Indoor/Outdoor Route Estimation Method Based on Global Map Matching Using BLE Beacons and GPS

Shinsuke Kajioka*, Takanori Saito† and Daisuke Yamamoto‡
Department of Computer Science
Nagoya Institute of Technology
Gokisocho, Showa, Nagoya, Aichi 466-8555, Japan
*kajioka@nitech.ac.jp, †t.saito.536@nitech.jp, ‡daisuke@nitech.ac.jp

***Abstract*—In recent years, devices, such as smartphones and Bluetooth Low Energy (BLE) beacons, equipped with BLE functionality have become increasingly widespread, resulting in a proliferation of services that use BLE beacons to estimate location. However, most existing research on location estimation using BLE beacons is limited to indoor locations, and few methods are available to accommodate both indoor and outdoor locations. This study proposes a method that combines data from BLE beacons and Global Positioning System (GPS) to estimate continuous path of human movement across both indoor and outdoor areas. This enables not only highly accurate estimation of indoor and outdoor routes but also flexible route estimation by expanding "indoor" and "outdoor" routes to include "semi-outdoor" routes: for example, supplementary BLE beacons are placed near high-rise buildings or in semi-outdoor areas where GPS signals are weak to improve the accuracy of estimation. For achieving such estimation, the proposed method joints maps on adjacent floors or indoor/outdoor. To validate the effectiveness of the proposed method, a prototype system for achieving this functionality is developed and experiment is conducted. The results of the experiment indicates that the proposed method produces better results of route estimation.**

***Index Terms*—map matching, BLE beacon, GPS, geographical information systems**

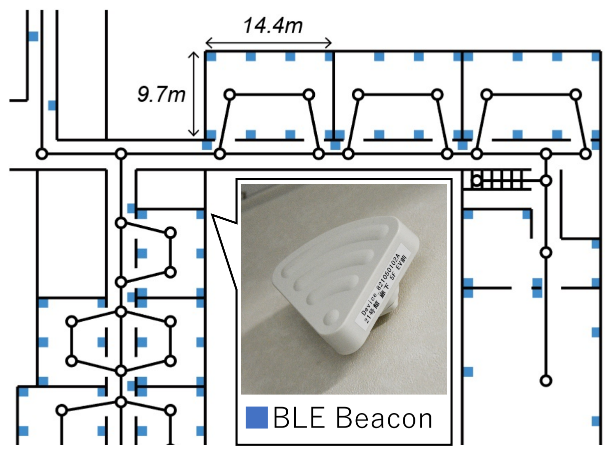

Fig. 1. A floor plan illustrated with locations of installed BLE beacons and candidate paths. A total of 1,600 BLE beacons are installed in all classrooms on campus and an attendance management system uses installed locations of BLE beacons.

I. Introduction

Devices, such as smartphones and beacons, equipped with Bluetooth Low Energy (BLE) functionality have become increasingly popular in recent years. As a result, BLE-based services, particularly location-estimation services using BLE beacons (hereafter referred to as "beacons") are steadily gaining popularity. As shown in Figure 1, the Nagoya Institute of Technology, to which the authors belong, has installed approximately 1,600 BLE beacons in all classrooms and hallways on campus [1], providing a smartphone service that can manage class attendance by receiving BLE beacon signals. BLE beacons are also expected to operate reliably even during power outages and disasters because of their low installation costs, stand-alone operation, low power consumption, and ability to operate for more than five years on dry cell batteries. Owing to these factors, studies estimating indoor location using BLE beacons have proliferated.

As a method of beacon-based location estimation, we have proposed a global map matching method [2] that uses beacons to estimate users' routes and stays. This is a route-and-stay estimation method that uses two networks, a route network for schematically representing a person's movement path and a beacon network for schematically representing the location relationship of beacons. By integrating and filtering beacon signals received using a smartphone, the system estimates users' routes and stays. On the contrary, most existing BLE beacon-based location estimation methods, including our method, are limited to indoor locations. BLE beacons have also received attention in estimating pedestrian traffic and can be used to analyze people flows for marketing, urban planning, evacuation drills, etc. For example, if there is an evacuation drill in an urban area, people will move to evacuation centers (parks, etc.) through various locations, such as building interiors, underground malls, and outdoors. If these indoor/outdoor routes could be estimated with high accuracy, this would contribute to better evacuation and urban planning.

