Proximity-Aware Location Based Collaborative Sensing for Energy-Efficient Mobile Devices

Jeongho Kwak, *Member, IEEE*, Jihwan Kim, *Member, IEEE*, and Song Chong, *Member, IEEE*

Abstract—A fundamental question in the study of location-based mobile sensing is how much energy can be saved while still guaranteeing the reliable localization accuracy. In this paper, we analyze key features of human proximity and find a motivation which implies that the energy-efficient and accurate localization is possible by sharing the locations of nearby mobile devices. From the location-based collaborative sensing idea, we formulate an optimization problem which aims to minimize the total number of location measurements for a given fairness criterion. Then, we propose a practical and distributed location sharing (DLS) protocol and an optimal parameter control algorithm (OWD) which makes the DLS protocol attain an asymptotic optimal performance. Via extensive simulations under various environments including real mobility traces, we verify that the proposed DLS+OWD policy significantly reduces the average power consumption of mobile devices with a higher fairness compared to the existing algorithms.

Index Terms—Distributed networks, location based collaborative sensing, short-range communications, distributed applications, algorithm/protocol design and analysis, mobile environments

1 Introduction

1.1 Motivation

With the advent of the Internet of Things (IoT) era, future mobile devices such as smartphones/pads are likely to gather more sensing data which involves location information. As an example, mobile crowd sensing (MCS), which collects sensing values including locations (e.g., CO_2, temperature, humidity, dust/ultraviolet, and brightness) from massive mobile devices, has been widely prevalent so that more mobile devices can exploit the collected information[2]. For example, in the collaborative street parking system [3], the mobile devices check vacant parking lots and inform other vehicles of them with location information. If many mobile devices recognize the vacant parking lot at the same location, the vacant sensing accuracy would increase. However, a lot of measurements of the same location values at the same positions would result in the high energy consumption of mobile devices [4], [5]. For instance, our surveys and measurements in Section 8 demonstrate that most of the recent smartphones consume a significant amount of GPS energy to estimate their locations with high accuracy on the outside. As an example, LG Nexus 4 smartphone consumes 69 percent more energy in case of GPS-on mode than that in case of standby mode [4].

- *J. Kwak is with the Department of Computer Science & Statistics, Trinity College Dublin, Dublin 2, Ireland. E-mail: jeongho.kwak@tcd.ie.*
- *J. Kim is with the Samsung Electronics, Suwon, Gyeongi 443-803, Korea. E-mail: kimji.netsys@gmail.com.*
- *S. Chong is with the School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Korea. E-mail: songchong@kaist.edu.*

Manuscript received 2 June 2016; revised 18 Nov. 2017; accepted 28 Apr. 2018. Date of publication 8 May 2018; date of current version 7 Jan. 2019. (Corresponding author: Jihwan Kim.)
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMC.2018.2833842

To reduce the localization energy, there have been several studies [5], [6], [7] which have used substitute information from WiFi/cellular infrastructures or adscititious sensors to consume a smaller amount of energy for obtaining the locations of sensing data. However, some of the energy-efficient localization schemes such as WPS (WiFi-based positioning system from SkyHookWireless [8]), or GSM-based positioning system sacrifice the localization. For example, the accuracy of the GPS sensor, WPS, and GSP are 8.8, 32.44 and 176.7 m in open sky environments, respectively [5]. Because many applications including the MCS (e.g., [2], [9], [10]) frequently require continuous outdoor location values while guaranteeing the high degree of accuracy, the localization solutions which satisfy both of localization accuracy and energy efficiency are required.

Short-range communication (SRC) such as classic Bluetooth, Zigbee or Bluetooth Low Energy (BLE) [11], [12] has been used to transmit/receive a small amount of data with low energy consumption. For instance, scanning power of classic Bluetooth is known as 96-192 mW [13], [14], which is much lower than that of GPS.[1] Moreover, because sleep mode power consumption is extremely low and the actual transition time from sleep mode to scanning mode is much shorter in BLE than in classic Bluetooth [15], [16], the average power consumption of the BLE can be much lower than the classic Bluetooth technology. Despite the excellent energy efficiency, the drawback of the SRC for communication is that it cannot be used for long-range communication. However, *we are inspired anyway by the "short range" of the SRC which can be an advantage to share the outdoor location values with high accuracy if the SRC is used to find proxy devices and broadcast the location values.*

1. Our measurements and surveys in Section 8 verify that the average GPS power consumption in case of continuous measuring is 560-1006.61 mW. However, note that this measurement can be different depending on the specification of GPS module.

In order to verify how many devices exist within an SRC coverage to share information among them, we first analyze the spatio-temporal human proximity in actual scenarios. Because the human beings are sociable in nature, they spend much time with other people in general. For example, students taking the same class are likely to be closely located on the same campus during a day. In addition, irrelevant people also happen to be closely located when they are at the crowded places such as downtown areas. A recent human mobility study (see [17] and references therein) reported that the movement of humans reflects their social proximity to other people.

Technically, we define two human proximity natures: (i) temporal proximity, i.e., how long a person stays with other people and (ii) spatial proximity, i.e., how close a person is to other people. According to the American time use survey in [18], a person enjoys a social relationship with acquaintances for 8.5 hours on average. Moreover, an analysis of the MIT reality mining data [18] shows that for 47 percent of people, once they meet other people, they enjoy the meeting for over a single hour. More specifically, we statistically analyze the proximity of humans based on real GPS trace data of students in the Korea Advanced Institute of Science and Technology (KAIST), Korea and North Carolina State University (NCSU), USA, campuses and observe two proximity natures in Section 2. In a perspective of temporal proximity, the analysis demonstrates that people stay within 10 m coverage during 6 min 40 sec in average. In a perspective of spatial proximity, we observe that approximately 5-10 people are usually located within a 10 m coverage which is the average Bluetooth range [18].

1.2 Main Idea & Contributions

The main idea of this paper is to share location values with proxy devices using short-range communication consuming a low amount of energy.[2] For example, let us assume that within a few meters, there are five devices which require to consecutively sense their CO_2 or temperature with own locations in order to send them to an MCS server. If one of the devices measures its outdoor location using GPS and share it with other four devices using SRC, then five devices obtain their locations by consuming low SRC energy of the five devices in addition to high GPS energy consumed by only one device while guaranteeing the localization accuracy of the GPS plus SRC range.[3]

However, there are some concerns about this idea as follows: *(i) who measures the location among proxy devices, and (ii) how long should the device measure the location?* If all devices obtain the location values with the minimum number of devices which directly measure the locations, they might attain the highest energy efficiency (if all devices consume the same energy to measure the locations). On the other hand, if the devices measure the location values during the same time portion, they probably share the location values with fair contributions to measure the locations.

Because smartphone users are fundamentally selfish, they might be reluctant to spend entire energy of their devices on localization. Therefore, the consideration of fairness among devices is equally important with the energy efficiency of them. However, both of the lowest numbers of location measurements which lead to total energy saving and complete fairness among all devices cannot be simultaneously satisfied. Hence, the answer to the aforementioned two questions is to strike a balance between a minimization of the number of location measurements and a fairness among all devices. Therefore, we formulate a long-term problem which aims to reduce the number of location measurements while considering the fairness among all devices. Note that our problem only targets the outdoor location measurement using GPS due to the fact that the measurement accuracy of the GPS drastically decreases in the indoor scenarios.[4]

As a solution of the proposed problem, we should find a sequence of localizing-on device sets so that the average localizing-on/off time portion of each device asymptotically approaches to an optimal solution where *"localizing-on device"* denotes a device which directly measures its location and broadcasts it to neighbor devices and *"localizing-off device"* denotes a device which does not directly measure the location and exploits the location values received from the localizing-on device. The notion of optimality was similarly considered in the concept of clustering in sensor networks [20], [21], [22]. However, the clustering cannot directly be applied to collaborative localization because of their different objectives (e.g., maximization of network lifetime) with our problem. In addition, our problem requires a global knowledge of the sets of proxy devices (i.e., a centralized solver is needed), and has to solve an NP-hard problem in the form of a maximum weighted independent set selection to obtain an optimal solution.

To resolve the above two challenges, we first present a distributed location sharing (DLS) protocol that operates by scanning location sharing devices and using a waiting time to perform localizing-off. And then, we provide an optimal mean waiting time decision (OWD) algorithm which makes the DLS protocol achieve an optimal localizing-on/off time portion in each device using only past localizing on/off statistics by invoking an idea of optimal CSMA-CA mechanism [23]. The key mechanism of the DLS protocol is that the proxy devices compete with each other to perform localizing-off with random waiting times for which the average value is updated by the OWD algorithm.

The main contributions of this paper are as follows.

1) We exhaustively investigate the key features of mobile devices (i.e., the proliferation of sensing including location, localization/SRC energy, and the spatial/temporal proximity of people), which provide an advantage for proxy devices to use location-based collaborative sensing applications.
2) We formulate an optimization problem that adjusts a tradeoff between the average number of location measurements using GPS and device fairness and develop a distributed location sharing protocol and an optimal waiting time decision algorithm which makes the DLS protocol achieve an optimality of the problem.

2. Note that a target of this idea is restricted to devices which consume higher localization energy, e.g., GPS than SRC.

3. Note that even if all devices do not want the same sensing value such as CO_2, the location value can be shared by all devices.

4. The indoor localization is out of the scope of our paper and refer to [19].

3) Via extensive simulations with various location topologies and a real mobility trace data at KAIST and NCSU campuses, we make the following key observations. (i) We verify that the DLS protocol and OWD algorithm (DLS+OWD) can obtain a near-optimal performance in terms of energy-efficiency and device fairness by comparing with an optimal centralized algorithm in case of fixed topology. (ii) The DLS+OWD achieves up to 27.2 percent power saving with 35.8 percent higher fairness, and uses 15-22 percent fewer numbers of message passing with proxy devices compared to heuristic algorithms based on clustering in sensor networks at fixed topologies. (iii) The DLS+OWD achieves 65.5 percent power saving compared to no collaboration, and 27.4 percent or more power saving with 25 percent higher fairness compared to the heuristic algorithms with real mobility trace data from KAIST and NCSU. These results show that our scheme adapts well even in unpredictably changing mobility environments.

