

Research Article

Leveraging Participatory Extraction to Mobility Sensing for Individual Discovery in Crowded Environments

Lin Wang,^{1,2} Jing Yang,¹ and Wenyan Liu^{1,2}

¹ School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China

² The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province, Qinhuangdao 066004, China

Correspondence should be addressed to Wenyan Liu; wylu@vip.163.com

Received 5 July 2013; Accepted 4 September 2013

Academic Editor: Yuan He

Copyright © 2013 Lin Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Neighbor discovery for moving individual is considered an important technology submitting to location-based service (LBS), which includes such things as recruitment flow of information, logical localization, and health monitoring. Based on the tradeoff between universality and accuracy of neighbor discovery, we propose the environmental characteristics participatory extraction method benefiting to mobile individual discovery. First, we fuse lightweight accelerometer, light sensors, and microphone collaboratively. Furthermore, support vector machine (SVM), Tanimoto coefficient, and Manhattan distance are used to calculate three kinds of fingerprint similarity, respectively, and then the principal component analysis based method reduces data dimension in order to obtain neighbor similarity rank. Finally, the experiment data are collected by 25 volunteers, and trace-driven simulations show that Euclidean distance error is below 4.69 and the convergence time is within 0.75 s.

1. Introduction

The proliferation of mobile devices and wireless has fostered the demand of context aware application, in which localization and emergency navigation are emerging applications with significant research and social values [1–3], at the same time, vehicular ad hoc networks (VANETs) have recently reemerged as one of promising research areas for safety and connectivity in road networks [4]. Urban sensing, participatory sensing, and moving individual clustering can provide rich contextual information for mobile applications such as social networking and location-based services (LBS) [5]. Moving individual neighbor discovery has thus become one of the key factors in mobile service quality. Among existing LBS technologies, combining radio frequency (RF) fingerprinting, based on Wi-Fi or cellular signals, with the inertial sensors (e.g., accelerometer, compass, and gyroscope) featured in smartphones has received great attention [1, 4, 6–9]. Sensors embedded in smartphones can detect and gather information regarding location, acceleration, orientation, rotation angle, and so forth. Some applications plan to share

this information in online social networks by recording users' behavior characteristics and the dynamics of everyday activities (such as commuting) and enabling search services associated with the users' physical environment. Mobile services closed to moving individual neighbor discovery include mobility-based recruitment, logical localization, and environment detection.

Based on some shortcomings and insufficiency of existing methods, we propose the environmental characteristics participatory extraction method for moving individual neighbor discovery by means of a smartphone based lightweight sensors. Smartphone has 3-axis accelerometer, light sensor, and microphone commonly, keeping open, which are participated by users passively. At the same time, users are willing to provide fingerprint characteristics on the premise of not revealing their privacy. We fuse lightweight accelerometer, light sensors and microphone collaboratively. We use support vector machine (SVM), Tanimoto coefficient, and Manhattan distance to calculate three kinds of fingerprint similarity, respectively, and then the principal component analysis is used to reduce data dimension in order to obtain neighbor

similarity rank, and the experiment data is collected by 25 volunteers to test the moving individual neighbor discovery method.

As a basic component which might be frequently invoked by LBS on energy-constrained mobile phones, it needs to meet several stringent design requirements.

High Accuracy. As a generic framework that many other applications would potentially rely on, it should accurately detect the neighbor similarity.

Real-Time. Mobile individual should receive neighbor similarity response promptly. An outdated detection result may be less valuable for many instantaneous applications.

Universality. It should avoid relying on an a priori knowledge or site survey, special sensors, or explicit user feedback to ensure wide applicability.

In this paper, we address these requirements by means of leveraging participatory extractive to mobility sensing. Our contributions can be summarized as follows. First, the proposed method uses the commercial inertial sensors on the smartphone to extractive ambient information without any label. Second, we design an iterative neighbor similarity algorithm which considers behavior similarity, spatial-temporal similarity, and background similarity. Third, we introduce the principal component analysis to excavate the critical factor evaluating the neighbor similarity rank. Finally, we recruit 25 volunteers to construct a crowded scenario, and trace-driven simulations demonstrate that Euclidean distance error is below 4.69 and the convergence time is within 0.75 s.

The remainder of the paper is organized as follows. The related work is introduced in Section 2. The problem description of moving individual neighbor discovery is discussed in Section 3. Section 4 details the system design process and fingerprint similarity measure and integration approach. The evaluation and performance results are described in Section 5. Some problems are discussed in Section 6. Section 7 concludes the paper.

2. Related Work

Improving accuracy and latency of neighbor discovery is a long-standing problem in ad hoc network (e.g., sensor network, social network, and delay-tolerant network) and has been approached from different directions. Techniques range from protocol optimizations for data collection in sensor network [10] or Wi-Fi access point scanning in wireless network [11] to techniques that leverage multiple built-in sensors on commercial smartphone [12–14] to classify the people activities.

Since most smartphones come with both Bluetooth and Wi-Fi and most current proximity-based applications need proximity detection to the cheaper Bluetooth radio, Searchlight [10] proposes an asynchronous discovery protocol based on periodic awake slots of probing. Since Searchlight employs a random component to shift each awake time, it will run hard when the service starts. Zifi [11] utilizes ZigBee radios to identify the existence of Wi-Fi networks

through unique interference signatures generated by Wi-Fi beacons. However, ZigBee infrastructure is an extra input with nonignorable overhead.

Extensive work have taken advantage of smartphones built-in GPS, GSM, and inertial sensors to sense contextual information, detecting people's daily behaviors, and achieving logical localization that provides real-time services for people living and working. TagSense [15] have introduced a smartphone-based approach to automatic image tagging system, IODetector [16] uses light sensors, magnetic sensors, and cellular network sensed signal to detect indoor/outdoor environment. However, all these belong to static recognition, without considering that we could make use of environmental characteristics participatory extraction to track moving individual and find its neighbors.

Collecting and mapping the Wi-Fi access points or GSM base stations named war-driving have been deployed in prior work. War-driving requires cost massive money and time. Moreover, war-driving data easily become unreliable and outdated over time.

