DOI: https://doi.org/10.1109/WPMC.2017.8301874.

Data Aggregation Among Mobile Devices for Upload Traffic Reduction in Crowdsensing Systems

Morihiko Tamai and Akio Hasegawa
Adaptive Communications Research Laboratories
Advanced Telecommunications Research Institute International
Kyoto 619-0288, Japan

***Abstract*—Crowdsensing systems exploit widely scattered mobile devices for large scale sensing applications. In crowdsensing systems, since there is a possibility that a huge amount of data is uploaded through cellular networks or Wi-Fi, a general framework for reducing the traffic volume is required to avoid overloading these network infrastructures. In this paper, we propose a data aggregation method among mobile devices to reduce traffic volume incurred by uploading of sensory data. By utilizing opportunistic contacts between mobile devices, sensory data are moved using short range communication over mobile devices so that aggregation performance is improved. To achieve effective data movement, our method utilizes contact history among mobile devices, and let the mobile devices that are frequently contacted by the other devices collect sensory data intensively from the other devices. Experiments show that our method can achieve high aggregation performance in various settings.**

I. Introduction

With recent proliferation and advances in mobile devices such as smartphones and tablets, mobile crowdsensing [1] has become a promising way for collecting various environmental data (e.g., noise level [2], traffic condition [3], etc.) from a large number of mobile devices, enabling a broad range of possibilities of optimization in our life. For example, a crowdsensing system for large scale spectrum monitoring by scattered mobile devices [4] can be a viable solution to improve spectrum utilization by identifying spectrum opportunities and achieving spectrum sharing, which is considered as one of key features in future 5G communications [5]. In crowdsensing systems, compared with the sensing infrastructures consisting of static sensor nodes, the system operators can gather sensory data covering wide area by scattered mobile devices without a large cost incurred for maintenance (e.g., battery exchange) of those devices.

In crowdsensing systems, mobile devices are required to communicate with base stations (BSs) or access points (APs) that are connected to the Internet to upload sensory data to the cloud servers using cellular networks (e.g., 3G, LTE) or Wi-Fi. However, it is preferable for users to save the amount of data that goes through cellular networks, since the communication cost incurred is relatively high [6]. Moreover, it is common for users to choose cellular data plans that limit the usable data volume per month. On the other hand, since Wi-Fi is required to cooperate with other communication systems such as Bluetooth, which are operating on the same frequency bands [7], it is also preferable to reduce the consumption of wireless resources in uploading sensory data using Wi-Fi. In addition, the coverage of Wi-Fi APs are relatively narrow [8], there are locations where only cellular networks are available in uploading sensory data.

In order to cope with such requirements, we propose a data aggregation method among mobile devices to reduce the traffic volume incurred by uploading of sensory data. Assuming the time and space correlation among sensory data, our method aggregates sensory data that are sampled in close proximity and timing into an aggregated data, which can be represented with fewer bytes. To facilitate the reduction of total amount of data being uploaded, by deferring immediate uploading of sensory data, we utilize the opportunistic contacts between mobile devices along with the mobility of users or vehicles. When two mobile devices are located in direct communication range, one side transmits sensory data to the other side using short range wireless technologies such as Bluetooth or Wi-Fi Direct. Then, the receiver side can collectively aggregate sensory data that are in own storage and that are received from the sender side, resulting in a further reduction in the amount of data compared to the case where aggregation is performed in each mobile device separately.

To improve the effectiveness of the reduction in data amount, when a pair of mobile devices meet, it is important to decide which one acts as the sender (i.e., the other one will be the receiver), which transmits all the sensory data in own storage to the receiver. In the proposed method, we assign a number (called *rank*) to each mobile device, and when a pair of devices meet each other, by comparing ranks of each other, the device that has a lower rank than the other will act as a sender. By doing so, we can expect that a relatively small number of mobile devices that have higher ranks than the others can collect the sensory data by receiving from many other devices that have relatively lower ranks, resulting in performance improvement in data aggregation.

