Information-intensive control method using estimated observation values based on correlation of observation values from multiple sensors

1st Yudai KOIKE
Shinshu University
Nagano City, Japan
23w2043b@shinshu-u.ac.jp

2nd Osamu TAKYU
Shinshu University
Nagano City, Japan
takyu@shinshu-u.ac.jp

***Abstract*—The authors have installed wireless terminals to monitor water levels in rivers and lakes in Shiojiri City, Nagano Prefecture, as a demonstration experiment of a wireless sensor network, and have been conducting environmental observations. Since the installation of the wireless sensor network, stable information aggregation has been confirmed.**

Wireless sensor networks for environmental observation require low power consumption and efficient use of wireless resources. Low power consumption is required because the environment in which they are installed has little power infrastructure, and the power supply is often powered by dry cell batteries or batteries. Efficient use of radio resources is also required because there may be other systems using the same frequency band.

We have shown through previous data analysis that the observed information is correlated. In this presentation, we use this correlation to predict future observations by machine learning, and present a study of information aggregation methods that control the frequency of information aggregation to improve the efficiency of power consumption and wireless resource utilization.

***Index Terms*—LPWA, LoRaWAN, IoT, Wireless Sensor Networks, Machine Learning.**

I. Introduction

In recent years, extreme weather events and the natural disasters that accompany them have been on the rise. Continuous environmental monitoring is necessary to minimize the damage caused by these disasters. The Internet of Things (IoT), which is becoming more and more popular every year, is an effective tool for environmental monitoring, in which wireless sensor networks are used to monitor the environment.

LPWAN (Low Pawer Wide Area Networks) is a communication method for wireless sensor networks. [2]LoRaWAN (Long Range Wide Area Networks) exists as a standard for IoT in this LPWAN, and LoRaWAN is attracting attention as a communication method that can be introduced early because of its low initial cost, including the use of unlicensed frequency bands. [3]

Often, the environments where LoRaWAN wireless sensor networks are installed are in mountainous areas, bridges, and other locations where human maintenance is time-consuming. Furthermore, the power infrastructure may be inadequate, and in such cases, the terminals are battery-powered for observation, so the system must operate for long periods of time. In addition, when packets are sent at the same timing and on the same channel in a multi-connected environment with multiple sensor nodes, information is lost due to packet collisions. Furthermore, there may be other systems using the same frequency band. Therefore, efficient use of radio resources is required.

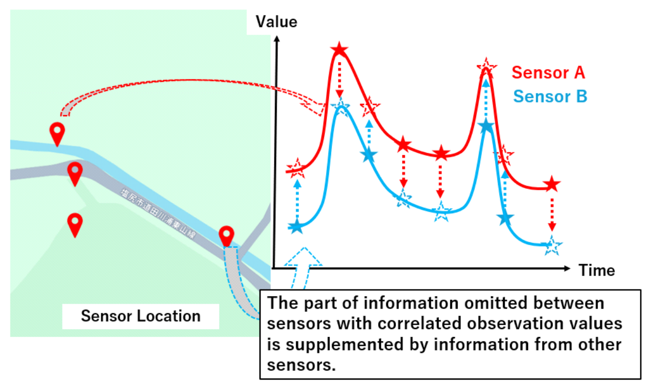

Fig. 1. Image

We have built a LoRaWAN-based system for monitoring water levels in rivers in Shiojiri City, Nagano Prefecture, and are conducting experiments using actual equipment. [4]

In this paper, we propose an "Information-intensive control method using estimated observation values based on correlation of observation values from multiple sensors" to achieve low power consumption and efficient use of wireless resources using the data obtained from this experiment.

II. Proposal method

It is known that the water level information obtained in this experiment is highly correlated with the variation of information for a certain combination of sensors. The actual sensor locations and the information aggregation of the proposed

This research and development was commissioned by SCOPE (Receipt No. JP235004002) of the Ministry of Internal Affairs and Communications.

