, C = 0.75, J = 0.67, M = 0.72

The rules depict the movement behaviours of different user-categories as well as footprint frequencies in different point-of-interests. On the other side, user-specific rules (R1 or R2) help to infer the interests of users and subsequently provides a better recommendation system. Figure 3 reports the execution time of mobility-based queries compared to a standalone system. In our proposed framework, we have used MapReduce paradigm to divide the computationally heavy task into different smaller parts and reduce the execution time. The standalone system refers to a single server where the task is computed sequentially. It is observed that the proposed cloud-fog-edge based system achieves better execution time.

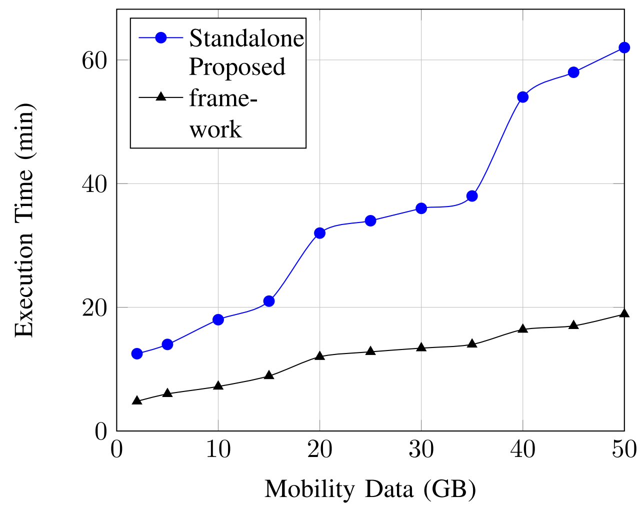

Figure 3: Computational Efficiency of proposed framework

	Kharagpur		Dartmouth	
User Category	Precision	Recall	Precision	Recall
Undergraduate Student	0.75	0.75	0.78	0.82
Graduate Student	0.846	0.95	0.9	0.82
Graduate Student (Research Student)	0.6315	0.545	NA	NA
Employees/Professors	0.555	0.625	0.62	0.56
Non-Residential Students	0.714	0.714	NA	NA

Table I: Precision/Recall values of experiment on Kharagpur and Dartmouth region [4]

V. Concluding Remarks

Mobile sensing data (GPS traces of moving agents, CDRs etc.) is known to provide significant insights about mobility dynamics as well as a strong and connected knowledge base for urban infrastructure. However, exploring and analyzing these huge volume of data is still a big challenge. The present work focuses on modelling human mobility behaviour both from temporal and spatio-temporal contexts and extracting mobility association rules from huge volume of accumulated GPS traces. Experimentation and initial results on real-life dataset of temporal activity log and GPS traces of individuals of an academic campus demonstrate the potential of the proposed modules. It has been shown that the proposed framework is capable to extract mobility association rules and resolve mobility-based queries effectively. The proposed approaches will act as a foundation of connected knowledge base between locations, people and trigger location-based services relevant to a given context.

Acknowledgement

The work is partially supported by TCS PhD Research Fellowship awarded to the student-author Shreya Ghosh. The parts of this doctoral symposium paper have been published in [1-12].

References

[1] S. Ghosh, A. Mukherjee, S. K. Ghosh, and R. Buyya, "Mobi-iost: mobility-aware cloud-fog-edge-iot collaborative framework for time-critical applications," *IEEE Transactions on Network Science and Engineering*, 2019.

[2] S. Ghosh and S. K. Ghosh, "Modeling individual's movement patterns to infer next location from sparse trajectory traces," in *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 693–698, IEEE, 2018.

[3] S. Ghosh, S. K. Ghosh, R. D. Das, and S. Winter, "Activity-based mobility profiling: A purely temporal modeling approach," in *Proceedings of The Web Conference 2018*, pp. 409–416, 2018.

[4] S. Ghosh and S. K. Ghosh, "Modeling of human movement behavioral knowledge from gps traces for categorizing mobile users," in *Proceedings of 26th International Conference on World Wide Web*, pp. 51–58, 2017.

[5] S. Sinha, M. K. Nirala, S. Ghosh, and S. K. Ghosh, "Hybrid path planner for efficient navigation in urban road networks through analysis of trajectory traces," in *Proceedings of 24th International Conference on Pattern Recognition (ICPR)*, pp. 3250–3255, IEEE, 2018.

