Highly Efficient Data Gathering with Tendency Prediction based on Position Information of Event in Wireless Sensor Networks

Osamu Takyu, Ryota Sugimoto
Shinshu University, Nagano, Japan
E-mail: takyu@shinshu-u.ac.jp

***Abstract*—Environmental monitoring with wireless sensor networks is attracting much attentions. If the event giving the impact to dramatically change the environment occurs, the suitable sensors should be requested to send the sensing information for the highly accurate environmental monitoring. This paper proposes the sensor selection scheme based on the location information of event. In the proposed scheme, the spatial tendency of sensing information is predicted from the location information of event and then the suitable sensors are selected.**

I. Introduction

The monitoring system of the physical environment with a lot of sensors through the internet makes discovering new aspects of the physical environment possible, such as the internet of things (IoT) and cyber physical system (CPS) [1][2]. The internet should deal with the explosive traffic by the access from a lot of sensors and thus the saturation of communication bandwidth is a serious problem. If each sensor works with the power battelly, a lot of access from sensor to data center causes the short life time[3] and the significant consumption of frequency resources [4].

Sensor selection for the data gathering is powerful for saving power consumption of the sensor and then enhancing the lifetime of it and reducing the consumption of the frequency resources. Even if some sensors do not send sensing data to the data center, the data center can compensate for the sensing data of them from the other sensors owing to the spatial correlation of sensing data between sensors. As a result, the heat map indicating the spatial distribution of sensing data is highly accurate to the true one, where the true heat map means that it is constructed by all the sensing data. Various kinds of sensor selection methods have been proposed, so far [5], [6]. However, these do not almost consider the time transition of the spatial correlation among the sensing data or the changing of it by any event or changing the environment. The transition of spatial correlation causes the degradation of accuracy of the heat map.

This paper proposes the sensor selection and the spatial compensation using the physical event. For example, the radio wave or the acoustic wave are spread over the physical environment from the source of it within a significant short time period. The reflection in the structure though the propagation of them occurs. Therefore, the stocastical tendency ofsensing data is strictly related to the position of the source of the radio wave or the acoustic wave.

This paper constructs the database indicating the data between the spatial correlation among the sensing data and the position of source of radio or acoustic wave. It is the pre-data gathering phase. After that, in the data gathering phase, the information of spatial correlation is obtained from the database in accordance with the position of the source of the radio or acoustic wave. The sensor selection for data gathering is decided in accordance with the obtained spatial correlation. This paper is a fundamental study and thus we assume the position of source is ideally specified and then it clarifies the accuracy of the heat map under the reduction of sensor nodes for data gathering. This paper evaluates it through the experimental evaluation with using the transiver of Low power wide are (LPWA)-LoRA as the radio sensor measuring the received signal strength indicator (RSSI)

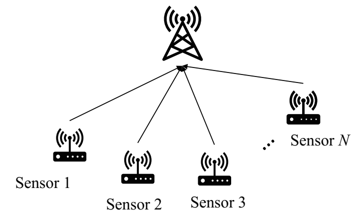


Fig. 1. System Overview of Wireless Sensor Networks

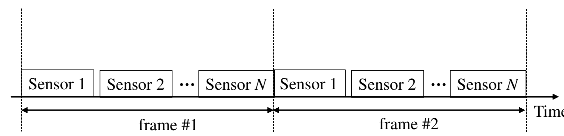


Fig. 2. Tiem Schedule for Data Gathering in WSN

II. System Overview

Figure 1 shows the overview of wireless sensor networks. There are N sensors and a data center in the monitoring field, where a data center is referred to as a fusion center (FC). Figure2 shows the time schedule for data gathering from each center to FC. Each center has an opportunity to send the sensing data to FC within a certain time period which is defined as a frame. Therefore, FC can gather the sensing data from all the sensors within a frame but the lifetime of all the sensors becomes short and the required frequency resources are large. This study saves the power consumption of each sensor and reduces the frequency resources by the sensor selection for data gathering.