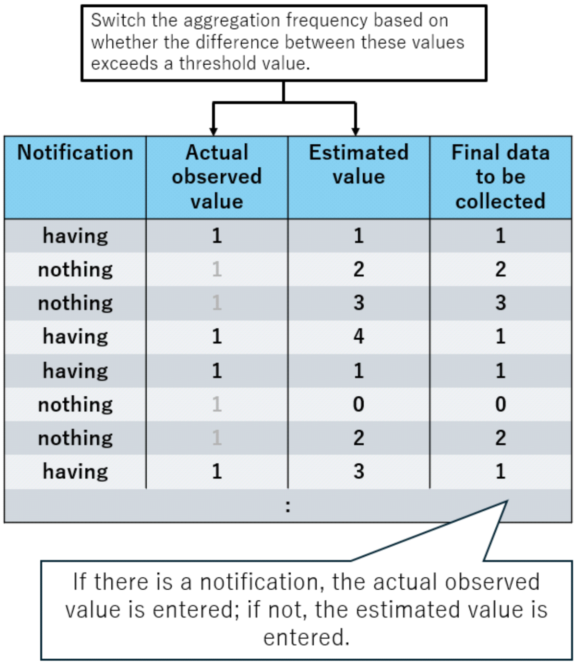

Fig. 2. Exsample of algorithm

method are shown in Figure 1. When considering the combination of two sensors A and B, if the information aggregation of sensor B is low frequency, the environmental changes during the period when the aggregation is not done will not be captured. However, the information variation of sensor B is learned by machine learning from the information of sensor A, which has a high correlation, and the missing information is estimated and supplemented. Furthermore, when an actual aggregation notification is received, as shown in Figure 2, the error between the estimated value and the actual observed value at the same time is compared. If the error exceeds the threshold value, the aggregation frequency is changed to a higher frequency and information is collected using the actual observed values. When the error is below the threshold value, the aggregation frequency is changed to a low frequency and information is collected using the estimated value. This allows the system to switch to aggregation when the error between the estimated value and the actual observed value increases, thus enabling it to cope with sudden fluctuations in observed values.

With the proposed method described above, it is expected that necessary information can be obtained while reducing redundant information aggregation.

The flow of the proposed method is shown below.

1) Obtain water level information for sensors A, B and period.
2) Learning the trend of information variation by random forest from the water level information of sensors A and B in month X
3) Estimate water level information for Sensor B for month X+1
4) Aggregate frequency is switched by an aggregate frequency control algorithm that thresholds the error between the actual observed value and the estimated value.
   a) Information aggregation is performed at a high frequency when the error exceeds a threshold value
   b) Information aggregation is performed at a low frequency when the error is below a threshold

Fig. 3. System Picture

III. Demonstrations we are conducting

The aggregation station and monitoring terminals are shown in Figure3. Furthermore, the locations where the water level monitoring terminals were installed are shown in Figure 4 and Figure 5. The numbers in the figures indicate the observation terminal numbers. The information collected by this system is stored on a server via the Internet. This enables remote monitoring. When a command is sent, instructions are broadcast from the aggregation station to the terminals. This makes it possible to change the aggregation frequency of each terminal so that it transmits once every 2 to 30 minutes in 2-minute increments.

Analysis of the aggregated water level information to date shows that certain sensor combinations have a high correlation with water level fluctuations. Table I shows the sensor combinations and correlation coefficients.

TABLE I
Sensor pairs and their correlation coefficients

Sensor pair	Correlation Coefficient
1-2	0.986
6-9	0.969
6-10	0.906
9-10	0.914

Fig. 4. Midoriko Lake

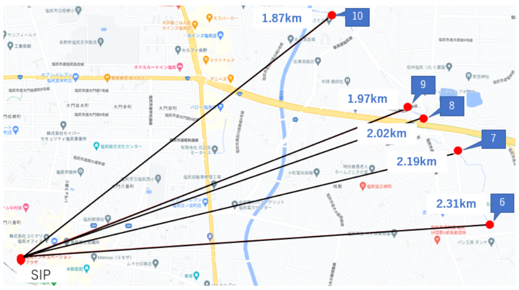

Fig. 5. Imoji river

IV. Simulation Result

For this simulation, we used August and September 2023 data from Sensor 1 and Sensor 2, two of the sensors used in the demonstration experiment, machine-learned in a random forest using the water level information from Terminal 1 and Terminal 2 in August to infer the water level of Terminal 2 in September from the water level information from Terminal 1 in September. The aggregation frequency was switched between once every 2 minutes and once every 30 minutes, and the threshold was set at 13 mm. As a conventional method, the aggregation frequency for the forecast period was set to aggregate once every 30 minutes, and the aggregation was thinned out, and the water level information at the thinned-out time was supplemented by the water level information at the immediately preceding time.