3. Problem Description

The tracking of individual movement plays an important role in human's health and daily life, especially the terrorists, people living with virus, and real-time tracking of bus in intelligent transportation. An example scenario is shown in Figure 1; the super user as shown in the pentagram is called target and moves forward along the purple route on the important traffic road, and the target meets with the standing, the walking, the running, the cycling, or the motoring at different times via Bluetooth. As a fastest growing LBS service, he wants to invite some people to participate the mobile gaming for waiting vehicle or digging same flavor person nearby for appropriate food.

Bluetooth does not need to anchor and fingerprint database comparing with global positioning system (GPS), GSM. Compared with Wi-Fi, using Bluetooth has several advantages. First, although lots of mobile devices are equipped with Wi-Fi modules nowadays, its popularity is still far beneath that of Bluetooth. Currently, there are total 3 billion Bluetooth devices in the market. Nearly 80% mobile phones and close to 100% of the smart phones are equipped with Bluetooth devices. Furthermore, Wi-Fi devices consume much more energy than that of Bluetooth. The energy consumed by Wi-Fi devices in idle status is 76 times more than that of Bluetooth in the same status [17]. Moreover, even the energy consumed by Wi-Fi devices in idle status is 5 times more than that of Bluetooth devices in scanning status which is the most energy-consuming part for Bluetooth devices. Bluetooth does not need to pay fees; the device name changes infrequently as well as finding surrounding equipment is default and does not need to establish and keep connection [18]. For the devices carried by mobile users, which is energy sensitive, energy-efficient Bluetooth is obviously a better option. Figure 2 shows the energy consumption, respectively, corresponding to different statuses in Table 1. Behavior, light, and background audio fingerprint are collected by means of

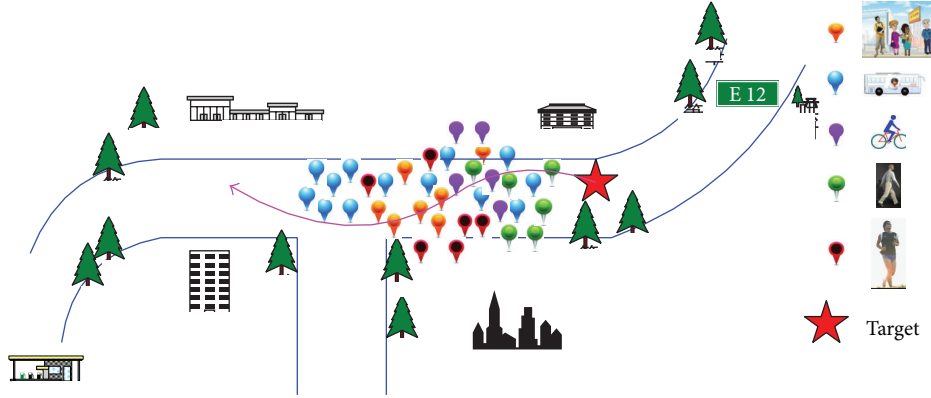


FIGURE 1: Example scenario.

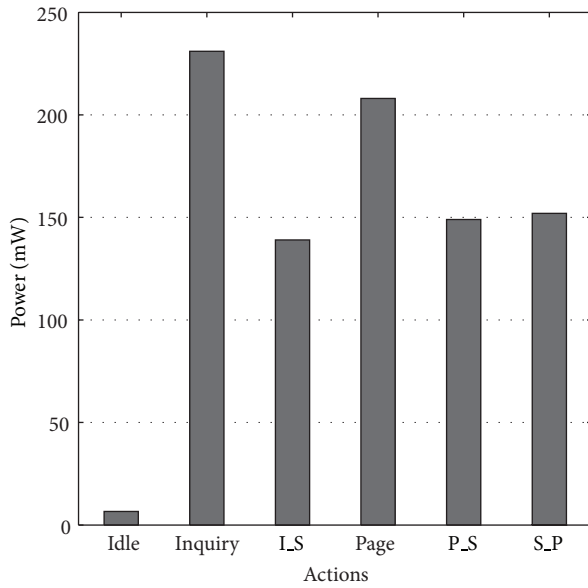


FIGURE 2: Bluetooth power versus status.

TABLE 1: Bluetooth time consumption versus status.

Actions	Duration	Interval
Idle	—	—
Inquiry	10.24 s	—
Inquiry_Scan	11.25 ms	2.56 s
Page	2.56 s	—
Page_Scan	11.25 ms	1.28 s/2.56 s
Send Association Packet	0.625 μ s	—

a smartphone based lightweight sensors, which are upload to a central server at the same time; then the server merges and calculates similarity with the target ones, so as to realize neighbor discovery of moving individual.

Before we present the design of moving individual neighbor discovery in detail, we formally define the relevant conception studied in this paper. There are the stationary, walking, cycling, running, and motoring, which are indicated

by clusters $C(1)$ – $C(5)$ in set U , and individuals $s(i)$ and $u(i)$ move randomly, when the target $s(i)$ is judged to certain cluster, as shown in Figure 3(a), the target $s(1)$ belongs to $C(1)$ and moves among the general ones $u(i)$, contacting with general users through Bluetooth during moving. Symbol and meaning are as shown in Table 2, and Algorithm 1 introduces the specific process.

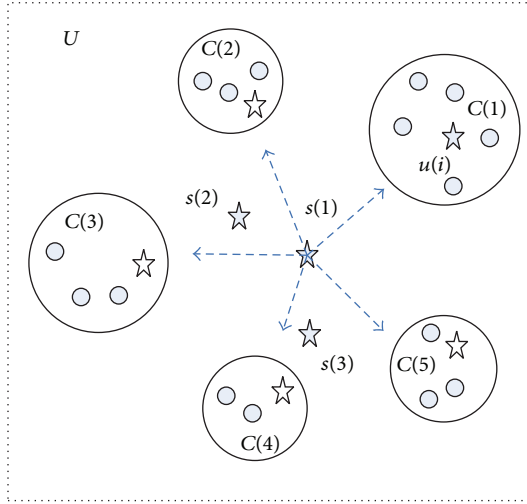
Definition 1. Neighbor ambiguity is shown in Figure 3(b) in $C(1)$, when there is only one person in the network. Then it is the neighbor; for example, there is only one neighbor, respectively, in T_a and T_c , so it does not include ambiguity neighbors. But if there is more than one person in the network, during the period T_b there are five neighbors, we could measure similarity and adopt principal component analysis method to find the optimal neighbor, suboptimal ones, and so on.

