Actual Consumption Estimation Algorithm for Occupancy Detection using Low Resolution Smart Meter Data

Shunichi Hattori and Yasushi Shinohara
Central Research Institute of Electric Power Industry, 2-11-1 Iwadokita, Komae-shi, Tokyo, Japan

Keywords: Occupancy Detection, Smart Meter, Electricity Consumption, Non-intrusive Load Monitoring.

Abstract: This paper proposes an actual consumption estimation algorithm that achieves highly accurate occupancy detection using electricity consumption data derived from smart meters. In Japan, electricity consumption data on households will soon be available because smart meters, which enable electric power companies to monitor how much electric power people are using in each household, have been installed in all households. Occupancy detection is a major technique that leverages electricity consumption data and can be applied to various services such as ambient assisted living, sales promotions, and peak load shifting. However, it is difficult to conduct high-accuracy occupancy detection using electricity consumption data automatically derived from smart meters because of their low resolution: 30-min intervals and 100 Wh increments. An actual consumption estimation algorithm is therefore proposed to generate data that reflects the characteristics of a household's state from low-resolution smart meter data. Occupancy detection is implemented using the estimated consumption data, which are generated by the proposed algorithm, and the results of experiments show that its performance is improved compared to the result obtained using raw smart meter data.

1 INTRODUCTION

In Japan, the restructuring of the electric power market began in the 1990s and full deregulation of the electric power market started in April 2016. Whole new services creating additional value for the sales departments are desired because this deregulation brings intensive competition from new entrants for the existing electric power companies.

At the same time, network-connected and remote-controlled electricity meters called *smart meters* have been installed in all households for the purpose of automated meter reading. The standard communication method with a smart meter is called the *A-route* and it enables electric power companies to monitor how much electric power people are using in each household. That is, electricity consumption data for all households are becoming available without any additional cost through the use of A-route. Analysis targeting the consumption data derived from the A-route is therefore a promising approach for developing useful and competitive services for electric power companies.

Occupancy detection (Chen et al., 2013; Kleiminger et al., 2015; Molina-Markham et al., 2010) leveraging electricity consumption data is known as one of the major non-intrusive load monitoring techniques (Hart, 1992). This technique was mainly developed in Europe and the United States for the purpose of efficient control of Heating, Ventilation, and Air Conditioning (HVAC) systems. Moreover, occupancy detection is also considered to be applicable in various useful services such as ambient assisted living (AAL), sales promotions, peak load shifting, and delivery route optimization. However, it is difficult to conduct high-accuracy occupancy detection using electricity consumption data derived from the A-route, because the A-route readout is low resolution data: 30-min intervals and 100 Wh increments (Komatsu and Nishio, 2015; Nomura et al., 2014). The data contain phantom demand variations because of these characteristics and they degrade the performance of occupancy detection using machine learning algorithms.

An *actual consumption estimation algorithm* is therefore proposed to leverage low-resolution A-route readout for occupancy detection. Occupancy detection is implemented using the estimated consumption data generated by the proposed algorithm and its accuracy, precision, and recall are improved better than those of results obtained using raw A-route readout.

Related work about non-intrusive load monitor-

Hattori S. and Shinohara Y.
Actual Consumption Estimation Algorithm for Occupancy Detection using Low Resolution Smart Meter Data.
DOI: 10.5220/0006129400390048
In *Proceedings of the 6th International Conference on Sensor Networks (SENSORNETS 2017)*, pages 39-48
ISBN: 421065/17

ing, occupancy detection, and its applications are reviewed in Section 2. Section 3 proposes the algorithm that estimates actual electricity consumption data from low-resolution A-route readout data. The feasibility of the proposed algorithm is discussed based on the performance of the occupancy detection in Section 4. Section 5 concludes the paper.

2 RELATED WORK

2.1 Non-intrusive Load Monitoring

The approach used to analyze the change of appliance-level electricity consumption from aggregated data is called disaggregation (Froehlich et al., 2011; Zoha et al., 2012). Disaggregation is categorized into non-intrusive load monitoring, which analyzes loads without any sensor or measuring equipment inside buildings except for the aggregated electricity consumption meter (Hart, 1992).

Disaggregation is a well-studied problem and there are many studies that deal with it based on various characteristics and methods such as harmonics (Nakano and Murata, 2007), neural networks (Chang et al., 2010), and factorial hidden Markov models (Batra et al., 2014; Kim et al., 2011), which is an expansion of the hidden Markov model.

It can be said that disaggregation is an approach that is similar to occupancy detection in terms of activity annotation on households or buildings. If such annotated data regarding appliances are obtained, the performance of occupancy detection algorithms should be improve. In addition, more detailed state information such as sleep, bathing, and cooking behavior might be classifiable. However, these approaches require electricity consumption data with high frequency, generally from 1 kHz to 100 MHz, because they try to extract the features of appliance-level consumption based on the characteristic pattern or electrical noise of each appliance (Zoha et al., 2012). This paper does not focus on this approach because the proposed method is designed for electricity consumption data with low resolution such as A-route readout.