In the rest of this paper, we begin with an analysis of human proximity using real GPS traces. In Section 3, we explain the technical background and system overview. In Sections 4 and 5, we develop a DLS protocol and an OWD algorithm, respectively. In Section 6, we explain several short-range communication technologies. In Section 7, we discuss privacy and security issues of the proposed scheme. Next, in Section 8, we evaluate proposed scheme under various environments. Finally, we conclude this paper in Section 9.

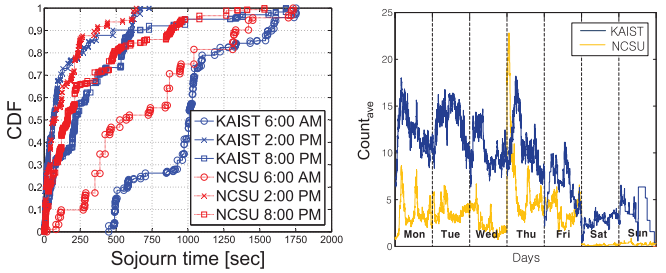
2 Human Proximity Analysis

In this section, we study human proximity characteristics by analyzing the GPS mobility traces of the students in two different campuses.

2.1 Analysis Setup

We had collected the GPS traces of 93 and 99 students with five seconds granularity for seven days in the KAIST and NCSU campuses, respectively. The areas of campuses are $2 \times 2\,\mathrm{km}^2$ in KAIST and $8.5 \times 8.5\,\mathrm{km}^2$ in NCSU, and the total numbers of students and faculties are approximately 10,000 in KAIST and 40,000 in NCSU, respectively. We assume that distributions of the experimental volunteers are the same as distributions of all the people in each campus, respectively. In other words, if we randomly draw a square in the campus, we regard that there exist average $\left(\lfloor \frac{\#\text{ of all people in the campus}}{\#\text{ of experimental volunteers}} \rfloor \times \#\text{ of experimental volunteers in the square}\right)$ number of people in the square. Equivalently, we scale down the distance between two volunteers to the factor of $\sqrt{\frac{\#\text{ of experimental volunteers}}{\#\text{ of all people in the campus}}}$ in this analysis. Self-similarity characteristics in human mobility [17] can support this assumption.

For the spatial and temporal proximity analyses, two metrics are defined, respectively. For the temporal proximity, sojourn time is defined as the continuous retention time of a device within 10m coverage from a certain user. For the spatial proximity, $\mathrm{Count}_{ave}(t)$ is defined as the average number of devices within 10 m coverage from a certain device in time t. In addition, we define an information sharing gain as the number of total devices divided by the number of devices who senses some information and shares it with proxy devices using the SRC. For example, as the information sharing gain increases, the average energy consumption for the information sensing decreases.



(a) CDF of sojourn time. (b) Number of neighbor devices.

Fig. 1. KAIST & NCSU trace analysis.

2.2 Trace Analysis

From GPS mobility trace analysis in KAIST and NCSU campuses, we make two interesting observations as follows.

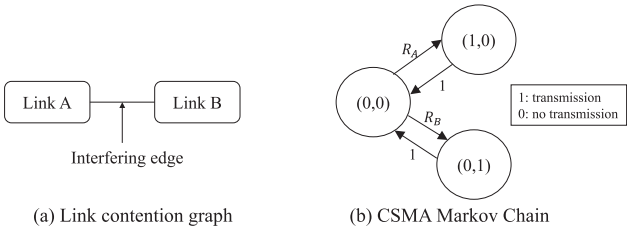
Temporal Proximity. Measured average sojourn times at morning (6:00AM), afternoon (2:00PM) and evening (8:00PM) are (KAIST,NCSU)=(1 min, 1min), (KAIST, NCSU)=(2min 30sec, 3min 30sec) and (KAIST,NCSU)=(16min 30sec, 9min 30sec), respectively. The reason why the sojourn time in the morning is relatively longer than that in the afternoon and evening probably because people more actively get around campus in the afternoon and evening than in the morning. The average total sojourn time is 6 minutes and 40 seconds.

Spatial Proximity. Fig. 1b shows that there exist approximately 10 Count_{ave} in KAIST, and 5 Count_{ave} in NCSU, respectively on weekdays. This implies that average 10 number of devices in KAIST, 5 number of devices in NCSU can collaboratively share some information within 10 m coverage if all devices want to share the information. However, please keep in mind that it does not mean that an information sharing gain is 10 or 5 because the number of information measuring devices varies in accordance with the SRC connectivity and the selection of sharing devices. One of our objectives is to maximize information sharing gain and we will handle this issue in the next section. We guess that the reason why the analysis in NCSU has a fewer number of Count_{ave} than that in KAIST is that the density of people in the NCSU campus ($40000\mathrm{persons}/75.25\mathrm{km}^2 = 531\mathrm{persons}/\mathrm{km}^2$) is lower than KAIST campus ($10000\mathrm{persons}/4\mathrm{km}^2 = 2500\mathrm{persons}/\mathrm{km}^2$).

In summary, we obtain a motivation from the temporal and spatial proximity analyses in both campus scenarios[5] that the average localization energy consumption of devices can be saved by collaboratively sharing the location values with proxy devices using the SRC.[6]

5. Although our analysis is carried out in the limited scenarios, the other environments such as crowded hotspot at downtown or the movement of soldiers with the same mission are expected to have the similar proximity properties.

6. Note that the results from this analysis are not directly used for the development of proposed protocol and algorithm, but only used for the motivation.



(a) Link contention graph (b) CSMA Markov Chain

Fig. 2. Example: CSMA operation mechanism.

3 Background and System Overview

In this section, we elucidate a background of location sharing idea and an overview of a proposed system which consists of a distributed location sharing protocol and an optimal mean waiting time decision algorithm in Sections 4 and 5, respectively.

3.1 Background - Optimal CSMA

We propose a location sharing scheme by invoking an idea from the optimal CSMA [23] in the field of wireless networks. Hence, we first introduce a concept and an operational principle of the optimal CSMA.

CSMA Model. We define a link and an interfering edge in wireless networks (see Fig. 2a as a simple example). A link is a pair of transmitter and receiver. An interfering edge means that two links are connected with each other; hence they cannot transmit at the same time (i.e., conflict). Otherwise, two links can transmit at the same time, i.e., there is no interference between two links.

Assumptions. To capture the essential features of the CSMA/CA, we make three simplifying assumptions. (i) If two links are conflicted, then either of the two links hears when the other one transmits. (ii) The sensing of conflict simultaneously occurrs. (iii) All links in the networks are always backlogged. Then, the CSMA/CA is operated by two key mechanisms: conflict sensing and competition for transmission among interfering links.

Operation of CSMA/CA. We explain the operation of the CSMA/CA[7] with a simple example in Fig. 2a. First, two transmitters are aware of the conflict among links which are connected. If there is no conflict, two transmitters of link A and link B make a decision of random backoff times which are exponentially distributed with mean $1/R_A$ and $1/R_B$, respectively where R_A and R_B are transmission aggressiveness of link A and link B, respectively. If there is no conflict during pre-determined backoff time, the transmitter transmits the data after pre-determined backoff time. We assume that the transmission time is exponentially distributed with mean 1. If the other transmitter at the other link transmits before pre-determined backoff time (so, conflict is sensed), then the transmitter suspends its backoff and resumes it after the conflicting transmission is finished. Ideally, we assume the probability that backoff times of different links are the same is zero, hence the collision would never happen.

Optimal CSMA. Typically, an optimal control of the WiFi networks using CSMA protocol is known as to be challenging due to its distributed nature. However, Jiang et al. [23] suggested distributed throughput optimal control of the CSMA just by controlling mean backoff time of each device. An operation of the CSMA can be modeled by continuous time Markov chain (CTMC) as shown in Fig. 2b. In the Fig. 2b, link A transmits the data with R_A of transition rate whereas link B transmits the data with R_B of transition rate. A stationary distribution of the CTMC can be expressed by a mean backoff time of each link due to the fact that transition rate of each link is related to the mean backoff time. Hence, throughput optimality[8] can be achieved via distributed control of the mean backoff time at each link.

3.2 System Overview

First, we suggest a distributed location sharing protocol where a key mechanism involves a neighbor finding (similar with the principle of the sensing in the CSMA) and a competition to turn off localizing using random waiting time (similar with the principle of the random backoff time in the CSMA). Then, we define an objective function which aims to minimize the average number of location measurements of all devices for a given fairness criterion. Next, we derive the stationary distribution of Markov chain of the network topology of which devices operate with the DLS protocol. Finally, we develop an OWD algorithm which finds a mean waiting time where the stationary distribution of the Markov chain is the same as an optimal solution to our objective.

4 Distributed Location Sharing Protocol

4.1 System Model

We consider a set of devices $\mathcal{N}$ where the devices require their continuous location values for any reasons (e.g., mobile crowd sensing service). We assume that the device can broadcast short messages using the SRC technologies [12], [24].[9] Moreover, we assume that the SRCs of devices are synchronized with each other [25]. We define the neighbors as two devices which can scan or find each other. Then, the SRC connectivity is represented by a conflict graph $\mathcal{G}$ in which each node represents each device and each edge between two devices represents each link if the corresponding two devices are neighbors. Note that the conflict graph $\mathcal{G}$ is time-varying in response to the human mobility. Let a device n be an L-leader (i.e., Localization leader) if it measures its location and broadcasts the location value. Each L-leader has L-members (i.e., Localization members) which receive and exploit L-leader's location value. We denote the sharing set of L-leader n by L_n, i.e., $L_n = \{m \in \mathcal{N} \mid n, m \text{ areneighbors}\}$. Note that an L-member can be involved in multiple sharing sets and some L-leaders can have an empty sharing set.

We consider a vector $x \in \{0,1\}^{|\mathcal{N}|}$ where nth element of x is 1 (i.e., $x_n = 1$) if device n is a L-leader, and $x_n = 0$, otherwise. We say that x is a feasible L-leader vector if it satisfies the following condition.

Definition 1 (Feasible L-leader set condition). *Because all devices should share location values, the feasible L-leader set condition of x is given by*

$$\bigcup_{n:x_n=1} (L_n \cup \{n\}) = \mathcal{N}. \tag{1}$$

7. From now on, we call CSMA/CA as CSMA for simplicity.

8. It is achieved when the system is optimally controlled so as to stabilize all link queues for all acceptable traffic arrivals.

