Estimating Indoor Population Density from Non-contact Sensor Data

Nobuyoshi Komuro
Institute of Management and Information Technologies, Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba, 263-8522 Japan
Email: kmr@facutly.chiba-u.jp

Abstract—With the spread of Covid-19, it is important to comprehend the indoor environment. It is also important to estimate indoor population density in order to avoid three Cs: Closed space, Crowded places, and Close-contact settings. This paper proposes a system for estimating indoor population density based on indoor environment data through Wireless Sensor Network (WSN). The proposed system collects indoor environment such as temperature, humidity, illuminace, CO_2 concentration, and dust level. Then, the proposed system estimates indoor population density from collected environment data. The proposed system estimates the indoor population density from indoor environment data by machine learning. This study also investigates whether sensor data are adequate for estimating indoor population density. The experimental results show that the proposed system achieves about 80 % or more estimation accuracy by using multiple types of sensors, thereby demonstrating the effectiveness of the proposed system. The proposed system is expected to measure the closed space also estimate crowded rooms, which is a helpful finding for preventing the spread of COVID-19.

Index Terms—Wireless Sensor Network (WSN), Internet of Things (IoT), indoor environment monitoring, machine learning, population density estimation

I. Introduction

With the spread of Covid-19, it is important to comprehend the indoor environment [1]-[6], such as CO2 and dust concentration, and also to avoid three Cs: Closed space, Crowded places, and Close-contact settings [1], [2], [5], [6]. The remote sensing technique is becoming popular for comprehending environmental condition [3], [4]. The remote sensing technique is a technique to comprehend global environmental information through sensing data from satellites. The CO_2 concentration was estimated by remote sensing data [3]. Also, the PM2.5 level is estimated by remote sensing data [4]. Although remote sensing technique is suitable to analyzing global environment data, it is difficult to comprehend indoor environment data.

It is also important to estimate indoor population density in order to avoid three Cs: Closed space, Crowded places, and Close-contact settings [2]. One of the effective approaches to estimate indoor population density is based on image and/or video data [6]. Yang *et al*. proposed an Artificial Intelligence (AI) based real-time social distancing detection and warning system using visual and audio devices [5]. Taylor *et al*. presents several studies that use non-contact sensors such as camera sensors and sound sensors to detect vital signs about COVID-19 [6]. These works have shown that the use of image and sound sensors were effective for supporting for infection prevention. However, from the perspective of privacy protection, it is desirable to estimate indoor environment without the use of image data.

The Internet of Things (IoT)-inspired data sensing is also expected for improving our life. The IoT has been acquiring much attention along with the improvement, miniaturization, and price reduction of wireless devices [7]-[18]. Many electric devices are connected to the Internet by adhering to the idea of the IoT. The IoT enables physical objects and/or space to communicate with each other. It likewise enables us to obtain various types of environmental data, which can be used for big data analysis. The IoT can also be utilized for various types of applications, such as smart home, smart building, smart health care, and smart rearing [15]-[18]. From the idea of Society 5.0, as proposed by Japanese Government, combining various types of data obtained using the IoT with machine learning and/or big data analysis enables us to solve social issues [19]. Hong and Otsuki developed a system which estimates human’s actions (Invasion/Indoor movement) based on the array sensor information [20]. In ref. [20], human’s action is estimated using the Support Vector Machine (SVM). Tao *et al*. developed a system that predicts the amount of wind power generating by using deep learning [21]. It is expected that the idea of Society 5.0 will enable us to estimate indoor population density without camera sensors. Komuro et al. reported that it was important to specify data set in estimating indoor working environment [22].

This paper proposes and builds an indoor population density estimation model based on collected indoor environment data such as temperature, illuminance, CO_2

Manuscript received September 25, 2021; revised February 17, 2022.
This work was partly carried out by the joint research program of CEReS, Chiba University (2020).
Corresponding author email: kmr@faculty.chiba-u.jp.
doi:10.12720/jcm.17.3.188-193

concentration, and house dust concentration. At the first step, this study develops wireless sensor nodes to be used in monitoring indoor environment. The developed system collected indoor environment data through the Wireless Sensor Network (WSN). The proposed system collects indoor environment data as big data. Then, the proposed system estimates indoor population density without camera sensors. In addition, this study investigates whether sensor data are effective for estimating indoor population density. Indoor population density is estimated from indoor environment data by machine learning method. The experimental results show the effectiveness of the proposed system.