At present, most research on location estimation using BLE beacons is for indoor use; however, there is little research on integration BLE beacons with Global Positioning System (GPS), which is used outdoors. Studies using BLE beacons typically estimate position considering the radio reception strength as input data, while studies using GPS typically

This work was supported by JSPS KAKENHI Grant Numbers JP19H04115 and JP21K19766.

 ISBN: 978-1-68558-079-7

estimate routes on the road network considering latitude and longitude as input data (routes are more abstract than a series of positions, and depending on the application, more useful). The combination of GPS and BLE beacons is inherently difficult because of the different data formats, methods, and estimation targets.

This study proposed a method that combines BLE beacon and GPS data to estimate an overall optimal path of human movement across both indoor and outdoor areas. This will not only enable highly accurate estimation of indoor and outdoor routes but also enable flexible route estimation by adding "semi-outdoor" to "indoor" and "outdoor" routes, for example, by placing supplementary BLE beacons near high-rise buildings or in semi-outdoor areas where GPS signals are weak, to improve estimation accuracy.

While it is standard practice to estimate a route after estimating location, we propose a method that estimates routes directly with high accuracy without estimating location. In other words, by specializing on route estimation, the system provides high accuracy and solves various problems that arise in practical use. This may contribute to the development of geographic information systems, including positioning information.

The rest of this paper is organised as follows: Section II shows related work on route estimation based on GPS data, Section III describes the definitions of the proposed system, Section IV discusses the proposed route estimation method, Section V shows evaluative experiment of the proposed method, and Section VI concludes this study.

II. Related Work

There has been significant research on route estimation based on GPS data. There are two main methods: one based on incremental map matching and the other based on global map matching.

Brakatsoulas et al. [3] provides an example of an incremental map-matching-based method. Two similarity indices, distance and orientation, were used to evaluate candidate links. Lou et al. [4] provides an example of a global map matching method. This study proposed ST-matching, which can perform high-accuracy map matching under low sampling rates with sampling intervals of 2 min or more. Score matching considers spatial and temporal scores and applies the Viterbi algorithm based on the scores for each candidate point and scores between candidate points. Newson and Krumm [5] proposed a globally optimal map matching based on a hidden Markov model.

In [6], a smartphone application was created to conduct a stamp rally at an event site, using booths scattered throughout the site as checkpoints. To realize user behavior analysis, they estimated indoor and outdoor routes using GPS and beacons. However, it lacked the high-level integration of the method proposed herein.

In [2], the authors conducted a study on beacon-based global map matching to estimate users' travel routes and stays. The authors proposed BST-matching, an improved version of ST-matching [4] for beacons, which determines candidate user locations at a given time based on beacon signal strength and uses the Viterbi algorithm for route estimation.

In [7], beacons were used for route estimation indoors, and GPS was used outdoors where beacons were not installed to extend the scope of possible route estimation.

Ito and Kawaguchi [8] described the process of estimating a user's indoor and outdoor travel paths using the optimal available positioning technology for the surrounding environment in which the user was moving. By defining in advance which positioning technologies are available in the road network and switching the information used for estimation, it becomes possible to use complex information for route estimation.

Several methods based on machine learning have also been proposed. Xiao et al. [9] proposed a method for location estimation using BLE beacons in 3D space using Auto Encoder. Urano et al. [10] proposed a method for indoor BLE beacon location estimation based on Long Short-Term Memory (LSTM). Unlike our work, these do not mention the fusion of BLE beacons and GPS.

