Journal of Information Processing Vol.23 No.2 94–104 (Mar. 2015)

[DOI: 10.2197/ipsjjip.23.94]

Invited Paper

Mobile Devices as an Infrastructure: A Survey of Opportunistic Sensing Technology

Takamasa Higuchi[1,a)] Hirozumi Yamaguchi[1,b)] Teruo Higashino[1,c)]

Received: September 12, 2014, Accepted: November 12, 2014

Abstract: Now that billions of people carry sensor-enabled mobile devices (e.g., smartphones), employing powerful capability of such commercial mobile products has become a promising approach for large-scale environmental and human-behavioral sensing. Such a new paradigm of scalable context monitoring is known as *opportunistic sensing*, and has been successfully applied to a broad range of applications. In this paper, we briefly introduce basic architecture and building blocks on which these emerging systems are based, and then provide a survey of recent progress in the opportunistic sensing technology.

Keywords: opportunistic sensing, participatory sensing, crowd behavior, environmental monitoring, crowd sourcing

1. Introduction

The market of smartphones and wearable devices is steadily expanding in a yearly basis. According to a recent survey, global smartphone shipments have reached one billion in 2013 [1], and the number of such new-generation mobile products in use will no doubt continue to increase at a substantial pace. Considering the fact that mobile phones today are equipped with a variety of sensors such as GPS, accelerometers, compasses, gyroscopes, barometers, cameras, microphones, light/proximity sensors and so on, they are no longer just a means of human communication but also powerful sensing tools, which can faithfully capture human behavior and the context in which the user is involved. The potential of sensor-enabled mobile phones have been proved in a huge number of epoch-making commercial mobile applications as well as extensive research work.

Since billions of people carry such mobile products with a rich set of sensors, each of which has access to the Internet via cellular or WiFi networks, we can see them as a huge sensor network that spans across the world. This enables a totally different way of observing and understanding physical phenomena in real world. Here, let us take crowd density sensing as an example. Measurement of crowd density had been, or is even still manually done by "human-wave tactics," recruiting a number of field workers who count the number of pedestrians passing a specific region. Obviously, this approach is far from efficient in terms of human effort and monetary cost. Another possible solution is to deploy sensor devices in the area of interest to automatically count the number of pedestrians. Vision-based pedestrian tracking using CCTV cameras [2], [3] falls into this category. As well as camera devices, other types of sensors such as passive infrared (IR) sensors [4], [5], laser range finders [6] and sensitive floors [7] have been successfully applied to crowd counting and tracking systems. Despite the minimal dependence on human intervention, they have a severe limitation on coverage of monitoring since they need pre-installed sensing infrastructure. The mobile sensing approach can effectively cope with such a coverage issue. For example, Weppner et al. [8] have recently proposed a Bluetooth-based solution for the crowd density sensing problem. A small proportion of mobile phone users periodically probe neighboring Bluetooth-enabled devices using a special mobile application that runs on their own mobile phones. The neighbor information is then reported to a centralized server with the corresponding location information, so that the server can aggregate the user-contributed data from multiple volunteers to roughly estimate crowd density in each pre-defined region. A strong advantage of this approach is that it is not dependent on any sensing infrastructure, except for a single central server. As far as there are a sufficient number of volunteers who contribute the local sensing results with their mobile phones, the system can recognize the spatial distribution of pedestrian crowds with minimal human intervention. Such a recently emerged paradigm of large-scale context monitoring is called *opportunistic sensing*, or *participatory sensing*. The idea of opportunistic sensing has been successfully applied not only for crowd density sensing but also a wide range of applications such as environmental sensing [9], monitoring of transportation systems [10], [11], [12] and observation of social network structure [13].

Continuous monitoring of human behavior and the current situation in the area of interest is not only being a core technology of pervasive computing, but also has plenty of potential to optimize our society as a whole. For example, the spatial distribution of pedestrian crowds at a city scale is valuable information for disaster control and urban planning as well as human navigation. In

1 Graduate School of Information Science and Technology, Osaka University, Suita, Osaka 565–0871, Japan
a) t-higuti@ist.osaka-u.ac.jp
b) h-yamagu@ist.osaka-u.ac.jp
c) higashino@ist.osaka-u.ac.jp

addition, trajectories and density of pedestrian crowds in a public building (e.g., a shopping mall) can also be effectively utilized for marketing, crowd control and intelligent energy management, in which air conditioning or lighting systems are adaptively controled according to the number of pedestrians in each region. Another promising application of such sensing technology is road traffic monitoring. Recent studies have shown that traffic conditions of a future time can be reasonably predicted by analyzing the current and historical traffic conditions that are observed by on-vehicle sensors [14]. Based on the prediction, driver navigation systems may actively control the traffic by globally optimizing the routes that are recommended to the users to reduce the total amount of fuel consumption and carbon emission in the entire city. We believe that opportunistic sensing technology is a key enabler of such *cyber-physical systems*, in which data from a huge number of heterogeneous sensors are analyzed to leverage the results for optimization of human behavior, social infrastructure and resource utilization.

In this paper, we first introduce basic architecture and building blocks, on which such emerging sensing systems are based. Then we provide a survey of recent progress in the opportunistic technology and discuss future challenges in this hot research topic. While there have been some survey papers on related technology, they just briefly review a broad range of topics on mobile-phone-based sensing in general [15], or provide tutorial and general discussion on opportunistic sensing [16]. Unlike these works, this paper focuses on introducing various successful instances from recent literature to show trends and state-of-the-art of opportunistic sensing technology.

2. Architecture

Basic architecture of opportunistic sensing systems is illustrated in **Fig. 1**. Each phone which contributes to the sensing task locally collects sensor measurement with its built-in sensors, and applies some preprocessing to the raw sensor data. A typical process in this phase is feature extraction, in which the sensor data are summarized into feature vectors with limited dimensions using some statistical criteria (e.g., mean, variance, etc.). This effectively reduces communication overhead accompanied by data uploading to a server, and mitigates the risk of privacy leakage from the contributed data. Because of the limited resources on mobile devices, it is highly desirable that computational complexity of the preprocessing task is as small as possible. Instead, the remaining data processing will be usually offloaded to the central server.

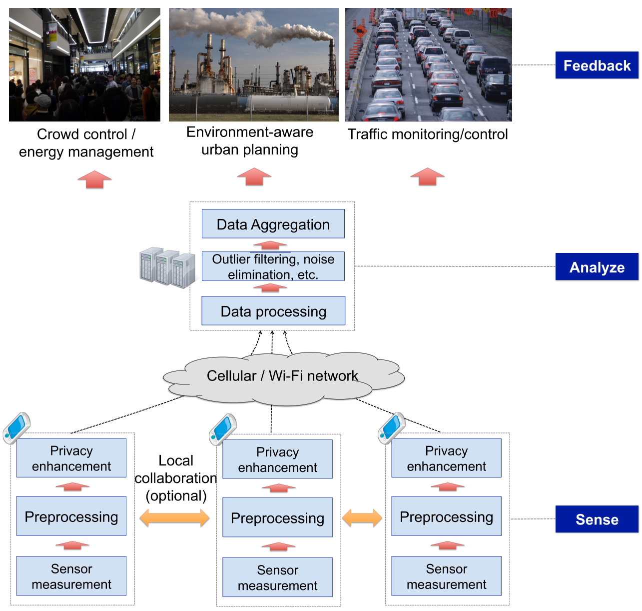

Fig. 1 Basic architecture of opportunistic mobile sensing.