II. Proposed Method

Fig. 1 shows the structure of the proposed system. The proposed system collects and saves indoor environment data. Sensor nodes measure environment data which includes temperature, humidity, illuminance, house dust concentration, and CO2 concentration. Sensor nodes send the measured data to the coordinator node. The coordinator node transfers the received data from sensor nodes to the data logger. The data logger logs sensor nodes' data and sends them to the cloud server. Based on the logged data, the indoor population density is estimated by machine learning.

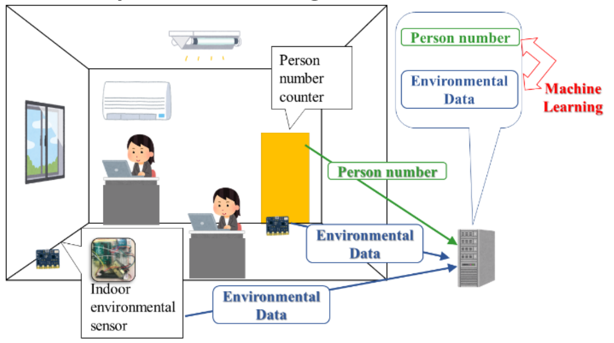

Fig. 1. System model

At the data collection phase, indoor environment data and person number are collected via developed sensors. The proposed system estimates the indoor population density with SVM, random forest algorithm, and Convolutional Neural Network (CNN), which are one of the supervised machine learning. Since random forest algorithm can obtain importance of each sensor type, this study focuses on the random forest algorithm in particular. The random forest algorithm creates decision trees on data samples, obtains the prediction from each of them, and selects the best solution through voting. The proposed system uses classification tree with the parameters in Table I for the decision tree. In order to create the decision tree of the random forest algorithm, the proposed system makes use not only of collected environment data from sensors but also person number counter, which is used as the correct answer. Environment sensor data are linked with the counted person number. At the development phase, first, random samples are selected from the collected data set. Next, a decision tree is created and grown for every sample. Estimation results are obtained from every decision tree. At the estimation phase, measured sensor data are encoded. Thereafter, prevailing data on population density is selected through a majority decision.

TABLE I: Decision Tree Parameters of the Proposed System

Criteria for measuring the quality of a split, criterion	gini
Maximum depth of decision tree	30
Minimum number of samples required to be at a leaf node	1
Minimum number of samples required to split an internal node	2
Number of trees in the forest, n_estimators	30
Randomness of the estimator, random_state	42

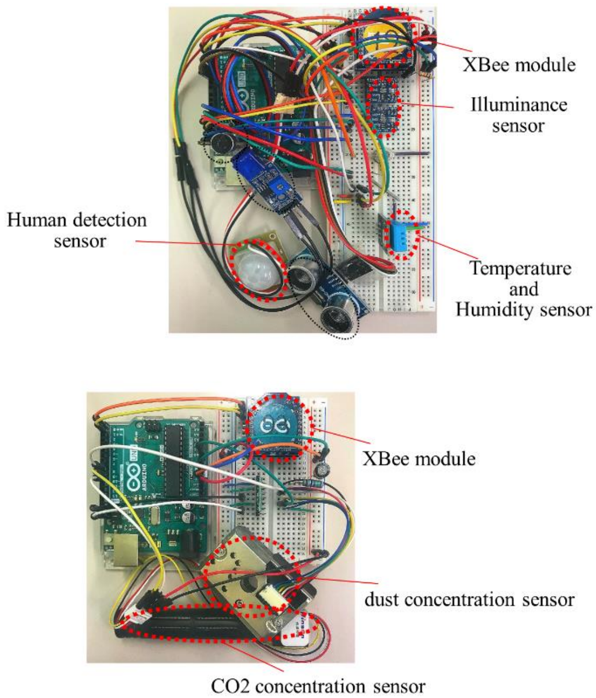

Fig. 2. Developed indoor environment sensor

A. Sensor Nodes

Fig. 2 shows the developed indoor environment sensors. Each sensor node is composed of environmental data measuring sensors, a wireless sensor module (XBee), and a one-board microcomputer. The operation of sensor nodes was carried out using the one-board microcomputer. Each sensor acquires outage voltage according to the measured value. The one-board microcomputer converts the obtained voltage to corresponding environment data values, which are temperature (Degree Celsius), humidity (%), illuminance (LUX), CO_2 concentration (PPM), and dust level (μg/m^3). In order to count the number of persons in the experimental room, each person presses "Enter"/"Exit" button when entering/leaving the room.

