Non-obstructive Room-level Locating System in Home Environments using Activity Fingerprints from Smartwatch

Seungwoo Lee[*, *], Yungeun Kim[*], Daye Ahn[**], Rhan Ha[**], Kyoungwoo Lee[*], Hojung Cha[*]**

[*] Department of Computer Science
Yonsei University
Seoul, Korea
{seungwoo, ygkim}@cs.yonsei.ac.kr
{kyoungwoo.lee, hjcha}@yonsei.ac.kr

[**] Department of Computer Science, Hongik University
Seoul, Korea
dayeahn@cs.hongik.ac.kr
rhanha@cs.hongik.ac.kr

[***] Korea Electronics Technology Institute
Seongnam-si, Korea
seungwoo.lee@keti.re.kr

ABSTRACT
Many smart home applications, such as monitoring for the elderly and home automation, require location information for individual occupants. Several techniques have been proposed for tracking occupants in a home environment. However, the current techniques do not provide a seamless in-home locating system owing to the occupants' device-free movement and the lack of cost-effective infrastructure for home location tracking. In this paper, we propose a home occupant tracking system that uses a smartphone and an off-the-shelf smartwatch without additional infrastructure. In our system, *activity fingerprints* are automatically generated from the microphone and the inertial sensors of the smartwatch, and location information is periodically obtained from the smartphone. We designed a hidden Markov model using the relationship between home activities and the room's location. Extensive experiments showed that our system tracks the location of users with 87% accuracy, even when there is no manual training for activities.

Author Keywords
In-home locating system; occupant tracking, machine learning; mobile sensing; context-aware computing.

ACM Classification Keywords
I.2.6 Artificial Intelligence: Learning; J.4 Computer Applications: Social and Behavior Sciences.

INTRODUCTION
The location information of an individual at home is a core component of pervasive computing applications such as life-logging [1], activities of daily living (ADLs) [2], monitoring of elderly individuals [3], and smart-home-related applications [4,5]. Extensive attempts have been made to acquire the location of individual users in indoor environments. However, there are no comprehensive solutions because most indoor locating systems sacrifice either cost-effectiveness or user convenience.

UbiComp '15, September 7-11, 2015, Osaka, Japan.
Copyright 2015 © ACM 978-1-4503-3574-4/15/09…$15.00.
http://dx.doi.org/10.1145/2750858.2804272

Several approaches have been attempted to track occupants at home using pre-installed infrastructure. For example, passive infrared (PIR) sensors [3,6] or camera sensors [7] are used to detect the movement of a person nearby. Radio-frequency identification (RFID) readers [8] and ultrasonic range finders [9] , or tomographic motion detection [10] are used to detect a person. Since these approaches are based on the deployment of human-detectable sensors at a target location, the approaches are costly and invade users' privacy [11]. Thus, infrastructure-based systems are usually targeted for specialized applications such as surveillance [12] and health monitoring [3] instead of casual use in a home environment.

Pedestrian dead reckon (PDR) methods are commonly used to track a user's location without requiring additional infrastructure and environmental information *a priori*. Conventional PDR systems are broadly classified into two categories: (1) PDR with mobile devices [13,14] and (2) PDR with wearable devices [15, 16, 17]. A PDR system with a mobile device targets personal devices such as smartphones to improve usability. A mobile-based PDR system focuses on estimating the location of a device, not a person; therefore, this system works only when the occupants carry the device. However, in a home environment, the user's location does not often coincide with the device's location because of the occupants' device-free movement [5,18]. For instance, an occupant may be playing a game in the living room while his or her smartphone is charging in another room. Meanwhile, since wearable devices such as watches or glasses are almost always attached to the human body, PDR with wearable sensors can update the user's location seamlessly. To reduce the accumulated error caused by the drift error of sensor measurements, wearable sensors are usually mounted on a user's foot [17,19] for zero-velocity update (ZUPT) [19]. However, wearing a dedicated device such as a foot-mounted sensor is a hindrance in daily life.

With the widespread use of diverse types of wearable devices, various types of wristbands and smartwatches are now commercially available [20,21,22,23]. A typical smartwatch is equipped with several sensors for activity or gesture recognition as well as a network module such as Bluetooth. Since conventional wearable devices cause issues for user convenience and resource limitations, the availability of a smartwatch, in particular, can change the paradigm in wearable computing because smartwatches are a general-purpose device that people are willing to wear and recharge. These types of devices enable a wide range of ubiquitous computing for context awareness in terms of practicability.

In this study, we propose a room-level locating system in home environments without the deployment of additional infrastructure. To accomplish this, the system uses a smartphone to obtain location information and a smartwatch to record *activity fingerprints* for inferring a user's location even when the user is not carrying a smartphone. The main idea is that an activity depends on the indoor space. Therefore, by using the activity fingerprints collected from the smartwatch, the proposed system infers the location of a user although the user does not often carry his or her smartphone in a home environment. The proposed system makes the following contributions:

- Without requiring dedicated devices or the deployment of a costly infrastructure, we propose a non-obstructive locating system using general-purpose smart devices such as a smartphone and a smartwatch.
- Our system is designed to minimize user annoyance and training costs. By combining the use of a smartphone and a smartwatch, the proposed system automatically estimates a user's location without requiring any user participation.
- To validate the feasibility and superiority of the proposed system, we evaluated our hidden Markov model (HMM)–based locating system by using real-world human activity datasets. To the best of our knowledge, our work is the first to leverage the activities of occupants to estimate their location in home environments.

RELATED WORK

Indoor Locating System with Mobile Devices

Among diverse approaches, the WiFi-based fingerprinting system is practical since existing WiFi access points (APs) are readily utilized [24]. Several issues should be considered for the practical use of received signal strength (RSS) fingerprint-based localization. The RSS map-building process typically requires an extensive and thorough site survey, which is usually done manually with specific hardware and software tools. Much effort has recently been spent on reducing the cost and complexity of building fingerprint maps. Some have focused on the effective method of constructing the map database itself [25] or reducing the training process, the core problem of off-line map building [26]. Rai et al. [27] proposed Zee that uses a map-based PDR for building radio maps based on mobile crowdsensing. Zee improves the accuracy of PDR with information extracted from floor plans, such as the location of walls, rooms, and obstacles. However, a detailed floor plan is not always available in practice, and the accuracy of PDR is low in large open spaces due to the absence of walls and obstacles.

Various sensors in smartphones can be used to indicate a user's indoor location. SurroundSense [28] builds a map using smartphones by extracting features that are found in a typical indoor space such as ambient sound, light, and color. The scheme focuses on extracting common features of indoor space. Thus, this scheme cannot be directly applied to the personalized approach because ambient features such as light and color may not be unique in home environments. In addition, in home environments, the location of a mobile device is not always guaranteed to correspond to the location of the device's owner.