Optionally, the user-contributed data may undergo further processing for privacy enhancement. It is known that, for some kind of widely-used feature values, the original sensor data can be roughly inferred by analyzing the feature vectors that are reported to a server [17]. This may incur a serious privacy threat for volunteers who contribute to the sensing service, especially if the original sensor data contains privacy-sensitive information (e.g., audio recordings during conversation). A variety of privacy enhancement approaches have been proposed to cope with such a threat and to incentivize the users to join the sensing service. Section 5.1 provides detailed discussion on the privacy issues in opportunistic sensing.

Some systems assume that neighboring mobile devices collaborate with each other via short-range wireless communication (e.g., Bluetooth) for locally aggregating sensor data to mitigate communication overhead accompanied by data uploading to a server [18], or to reduce battery consumption by strategically balancing their sensing tasks [19].

After the preprocessing followed by the optional privacy enhancement process, the mobile phone uploads the resulting information to a server via wireless network. Most of the existing mobile sensing systems assume that the user-contributed information is collected to a server via cellular network. However, the cellular-based communication often incurs some monetary expenses to the users, which would be a discouraging factor for voluntary contribution. If the application allows some delays in collecting sensor data, each mobile phone can temporarily store the data in its local storage, and upload them when it has access to a WiFi access point. Such an opportunistic networking approach effectively reduces “cost for contribution” and lowers the threshold for user’s participation.

Finally, the server further analyzes the user-contributed data to extract the information required by applications. Outlier filtering and noise elimination algorithms can be applied in this phase to enhance accuracy and robustness of the sensing results. Then, local information from multiple mobile phones are aggregated by appropriate decision fusion rules (e.g., majority voting) to derive the final estimate. The order of the three processes at server side can be changed depending on characteristics of the user-contributed data and requirement of the applications. Then the results of the analysis can be fed back to our society in various manners (e.g., crowd control, energy management, urban planning and traffic monitoring/control), as illustrated in Fig. 1.

3. Building Blocks

Opportunistic sensing systems are grounded on existing works on (individual) mobile sensing, such as localization, activity recognition, proximity sensing, sound recognition and visual scene understanding. This section briefly surveys these building blocks.

3.1 Localization

A significant proportion of mobile crowd sensing systems associate user-contributed sensor information with the corresponding geo-locations. For that purpose, the systems often require current locations of mobile phones along with a variety of sensor data. Although GPS is widely used for location acquisition of mobile phones, GPS receivers are extremely power-hungry and thus their continuous activation usually incurs unacceptable battery consumption. An effective way to mitigate battery consumption is strategically duty-cycling GPS receivers, so that the location errors are bounded below a tolerable level. EnTracked [20] employs accelerometers to distinguish movement state (i.e., stationary or in-motion) of the users, while estimating user's velocity based on the estimated positions obtained by GPS. Then it combines these information to determine the timing of the next position update. EnLoc [21] interpolates between consecutive location readings from GPS by predicting user's movement based on the past mobility patterns, which contributes to reduce instantaneous errors. RAPS [22] dynamically activates GPS when the expected position uncertainty exceeds a designated level and GPS fix is likely available according to the history of localization attempts at each area. Pedestrian dead reckoning (PDR) [23], [24], [25] is an alternative solution for energy-efficient localization, in which trajectory of a phone holder is estimated by detecting his walking steps and moving direction using accelerometers, compasses and/or gyroscopes. To cope with rapid accumulation of position errors due to sensor noise and irregular human motion, CompAcc [25] extracts possible moving paths from Google Maps and matches noisy PDR traces with those reference paths. Combination of PDR with the duty-cycled GPS is also a reasonable solution to achieve desired accuracy at smaller energy cost.

Another drawback of GPS is a lack of availability in indoor environments, where radio signals from GPS satellites are obstructed by the structure of buildings. WiFi-based localization [26], [27], [28] has been well investigated to complement GPS and widely used for indoor positioning of mobile/wearable devices. It basically requires a calibration phase, in which received signal strength from multiple WiFi access points are measured at each location in the building to construct a radio map. Once the above calibration is completed, location of mobile devices can be identified by matching the radio signature observed by the device with those in the radio map. Furthermore, the recent trend of configuration-free localization mitigates effort in the calibration process [29], [30]. As well as radio signals from WiFi access points, ambient sound [31], GSM [32], [33] and FM broadcast radios [34], [35] have also been proved to serve as effective location signatures in indoor environments. SurroundSense [36] employs environmental signatures like ambient sound, acceleration, color and light, which can be sensed by cameras and microphones in mobile phones. Combining these optical, acoustic and motion attributes with WiFi signatures, it can robustly distinguish adjacent locations that are separated only by a wall. While most of such fingerprint-based localization systems successfully work with off-the-shelf mobile devices, they have a drawback of vulnerability to environmental change. Thus the reference fingerprints need to be regularly updated to maintain reasonable positioning accuracy. PDR also provides an alternative solution for indoor positioning. In Ref. [37], errors in the PDR traces are mitigated by fusing received signal strength from WiFi access points and floor map information. References [23], [38] introduce particle-filter-based approaches that match the original PDR traces with floor maps to provide accurate indoor positioning.

Unfortunately, there has not been any localization technology that can provide fine-grained location in a ubiquitous manner regardless of the phone's context (e.g., indoor or outdoor, etc.) and availability of infrastructure. Currently, the most reasonable and widely accepted approach for location acquisition is to combine different types of techniques according to the surrounding situation and accuracy requirement by applications.

3.2 Activity Recognition Using Motion Sensors

Physical activities (e.g., walking, running, sitting, driving, etc.) of people in the area of interest strongly help understanding the situation and characteristics of the environment. So far, a considerable amount of research effort has been made to capture such human behavior by accelerometers, compasses and gyroscopes, which are commonly available in commercial mobile devices. Activity recognition systems are basically composed of two phases; (i) feature extraction from raw sensor data and (ii) inference by machine-learning-based classifiers, which take the feature values as input and output an estimated activity.

In order to fully extract information from raw sensor data, feature values are usually defined both in time and frequency domains. Typical features in the time domain include mean, standard deviation, variance, interquartile range (IQR), mean absolute deviation (MAD), correlation between axes, entropy and kurtosis [39], while frequency-domain features (e.g., energy in specific frequency bands) are extracted by applying Fourier Transform or Discrete Cosine Transform to the raw sensor data. Selection of feature values is a key process in design of activity recognition systems, since inappropriate features never contribute to accuracy improvement, or even may harm the system performance.

After the feature extraction is completed, the system estimates the current activity of the target person using activity classifiers, which are built beforehand based on a set of training data (i.e., feature values annotated with the corresponding activities). The classification algorithm should be carefully selected considering accuracy requirement and constraint on computational resources. k-NN [40] calculates distance between the observed feature values and those in the training data set to find k nearest instances. Then the estimated activity is determined based on the majority of corresponding activities for those k training data. Despite the ease of implementation, classification by k-NN is computationally expensive; its computational complexity increases in proportion to the number of training data. Decision tree [41] is one of the most popular approaches because of its small computational complexity in the inference phase. It organizes a set of inference rules in the form of hierarchical tree structure, in which each node is associated with a condition for the feature values. Support Vector Machine (SVM) [42] is another popular approach to feature classification for activity recognition. SVMs employ kernel functions to project training data to a higher dimensional space in order to find a linear decision boundary to separate those data. The hidden Markov model (HMM) [43] is a probabilistic model, in which a sequence of observable variables (i.e., feature values) is generated by a sequence of hidden Markov states. Thus it provides a reasonable solution if sensor data has some temporal dependencies. Bayesian approaches provide an effective way to cope with uncertainty in the feature values. A common idea is to find an activity that maximizes a posteriori probability, which is calculated based on a pre-defined likelihood model. Naïve Bayes [44] is the simplest form of such Bayesian approaches, in which each feature value is assumed to be conditionally independent, given a specific activity. While this assumption may not necessarily hold in practical scenarios, it has been shown that Naïve Bayes can achieve reasonable classification performance in a variety of applications.