9. The energy consumption issue of the SRC will be addressed in Section 6.

This definition means that every device is included in at least one sharing set or it directly measures its location. Then, we can define a set of feasible L-leaders $\mathcal{I}$ where $(x^i, i \in \mathcal{I})$ satisfies the condition (1) for a given conflict graph $\mathcal{G}$.

4.2 Distributed Location Sharing Protocol

We describe a distributed collaborative localization protocol, called distributed location sharing protocol. Similar with the CSMA protocol [23], our DLS protocol is operated by a competition among neighbors for localizing-off[10] where the priority of the localizing-off is given by randomly chosen waiting time of each device.

Protocol Idea. We inherit the idea from well-known CSMA[23] for the distributed operation, where the key operations are carrier sensing and random waiting time. Our protocol selects an L-member set (it is a complementary L-leader set) by individual operation of each device to find neighbors and make a decision of random waiting time.

1) Neighbor finding. Each device listens to the location value broadcasted from L-leaders for checking competition devices which are localizing-on.[11] We assume that each device is able to know the number of competition devices by the SRC.[12] The existence of neighboring devices can be easily known by device discovery protocol in Bluetooth or beacon message in BLE [12], [26].

2) *Random waiting time*. If a device senses localizing-on devices among the neighbors, the device competes with the neighbors for localizing-off with randomly selected waiting time between 0 and maximum waiting time. For example, the device which selects the shortest waiting time among competitors is the first winner.

Practical Issues. Next, we consider four detailed issues to design the DLS protocol. *(i) Which messages should be contained in the short-range communication?* We consider four types of messages in the SRC: location value, localizing-on, localizing-off and NO message. The location value is broadcasted by L-leaders every t_{loc}. The localizing-on (or localizing-off) message means that a certain device informs neighbors of that the device will measure (or stop measuring) its location value. The NO message is used to satisfy the feasibility condition which states that all devices should have their location values for all times. *(ii) How to detect neighbor topology?* Because the neighbors compete with each other based on listening to the localizing-on, the localizing-off and the location value messages, they do not need additional messages for topology detection. The devices are able to know the number of neighbors or competitors based on the location values received during past t_{loc}. *(iii) Which location value is selected if a certain device receives several location values from several L-leaders?* Because the devices do not measure signal strengths of SRC, they do not know which one is closer. Therefore, the devices who received location values from several devices calculate the location value as the average value of all received location values. For instance, a location value would be more accurate if several location values are averaged out. *(iv) What happens if the SRC links are broken while the location sharing?* The localizing-off devices directly begin the localizing-on[13] and broadcast their location values, then they naturally participate in the competition under the newly determined SRC connectivity.

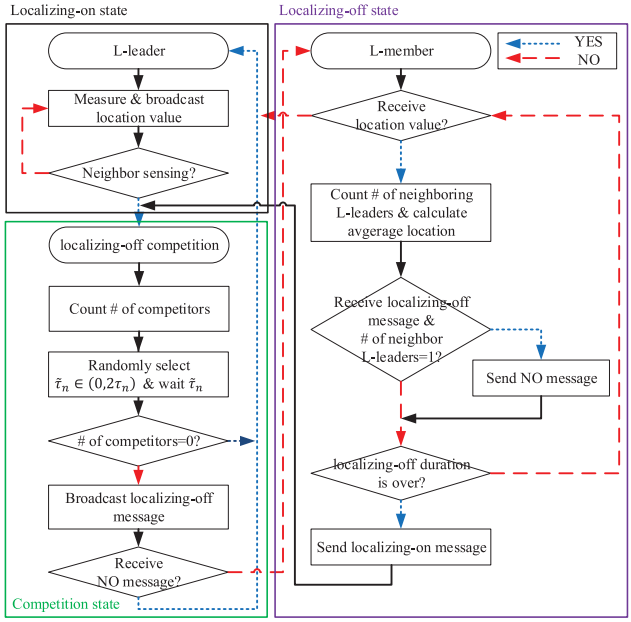


Fig. 3. Flowchart of a DLS protocol.

Operation of the DLS Protocol.[14] The DLS protocol is operated by a flowchart as shown in Fig. 3. We consider three states of each device, each of which represents the localizing-on state, the localizing-off state, and the competition state. At the localizing-on state, each device measures a location and broadcasts the location value every t_{loc}. If the device listens to the location value or the localizing-on messages from neighbors during t_{loc} (neighbor finding), the device moves to the localizing-off state. At the localizing-off state, each device listens to the location values from its L-leaders. Then the device calculates an average location value based on the received location values. If the device receives the localizing-off message when the device's L-leader is only one, the device broadcasts NO message. If the localizing-off duration is finished, the device broadcasts the localizing-on message and moves to the localizing-off competition state. If the device does not receive any location value, the device moves to the localizing-on state. At the competition state, each device checks the number of neighbors which are L-leaders by receiving the localizing-on messages or the location values during t_{loc}. Then, the device randomly selects waiting time $\tilde{\tau}_n$ between 0 and the maximum waiting time τ_n and waits for $\tilde{\tau}_n$. If the number of competitors becomes zero, the device moves to the localizing-on state, otherwise, it broadcasts the localizing-off message. If the device receives NO message after broadcasting the

10. Recall that the localizing-off means that a device does not directly measure its location, but it exploits the location values received from localizing-on devices among the neighbors.

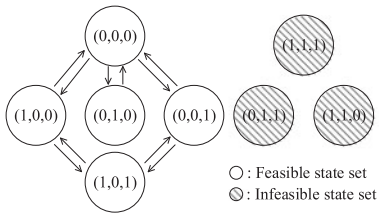
11. Recall that the localizing-on means that a device directly measures its location and share it with neighbors using the SRC.

12. It is similar with interference sensing or measurement in CSMA protocol [23].

13. In this situation, turning-on time of assisted GPS in hot state mode is known for within 1 second [27].

14. Note that since the competition time is negligible to obtain the location information, and the location information is shared by localizing-on devices every time interval of the SRC, the maximum allowable delay for the location sharing is the same as the time interval of SRC.

(a) Device topology.



(b) Markov states.

Fig. 4. Examples: Device topology and Markov states. An arrow between devices represents that two devices are neighbors. In Markov states, 0 denotes localizing-on, 1 denotes localizing-off.

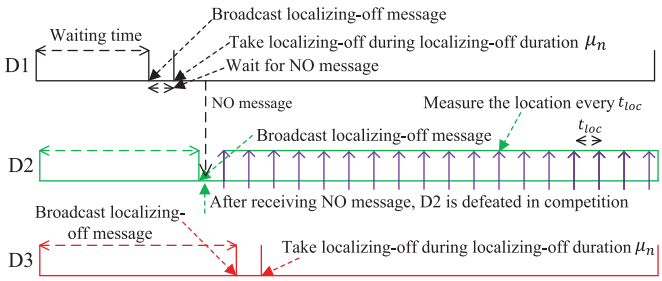


Fig. 5. An example: An operation of DLS protocol.

localizing-off message, the device is defeated in the competition and moves to the localizing-on state. We present a simple example of the DLS protocol as follows.

An Example of the DLS Protocol. In Fig. 4a, D1, D2 are neighbors of each other and D2, D3 are neighbors of each other. From this SRC connectivity, we are able to model Markov states as shown in Fig. 4b. From the feasible L-leader set condition (1), the probabilities to the infeasible state sets from other state sets must be zero. To that end, we introduce "NO message" to prevent going to the infeasible state sets. For instance, assume that waiting times of D1, D2 and D3 are initially selected by the order of D3 $>$ D2 $>$ D1 as shown in Fig. 5. Then, D1 first broadcasts a localizing-off message. Since D2 is in the localizing-off competition state, D2 does not broadcast the NO message. After waiting time of D1, D1 broadcasts a localizing-off message and moves to the localizing-off state. Next, after waiting time of D2, D2 broadcasts a localizing-off message. Since D1 does not measure the location, D1 broadcasts a NO message to D2 in order to prevent going to the infeasible state. After then, D2 is defeated in competition and moves to the localizing-on state. Finally, after waiting time of D3, D3 broadcasts a localizing-off message and waits for the NO message, and then D3 stops its location measurement for the localizing-off duration.

4.3 Message Passing

Our measurement in Section 8 shows that an average classic Bluetooth scanning power is 148.3 mW for 1-second scanning interval, which cannot be ignored compared to GPS (316.93 mW for 10 seconds interval in our measurement result (Section 8)). Hence, we check the message passings using the SRC of the DLS protocol compared to the lowest ID (Low-ID) [20], highest connectivity (High-C) [21] and HEED [22] algorithms which are well-known clustering algorithms in the context of sensor networks.[15] Table 1 summarizes the type of message passing where (O,X) means that the type of message passing is used or not. The Low-ID and High-C and HEED algorithms need to broadcast a cost and an L-leader declaration (i.e., cluster head (CH) value) per each L-leader selection phase. The DLS protocol needs to broadcast the localizing-on/off indicator per each localizing-off duration, but sending NO message can be changed in accordance with the feasible L-leader set condition. Consequentially, the number of message passings of the proposed location sharing scheme would be smaller than that of the existing algorithms, and it is verified in Section 8 via simulations in various topologies.

TABLE 1
Message Passing of Considering Algorithms

Algorithms	Message passing				
	Location	NO	Cost	localizing-on/off	Cluster head
DLS+OWD	O	O	X	O	X
Low-ID	O	X	O	X	O
High-C	O	X	O	X	O
HEED	O	X	O	X	O

5 Optimal Mean Waiting Time Decision Algorithm

If all waiting times of all devices are made by totally random manner, the energy consumption of all devices from the operation of the DLS protocol would depend on the topology of devices. Hence, it may not result in the minimization of the number of location measurements of all devices for a given fairness criterion. In this section, we suggest an optimal mean waiting time decision algorithm where the objective is to minimize the average number of location measurements with a consideration of fairness among devices.