The results for the period when the water level fluctuation was large during the estimation period of Sensor B are shown in Figure 6 and Figure 7. It can be understood that the conventional method is not able to follow the water level fluctuation even if it occurs during the period when no aggregation is performed, but the proposed method is able to follow the water level fluctuation by aggregation when it exists. The numerical results of the proposed method for the entire speculation period of Sensor B show that the MAE with the true value is 0.8 mm and the maximum error with the true value is 26.1 mm. In addition, the aggregation was performed once every 30 minutes, but the aggregation frequency control algorithm resulted in an aggregation once every 2 minutes, with a transmission rate of 0.2. Here, the transmission rate is the value obtained by dividing the number of times information is actually notified by the number of times it is notified when aggregation is performed at the highest frequency. In the conventional method, when the transmission rate is set to 0.2, the maximum error from the true value is 77mm. From the above, we confirmed that the proposed method reduces the number of information notifications to 80

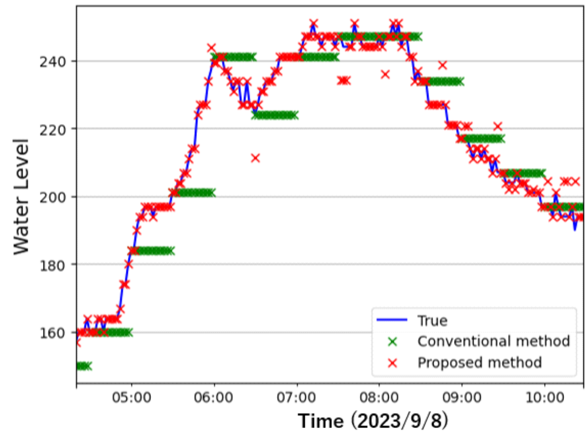

Fig. 6. Results for 9/8, 2024

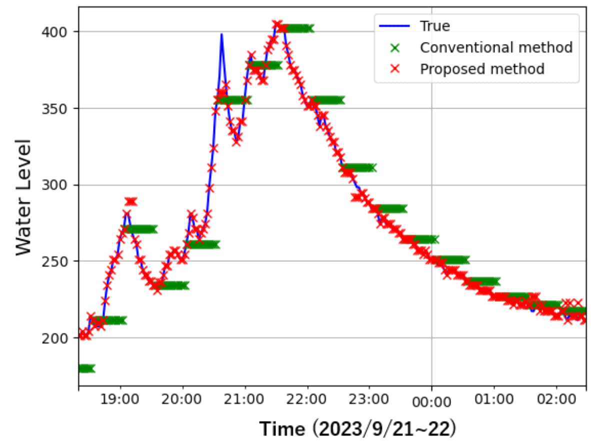

Fig. 7. Results for 9/21-22, 2024

Acknowledgment

This research and development was commissioned by SCOPE (Receipt No. JP235004002) of the Ministry of Internal Affairs and Communications.

References

[1] Y.Fuwa, "The Use of ICT Technology in Typhoon No. 19 Disaster in Nagano, Japan in 2019 and Future Prospects" ICICE Communication Society Magazine,no.59,pp.182-192, 2021-Winter

[2] B. A. Homssi et al., "A Framework for the Design and Deployment of Large-Scale LPWAN Networks for Smart Cities Applications," in IEEE Internet of Things Magazine, vol. 4, no. 1, pp. 53-59, March 2021.

[3] V. A. Stan, R. S. Timnea and R. A. Gheorghiu, ”Overview of high reliable radio data infrastructures for public automation applications: LoRa networks,” 2016 8th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Ploiesti, Romania, 2016, pp. 1-4.

[4] Y.Koike et al. , IEEE ICUFN 2023.