2.2 Occupancy Detection

Occupancy detection is a technique that estimates the occupancy state of commercial or residential buildings (e.g., whether a household is occupied or unoccupied). This technique was mainly developed for the purpose of energy-efficiency optimizations for HVAC control and can be categorized into two approaches: intrusive and non-intrusive. In the former approach, various methods using environmental sensors such as motion (passive infrared), CO_2, acoustic information, and doors (magnetic contact) have been proposed (Nguyen and Aiello, 2013). In contrast, electricity consumption data are mainly employed in the latter approach (Chen et al., 2013; Kleiminger et al., 2015; Molina-Markham et al., 2010).

Non-intrusive occupancy monitoring has become a well-studied problem since Molina-Markham et al. suggested that occupancy state can be classified based on aggregated electricity consumption data (Molina-Markham et al., 2010). The advantage of the non-intrusive approach is that it detects occupancy without the need to install any sensors in buildings. Nowadays, it is a major research field in some regions of Europe and the United States, where smart meters have been installed before the rest of the world.

Fig. 1 shows the occupancy detection procedure using electricity consumption data based on classification criteria. Occupancy is generally detected using the following steps.

Step 1: Extract features of a training set $T = \{t_1, t_2, \dots\}$ for discriminating occupancy states and generate a classification criteria based on the features.

Step 2: Input test set $T' = \{t'_1, t'_2, \dots\}$ to the classification criteria generated in Step 1.

Step 3: Obtain the set of classified occupancy states for T'.

Using the procedure shown in Fig. 1, Chen et al. proposed a threshold-based method that classifies occupancy state using simple criteria such as maximum value, standard deviation, and range (the change between maximum and minimum value) (Chen et al., 2013). Beckel et al. employed machine learning algorithms such as support vector machine (SVM) and hidden Markov models for state classification (Kleiminger et al., 2015). They also provide the ECO data set,[1] which includes pairs of 1 Hz electricity consumption data and the occupancy state ground truth of five Swiss households (Beckel et al., 2014).

These studies target electricity consumption data whose frequency ranges from 1-s to 1-min intervals. A method targeting low-resolution data, such as smart meter data in Japan, has not yet been studied. The detailed characteristics of smart meter data communication in Japan, which is called A-route, are described in the next section.

[1] https://www.vs.inf.ethz.ch/res/show.html?what=eco-data [Accessed 4 October 2016]

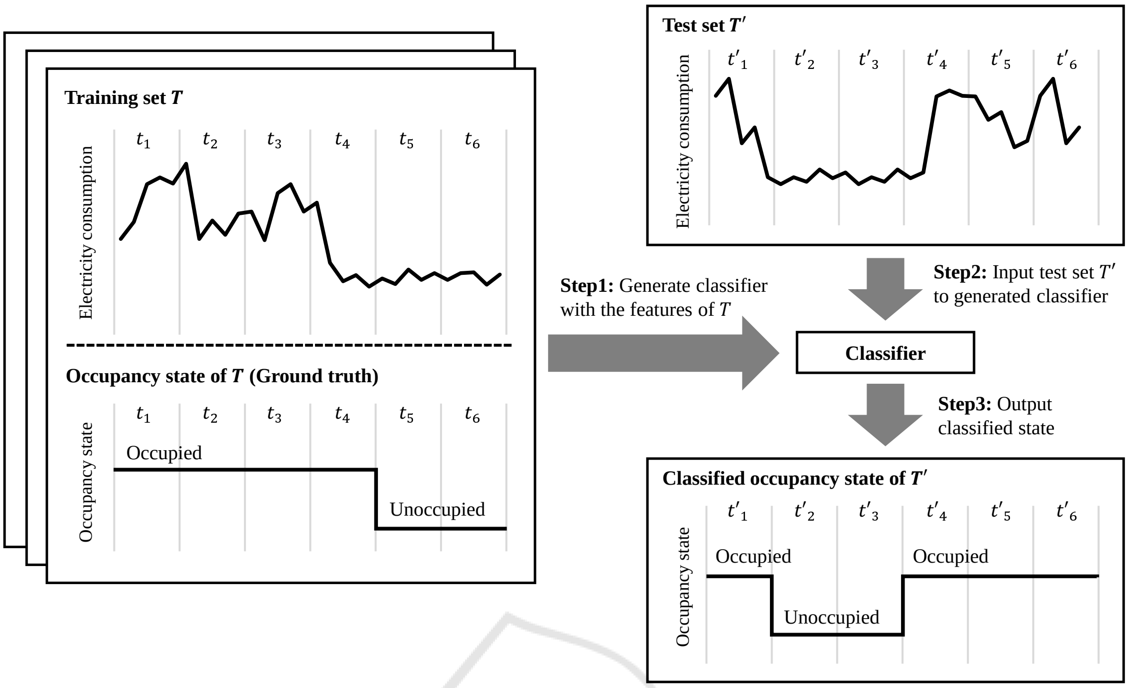

Figure 1: Procedure of occupancy detection from electricity consumption data based on classification criteria.