5.1 Optimization Problem

Our two questions in introduction section for collaborative localization were in the following: (i) which device measures its location among nearby devices, and (ii) how long the device should measure the location. To address these two questions, our objective is to find the time portion of L-leaders of each device, $\theta_n \in \boldsymbol{\theta}$, which is a solution of an optimization problem with the constraints on feasible L-leader condition. All feasible vector sets in $\boldsymbol{\theta}$ satisfying feasible L-leader sets are given by

$$\mathcal{F} = \left\{ \boldsymbol{\theta} \in R^{|\mathcal{N}|} | \theta_n \geq \sum_{i \in \mathcal{I}} x_n^i \phi_i, \forall n, \sum_{i \in \mathcal{I}} \phi_i = 1, \ \phi_i \geq 0 \right\}, \quad (2)$$

15. Detailed explanations of these algorithms will be shown in Section 8.

TABLE 2
Decomposed Problems and Solutions

	Decomposed problems	Solutions for all n
[DP]	$\max_{\lambda_n} \lambda_n\left(\sum_{i\in\mathcal{I}} x_n^i\phi_i^* - \theta_n^*\right),\ \forall n$	$\lambda_n(t+T) = \lambda_n(t) + \beta(\rho_n(t) - \theta_n^*(t))$
[PP1]	$\min_{\boldsymbol{\theta}} \sum_{n\in\mathcal{N}} \left(\theta_n^\alpha - \lambda_n^*\theta_n\right)$	$\theta_n^*(t) = \begin{cases} (\lambda_n(t)/\alpha)^{1/(\alpha-1)}, & \text{if } \alpha > 1, \\ \rho_n(t), & \text{if } \alpha = 1, \end{cases}$
[PP2]	$\min_{\boldsymbol{\phi}} \sum_{i\in\mathcal{I}} \left(\phi_i\left(\sum_{n\in\mathcal{N}} \lambda_n^* x_n^i\right)\right),\ \text{s.t.} \sum_{i\in\mathcal{I}} \phi_i = 1$	Centralized, NP-hard [29]

where $\phi_i \in \boldsymbol{\phi}$ is the time portion of x^i. The optimization problem **(P)**[16] is chosen such that

$$(\mathbf{P}): \quad \min_{\boldsymbol{\theta}} \sum_{n\in\mathcal{N}} \theta_n^\alpha, \tag{3}$$

$$\text{subject to } \boldsymbol{\theta} \in \mathcal{F}, \tag{4}$$

where $\alpha \geq 1$ is the fairness parameter. When α is 1, we only consider a minimization of the total number of location measurements. As α increases, the problem **(P)** tends to select less elected devices as L-leader to minimize the objective function, which means higher fairness is enforced. Our design principle of the objective **(P)** is based on the philosophy which states that every device has the same contribution to the entire network by sharing its location for the same time period. In other words, even if different devices consume a different amount of energy to measure their locations due to the difference of each device's specification, their contributions to share the location values measured by GPS for the same time period are the same. Hence, all selfish users regard that the location sharing policy is fair if the contributions of them to share the locations are the same. For the special case, if all devices consume the same energy for measuring their locations for the same time duration using the GPS, then the objective **(P)** aims to minimize total energy consumption of all devices with $\alpha = 1$.

5.2 Optimal Solutions

Our problem **(P)** is challenging due to the feasibility of localization in (2), hence we decompose the problem **(P)** into three subproblems by employing the Lagrange duality in convex optimization [28] in the following.

Problem Decomposition. First, we define Lagrangian function of **(P)** and corresponding dual function as follows:

$$L(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\lambda}) = \sum_{n\in\mathcal{N}} \left(\theta_n^\alpha + \lambda_n\left(\sum_i x_n^i\phi_i - \theta_n\right)\right), \tag{5}$$

$$g(\boldsymbol{\lambda}) = \inf_{\boldsymbol{\theta},\boldsymbol{\phi}} \ L(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\lambda}), \tag{6}$$

where $\lambda_n \in \boldsymbol{\lambda}$ is a dual variable to satisfy the constraint (2). Then, minimizing (5) is the same problem with **(P)**, and we call it as a primal problem. From the fact, we can decompose the problem into primal problem 1 **[PP1]** and primal problem 2 **[PP2]** in Table 2 to find variables $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$, respectively. According to Lagrange duality [28], an optimal value of **(P)** is the same as an optimal value of the dual problem which finds dual variables $\boldsymbol{\lambda}$ to maximize $g(\boldsymbol{\lambda})$, hence it is reduced to the dual problem **[DP]** in Table 2. By iteratively solving the three decomposed problems in the following, we are able to find an optimal L-leader time portion of each device. Let $t = kT$ for some nonnegative integer $k = 0, 1, 2, \ldots$ and $T > 0$ is a constant.

1) *Dual problem [DP] solution:* By using the form of distributed gradient method [30] in **[DP]**, dual variable λ_n of each device n, namely *virtual queue*,[17] can be updated every $t = kT$ in the following:

$$\lambda_n(t+T) = \lambda_n(t) + \beta(\rho_n(t) - \theta_n^*(t)), \tag{7}$$

where $\rho_n(t)$ denotes the actual localizing-on ratio of device n for the last T, which is equivalent to $\sum_{i\in\mathcal{I}} x_n^i\phi_i$ in **[DP]**, $\theta_n^*(t)$ is a solution of **[PP1]**, and $\beta > 0$ is a constant.

2) *Primal problem1 [PP1] solution:* We can derive $\theta_n^*(t)$ by applying Karush-Kuhn-Tucker (KKT) conditions [28][18] because **[PP1]** is a convex function. The derived $\theta_n^*(t)$ is in the following:

$$\theta_n^*(t) = \begin{cases} (\lambda_n(t)/\alpha)^{1/(\alpha-1)}, & \forall n \in \mathcal{N}, \text{ if } \alpha > 1, \\ \rho_n(t), & \forall n \in \mathcal{N}, \text{ if } \alpha = 1. \end{cases} \tag{8}$$

3) *Primal problem2 [PP2] solution:* **[PP2]** is a problem to find an L-leader set i which has the minimum sum of virtual queues in all feasible L-leader sets at time $t = kT$. Therefore, while **[DP]** and **[PP1]** can be solved by distributed manners, **[PP2]** is a centralized problem which is NP-hard [29], and a solver of the problem should know all SRC connectivity information $\mathcal{G}$ in order to find $\boldsymbol{\phi}$. So in the next section, we devise a fully distributed algorithm which contains the distributed solution of **[PP2]**.

Solution Mechanism. Dynamics in (7) can be considered as a queueing system in which the arrival quantity is ρ_n and the departure quantity is θ_n^*. When $\alpha > 1$, the virtual queue λ_n is operated so as to achieve fairness among devices. For instance, λ_n increases if device n is sufficiently selected as the L-leader compared to θ_n^*, i.e., $\rho_n > \theta_n^*$. Then, the device is less selected since the solution of **[PP2]** finds devices

16. Since actual $\mathcal{G}$ varies in response to the mobility of devices, we should know the future mobility events to solve this problem. Unfortunately, estimating future mobility events is intractable, hence we assume that pursuing the optimal solution under the current $\mathcal{G}$ is the best in the current status. Even in this assumption, our simulation results show enough energy saving under the real mobility scenario.

17. The dual variable λ_n is updated every time slot like a queue. However, it is not a real queue, so it is called the virtual queue.

18. This condition can be applied by differentiating (5).

whose sum of virtual queues is the minimum value. Then, the localizing-on ratio of the device decreases. This mechanism forms a negative feedback of (7). That is, if α is larger, the departure value θ_n^* becomes smaller by (8), so $\lambda_n(t)$ increases at an even small increment of localizing-on ratio, which affects a decrement of localizing-on ratio. Therefore, more fairness is enforced. When $\alpha = 1$, the virtual queue λ_n is the same for all time and all devices. Therefore, $\boldsymbol{\phi}$ in **[PP2]** is selected when the number of L-leaders satisfying a feasible L-leader set condition is the minimum value (i.e., a total number of location measurements is minimized). This means that the solution absolutely does not consider fairness among devices.

5.3 Optimal Mean Waiting Time Decision Algorithm

The DLS protocol keeps being operated by a fully distributed mechanism depending on only the maximum waiting time, which is a double of mean waiting time for a given localizing-off duration. In this section, we design an optimal mean waiting time decision algorithm which achieves an optimality of the **(P)** problem. In the OWD algorithm, we inherit the distributed solutions of **[PP1]** and **[DP]** and develop a distributed optimal solution for remained **[PP2]**. To that end, we first derive a stationary distribution of Markov chain model as shown in Fig. 4b. Next, we prove that the stationary distribution, which is a control parameter of **[PP2]**, can be expressed by the specific form of equation related to a mean waiting time of each device. Moreover, we prove that Markov chain visits only and all the optimal L-member set of **[PP2]** with the satisfaction of the specific form of the equation. Then, the decomposed three problems are all distributed but coupled with each other. Hence, we can obtain an optimal solution to our problem by iteratively solving the three decomposed problems. This algorithm controls a mean waiting time of each device with the past localizing-on statistics, which makes our system find an optimal localizing-on time portion of each device.

Distributed Solution of [PP2]. From now on, instead of **[PP2]**, we use the complementary maximization form as follows.

Lemma 1. *Denote by y_n^i the localizing-off indicator of device n in L-leader set i, i.e., $y_n^i = 1$ if device n is in L-member state, and $y_n^i = 0$ otherwise. Then, **[PP2]** is equivalent to*

$$\max_{\boldsymbol{\phi}} \sum_i \left(\phi_i \left(\sum_{n \in \mathcal{N}} \lambda_n y_n^i \right) \right), \qquad s.t. \sum_i \phi_i = 1. \tag{9}$$

Proof. The proof is presented at Appendix. □

Let $\tau = [\tau_1, \dots \tau_{|\mathcal{N}|}]^T$ be a vector of the mean random waiting time, and $\mu = [\mu_1, \dots \mu_{|\mathcal{N}|}]^T$ be a vector of the mean localizing-off duration. If each device n runs the DLS protocol in Fig. 3, the L-member selection procedure follows a Markov chain in which a state is an L-member set as shown in Fig. 4b. Consider a state y^i and a device n which is a L-leader, i.e., $y_n^i = 0$, and has at least one neighboring L-leader. Then, state y^i transits to state $y^i + e_n$ with rate $1/\tau_n$, and state $y^i + e_n$ transits to state y^i with rate $1/\mu_n$, where e_n is a vector with cardinality $|\mathcal{N}|$ whose nth value is 1 and other values are 0s. If the device n has no neighboring L-leader, then state y^i cannot transit to state $y^i + e_n$ due to the NO message in Fig. 3. Thus, similar to the CSMA Markov chain in [23], the Markov chain of the L-member selection is reversible, and the stationary distribution is given by

$$\phi_i(\tau) = \frac{\prod_{n: y_n^i = 1} \frac{\mu_n}{\tau_n}}{\sum_{i' \in \mathcal{I}} \prod_{n': y_{n'}^{i'} = 1} \frac{\mu_{n'}'}{\tau_{n'}}}. \tag{10}$$

This equation shows that an L-member set i is more frequently visited in the Markov chain if it contains devices with short waiting time. We assume that μ_n of all devices are constants for simplicity. Then, we control a parameter τ to solve **[PP2]** for given $\boldsymbol{\lambda}$ in the following:

$$\tau_n = \mu_n \exp(-B\lambda_n), \qquad \forall n \in \mathcal{N}, \tag{11}$$

where $B > 0$ is a constant. Then, the following result is an immediate consequence of the rule in (11).