Definition 2. The person who gets the highest score in principal component analysis method is called optimal neighbor.

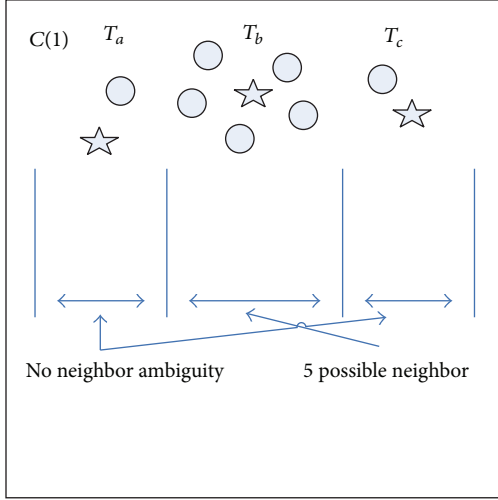
4. Moving Individual Neighbor Discovery

In this section, we first introduce the system architecture and the design details for each component. Then we specify how to aggregate the outputs obtained from each component and generate a final decision, which is provided to upper layer applications through a service interface.

4.1. Multidimensional Sensing Data Extraction. Figure 4 illustrates the system architecture of moving individual neighbor discovery. To meet stringent design requirements, it utilizes a series of lightweight sensors for the neighbor discovery. It primarily makes use of three types of lightweight detectors: behavior detector, light detector, and audio detector. Behavior detector uses accelerometer to collect behavior data. To meet real-time, we adopt SVM to detect behavior, based on the tradeoff between precision and convergence speed, this paper experiments with the appointed and random training set, the search, and empirical parameters. Light detector adopts light sensors to sense ambient light signals. It also utilizes



(a)



(b)

FIGURE 3: Scene. (a) Model and (b) neighbor ambiguity.

other two lightweight sensors, the proximity sensor and the system time clock, to assist the detection. Audio detector captures ambient audio signals to determine the surrounding environment type. In the rest of this section, we will describe the design details of each component.

4.2. Moving Environment Characteristics Determination

4.2.1. Moving Individual Behavior. Selection of sensor which senses individual behavior has a great impact on energy consumption and recognition accuracy of smartphones in behavior detector. Bluetooth sensors are not ubiquitous in outdoor settings. Static Bluetooth beacons mainly exist in indoor settings. Relying on Bluetooth signals exhibited by

TABLE 2: Symbol and meaning.

Symbol	Meaning
U	User set
$C(i)$	Clusters of different behaviors
$s(i)$	Super user
$u(i)$	General user
Type	User type 0-super user; 1-general user
a, l, s	User attributes: behavior, light, and audio
$Similar(1)$	Behavior similarity 0-1
$Similar(2)$	Light similarity 0-1
$Similar(3)$	Audio similarity 0-1
Score	Similarity score
Rank	Rank of similarity score

Input: User set U

Output: score, rank

- (1) Judge the number of super user $s(i)$ according to type of user in set U ;
- (2) Determine the super user belongs to which kinds of movement behavior $C(i)$;
- (3) According to the step (4)–(6) cluster the super user one by one;
- (4) To extract behavior, light audio fingerprint of general users;
- (5) Calculate behavior, light and audio similarity;
- (6) Data dimension reduction by means of principal component analysis method, get integrated similarity score and rank;
- (7) Repeat step (4)–(6) until neighbors of all the super user are sorted.

ALGORITHM 1: Moving individual neighbor discovery.

mobile phones has its own problems. For instance, it is difficult to distinguish whether an individual is moving or if the environment around them is changing (other people carrying devices are moving). Hence, Bluetooth does not provide the accuracy needed for the modes being distinguished [19]. In terms of GSM and Wi-Fi, these features act as a proxy for the speed of an individual, and when speed is similar, as is the case with slow moving traffic, biking, and walking, these features are not as discriminative, and they depend on the density of network end points (cell towers and access points). Accurate human behavior recognition, supporting both indoor (underground) and outdoor mobility [20], is best provided by a combination of GPS (for computing velocity), Wi-Fi localization (for detecting location changes where GPS is unavailable), and accelerometer (for detecting movements directly). However, constantly sampling these sensors will quickly drain the batteries of even the latest generation of smartphones. Thus, for smartphone applications, power efficiency is of great concern, so behavior recognition must be designed to expend as little energy as possible. At the same time, we need to propose a solution that does not depend on the hardware facilities to improve scalability. This paper

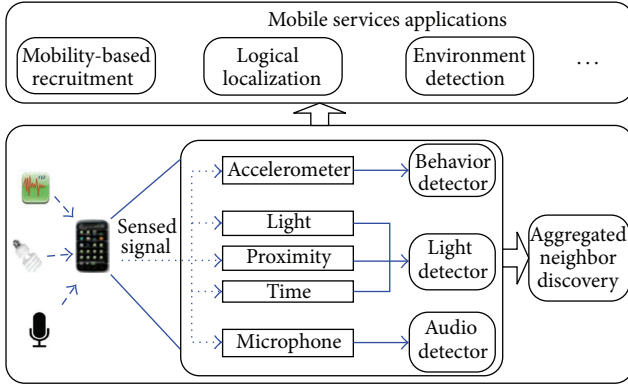


FIGURE 4: Moving individual neighbor discovery system architecture.

proposes that we use low-power accelerometer to collect data and detect moving individual behavior.

This paper divides outdoor user behaviors into five categories: stationary, walking, running, cycling, and motor-ing, which is a multipattern recognition problem actually. SVM and neural networks can be used for pattern recognition. Compared with neural networks, SVM solves the high-dimensional data model construction problems more effectively under the limited sample condition, and it has some advantages including generalization ability, converging to global optimum, and dimension insensitivity. Compared with traditional method, SVM has obvious advantages, and the idea of optimal separating hyperplane can improve the classification accuracy, which is a convex quadratic programming problem, avoiding local optima and introducing kernel function theory, and overcoming the “curse of dimensionality” problem. So we choose SVM to recognize human behavior.

