https://doi.org/10.48009/3_iis_2020_41-52

Issues in Information Systems
Volume 21, Issue 3, pp. 41-52, 2020

CAN ANYONE PROTECT THE PRIVACY IN THE IOT ERA: EVIDENCE FROM THE SMART GRID

Xingzhi Li, Chongqing University, lixingzhi@cqu.edu.cn
Zhi Xiao, Chongqing University, xiaozhicqu@163.com
Wenfeng Zhang, Chongqing University, zhangwenfeng44@hotmail.com
Du Ni, Chongqing University, mt15nn@mail.wbs.ac.uk

ABSTRACT

The emerging Internet of Things (IoT) is rapidly changing people's life and work by integrating massive data collected by various smart devices. Thus, privacy protection for heavy IoT-device users becomes one of the most concerned problems in this field. However, it is found that a simple IoT device may give away people's privacy by IoT-based data mining methods in this paper. It shows that with the help of transfer learning based on the energy disaggregation data from the IoT-based smart grid, even ordinary low-frequency grid load data could effectively reveal the resident's privacy like family income or family member's age. The results show that the pre-trained feature extractors for people's electricity consumption patterns can help infer their privacies, which in turn proves that people's private traits will affect their usage patterns of the electrical appliances. The results also warn that privacy protection should be concerned even by people who use only the simplest and the most common IoT devices (for example, a smart electric meter, which is used by more than 55% of families in the U.S.).

Keywords: Internet of Things (IoT), Privacy protection, Smart grid, Transfer learning

INTRODUCTION

Internet of Things is a collection of "things" embedded with electronics, software, sensors, actuators, and connected via the Internet to collect and exchange data with each other (Yang et al., 2017). It will influence every aspect of people's daily activities by promoting efficient resource utilization, minimizing human efforts, and making better decisions. However, the information collected by IoT devices may lead to undesirable damage to user's privacy. Numerous IoT privacy threats that even involve major Internet service providers such as Microsoft, Google, and Facebook have been reported in the past several years (Porambage et al., 2016). Many studies on IoT privacy protection have been done by designing safe IoT architectures (Hou et al., 2019), developing data encryption algorithms (Wang et al., 2019), and constructing better IoT frameworks (Ammar et al., 2018). Most of these studies focus on information security in a complex IoT system. But, in this paper, proofs from the smart grid will show that a simple and common IoT device may result in privacy risks.

The smart grid is a part of the IoT system. It is a bidirectional electric system that uses information, cyber-secure communication technologies, and computational intelligence in an integrated fashion across electricity generation, transmission, substations, distribution, and consumption (Gharavi and Ghafurian, 2011). It benefits both energy providers and users: for energy providers, it lowers the operational cost and facilitates the grid fault detection; for users, it provides more accurate electric bills and increases the energy reliability. However, there are also privacy risks in smart grids like most other IoT systems. Non-intrusive Load Monitoring (NILM) is an underlying risk to privacy in the smart gird (Mustafa et al., 2019). NILM is a technique that uses different kinds of energy disaggregation tools (Mocanu et al., 2016) to monitor the load use of electric appliances by the real-time main electricity load of a family. It is beneficial for academic purposes, like studying residents' energy usage behaviors (Li et al., 2020) and electricity load prediction (Welikala et al., 2017).

However, data owners can infer the user's consumption patterns by NILM algorithms with high-frequency load data sampled in seconds (D'Incecco et al., 2020) or milliseconds (Liu et al., 2019). Some people may think that the electricity consumption patterns are trivial, or their electric load data are not recorded in such a high frequency. This study will show that, with the transfer learning methods based on NILM, the household income level and the age

structure can be revealed based on low-frequency load data (the total value of the family use sampled in hours), which are collected by the widely deployed smart electric meter (e.g., more than 76 million smart electric meters have been installed with more than 55% coverage for residents in the U.S. by 2018 (U.S. Energy Infomation Administration, 2019) and with more than 30% coverage in Great Britain by 2019 (Kerai, 2020)). Due to different bargaining powers and price sensitivities of different groups of people, energy providers may use these kinds of private information in electricity pricing, which leads to price discrimination.

Table 1. Researches on IoT privacy issues and NILM

Authors	Year	Article type	Topic	Features
Y. Yang, et al.	2016	Literature review	IoT security and privacy issues	Related IoT privacy issues to the security problems
P. Porambage, et al.	2016	Literature review	IoT privacy issues	Systematically reviewed the user privacy issues and data mining privacy issues
J. Hou et al.	2019	Literature review	IoT privacy issues	IoT architectures
X. Wang et al.	2019	Literature review	IoT privacy issues	Encryption algorithms
M. Ammar et al.	2018	Literature review	IoT privacy issues	IoT frameworks
D. Mocanu et al.	2016	Research article	NILM	Machine learning for NILM
Y. Liu et al.	2019	Research article	NILM	Transfer learning for NILM with high frequency data
M. D'Incecco et al.	2020	Research article	NILM	Transfer learning for NILM with minute-level data

Three main contributions are made in the study. First, a NILM for incomplete low-frequency electricity load data is proposed and verified. Second, the pre-trained NILM model's feature extractors are found to be suitable for transferring to the privacy inference model, which in turn indicates that the resident's private traits are strictly related to their electricity consumption patterns. Third, the study proves that a simple IoT device like the widely used smart electric meter can be an underlying risk to people's privacy; thus, more studies should be done for privacy protection in the IoT era.