Table 1: Specification and Characteristics of Smart Meter Data in Japan.

Route	Specification	Advantages	Disadvantages
A-route	• 30-min intervals • 100 Wh increments	• Needs no dedicated equipment • Data on all households are available	• Low-resolution data • Contain phantom power variations
B-route	• From 5-s to 5-min intervals (depending on equipment) • 1 W increments	• High-resolution data	• Need dedicated equipment to obtain consumption data

2.3 Smart Meters

A smart meter is a network-connected and remote-controlled digital meter that monitors electricity consumption online for meter reading. These meters have been installed in some regions of Europe and the United States. In Japan, smart meters will be installed in all households as of 2023. Although the main purpose of the installation is automated meter reading for billing, smart meters also enable electric power companies to obtain time series data regarding electricity consumption for each household. Occupancy detection targeting households employing smart meter data is therefore a promising approach to developing various competitive services.

There are two communication routes, called *A-route* and *B-route*, that are used with smart meters installed in Japan. Table 1 shows the specification and characteristics of these routes. In A-route, electricity consumption data of all households can be obtained automatically through the power distribution sectors of the electric power companies, although they are low frequency and coarse sampling data (Komatsu and Nishio, 2015; Nomura et al., 2014). In contrast, the route that acquires electricity consumption data through the home area network is called B-route. Whereas data with higher resolution are derived from B-route, dedicated equipment such as a home energy management system (HEMS) is required to utilize this route. Hence, the utilization of B-route is currently totally impractical because HEMS is generally expensive equipment (it costs approximately USD 1,500 as of 2016).

Hence, data acquisition using A-route is easy and economical because it needs no extra equipment such as HEMS. However, utilization of A-route data must take into account its specification: low frequency and coarse sampling. As shown in Table 1, A-route is 30-min intervals and 100 Wh increments consumption data. In A-route, the fractions of meter readouts less than 100 Wh are rounded off and carried over to the next readout in 30 min.

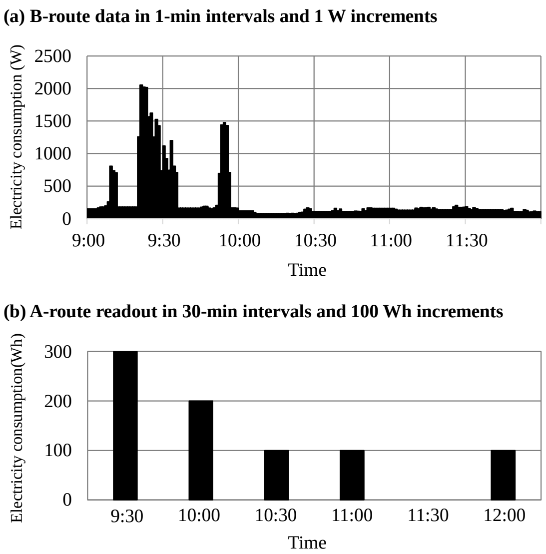

Figure 2: Example of smart meter data: (a) B-route and (b) A-route data.

Figs. 2(a) and (b) show an example of electricity consumption data over the same period and household via B- and A-routes, respectively. The household can be assumed unoccupied after 10:00 because the consumption and its variation are relatively low compared to those before 10:00. However, the A-route readout shown in Fig. 2(b) oscillates between values of 0 and 100 Wh after 10:30 because of the round off and carry over. These characteristics are fatal for occupancy detection, which classifies occupancy state based on features such as standard deviation, because this phantom fluctuation tends to be regarded as the characteristics of an occupied state. A method for estimating actual consumption which reduces the phantom fluctuation during the unoccupied state is therefore needed to detect occupancy with high accuracy using smart meter data.

2.4 Applications of Occupancy Detection

Although occupancy detection has mainly been applied to building automation systems such as HVAC control in Europe and the United States (Nguyen and Aiello, 2013), some additional applications are expected.

AAL.
AAL is a key technology for tackling the problems of an aging society. Rashidi et al. indicated that it is imperative to develop AAL technologies that help older adults in their own homes (Rashidi and Mihailidis, 2013). Non-intrusive AAL is expected to be based on occupancy detection because occupancy detection enables us to monitor the daily behavior of older adults without installing any equipment or medical devices.

Sales Promotion.
Advertisement delivery systems are another promising way to exploit occupancy detection technology. Understanding the lifestyle in each household could enable just-in-time ad serving, although it might elicit a negative response regarding privacy issues.

Peak Load Shifting.
Mitigating electrical peak demand has been a critical issue for electric power companies because it improves the efficiency of electric power generation and reduces generation capacity requirements. For example, some Japanese electric power companies conducted experiments that distributed coupons to people as an incentive to go to commercial facilities when electricity demand is close to the limit of the existing power supply.[2] Occupancy detection is expected to contribute efficient encouragement to people who are always at home during peak time.

Delivery Route Optimization.
Redelivery due to absence accounts for a quarter of the total delivery distance in Japanese courier companies.[3] Delivery route optimization employing household occupancy states is a promising application to reduce redelivery cost and CO_2 emissions.