The basic idea of SVM is to construct an optimal classification function for two types of problems, to make them separate without error, keeping maximum interval between them. We use Lagrange to solve the quadratic programming problem. Only a few Lagrange multipliers are not equal to zero in the solution; the corresponding sample is support vector. Finally we achieve classification of entire acceleration data set by training support vectors.

For linearly separable classification problems, the decision function is as follows:

$$d(x) = \text{sgn}(\langle w, x \rangle + b)$$

$$= \text{sgn}\left(\sum_{i=1}^l a_i y_i \langle x, x_i \rangle + b\right), \quad (1)$$

where l is the number of sample, $x = [x_1, x_2, \dots, x_n]^T$ is input feature vector, accelerometer $|a\{x, y, z\}|$, and L_2 -norm, $w = [w_1, w_2, \dots, w_n]^T$ is weight vector, value of $|a\{x, y, z\}|$ and L_2 -norm $\|A\|_2 = \sqrt{(A_x)^2 + (A_y)^2 + (A_z)^2}$, b is offset, $y = \text{sgn}(x)$ is sign function, $d(x) = 0$ is decision hyperplane, when $d(x) \geq 0$, x is adjudged as positive categories otherwise is negative ones. The separated interval is $2/\|w\|$, and to achieve

the maximization of classification interval, it requires $\|w\|$ to be minimal; namely, $|\langle w, w \rangle|$ is minimal. Using Lagrange to solve this optimization problem, the weight vector $w^* = \sum_{i=1}^l y_i a_i^* x_i$ can achieve the maximum hyperplane of classification interval, where $a = (a_1, a_2, \dots, a_n)^T$, and $a_i \geq 0$ is the Lagrange multiplier. Finally, we obtain b by formula (1).

For linearly inseparable classification problems, we adopt Kernel function theory which can covert low-dimensional linearly inseparable into high-dimensional space linearly separable. We use inner product kernel function $K(x_1, x_2)$, replaces inner product $\langle x_1, x_2 \rangle$, and then use the same theory with linearly separable problem to classify. This paper selects radial basis kernel function (RBF):

$$K(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{\sigma^2}\right), \quad (2)$$

where σ is referred to as the expansion constant of radial basis function. In terms of five categories, first, we send the eigenvector of training set into SVM, and then we use the trained support vector machine to classify test set.

In this experiment, the precision is seen as behavior similarity between general and super user *Similar(1)*. The selection of training set and parameters mainly depend on the experience and experimentation, and we need to find a unified guidance method in order to shorten training time and achieve higher accuracy. This paper experiments with the appointed and random training set, the search, and empirical parameters as described in Section 5.

4.2.2. Space-Time Characteristics. Light detector adopts light sensors to capture ambient light signals. It also utilizes other two lightweight sensors, the proximity sensor and the system time clock, to assist the detection. Common similarity measure methods include Tanimoto coefficient [9] and cosine similarity. Tanimoto coefficient is also called generalized Jaccard coefficient, is an extension of cosine similarity, at the same time, and is widely used to calculate document data similarity. It requires less shared items than cosine similarity. So we use Tanimoto coefficient to measure the similarity *Similar(2)* as show in (3), here the light intensity is converted to vector space. Besides, super user could detect many general ones, calculating similarity, respectively:

$$T(x, y) = \frac{x \cdot y}{\|x\|^2 + \|y\|^2 - x \cdot y}$$

$$= \frac{\sum x_i y_i}{\sqrt{\sum x_i^2} + \sqrt{\sum y_i^2} - \sum x_i y_i}, \quad (3)$$

where y_1 is light fingerprint scanned by super user, y_2 is light fingerprint scanned by general ones, and $T(y_1, y_2)$ is greater, which shows that the degree of similarity is greater between two users. As shown in Figure 5, it is light intensity of super and general ones.

4.2.3. Background Sound. We have detected behavior and light similarity, while the ambient sound in a place can be

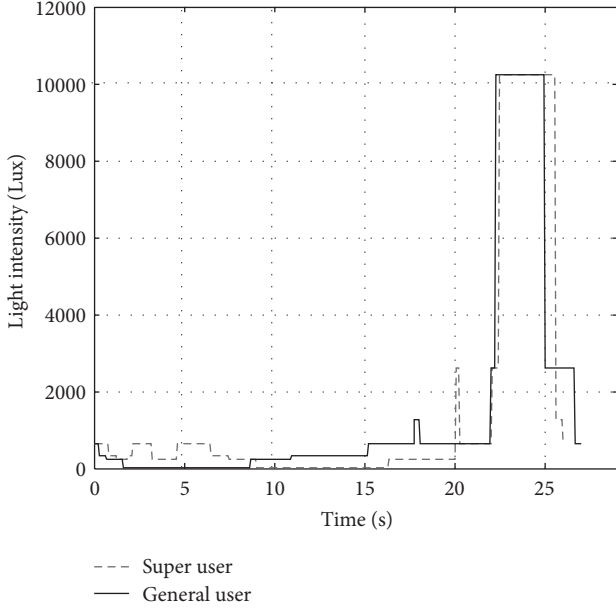


FIGURE 5: Light fingerprint.

suggestive of the type of place [14]. Some places play loud music in the background versus others that are quieter. We recorded the ambient sound of 3 types of place using the phone microphone. Sample frequency is set to 20 Hz. First, we try to distinguish the type of environment by comparing probability density function or box plot of audio signal, as shown in Figures 6 and 7; however, both speech and noise have overlapping part (as shown in dotted box). Then we have observed that signal energy under different sample windows presents a clear step, as shown in Figure 8, so we use Manhattan distance of audio signal energy to distinguish different environments. Note that the energy E of a time domain signal $x(n)$ is defined [13] by

$$E = \sum_n |x(n)|^2. \quad (4)$$

Manhattan distance is mostly used for space distance measure of multidimensional data. Manhattan distance between two n -dimensional vectors $x(x_1, x_2, \dots, x_n)$ and $y(y_1, y_2, \dots, y_n)$ is shown in (5). When Manhattan distance represents similarity, generally we use the following formula to conversion, whose range is between $[0, 1]$; the smaller the distance, the greater the similarity. It has less calculation than Euclidean distance, and performance is relatively higher. We calculate the average similarity $Similar(3)$ of signal energy under different sample windows:

$$d(x, y) = \sum_{k=1}^n |x_k - y_k|, \quad (5)$$

$$similar(x, y) = \frac{1}{d(x, y)}.$$

4.3. Moving Individual Fingerprint Similarity. We have obtained fingerprint similarity between target and neighbors

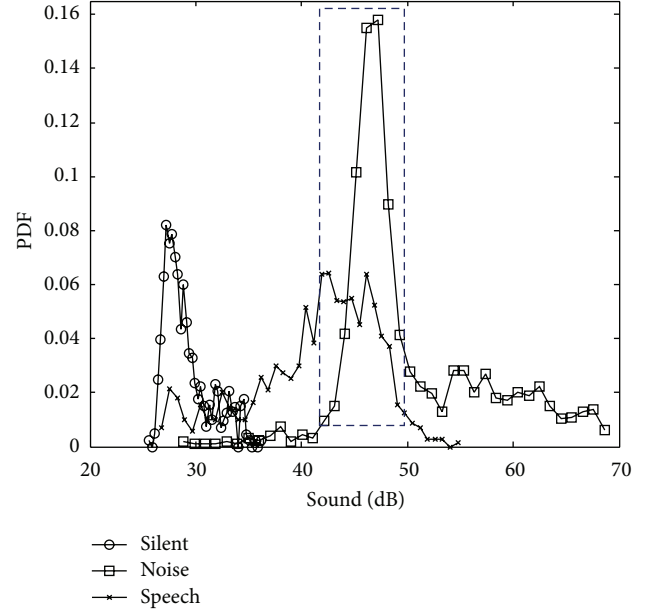


FIGURE 6: Probability density function of audio signal.

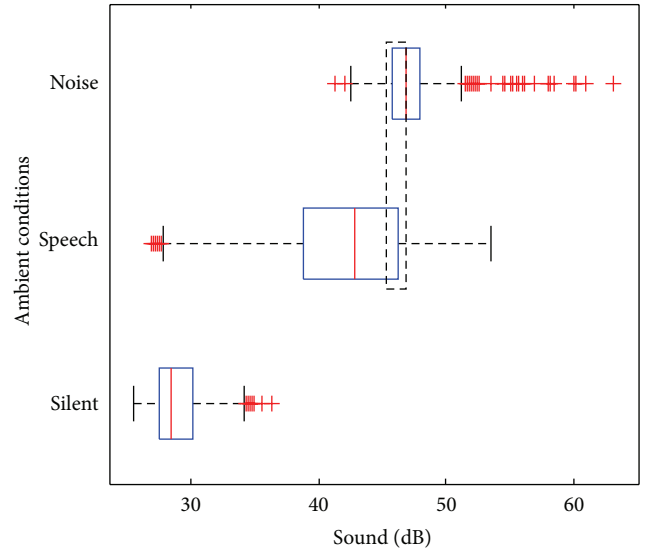


FIGURE 7: Box plot of audio signal.

through the above three kinds of detector. Further we need to put into three kinds of fingerprint similarity to a new integrated similarity variable. Usually, samples are mapped from input space to a low-dimensional space by linear or nonlinear transformation in order to gain a compact low-dimensional representation of original data set. According to the problem characters, this paper adopts principal component analysis method for data dimension reduction.

The mathematical model of principal component analysis is as follows:

$$F = AX, \quad (6)$$

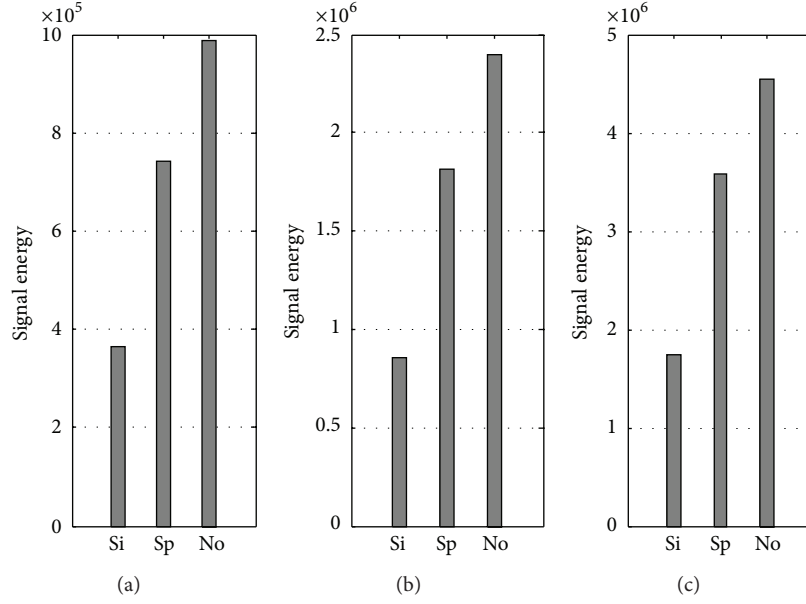


FIGURE 8: Audio signal energy versus sample window (Si: silent; Sp: speech; No: noise).

where

$$F = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_p \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}, \quad (7)$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix},$$

where X is fingerprint similarity, A is coefficient matrix of principal component, and F is principal component.

The proposed model needs to satisfy the following conditions:

- (1) F_i, F_j is unrelated ($i \neq j, i, j = 1, 2, \dots, p$);
- (2) the variance of F_1 is greater than F_2 , and is greater than F_3 and so on;
- (3) $a_{k1}^2 + a_{k2}^2 + \cdots + a_{kp}^2 = 1, k = 1, 2, \dots, p$.

Through data dimension reduction, 3-dimensional fingerprint similarity X is represented by the principal component F_j and score, but not one-to-one correspondence. The principal component of the proposed model contains information from more to less, and final score represents rank of similarity.

Figure 9(a) introduces fingerprint similarity distribution box plot. Pareto chart in Figure 9(b) shows variance contribution rates of principal component, with a rectangular coordinate system. The left ordinate indicates frequency number, the right indicates frequency, the analysis line represents cumulative frequency, and horizontal axis indicates influence factors

of similarity, according to the influence degree (occurrence frequency) arranged from left to right. Principal component comprehensive score and rank are shown in Figure 10.