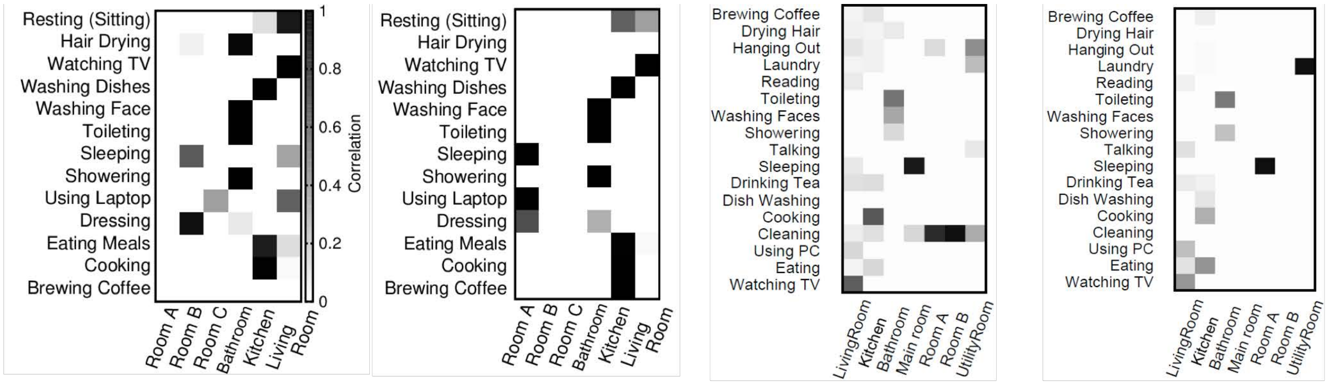
In-home Tracking System

Active research has been conducted to detect or track individuals in home environments. To detect the movement of occupants, low-cost motion sensors [3,12] such as PIR and ultrasonic sensors are commonly used. RF transmitters for detecting tomographic motion can be used to track occupants [10]. However, such sensors typically detect space occupation. Therefore, this scheme is not applicable to a personalized application because counting and identifying multiple occupants is simply not possible. The Doorjamb [9] tracking system was proposed to track multiple occupants using doorway-mounted range finders to measure the height of the person passing through for identification. Doorjamb provides room-level tracking without privacy-invasive sensors or requiring devices to be carried but fails to identify users if the occupants have similar heights.

Conversely, many studies [8,13,14,15] have focused on directly tracking physical devices instead of humans, assuming that a user's location corresponds to the location of the carried or worn devices. Inertial sensors or network modules in devices are used for locating purposes. For example, Doormat [8] uses wearable RFID tags. With a device-based tracking system, identifying multiple people is not challenging because all tags have a unique network or device number. However, carrying devices all the time or wearing burdensome foot-mounted sensors is less likely to be applicable in home environments. To address these concerns, our system is designed to track people using a non-intrusive wristwatch-type wearable device.

Machine Learning for Activity Recognition

Smartphones or wearable devices are often used to recognize users' activity. Accelerometers are a commonly used type of sensor for activity recognition with wearable sensors [29]. Some studies developed wearable devices with three-axis accelerometers [29,30] for detecting physical activities such as walking, turning into a corridor,



(a) user A (homemaker, home A) (b) user B (worker, home A) (c) user C(homemaker, home B) (d) user D (student, home B)

Figure 1. Correlation between activities and spaces.

or climbing up the stairs, another [17] detected falls and the movements of a user after a fall, and yet another [2] monitored a user's activity levels for fitness purposes using smartphones. Some recent commercial products [20] have been developed to track users' activity, such as the number of steps walked, quality of sleep, and calories burned using inertial measurement unit (IMU) sensors. In addition, microphones can be used to monitor the user's activities. BodyScope [31] recognizes activities such as eating, drinking, and speaking by analyzing sounds using an acoustic sensor around the neck.

PRELIMINARIES

As a preliminary experiment to analyze the relation between activities and location (i.e., a room), we collected the daily activities and corresponding locations of two users (A and B) in the same home for 3 weeks using a smartphone application. The participants manually annotated their current activities and indoor location periodically through the smartwatch application.

Relationship between Activity and Space

Users' activities are influenced by the space in which they are located. For example, activities such as "watching TV" and "sitting on the sofa" typically occur in a living room. Similarly, activities related to meals, such as "eating" and "drinking," typically occur at the dinner table in a dining room. To show the relationship between activities and space, we calculated the correlation between activities and locations as follows:

$$C_{a_i,l_j} = \frac{\#\ of\ labeled\ location\ is\ j\ in\ fingerprints\ for\ activity\ i}{\#\ of\ generated\ fingerprints\ for\ activity\ i}.$$

Figure 1 shows the results of the correlation for four users. The high value of the correlation coefficient indicates that the activity usually occurred at a corresponding location. This seems to mean that home activities are usually positively correlated with a specific space. A location that corresponds to the activity of user A or C may differ from that of user B or D, however, even if they live together. In addition, some activities are not clearly limited to one location. For instance, user A often used a laptop in either room C or the living room, whereas the PC/laptop activity of user B was entirely limited to room A. Homemakers (users A and C) visited almost all rooms in the house to clean, including almost all shared spaces, whereas other occupants (users B and D) usually visited only a few rooms, namely, their own rooms, the living room, and the bathroom. Consequently, the correlation between a space and an activity depends on the occupants, the activity type, and the appliances installed in the home environment.

Challenges of an Activity-based Locating System

The preliminary experiment showed that we can infer the location of people in a home environment by monitoring their activities. However, there are certainly challenges in recognizing activities in a home environment.

Limitation in Using a Wristwatch-type Device

The accelerometer of a smartwatch recognizes the features of a physical activity by measuring the acceleration force of the wrist. However, wearing a watch on a wrist is not enough for activity recognition. For example, people may wear the watch on their non-dominant hand. Certain activities are not distinguishable by using only the accelerometer measurement. To overcome this problem, previous studies have employed multiple accelerometers mounted on various parts of the body, such as the hip, thigh, wrist, and foot [29], to monitor activities more precisely. However, attaching multiple sensors to the human body is not practical owing to the uncomfortable user experience. To improve the precision of recognition, we exploited the acoustic features of a room or activities extracted by using the microphone on the smartwatch.

Diverse Patterns of Individual Behavior

The dependency of an activity on an indoor space varies according to the individual and the type of activity, as shown in Figure 1. Moreover, even the same person can act differently as people have different physical characteristics (i.e., age, height, weight) and habits. Therefore, a personalized learning mechanism is required for classifying human activities.