In this scheme, a challenging question is how to assign ranks to mobile devices to facilitate effective data movement over the opportunistic network among mobile devices. Our method utilizes contact history among mobile devices and extracts the values of a centrality measure of the devices, which represent importance of each device in the network. By assigning a rank to each device based on its value of centrality measure, the devices that are frequently contacted by the other devices tend to be a receiver of sensory data, which is beneficial for

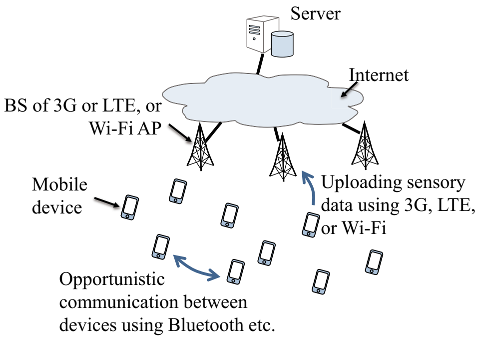
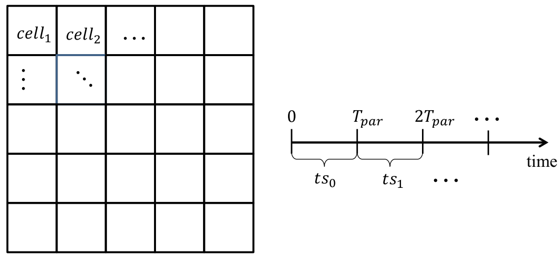


Fig. 1. Crowdsensing system.



(a) (b)

Fig. 2. Partitioning in (a) space domain and (b) time domain.

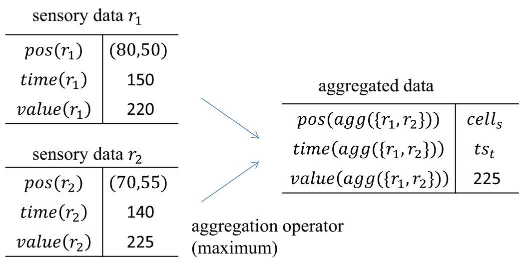


Fig. 3. Data aggregation.

effective data movement.

We conduct simulation-based experiments to show the effectiveness of the contact history-based rank assignment. As a result, we find that our method can achieve high aggregation performance compared to the method that assigns ranks randomly. We also show that dynamic rank assignment according to changes in contact graph is effective for performance improvement in data aggregation.

II. System Model

Fig. 1 shows the overview of our target environment. In Fig. 1, mobile devices represent portable devices such as smartphones and tablets that are carried by the users or vehicles. Mobile devices sample sensory data (e.g., noise level, signal strength in a frequency band, etc.) with embedded sensors or external sensors connected to a mobile device at various locations and timings. Sensory data are stored in the local storage of a mobile device for a certain period by deferring uploading to the server. During the period, when a pair of mobile devices meet each other, i.e., when they are within the range of direct communication of a short range wireless interface, one side transmits all sensory data in its storage to the other side. After some time elapsed, each device uploads all the sensory data in own storage to the server via cellular network or Wi-Fi.

In the following, the set of mobile devices are denoted by $U = \{u_1, u_2, ..., u_N\}$, where N is the number of mobile devices. Each mobile device u_i has an unique identifier, which is denoted by $id(u_i)$. Each mobile device performs sensing periodically. The sampling period is denoted by T_{sen} (e.g., 10 sec). Each mobile device periodically, e.g., every 3 hours, uploads all the sensory data in its storage to the server. Each sensory data contains the time and the position (e.g., latitude and longitude) at which the sensing is performed, and the value obtained by a sensor. For a sensory data r, we denote the position, the time, and the value by $pos(r)$, $time(r)$, and $value(r)$, respectively.

We assume that sensory data obtained in close proximity and timing have a significant correlation and they can be aggregated by applying aggregation operators (such as maximum, minimum, or average) to eliminate redundancy. To determine the range of sensory data that can be aggregated together, we consider the partitioning of space and time as follows. In space domain, the whole two dimensional area is divided into grid cells, denoted as $cell_i$, as shown in Fig. 2 (a). Similarly, the time is divided into slots, denoted as ts_i, of equal duration T_{par}, as shown in Fig. 2 (b).

A set of sensory data r_i $(i = 1, 2 \ldots, M)$ can be aggregated if and only if there exist a cell $cell_s$ and a timeslot ts_t that satisfy the following condition:

$$pos(r_i) \in cell_s \wedge time(r_i) \in ts_t, \quad \forall i. \tag{1}$$

We denote the aggregated data by $agg(\{r_1, \ldots, r_M\})$. As an example, in Fig. 3 we show a case where two sensory data are aggregated by the maximum operator. By aggregating M sensory data, the total size of sensory data is reduced to approximately $1/M$ of the original size.

Each mobile device periodically broadcasts *hello packet* with a period of T_{hel} to inform its presence to the nearby devices. When a hello packet is received at a mobile device, it establishes the connection with the sender of the hello packet, and exchanges *device information*, which includes the device identifier and the value of *rank* (described in Section III). After that, based on the device information exchanged, one device will act as the *data sender* and the other will be the *data receiver*. Finally, the data sender transmits all of the sensory data in its storage to the data receiver, and the connection between the devices is closed. It is possible to further transmit sensory data received from one device to the other devices. For example, suppose that mobile devices u_1 and u_2 met each other and u_2 received sensory data from u_1. After some time, when u_2 meets another device u_3, u_2 is able to transmit sensory data sampled itself together with the data received from u_1 to u_3.