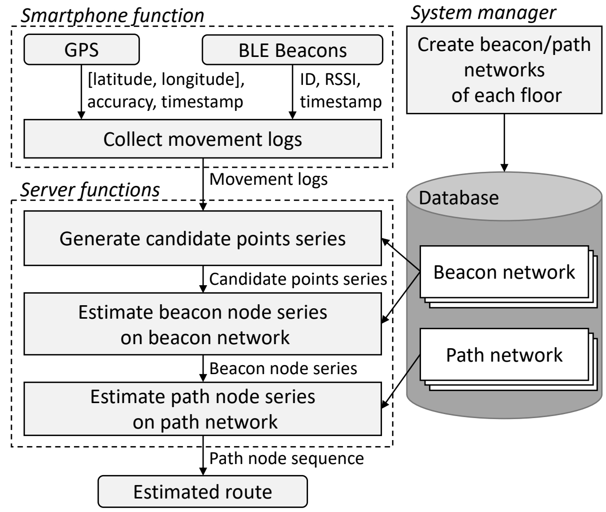

Fig. 2. Route estimation procedure of the proposed system. A smartphone function collects both GPS data and BLE beacons as movement logs. Server functions process the movement logs and derive a path node sequence as an estimated route. System manager should create beacon/path networks beforehand.

III. Proposed System

This section presents the definitions of the movement log and the path and beacon network.

A. Configuration of the Proposed System

Figure 2 shows the procedure structure of the proposed system. System manager should prepare beacon and path networks of each floor or exterior beforehand. First, the log collection function collects movement logs from GPS and BLE beacon signals on a smartphone. The collected movement logs are stored and carried them to the route estimation server.

TABLE I
BEACON LOG FORMAT.

ItemName	Type	Description
Beacon_ID	integer	Beacon ID
RSSI	double	RSSI
Timestamp	ISODate	Timestamp

TABLE II
GPS LOG FORMAT.

ItemMame	Type	Description
latlon	Position	Location coordinates expressed as latitude and longitude
Accuracy	double	Accuracy
Timestamp	ISODate	Timestamp

The route estimation server converts the movement logs into a time-series of sets of candidate points, which is then converted into a series of beacon nodes in the beacon network. Subsequently, the series of beacon nodes on the beacon network is transformed into a series of path nodes on the path network, and the estimated route of user movement is output.

B. Movement Logs

In this study, to estimate routes, BLE beacon data and GPS data are stored as movement logs using smartphones carried by users.

The movement log has two fragments of information in chronological order: the beacon log and GPS log. The beacon log format is presented in Table I, and the GPS log format is presented in Table II. The beacon log consists of the beacon's ID, Received Signal Strength Indicator (RSSI), and a timestamp. The GPS log consists of latitude and longitude coordinates, accuracy, and a timestamp.

C. Path and Beacon Network

Similar to our previous method [2], this study uses two networks, path and beacon, as shown in Figure 3. A beacon network represents beacon installation as a graph, with nodes located at beacon positions. The path network represents the actual candidate routes that users will take, regardless of the beacon position.

The conventional method deals only with indoor path and beacon networks, whereas the proposed method necessitates the construction of both indoor and outdoor networks. In this study, indoor and outdoor networks were created separately and connected using stairway nodes and ingress/egress nodes, as described below. These nodes function as connecting nodes in the conventional method [2].

The network is edited using our previously developed indoor/outdoor network editing system [11].

Stairway nodes are used to connect building floors. An example of a stairway node is shown in Figure 4. The corresponding stairway nodes on each adjacent floor have the same coordinate. They are installed at the landing of a staircase in an indoor map and have the effect of connecting maps on adjacent floors.