FC construct the heat map indicating the spatial distribution of sensing data every time frame. If some sensors are not selected as the target for data gathering and thus they do not send the sensing data to the FC, the FC compensates the unsent sensing data by the sensing data from the other sensor.

We assume the electromagnetic wave, such as radio and optic, as the monitoring target. The proposed sensor selection and data compensation can be applied to the monitoring system whose source is specified. This paper assumes the radio signal strength indicator (RSSI) is considered as the sensing data.

III. Proposed Sensor Selection and Data Compensation

The proposed sensor selection and data compensation are composed of the three stages. The first stage is the pre-measurement stage. The correlation of RSSIs among the sensors are evaluated and the data set between the correlation of them and the position information of the radio source is constructed, where the server memorizing the data set is referred to as a database. The second phase is an actual data selection. In accordance with the position information of the radio source and the database, the FC selects the sensor as the target for data gathering. The third phase is data compensation. The FC compensates the RSSI which is not sent by the sensor by that from the other sensor. For the higher accuracy of compensation, the correlation of RSSIs among the sensors, which is measured in the first phase, makes the suitable RSSI for data compensation selected.

A. Pre-measurement Stage: Analysing Correlation of RSSIs

The ranking of sensor selection based on the correlation of RSSIs among the sensors is constructed as follows.

1) Measuring the RSSI of all the sensors as well as the position information of the radio resource, where the radio resource is the same as the transmitter of the wireless communication system. We define the kth position ($k \in 1, 2, \ldots, K$) of radio resource, where K is the total number of positions for the pre-measurement stage.
2) In the kth position of radio resource, the measured RSSIs of all the sensors are sorted in the ascending order. The sorted RSSIs set is defined as the measured sequence. The RSSI of the nth rank in the kth position of radio resource is defined as $a_{kn}(n \in 1, 2, ldots, N)$. As a result, the measured sequence, A_k, is given as

$$A_k = [a_{k1} a_{k2} \ldots a_{kN}] \tag{1}$$

After that, the difference sequence of RSSIs is derived as

$$B_k = A_k - \hat{A}_k, \tag{2}$$

where $\hat{A}_k$ is the A_k shifted by one RSSI and it is derived as

$$\hat{A}_k = [0, a_{k1} \ldots a_{kN-1}]. \tag{3}$$

3) Decision of Priority Ranking: After the difference sequence of RSSIs, B_k, is sorted in the ascending order, the priority rank for deciding the selection of sensor is constructed in accordance with the sorted B_k. The larger the difference of RSSI is, the higher the priority rank of the sensor is. If the measured RSSI is under the detection sensitivity, η, the sensor is assigned to the lowest priority rank. As a result, gathering the RSSI under the detection sensitivity is avoided.
4) When the location of the radio source is changed, the priority ranking is constructed. As a result, the data sets between the location of the radio source and A_k, B_k, and priority rank are constructed and then these are recorded in the database.

B. Data Gathering Stage

Before deciding the sensor selections, the proposed selection scheme needs the position information of the radio source. The multi-stage data gathering of sensor networks has been proposed [7]. In this data gathering, the position of the radio source is estimated at an early stage. In this paper, we assume the position information of radio source is given.

1) Transmission rate, β, is decided, where it is defined as the number of selected sensors normalized by the number of all the sensors. As β becomes smaller, the selected sensors for data gathering become fewer. Therefore, the consumption of power in sensors and the frequency resource are more significantly reduced. However, the difference between the constructed heat map and the actual one becomes larger. Therefore, the β can control the tradeoff between the consumption of power and frequency resources and the accuracy of heatmap.
2) The radio resource in the pre-measured stage is selected in accordance with the nearest to the actual radio resource. After that, from the database, the A_k, B_k, and priority ranking connected to the selected radio resource are extracted.
3) From the β, the required number of selected sensors is decided and then the sensors with top L are selected, where L is as large as the required number of selected sensors.

C. Spatial Compensation of RSSIs Stage

When the RSSI is not sent to the FC because of non-selecting of the sensor, it is compensated by the RSSI of the other sensor. The compensation is performed as follows.