3.3 Proximity Sensing

There has been a body of research on capturing social relationship among people (e.g., friend, families, etc.) based on sensor data contributed by a crowd of mobile phone users. Proximity patterns, or encounter information between mobile devices take an important role in discovering such social network structure.

Because of its ease of implementation and availability in a wide range of commercial products, the device discovery mechanism of Bluetooth has been successfully utilized for phone-to-phone proximity sensing in a variety of ubiquitous computing applications [13], [45], [46]. While Bluetooth has a shorter radio transmission range (typically about 10 m) compared to the ad-hoc mode of WiFi, and thus basically consumes lower energy, continuous proximity sensing still rapidly drains battery in mobile devices. eDiscovery [47] copes with this problem by adaptively adjusting the frequency of probing based on the number of neighbors that are detected by the recent probes.

Use of audio tones is also a reasonable option to obtain accurate proximity information. Since ambient noise usually concentrates in lower frequency components below 2 kHz, audio tones with higher frequencies can be robustly detected by applying Fast Fourier Transform (FFT) to the recorded audio samples. Reference [48] reports that the amplitude of audio signals steeply declines around 5 m from the transmitter phone, which effectively mitigates false positives in proximity sensing.

As well as just detecting phone-to-phone proximity, recent literature shows that audio tones also enable accurate distance measurement between mobile phones. Peng et al. [49] and Liu et al. [50] estimate distance between a pair of mobile phones based on propagation delays of audio beacons that are transmitted by each phone. Qiu et al. [51] enable 3-dimensional relative localization between smartphones equipped with two microphones. While they incur some additional overhead to cope with clock offset between the phones, they enable contact probing with much finer resolution and help accurate detection of social interactions among mobile phone users.

3.4 Sound Recognition

Sound recorded by microphones in mobile phones also contains a substantial amount of cues about human behavior, social interaction and current situation in the environment. For example, energy in the audio spectrum reflects the level of ambient noise, which has strong correlation with crowd density in the environment [52]. Speech recognition [53] and speaker identification [54] techniques further analyze detailed characteristics of the audio clips to capture people's verbal communication. As the first step of audio analysis, an audio clip is usually divided into a sequence of fixed-length frames. Then FFT is applied to these frames to calculate power of each frequency component (i.e., audio spectrum), from which a variety of features are extracted. SoundSense [55] is a pioneering work that enables detailed analysis of audio clips with the limited computational resources of mobile phones. It effectively combines the following 8 features from existing audio analysis literature to classify audio frames into three categories; speech, music and ambient sound:

(1) Zero Crossing Rate (ZCR) [56]: The number of zero crossings within a single frame in the time domain waveform.
(2) Low Energy Frame Rate [57]: The number of frames in a frame window whose root mean square (RMS) power is less than 50% of the average RMS power over the entire window.
(3) Spectral Flux (SF) [57]: L2-norm of the spectral amplitude difference vector of two adjacent frames.
(4) Spectral Rolloff (SRF) [58]: The frequency bin below which 93% of the distribution is concentrated.
(5) Spectral Centroid (SC) [58]: Centroid of the spectral power distribution.
(6) Bandwidth [58]: The width of the range of the frequencies that the signal occupies.
(7) Normalized Weighted Phase Deviation [59]: The phase deviations of the frequency bins in the spectrum weighted by their magnitude.
(8) Relative Spectral Entropy (RSE) [60]: The Kullback-Leibler divergence between the frequency domain of the current window of frames and previous 500 frames.

SoundSense constructs a decision tree based on all the features above to assign an estimated category to each audio frame. Finally it also applies smoothing to the sequence of outputs from the decision-tree-based classifier to correct instantaneous errors in the classification results.

For further analysis beyond sound classification, Mel Frequency Ceptral Coefficient (MFCC) [61] has been utilized in a variety of sound recognition applications (e.g., speech recognition). Human sound perception has fine-grained resolution for low frequency components, while being less sensitive to higher frequencies. MFCC divides the audio spectrum into frequency bins with different sizes to mimic this characteristic. SoundSense, which is mentioned above, also employs MFCC to find audio events from the set of classified audio frames. Despite its powerful audio representation capability and applicability to various types of analysis, calculation of MFCC features is computationally intensive and not suitable for real-time processing on mobile devices.

3.5 Visual Scene Understanding

Utilizing cameras on mobile devices is also a powerful means of situation recognition. Scene understanding based on images and videos have been a central research topic in the computer vision community. One of the important problems which are closely related to opportunistic sensing is human detection from monocular images. A number of techniques have been proposed in terms of features, models and general architectures [62] for robust human detection. As well as human bodies, other objects or scenes in general can be recognized once their visual model is constructed [63], [64], [65]. Most of existing scene recognition models basically assume outdoor environment, and accurate indoor scene recognition have been an open problem for a long period of time. Quattoni et al. [66] recently developed a scheme to cope with this challenging problem, which can classify indoor scenes into pre-defined categories (e.g., office, store, etc.).

In addition to scene recognition, image analysis technology has also been successfully applied to human activity recognition, which we have discussed in Section 3.2. Detailed analysis and a survey in this direction can be found in Ref. [67].

4. Applications

The idea of opportunistic sensing has been successfully applied to a wide range of applications. This section picks up some of successful instances from the recent literature on opportunistic sensing technology.

4.1 Environmental Sensing

Advantage of opportunistic sensing is its ease of achieving wide coverage. Thus it is in nature suitable for large-scale environmental monitoring. Ear-phone [9] is a participatory sensing system that monitors environmental noise pollution in urban areas, aiming at better awareness of noise levels and aid in the development of pollution mitigation strategies. It employs microphones in mobile phones to measure a loudness characteristic and builds a noise map by collecting the data to a centralized server with the phone's current location (obtained by GPS). To recover an accurate noise map from user-contributed data that are typically incomplete and randomly distributed in space and time, it effectively applies a set of algorithms for compressive sensing. Common Sense [68] applies the idea of opportunistic sensing to large-scale air quality monitoring. It employs a custom external sensor device equipped with commercial carbon monoxide, nitrogen oxides, ozone gas sensors as well as light, temperature, relative humidity, and orientation sensors. The device also has an 802.15.4 interface to collaborate with other sensors via a local low-energy wireless network, and also can connect to mobile phones via Bluetooth for visualization and data uploading purposes. The collected data are annotated with the corresponding location and a timestamp, and then uploaded to a server for dissemination, visualization and analysis over the web. MAQS [69] enables fine-grained indoor air quality sensing by integrating smartphones and portable CO_2 sensors. Instead of using sensors for various types of air pollutants, it proposes a novel air quality sensing method based on air exchange rates (i.e., how quickly the air is cycled through a room), which can be detected only by CO_2 sensors. Since people in the same room share similar air quality, it introduces zone-based collaborative sensing, in which multiple phones in vicinity share a single CO_2 sensor to save energy consumption. Unlike other opportunistic sensing systems, MAQS does not employ any centralized server. Instead, it shares the sensing data via the local network between the sensors and mobile phones.

4.2 Transportation

Monitoring of road traffic and public transportation systems is another promising applications that fully exploit high spatio-temporal coverage of opportunistic sensing.