Theorem 1. *Fix $\boldsymbol{\lambda}$ and consider the DLS protocol under the waiting time satisfying (11). Then, in steady state, the optimal L-leader set of **[PP2]** is visited only and all as B goes to infinity.*

Proof. The proof is presented at Appendix. □

Description of OWD Algorithm. Now, we are able to describe our optimal mean waiting time decision algorithm by iteratively solving the three decomposed problems in Table 2 as follows.

OWD: Optimal Mean Waiting Time Decision Algorithm

Let $t = kT$ for some nonnegative integer $k = 0, 1, 2, \dots$ and $T > 0$ is a constant. Every kT, each device updates mean waiting time $\tau_n(kT)$ by the following mechanisms for all n.

$$\tau_n(t+T) = \mu_n \exp(-B\lambda_n(t)), \tag{12}$$

$$\lambda_n(t+T) = \begin{cases} \left[\lambda_n(t) + \beta \left(\rho_n(t) - \left(\frac{\lambda_n(t)}{\alpha} \right)^{\frac{1}{\alpha - 1}} \right) \right]^+, & \text{if } \alpha > 1, \\ \lambda_n(t), & \text{if } \alpha = 1, \end{cases} \tag{13}$$

where $\rho_n(t)$ is the localizing-on ratio of device n for T.

The mean waiting time τ_n is controlled to stabilize its virtual queue. For instance, the virtual queue $\lambda_n(t)$ increases when the device n is sufficiently selected as L-leader by (13). Then, $\tau_n(t)$ decreases by (12), so the device n tries to do not perform the localization with high probability. This forms a negative feedback mechanism. Note that the DLS protocol continuously operates even though mean waiting time and virtual queue are updated by OWD algorithm every T.

The computational complexity of OWD algorithm is $\mathcal{O}(1)$ because there is no loop to update mean waiting time. In addition, because the operation of DLS protocol is similar to CSMA protocol [23], it has also low computational complexity. As a result, each device may not consume noticeable energy to operate the DLS+OWD policy. Therefore, we can ignore the energy consumption for local computation.

Comparison with Optimal CSMA. Although our DLS +OWD policy is similarly developed with the optimal CSMA, there are fundamental differences in the following. *(i) Model:* In the CSMA, a link between two devices is

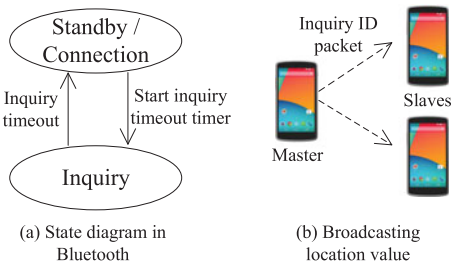


Fig. 6. State diagram in classic Bluetooth and an example for broadcasting location value.

regarded as a node and a set of two interfering links is regarded as an edge. However, in our model, a device is regarded as a node and a set of two devices within the range of SRC is regarded as an edge. *(ii) Neighboring device finding:* In the CSMA, devices are aware of the interference from other links. However, in our model, devices listen to the messages from neighboring devices instead of interference. *(iii) Feasibility:* In the CSMA, two nodes connecting with one edge, i.e., two interfering links, cannot be simultaneously activated. However, in our model, at least one neighbor device or itself should perform localizing-on. *(iv) Objective:* In the optimal CSMA, the objective is to maximize the sum of utilities of all links. However, the objective of our DLS +OWD policy is to minimize the sum of disutilities, i.e., the number of location measurements and fairness costs of all devices. *(v) Hidden terminal problem:* In the optimal CSMA, the hidden terminal situation leads to serious packet collision problem, but the similar situation in our framework is not a serious problem because our framework is based on the application layer. Hence, we can just average out the simultaneously received data.

6 Short Range Communication

In this section, we discuss the short-range communication which is essential for our DLS protocol.

Classic Bluetooth [11] or Bluetooth Low Energy [12] are suitable for the short-range communication in collaborative localization due to its coverage and energy-efficiency. The classic Bluetooth system provides a device discovery protocol [24] which just discovers or scans proxy devices before the connections to other devices. The device discovery protocol is operated at an inquiry state in Fig. 6a. In addition, the BLE provides similar proximity profile by scanning proxy devices [12]. Moreover, the device discovery protocol provides a class of device (CoD) which classifies collaborative localization devices with other devices. We can exchange the location values or side information such as *NO message* or *localizing-on/off* message among devices by an inquiry ID packet header in Fig. 6b.[19]

Various energy-efficient SRC technologies have been recently released such as the BLE [11], [12] and the backscatter communications [33] which are promising state-of-the-art technologies, especially for the IoT applications. Due to the fact that the sleep mode power consumption of the BLE is extremely low and the actual transition time from a sleep mode to a scanning mode is much shorter than that in the classic Bluetooth [15], [16], the average power consumption of the BLE can be much lower than the classic Bluetooth technology. In addition, the backscatter communications using WiFi, FM or TV signals with the low data rate and the short coverage are known to consume extremely low energy (e.g., few tens of microwatts of power) [33]. Hence, these state-of-the-art technologies also would be well matched with our scheme which periodically transmits a very small amount of data with short-range coverage. However, note that some of the latest GPS modules such as Digikey [34] consume extremely low energy that can be compared to the energy consumption of Bluetooth. Therefore, our proposed location sharing scheme allows each mobile device to decide whether to participate in the location sharing or not depending on the energy efficiency of their GPS and SRC modules.

19. The operation of BLE for scanning devices is similar with this.

7 Privacy and Security Issues

The location sharing among proxy devices can be provoked privacy and reliability problems. Indeed, leakage of identity is fatal for security and privacy. Therefore, in this section, we summarize security and privacy issues which can arise in the proposed location sharing scheme, and provide possible solutions and limitations.

First, an adversary can infer the user's real identity from its location traces. Mobile users are exposed to potential privacy threats if the adversary chases the trajectory of the user. Therefore, we suggest to change each user's *fake ID* frequently or using different fake ID for each competition to prevent the malicious tracking of location trajectory, which is suggested by security studies in D2D communications, e.g., [35], [36]. Indeed, in order to operate our DLS protocol properly, we only need to know *the number of proxy devices*, and we do not need to know *each identity of each proxy device*. Changing the fake ID of each device frequently or using different fake ID for each competition does not affect the performance of our location sharing scheme.

Second, because our location sharing scheme is operated by fully distributed manner, some selfish users may not be willing to send GPS location data to proxy devices while receiving the data from the proxy devices. This free-riding users would collapse fairness and sharing reliability. As a way to prevent this free-riding, several papers (e.g., see [36] and references therein) suggested an entity authentication for new devices or a consideration of cooperation degree, which is defined as the number of localizing-on of the corresponding device. Whenever a new user requests an authentication, the centralized coordinator can check the cooperation degree of the requested user, and give an authentication key to the user only if the cooperation degree of the user is higher than a pre-determined threshold.

Third, malicious users can disturb other proxy users by transmitting false location information. In order to prevent this attack, the users need to check the difference between the previous location value and received location value. If the difference between the just received value and previous value is larger than a pre-determined threshold, the received value is highly likely to be a false value. Another way to prevent false location values is to make a decision of more than one loser in each competition, and then averaging out all received location values. Thereby, we can reduce

TABLE 3
GPS, Bluetooth, and BLE Power Consumption of Various Devices

GPS		Bluetooth scanning/discoverable		BLE scanning/sleep	
Smartphone	Power consumption	Smartphone	Power consumption	BLE module	Power consumption
LG Nexus 4 [4]	687.25mW	Galaxy S2 [14]	192mW/159mW	TI CC2650 [15]	108mW/1,2uW
Samsung GT-19100 [31]	560mW	HTC Desire [14]	96mW/65mW	Cypress AN92584 [16]	25.5mW/3.9uW
Nexus S [32]	1006.61mW	Google Nexus S [13]	120mW/95mW		
Xiaomi Hongmi [32]	928mW				

the risk of false location transmission attack by trading more energy consumption of the localization. We evaluate an impact of the location reliability on the system performance in Section 8.

Finally, recent mobile users are able to restrict per-app access to GPS information due to the privacy issues. Therefore, we recommend that the proposed location sharing scheme is implemented in mobile operating systems (OSs) to access the GPS information. Although we summarized the several issues and possible solutions regarding privacy issues in our location sharing framework, the proposed location sharing framework still has inevitable weaknesses to the other privacy issues. For example, a collaboration of three malicious users enables to chase one specific user's trajectory by triangulation method. Moreover, provided solutions to prevent the privacy problems make the mobile devices consume more energy or require a partial support of centralized coordinator. Therefore, more powerful privacy-preserving methods in the proposed location sharing framework can be studied as a future work.

8 Evaluation

In this section, we evaluate the performance of the DLS protocol and the OWD algorithm under fixed toy topologies and real mobility traces in KAIST and NCSU.