4) If the sensor has the lowest rank because of the measured RSSI under the detection sensitivity, η, η is considered as the measured RSSI.
5) In accordance with the sensor order of measured sequence A_{k^*}, the gathered RSSI are sorted, where k^* means the position number of the nearest radio resource to the actual one. Note that in this sorting, the ungathered RSSI is blank. After that, the gathered RSSI of the lower rank and nearest sensor is assigned to the blank.

In the compensation, the RSSI of minimum difference sensor is used for the compensation and thus the difference between the compensated RSSI and the actual one is smaller.

[a]LPS8 LoRaWAN Gateway [b]LHT65 LoRaWAN Sensor

Fig. 3. LPWA Systems in LoRa

TABLE I
PARAMETERS OF LORAWAN

Minimum Sensitivity	-127dBm
Transmit Power	10mW
Center Frequency	923.6MHz
Spreading Factor	7
Frequency Bandwidth	125kHz
Trnsmission Period	20s

IV. EXPERIMENTAL EVALUATION AND COMPUTER SIMULATION

A. Overview of Experimental Evaluation

Figure 3 shows the equipment of experimental evaluation. This paper considers the LoRa of 920MHz bands as a radio monitoring system, where the sensor and the FC of LoRa are considered as the radio sensor and the radio resource, respectively. 32 sensors are uniformly set over the monitoring area. When the FC broadcasts a signal to all the sensors, each sensor measures the RSSI. In this experiment, when each sensor accesses the FC, the FC measures the RSSI for simplicity. Owing to the radio symmetry, the measured RSSI of the access from the FC to the sensor is the same as that from the sensor to the FC.

Table I shows the parameters of LoRaWAN and the position of FC is measured by GPS.

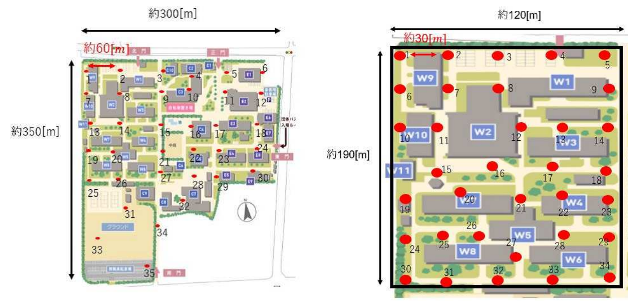
Figure 4 shows the measurement field in experimental evaluation. There are two fields and Figure 4 (a) and (b) shows two fields. Figure 4 (b) is 1/3 narrower than that (b). The distances between two sensors are 60m in Figure 4 (a) and 30m in Figure 4 (b), respectively. Figure 5 shows the location of sensors.

Figure 6 shows the image of FC. FC stands on the carriage which is 80 cm high. There are a mobile battery and a mobile wifi connecting the internet from the FC. Figure 7 shows the positions of FC in pre-measurment stage and data gathering stage.

The minimum sensitivity ,η, is -135 dBm.

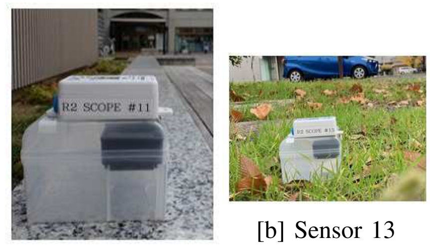
B. Evaluation Results

Figures 8 and 9 show the performance between the measured RSSI and the location of FC in 15 and 21 sensors and in 20 and 26, respectively. In the former and the latter, the positions of each sensor are given as figure 4 (a) and (b), respectively. From figure 4, the relationship between 15 and 21 sensors is line of sight and there are a few structures around them. There are some structures around 20 and 26 sensors. Therefore, the reflections of radio waves frequently occur. In figure 11 there are some common data between 15 and 21 sensors but in figure 15, there are a lot of different data between 20 and 26 sensors but in some specific position of FC, the common data between them is confirmed. If the location of FC is specified, the correlation relation among sensors can be predicted.