B. Measurement System Construction

Environment measurement devices are composed of the developed sensors, one-board microcomputer, and

XBee router. Star topology sensor network was constructed in the experimental room. There are one coordinator node and 8 sensor nodes. Table II shows the information of the equipped sensors of each node. Figures 2 and 3 show the layout of the experimental room and the deployment of sensor nodes, respectively.

Each sensor node measures environment data every 10 minutes. In order to avoid data collision due to concurrent transmission, each sensor node sends data by shifting 75 seconds.

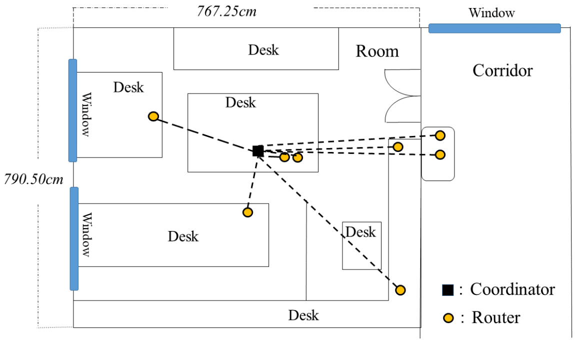

Fig. 3. Layout in the experiment room

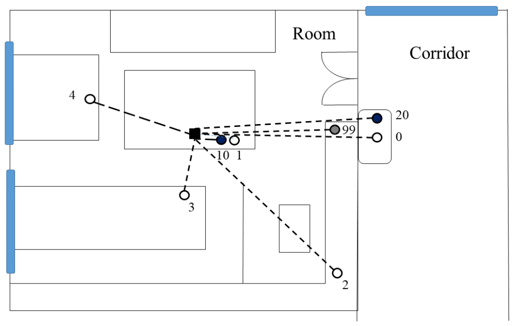

Fig. 4. Deployment of sensors

TABLE II: EQUIPPED SENSORS OF EACH NODE

Sensor ID	Equipped Sensor
0	Temperature & Humidity and Illuminance
1	Temperature & Humidity, Illuminance, and Human detection
2	Temperature & Humidity, Illuminance, and Human detection
3	Temperature & Humidity, Illuminance, and Human detection
4	Temperature & Humidity, Illuminance, and Human detection
10	CO2 concentration and Dust concentration
20	CO2 concentration and Dust concentration
99	Person number counter

C. Population Density Estimation

The population density in an experimental room was estimated from obtained environmental data. Population density was estimated by machine learning. This study estimates population density from the indoor environmental data. The environmental data (temperature, humidity, illuminance, house dust concentration, and CO_2 concentration) and the person number were logged for 13 days. 70 % of the logged environmental data were used for training data, and 30 % of the logged data were used for test data.

III. RESULTS

A. Environment Data

Figures 5 to 7 show the graphs of measured environmental data by each sensor, which are CO2 concentration, house dust concentration, and indoor person number, respectively. It is seen from Figs. 5 to 7 that the proposed system allows us to comprehend indoor environment in real-time.