CarTel [10] provides a novel mobile sensing platform to collect, process, deliver and visualize data from a collection of sensors equipped in vehicles. Cars temporarily store collected sensor data in their local databases, and opportunistically deliver them to a web-based portal when connection with e.g., WiFi access points is established. The portal visualizes these sensor data in a user-friendly manner, and allows mobile sensing applications to query it without awareness of spatial distribution and mobility of real sensors. The authors also provide a variety of case studies using CarTel, including analysis of driving patterns and road traffic (i.e., commute time analysis and traffic hot spot detection using collected GPS traces, as well as image acquisition from in-vehicle cameras). Nericell [11] enables monitoring of road and traffic condition using off-the-shelf smartphones instead of dedicated on-board sensors. To this goal, it provides a collection of sensing components to detect quality of the roads (e.g., potholes and bumps) and driving behavior (e.g., braking and honking) using the phone's accelerometer, microphone, GSM radio and/or GPS receiver. In order to reduce communication overhead, the analysis of sensor data is performed locally on the phone using energy-efficient and computationally inexpensive algorithms. Zhang et al. [70] estimate drivers' refueling activity in urban area based on GPS traces collected from taxicabs. By analyzing the spatio-temporal trajectory information, their system provides real-time estimates of waiting times of gas stations and an indicator of overall gas usage. These information can be effectively utilized for user refueling recommendation, planning of gas station deployment and macro-scale economic decisions based on energy consumption analysis. GreenGPS [71] maps fuel consumption on city streets to allow drivers to find the most fuel-efficient routes. It obtains fuel consumption measurements using the OBD-II interface (i.e., a standard interface to vehicle on-board sensors), and collects them to a centralized server by an opportunistic sensing approach. The authors show through their field experiments that the routes recommended by the GreenGPS system enable on average 10% savings in fuel consumption, compared to the routes constructed by different criteria. ParkNet [72] collects occupancy of road-side parking space by opportunistic sensing. It assumes that each vehicle is equipped with a GPS receiver and a ultrasonic rangefinder, which is used to sense availability of parking spot during driving on a city road. The data is then aggregated at a central server to build a real-time map of parking availability. Janecek et al. [73] leverage anonymized signaling data collected from a cellular mobile network to infer vehicle travel times and road congestion. They first estimate travel times and detect a presence of congestion using coarse-grained signaling data that can be collected from all the mobile phones, and then perform detailed analysis on fine-grained data from "active" phones (i.e., mobile phones engaged in a voice call or data connection) to localize the congestion. An advantage of their system is that it does not require any active sensing (e.g., uploading GPS traces) by mobile phone users and can use mobile cellular network as a large-scale mobility sensor.

Zhou et al. [12] developed a system to predict waiting time at bus stops based on opportunistic sensing by bus passengers. It automatically recognizes whether the user is on a bus by locally analyzing audio recordings from microphones in his/her mobile phones. Then the system identifies the current location of the bus based on a sequence of cell towers which the phones connect to. This effectively mitigates energy consumption on mobile phones, compared to the use of GPS. Maekawa et al. [74] enable car-level localization and congestion estimation for intelligent navigation of railway passengers. They collect radio connectivity and received signal strength (RSS) from neighboring phones by Bluetooth-based proximity sensing. The basic idea for localization is that the RSS steeply declines at boundaries of adjacent passenger cars. By clustering the network topology based on the RSS values, the system can recognize relative car-level location between passengers, which can then be converted to absolute position with support of a small number of passengers who manually report their correct location. The RSS information is also utilized for car congestion estimation since the radio signal is more likely to be obstructed by human bodies as the number of passengers in the car increases.

4.3 Crowd Behavior and Social Relationship

We have shown in Sections 3.1 and 3.2 that low-cost sensors in mobile devices have been successfully utilized to recognize location and activity of individual mobile phone users. Aggregating the behavior information of those individual users by the framework of opportunistic sensing, the system can obtain behavior of a collection of people (i.e., crowds) as well as social interactions between multiple phone users.

As mentioned in Section 1, Weppner et al. [8] proposed a crowd density estimation system based on Bluetooth-based proximity sensing by a small proportion of mobile phone users. A major challenge of this approach is that the ratio of people who carry a discoverable Bluetooth device may be low in practical scenarios. To cope with the issue, they define relative features that do not directly depend on the absolute number of devices in the environment and introduce machine-learning-based approach to classify the crowd density in the environment into multiple categories. Nishimura et al. [52] proposed a participatory sensing system, which is capable of detecting smoothness of pedestrian flows as well as crowd density, aiming to build a congestion map in public space (as in **Fig. 2**). Built-in accelerometers in mobile phones are leveraged to detect characteristics of pedestrian's walking motion which changes according the presence of congestion and smoothness of the surrounding pedestrian flow. In addition, magnitude of ambient audio noise is also detected by

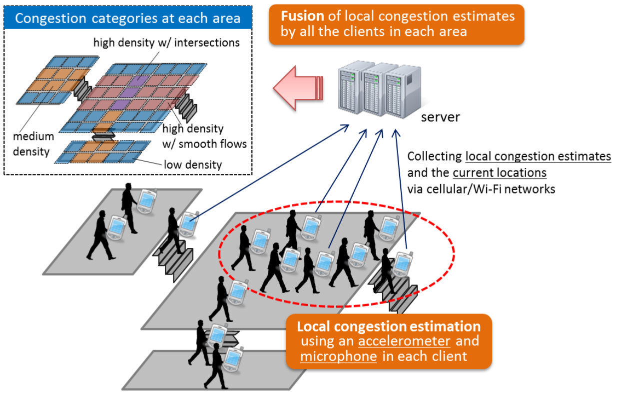

Fig. 2 Crowd density estimation by opportunistic sensing [52].

microphones to support the crowd density recognition. By incorporating the information from two different sensors and then aggregating the reports from multiple mobile phones, the system classifies the situation in pre-defined areas into four categories. Kannan et al. [75] developed a distributed system to count the number of mobile phone users in the area of interest. It employs audio tones, in which the list of detected phones is encoded into a set of frequency components. By repeatedly exchanging and updating these tones between neighboring devices using microphones and speakers in mobile phones, it accurately counts the number of unique phones in the environment. Isaacman et al. [76] modeled macro-scopic human mobility between different metropolitan areas by extensive analysis of Call Detail Records (CDRs) from a cellular network. Since human mobility substantially depends on the geography of the city people live in, the mobility model should take into account both the geography and individual user mobility patterns. To this goal, they sample the spatial and temporal probability distributions from CDR to build a model which is capable of generating sequences of locations and times of synthetic people. Such a mobility model can be effectively utilized for evaluation of mobile computing systems; e.g., to predict geographical coverage of opportunistic sensing applications. Kjærgaard et al. [77] detect pedestrian flocks (i.e., a group of pedestrians who move together for a certain period of time) in indoor environments. Their system extracts features from acceleration, compass and WiFi measurements and applies a clustering algorithm to these features to find the pedestrian flocks. The analysis of such flocking behavior helps evacuation management and development of socially-aware applications.

4.4 Situation Recognition and Semantic Reasoning

Opportunistic sensing is also an effective approach to understanding characteristics of geographic region or specific location. ConvenienceProbe [78] analyzes trajectory data offered by mobile phone users to identify retail trade areas, which is critical information for determining the optimal store location, finding competing stores and planning outdoor advertisements. Chon et al. [79] designed a framework for automatically recognizing categories of places (e.g., store, restaurant, etc.) based on opportunistically captured images and audio clips from smartphones. For this purpose, they fully exploit existing visual/audio analysis technology such as scene recognition, optical character recognition and speech recognition to find "hints" for categorizing the places (e.g., words spoken by people, text written on signs, and objects recognized in the environment). These hints are associated with the corresponding location data obtained by GPS or WiFi localization, and then analyzed by a centralized server. The characteristics of previously unseen places are learned using topic models, in which a collection of hints extracted from raw sensor data are grouped by the visited places to find their distributions. Once a topic model is established, it can be used to categorize the subsequent sensor data. iSee [80] detects and localizes specific events (e.g., presence of smokers or graffiti in public places) by participatory sensing with mobile phones. Unlike other systems introduced above, iSee employs users' manual input as *virtual sensors*. When a user finds a designated event, he reports it using a mobile application by swiping on his smartphone's touchscreen in the direction of the event, as well as offering the current GPS reading. The presence and locations of the events are then identified at a server by accumulating the reports from a number of volunteers over time.