The remainder of this paper is organized as follows. Section 2 provides detailed descriptions of data sources, data format, research framework, and methods. Section 3 provides the computational results. Section 4 concludes the study and presents some ideas for further studies.

DATA AND METHODS

Data Sources

Research data in the study are acquired from Pecan Street Inc. (Pecan Street Inc., 2019). Pecan Street Inc. runs a testbed, including 1115 volunteered homes and businesses in a smart grid since 2009. It is officially patronized by the U.S. Department of Energy, and its data have supported many academic studies. Participants in this project are mainly distributed in Pecan Street, Austin, Texas.

Both smart electric meter record data and field survey data are used in the study. Smart meter record data includes hourly records of 62 kinds of smart electric meters (recording the main load and loads for 61 different appliances) from 328 households located in Pecan Street, Austin ranging from December 1st, 2016 to February 28th, 2018. The data are formatted by Pecan Street Inc as CSV files. Most of them are incomplete. The surveys are four randomly sampled surveys among the project done by Pecan Street Inc in 2012, 2013, 2014, and 2017. Respond rate of each survey is above 90%. The surveys contain household income and household age structure, which are the focus of the paper. The hourly electricity records and the surveys use the same set of household IDs in the database so they can be matched.

In the study, the households are divided into two groups by income: households with a higher income (above or equal to the mode) and households with a lower income (below the mode). Likewise, the households can also be divided into two groups by age structure: households of older residents (residents of the household are all above 65) and households of younger residents (at least one member of the household is below 65). Then the task is to classify households based on its main load.

Methods

There are only 225 (242) labeled samples of income levels (age structure) for model training and testing. The sample size is small for sequence classification (to reveal which income group or age group the household belongs to). Hence, a transfer learning approach is proposed to fix this problem.

Transfer learning is a technique that aims to improve the model's predictive ability for a target domain based on the knowledge from a related learning domain; it would significantly improve the performance of learning by avoiding excessive data-labeling efforts if done successfully (Pan and Yang, 2009). In this case, considering the close relationship between household electricity consumption behaviors and household income and age structure (Li et al., 2020), a sequence-to-sequence (seq2seq) NILM model is built and pre-trained as a base model to extract features of household consumption behaviors. The first two layers of the seq 2seq NILM model are transferred to a classification model of household income or age structure.

A 5-fold cross-validation process is adopted for model evaluation. In the first place, the labeled households are randomly shuffled. Then in each loop of the cross-validation process, the hourly electricity usage records are divided into three parts: the unlabeled records, the labeled training records of 80% households, and the labeled test records of the remaining 20% households. The base NILM model is trained with the unlabeled records and labeled training records, and the privacy classification model is trained with labeled training records. The performance of the whole model is evaluated on the test set. The framework of the proposed methods is illustrated in Fig. 1.

Details of the base seq2seq NILM model and the privacy classification model are presented in the following part.

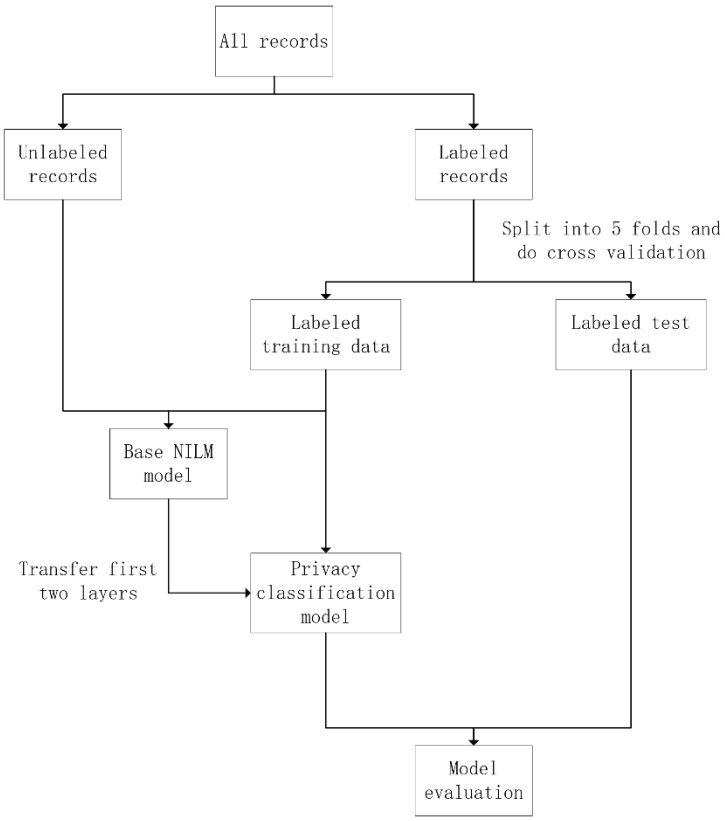

Figure 1. The framework of the proposed methods in this paper..