Training Costs of a Locating System

To estimate the user's location through activity recognition, the correlation between an activity and a space should be learned in advance. Moreover, this learning process should be performed in a personalized way owing to the user

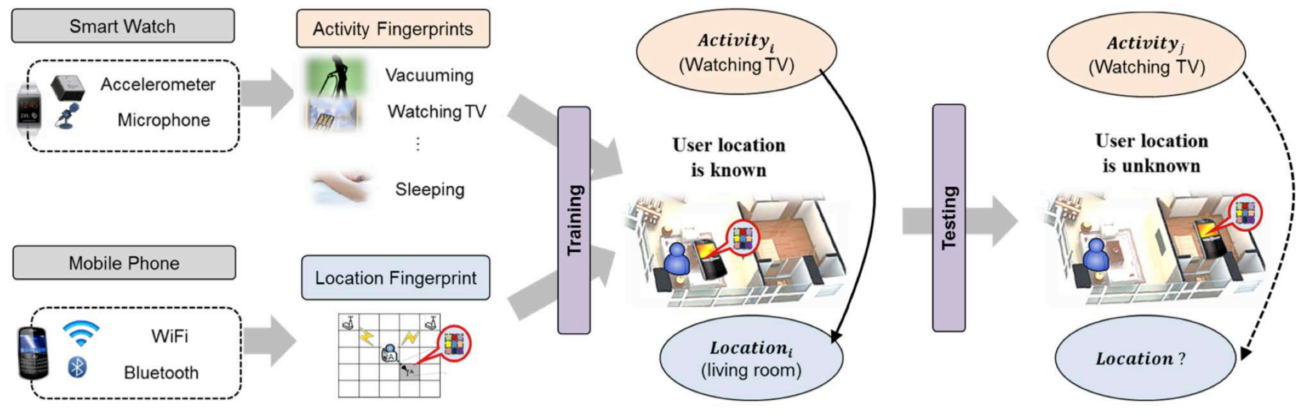


Figure 2. Activity-based location system scenario.

dependency. However, since the manual annotation of the current location of each user requires a significant effort, a learning method that autonomously estimates the user's location is required. We implemented an automatic learning method through the use of a watch device and a smartphone. We designed a home locating system that autonomously extracts the features of users' activities and learns the correlation between spaces and activities in the users' daily lives.

SYSTEM DESIGN

This section presents the details of the in-home tracking system using a smartwatch and a smartphone.

System Overview

Our system is designed to locate occupants in a home environment without requiring additional infrastructure. To monitor activities and movements in daily life, we selected off-the-shelf smartwatches that have become general-purpose wearable devices. Here, we assume that the users always wear their smartwatches except for a few situations (e.g., while charging the devices, sleeping, and taking a shower).

The system tracks the occupant in the following way. The smartwatch monitors a user's movement of a user and a smartphone periodically performs Wi-Fi positioning using the Wi-Fi RSS fingerprint to estima the "room" location of the device. Distances are measured and data are exchanged between a smartphone and a smartwatch periodically via Bluetooth (BT) communication. Based on the distance measurement, if an occupant is considered near his or her smartphone, the current location of the user is determined as the estimated location of the smartphone from the Wi-Fi positioning system. The activity fingerprints with room location information are added to the training samples to learn an inference model that considers the correlation between the space and the activities. However, when the occupant is considered away from the phone, the system estimates the occupant's location from the inference model that finds indoor space that is highly correlated to the activity observed by the smartwatch as shown in Figure 2.

System Architecture

Figure 3 illustrates the architecture of the proposed system. This system consists of the following components.

Feature Extraction from Wearable Devices

First, this component preprocesses raw sensor data to detect the movement of the user from the accelerometer sensor inputs. Then, the component extracts the activity fingerprints from various sensor data. The extracted features are transmitted to the smartphone by Bluetooth (BT) communication.

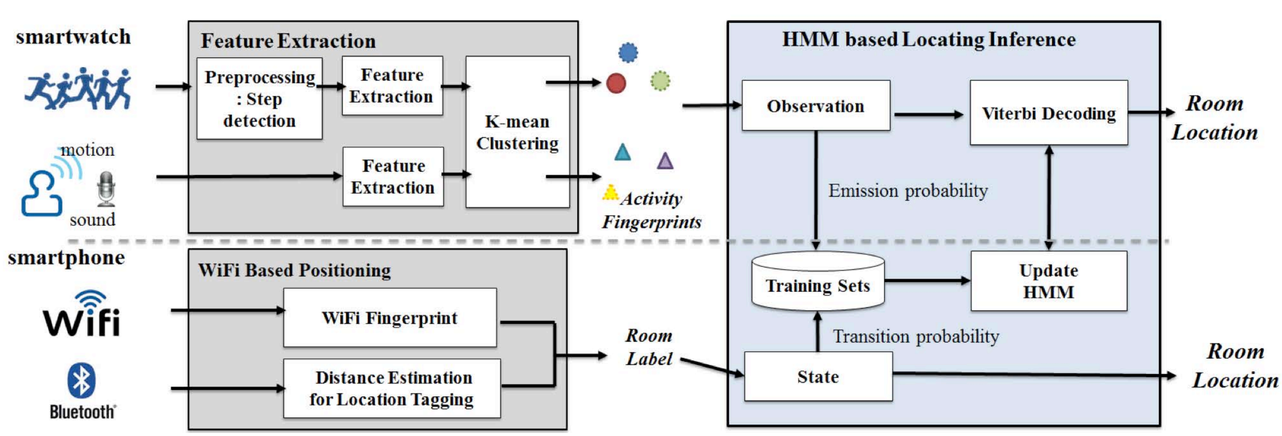


Figure 3. System architecture.

Estimating Location Using Smartphones

The smartphone periodically senses the signal strength from nearby Wi-Fi APs and obtains their locations based on the Wi-Fi fingerprint technique. The device estimates the distance using the paired smartwatch to determine whether the user is co-located.

Locating Occupants Based on the Hidden Markov Model

This component is composed of two parts. The first part is the training phase of the inference model when the smartphone is co-located with the user in the same space. With sufficient training samples, the system automatically constructs and updates the HMM-based inference model. In the second part, the system estimates users' location by using activity fingerprints obtained from the smartwatch.

ACTIVITY FINGERPRINT GENERATION

Preprocessing: Moving State Recognition

We focus on locating occupants in a home environment by activity recognition. Our idea is based on the precondition that the recognized activities should be performed in a relevant space. Therefore, we detect the occupants' moving state to filter out the activities while they are moving around (e.g., vacuuming and cleaning). By estimating the moving distance, the system can predict when the user's location is updated, as long as the user's movements indicate a change in his or her location. We categorize the moving state as one of three types: "stationary," "moving," and "unknown" (when the user takes off the watch). The moving distance and the moving state are estimated based on step detection [17,19,32]. We adopted the step detection algorithm proposed in [32,33]. The steps are counted based on the peak detection [33] of the acceleration in a smartwatch that is used to mark the candidate stepping points. The technical details are found in [32,33]. We define ΔT_{min} and ΔT_{max}, which are the minimum interval and maximum intervals between consecutive acceleration peaks. If more than two peaks are measured within ΔT_{min}, only the first peak is estimated as a step, and the other peaks are discarded to prevent false positive detection of steps. We set ΔT_{min} at 0.33 s because normal users walk fewer than three steps per second at home. If more than two steps are counted within ΔT_{max} (e.g., 2 s), the moving state is determined to be "moving"; otherwise, the state is determined to be "stationary" if a peak is not detected during ΔT_{max}. When there is no change in the measurements (e.g., the standard deviation of measurement is less than 0.05 g) during ΔT_{max}, the system considers that the smartwatch has been taken off, and the moving state is considered "unknown." Activity monitoring with the smartwatch starts when the moving state changes from "moving" to "stationary."