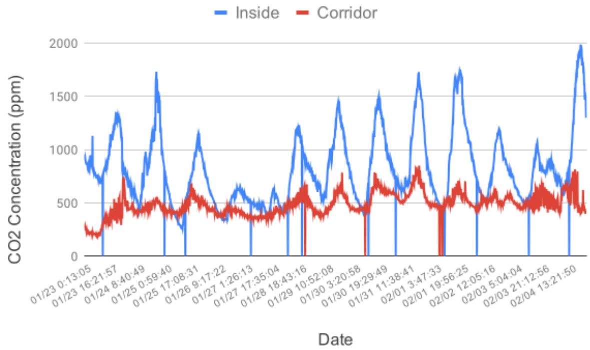

Fig. 5. CO_2 Concentration

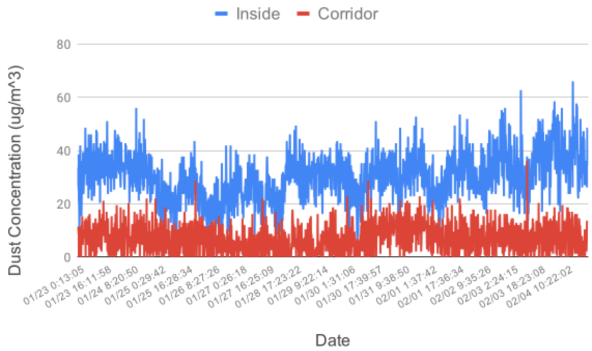

Fig. 6. Dust concentration

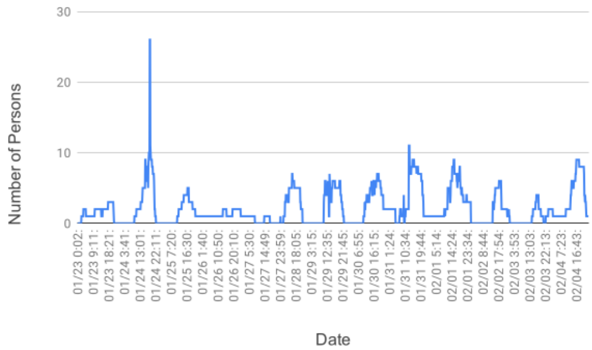

Fig. 7. Dust concentration

B. Data Loss Ratio

Table III shows the data loss ratio of each sensor node. Since processing speed and clock speed of each one-board microcomputer is slightly different, transmission timing of each sensor node changed, which caused concurrent transmission between sensor nodes. Therefore, data losses occur.

TABLE III: DATA LOSS RATIO

Sensor ID	Data Loss Ratio
0	0.49 %
1	2.04 %
2	0.77 %
3	0.71 %
4	0.77 %
10	0.66 %
20	0.77 %
99	0.77 %

C. Person Number Estimation

Table IV shows the estimation accuracy of the number of persons in the experimental room. The indoor person number was estimated by Random Forest algorithm (RF), SVM, and CNN. It is seen from Table IV that the estimation accuracy differs by the kinds and the number of input sensor data. Using multiple kinds of sensors improves the estimation accuracy. SVM can achieves about 75 % estimation accuracy by using temperature ad and humidity sensors. However, because of the effect of over learning, the estimation accuracy of SVM deteriorates as the number of sensor types increases. On the other hand, CNN can achieve over 80 % by using three or more appropriately selected sensor types. And RF can also achieve over 80 % estimation accuracy by using three or more types of sensors. In addition, RF can achieve over 85 % estimation accuracy by using five types of sensors (all types of sensors). Although it took several minutes to obtain the results with CNN, it took about one seconds to get ones with RF. In addition, RF can obtain importance of each sensor data. Therefore, RF is suitable for estimation in the proposed system.

The importance of types of sensors is presented in Fig. 8. Clearly, the importance of CO_2 concentration ranks the highest among the sensors. This result implies that the person number can be affected by CO_2 concentration. This is because the more people there are, the more CO_2 is emitted by human breathing. However, as we can see from Table IV, if we only use CO_2 sensor(s), we do not obtain high estimation accuracy. It is seen from Fig. 8 that the importance of illuminance sensor shows the second highest and that of humidity sensor shows the third highest. It is also seen from Table IV and Fig. 8 that the importance of house dust concentration for estimating indoor population density.

From above discussions, combining big data from environment sensors and machine learning techniques is useful for estimating the indoor population density. The estimation accuracy can be improved by using not only CO2 concentration data but also multiple types of environmental data. Since the estimation accuracy depends on the input sensor data, the author should investigate the types of sensing data to obtain higher estimation accuracy in more details.