The actual consumption estimation algorithm in the next section, which detects occupancy with high accuracy using electricity consumption data with low resolution, is proposed as a fundamental technique for realizing these applications.

3 ACTUAL CONSUMPTION ESTIMATION ALGORITHM

In order to estimate occupancy state from A-route data more accurately, an actual consumption estimation algorithm is proposed. This algorithm estimates actual electricity consumption based on cumulative consumption. Figs. 3 and 4 show examples of actual consumption and A-route readout in 30-min intervals, respectively. Although these two figures show the same period and household consumption, A-route

[2] http://www.ft.com/cms/s/0/5670af9e-0b22-11e4-9e55-00144feabdc0.html [Accessed 4 October 2016]

[3] http://www.mlit.go.jp/report/press/tokatsu01_hh_000234.html [Accessed 4 October 2016, written in Japanese]

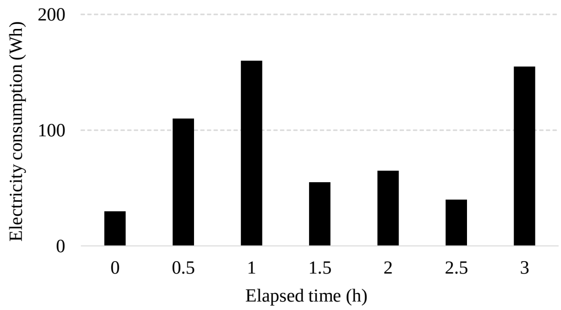

Figure 3: Actual electricity consumption in 30-min and 1 Wh increments.

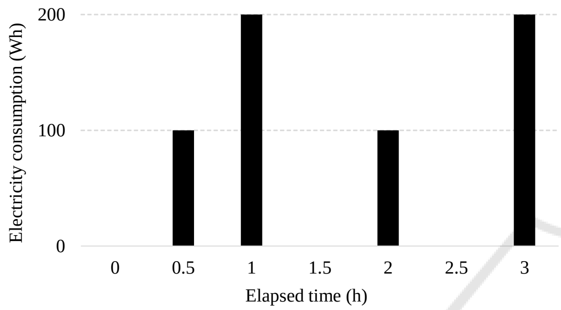

Figure 4: A-route readout in 30-min and 100 Wh increments.

readout data differ from the actual consumption because of the characteristics described in Section 2.3.

In this algorithm, the actual consumption data are estimated using the cumulative consumption of the A-route readout. In Fig. 5, which is converted into cumulative consumption from Fig. 4, two stair-like solid lines are drawn with cumulative A-route readout and the values with 100 Wh added. The lower line shows the lower limit of the A-route readout range. In addition, the upper line shows its upper limit. That is, the area that consists of these two lines can be regarded as the range of true cumulative electricity consumption, which monotonically increases. The dashed line which is the minimal distance within the range area is calculated, as shown in Fig. 5. The estimated actual consumption data are finally calculated as each increase of the dashed line in 30-min intervals.

This problem can be formulated as shown below. In these formulas, the minimum increment of consumption (100 Wh) is denoted by l, A-route readout is x, and estimated consumption is y, respectively.

$$Minimize \quad \sum_{t=1}^{T} \sqrt{1+y_t^2} \tag{1}$$

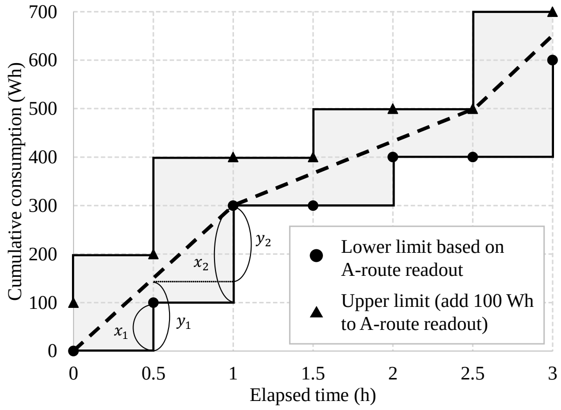

Figure 5: True cumulative electricity consumption and estimated cumulative consumption.

$$subject\ to \quad y_t \geq 0$$
$$\sum_{s=1}^{t} x_s \leq \sum_{s=1}^{t} y_s \leq \sum_{s=1}^{t} y_s + l$$
$$\sum_{t=1}^{T} y_t = \sum_{t=1}^{T} x_t$$

Equation (1) shows that the distance of the estimation line consists of elapsed time and estimated consumption y_t. The distance of the estimation line can be calculated as the sum of the oblique lines, which show the estimated consumption such as y_1 and y_2 in Fig. 5. The dashed line minimizing this distance is determined as the estimated cumulative consumption in the proposed algorithm.

Each constrained condition corresponds to non-negativity, the range of actual consumption, and the consistency of the cumulative consumption. Therefore, the dashed line that fulfills these conditions corresponds to non-negativity cumulative consumption, which is in the area between the two solid lines in Fig. 5.