Activity Feature Extraction

This section describes the details of feature extraction from a smartwatch. Activity features consist of two types of information: kinetic features and acoustic features. The kinetic features represent the kinetic arm's movement during the activity, and the acoustic features represent the ambient sound in the location as well as the sound that occurred with the activity.

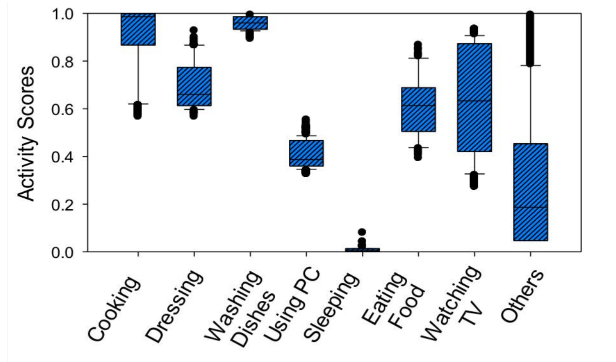


Figure 4. Activity scores according to home activities.

Kinetic Feature Extraction

Measurements from the accelerometer capture the motion characteristics of postures and repetitive arm movement. We used a sliding window of about 5 s to monitor the activities, and the measurement is collected at a rate of more than 50 Hz. The window is 50% overlapped with previous sliding windows, which have shown success in previous research [29]. Once the sliding window is filled with measurements, the following kinetic features are extracted.

Activity Score (AS). We estimate the activity level of the occupants by monitoring how frequently they use their arms during a period (e.g., recent 10 min). The activity level represents whether the occupants are in a static state, such as sitting on a chair and sleeping, or in a dynamic state that involves complex movements (e.g., cooking). The activity score is also a key feature to distinguish sleeping activities from other static activities, as illustrated in Figure 4. AS is estimated as follows:

$$Activity\ score = \frac{T_{active}}{T_{active} + T_{idle}}.$$

Mean and Standard Deviation Vector. Three-dimensional mean and standard deviation vectors are calculated from the acceleration signals in the window (5-s windows that contain 512 frames in our implementation). With N acceleration measurement samples, the mean and standard deviation vectors are represented as follows:

$$M_t(a) = \frac{1}{N}\sum_{j=t-N-1}^{t} a_i(j)\,, (i = x, y, and\ z),$$

$$S_t(a) = \left[\frac{1}{N-1}\sum_{j=t-N-1}^{t}(a_i(j) - a_i)\right]^{1/2}.$$

Tilt Angle. The tilt angle is defined as the angle between the positive z-axis and the gravitational vector g. The tilt angle is used to differentiate among certain postures.

Correlation Vector. The vector that consists of the correlation between each pair of axis as follows:

$$Cor\ (t) = \left[r_{xy}(t), r_{yz}(t), r_{zx}(t)\right],$$

where $r_{xy}(t) = \frac{1}{N-1}\sum_{t=1}^{N}\frac{a_i(t)-a_i}{s_i}\frac{a_j(t)-a_j}{s_j}$.

Power Spectrum Density of the Frequency Band (PSD). PSD is a frequency-domain plot of the power per Hertz versus the frequency. The PSD feature represents repetitive movements in the frequency domain. We divide the PSD into four frequency sections such as 0–1, 1–2, 2–4, and 4–8 Hz.

Energy. To capture short-term energetic movements, the energy feature is calculated as shown in [34]. The energy is the sum of the squared discrete fast Fourier transform (FFT) component magnitudes of the signal. The sum is divided by the window length for normalization. The energy is calculated as follows:

$$Energy(w) = \frac{\sum_{i=1}^{i=n} x_i^2}{|w|},$$

where x_i is the FFT component within the window w_k.

Acoustic Feature Extraction

Our system performs acoustic recognition using the microphone embedded in the smartwatch. The acoustic sound of an activity or a unique ambient sound in a space (e.g., operating noise of appliances) can be used to infer home activities. For example, we notice that someone is taking a shower by the sound of water. We record the sound when a user is in the stationary state for 3 s every 180 s using the smartwatch's microphone. Then, the root mean square (RMS) and the Mel-frequency cepstral coefficients (MFCCs) [35], which are commonly used for speech sound signal processing, are extracted as acoustic features. The *mel* is a basic unit of measure based on the perceived frequency of human ears. The MFCC is based on a linear cosine transform of a log-scale power spectrum on a Mel-frequency filter bank. We used 25-Mel filter banks and calculated the first 12 coefficients for acoustic features.

Clustering and Generation of Activity Fingerprints

We generate activity fingerprints based on the extracted kinetic features and acoustic features. The challenge is to automatically generate annotation on learning data without user participation. To address this challenge, we clustered the extracted feature vectors in order to categorize sensing data according to human activity. We first reduced the dimension of vectors by using linear discriminant analysis (LDA) [36] dimensionality reduction since the extracted features consist of high-dimensional vectors (37 dimensions for kinetic features and 12 dimensions for acoustic features). The *k*-means clustering algorithm [37] is then used to categorize the feature vectors based on similarity without manual labeling. In the *k*-means clustering algorithm, the vectors are divided into exclusive groups of *k,* and *k* seed vectors are initially selected from the features. We define the clustered result of the feature vectors as *activity fingerprints*. An important factor in the *k*-means algorithm is that the results depend on the *k* value, which should be known *a priori*. We empirically determined the *k* value that can be adaptively decided according to the home environment.

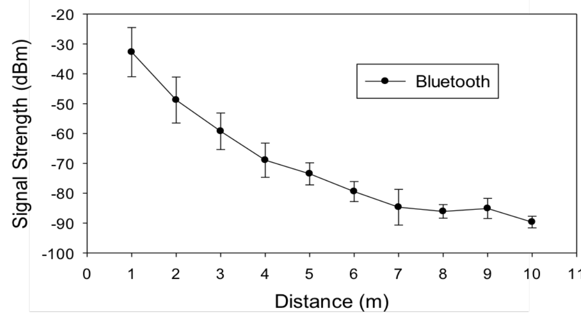


Figure 5. Bluetooth RSS according to distance (line-of-sight).

Acquisition of Room Location

The Wi-Fi-based positioning system (WPS) is widely used for indoor tracking because it provides an accurate position (e.g., at the room level) of a mobile device without additional infrastructure. However, wearable devices such as a smartwatch do not have Wi-Fi capability owing to the limited battery capacity and compact form factor. We therefore find the location of occupants using their smartphones.

Wi-Fi-based Positioning System Using a Smartphone

Since we assume that each occupant has a personal smartphone, we use a Wi-Fi-Based Positioning System (WPS) to obtain the location of the occupants while they are carrying their smartphones. If the occupant stays in a certain space at home, the smartphone measures the RSS from the surrounding Wi-Fi APs. The system then determines that an occupant is located in place *a* at time *t* if $s(f_t, f^a) \leq \rho$, where the similarity function $s(\cdot)$ is based on the Tanimoto coefficient [38], f_t is the RSS vector measured at time *t* by the mobile device, f^a is a Wi-Fi fingerprint trained at place *a* in advance, and ρ is a given threshold.