5. Research Challenges

5.1 Privacy

User-contributed sensor data in opportunistic sensing systems often contains some privacy sensitive information. Basically, the users put some trust in system administrators who operate a centralized server that they never reveal the contributed data to other parties, nor use the data for other purposes than those originally stated. However, once the data is exposed to a malicious party by some reason, it incurs serious privacy threats. For example, many opportunistic sensing systems require the data to be annotated with the corresponding location information, from which the user's home, work place and daily routines can be easily inferred. Audio recordings and pictures can contain a rich amount of private information, including conversation, the surrounding situation and other persons who the user is interacting with. A simple solution to mitigate these threats is that sensing applications on mobile phones allow users to configure types, sampling frequency and granularity of sensor data they contribute to the server. Users may also explicitly specify sensitive sensor data that they do not intend to upload (e.g., locations of their home or work place). However, the resulting privacy level significantly depends on the knowledge of the users themselves on possible attacks by malicious users, and thus it may still happen that privacy-sensitive information is revealed in a way that the users have never intended. Alternative ways of privacy enhancement are to put bogus data or random noise onto the sensor data before uploading to a server, so that the statistical trends of a collection of user-contributed data are still preserved [17]. A common challenge of these approaches is to ensure a certain security level while minimizing the impact on performance of the system. While users can also control their privacy levels by adjusting the amount of noise that is added in the perturbation process, larger noise leads to less accuracy or coarse granularity in inference by the server. Finding the best trade-off point between privacy and performance is still an open problem. The interested readers may refer to Ref. [81], which provides extensive survey on state-of-the-art of privacy preservation technology for mobile sensing systems.

5.2 Data Collection

Utility of opportunistic sensing systems is directly dependent on the number and spatial distribution of volunteers who contribute to the sensing service with their mobile devices. This fact has motivated the research community to develop incentive mechanisms, so that the users willingly help sensing tasks and contribute valuable information. A possible way to stimulate active participation of users is to pay a small amount of money as a reward for one's contribution. Yang et al. [82] modeled an incentive mechanism under limited budget for the micropayment by game theory, and showed how to compute the equilibrium, at which the utility of the platform is maximized while none of the users can improve its utility by unilaterally deviating from its current strategy. They also proposed an auction-based incentive mechanism where users have more control over the payment they will receive. As well as the amount of data, its quality is also an important factor to maintain performance of the system, since some malicious users may upload bogus or significantly noisy data to degrade accuracy of inference by the system. To address this issue, Huang et al. [83] designed a reputation system, where the application server can evaluate the trustworthiness of contributing devices so that corrupted sensor data or malicious contributions are identified. The results can be utilized to enhance robustness of the systems e.g., by lowering the weights of malicious data in the computation of summary statistics. While these works have made important progress in the data collection problem, there still remain a number of problems to be solved such as efficient collaboration of incentive and reputation models and reasonable balancing between incentives/reputations and privacy.

5.3 Data Analysis

Intelligent analysis of sensor data is also an important challenge. Classification is one of the most popular approaches for recognizing the current context based on sensor measurements. While we have shown some of major classification algorithms in Section 3.2 as enablers of human activity recognition, they can also be applied to a wider range of applications. To build a machine learning model for such classification tasks, the system needs a sufficient amount of training data, which is composed of sensor measurements associated with the corresponding context labels (e.g., human activity, sound events, location categories, etc.). Traditionally, such annotated training data have been collected by developer of the system through some measurement campaigns. However, it requires substantial human burden, and may not scale as the system aims to support plenty of contexts for fine-grained situation recognition. In addition, robustness of the resulting model usually depends on the amount and diversity of the training data, which would be limited in such a traditional approach. A possible solution to build a robust model with minimal cost would be collaborative learning, in which mobile users contribute their sensor data as training samples [17], [84]. As with opportunistic sensing, this approach also incurs privacy concern and requires incentive/quality control mechanisms for data collection. Although Pickle [17] successfully addresses the privacy issue for some of popular machine learning algorithms, further analysis should be made to apply the idea of collaborative learning to practical sensing systems.

Automatically detecting meaningful patterns from collected sensor data would also be a promising direction of opportunistic sensing technology. Most of existing systems rely on models and classifiers that are trained beforehand, and thus are not able to recognize new objects, situations and events, which have not been seen in the training phase. Some recent works on mobile sensing try to remove this limitation by analyzing similarity between the user-contributed data. A common idea on which they are grounded is that the meaningful measurement patterns should be observed multiple times or by a number of mobile users. SoundSense [55] employs unsupervised adaptive learning techniques to discover and learn new types of sound events from the phone's audio recordings. Then it ranks potential sound events with respect to frequency and duration of occurrences to identify significant sounds for each user. UnLoc [85] applies a clustering algorithm to the sensor data collected from a number of mobile phones to automatically find identifiable ambient signatures in a building. For example, an elevator imposes a distinct pattern on a smartphone's accelerometer, while a corridor-corner may overhear a unique set of WiFi access points. They envision such ambient signatures as "landmarks" and use them to correct errors of indoor localization. Such automatic analysis mechanisms allow the sensing system to evolve its own capability in an autonomous manner and dramatically enhance their adaptability.

Another challenge is effective utilization of historical sensor data. Since the number of volunteers who contribute to opportunistic sensing would be usually limited in practical scenarios, it may often happen that the sensor data collected from a specific area is extremely few in a certain time range. It significantly harms performance of the system especially if their purpose is real-time monitoring of some physical phenomena (e.g., traffic flow, crowd density, noise level, etc.). In terms of robustness of the sensing systems, it must be desirable that reasonable accuracy can be maintained even with the limited spatio-temporal coverage of user-contributed data. In addition, some applications require future conditions in designated locations rather than the current status. For example, users of vehicle navigation systems would like to know traffic conditions at the time when the computed routes are actually driven. An effective way to address these requirements is to build an empirical model based on historical sensor data. Herring et al. [14] estimate real-time traffic flows and predict future traffic conditions based on spatially and temporally sparse GPS trajectories collected from a small amount of probe vehicles. They build historical distributions of travel time on each road segment and adaptively update these distributions using a probabilistic framework that represents dependency among the road segments. Yuan et al. [86] combine GPS logs collected from taxicabs and other context information available on the Internet (e.g., weather conditions) to build a detailed traffic model, which takes various external factors into account. The basic idea of utilizing historical data would be also effective for other types of sensing systems such as crowd flow sensing and environmental monitoring. In order to fully take advantage of such historical data, appropriate statistical models should be designed, considering requirement of applications (e.g., spatial and temporal granularity, accuracy, etc.), characteristics of the target phenomena to be monitored as well as types, acquisition frequency and quality of the sensor data.

6. Conclusion

In this paper, we have provided a survey of state-of-the-art on opportunistic sensing, in which a large number of mobile phones or other types of mobile devices contribute their sensor data via networks to achieve real world sensing at large scale. We have first briefly introduced basic architecture and building blocks on which the opportunistic sensing systems grounded, and then introduced some of important instances from the recent literature. Finally, we have discussed future challenges on this attracting research topic.

The ease of achieving wide coverage with minimal infrastructural cost is a strong advantage of opportunistic sensing, and would continue to stimulate further investigation and development of a plenty of potential novel applications. On the other hand, it may not be able to completely replace the existing infrastructure-based sensing systems, since its resolution of monitoring is in general less precise than those with pre-installed sensing infrastructure. We believe that seamless collaboration of pervasive opportunistic sensing and fine-grained sensing by infrastructure-based systems which are opportunistically available in the environment would be a reasonable direction of future real-world sensing.