This problem equals the problem shown in formula (2) and its dual problem in formula (3). Formula (3) is a convex optimization problem called the total variation optimization problem.

$$Minimize \quad \sum_{t=1}^{T} y_t^2 \tag{2}$$

$$subject\ to \quad y_t \geq 0$$
$$\sum_{s=1}^{t} x_s \leq \sum_{s=1}^{t} y_s \leq \sum_{s=1}^{t} y_s + l$$
$$\sum_{t=1}^{T} y_t = \sum_{t=1}^{T} x_t$$

$$Minimize \quad \frac{1}{2}\sum_{t=1}^{T}(y_t - x'_t)^2 + \frac{1}{2}\sum_{t=1}^{T}|y_t - y_{t-1}|^2 \quad (3)$$

$$x'_0 = x_0 + \frac{1}{2}, \quad x'_t = x_t \qquad (t > 0)$$

Optimized cumulative consumption (the minimal-length dashed line) can be calculated with the algorithm shown below. The algorithm sequentially updates the upper and lower lines, which indicate the upper and lower limits of acceptable gradients. The dashed line is fixed when both lines depart from the range of true cumulative consumption. Finally, the starting point is updated based on the end-point of the fixed dashed line.

(1) The starting point and upper and lower lines are determined (see Fig. 6(a)).

(1-1) Starting point: the terminal of the determined dashed line.

(1-2) Upper line: A line connecting the starting point and upper limit of the range of true cumulative consumption.

(1-3) Lower line: A line connecting the starting point and lower limit.

(2) Increment the time by 30 min and update the upper and lower lines (see Fig. 6(b)).

(2-1) When the upper line is larger than the upper limit, update the upper line.

(2-2) When the lower line is smaller than the lower limit, update the lower line.

(3) If both lines depart from the range of true cumulative consumption, the dashed line and starting point are updated (see Figs. 6(c), (d), and (e)).

(3-1) When the lower line is smaller than the upper cumulative consumption, the line connecting the starting point and contact point with the lower limit of the area are determined as the dashed line.

(3-2) When the upper line is larger than the lower cumulative consumption, the line connecting the starting point and contact point with the upper limit of the area are determined as the dashed line.

(3-3) The contact point is determined as the new starting point. The upper and lower lines are updated.

(4) Return to Step 2 unless the time is terminated.

Fig. 7 shows an example of the estimated consumption data, which are generated by the proposed algorithm from the A-route readout in Fig. 4. It can be seen that the estimated data in Fig. 7 are closer to the actual data in Fig. 3 than the A-route readout in Fig. 4.

4 EXPERIMENTS

4.1 Outline of Experiments

Occupancy detection using actual electricity consumption data was conducted to consider the feasibility of the proposed algorithm. The ECO data set (Beckel et al., 2014) referred to in Section 2.2 was used for evaluating performance. This data set consists of the 1 Hz electricity consumption data and occupancy state on five Swiss households from June 2012 to January 2013.

In the experiment, the following four types of electricity consumption data were prepared from the ECO data set for the performance comparison.

B-route.
Actual instantaneous data in 1-min intervals and 1 W increments. The B-route data is the consumption data with the highest resolution in this experiment, and it corresponds to data that could be derived from dedicated equipment such as HEMS.

30-min/1 Wh.
Actual watt-hour data in 30-min intervals and 1 Wh increments. These data correspond to the answer data that the proposed algorithm described in Section 3 aims to estimate.

A-route.
Watt-hour values in 30-min intervals and 100 Wh increments. This data corresponds to the readout derived from a smart meter via A-route.

Estimated.
Watt-hour estimated data in 30-min intervals and 1 Wh increments. The data mean estimated consumption data generated from the A-route readout by the proposed algorithm.

Occupancy detection was conducted every hour. Random forests and SVM, which are known as high-accuracy machine learning algorithms, were employed for classification. The explanatory variables shown in Table 2 are employed as the feature for occupancy detection. Note that the temporal resolution is different from the others in the B-route data.

Although occupancy detection was conducted every hour, its state (occupied or unoccupied) is in 1-s intervals in the data set. The hourly state is defined as follows.

1. When the entire hourly state is unoccupied, the state is regarded as unoccupied.
2. When the hourly state is mixed occupied and unoccupied states, the state is regarded as occupied.
3. When the entire hourly state is occupied, the state is regarded as occupied.