Conditional Use of Location Information

A smartphone periodically collects RSS vectors (e.g., every 5 min) to obtain room information. However, in a home environment, it is not guaranteed that the location of a mobile device is always the same as the location of the device owner. To filter out the case in which a user is in a different indoor space from where the smartphone is placed, we use the signal strength of the BT communication between the smartphone and the smartwatch. Figure 5 shows the signal strength attenuation with the distance for BT communication in an indoor space. The attenuation rate of the BT signal is much higher than that of the Wi-Fi signal, and the signal variation is small because the BT transmission range is short (10–15 m). We assume that the smartphone and the smartwatch are in the same space when the signal strength during the data exchange via BT is larger than –72 dBm, which indicates a distance closer than 4.5 m. In our experiments, false positive cases (the device owner and the smartphone are co-located in the same space, but the strength of the BT signal is lower than –72 dBm)

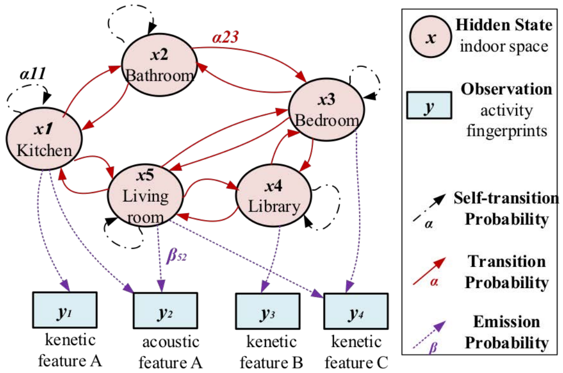


Figure 6. Graphical representation of the HMM for the home locating system.

frequently occurred because of non-line-of-sight situations. However, false negative cases (the BT signal is strong when the owner's location does not coincide with location of the device) were rare. Therefore, we use the location of a smartphone obtained from the WPS to estimate the owner's location only when the BT signal is stronger than the threshold value (–72 dBm).

INFERENCE OF ROOM LOCATION USING ACTIVITY FINGERPRINTS

This section presents the inference on the occupant's room location. The scheme comprises two phases: the training phase and the estimating phase. In the training phase, the relationship between the activity and the indoor space is learned when an occupant is co-located with a smartphone. When an occupant is away from a mobile device owing to device-free movement, the estimation phase infers the occupant's location using activity fingerprints, since the location cannot be directly determined using the WPS with a mobile device.

Training Samples for Learning

The system requires a training process to link the activity fingerprint with indoor space. In conventional work on machine learning and recognition systems, users manually annotate the type of activity that they have performed for the training [17, 29, 30, 31]. However, our system does not require manual annotation for the training since the system focuses on the occupants' location, not on the type of activity. When a user is co-located with a smartphone, the system uses the estimated location obtained from the WPS to generate training samples for the inference model.

The Hidden Markov Model

We employ the HMM to model user mobility at home. An HMM is a generative probabilistic model with hidden states x and observations y. The states x_i emits an observation with a conditional probability distribution known as emission probability α. Transitions between hidden states are governed by a set of probabilities called transition probability β. Figure 6 shows the graphical representation of our locating system based on activity fingerprints. A set of hidden state x and observations y denote the room location of the user and activity fingerprints, respectively. In the HMM, since the states are hidden and observations are known, the main goal of the HMM is to determine the hidden state sequence $(x_1, x_2, \dots, x_t)$ that corresponds to the observation sequence $(y_1, y_2, \dots, y_t)$. In our cases, with the sequence of observed activity fingerprints, the system estimates a set of room locations the user previously visited. To construct the HMM for the locating system, the transition and the emission probability distribution should be estimated with training samples.

Transition Probability

Transition probability indicates the possibility of the transition from state x_i to state x_j, represented as α_{ij}. To estimate α_{ij}, we use the historically-observed movements through the Wi-Fi-based positioning system. The transition probability from α_{ij} room location r_i to the room r_j is calculated as

$$\alpha_{ij} = p\left(x_t = r_j \middle| x_{t-1} = r_i \right) = \frac{T(i,j)}{\sum_{k=1}^{N} T(i,k)},$$

where $T(i,j)$ is the total number of transitions from x_i to x_j in the training samples and N is the number of rooms.

Emission Probability

The emission probability β_{ij} for a given location r_i and the corresponding activity fingerprints af_j represent the probability of observing that activity fingerprint af_j conditioned on the user being in room r_i. The emission probability β_{ij} is calculated as

$$\beta_{ij} = p\left(y_t = af_j \middle| x_t = r_i \right) = \frac{E(i, af_j)}{\sum_{k=1}^{M} E(i, af_k)},$$

where $\mathrm{E}(i, a_k)$ is the total number for emitting af_k of at state x_i and M is the number of clusters of feature vectors.

Estimate User's Location

Completing the HMM training, the user location is obtained in the estimation phase when the system detects that the user is not co-located with his or her smartphone. We use Viterbi decoding [39] to estimate the most-likely hidden states in the HMM from the observed activity fingerprints.

Viterbi Decoding

The Viterbi decoding is a dynamic programming-based algorithm for selecting the best state sequence that maximizes the likelihood of the state sequence from the given observation sequence. Given an HMM with the state space X and the initial probability π, the probability of the most-likely state x_k at time t is calculated by the recurrence relations, expressed as

$$V_{k,}(t) = \beta_{kt} \cdot \max\left(\alpha_{uk}(t) \cdot V_u(t-1)\right),$$
$$V_k(0) = \beta_{k0} \cdot \pi$$

where α_{uk} is the transition probability from state x_u to state x_k, β_{kt} is the emission probability of activity fingerprint af_t observed at time t on room location r_k.

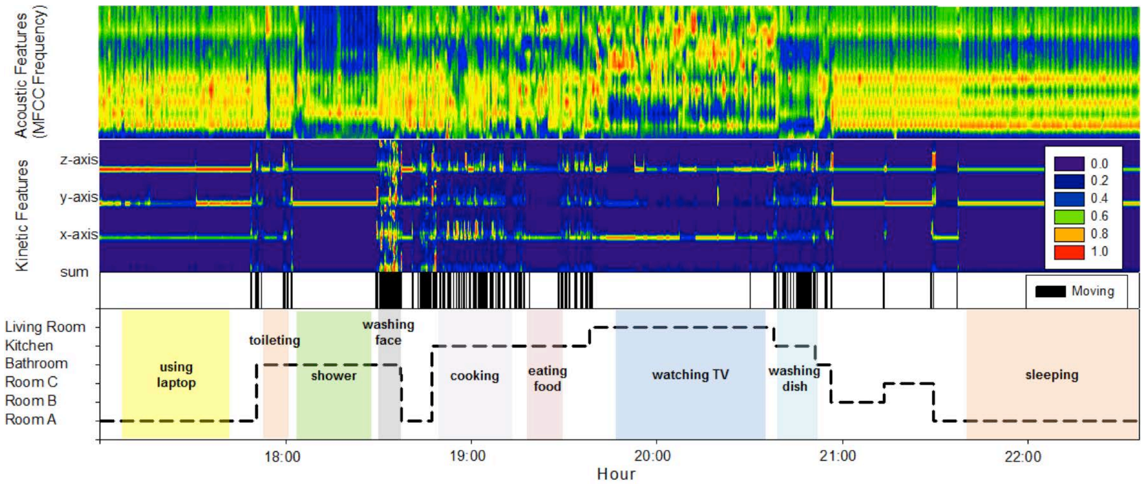


Figure 7. Trace of daily activity at home.