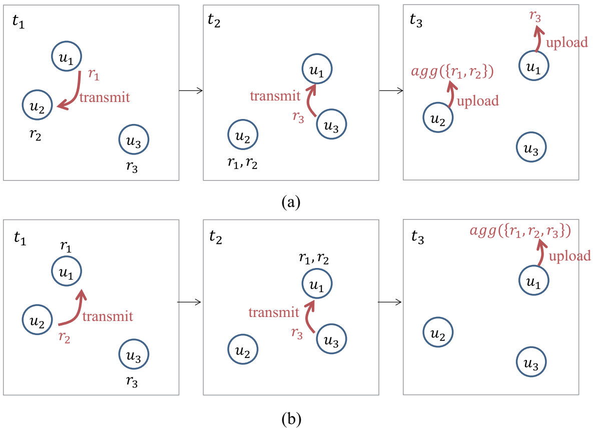


Fig. 4. Example of data movement on opportunistic network. (a) u_1 transmits data to u_2. (b) u_2 transmits data to u_1.

In this paper, we assume that aggregation operators are performed at the mobile devices only immediately before the uploading of sensory data to the server. In this sense, all sensory data are transmitted from one device to the other in its original (or raw) form. This is essential if we want to use *median* as the aggregation operator, because in general it is not possible to further aggregate sensory data that include data once aggregated before in calculating accurate value of median [9]. Nevertheless, aggregation can be effective not only in uploading but also in communication between devices if we need to reduce the communication overhead to alleviate the problems such as battery consumption for short range communication. Such a situation is considered in our future work.

III. Rank assignment

In this section, we describe how to move sensory data between mobile devices over the opportunistic network in order to improve aggregation performance. We first describe why it is important to decide the appropriate movement direction of sensory data when a pair of mobile devices meet. Then, we describe a basic procedure to move sensory data among mobile devices by introducing ranks assigned to each mobile device. Finally, we describe how to improve aggregation performance by assigning ranks by utilizing contact history among mobile devices.

Fig. 4 shows two cases where opportunistic communications occur based on the mobility of mobile devices. In the figure, we consider a situation where three mobile devices exist and each device holds one sensory data, denoted by r_1, r_2, and r_3 in its storage. In Fig. 4 (a), at time t_1, u_1 and u_2 meet each other, and u_1 transmits sensory data to u_2. Then, at time t_2, u_1 and u_3 meet each other, and u_3 transmits sensory data to u_1. Finally, at time t_3, u_1 and u_2 upload r_3 and $agg(\{r_1, r_2\})$ to the server, respectively. In this case, the amount of data uploaded is equivalent to the amount of two sensory data. On the other hand, in Fig. 4 (b) at time t_1, u_2 transmits sensory data to u_1. Finally, at time t_3, u_1 uploads $agg(\{r_1, r_2, r_3\})$ to the server. In this case, the amount of data uploaded is equivalent to the amount of single sensory data and is reduced compared to the previous case. The above example suggests the importance of deciding the movement direction of sensory data in terms of improvement in aggregation performance.

A. Basic rank assignment

In order to decide movement direction of sensory data, we assign a rank, which is a number, to each mobile device. The rank of mobile device u_i is denoted by $rank(u_i)$. Then, when a pair of mobile devices meet, the mobile device that has a lower rank transmits all the sensory data in its storage to the other mobile device. As the data movement based on the ranks is repeatedly performed when a pair of mobile devices meet, we can expect that sensory data are intensively collected by a relatively small number of mobile nodes that have higher ranks than the others, which results in improvement in aggregation performance.

In this scheme, it is important to determine how to assign ranks to mobile devices. One of the simplest approaches is to use the identifier value $id(u_i)$ as the rank of the mobile device u_i. We call this *ID-based ranking*. Later in Section III-B, we describe another approach to rank assignment.

Algorithm 1 Determining movement direction of sensory data

procedure `alg1_sub`(x_1, x_2)
  if $x_1 < x_2$ **then**
    transmit all the sensory data in its storage to u_{peer}.
  else
    receive sensory data from u_{peer}.
  end if
procedure `alg1_main`(u_{self}, u_{peer})
  if $rank(u_{self}) \neq rank(u_{peer})$ **then**
    call `alg1_sub`($rank(u_{self})$, $rank(u_{peer})$).
  else
    call `alg1_sub`($id(u_{self})$, $id(u_{peer})$).
  end if

Algorithm 1 summarizes the general framework for determining the movement direction of sensory data. When a pair of mobile devices meet, after exchanging their device information including ranks, each mobile device executes `alg1_main`, where u_{self} and u_{peer} represent itself and the peer, respectively. If $rank(u_{self})$ and $rank(u_{peer})$ have different values, the mobile device that has a lower rank transmits its sensory data to the peer. Otherwise, the same movement direction as ID-based ranking is adopted.