TABLE IV: ESTIMATION ACCURACY FOR THE TYPES OF SENSORS

Types of Sensors	SVM	RF	NN
Temperature	0.675	0.669	0.632
Illuminance	0.502	0.643	0632
CO_2	0.373	0.400	0.476
Temperature and Humidity	0.749	0.749	0.715
Temperature and Illuminance	0.547	0.732	0.706
Humidity and Illuminance	0.563	0.781	0.768
Temperature and CO_2	0.460	0.704	0.595
Temperature, Humidity, and Illuminance	0.584	0.819	0.795
Temperature, Humidity, and CO_2	0.527	0.803	0.698
Humidity, Illuminance, and CO_2	0.426	0.831	0.817
Temperature, Humidity, Illuminance, and CO_2	0.456	0.843	0.824
Temperature, Humidity, Illuminace, CO_2, and Dust	0.454	0.853	0.839

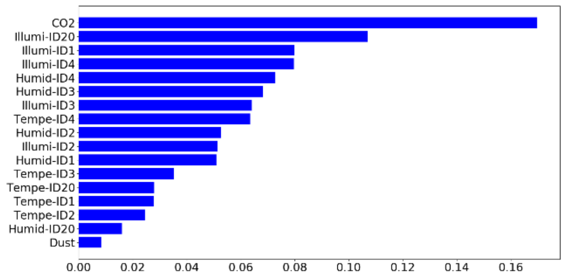

Fig. 8. Importance of sensor types

IV. CONCLUSION

This paper proposed and built an indoor population density estimation model based on collected indoor environment data through WSN. At the first step, this study developed wireless sensor nodes to be used in monitoring indoor environment. The proposed system collected indoor environment data, such as temperature, humidity, illuminace, CO_2 concentration, and dust level, through the WSN. Then the proposed system estimated indoor population density without image data from camera sensors. In addition, this study investigated whether sensor data are effective for estimating indoor population density. The experimental results showed that the proposed system achieved 80 % or more estimation accuracy by using multiple type of sensors, thereby demonstrating the effectiveness of the proposed system.

The proposed system is expected to measure the closed space also estimate crowded space, which is a useful finding for future research approaches. It is also expected that the obtained results contribute to build clean air environment and also to protect COVID-19 infection. These are the contributions of this study to global innovation. Future works include the investigation of sensor data types for obtaining higher estimation accuracy, the reduction of data losses, and long term experiment that includes seasonality.

Conflict of Interest

The author declares no conflict of interest.

Author Contributions

N.K. suggested the basic concept of this study and also suggested the algorithm demonstrated in this study. N.K. designed and performed the experiments. He also analyzed the experiment data.

References

[1] M. L. A. Torres, F. A. T. Canhisares, and V. C. Quintao, "Management of CO_2 absorbent while using the anesthesia machine as a mechanical ventilator on patients with COVID-19," *Brazilian Journal of Anesthesiology*, vol.70, no. 2, pp. 184–185, Apr. 2020.

[2] K. Shimizu, G. Wharton, and H. Sakamoto, "Resurgence of covid-19 in Japan," *British Medical Journal*, vol. 370, no. 322, pp. 1–2, Aug. 2020.

[3] C. Schutze, P. Dietrich, A. Schossland, *et al*., "Application of monitoring methods for remote detection of atmospheric CO2 - concentration levels during a back-production test at the ketzin pilot site," *Energy Procedia*, vol. 76, pp. 528–535, Aug. 2015.

[4] C. Lin, Y. Li, Z. Yuan, A. K. H. Lau, C. Li, and J. C. H. Fung, "Using satellite remote sensing data to estimate the high-resolution distribution of ground-level PM2.5," *Remote Sensing of Environment*, vol. 156, pp. 117–128, Jan. 2015.

[5] D. Yang, E. Yurtsever, V. R. K. A. Redmill, and U. Ozguner, "A Vision-based social distancing and critical density detection system for COVID-19," arXiv:2007.03578v2, July 2020.

[6] W. Taylor, Q. H. Abbasi, K. Dashtipour, S. Ansari, S. A. Shah, A. Khalid, and M. A. Imran, "A review of the state of the art in Non-ContactSensing for COVID-19," *MDPI Sensors*, vol. 20, no. 5665, pp. 1–19, Oct. 2020.

[7] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data gathering algorithm in sensor networks using energy metrics," *IEEE Trans. Parallel Distib. Syst.*, vol. 13, no. 9, pp. 924–935, Sept. 2002.