EVALUATION

In this section, we evaluate the activity fingerprint–based locating system. Our primary results show that a user's activities at home can indicate the user's location without the use or need for additional infrastructure.

Implementation and Data Collection

We implemented the proposed system with an Android smartphone and a smartwatch, the Samsung Galaxy Gear [21]. In particular, the Gear has various sensors, such as a three-axis accelerometer, a gyroscope, a microphone, a camera, and a Bluetooth module. The smartphone performs a Wi-Fi scan periodically (every 5 min) to infer the location of occupants using the Wi-Fi fingerprint3ing technique, while the smartwatch continuously monitors its various sensors to collect the activity fingerprints.

To estimate the feasibility of the activity fingerprints–based locating system, we first collected experimental data in the testbed. The experiment was conducted with three participants for 3 days. We trained activities according to room location and tested the system in a personalized way. To collect ground truth, training was done individually in accordance with a given schedule. For the experiment in the wild, we monitored four occupants for 3 to 5 weeks in their home environments. To collect the ground truth on location, we asked the occupants to annotate their locations and current activities in their smartwatches by sending message requests whenever the occupants' status changed from "in motion" to "stationary" or whenever the duty cycle time expired (every 15 min).

Activity Fingerprints of Daily Log

Figure 7 shows the tracing of the kinetic and acoustic features and the actual activities and locations of an occupant at home. We visualized the acoustic and kinetic features by normalizing each component with a range from zero to one.

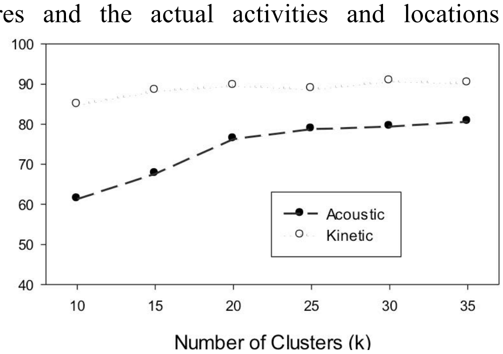


Figure 8. Accuracy of prediction according to the

The occupants performed various activities while moving from room to room, and we observed the unique features of each activity. Some activities, such as using a laptop, were distinguishable from other activities by kinetic features, but they did not show a clear difference from other activities in their acoustic features. In contrast, several activities showed a clear difference in acoustic features. In the case of taking a shower, for instance, kinetic features were not useful, since the smartwatch is taken off during the shower; the acoustic feature helps, however, as it senses the sound of water. Consequently, acoustic features can complement a kinetic feature for recognizing activity in the home environment.

Parameter Evaluation with Manually Annotated Data

To evaluate the efficiency of activity fingerprinting, we first selected the parameter (input k in the clustering algorithm) of our system and then evaluated the performance of the proposed system using the HMM inference model constructed with ground truth information.

Figure 8 shows the influence of the number of clusters that are used to generate the activity fingerprints. In general, a large number of clusters increases the accuracy of estimating locations, because finely divided activity fingerprints can better describe human activities. When the number of clusters was larger than 20, however, no significant differences in accuracy were observed. We thus determined the ideal number of cluster value k is 25, because 25 kinds of activity fingerprints are enough to represent human activities at home.

As shown in Figure 9**,** the precision of using only a single feature was 87% with kinetic features and 76% with acoustic features. When the two features were used together, the precision and the recall increased to more than 90%. The precision as well as the recall were higher than 90% in

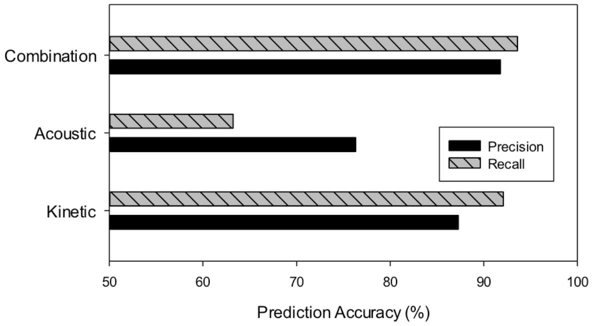


Figure 9. Accuracy of location prediction according to feature extraction.

all places except for the bathroom and the kitchen, where the precision was 83.9% and 85.0%, respectively.

Inference of Location Without Manual Annotation

Wi-Fi Fingerprint-Based Positioning

Our system automatically labels the locations of activity fingerprints using the Wi-Fi fingerprinting technique. We therefore evaluated the accuracy of the Wi-Fi fingerprinting system in a home environment. An occupant collected three RSS measurements in each room to construct a radio DB. We then estimated the location of the occupants with the RSS measurements collected during the experiment. Figure 10a shows the similarity of the normalized RSS vectors collected in Home A. In most cases, the RSS measurements in one room seem to be distinct from the measurements in other rooms. Figure 10b shows the accuracy of the Wi-Fi-based positioning system according to the room location. In our experiments, the Wi-Fi-based location system yielded 91% (Home A) and 88% (Home B) accuracy, respectively.

Impact of Training Periods

We further evaluated the influence of training periods on predicting location in order to analyze the required time and the number of training samples for predicting location accurately. Figure 11 shows the relation of the training period and prediction accuracy according to individuals. With training samples collected over 5 days, the system accurately inferred the user's location. Especially in the cases of users B and D, the use of training samples that were collected for only 2 days was enough to learn their home activities and corresponding locations (higher than 80% accuracy) at home. This is because users B and D are students, and they have very simple activities and mobility patterns at their home, as shown in Figure 1.

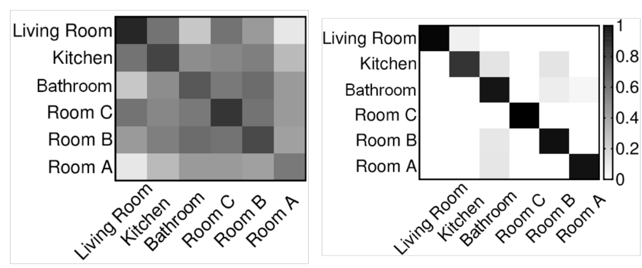
Comparison with Support Vector Machine–Based Classification

Finally, we evaluated the performance of our system with training samples from automatic annotation using the WPS. To show the superiority of the proposed system, we compared the performance of the proposed system with a well-known classification method. As a comparison model, we used the support vector machine (SVM) method, which is a discriminative classifier; the method is widely used for activity recognition and audio classification [29,30]. Since two kinds of feature vectors exist in our system, we used two SVM-based classifiers, for kinetic features and for acoustic features. Each classifier independently classifies the occupant location with the feature vectors. We estimated the location by maximizing the likelihood based on the posterior probability of two classifiers as follows:

$$L(t) \approx \arg_{r_i} max\ p(L = r_i|C_k) \cdot p(L = r_i|C_a)$$

		Prediction						**Recall (%)**
		Main Room	Room A	Room B	Kitchen	Living Room	Bathroom	
Indoor Spaces	Room 1	8964	4	17	140	92	144	95.8
	Room 2	5	6642	3	99	192	161	93.5
	Room 3	15	3	2545	3	1	7	98.9
	Kitchen	122	161	0	3792	260	141	84.7
	Living Room	219	440	43	209	14777	20	94.1
	Bathroom	517	62	2	165	358	3170	74.2
Precision (%)		91.1	90.8	97.5	86.0	94.2	87.0	

Table 1. The confusion matrix of the location prediction with cross validation



(a) RSS similarity (b) Prediction accuracy

Figure 10. Wi-Fi-based positioning system (Home A).

where r_i is the room location and C_k and C_a are the SVM classifiers for the kinetic features and the acoustic features, respectively.