8.1 Evaluation Under Fixed Toy Topologies

Real Power Measurement of a Smartphone. First, we survey the power consumption of various devices in [4], [12], [13], [14], [31], [32] as shown in Table 3. Note that they have used real power monitoring devices such as Monsoon power monitor or multimeter and a power monitoring application to measure the power consumption. In summary, according to our measurement and surveys, power consumption of GPS is definitely higher than that of the classic Bluetooth power for various mobile devices, and the power consumption of SRC can be further reduced by using more advanced technology such as BLE due to its extremely low power consumption in sleep mode.

Next, we measure real GPS and classic Bluetooth device power consumptions of a popular smartphone. To that end, we use a Nexus S google reference smartphone and Monsoon power monitor[20] to measure the power consumption of the smartphone. The scanning interval of the classic Bluetooth is set to be 1-second and the smartphone measures GPS location every 10 seconds. Then, every user can update its location at most every 11 seconds, hence it would not severely affect user experience. We measure the power consumption 10 times and take the average of the results. As a result of the measurement, we found that the average GPS power consumption is 316.9 mW while the average Bluetooth power consumption is 148.3 mW. We use these measured values to execute the simulations.

Simulation Setup. We consider five topologies, each of which has different SRC connectivity among devices as shown in Fig. 7. Symmetric, Star, and Full-C (Fully connected) are the basic topologies that each device can easily be aware of an energy minimal localizing-on/off solution by exchanging the number of neighboring devices with each other. However, in Complex1 and Complex2 topologies, each device cannot be aware of an energy minimal localizing-on/off solution of all devices. For example, in Complex2 topology, D1 should measure its location because the device has the maximum number of links among neighboring devices. D6 and D7 should measure their locations for feasibility condition. However, it is not an energy minimal localizing-on/off solution. The energy minimal localizing-on/off solution is that D4 and D5 measure their locations, respectively.

In addition, the power consumption of SRC of every device for 1 seconds scanning interval is set to be 148.3 mW and the GPS power consumption of every device for 10 seconds localization interval is set to be 316.93 mW where all of them are measured by ourselves. Moreover, in order to demonstrate the impact of using more energy-efficient SRC, we compare the simulation results between using classic Bluetooth and BLE as SRC. We assume the average power consumption of a BLE module called TI CC2650 with duty cycle 50 percent and 1 second scanning interval is 54 mW given by [15]. In the OWD algorithm, β, B, T and an initial virtual queue are set to be 0.1, 0.1, 60 seconds and 50, respectively. We establish the localizing-off duration as 10 seconds, and the location sharing period is 1 seconds, i.e., the L-leaders broadcast their location values every 1 second. Actually,

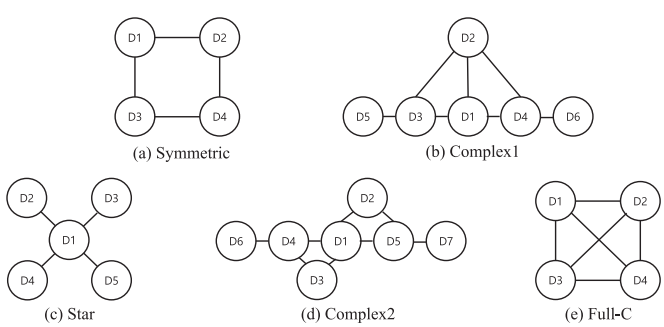
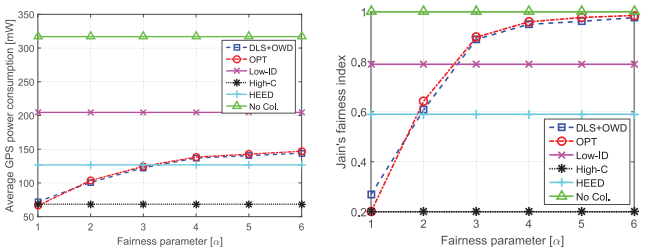


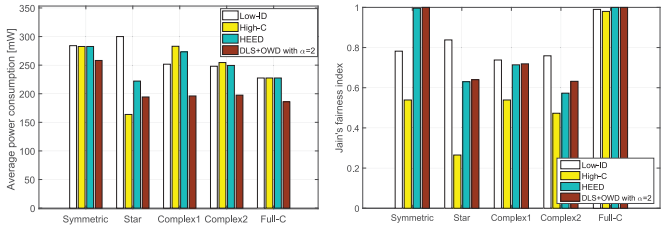
Fig. 7. Toy topologies for evaluations.

20. It is widely used for many studies related to a smartphone power consumption (see e.g., [37], [38]).



(a) Average GPS power versus fairness parameter α. (b) Jain's fairness index versus fairness parameter α.

Fig. 8. Optimality of proposed DLS+OWD policy.



(a) Average power consumption (b) Jain's fairness index

Fig. 9. Performance comparison with other algorithms.

turning-on time of A-GPS[21] in hot start mode is known for within 1 second [27]. However, we assume that it can be negligible for our localizing-off duration setting (10 seconds). Also, since average sojourn time defined in Section 2 is approximately more than 1 minute in general time zone, it can be regarded as a reasonable localizing-off duration. The average GPS and Bluetooth power consumption and Jain's fairness index [39] are considered as the performance metrics. The Jain's fairness index is calculated as follows:

$$\mathcal{J}(a_1, a_2, \ldots a_n) = \frac{\left(\sum_{i=1}^{n} a_i\right)^2}{n \cdot \sum_{i=1}^{n} a_i^2}, \tag{14}$$

where a_i denotes average GPS power consumption of device i over all time periods, and n denotes the number of mobile devices which share location information. In most of the studies on networks, Jain's fairness index has been used for a metric of throughput fairness, but in this paper, we use the Jain's fairness index as a fairness metric of average GPS power consumption of devices [40].

Optimality Verification. We verify an optimality of the DLS +OWD scheme under a star topology (Fig. 7c).[22] For comparison, we consider an optimal centralized solution (OPT) which finds optimal $\boldsymbol{\phi}$ in **[PP2]** problem with a global knowledge of SRC connectivity. Since our problem **(P)** considers only localization power consumption,[23] Fig. 8 depicts the average GPS power consumption and Jain's fairness index of DLS+OWD as a function of fairness parameter α. The DLS +OWD comes close to the OPT algorithm for all fairness parameters ($\alpha = 1, \ldots, 6$) in both terms of the average GPS power consumption and the fairness.[24] The tradeoff between average GPS power consumption and fairness parameter of our DLS+OWD scheme comes from the fact that as the fairness parameter increases, the difference of θ_n^{α} among devices which have different θ_n becomes larger. Hence, the objective function renders the selections of the power-optimal device sets (e.g., D1 in the star topology) reduced.

Comparison with Other Algorithms. We compare the performance of the DLS+OWD scheme with the other clustering algorithms (Low-ID [20], High-C [21] and HEED [22]) in five different topologies as shown in Fig. 7. In the heuristic algorithms based on clustering in sensor networks, devices exchange costs (random numbers (Low-ID), an inverse of the number of connectivity (High-C) and inverse of the number of connectivity plus residual energy (HEED)) with neighbors and select a device having the lowest cost as a cluster head which measures location. Fig. 9 shows GPS +Bluetooth power consumption and the Jain's fairness index. (i) For all topologies, the Low-ID algorithm tends to select a cluster head with fair time portion for all devices, so the average power consumption is high even though the fairness is high. The High-C algorithm always selects the same devices as cluster heads, hence the fairness is very low. The HEED algorithm tries to increase fairness than the High-C algorithm, but it still has lower performance than the DLS+OWD scheme with α=2. (ii) Especially, in Figs. 7b and 7d (complex topologies), the DLS+OWD scheme with α=2 outperforms other algorithms (up to 27.2 percent power saving with 35.8 percent higher fairness). This is because our OWD algorithm tries to find an optimal L-leader set which minimizes the sum of a square of GPS power when α=2 whereas other algorithms cannot find this optimal L-leader set. Table 4 shows the normalized average number of messages of our scheme while the other algorithms use the constant number of messages. (iii) For all topologies, the DLS+OWD scheme uses 15-22 percent of fewer numbers of message passings than the other algorithms. It means that the DLS+OWD scheme saves the Bluetooth power compared to the heuristic algorithms as well as the GPS power.

Impact of Different SRC Technologies. We evaluate an impact of different SRC communications on the average power consumption of GPS+SRC when the proposed DLS +OWD scheme is adopted. Fig. 10 depicts the average power consumption of GPS+SRC when the classic Bluetooth and BLE are used for the SRC, respectively. Both results in two different topologies demonstrate that using more

TABLE 4
Normalized Average Number of Message Passings among Devices

Topologies	Normalized average number of message passings	
	Low-ID, High-C, HEED	DLS+OWD with α=2
Symmetric	1	0.8481
Star	1	0.7785
Complex1	1	0.8536
Complex2	1	0.7978
Full-C	1	0.8284

21. It is GPS technology with help of cellular networks to get locations more quickly.

22. We choose the star topology so as to clearly show the impact of the fairness metric.

23. We simplify the objective under the assumption that the communication energy using Bluetooth is much lower than localization energy using GPS.

24. Other algorithms do not have control parameter α, i.e., they ignore the tradeoff between energy efficiency and fairness, so the performances of other algorithms have the same values for all α, respectively.

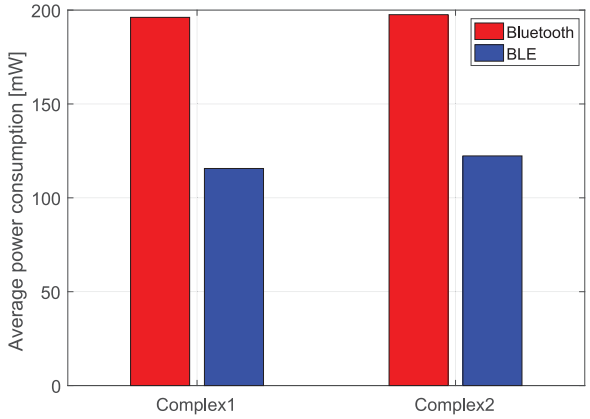


Fig. 10. Evaluation of different SRC technologies in Complex1 and Complex2 topologies.