Acknowledgments This work was supported in part by the KDDI foundation and the CPS-IIP Project (FY2012–FY2016) in the research promotion program for national level challenges by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

References

[1] IDC Corporate USA: Worldwide Smartphone Shipments Top One Billion Units for the First Time, available from ⟨http://www.idc.com/getdoc.jsp?containerId=prUS24645514⟩.

[2] Teixeira, T., Jung, D. and Savvides, A.: Tasking Networked CCTV Cameras and Mobile Phones to Identify and Localize Multiple People, *Proc. 12th ACM International Conference on Ubiquitous Computing* (*Ubicomp '10*), pp.213–222 (2010).

[3] Fleuret, F., Berclaz, J., Lengagne, R. and Fua, P.: Multicamera People Tracking with a Probabilistic Occupancy Map, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.30, No.2, pp.267–282 (2008).

[4] Song, B., Choi, H. and Lee, H.S.: Surveillance Tracking System Using Passive Infrared Motion Sensors in Wireless Sensor Network, *Proc. 2008 International Conference on Information Networking* (*ICOIN '08*), pp.1–5 (2008).

[5] Zappi, P., Farella, E. and Benini, L.: Tracking Motion Direction and Distance With Pyroelectric IR Sensors, *IEEE Sensors Journal*, Vol.10, No.9, pp.1486–1494 (2010).

[6] Zhao, H. and Shibasaki, R.: A novel system for tracking pedestrians using multiple single-row laser-range scanners, *IEEE Trans. Systems, Man and Cybernetics, Part A: Systems and Humans*, Vol.35, No.2, pp.283–291 (2005).

[7] Sousa, M., Techmer, A., Steinhage, A., Lauterbach, C. and Lukowicz, P.: Human tracking and identification using a sensitive floor and wearable accelerometers, *Proc. 11th IEEE International Conference on Pervasive Computing and Communications* (*PerCom '13*), pp.166–171 (2013).

[8] Weppner, J. and Lukowicz, P.: Bluetooth based collaborative crowd density estimation with mobile phones, *Proc. 11th IEEE International Conference on Pervasive Computing and Communications* (*PerCom '13*), pp.193–200 (2013).

[9] Rana, R.K., Chou, C.T., Kanhere, S.S., Bulusu, N. and Hu, W.: Ear-phone: An End-to-end Participatory Urban Noise Mapping System, *Proc. 9th ACM/IEEE International Conference on Information Processing in Sensor Networks* (*IPSN '10*), pp.105–116 (2010).

[10] Hull, B., Bychkovsky, V., Zhang, Y., Chen, K., Goraczko, M., Miu, A., Shih, E., Balakrishnan, H. and Madden, S.: CarTel: A Distributed Mobile Sensor Computing System, *Proc. 4th International Conference on Embedded Networked Sensor Systems* (*SenSys '06*), pp.125–138 (2006).

[11] Mohan, P., Padmanabhan, V.N. and Ramjee, R.: Nericell: Rich Monitoring of Road and Traffic Conditions Using Mobile Smartphones, *Proc. 6th ACM Conference on Embedded Network Sensor Systems* (*SenSys '08*), pp.323–336 (2008).

[12] Zhou, P., Zheng, Y. and Li, M.: How Long to Wait?: Predicting Bus Arrival Time with Mobile Phone Based Participatory Sensing, *Proc. 10th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '12*), pp.379–392 (2012).

[13] Do, T.M.T. and Gatica-Perez, D.: GroupUs: Smartphone Proximity Data and Human Interaction Type Mining, *Proc. 15th Annual International Symposium on Wearable Computers* (*ISWC '11*), pp.21–28 (2011).

[14] Herring, R., Hofleitner, A., Abbeel, P. and Bayen, A.: Estimating arterial traffic conditions using sparse probe data, *Proc. 13th IEEE International Conference on Intelligent Transportation Systems* (*ITSC '10*), pp.929–936 (2010).

[15] Lane, N.D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T. and Campbell, A.T.: A Survey of Mobile Phone Sensing, *Comm. Mag.*, Vol.48, No.9, pp.140–150 (2010).

[16] Ganti, R., Ye, F. and Lei, H.: Mobile crowdsensing: Current state and future challenges, *IEEE Communications Magazine*, Vol.49, No.11, pp.32–39 (2011).

[17] Liu, B., Jiang, Y., Sha, F. and Govindan, R.: Cloud-enabled Privacy-preserving Collaborative Learning for Mobile Sensing, *Proc. 10th ACM Conference on Embedded Network Sensor Systems* (*SenSys '12*), pp.57–70 (2012).

[18] Higuchi, T., Yamaguchi, H., Higashino, T. and Takai, M.: A Neighbor Collaboration Mechanism for Mobile Crowd Sensing in Opportunistic Networks, *Proc. 2014 IEEE International Conference on Communications* (*ICC 2014*), pp.42–47 (2014).

[19] Lee, Y., Ju, Y., Min, C., Kang, S., Hwang, I. and Song, J.: CoMon: Cooperative Ambience Monitoring Platform with Continuity and Benefit Awareness, *Proc. 10th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '12*), pp.43–56 (2012).

[20] Kjærgaard, M.B., Langdal, J., Godsk, T. and Toftkjær, T.: EnTracked: Energy-efficient robust position tracking for mobile devices, *Proc. 7th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '09*), pp.221–234 (2009).

[21] Constandache, I., Gaonkar, S., Sayler, M., Choudhury, R. and Cox, L.: EnLoc: Energy-Efficient Localization for Mobile Phones, *Proc. 28th Annual Joint Conference of the IEEE Computer and Communications Societies* (*INFOCOM '09*), pp.2716–2720 (2009).

[22] Paek, J., Kim, J. and Govindan, R.: Energy-efficient rate-adaptive GPS-based positioning for smartphones, *Proc. 8th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '10*), pp.299–314 (2010).

[23] Krach, B. and Robertson, P.: Integration of foot-mounted inertial sensors into a Bayesian location estimation framework, *Proc. 5th International Workshop on Positioning, Navigation and Communication* (*WPNC '08*), pp.55–61 (2008).

[24] Steinhoff, U. and Schiele, B.: Dead Reckoning from the Pocket — An Experimental Study, *Proc. 8th International Conference on Pervasive Computing and Communications* (*PerCom '10*), pp.162–170 (2010).

[25] Constandache, I., Choudhury, R.R. and Rhee, I.: Towards Mobile Phone Localization without War-Driving, *Proc. 29th Annual Joint Conference of the IEEE Computer and Communications Societies* (*INFOCOM '10*), pp.1–9 (2010).

[26] Bahl, P. and Padmanabhan, V.N.: RADAR: An in-building RF-based user location and tracking system, *Proc. 19th Annual Joint Conference of the IEEE Computer and Communications Societies* (*INFOCOM '00*), pp.775–784 (2000).

[27] Yin, J., Yang, Q. and Ni, L.M.: Learning adaptive temporal radio maps for signal-strength-based location estimation, *IEEE Trans. Mobile Computing*, Vol.7, No.7, pp.869–883 (2008).

[28] Youssef, M. and Agrawala, A.: The Horus WLAN location determination system, *Proc. 3rd International Conference on Mobile Systems, Applications, and Services* (*MobiSys '05*), pp.205–218 (2005).

[29] Chintalapudi, K.K., Iyer, A.P. and Padmanabhan, V.: Indoor Localization Without the Pain, *Proc. 16th International Conference on Mobile Computing and Networking* (*MobiCom '10*), pp.173–184 (2010).

[30] Lim, H., Kung, L.-C., Hou, J.C. and Luo, H.: Zero-Configuration, Robust Indoor Localization: Theory and Experimentation, *Proc. 25th IEEE International Conference on Computer Communications* (*INFOCOM '06*), pp.1 –12 (2006).