[6] S. Ghosh and S. K. Ghosh, "Exploring human movement behaviour based on mobility association rule mining of trajectory traces," in *Proceedings of 17th International Conference on Intelligent Systems Design and Applications*, pp. 451–463, Springer, 2017.

[7] S. Ghosh and S. K. Ghosh, "Thump: Semantic analysis on trajectory traces to explore human movement pattern," in *Proceedings of the 25th International Conference on World Wide Web*, pp. 35–36, International World Wide Web Conferences Steering Committee, 2016.

[8] S. Ghosh, J. Das, and S. K. Ghosh, "Locator: A cloud-fog-enabled framework for facilitating efficient location based services," in *Proceedings of 12th International Conference on COMmunication Systems & NETworkS (COMSNETS)*, pp. 87–92, IEEE, 2020.

[9] S. Ghosh, S. K. Ghosh, and R. Buyya, "Movcloud: A cloud-enabled framework to analyse movement behaviors," in *Proceedings of 11th IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*, pp. 239–246, IEEE, 2019.

[10] S. Ghosh and S. K. Ghosh, "Traj-cloud: a trajectory cloud for enabling efficient mobility services," in *Proceedings of 11th International Conference on Communication Systems & Networks (COMSNETS)*, pp. 765–770, IEEE, 2019.

[11] S. Ghosh, A. Chowdhury, and S. K. Ghosh, "A machine learning approach to find the optimal routes through analysis of gps traces of mobile city traffic," in *Recent Findings in Intelligent Computing Techniques*, pp. 59–67, Springer, 2018.

[12] S. Ghosh and S. K. Ghosh, "Exploring the association between mobility behaviours and academic performances of students: a context-aware traj-graph (ctg) analysis," *Progress in Artificial Intelligence*, vol. 7, no. 4, pp. 307–326, 2018.

[13] S. Spaccapietra, C. Parent, M. L. Damiani, J. A. de Macedo, F. Porto, and C. Vangenot, "A conceptual view on trajectories," *Data & knowledge engineering*, vol. 65, no. 1, pp. 126–146, 2008.

[14] C. Parent, S. Spaccapietra, C. Renso, G. Andrienko, N. Andrienko, V. Bogorny, M. L. Damiani, A. Gkoulalas-Divanis, J. Macedo, N. Pelekis, *et al.*, "Semantic trajectories modeling and analysis," *ACM Computing Surveys*, vol. 45, no. 4, p. 42, 2013.

[15] R. H. Güting, F. Valdés, and M. L. Damiani, "Symbolic trajectories," *ACM Transactions on Spatial Algorithms and Systems*, vol. 1, no. 2, p. 7, 2015.

[16] B. Du, C. Liu, W. Zhou, Z. Hou, and H. Xiong, "Catch me if you can: Detecting pickpocket suspects from large-scale transit records," in *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 87–96, ACM, 2016.

[17] G. Pan, G. Qi, Z. Wu, D. Zhang, and S. Li, "Land-use classification using taxi gps traces," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 113–123, 2013.

[18] C. Renso, M. Baglioni, J. A. F. de Macedo, R. Trasarti, and M. Wachowicz, "How you move reveals who you are: understanding human behavior by analyzing trajectory data," *Knowledge and Information Systems*, vol. 37, no. 2, pp. 331–362, 2013.

[19] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[20] J. Zhang and J. Huan, "Inductive multi-task learning with multiple view data," in *18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 543–551, 2012.

[21] W. Dai, G.-R. Xue, Q. Yang, and Y. Yu, "Transferring naive bayes classifiers for text classification," in *AAAI*, vol. 7, pp. 540–545, 2007.

[22] L. Nie, L. Zhang, Y. Yang, M. Wang, R. Hong, and T.-S. Chua, "Beyond doctors: Future health prediction from multimedia and multimodal observations," in *23rd ACM International Conference on Multimedia*, pp. 591–600, 2015.

[23] Y. Wei, Y. Zheng, and Q. Yang, "Transfer knowledge between cities," in *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1905–1914, 2016.

[24] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proceedings of the ACM international joint conference on pervasive and ubiquitous computing*, pp. 3–14, ACM, 2014.