B. Contact history-based rank assignment

It is known that the human mobility exhibits periodic behavior and can be predicted with high reliability, because people tend to move between predetermined locations such as homes, workplaces, and schools in their daily lives [10], [11]. Exploiting this property, we present a method that utilizes the contact history among mobile devices for rank assignment.

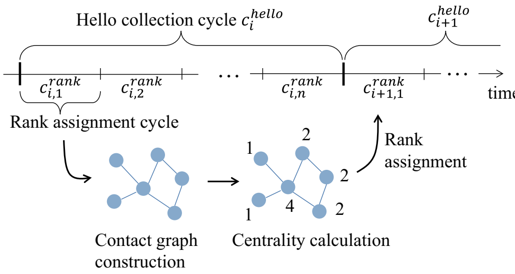


Fig. 5. Procedure for assigning ranks based on contact history.

Fig. 5 shows the overview of the procedure for rank assignment based on contact history. First, each mobile device collects hello packets sent from the nearby devices during each *hello collection cycle*, while moving according to the human's or the vehicle's movement in daily life. As shown in Fig. 5, a hello collection cycle contains multiple *rank assignment cycles*, and for each rank assignment cycle the list of the identifiers of the devices from which at least one hello packet is received during the rank assignment cycle is recorded at each mobile device. By this, in each hello collection cycle, a *contact history* is created at each mobile device so that it contains a set of lists of device identifiers, where each list contains the identifiers of the devices contacted during one rank assignment cycle. At the end of each hello collection cycle, each mobile device uploads own contact history to the server. The length of a hello collection cycle is set to be able to capture the periodicity in the mobility pattern of the mobile devices, e.g., 24 hours, while the length of a rank assignment cycle is set to a smaller value, e.g., 1 hour.

Second, at the server side, based on the contact histories received from the all mobile devices, a *contact graph*, in which each edge represents a contact between two mobile devices during one rank assignment cycle, is constructed for each of the rank assignment cycles. Next, for each contact graph, the centrality values, which represent the importance of each node in the graph, are calculated. Various centrality measures are proposed in the literature [12]. Among them, in this paper, we adopt degree centrality, which is the simplest and best known centrality[1]. Degree centrality for a given node u_i is calculated as:

$$degree(u_i) = \sum_{j=1, i \neq j}^{N} e_{ij}, \tag{2}$$

where e_{ij} equals to 1 if an edge exists between u_i and u_j in a contact graph, and otherwise equals to 0. We denote by $rank_{i,j}(u_k)$ the rank of mobile device u_k assigned for opportunistic communication in jth rank assignment cycle (denoted as $c_{i,j}^{rank}$) in the ith hello collection cycle (denoted as c_i^{hello}). Then, the rank that will be used during jth rank assignment cycle in the $i+1$th hello collection cycle at mobile device u_k is calculated as follows:

$$rank_{i+1,j}(u_k) = degree(u_k), \tag{3}$$

where $degree(u_k)$ is calculated on the contact graph in $c_{i,j}^{rank}$.

Finally, the mobile devices are informed of their ranks to be used in the each jth rank assignment cycle in the $i+1$th hello collection cycle, before the $i+1$th hello collection cycle starts.

As described above, in contact history-based rank assignment, a rank is assigned for each rank assignment cycle, and the rank of a mobile device is dynamically changed during one hello collection cycle. The rationale behind this is that the use of a fixed value as the rank during one hello collection cycle is not appropriate, because the mobility pattern of a mobile device varies with time of day (e.g., morning, afternoon, night, etc.). We see this is effective for improving aggregation performance later in Section IV.

[1]We also investigated the aggregation performance of closeness and betweenness centralities, and we could not find a significant difference compared to degree centrality in our experiments.

TABLE I
SIMULATION PARAMETERS.