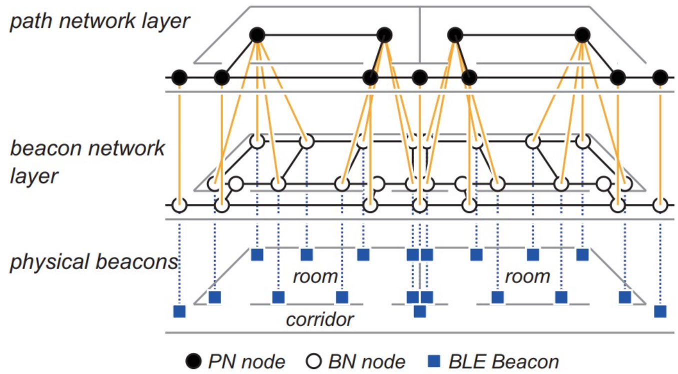

Fig. 3. Path and beacon network. "PN node" and "BN node" indicate a node on path network layer and beacon network layer, respectively.

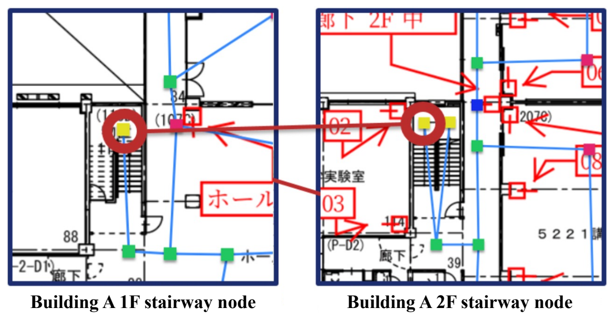

Fig. 4. Example of a stairway node. The red circle indicates the stair node connecting the first and second floors.

Ingress/egress nodes are nodes that connect indoor networks to outdoor networks. An example of corresponding ingress/egress nodes are shown in Figure 5. The corresponding ingress/egress nodes on the same location of each map have the same coordinate. Ingress/egress nodes are positioned between building interiors and exteriors to connect corresponding maps on adjacent indoor/outdoor.

IV. ROUTE ESTIMATION METHOD

In this section, we propose a method to estimate indoor and outdoor routes taken by users using GPS and BLE beacon movement logs as well as the path and beacon network.

A. Combination and Filtering of Movement Logs

In general, BLE beacon signals and GPS signals are prone to interference from buildings and obstacles. Therefore, their accuracy is likely to vary based on location and time of day.

Therefore, in this section we present a method for combining and filtering movement logs as a preprocessing step before performing the estimation.

When combining movement logs, logs are arranged in chronological order from beginning to end, and from the resulting divided i logs, only the information with the highest accuracy is retained. The combination process is shown below: Note that $i = 3$ in this study.

1) The movement log was split into i parts in chronological order.

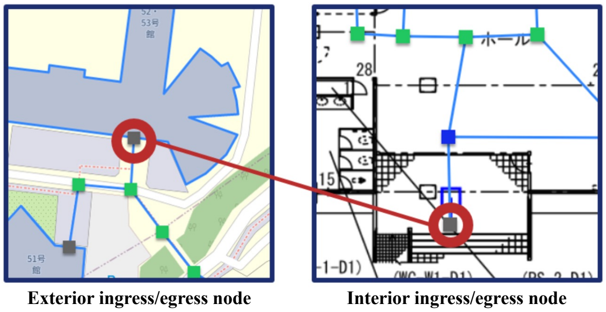

Fig. 5. Example of an ingress/egress node. The red circle indicates the ingress/egress node connecting the outdoor and indoor.

2) If there were multiple beacon logs with the same major value among the divided i beacon logs, the one with the highest RSSI value was selected.
3) From the separated i GPS logs, the log with the highest accuracy was selected.
4) The i movement logs were combined, and the time after combination was considered to be the newest time.

Subsequently, we explain filtering: Filtering removes movement logs below a specified threshold from the combined movement logs. The filtering procedure is shown below: The $RSSI_{type}$ value was set to -75 when $type$ is ‘indoor’ and -79 when $type$ is ‘corridor’.