[31] Tarzia, S.P., Dinda, P.A., Dick, R.P. and Memik, G.: Indoor Localization Without Infrastructure Using the Acoustic Background Spectrum, *Proc. 9th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '11*), pp.155–168 (2011).

[32] Varshavsky, A., Pankratov, D., Krumm, J. and de Lara, E.: Calibree: Calibration-free localization using relative distance estimations, *Proc. 6th International Conference on Pervasive Computing* (*Pervasive '08*), pp.146–161 (2008).

[33] Varshavsky, A., de Lara, E., Hightower, J., LaMarca, A. and Otsason, V.: GSM indoor localization, *Pervasive and Mobile Computing, Elsevier*, Vol.3, No.6, pp.698–720 (2007).

[34] Chen, Y., Lymberopoulos, D., Liu, J. and Priyantha, B.: FM-based indoor localization, *Proc. 10th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '12*), pp.169–182 (2012).

[35] Popleteev, A., Osmani, V. and Mayora, O.: Investigation of indoor localization with ambient FM radio stations, *Proc. 10th IEEE International Conference on Pervasive Computing and Communications* (*PerCom '12*), pp.171–179 (2012).

[36] Azizyan, M., Constandache, I. and Roy Choudhury, R.: SurroundSense: Mobile phone localization via ambience fingerprinting, *Proc. 15th Annual International Conference on Mobile Computing and Networking* (*MobiCom '09*), pp.261–272 (2009).

[37] Wang, H., Lenz, H., Szabo, A., Bamberger, J. and Hanebeck, U.: WLAN-Based Pedestrian Tracking Using Particle Filters and Low-Cost MEMS Sensors, *Proc. 4th Workshop on Positioning, Navigation and Communication* (*WPNC '07*), pp.1–7 (2007).

[38] Woodman, O. and Harle, R.: Pedestrian Localisation for Indoor Environments, *Proc. 10th International Conference on Ubiquitous Computing* (*UbiComp '08*), pp.114–123 (2008).

[39] Lara, O. and Labrador, M.: A Survey on Human Activity Recognition using Wearable Sensors, *IEEE Communications Surveys Tutorials*, Vol.15, No.3, pp.1192–1209 (2013).

[40] Jatoba, L.C., Grossmann, U., Kunze, C., Ottenbacher, J. and Stork, W.: Context-aware mobile health monitoring: Evaluation of different pattern recognition methods for classification of physical activity, *Proc. 30th Annual IEEE International Conference on Engineering in Medicine and Biology Society* (*EMBS '08*), pp.5250–5253 (2008).

[41] Bao, L. and Intille, S.S.: Activity recognition from user-annotated acceleration data, *Proc. 2nd International Conference on Pervasive Computing* (*Pervasive '04*), pp.1–17 (2004).

[42] Anguita, D., Ghio, A., Oneto, L., Parra, X. and Reyes-Ortiz, J.L.: Human Activity Recognition on Smartphones Using a Multiclass Hardware-friendly Support Vector Machine, *Proc. 4th International Conference on Ambient Assisted Living and Home Care* (*IWAAL '12*), pp.216–223 (2012).

[43] Lee, Y.-S. and Cho, S.-B.: Activity Recognition Using Hierarchical Hidden Markov Models on a Smartphone with 3D Accelerometer, *Proc. 6th International Conference on Hybrid Artificial Intelligent Systems* (*HAIS '11*), pp.460–467 (2011).

[44] Tapia, E., Intille, S. and Larson, K.: Activity Recognition in the Home Using Simple and Ubiquitous Sensors, *Proc. 2nd International Conference on Pervasive Computing* (*Pervasive '04*), pp.158–175 (2004).

[45] Higuchi, T., Yamaguchi, H. and Higashino, T.: Clearing a Crowd: Cotext-Supported Neighbor Positioning for People-Centric Navigation, *Proc. 10th International Conference on Pervasive Computing* (*Pervasive '12*), pp.325–342 (2012).

[46] Vu, L., Do, Q. and Nahrstedt, K.: Jyotish: A novel framework for constructing predictive model of people movement from joint WiFi/Bluetooth trace, *Proc. 9th IEEE International Conference on Pervasive Computing and Communications* (*PerCom '11*), pp.54–62 (2011).

[47] Han, B. and Srinivasan, A.: eDiscovery: Energy Efficient Device Discovery for Mobile Opportunistic Communications, *Proc. 20th IEEE International Conference on Network Protocols* (*ICNP '12*), pp.1–10 (2012).

[48] Constandache, I., Bao, X., Azizyan, M. and Choudhury, R.R.: Did you see Bob?: Human localization using mobile phones, *Proc. 16th International Conference on Mobile Computing and Networking* (*MobiCom '10*), pp.149–160 (2010).

[49] Peng, C., Shen, G., Zhang, Y., Li, Y. and Tan, K.: BeepBeep: A High Accuracy Acoustic Ranging System using COTS Mobile Devices, *Proc. 5th ACM Conference on Embedded Networked Sensor Systems* (*SenSys '07*), pp.1–14 (2007).

[50] Liu, H., Gan, Y., Yang, J., Sidhom, S., Wang, Y., Chen, Y. and Ye, F.: Push the Limit of WiFi Based Localization for Smartphones, *Proc. 18th Annual International Conference on Mobile Computing and Networking* (*Mobicom '12*), pp.305–316 (2012).

[51] Qiu, J., Chu, D., Meng, X. and Moscibroda, T.: On the feasibility of real-time phone-to-phone 3D localization, *Proc. 9th ACM Confer-*

ence on Embedded Networked Sensor Systems (*SenSys '11*), pp.190–203 (2011).

[52] Nishimura, T., Higuchi, T., Yamaguchi, H. and Higashino, T.: Detecting Smoothness of Pedestrian Flows by Participatory Sensing with Mobile Phones, *Proc. 18th International Symposium on Wearable Computers* (*ISWC '14*) (2014).

[53] Deng, L. and Huang, X.: Challenges in Adopting Speech Recognition, *Comm. ACM*, Vol.47, No.1, pp.69–75 (2004).

[54] Reynolds, D.: An overview of automatic speaker recognition technology, *Proc. 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing* (*ICASSP '02*), Vol.4, pp.4072–4075 (2002).

[55] Lu, H., Pan, W., Lane, N.D., Choudhury, T. and Campbell, A.T.: SoundSense: Scalable sound sensing for people-centric applications on mobile phones, *Proc. 7th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '09*), pp.165–178 (2009).

[56] Saunders, J.: Real-time Discrimination of Broadcast Speech/Music, *Proc. 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing* (*ICASSP '96*), pp.993–996 (1996).

[57] Scheirer, E. and Slaney, M.: Construction and Evaluation of a Robust Multifeature Speech/Music Discriminator, *Proc. 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing* (*ICASSP '97*) (1997).

[58] Li, D., Sethi, I.K., Dimitrova, N. and McGee, T.: Classification of General Audio Data for Content-based Retrieval, *Pattern Recogn. Lett.*, Vol.22, No.5, pp.533–544 (2001).

[59] Dixon, S.: Onset detection revisited, *Proc. 9th International Conference on Digital Audio Effects* (*DAFx '06*), pp.133–137 (2006).

[60] Basu, S.: A linked-HMM model for robust voicing and speech detection, *Proc. 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing* (*ICASSP '03*), Vol.1, pp.816–819 (2003).

[61] Davis, S. and Mermelstein, P.: Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences, *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.28, No.4, pp.357–366 (1980).

[62] Enzweiler, M. and Gavrila, D.M.: Monocular Pedestrian Detection: Survey and Experiments, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.31, No.12, pp.2179–2195 (2009).