[a] 60m distance between sensors [b] 30m distance between Sensors

Fig. 4. Location of Sensors



[a]Sensor 11 [b] Sensor 13

Fig. 5. Appearance of Sensors

We use the MAPE (Mean Absolute Percentage Error) as the accuracy of the restored data from the actual one. It is defined as follows.

$$MAPE = E\left[\frac{1}{n}\sum_{k=1}^{n}\frac{|a_i - \acute{a_i}|}{a_i}\right], \quad (4)$$

where a_i is the actual RSSI of the ith sensor, which is the non selected sensor for data gathering and a_i' is the compensated RSSI of it.

We consider the following conventional selection for comparison. In the pre-measurement stage, the average correlation of RSSI amond sensors are calculated. The relationship between the sensor and the other sensor having the largest correlation is clarified. In the data gathering stage, the sensors are selected for data gathering as the transmission rate, β, is satisfied. In compensation stage, the RSSI which is not sent to FC is compensated by the RSSI of the sensor with largest correlation. From the fundamental study, the location of FC is ideally specified.

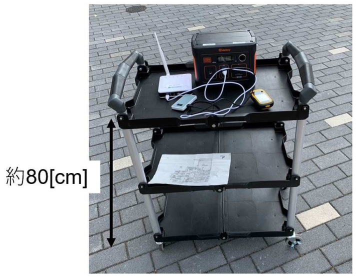
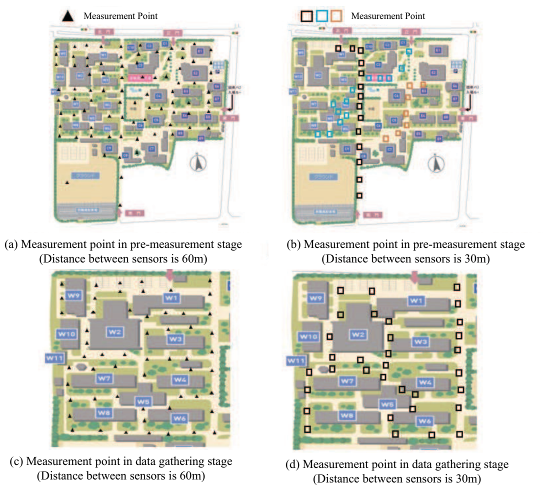


Fig. 6. Image of FC



(a) Measurement point in pre-measurement stage (Distance between sensors is 60m)

(b) Measurement point in pre-measurement stage (Distance between sensors is 30m)

(c) Measurement point in data gathering stage (Distance between sensors is 60m)

(d) Measurement point in data gathering stage (Distance between sensors is 30m)

Fig. 7. Positions of FC

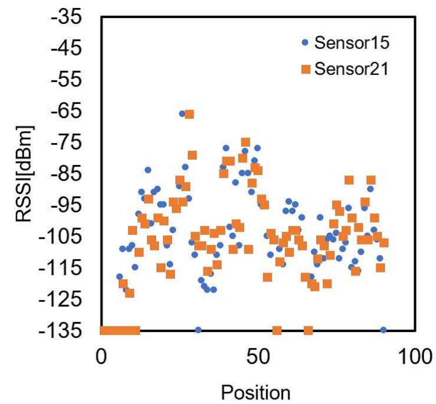


Fig. 8. Measurement Results of RSSI in Sensor 15 and 21

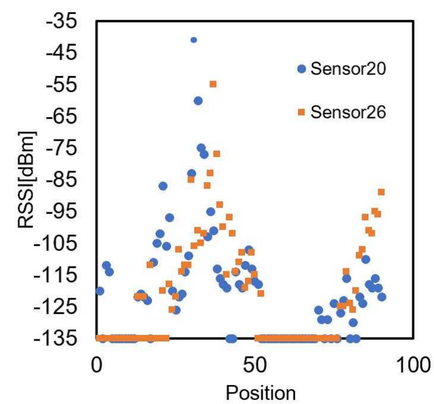


Fig. 9. Measurement Results of RSSI in Sensor 20 and 26

Figure 10 shows the cumulative distribution function (CDF) of MAPE in transmission rate,$\beta = 43\%$, where the monitoring environment is figure 4 (a). From this figure, the MAPE with the proposed selection and compensation is smaller than that with the conventional one. In CDF=0.5, the former is 6% smaller than the latter. Therefore, the proposed selection and compensation achieve the better recovery of the heatmap even under the fewer data gathering.