5. Evaluation

We implement and evaluate a prototype system on the Android platform with three different types of mobile phones (Samsung Galaxy S2 i9100, HUAWEI T8828, and HTC Sensation G14) to collect behavior, light, and audio signal. Instructions were given as to the duration of each activity, and 25 volunteers were advised to perform each behavior, including stationary, walking, running, cycling, and motor-ing. The volunteers concentrated on one transportation mode at a time and performed all five consecutively during their data collection session. The total amount of data collected across all 25 individuals was 125 hours, comprised of 1 hour of data per behavior. The results of the assessment focus on three factors: (1) to do in-depth analysis for training set and parameters selection of behavior detection; (2) to compare similarity sorting error for different numbers of general user and analyze precision, universality, and scalability of neighbor discovery; (3) the running cost of moving individuals neighbor discovery.

5.1. Training Set and Parameters Selection. Training set and parameters selection of behavior detection have an important effect on precision and convergence speed, and this paper adopts combination of the appointed, random training set and empirical, search parameter. Then we calculate average precision and convergence time of 20 groups experiments, as shown in Figure 11. Although the method of parameters optimization can improve the precision, the convergence time is exponential growing, based on the balance between

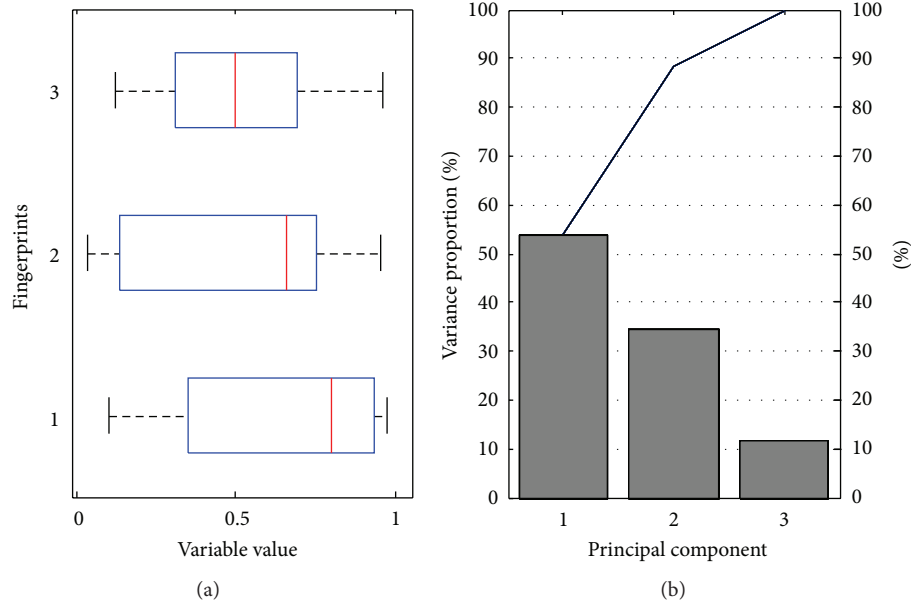


FIGURE 9: PCA principal component analysis. (a) Box plot of variable distribution and (b) Pareto chart.

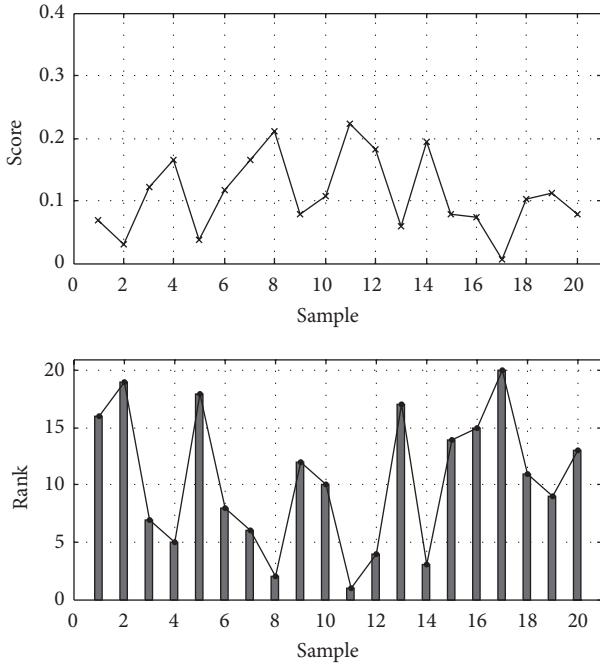


FIGURE 10: Score and rank of PCA.

precision and convergence time. We use the combination of appointed training set empirical parameters.

5.2. Number of User Impact. We evaluate precision and scalability of moving individual neighbor discovery by Euclidean distance error of similarity sorting. Five groups of experiments, respectively, include similar points 5, 10, 15, 20, and 25. The result compared with the real similarity sorting is shown in Figure 12. Euclidean distance error is maximum,

only 4.69 when the similar points is 25, while it is minimum, only 2, when similar points is 5. Figure 13 shows the real and measuring comparison when similar points is 20; therefore, number of users has little impact on experiment.

Besides behavior, light and sound information has been considered in this paper. There are many other properties, such as wear, hobbies, weather, and building materials. About scalability of user attributes, multidimensional data can achieve data dimensionality reduction through principal component analysis method. Experiments are implemented on three different types of mobile phones built-in ubiquitous sensors, without prior knowledge, site survey, and explicit feedback, which satisfy universality.

5.3. Running Cost. Finally we evaluate time overhead for moving individual neighbor discovery. The behavior, light, and sound detection and principal component analysis conduct 20 groups of experiments, respectively. Figure 14 introduces time consumption comparison of three components and principal component analysis. The results show that each component consumes time that is all below 0.2 s, and dimension reduction is basically all within 0.3 s; only the first experiment data needs to be trained, as you can see it consumes longer time, but also under 0.45 s, and the total elapsed time is under 0.75 s. From the above experimental analysis, we can see that our method satisfies real-time.