Parameter	Value
Number of mobile devices	225
Simulation time	5000 s
Communication range	50 m
Area size	1.5 km $\times$ 1.5 km
Mobility model	Random Waypoint or Random Walk
Movement speed	$[0.5, 1.5]$ m/s
Pause duration in Random Waypoint	$[60, 300]$ s
Movement duration until next change of movement direction in Random Walk	60 s
Sensing period	1 s
Hello period	10 s
Cell width	250 m
Timeslot length	100 s
Length of hello collection cycle	5000 s
Length of rank assignment cycle	200 s

IV. EXPERIMENTAL EVALUATION

We conduct simulations to evaluate the aggregation performance of the rank assignment methods. In the experiments, 225 mobile devices are deployed within a 1.5 km $\times$ 1.5 km area, where the initial position of each mobile device is determined randomly, as shown in Fig. 6. Each mobile device moves according to Random Waypoint model or Random Walk model [13], with a speed randomly selected from $[0.5, 1.5]$ m/s. The sensing period T_{sen} is set to one second, where one sensory data is generated at each mobile device every T_{sen}. The width of a cell and the length of a timeslot are set to 250 m and 100 s, respectively. Each mobile device sends a hello packet with a period (T_{hel}) of 10 s. The communication range between mobile devices is set to 50 m. The simulation settings are shown in Table I. For each experiment, the simulation is executed for 20 times with different random seeds and the results are averaged over the 20 runs.

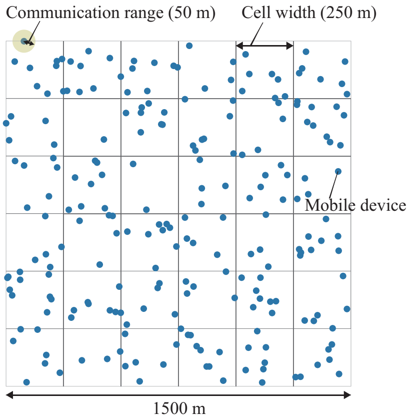


Fig. 6. Simulation area.

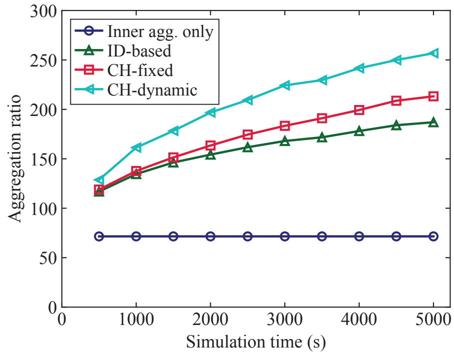


Fig. 7. Aggregation ratio.

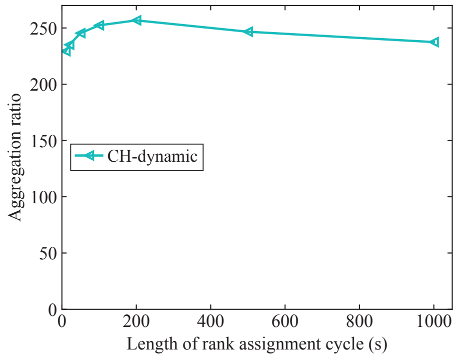


Fig. 8. Aggregation ratio vs. length of rank assignment cycle.

In the experiments, to measure aggregation performance, we assume that the amount of data of a single sensory data is regarded as 1. We also assume that the amount of data of an aggregated data generated from multiple sensory data has the same amount of data as a single sensory data. Thus, by aggregating $n(\in \mathcal{N})$ sensory data, the amount of data of the aggregated data is regarded to be 1. Similar to a measure used in the area of data compression [14], we introduce *aggregation ratio*, which measures aggregation performance, as follows:

$$\text{aggregation ratio} = \frac{\text{amount of data before aggregation}}{\text{amount of data after aggregation}}. \quad (4)$$

A. Aggregation performance

We first examine aggregation performance of various rank assignment methods. In this experiment, we compared the following methods:

(a) *Inner aggregation only*. A mobile device is forbidden to communicate with other devices, and performs aggregation only on sensory data held in own storage.

(b) *ID-based ranking*. Data movement is performed based on the procedure in Section III-A. Before starting a simulation, each mobile devices is assigned randomly selected unique identifier in $\{1, \ldots, N\}$, which is fixed throughout the simulation.

(c) *Contact history-based ranking*. Data movement is performed based on the procedure in Section III-A. In the experiment, at first, according to the procedure in Section III-B, a contact history is recorded at each mobile device during a 5000 seconds simulation, and the ranks of mobile devices in 5000 seconds are calculated. Then, by setting the same random seed for the mobility model, another 5000 seconds simulation with the same mobility is executed while data movement is performed based on the calculated ranks. To investigate the effectiveness of dynamic rank assignment per each rank assignment cycle presented in Section III-B, we compared two contact history-based ranking methods: *CH-fixed* and *CH-dynamic*, where one fixed rank is calculated for each mobile device from a 5000 seconds simulation in CH-fixed, while a sequence of ranks, each of which is assigned during one rank assignment cycle, is calculated for each mobile device in CH-dynamic.