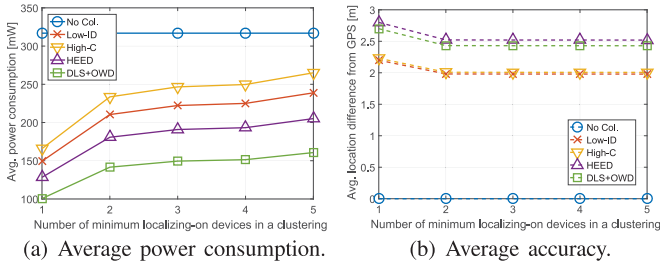
energy efficient SRC, i.e., BLE saves more than 40 percent of total average power consumption of GPS+SRC compared to use the classic Bluetooth as the SRC.

8.2 Verification under Real Mobility Trace

We verify the performance of the DLS+OWD scheme and the heuristic algorithms under real mobility trace in KAIST and NCSU campuses.

Simulation Setup. We delegate mobility trace data of 93 students in KAIST campus and 99 students in NCSU campus at Monday to Thursday PM 1:00 to PM 8:00 (for seven hours). Assumptions are the same as the human proximity analysis in Section 2. Fairness parameter α is set to be 2. We consider the average power consumption (GPS+Bluetooth) of all devices, Jain's fairness index and the average location difference from GPS, i.e., the degree of accuracy as performance metrics. Used parameters in this simulation are the same as the previous simulations in Section 8.1 except that classic Bluetooth scanning interval is 10 seconds (hence, Bluetooth power is set to be 56.15 mW).

Performance Comparison with Other Algorithms. Fig. 11 depicts average power consumption, Jain's fairness index and average location difference from GPS. The observations are as follows. (i) Our DLS+OWD with α=2 reduces average power consumption by 65.5 percent (in KAIST), 23 percent (in NCSU) compared to no collaboration, and by 27.4 percent (in KAIST), 10 percent (in NCSU) compared to HEED under similar fairness. (ii) Location differences from GPS of all algorithms are within 5 m. This implies that the estimated location using our scheme is not far from real location even though we do not consider accuracy for designing DLS+OWD.



(a) Average power consumption. (b) Average accuracy.

Fig. 12. Impact of location reliability in KAIST campus.

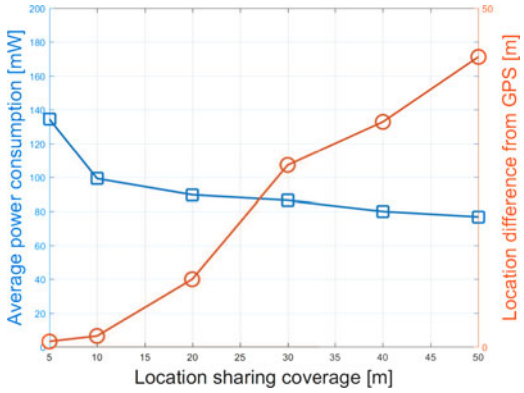
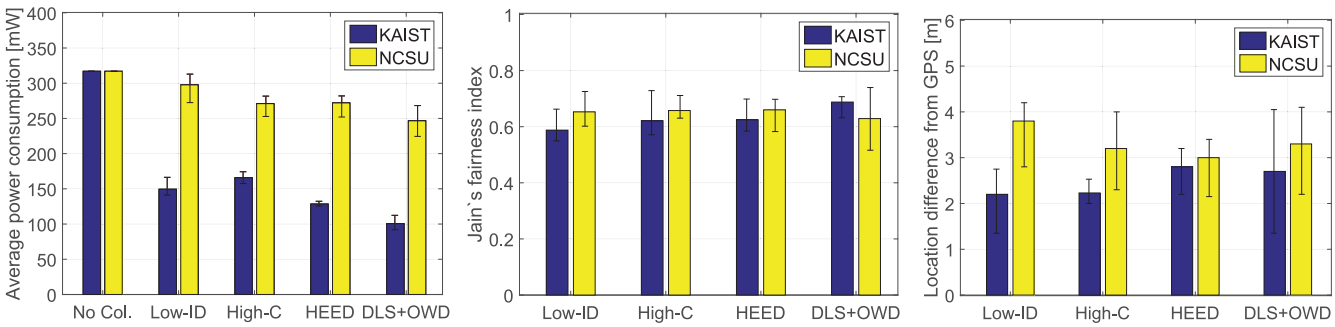


Fig. 13. Average power and accuracy of DSS+OWD versus sensing information sharing coverage in KAIST trace.

Impact of Location Reliability. Fig. 12 depict average power consumption and location accuracy versus the number of minimum localizing-on devices in an SRC clustering under the KAIST campus trace. As the number of minimum localizing-on devices in an SRC clustering increases to enhance the reliability of the localization, the average power consumption of the proposed DLS+OWD scheme slightly increases, but still less than all other algorithms. In addition, there are no notable differences in location accuracy for all cases.

Impact of Location Sharing Coverage on Average Power and Accuracy. Fig. 13 shows the average GPS+Bluetooth power consumption and location difference from GPS as a function of location sharing coverage in KAIST trace. Average power consumption for GPS+Bluetooth does not much decrease for longer coverage than 10 m whereas localization error (i.e., location difference from GPS) significantly increases for that coverage because increasing ratio of the number of proxy people is reduced when the coverage is over 10 m in



(a) Average power consumption. (b) Jain's fairness index. (c) Average accuracy.

Fig. 11. Performance under real mobility trace: A black line denotes a variation of each bar.

mobility trace of KAIST. This result means that 10 m spatial coverage (e.g., average Bluetooth coverage [18]) for collaborative localization is practically enough to bring a significant degree of accuracy and energy efficiency.

9 Conclusion

In this paper, we first analyzed temporal and spatial human proximities using GPS traces in real campuses, of which results give us a motivation which implies that the collaborative localization among nearby mobile devices can achieve significant power saving. Next, we suggested a distributed collaborative location sharing protocol and developed an optimal parameter decision algorithm to strike a balance between average power saving and device fairness. Finally, via real trace-driven simulations, we observed that the proposed protocol and algorithm have significant power saving gain even under unpredictably changing proximity environments when mobile devices use Bluetooth and GPS. We note that if new low energy technologies such as BLE and backscatter are exploited as the SRC, overheads to use proposed scheme would be significantly reduced. Our work exploited the fundamental human proximity nature and the collaboration among smartphones/pads. We believe that the collaboration study among mobile devices would be more important under future mobile trend where more energy, processing and storage burdens are given to mobile devices.

Appendix

Proof of Lemma 1

Proof. By the definition of x_n^i and y_n^i, $y_n^i = 1 - x_n^i, \forall n, i$. Then, from **[PP2]**, we have

$$\min_{\boldsymbol{\phi}} \sum_i \left(\phi_i \left(\sum_{n \in \mathcal{N}} \lambda_n x_n^i \right) \right) \tag{15}$$

$$= \min_{\boldsymbol{\phi}} \sum_i \left(\phi_i \left(\sum_{n \in \mathcal{N}} \lambda_n - \sum_{n \in \mathcal{N}} \lambda_n y_n^i \right) \right) \tag{16}$$

$$= \min_{\boldsymbol{\phi}} \sum_i \phi_i \sum_{n \in \mathcal{N}} \lambda_n + \max_{\boldsymbol{\phi}} \sum_i \phi_i \sum_{n \in \mathcal{N}} \lambda_n y_n^i \tag{17}$$

$$= \sum_{n \in \mathcal{N}} \lambda_n + \max_{\boldsymbol{\phi}} \sum_i \phi_i \sum_{n \in \mathcal{N}} \lambda_n y_n^i. \tag{18}$$

Since $\sum_n \lambda_n$ is independent of i and $\sum_i \phi_i = 1$, the first term of (18) is $\sum_n \lambda_n$ (constant in terms of ϕ). Therefore, (18) can be represented by (9). This completes the proof. □

Proof of Theorem 1

Proof. If the DLS protocol runs with a waiting time satisfying (11), a time fraction (or stationary distribution) of L-member set i, ϕ_i, is given by

$$\phi_i(\tau) = \frac{\exp(B \sum_{n \in \mathcal{N}} \lambda_n y_n^i)}{\sum_{i' \in \mathcal{I}} \exp(B \sum_{n \in \mathcal{N}} \lambda_n y_n^{i'})}. \tag{19}$$

Let $\mathcal{I}^*$ be a set of i^* in which $U(i^*) = \sum_{n \in \mathcal{N}} \lambda_n y_n^i$ is to be a global maximum. Note that localizing-on/off states in set $\mathcal{I}^*$ is the solutions of **[PP2]** from Lemma 1. Next, we divide numerator and denominator of (19) by e^{kB} where $k = \max_i U(i)$, then, we have

$$\begin{aligned} \phi_i(\tau) &= \frac{e^{(B(U(i)-k))}}{\sum_{i' \in \mathcal{I}^*} e^{(B(U(i)-k))} + \sum_{i' \notin \mathcal{I}^*} e^{(B(U(i)-k))}} \\ &= \frac{e^{(B(U(i)-k))}}{|\mathcal{I}^*| + \sum_{i' \notin \mathcal{I}^*} e^{(B(U(i)-k))}}. \end{aligned} \tag{20}$$

As B goes to infinity, the time fraction of L-member set i is given by

$$\lim_{B \to \infty} \phi_i(\tau) = \begin{cases} \frac{1}{|\mathcal{I}^*|}, & \text{if } i \in \mathcal{I}^*, \\ 0, & \text{if } i \notin \mathcal{I}^*, \end{cases} \tag{21}$$

since $e^{B(U(i)-k)}$ goes to 0 if $U(i) < k$ and 1 if $U(i) = k$. Thus, the Markov chain visits only and all the optimal L-member set of **[PP2]** equally likely. □

Acknowledgments

This work was supported by the ICT R&D program of MSIP/IITP. [R-20161130-004520, Research on Adaptive Machine Learning Technology Development for Intelligent Autonomous Digital Companion]. This project has also received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 713567. Some part of this work was presented at WiOpt 2013 [1].

References

[1] J. Kwak, J. Kim, and S. Chong, "Energy-optimal collaborative GPS localization with short range communication," in *Proc. 11th Int. Symp. Workshops Model. Optimization Mobile Ad Hoc Wireless Netw.*, May 2013, pp. 256–263.

[2] H. Ma, D. Zhao, and P. Yuan, "Opportunities in mobile crowd sensing," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 29–35, Aug. 2014.