6. Discussion

According to different living habits and customs, the users may place their phones differently such as in hand, in pocket, in purse, or in a backpack, which causes the inertial deviation for behavior analysis. When we collect light and audio fingerprint, we must ensure that the light sensor and

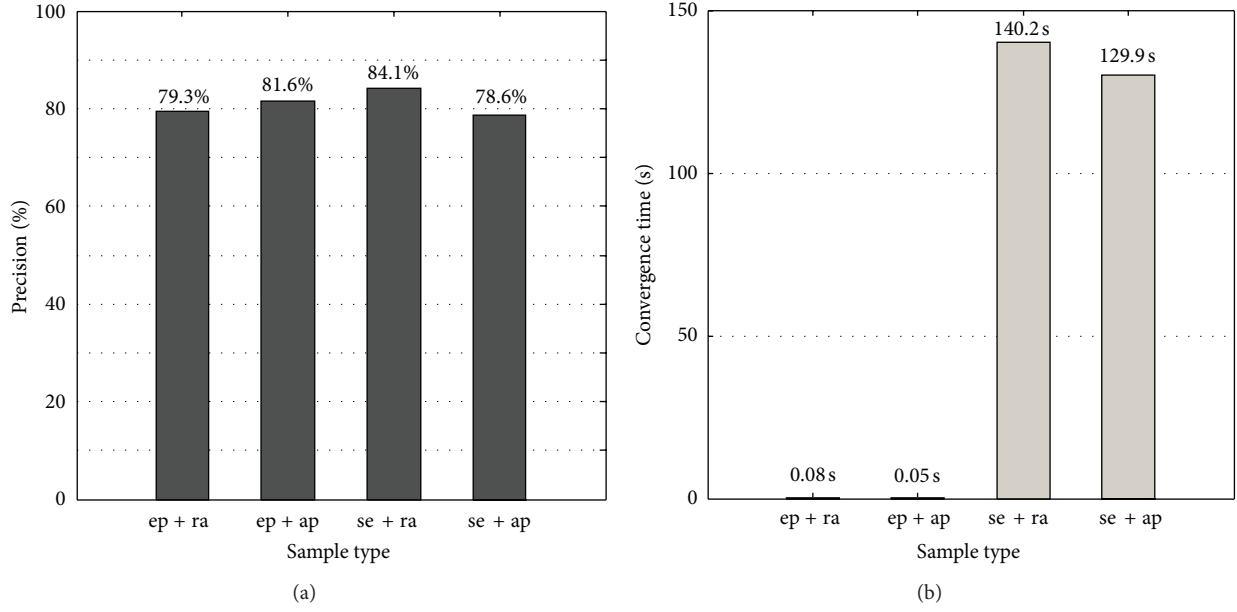


FIGURE 11: Precision of behavior recognition and convergence time (ep: empirical parameter, se: search parameter, ra: random training set, and ap: appointed training set). (a) Precision; (b) convergence time.

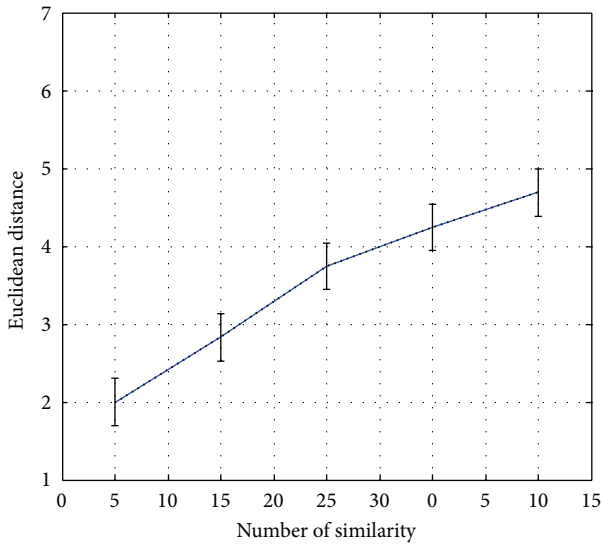


FIGURE 12: Euclidean distance error versus number of similarity.

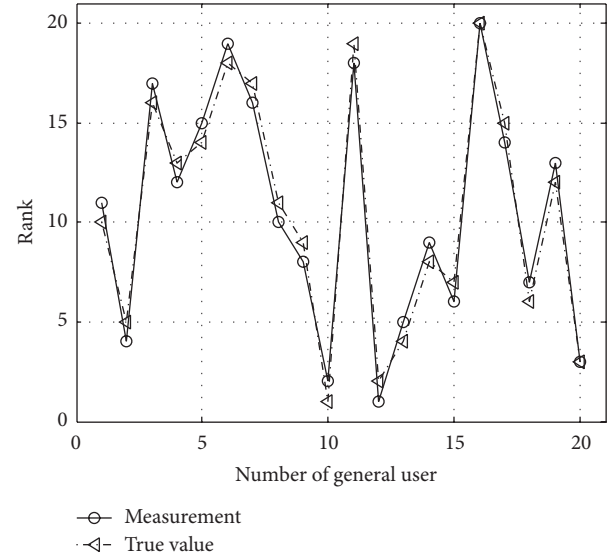


FIGURE 13: Comparison between measurement and true value.

the microphone are effective, so in this context, we do not need to consider that the positions of smartphones have an impact on behavior detection.

Once the number of super user increases, we need to judge all super users belonging to certain clusters again, so that computational complexity is greatly increasing. In further work, we consider that at the first time, we judge whether the new super user belongs to the same cluster with the original one, if they belong to the same cluster, so no need to redefine, while we can also increase the super user agency, partly clustering, in order to reduce computational

complexity, so as to achieve the purpose of providing services to each other.

7. Conclusion

We present the design and implementation of a moving individual neighbor discovery system, which efficiently takes input from accelerometer, light, and microphone to derive the neighbor discovery information. We use the appropriate similarity measure and principal component analysis method to calculate and integrate similarity information.