Fig. 7 shows the aggregation performance of each method. In CH-dynamic, the length of a rank assignment cycle is set to 200 seconds. In the experiment, Random Waypoint is used as the mobility model for the all mobile devices. From Fig. 7, we find that the aggregation ratios of ID-based, CH-fixed, and CH-dynamic increase with the time elapsed. This suggests that we can expect improved aggregation performance by deferring uploading of sensory data, by setting a larger value to the uploading period. We also find that the contact history-based ranking methods achieve higher aggregation ratio than ID-based ranking. Besides, the performance improvement in CH-dynamic is more obvious compared to CH-fixed.

B. Impact of length of rank assignment cycle

Next, we investigate the impact of the length of a rank assignment cycle on aggregation performance. In the experiment, we let all mobile devices use Random Waypoint, and measure aggregation ratio of CH-dynamic with different lengths as a rank assignment cycle.

The results are shown in Fig. 8. Each aggregation ratio in the figure is measured at the simulation time of 5000. From

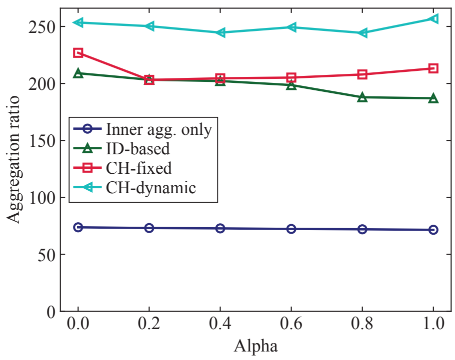


Fig. 9. Aggregation ratio vs. mobility model.

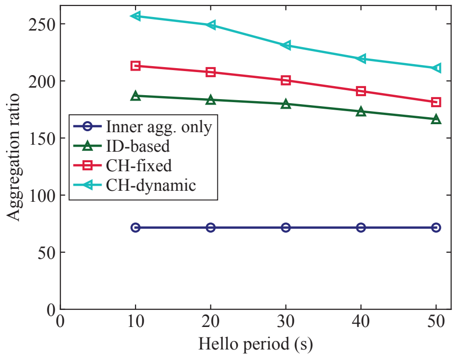


Fig. 10. Aggregation ratio vs. hello period.

the figure, we can see that aggregation ratio changes with different lengths of a rank assignment cycle, and achieves the highest aggregation ratio when the length is 200 seconds. This results suggest that by appropriately setting the length of a rank assignment cycle, we can improve aggregation performance. However, it is not trivial to find the appropriate setting, since it depends on the mobility pattern of mobile devices.

C. Impact of mobility model

Next, we investigate the impact of the mobility of mobile devices on aggregation performance. In the experiment, αN mobile devices move using Random Waypoint, while the remaining devices move using Random Walk, where $N = 225$ and α is varied from 0.0 to 1.0.

The results are shown in Fig. 9. Each aggregation ratio in the figure is measured at the simulation time of 5000. From the figure, we can see that contact history-based rank assignment methods can improve aggregation performance in various mobility settings compared to ID-based ranking. We also see that CH-dynamic achieves significantly higher aggregation ratio than the other methods.

D. Impact of hello period

Finally, we investigate the impact of hello period on aggregation performance. In the experiment, we let all mobile devices use Random Waypoint. Hello period is varied from 10 to 50 seconds, while measuring aggregation ratio of each method.

The results are shown in Fig. 10. Each aggregation ratio in the figure is measured at the simulation time of 5000. From the figure, we can see that aggregation ratio gradually decreases with the increase of hello period in ID-based ranking and contact history-based ranking, because the opportunity of communication between devices is decreased. We also find that decreasing rate in CH-dynamic is larger than the other methods, which suggests that CH-dynamic can effectively exploit contact opportunity compared to the other methods.

V. Related work

A. Wireless sensor networks

Various techniques that efficiently forward sensory data to the sink nodes in the traditional wireless sensor networks (WSNs) are proposed in the literature [15]. In WSNs, since the nodes are powered by batteries, it is critical to reduce energy consumption to extend network lifetime. Since the amount of energy consumed by communication between nearby sensor nodes is relatively high, data aggregation is one of the effective ways for reducing communication overhead. In [16], [17], [18], data forwarding protocols are proposed, which construct spanning trees over WSNs and forward sensory data from the leaf nodes to the root (sink) nodes, while applying aggregation operators (e.g., maximum, average, etc.) to the sensory data.

On the other hand, in [19], [20], [21], the authors used entropy encoding such as Huffman coding for reducing the amount of sensory data. In these methods, although the sensory data can be compressed without loss of information, the degree of reduction in data size is relatively small.