1) The type of the beacon location was obtained from the information in the movement log, based on a correspondence table.
2) If the RSSI in the beacon log was lower than that of the threshold $RSSI_{type}$, the information was considered unreliable and was removed from the movement log.
3) If the accuracy in the GPS log was lower than that of the threshold ACC_{type}, it was removed from the movement log.

B. Creating a Candidate Point Cloud Series

This section describes the method for converting the movement logs to candidate point cloud series using the combination and filtering processes. A candidate point set represents a set of nodes in the beacon network that can exist at time t, and a candidate point set series represents a series of data that arranges candidate point sets in chronological order.

Two types of movement logs are included: beacon and GPS. The candidate point cloud therefore includes two groups of candidate points, one generated from beacon logs and the other from GPS logs.

The following is the procedure for converting beacon logs to candidate point clouds: At time t, beacon nodes corresponding to the top k beacons with the highest RSSI form the candidate point group.

The following is the procedure for converting GPS logs to candidate point clouds: Beacon nodes are obtained within a circle of radius R centered on the latitude and longitude coordinates estimated using GPS at time t. Subsequently, the top k beacon nodes closest to the estimated location are selected as the candidate point group at that time.

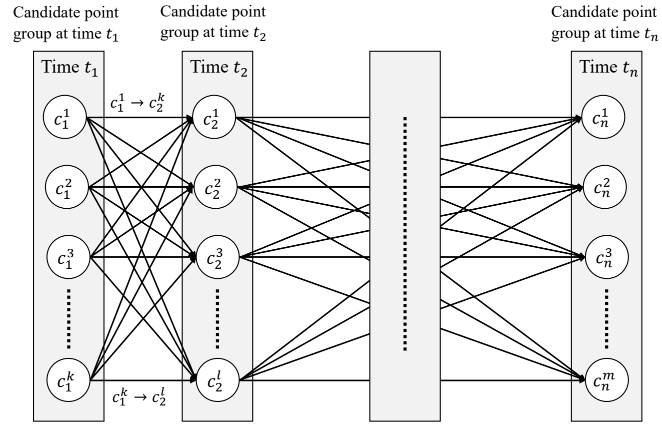

Fig. 6. Example candidate point graph. The graph is used for determining estimated path.

Finally, the set of candidate points generated from both beacon and GPS logs are combined to determine the set of candidate points at time t.

C. Estimating Routes From Candidate Point Cloud Series

In this section, we propose a method based on BST matching [2] for estimating a user’s travel path from a candidate point cloud series and beacon network.

Figure 6 shows an overview of determining estimated path. The scores between each candidate point and candidate point at time t are computed. For each candidate point group, the path with the lowest overall score P is the path obtained by the Viterbi algorithm; this is the estimated path.

Score P is expressed as follows:

$$P = argmax\,(c_1, c_2, \cdots, c_n) \left\{ N\left(c_1^j\right) + \sum_{i=2}^{n} F\left(c_{i-1}^t \rightarrow c_i^s\right) \right\}$$

where $N\left(c_i^j\right)$ and $F\left(c_{i-1}^t \rightarrow c_i^s\right)$ represent functions that computes scores of candidate points and scores between candidate points, respectively. By calculating beacon-based and GPS-based candidate points together, the candidate point with the highest estimation accuracy can be selected for each point in time, regardless of the type. Candidate point score $N\left(c_i^j\right)$ represents the proximity between the observation point and candidate point c_i^j. The higher the proximity, the higher the score, and vice versa. The distance d_B between the beacon and observer can be obtained from the RSSI of the beacon received by the observer using the Friis transmission equation, defined as follows:

$$d_B = 10^{\frac{1}{20}(RSSI_{max} - RSSI)}$$

where $RSSI_{max}$ denotes the RSSI value when a radio wave is received at a distance of $1\,\mathrm{m}$ from the beacon. This value varies based on the location of the beacon; in this study, it is set at -55 for beacons in rooms and -60 for beacons in corridors.

TABLE III
NUMBER OF BEACONS INSTALLED IN THE EXPERIMENTAL ENVIRONMENT.