[63] Malisiewicz, T., Gupta, A. and Efros, A.A.: Ensemble of exemplar-SVMs for Object Detection and Beyond, *Proc. 2011 International Conference on Computer Vision* (*ICCV '11*), pp.89–96 (2011).

[64] Torralba, A., Fergus, R. and Freeman, W.: 80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.30, No.11, pp.1958–1970 (2008).

[65] Xiao, J., Hays, J., Ehinger, K., Oliva, A. and Torralba, A.: SUN database: Large-scale scene recognition from abbey to zoo, *Proc. 2010 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR '10*), pp.3485–3492 (2010).

[66] Quattoni, A. and Torralba, A.: Recognizing indoor scenes, *Proc. 2013 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR '13*), pp.413–420 (2009).

[67] Turaga, P., Chellappa, R., Subrahmanian, V.S. and Udrea, O.: Machine Recognition of Human Activities: A Survey, *IEEE Trans. Cir. and Sys. for Video Technol.*, Vol.18, No.11, pp.1473–1488 (2008).

[68] Willett, W., Aoki, P., Kumar, N., Subramanian, S. and Woodruff, A.: Common Sense Community: Scaffolding Mobile Sensing and Analysis for Novice Users, *Proc. 8th International Conference on Pervasive Computing* (*Pervasive '10*), pp.301–318 (2010).

[69] Jiang, Y., Li, K., Tian, L., Piedrahita, R., Yun, X., Mansata, O., Lv, Q., Dick, R.P., Hannigan, M. and Shang, L.: MAQS: A Personalized Mobile Sensing System for Indoor Air Quality Monitoring, *Proc. 13th International Conference on Ubiquitous Computing* (*UbiComp '11*), pp.271–280 (2011).

[70] Zhang, F., Wilkie, D., Zheng, Y. and Xie, X.: Sensing the Pulse of Urban Refueling Behavior, *Proc. 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (*UbiComp '13*), pp.13–22 (2013).

[71] Ganti, R.K., Pham, N., Ahmadi, H., Nangia, S. and Abdelzaher, T.F.: GreenGPS: A Participatory Sensing Fuel-efficient Maps Application, *Proc. 8th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '10*), pp.151–164 (2010).

[72] Mathur, S., Jin, T., Kasturirangan, N., Chandrasekaran, J., Xue, W., Gruteser, M. and Trappe, W.: ParkNet: Drive-by Sensing of Road-side Parking Statistics, *Proc. 8th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '10*), pp.123–136 (2010).

[73] Janecek, A., Hummel, K.A., Valerio, D., Ricciato, F. and Hlavacs, H.: Cellular Data Meet Vehicular Traffic Theory: Location Area Updates and Cell Transitions for Travel Time Estimation, *Proc. 2012 ACM Conference on Ubiquitous Computing* (*UbiComp '12*), pp.361–370 (2012).

[74] Maekawa, Y., Uchiyama, A., Yamaguchi, H. and Higashino, T.: Car-level Congestion and Position Estimation for Railway Trips Using Mobile Phones, *Proc. 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (*UbiComp '14*) (2014).

[75] Kannan, P.G., Venkatagiri, S.P., Chan, M.C., Ananda, A.L. and Peh, L.-S.: Low Cost Crowd Counting Using Audio Tones, *Proc. 10th ACM Conference on Embedded Networked Sensor Systems* (*SenSys '12*), pp.155–168 (2012).

[76] Isaacman, S., Becker, R., Cáceres, R., Martonosi, M., Rowland, J., Varshavsky, A. and Willinger, W.: Human Mobility Modeling at Metropolitan Scales, *Proc. 10th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '12*), pp.239–252 (2012).

[77] Kjærgaard, M.B., Wirz, M., Roggen, D. and Tröster, G.: Detecting Pedestrian Flocks by Fusion of Multi-modal Sensors in Mobile Phones, *Proc. 2012 ACM Conference on Ubiquitous Computing* (*UbiComp '12*), pp.240–249 (2012).

[78] You, C.-W., Kao, H.-L., Ho, B.-J., Chen, Y.-H., Wang, W.-F., Bei, L.-T., Chu, H.-H. and Chen, M.-S.: ConvenienceProbe: A Phone-Based System for Retail Trade-Area Analysis, *IEEE Pervasive Computing*, Vol.13, No.1, pp.64–71 (2014).

[79] Chon, Y., Lane, N.D., Li, F., Cha, H. and Zhao, F.: Automatically Characterizing Places with Opportunistic Crowdsensing Using Smartphones, *Proc. 2012 ACM Conference on Ubiquitous Computing* (*UbiComp '12*), pp.481–490 (2012).

[80] Ouyang, R.W., Srivastava, A., Prabahar, P., Choudhury, R.R., Addicott, M. and McClernon, F.J.: If You See Something, Swipe Towards It: Crowdsourced Event Localization Using Smartphones, *Proc. 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (*UbiComp '13*), pp.23–32 (2013).

[81] Christin, D., Reinhardt, A., Kanhere, S.S. and Hollick, M.: A Survey on Privacy in Mobile Participatory Sensing Applications, *J. Syst. Softw.*, Vol.84, No.11, pp.1928–1946 (2011).

[82] Yang, D., Xue, G., Fang, X. and Tang, J.: Crowdsourcing to Smartphones: Incentive Mechanism Design for Mobile Phone Sensing, *Proc. 18th Annual International Conference on Mobile Computing and Networking* (*Mobicom '12*), pp.173–184 (2012).

[83] Huang, K.L., Kanhere, S.S. and Hu, W.: Are You Contributing Trustworthy Data?: The Case for a Reputation System in Participatory Sensing, *Proc. 13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems* (*MSWIM '10*), pp.14–22 (2010).

[84] Miluzzo, E., Cornelius, C.T., Ramaswamy, A., Choudhury, T., Liu, Z. and Campbell, A.T.: Darwin Phones: The Evolution of Sensing and Inference on Mobile Phones, *Proc. 8th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '10*), pp.5–20 (2010).

[85] Wang, H., Sen, S., Elgohary, A., Farid, M., Youssef, M. and Choudhury, R.R.: No need to war-drive: Unsupervised indoor localization, *Proc. 10th International Conference on Mobile Systems, Applications, and Services* (*MobiSys '12*), pp.197–210 (2012).

[86] Yuan, J., Zheng, Y., Xie, X. and Sun, G.: Driving with Knowledge from the Physical World, *Proc. 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (*KDD '11*), pp.316–324 (2011).

Takamasa Higuchi received his B.E., M.E. and Ph.D. degrees in Information and Computer Science from Osaka University in 2010, 2012 and 2014, respectively. Since 2014, he has been a Specially Appointed Assistant Professor at Osaka University. His current research interests include mobile sensing, location/situation awareness, and network protocols for pervasive computing. He is a member of IEEE and IPSJ.

Journal of Information Processing Vol.23 No.2 94–104 (Mar. 2015)

Hirozumi Yamaguchi received his B.E., M.E., and Ph.D. degrees in Information and Computer Science from Osaka University, Japan. He is currently an associate professor at Osaka University. His current research interests include design and implementation of distributed systems and communication protocols, especially protocols and services on mobile ad hoc networks and application layer multicast. Also, he is working on some issues of wireless network simulations. He is a member of IEEE and IPSJ.

Teruo Higashino received his B.S., M.S., and Ph.D. degrees in Information and Computer Science from Osaka University, Japan in 1979, 1981 and 1984, respectively. He joined the faculty of Osaka University in 1984. Since 2002, he has been a professor in Graduate School of Information Science and Technology at Osaka University. His current research interests include design and analysis of distributed systems, communication protocol and mobile computing. He is a senior member of IEEE, a fellow of IPSJ, and a member of ACM and IEICE.