Figure 11 shows the performance between the transmission rate, β and the average MAPE, where the monitoring environment is figure 4 (a). From this figure, the proposed selection and compensation achieves the 7% smaller average MAPE than the conventional one. If the transmission rate is larger than 10%, the average MAPE with the proposed selection and compensation is smaller than 10 $\%$.

The proposed scheme adaptively changes the selection and compensation rules in accordance with the changing spatial correlation of RSSI decided by the position of FC. As a result, the data compensation with using the higher correlated data is achieved.

Figure 12 shows the CDF of MAPE, where the monitoring environment is figure 4 (b). The transmission rate, β, is 50 $\%$. From this figure, the proposed selection and compensation is better than the conventional one. Compared to the figure 10, the MAPE of the proposed selection and compensation is improved. This is because the correlation of RSSIs among sensors becomes larger and thus the better compensation of RSSIs is achieved. In the conventional scheme, the improvement of MAPE from figure 10 to figure 12 is not so large. Therefore, the adaptive selection and compensation to the position of radio resource is required.

Figure 13 shows the relationship between the transmission rate and the average MAPE, where the monitoring environment is figure 4 (b). When the transmission rate is larger than 50%, the proposed selection and compensation achieves a smaller average MAPE than 5%. Therefore, it can achieve better recovery of the heatmap of RSSI under the smaller sensors for data gathering.

C. Impact of Position Estimation Error

We clarify the impact of the position estimation error of FC. We consider the second nearest position of the pre-measured FC from the actual FC. The two types of error models are considered. In the first one, the nearest location and the second nearest location of FC are estimated as the location of FC by

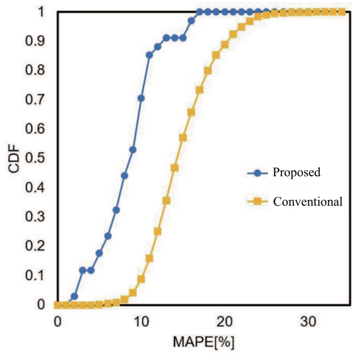


Fig. 10. CDF of MAPE in figure 4 (a)

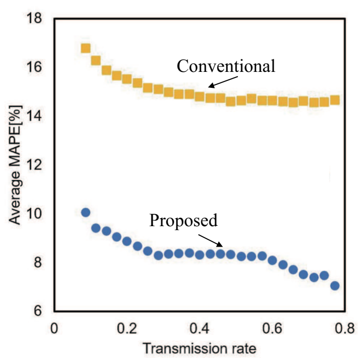


Fig. 11. Performance between Transmission Rate and Average MAPE in figure 4 (a)

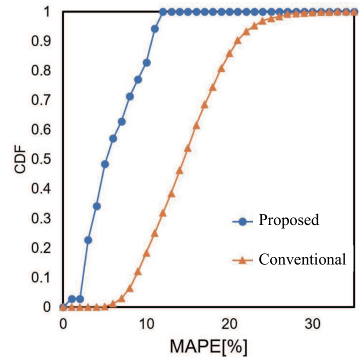


Fig. 12. CDF of MAPE in figure 4 (b)

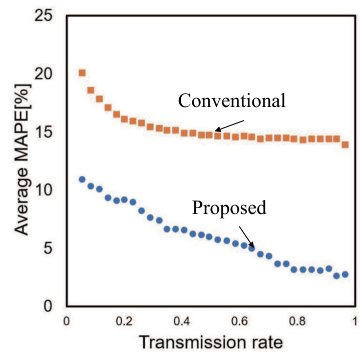


Fig. 13. Performance between Transmission Rate and Average MAPE in figure 4 (b)

50 % and 50 %. In the second one, the second nearest location of FC is estimated as the location of FC by 100 %.