As shown in Figure 12, the proposed system outperformed the SVM-based classification with any sensor selection. In SVM-based classification, similar to the results to the proposed system, using a single kinetic feature outperformed the acoustic-only features, because acoustic features are prone to corruption by ambient noise or other sounds generated by external environmental factors. For example, if a user eats a meal while other occupants are watching TV in the living room, the acoustic features of an "eating meal" activity are combined with the TV sound generated by other occupants. In addition, compared with the results after training from manually annotated samples in Figure 9, there was a loss of 6% of prediction accuracy because of the increased ambiguity of the training samples. The overall accuracy of locating occupants was still around 87%, however. These results indicate that the proposed

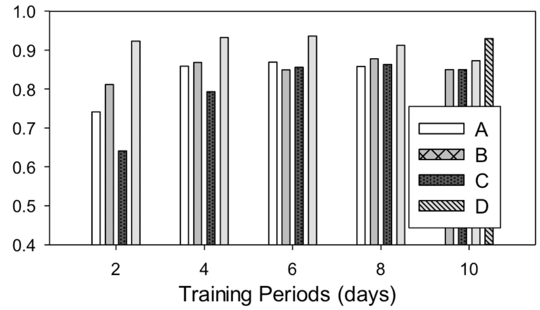


Figure 11. Impact of the collecting period for training samples.

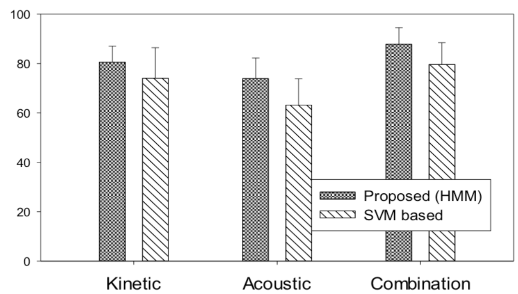


Figure 12. The prediction accuracy of the proposed system without manual annotation.

system provides an accurate home locating system without the need for manual annotation from users.

Impact on Battery Lifetime

A smartwatch or a wristband has limited battery capacity owing to its small form. As our application is based on measurements and calculations that use several hardware components, the application obviously shortens the battery life of these devices. In our experiments, the running time of our application with the Galaxy Gear was around 5.3 h. To extend the battery life of the smartwatch and reduce delays in recognition, we applied adaptive duty cycling to generate activity fingerprints. We noted that some home activities, such as sleeping, resting, and watching TV, are static and account for a significant portion of time spent at home. Therefore, unless user movements or changes in activities are detected, duty cycling for generating activity fingerprints is gradually incremented (e.g., duty cycling two times). If user movements or changes in activities are detected, the duty cycling is reset to the initial duty value (e.g., 180 seconds for acoustic sampling). By using simple adaptive duty cycling, the running time can be extended to 8.7 h. This result indicates that optimized and adaptive sampling of sensors and BT communication can further extend the lifetime of smartwatches.

DISCUSSION

In this section, we describe potential limitations of our system, along with future research directions.

Distinguishability of Similar Activities

Several smartwatches and wristbands with various types of sensors are now commercially available. Some smartwatches [22,23] are equipped with a three-axis magnetometer sensor but no microphone sensor. To analyze the influence of the direction (or orientation) feature from magnetometers on activity recognition, we performed additional experiments to extract a direction feature for home activities. To collect direction measurements from the magnetometer sensor, we employed the magnetometer-enabled LG G-watch [23], which is implemented on the Android Wear platform. Next, we compared the kinetic features from the magnetometer for home activities that have similar kinetic features as measured by accelerometers. By combining direction features with a magnetometer sensor, similar (or even the same) activities in different locations can be distinguished, because the activities rely on the bearing the deployed furniture and home appliances related to the activities.

Training Costs and Training Periods

Since our system is to be used during the course of normal daily life, we have made an effort to minimize the training cost, which is highly related to user annoyance. However, the one thing required *a priori* is a Wi-Fi radio database, since the smartphone has to compare current fingerprints with the fingerprints of each room to automatically generate training samples. In general, constructing a Wi-Fi map is laborious, but that was not the case in our work. Note that a house consists of a small amount of indoor space, and unlike the common WiFi fingerprinting system for indoor localization, our system focused on distinguishing indoor space only instead of coordination of the user's location. Therefore, the training phase can be simplified. For example, our system collects Wi-Fi fingerprints only at the center or staying points of every room because the system uses the Wi-Fi positioning system to distinguish room space, not to estimate an exact location in the room. This simplifies the radio map construction. In our experiments, the training phase was completed within a short time, i.e., within 10 min, by scanning the Wi-Fi three to five times at the center or stay-point of each room and by annotating the space.

CONCLUSION

An occupant's activities at home depend on the characteristics of the indoor space. Based on this insight, we proposed an occupant locating system that tracks an individual occupant in a home environment. By using an HMM with activity fingerprints, our system estimates the location of each occupant without requiring privacy-invasive sensors, burdensome wearable sensors, or high-cost infrastructure. Our evaluation demonstrates that the proposed system successfully estimates an occupant's location without requiring user participation. The proposed system overcomes several constraints of conventional indoor tracking systems, such as the laborious training phase of supervised learning, the need to carry a device at all times, or the need for installation of additional sensors at target locations. We therefore believe that our approach can be applied to location-based home applications such as in-home activity monitoring and occupant-driven home energy management.

ACKNOWLEDGMENTS
This work was supported by Samsung Electronics Co., the National Research Foundation of Korea grant funded by the Korean government, Ministry of Education, Science and Technology (No. NRF-2012R1A1A3006105), and the Energy Efficiency & Resources, Korea Institute of Energy Technology Evaluation and Planning, Korea (No. 2013T100100019).