Location	Number of installed beacons
Bldg. 52, 1st floor	11
Bldg. 52, 2nd floor	32
Bldg. 23, 1st floor	25
Bldg. 23, 2nd floor	22

The beacon-based score $N\left(c_i^j\right)$ of a candidate point c_i^j is the reciprocal of distance d_B and is expressed as follows:

$$N\left(c_i^j\right) = \frac{1}{d_B}$$

The GPS-based score $N\left(c_i^k\right)$ of a candidate point c_i^k is the reciprocal of distance d_G between the estimated position at time t_i and the candidate point c_i^k, expressed as follows:

$$N\left(c_i^k\right) = \frac{1}{d_G}$$

Intra-candidate score $F_t\left(c_{i-1}^t \rightarrow c_i^s\right)$ indicates the travel speed between time t_{i-1} to t_i on the estimated path at average walking speed v_a. This value approaches zero as average walking speed increases.

$$F_t\left(c_{i-1}^t \rightarrow c_i^s\right) = \begin{cases} 1 & (v \leq v_a) \\ \dfrac{v_a t}{dist} & (v > v_a) \end{cases}$$

where $dist$ denotes the distance between two points on the network and t denotes elapsed time. v_a was set at $2.0\,\mathrm{m/s}$ for this study. This suppresses estimation of routes that would be considered unnatural at walking pace.

V. EVALUATIVE EXPERIMENT

In this section, we conducted an experiment to evaluate the accuracy of route estimation and generation time for validating the efficacy of the proposed method.

A. Experiment Conditions

In this experiment, participants were asked to carry a smartphone and walk along a designated route to verify the accuracy of the actual route and estimated route. The experiment was conducted in Buildings 52 and 23 of the Nagoya Institute of Technology. The Buildings 52 and 23 have 3 and 4 floors, respectively. Table III lists the total number of beacons installed in each floor. The routes used in the experiment spanned indoor and outdoor areas: three routes were set up, as shown in Figure 7. Movement logs were collected 6 times for each route.

The following five methods of comparison were used:

m1 (Proposed method): Path estimation based on global map matching using beacon and GPS

m2 Route estimation that prioritizes beacons with ancillary use of GPS

m3 Route estimation that prioritizes GPS with ancillary use of beacons

m4 Route estimation using beacons only

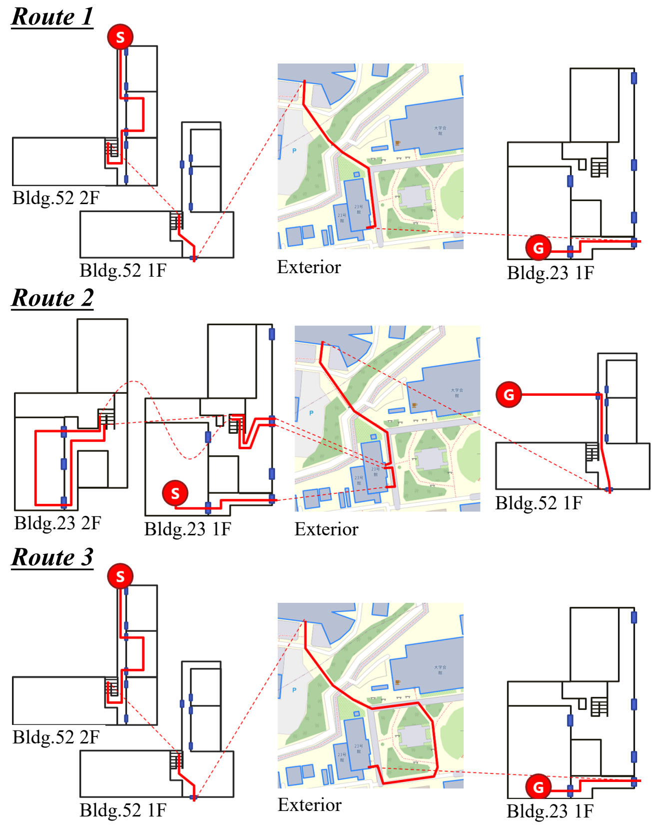

Fig. 7. Three routes used in the experiment.

m5 Route estimation using GPS only

m1 is our proposed method, which optimally uses GPS and beacons together. By contrast, **m2** and **m3** use both beacon and GPS data, however give preference to one type over the other. **m4** and **m5** use only one of the data types.