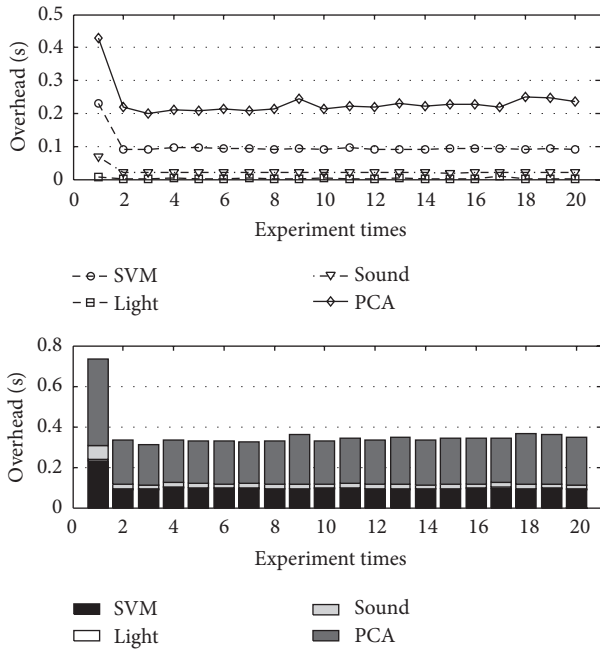


FIGURE 14: Convergence time of moving individual neighbor discovery.

Experimental results prove the validity of this method. Our method cannot only be used in services recommendation applications, recommending the same types of services to community groups, for example, a gym push fitness programs for persons who do exercises regularly. The experiment data was collected from 25 volunteers, and the results show that Euclidean distance error is below 4.69 and the convergence time is within 0.75 s.

Acknowledgments

This research is supported in part by National Natural Science Foundation of China under Grant nos. 61303233 and 61272466 and the Science and Technology Research and Development Program of Qinhuangdao nos. 2012021A045 and 2012021A046. The authors sincerely appreciate the suggestions and feedback from anonymous reviewers.

References

- [1] L. Wang, Y. He, Y. Liu, W. Liu, J. Wang, and N. Jing, "It is not just a matter of time: oscillation-free emergency navigation with sensor networks," in *Proceedings of the 33rd IEEE Real-Time Systems Symposium (RTSS '12)*, pp. 339–348, 2012.
- [2] Z. Yang and Y. Liu, "Quality of trilateration: confidence-based iterative localization," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 5, pp. 631–640, 2010.
- [3] Z. Yang, Y. Liu, and X.-Y. Li, "Beyond Trilateration: on the Localizability of Wireless Ad Hoc Networks," *IEEE/ACM Transactions on Networking*, pp. 1806–1814, 2010.
- [4] J. Jeong, S. Guo, Y. Gu, T. He, and D. H. C. Du, "Trajectory-based data forwarding for light-traffic vehicular ad hoc networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 5, pp. 743–757, 2011.
- [5] Z. Yang and Y. Liu, "Understanding node localizability of wireless ad-hoc networks," in *IEEE Transactions on Mobile Computing (TMC '10)*, pp. 1249–1260, March 2010.
- [6] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen, "Zee: zero-effort crowdsourcing for indoor localization," in *Proceedings of the 18th annual international conference on Mobile computing and networking (MOBICOM '12)*, pp. 293–304, 2012.
- [7] C. Wu, Z. Yang, Y. Liu, and W. Xi, "WILL: wireless indoor logical localization without site survey," in *Proceedings of the 31th IEEE International Conference on Computer Communications (INFOCOM '12)*, pp. 64–672, 2012.
- [8] Z. Yang, C. Wu, and Y. Liu, "Locating in fingerprint space: wireless indoor localization with little human intervention," in *Proceedings of the 18th annual international conference on Mobile computing and networking (MOBICOM '12)*, pp. 269–280, 2012.
- [9] D. H. Kim, Y. Kim, D. Estrin, and M. B. Srivastava, "SensLoc: sensing everyday places and paths using less energy," in *Proceedings of the 8th ACM International Conference on Embedded Networked Sensor Systems (SenSys '10)*, pp. 43–56, November 2010.
- [10] M. Bakht, M. Trower, and R. Kravets, "Searchlight: helping mobile devices find their neighbors," in *Proceedings of the 3rd ACM SOSP Workshop on Networking, Systems, and Applications on Mobile Handhelds (MobiHeld '11)*, pp. 185–196, October 2011.
- [11] R. Zhou, Y. Xiong, G. Xing, L. Sun, and J. Ma, "ZiFi: wireless LAN discovery via ZigBee interference signatures," in *Proceedings of the 16th Annual Conference on Mobile Computing and Networking (MobiCom '10)*, pp. 49–60, September 2010.
- [12] Y. Chon, E. Talipov, H. Shin, and H. Cha, "Mobility prediction-based smartphone energy optimization for everyday location monitoring," in *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems (SenSys '11)*, pp. 82–95, November 2011.
- [13] Y. Wang, J. Lin, M. Annavaram et al., "A framework of energy efficient mobile sensing for automatic user state recognition," in *Proceedings of the 7th ACM International Conference on Mobile Systems, Applications, and Services (MobiSys '09)*, pp. 179–192, June 2009.
- [14] M. Azizyan, I. Constandache, and R. R. Choudhury, "Surround-Sense: mobile phone localization via ambience fingerprinting," in *Proceedings of the 15th Annual ACM International Conference on Mobile Computing and Networking (MobiCom '09)*, pp. 261–272, September 2009.
- [15] C. Qin, X. Bao, R. Roy Choudhury, and S. Nelakuditi, "TagSense: a smartphone-based approach to automatic image tagging," in *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services (MobiSys '11)*, pp. 1–14, July 2011.
- [16] P. Zhou, Y. Zheng, Z. Li, M. Li, and G. Shen, "IODetector: a generic service for indoor outdoor detection," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems (SENSYS '12)*, pp. 113–126, 2012.
- [17] G. Ananthanarayanan and I. Stoica, "Blue-Fi: enhancing Wi-Fi performance using bluetooth signals," in *Proceedings of the 7th ACM International Conference on Mobile Systems, Applications, and Services (MobiSys '09)*, pp. 249–261, June 2009.
- [18] T. Wang, W. Jia, and P. Shen, "An efficient way for relative mobile localization," in *Proceedings of the 32nd International Conference on Distributed Computing Systems Workshops (ICDCS '12)*, pp. 209–215, 2012.

- [19] S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava, "Using mobile phones to determine transportation modes," *ACM Transactions on Sensor Networks*, vol. 6, no. 2, article 13, 2010.
- [20] A. Thiagarajan, J. Biagioni, T. Gerlich, and J. Eriksson, "Cooperative transit tracking using smart-phones," in *Proceedings of the 8th ACM International Conference on Embedded Networked Sensor Systems (SenSys '10)*, pp. 85–98, November 2010.