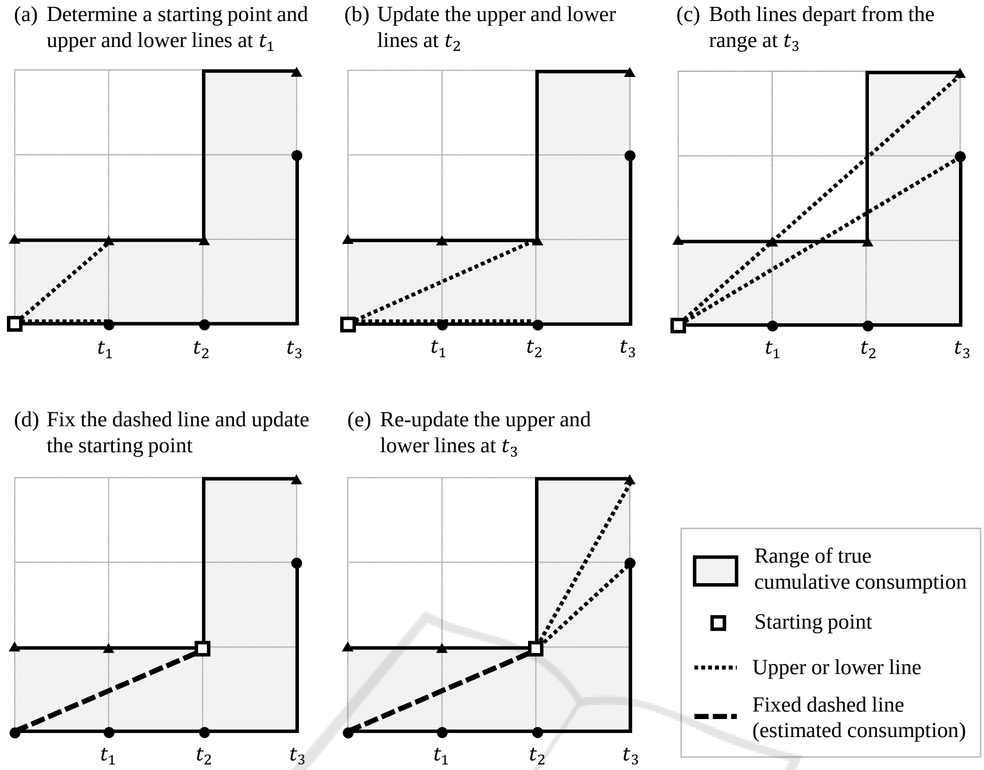

Figure 6: Procedure of the actual consumption estimation algorithm.

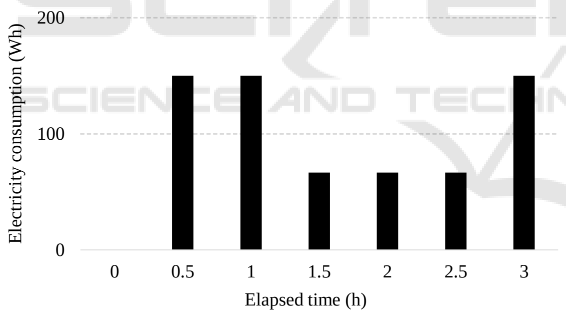

Figure 7: Estimated actual consumption data in 30-min intervals from A-route readout.

Table 2: Explanatory Variables for Occupancy Detection.

Variable	Description
id	Unique ID of household
mean	Average consumption in the target interval
max	Maximum consumption in the target interval
min	Minimum consumption in the target interval
range	Difference between maximum and minimum values
std	Standard deviation in the target interval
hour	Hour when occupancy detection is conducted
temp	Average temperature in the target interval
season	Summer or winter (dummy variable)

That is, the state becomes *occupied* when it includes an occupied state of at least 1 s. The state is used as an objective variable in occupancy detection.

In this experiment, the target hours of occupancy detection are from 6:00 to 21:59 because almost all the states during night and early-morning are occupied in the ECO data set.

4.2 Results of Actual Consumption Estimation

The actual consumption estimation algorithm described in Section 3 was applied to the A-route readout, which was prepared from the ECO data set. Fig. 8 shows the visualized results of the four types data including B-route, 30-min/1 Wh, A-route and estimated consumption. As described in Section 4.1, 30-min/1 Wh consumption data shown in Fig. 8(b) corresponds to the answer that the proposed algorithm aims to estimate. Fig. 8(c) shows that the A-route readout includes sequentially iterated values consisting of 0 and 100 Wh when consumption is relatively low, as described in Section 2.3. In contrast, Fig. 8(d) shows the estimated result, which suppresses the fluctuations and represents the actual consumption more precisely than the raw A-route readout, although it is smoother than the 30-min/1 Wh data.

As described in Section 3, the estimated consumption data are generated based on the dashed line, which is the minimal distance within the range of true cumulative consumption. Given these conditions, the