Precision, Recall, and F-score for estimation accuracy and generation time were used as evaluation metrics. Considering V_1 as the set of links in the routing network estimated from movement logs and V_2 as the set of links actually traversed, Precision P_V, Recall R_V, and F-score F_V are obtained as follows:

$$P_V = \frac{|V_1 \cap V_2|}{|V_1|}, \quad R_V = \frac{|V_1 \cap V_2|}{|V_2|}, \quad F_V = \frac{2}{\left(\frac{1}{P_V} + \frac{1}{R_V}\right)}$$

B. Experiment Results

We conducted an experiment and got results. Figure 8 shows the averages of the experiment results. The proposed method, **m1**, with an average F-score of 0.87, outperformed the other methods (0.84 for **m2**, 0.83 for **m3**, 0.76 for **m4**, and 0.67 for **m5**). The proposed method, **m1**, also yielded the best results for precision and recall.

m1, **m2**, and **m3** are methods that use both GPS and beacon data, however **m1** yielded the best results. This suggests that, compared to the methods prioritizing beacons (**m2**) and GPS

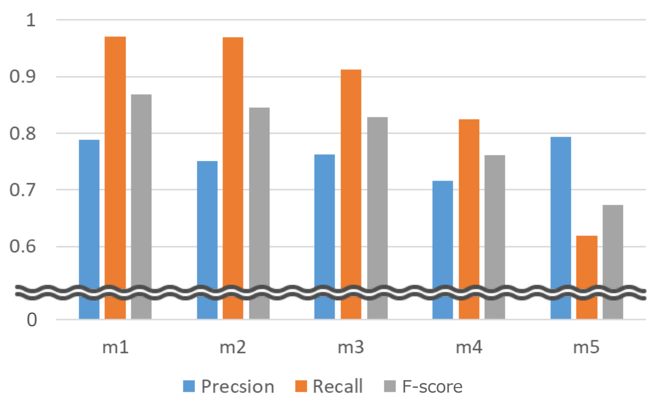

Fig. 8. Average results from the experiment.

TABLE IV
RESULTS OF A T-TEST. COMPARISON OF EACH METHOD AND **M1**.

	Route 1	Route 2	Route 3
m2	0.007	0.051	0.019
m3	0.054	0.009	0.092
m4	0.007	0.003	0.002
m5	0.000	0.012	0.000

(**m3**), it is more effective to use data based on the evaluation equation, as in our proposed method.

A t-test was conducted for each route to observe if there was a significant difference between **m1** and the other methods. Table IV lists the results of these t-tests. In all cases, the results were below the 0.1 significance level, supporting the superiority of the proposed method **m1**. The significance level was not below 0.05 for several items. This can be attributed to the fact that only six movement logs were collected. In future, we would like to increase the number of experiments, repeating the experiment to show significance.

Tables V and VI list the execution time results and the number of nodes at that time, respectively. Table V shows that **m4**, which uses only beacons for estimation, took the shortest computation time, while **m1** took the longest. Note that **m4** also has the smallest number of nodes in the candidate point graph used for estimation of all methods, while **m1** has the highest number of nodes, as shown in Table VI. Thus, it can be concluded that the computation time required was based on the number of nodes used for estimation.

Of the five estimation methods, **m1**, **m3**, and **m5**, which required longer computation timed, are likely to generate large numbers of candidate points using GPS data. The number of nodes in the candidate point graph therefore increased compared to **m2** and **m4**, which have shorter computation times, resulting in longer computation times.