[3] X. Chen, E. Neto, and M. Ripeanu, "Crowdsourcing for on-street smart parking," in *Proc. ACM Int. Symp. Des. Anal. Intell. Veh. Netw. Appl.*, Oct. 2012, pp. 1–8.

[4] L. Tawalbeh and R. Mehmood, "Greener and smarter phones for future cities: Characterizing the impact of GPS signal strength on power consumption," *IEEE Access*, vol. 4, pp. 858–868, Feb. 2016.

[5] J. Paek, J. Kim, and R. Govindan, "Energy-efficient rate-adaptive GPS-based positioning for smartphones," in *Proc. Int. Conf. Mobile Syst. Appl. Serv.*, Jun. 2010, pp. 299–314.

[6] I. Constandache, R. Choudhury, and I. Rhee, "Towards mobile phone localization without war-driving," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–9.

[7] M. Azizyan, I. Constandache, and R. Choudhury, "SurroundSense: Mobile phone localization via ambience fingerprinting," in *Proc. Annu. Int. Conf. Mobile Comput. Netw.*, Sep. 2009, pp. 261–272.

[8] Skyhook wireless, (2018). [Online]. Available: http://www.skyhookwireless.com/

[9] S. Kim and E. Paulos, "inAir: Measuring and visualizing indoor air quality," in *Proc. Int. Conf. Ubiquitous Comput.*, Sep. 2009, pp. 81–84.

[10] Google service: Google Now, (2018). [Online]. Available: http://www.google.com/landing/now/

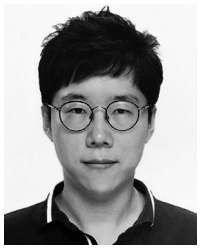
[11] H. Labiod, H. Afifi, and C. Santis, *WiFi, Bluetooth, Zigbee, and WiMAX*. Berlin, Germany: Springer, 2007.

[12] J. Liu, C. Chen, Y. Ma, and Y. Xu, "Energy analysis of neighbor discovery in bluetooth low energy," in *Proc. IEEE Veh. Technol. Conf.*, Sep. 2013, pp. 1–5.

[13] S. Tarkoma, M. Siekkinen, E. Lagerspetz, and Y. Xiao, *Smartphone Energy Consumption: Modeling and Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2014.

[14] R. Friedman, T. Haifa, A. Kogan, and Y. Krivolapov, "On power and throughput tradeoffs of WiFi and Bluetooth in smartphones," *IEEE Trans. Mobile Comput.*, vol. 12, no. 7, pp. 1363–1376, Jul. 2013.

[15] J. Lindh, C. Lee, and M. Hernes, "Measuring bluetooth low energy power consumption," (Jan. 2017). [Online]. Available: http://www.ti.com/lit/an/swra478c/swra478c.pdf
[16] U. Agarwal, K. Patel, Vikram, P. Reddy, and S. Santosh, "Designing for low power and estimating battery life for BLE applications," (2017). [Online]. Available: http://www.cypress.com/file/140991/download
[17] K. Lee, S. Hong, S. Kim, I. Rhee, and S. Chong, "SLAW: Self-similar least action human walk," *IEEE/ACM Trans. Netw.*, vol. 20, no. 2, pp. 515–529, Apr. 2012.
[18] Y. Lee, Y. Ju, C. Min, S. Kang, I. Hwang, and J. Song, "CoMon: Cooperative ambience monitoring platform with continuity and benefit awareness," in *Proc. Int. Conf. Mobile Syst. Appl. Serv.*, Jun. 2012, pp. 43–56.
[19] A. Yassin, T. Nasser, M. Awad, A. Dubai, R. Liu, C. Yuen, R. Raulefs, and E. Aboutanios, "Recent advances in indoor localization: A survey on theoretical approaches and applications," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1327–1346, Apr.–Jun. 2017.
[20] D. Baker and A. Ephremides, "A distributed algorithm for organizing modeling radio telecommunication networks," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst.*, Apr. 1981, pp. 476–483.
[21] A. Parekh, "Selecting routers in ad-hoc wireless networks," in *Proc. SBT/IEEE Int. Telecommun. Symp.*, Aug. 1994, pp. 420–424.
[22] O. Younis and S. Fahmy, "HEED: A hybrid energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Trans. Mobile Comput.*, vol. 3, no. 4, pp. 366–379, Oct.–Dec. 2004.
[23] L. Jiang and J. Walrand, "A distributed CSMA algorithm for throughput and utility maximization in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 960–972, Jun. 2010.
[24] F. Naya, H. Noma, R. Ohmura, and K. Kogure, "Bluetooth-based indoor proximity sensing for nursing context awareness," in *Proc. IEEE Int. Symp. Wearable Comput.*, Oct. 2005, pp. 212–213.
[25] S. Sridhar, P. Misra, G. Gill, and J. Warrior, "CheepSync: A time synchronization service for resource constrained Bluetooth LE advertisers," *IEEE Commun. Mag.*, vol. 54, no. 1, pp. 136–143, Jan. 2016.
[26] Bluetooth for device discovery networking guide, (2018). [Online]. Available: http://www.libelium.com/documentation/waspmote/
[27] LR9101 GPS module specification sheet rev.0.6, (2006). [Online]. Available: ftp://leadtek.co.jp/gps/9101/Leadtek9101_V06_092606.pdf
[28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[29] P. Agarwal and C. Proopiuc, "Exact and approximation algorithms for clustering," in *Proc. 9th Annu. ACM-SIAM Symp. Discrete Algorithms*, Jan. 1999, pp. 658–667.
[30] R. Gibbens and F. Kelly, "Resource pricing and the evolution of congestion control," *Automatica*, vol. 35, no. 12, pp. 1969–1985, 1999.
[31] X. Li, X. Zhang, K. Chen, and S. Feng, "Measurement and analysis of energy consumption on Android smartphones," in *Proc. IEEE Int. Conf. Inf. Sci. Technol.*, Apr. 2014, pp. 242–245.
[32] S. Datta, C. Bonnet, and N. Nikaein, "Minimizing energy expenditure in smart devices," in *Proc. IEEE Conf. Inf. Commun. Technol.*, Apr. 2013, pp. 712–717.
[33] D. Bharadia, K. Joshi, M. Kotaru, and S. Katti, "BackFi: High throughput WiFi backscatter," in *Proc. ACM Conf. Special Interest Group Data Commun.*, Aug. 2015, pp. 283–296.
[34] Digikey, (2018). [Online]. Available: https://www.digikey.com/products/en/rf-if-and-rfid/rf-receivers/870?k=rfid
[35] X. Gong, X. Chen, K. Xing, D. Shin, M. Zhang, and J. Zhang, "Personalized location privacy in mobile networks: A social group utility approach," in *Proc. IEEE INFOCOM*, Apr. 2015, pp. 1008–1016.
[36] M. Haus, M. Waqas, A. Ding, Y. Li, S. Tarkoma, and J. Ott, "Security and privacy in Device-to-Device (D2D) communication: A review," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1054–1079, Jan. 2017.
[37] J. Kwak, Y. Kim, J. Lee, and S. Chong, "DREAM: Dynamic resource and task allocation for energy minimization in mobile cloud systems," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2510–2523, Dec. 2015.
[38] J. Kwak, O. Choi, S. Chong, and P. Mohapatra, "Processor-network speed scaling for energy-delay tradeoff in smartphone allocations," *IEEE/ACM Trans. Netw.*, vol. 24, no. 3, pp. 1647–1660, Jun. 2016.
[39] R. Jain, *The Art of Computer Systems Performance Analysis*. Hoboken, NJ, USA: Wiley, 1991.
[40] H. Shi, R. Prasad, E. Onur, and I. Niemegeers, "Fairness in wireless networks: Issues, measures and challenges," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 5–24, Feb. 2014.



Jeongho Kwak (S'11-M'15) received the BS degree (summa cum laude) in electrical and computer engineering from Ajou University, Suwon, South Korea, in 2008, and the MS and PhD degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2011 and 2015, respectively. He joined the CONNECT, Trinity College Dublin, Ireland, where he is currently a postdoctoral research fellow. Previously, he was a postdoctoral researcher with KAIST, Korea, INRS, and Western University, Canada. He is a recipient of the Marie Sklodowska-Curie Research Fellowship from the European Union in 2017. His research interests lie on storage/computing/networking resource orchestration in hybrid edge network architecture. He is a member of the IEEE.

Jihwan Kim (S'11-M'15) received the BS, MS, and PhD degrees all in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2009, 2011, and 2015, respectively. He is currently a senior engineer with Samsung Electronics. His research interests include the areas of WLANs, cellular network optimization, and 5G network design. He is a member of the IEEE.

Song Chong (M'93) received the BS and MS degrees from Seoul National University, and the PhD degree from the University of Texas at Austin, all in electrical and computer engineering. He is the ICT endowed chair professor in electrical engineering at the Korea Advanced Institute of Science and Technology (KAIST). He is the head of the Computing, Networking, and Security Division in electrical engineering and the founding director of the KAIST 5G Mobile Communications & Networking Research Center. Prior to joining KAIST, he was with AT&T Bell Laboratories, Holmdel, New Jersey, as a member of technical staff. His current research interests include wireless networks, mobile systems, distributed algorithms, optimization, performance evaluation, and machine learning. He has served on the editorial boards of the *IEEE/ACM Transactions on Networking*, the *IEEE Transactions on Mobile Computing*, and the *IEEE Transactions on Wireless Communications*, and the program committee of a number of leading international conferences including IEEE INFOCOM, ACM MobiCom, ACM CoNEXT, ACM MobiHoc, IEEE ICNP, and ITC. He is the steering committee chair of IEEE WiOpt and was the general chair of IEEE WiOpt 2009 and the program co-chair of IEEE SECON 2015. He received the 2013 and 2016 IEEE William R. Bennett Prize Paper Awards, given to the best original paper published in the *IEEE/ACM Transactions on Networking* in the previous three calendar years, the 2013 IEEE SECON Best Paper Award, the 2016 KAIST Grand Prize Technology Innovation Award, and the 2016 Haedong Grand Prize Research Award, given to the Korean scholar who has made the most significant contribution to the advancement of communications research over the last 10 years. He is a member of the IEEE.