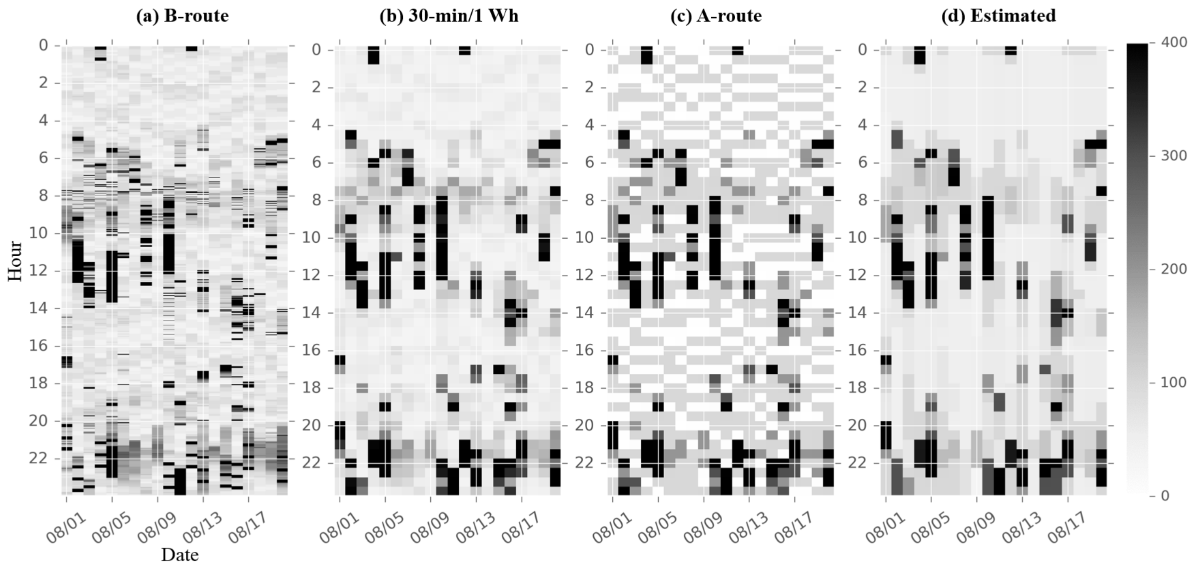

Figure 8: Visualized consumption data of the ECO data set using (a) B-route, (b) 30-min/1 Wh, (c) A-route, and (d) Estimated data.

proposed algorithm presupposes the continuation of nearly equal consumption. For that reason, the proposed algorithm tends to smooth the estimated results when the actual consumption is low or surges quickly.

4.3 Performance of Occupancy Detection

Occupancy detection was conducted based on the conditions described in Section 4.1. The variables shown in Table 2 were employed as features and the occupancy state was classified into a binary class: the household is occupied or unoccupied.

Generally, it is known that classifying a minority class is harder than the majority one when machine learning-based classification is conducted on imbalanced data sets (Japkowicz et al., 2000). In the ECO data set, classifying an unoccupied state is difficult because it accounts for only about 20% of the data. Precision and recall of the unoccupied state are therefore employed as the performance criteria in addition to accuracy, which shows the total rate of classification for occupied and unoccupied states.

Figs. 9 and 10 respectively show the performances of occupancy detection using random forests and SVM methods, which employ four types of electricity consumption data: B-route, 30-min/1 Wh, A-route, and estimated consumption. The performances of B-route are basically highest, followed by estimated consumption, 30-min/1 Wh, and A-route in descending order. The accuracy, precision, and recall of the estimated consumption data outperformed the 30-min/100 Wh and A-route data results for both ma-

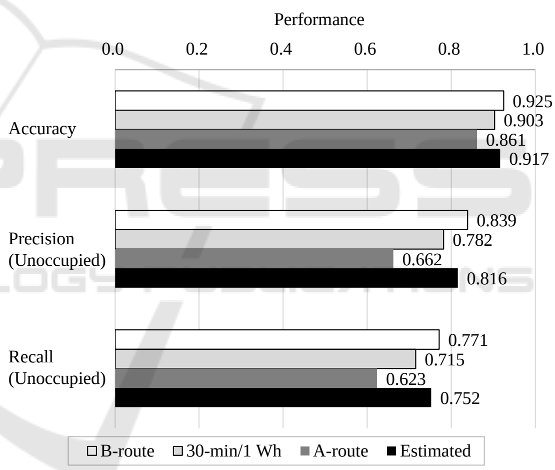

Figure 9: Performance of occupancy detection by random forests.

chine learning algorithms. Moreover, these results are comparable to the results obtained with B-route data, which corresponds to high resolution consumption data.

The contribution ratio of features by the random forests classifier is shown in Table 3. The contribution ratio is used to evaluate the efficiency of each explanatory variable. This table shows the variables that are independent of electricity consumption, such as *hour* and *temp*, have relatively high ratios in the results obtained with A-route, whereas *max* is the highest ratio in the results obtained with the others. As described in Section 2.3, A-route readout contains phantom variation that consists of alternating 0 and 100 Wh values. The discrepancy between the actual consump-

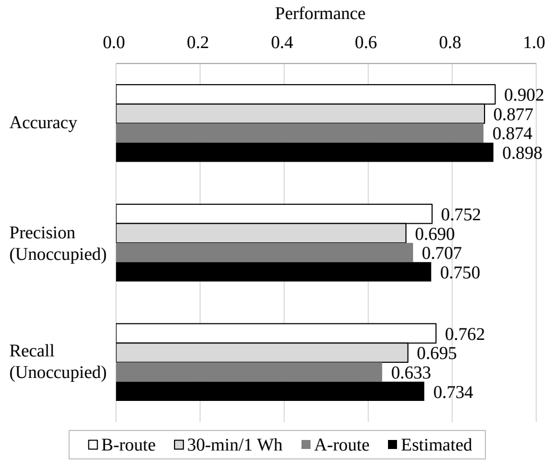

Figure 10: Performance of occupancy detection by SVM.