VI. CONCLUSION AND FUTURE WORK

In this study, we proposed a route estimation method based on global map-matching that uses both BLE beacons and GPS. By using BLE beacons with GPS, it is possible to estimate routes that cross indoor and outdoor areas. We also implemented a prototype system for achieving this feature and conducted an evaluative experiment using the prototype system. The results of the experiment indicated that using the BLE beacon and GPS-based scores produced better results. In the future, we will attempt to increase the accuracy of the method. In addition, because the evaluation in this study was conducted using only a small amount of data, we want to conduct a larger-scale demonstration experiment in future.

TABLE V
COMPUTATIONAL TIME TAKEN TO ESTIMATE.

[ms]

	m1	**m2**	**m3**	**m4**	**m5**
Route 1	13575	5993	12205	1164	11701
Route 2	24889	8326	19315	1018	23903
Route 3	23655	14663	20383	841	20423

TABLE VI
NUMBER OF NODES IN THE CANDIDATE POINT GRAPH USED FOR ESTIMATION.

	m1	**m2**	**m3**	**m4**	**m5**
Route 1	508	331	469	110	398
Route 2	937	476	848	152	785
Route 3	803	607	757	92	671

REFERENCES

[1] S. Kajioka, T. Mori, T. Uchiya, I. Takumi, and H. Matsuo, “Experiment of indoor position presumption based on RSSI of bluetooth LE beacon,” in *Proceedings of the IEEE 3rd Global Conference on Consumer Electronics (GCCE 2014)*, pp. 337–339, 2014.

[2] D. Yamamoto, R. Tanaka, S. Kajioka, H. Matsuo, and N. Takahashi, “Global map matching using BLE beacons for indoor route and stay estimation,” in *Proceedings of the 26th ACM SIGSPATIAL*, pp. 309–318, 2018.

[3] S. Brakatsoulas, D. Pfoser, R. Salas, and C. Wenk, “On map-matching vehicle tracking data,” in *Proceedings of VLDB*, pp. 853–864, 2005.

[4] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang, and Y. Huang, “Map-matching for low-sampling-rate gps trajectories,” in *Proceedings of SIGSPATIAL GIS 2009*, pp. 352–361, 2009.

[5] P. Newson and J. Krumm, “Hidden Markov map matching through noise and sparseness,” in *Proceedings of the ACM SIGSPATIAL 2009*, pp. 336–343, 2009.

[6] K. Urano, K. Hiroi, K. Kaji, and N. Kawaguchi, “A location estimation method using BLE tags distributed among participants of a large-scale exhibition,” in *Adjunct Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing Networking and Services*, pp. 124–129, 2016.

[7] M. Ficco, F. Palmieri, and A. Castiglione, “Hybrid indoor and outdoor location services for new generation mobile terminals,” *Personal and Ubiquitous Computing*, vol. 18, pp. 271–285, 2014.

[8] S. Ito and N. Kawaguchi, “Wireless LAN based hybrid positioning system using bayesian inference and access point selection,” *Transactions of the Institute of Electrical Engineers of Japan. C*, vol. 126, no. 10, pp. 1212–1220, 2006.

[9] C. Xiao, D. Yang, Z. Chen, and G. Tan, “3-D BLE indoor localization based on denoising autoencoder,” *IEEE Access*, vol. 5, pp. 12751–12760, 2017.

[10] K. Urano, K. Hiroi, T. Yonezawa, and N. Kawaguchi, “An end-to-end BLE indoor localization method using LSTM,” *Journal of Information Processing*, vol. 29, pp. 58–69, 2021.

[11] S. Kajioka, T. Saito, D. Yamamoto, and N. Takahashi, “A web-based geospatial adjacency graph construction tool for indoor/outdoor beacons,” in *Proceedings of the IEEE 10th Global Conference on Consumer Electronics (GCCE 2021)*, pp. 1–2, 2021.

 ISBN: 978-1-68558-079-7