Figure 14 shows the performance between CDF and MAPE, where the transmission rate is 43%. The type two error causes the MAPE degradation by two percentage points.

Figure 15 shows the relationship between the transmission rate and the average MAPE under the two kinds of position estimation error. The 1st type and the 2nd type of errors cause the MAPE degradation by 0.5 and 1.0 percentage point. As the estimation error of position of FC is smaller than 20 meters, the second nearest position of FC can be selected. If the 1.0 percentage point error of MAP is accepted, the required estimation of position of FC should be smaller than 20 meters.

V. Conclusions

This paper proposes the sensor selection and the compensation for highly accurate measuring heatmap and reducing power and frequency resources in wireless sensor network.

Acknowledgement

A part of this research project is sponsored by Ministry of Internal Affairs and Communications in Japan under the project name of Strategic Information and Communications R&D Promotion Programme (SCOPE JP 205004001) .

References

[1] Y. Zhou, F. R. Yu, J. Chen and Y. Kuo, "Cyber-Physical-Social Systems: A State-of-the-Art Survey, Challenges and Opportunities," in IEEE Communications Surveys & Tutorials, vol. 22, no. 1, pp. 389-425, Firstquarter 2020.

[2] D. Ratasich, F. Khalid, F. Geissler, R. Grosu, M. Shafique and E. Bartocci, "A Roadmap Toward the Resilient Internet of Things for Cyber-Physical Systems," in IEEE Access, vol. 7, pp. 13260-13283, 2019.

[3] O. Said, Z. Al-Makhadmeh and A. Tolba, "EMS: An Energy Management Scheme for Green IoT Environments," in IEEE Access, vol. 8, pp. 44983-44998, 2020.

[4] G. Hattab and D. Cabric, "Spectrum Sharing Protocols based on Ultra-Narrowband Communications for Unlicensed Massive IoT," 2018 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), Seoul, Korea (South), 2018, pp. 1-10.

[5] W. Zhao, Y. Han, H. Wu and L. Zhang, "Weighted Distance Based Sensor Selection for Target Tracking in Wireless Sensor Networks," in IEEE Signal Processing Letters, vol. 16, no. 8, pp. 647-650, Aug. 2009.

[6] Z. Pengfei, T. K. Boon and W. Yixin, "Sensor Selection in Wireless Sensor Networks for Structural Health Monitoring," 2019 IEEE SENSORS, Montreal, QC, Canada, 2019, pp. 1-4.

[7] T. Kobayashi, K. Yamamoto, O. Takyu and Y. Fuwa, "Verification of Interference Estimation by a Radio Sensor Considering Detection Sensitivity for Dynamic Spectrum Access in Spatial White Space," 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2020, pp. 720-724.

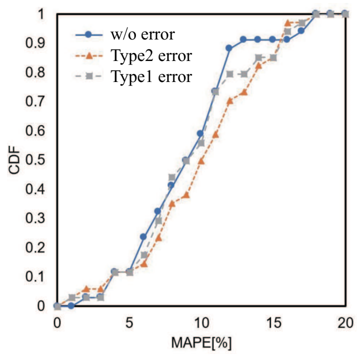


Fig. 14. CDF of MAPE with estimation error of location of FC

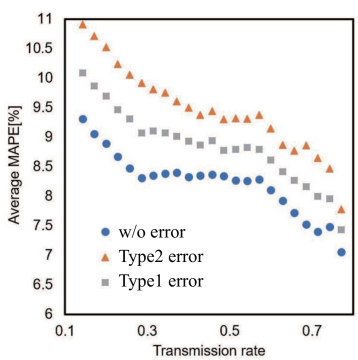


Fig. 15. Peformance between transmission rate and average MAPE with estimation error of location of FC