Table 3: Contribution Ratio of Each Explanatory Variable by Random Forests.

Variable	B-route	30-min/1 Wh	A-route	Estimated
id	0.103	0.131	0.049	0.121
mean	0.183	0.165	0.149	0.173
max	**0.211**	**0.236**	0.072	**0.208**
min	0.149	0.138	0.116	0.130
range	0.129	0.084	0.028	0.032
std	0.093	0.088	0.033	0.023
hour	0.056	0.074	0.232	0.147
temp	0.005	0.073	**0.316**	0.160
season	0.070	0.009	0.006	0.007

tion and A-route readout tends to be large when consumption is small. As a result, its contribution ratio of explanatory variables regarding electricity consumption decreased because it does not reflect the characteristics of occupied and unoccupied household well. In contrast, the estimated consumption data are considered to reflect the characteristics appropriately because its ratio of explanatory variables regarding electricity consumption is higher than that of A-route.

5 CONCLUSION

This paper proposed an actual consumption estimation algorithm that detects occupancy with high accuracy using electricity consumption data with low resolution. The proposed algorithm estimates actual consumption using the cumulative consumption of the A-route readout and the segmented line that exists within the range of true cumulative consumption. The experimental results of occupancy detection using the consumption data estimated by the proposed algorithm show an improvement in performance compared to the result obtained with raw A-route readout. The results also show that the estimated consumption data reflects the characteristics of occupied and unoccupied states appropriately.

The proposed algorithm is expected to be useful for various tasks such as profile analysis of household attributes based on A-route data. Future work also includes occupancy detection targeting Japanese domestic households and efficient feature selection for improving performance.

REFERENCES

Batra, N., Kelly, J., Parson, O., Dutta, H., Knottenbelt, W., Rogers, A., Singh, A., and Srivastava, M. (2014). Nilmtk: An open source toolkit for non-intrusive load monitoring. In *Proceedings of the 5th International Conference on Future Energy Systems*, pages 265–276. ACM.

Beckel, C., Kleiminger, W., Cicchetti, R., Staake, T., and Santini, S. (2014). The ECO data set and the performance of non-intrusive load monitoring algorithms. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, pages 80–89. ACM.

Chang, H.-H., Lin, C.-L., and Lee, J.-K. (2010). Load identification in nonintrusive load monitoring using steady-state and turn-on transient energy algorithms. In *Computer Supported Cooperative Work in Design (CSCWD), 2010 14th International Conference on*, pages 27–32. IEEE.

Chen, D., Barker, S., Subbaswamy, A., Irwin, D., and Shenoy, P. (2013). Non-intrusive occupancy monitoring using smart meters. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, pages 1–8. ACM.

Froehlich, J., Larson, E., Gupta, S., Cohn, G., Reynolds, M., and Patel, S. (2011). Disaggregated end-use energy sensing for the smart grid. *IEEE Pervasive Computing*, 10(1):28–39.

Hart, G. W. (1992). Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891.

Japkowicz, N. et al. (2000). Learning from imbalanced data sets: A comparison of various strategies. In *AAAI Workshop on Learning from Imbalanced Data Sets*, volume 68, pages 10–15.

Kim, H., Marwah, M., Arlitt, M. F., Lyon, G., and Han, J. (2011). Unsupervised disaggregation of low frequency power measurements. In *SDM*, volume 11, pages 747–758. SIAM.

Kleiminger, W., Beckel, C., and Santini, S. (2015). Household occupancy monitoring using electricity meters. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 975–986. ACM.

Komatsu, H. and Nishio, K. (2015). How can smart data analytics improve the way of information provision? –the trend of methods and an exploratory analysis–. In *Proceedings of 8th International Conference on*

Energy Efficiency in Domestic Appliances and Lighting.

Molina-Markham, A., Shenoy, P., Fu, K., Cecchet, E., and Irwin, D. (2010). Private memoirs of a smart meter. In *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-efficiency in Buildings*, pages 61–66. ACM.

Nakano, Y. and Murata, H. (2007). Non-intrusive electric appliances load monitoring system using harmonic pattern recognition-trial application to commercial building. In *International Conference on Electrical Engineering*.

Nguyen, T. A. and Aiello, M. (2013). Energy intelligent buildings based on user activity: A survey. *Energy and Buildings*, 56:244–257.

Nomura, K., Kashiwagi, T., Yamashita, T., and Kawade, T. (2014). Solution to utilization of smart meter data. *FUJITSU Sci. Tech. J*, 50(2):58–66.

Rashidi, P. and Mihailidis, A. (2013). A survey on ambient-assisted living tools for older adults. *IEEE Journal of Biomedical and Health Informatics*, 17(3):579–590.

Zoha, A., Gluhak, A., Imran, M. A., and Rajasegarar, S. (2012). Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors*, 12(12):16838–16866.