REFERENCES
1. J. Rekimoto, T. Miyaki, and T. Ishizawa. 2007. Lifetag: Wi-Fi-based continuous location logging for life pattern analysis. In *Proceedings of the 3rd International Conference on Location- and Context-Awareness*, 35-49.
2. I. Anderson, J. Maitland, S. Sherwood, L. Barkhuus, M. Chalmers, M. Hall, B. Brown, and H. Muller. 2007. Shakra: Tracking and sharing daily activity levels with unaugmented mobile phones. *Mobile Netw Appl* 12, 2-3: 185-199.
3. A. R. Kaushik and B. G. Celler. 2007. Characterization of PIR detector for monitoring occupancy patterns and functional health status of elderly people living alone at home. *Technol Health Care* 15, 4: 273-288.
4. P. Sánchez, M. Jiménez, F. Rosique, B. Álvarez, and A. Iborra. 2011. A framework for developing home automation systems: from requirements to code. *J Syst Software* 84, 6: 1008-1021.
5. S. Lee, D. Ahn, S. Lee, R. Ha, and H. Cha. 2014. Personalized energy auditor: Estimating personal electricity usage. In *The 11th IEEE International Conference on Pervasive Computing and Communications* (PerCom 2014), 44-49.
6. V. L. Erickson, M. A. Carreira-Perpinan, and A. E. Cerpa. 2011. OBSERVE: Occupancy-based system for efficient reduction of HVAC energy. In *Proceedings IPSN 2011*, 258-269.
7. A. Kamthe, L. Jiang, M. Dudys, and A. Cerpa. 2009. Scopes: Smart cameras object position estimation system. *Wireless Sensor Networks* 279-295.
8. J. Ranjan, Y. Yao, and K. Whitehouse. 2013. An RF doormat for tracking people's room locations. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 797-800.
9. T. W. Hnat, E. Griffiths, R. Dawson, and K. Whitehouse. 2012. Doorjamb: Unobtrusive room-level tracking of people in homes using doorway sensors. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, 309-322.
10. J. Wilson and N. Patwari. 2011. See-through walls: Motion tracking using variance-based radio tomography networks. *IEEE Trans Mobile Comput* 10, 5: 612-621.
11. E. K. Choe, S. Consolvo, J. Jung, B. Harrison, and J. A. Kientz. 2011. Living in a glass house: A survey of private moments in the home. In *Proceedings of the 13th International Conference on Ubiquitous Computing*, 41-44.
12. B. Song, H. Choi, and H. S. Lee. 2008. Surveillance tracking system using passive infrared motion sensors in wireless sensor network. In *Proceedings of International Conference on Information Networking* (ICOIN 2008), 1-5.
13. U. Steinhoff and B. Schiele. 2009. Dead reckoning from the pocket—an experimental study. In *Proceedings of IEEE International Conference on Pervasive Computing and Communications* (PERCOM 2009), 162-170.
14. Y. Kim, Y. Chon, and H. Cha. 2012. Smartphone-based collaborative and autonomous radio fingerprinting. *IEEE Trans Syst Man Cy C* 42, 1: 112-122.
15. M. Hardegger, et al. 2012. ActionSLAM: Using location-related actions as landmarks in pedestrian SLAM. In *Proceedings of International Conference on Indoor Positioning and Indoor Navigation* (IPIN 2012), 1-10.
16. M. Hardegger, G. Tröster, D. Roggen. Improved ActionSLAM for long-term indoor tracking with wearable motion sensors. 2013. In *Proceedings of the 17th Annual International Symposium on Wearable Computers*, 1-8.
17. O. Woodman and R. Harle. 2008. Pedestrian localisation for indoor environments. In *Proceedings of the 10th ACM international conference on Ubiquitous computing* (Ubicomp 2008), 114-123.
18. T. W. Hnat, V. Srinivasan, J. Lu, T. I. Sookoor, R. Dawson, J. Stankovic, and K. Whitehouse. 2011. The hitchhiker's guide to successful residential sensing deployments. In *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems* (Sensys '11), 232-245.
19. E. Foxlin. 2005. Pedestrian tracking with shoe-mounted inertial sensors. *IEEE Comput Graph* 25, 6: 38-46.
20. Fitbit, Inc. Fitbit Force. Retrieved July 14, 2015 from http://www.fitbit.com/
21. Samsung, Inc. Samsung Galaxy Gear. Retrieved July 14, 2015 from http://www.samsung.com/nl/consumer/mobile-phone/wearable/galaxy-gear/SM-V7000ZKAPHN
22. Sony, Inc. Sony Smartwatch2. Retrieved July 14, 2015 from http://www.sonymobile.com/global-en/products/accessories/smartwatch-2-sw2/

23. LG, Inc. LG G-watch. Retrieved July 14, 2015 from http://www.lg.com/global/gwatch/index.html#main
24. P. Bahl and V. N. Padmanabhan. 2000. RADAR: an in-building RF-based user location and tracking system. In *Proceedings INFOCOM 2000*, 775-784.
25. S. Nattapong and V. K. Prashant. 2008. On clustering RSS fingerprints for improving scalability of performance prediction of indoor positioning systems. In *Proceedings of the First ACM International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments*, September 19-19, 2008, San Francisco, California, USA.
26. A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J. Tabert, P. Powledge, G. Borriello, and B. Schilit. 2005. Place Lab: Device positioning using radio beacons in the wild. In *Proceedings of the third International Conference on Pervasive Computing*, 116-133.
27. A. Anshul Rai, K. K. Chintalapudi, P. Venkat, and R. Sen. 2012. Zee: Zero-effort crowdsourcing for indoor localization. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking* (MobiCom), 369-378.
28. Martin Azizyan, Ionut Constandache, and Romit Roy Choudhury. 2009. SurroundSense: Mobile phone localization via ambience fingerprinting. In *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking*, 261-272.
29. L. Bao and S. Intille. 2004. Activity recognition from user-annotated acceleration data. In *Proceedings of Pervasive Computing*, 1-17.
30. J. Lester, T. Choudhury, and G. Borriello. 2006. A practical approach to recognizing physical activities. In *Proceedings of Pervasive Computing*, 1-16.
31. K. Yatani and K. N. Truong. 2012. BodyScope: a wearable acoustic sensor for activity recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (Ubicomp 2012), 341-350.
32. S. Lee, B. Kim, H. Kim, R. Ha, and H. Cha. 2011. Inertial sensor-based indoor pedestrian localization with minimum IEEE 802.15.4a configuration. *IEEE Trans Industrial Informatics* 7, 3: 455-466.
33. M. Mladenov and M. Mock. 2009. A step counter service for Java-enabled devices using a built-in accelerometer. In *Proceedings of the International Workshop on Context-Aware Middleware and Services* (CAMS 2009), 1-5.
34. N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman. 2005. Activity recognition from accelerometer data. *AAAI* 5: 1541-1546.
35. Z. Fang, et al. 2001. Comparison of different implementations of MFCC. *J Comput Sci Technol* 16, 6: 582-589.
36. H. Abdi. 2007. Discriminant correspondence analysis. In *Encyclopedia of Measurement and Statistics*. Sage, Thousand Oaks, CA, 270-275.
37. J. B. MacQueen. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of Berkeley Symposium on Mathematical Statistics and Probability*, 1, 281-297.
38. D. Rogers and T. Tanimoto. 1960. A computer program for classifying plants. *Science* 21: 1115-1118.
39. A. Viterbi. 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans Inf Theory* 13, 2: 260–269.