[8] X. Fan and Y. Song, "Improvement on LEACH protocol of wireless sensor network," in *Proc. International Conference on Sensor Technologies and Applications*, Oct. 2007, pp. 260–264.

[9] N. Komuro, H. Habuchi, and T. Tsuboi, "Nonorthogonal CSK/CDMA with received-power adaptive access control scheme," *IEICE Trans. Fundamentals*, vol. E91-A, no. 10, pp. 2779–2786, Oct. 2008.

[10] K. Kobayashi, A. Nakamura, K. Ohno, and M. Itami, "Improving performance of DS/SS-IVC scheme based on location oriented PN code allocation," *IEICE Trans. Fundamentals*, vol. E99-A, no. 1, pp. 225–234, Jan. 2016.

[11] C. Y. Luo, N. Komuro, K. Takahashi, H. Kasai, H. Ueda, and T. Tsuboi, "Enhancing QoS provision by priority scheduling with interference drop scheme in multi-hop ad hoc network," in *Proc. IEEE Global Communication Conference (GLOBECOM)*, Dec. 2008, pp. 1321–1325.

[12] J. Ma, H. Sekiya, A. Nagasaki, N. Komuro, and S. Sakata, "MAC protocol for Ad Hoc networks using smart antennas for mitigating hidden and deafness problems," IEICE *Transactions on Communications*, vol. E95-B, no. 11, pp. 3545–3555, Nov. 2012.

[13] C. T. Sony, C. P. Sangeetha, and C. D. Suriyakala, "Multi-hop LEACH protocol with modified cluster head selection and TDMA schedule for wireless sensor networks," in *Proc. Global Conference on Communication Technologies (GCCT)*, Apr. 2015, pp. 539–545.

[14] A. P. Plageras, K. E. Psannis, K. E. C. Stergiou, H. Wang, and B. B. Gupta, "Efficient IoT-based sensor BIG data collection-processing and analysis in smart buildings," *Future Generation Computer Systems*, vol. 82, pp. 349–357, May 2018.

[15] S. D. T. Kelly, N. K. Suryadevara, and S. C. Mukhopadhyay, "Towards the implementation of IoT for environmental condition monitoring in homes," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3846–3853, Oct. 2013.

[16] J. Byun, B. Jeon, J. Noh, Y. Kim, and S. Park, "An intelligent self-adjusting sensor for smart home services based on ZigBee communications," *IEEE Trans. Consumer Electronics*, vol. 58, no. 3, pp. 794–802, Mar. 2012.

[17] K. Gill, S. H. Yang, F. Yao, and X. Lu, "A ZigBee-based home automation system," *IEEE Trans. Consumer Electronics*, vol. 55, no. 2, pp. 422–430, Feb. 2009.

[18] Z. Weixing, D. Chenyun, and H. Peng, "Environmental control system based on IoT for nursery pig house," *Trans. Chinese Society of Agricultural Engineering*, vol. 28, no. 11, pp. 177–182, June 2012.

[19] Y. Shiroishi, K. Uchiyama, and N. Suzuki, "Society 5.0: For human security and well-being," in *Computer*, vol. 51, no. 7, pp. 91–95, July 2018.

[20] J. Hong and T. Ohtsuki, "A state classification method based on space-time signal processing using SVM for wireless monitoring system," *Proc. IEEE PIMRC*, Sept. 2011.

[21] Y. Tao, H. Chen, and C. Qiu, "Wind power prediction and pattern feature based on deep learning method," in *Proc. IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, Dec. 2014, pp. 1–4.

[22] N. Komuro, T. Hashiguchi, K. Hirai, and M. Ichikawa, "Development of wireless sensor nodes to monitor working environment and human mental conditions," in *Proc. International Conference on IT Convergence and Security, Lecture Notes in Electrical Engineering*, vol. 712, Dec. 2020.

Nobuyoshi Komro was born in Kitaibaraki, Ibaraki, Japan in 1977. He received the B.E., M.E., and Ph. D degrees in Information Science from Ibaraki University, Hitachi-shi, Japan, in 2000, 2002, and 2005, respectively. He is now with Chiba University as an Associate Professor with the Institute of Management and Information Technologies. He was with Rutgers University as a visiting scholar from May 2012 to April 2013. His research interests include sensor network protocol and wireless/optical CDMA system.