Since data forwarding protocols for WSNs described above do not consider mobility of nodes, it is difficult to effectively apply those protocols to crowdsensing systems.

B. Crowdsensing systems

In the literature, various crowdsensing systems or participatory sensing systems, which exploit widely scattered mobile devices for sensing applications such as environmental monitoring, are proposed [2], [3], [22], [23], [24]. In these systems, although there is a possibility that a huge amount of data is uploaded through cellular networks or Wi-Fi, a general framework for reducing the traffic volume have not been explored enough.

On the other hand, in [25], [26], crowdsensing systems that exploit communications among mobile devices are proposed, aiming to preserve user privacy (e.g., GPS coordinates) by anonymizing sensory data. In these systems, when mobile devices meet each other, data aggregation or data exchange is performed to anonymize sensory data, while preserving data quality as high as possible.

In [27], the authors present COUPON, which is a crowdsensing system utilizing opportunistic communication among mobile devices. Similar to our approach, by performing data aggregation (or fusion) on mobile devices, the amount of sensory data is reduced. In COUPON, similar to the traditional WSNs consisting of static sensor nodes, there are special mobile devices called *sink*, which collect sensory data from the other mobile devices and upload sensory data to the server via cellular networks. In this setting, the authors propose data forwarding protocols that are based on Binary Spray-and-Wait [28], which is one of data forwarding protocols for Delay Tolerant Networks (DTNs), and are extended to support data fusion, aiming to reduce the amount of sensory data, which result in low delay and low energy consumption in data forwarding.

On the other hand, our target environment is different from [27], because we suppose that any mobile device can directly upload sensory data to the server periodically through cellular networks or Wi-Fi, where there are no special sink devices in the network. In addition, the purpose of data aggregation is traffic reduction for cellular networks or Wi-Fi, and efficiency in data forwarding among mobile devices is not intended in our system design. Thus, data forwarding protocols such as Spray-and-Wait are not effective in our target environment, because the protocols produce multiple copies of sensory data in the network, resulting in inefficiency in reducing traffic volume in uploading.

VI. Conclusion

In this paper, we proposed rank assignment methods for effective data aggregation to achieve reduction in traffic volume incurred by uploading of sensory data in crowdsensing systems. In the proposed method, ranks, which are extracted by analyzing contact history among mobile devices, are assigned to the mobile devices so that aggregation performance is improved. Through simulation studies, we found that the contact history-based rank assignment methods are able to improve aggregation performance, especially when a sequence of ranks, each of which is used in one rank assignment cycle, is used in each mobile device.

In the future, we are interested in extending data aggregation scheme to also reduce the overhead incurred by device to device communication. We will also explore the method to find appropriate length of rank assignment cycle automatically.

Acknowledgments

This research was conducted under a contract of R&D for radio resource enhancement, organized by the Ministry of Internal Affairs and Communications, Japan.

References

[1] R. K. Ganti, F. Ye, H. Lei: "Mobile Crowdsensing: Current State and Future Challenges," *IEEE Communications Magazine*, Vol. 49, No. 11, pp. 32–39, 2011.

[2] R. K. Rana, C. T. Chou, S. S. Kanhere, N. Bulusu, W. Hu: "Ear-Phone: An End-to-End Participatory Urban Noise Mapping System," In *Proc. of IPSN'10*, pp. 105–116, 2010.

[3] B. Pan, Y. Zheng, D. Wilkie, C. Shahabi: "Crowd Sensing of Traffic Anomalies based on Human Mobility and Social Media," In *Proc. of SIGSPATIAL'13*, pp. 344–353, 2013.

[4] T. Zhang, A. Patro, N. Leng, S. Banerjee: "A Wireless Spectrum Analyzer in Your Pocket," In *Proc. of HotMobile'15*, pp. 69–74, 2015.

[5] J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C.K. Soong, J. C. Zhang: "What Will 5G Be?," *IEEE Journal on Selected Areas in Communications*, Vol. 32, No. 6, pp. 1065–1082, 2014.

[6] J. Lee, Y. Yi, S. Chong, Y. Jin: "Economics of WiFi Offloading: Trading Delay for Cellular Capacity," *IEEE Trans. on Wireless Communications*, Vol. 13, No. 3, pp. 1540–1554, 2014.

[7] S. Rayanchu, A. Patro, S. Banerjee: "Airshark: Detecting Non-WiFi RF Devices using Commodity WiFi Hardware," In *Proc. of IMC'11*, pp. 137–154, 2011.

[8] K. Lee, J. Lee, Y. Yi, I. Rhee, S. Chong: "Mobile Data Offloading: How Much Can WiFi Deliver?," *IEEE/ACM Trans. on Networking*, Vol. 21, No. 2, pp. 536–550, 2013.

[9] P. Jesus, C. Baquero, P. S. Almeida: "A Survey of Distributed Data Aggregation Algorithms," *IEEE Communications Surveys & Tutorials*, Vol. 17, No. 1, pp. 381–404, 2014.

[10] C. Song, Z. Qu, N. Blumm, A.-L. Barabási: "Limits of Predictability in Human Mobility," *Science*, Vol. 327, pp. 1018–1021, 2010.

[11] D. Karamshuk, C. Boldrini, M. Conti, A. Passarella: "Human Mobility Models for Opportunistic Networks," *IEEE Communications Magazine*, Vol. 49, No. 12, pp. 157–165, 2011.

[12] A. Landherr, B. Friedl, J. Heidemann: "A Critical Review of Centrality Measures in Social Networks," *Business & Information Systems Engineering*, Vol. 2, No. 6, pp. 371–385, 2010.

[13] T. Camp, J. Boleng, V. Davies: "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing*, Vol. 2, No. 5, pp. 483–502, 2002.

[14] S.K. Mukhopadhyay, S. Mitra, M. Mitra: "A lossless ECG data compression technique using ASCII character encoding," *Computers and Electrical Engineering*, Vol. 37, No. 4, pp. 486–497, 2011.

[15] N. A. Pantazis, S. A. Nikolidakis, D. D. Vergados: "Energy-Efficient Routing Protocols in Wireless Sensor Networks: A Survey," *IEEE Communications Surveys & Tutorials*, Vol. 15, No. 2, pp. 551–591, 2013.

[16] S. Madden, M. J. Franklin, J. M. Hellerstein, W. Hong: "TAG: a Tiny AGgregation service for ad-hoc sensor networks," In *Proc. of OSDI'02*, pp. 131–146, 2002.

[17] S. Madden, M. J. Franklin, J. M. Hellerstein, W. Hong: "The Design of an Acquisitional Query Processor For Sensor Networks," In *Proc. of SIGMOD'03*, pp. 492–502, 2003.

[18] N. Shrivastava, C. Buragohain, D. Agrawal, S. Suri: "Medians and Beyond: New Aggregation Techniques for Sensor Networks," In *Proc. of SenSys'04*, pp. 239–249, 2004.

[19] K. C. Barr, K. Asanović: "Energy-Aware Lossless Data Compression," *ACM Trans. on Computer Systems*, Vol. 24, No. 3, pp. 250–291, 2006.

[20] E. P. Capo-Chichi, H. Guyennet, J.-M. Friedt: "K-RLE : A new Data Compression Algorithm for Wireless Sensor Network," In *Proc. of SENSORCOMM'09*, pp. 502–507, 2009.

[21] F. Marcelloni, M. Vecchio: "A Simple Algorithm for Data Compression in Wireless Sensor Networks," *IEEE Communications Letters*, Vol. 12, No. 6, pp. 411–413, 2008.

[22] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, R. A. Peterson: "People-Centric Urban Sensing," In *Proc. of WICON'06*, Article No. 18, 2006.

[23] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M.B. Srivastava: "Participatory Sensing," In *Proc. of World Sensor Web Workshop (WSW'06)*, 2006.

[24] N. D. Lane, S. B. Eisenman, M. Musolesi, E. Miluzzo, A. T. Campbell: "Urban Sensing Systems: Opportunistic or Participatory?," In *Proc. of HotMobile'08*, pp. 11–16, 2008.

[25] D. Christin, J. Guillemet, A. Reinhardt, M. Hollick, S. S. Kanhere: "Privacy-preserving Collaborative Path Hiding for Participatory Sensing Applications," In *Proc. of MASS'11*, pp. 341–350, 2011.

[26] K. Xing, Z. Wan, P. Hu, H. Zhu, Y. Wang, X. Chen, Y. Wang, L. Huang: "Mutual Privacy-Preserving Regression Modeling in Participatory Sensing," In *Proc. of Infocom'13*, pp. 3039–3047, 2013.

[27] D. Zhao, H. Ma, S. Tang, X.-Y. Li: "COUPON: A Cooperative Framework for Building Sensing Maps in Mobile Opportunistic Networks," *IEEE Trans. on Parallel and Distributed Systems*, Vol. 26, No. 2, pp. 392–402, 2015.

[28] T. Spyropoulos, K. Psounis, C. S. Raghavendra: "Efficient Routing in Intermittently Connected Mobile Networks: The Multiple-Copy Case," *IEEE/ACM Trans. on Networking*, Vol. 16, No. 1, 2008.