Seq2seq NILM Model

A seq2seq model is a machine model that converts sequences from one domain to sequences in another domain. It is widely used in sequence processing fields like natural language translation and signal processing. Researchers prefer sequence-to-point (seq2point) in previous real-time NILM studies (Nalmpantis and Vrakas, 2018) since they need to infer the disaggregated load use at time t based on historical main load records at time $t, t-1, t-2, \ldots$. However, the purpose of this studies is not to monitor load use in real-time. NILM here works as a feature extractor for household's electricity consumption patterns. Then the task is to convert historical sequences of the main load to sequences of the historical disaggregated load. Thus, a seq2seq model is valid here for its higher efficiency in model training.
The original load records are very long sequences and vary in length. Thus, they are truncated by week number (from Monday to Sunday) into the same length of 168 (24×7). The seq2seq deep learning model structure is presented in Fig 2, a.

The first three layers extract data features of the electricity load sequence (see Fig 2 b., c., d.). The following two convolutional layers increase the model's generalization ability. Owning to the fact that few participants in the project installed all the smart meters monitoring 61 kinds of appliances' usage, the records are incomplete in most cases. Let the missing data be -1 in the data preprocessing stage (since the real electricity use cannot be negative, -1 can be used as a unique mark for missing data). Therefore, the loss function is defined as $\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\frac{(sign(y_{nt})+1)}{2}(y_{nt}-\hat{y}_{nt})^2$, where N is 62, denoting the 62 channels of the output of the model (including 61 disaggregated energy use and the aggregated energy use); T is 168, denoting the length of the truncated sequences, y_{nt} denotes the true value of the n-th energy use record at time t; $\hat{y}_{nt}$ denotes the corresponding predicted value, and $\frac{(sign(y_{nt})+1)}{2}$ takes 0 when $y_{nt} < 0$ (indicating that the true value is missing) and takes 1 when $y_{nt} \geq 0$.

It should be noted that the purpose of the paper is not to find the best NILM model but to discover the hidden risk against people's privacy. Thus, fine-tuning is beyond the study's scope, and the model's parameters might be further optimized in the future. However, in the result part, it will be proved that even this basic model performs well in disaggregating incomplete low-frequency data.

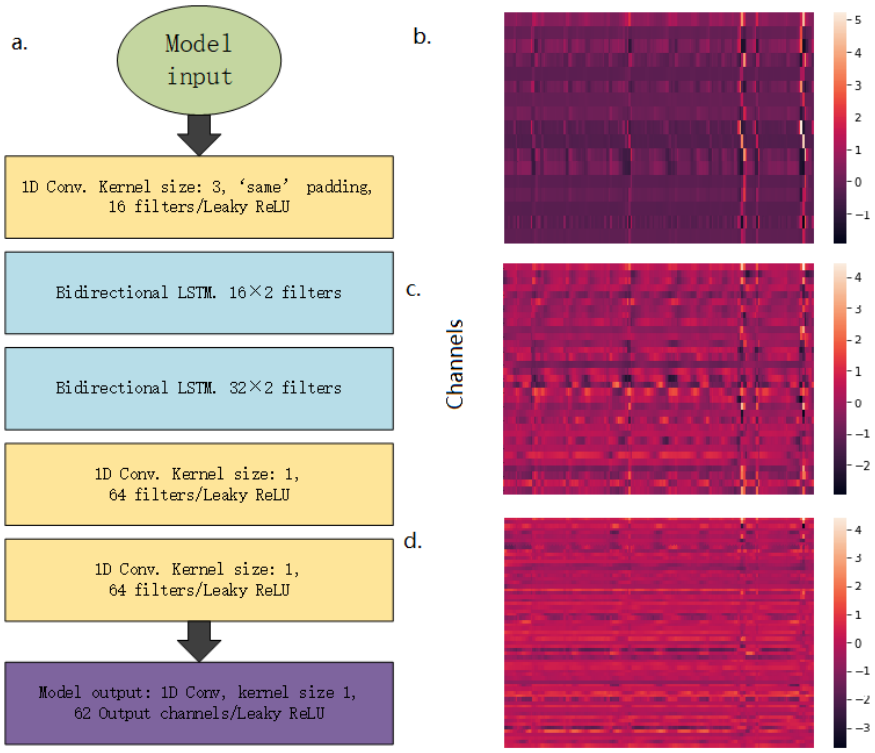

Figure 2. Seq2seq NILM model structure and visualized example of its extracted features.

Transfer Learning for Privacy Classification

Once the base model is well pre-trained, the deep learning model for the essential problem in the study can be carried out. In the first place, a classification model needs to be proposed. To utilize the pre-trained base model, the new proposed model shall have the same structure as the base model does in the feature extraction layers. Meanwhile, the model needs to balance the generalization ability (to prevent underfitting) and the structure risk (to prevent overfitting). Based on these rules, the structure of the target model is designed as model II in Fig. 3.

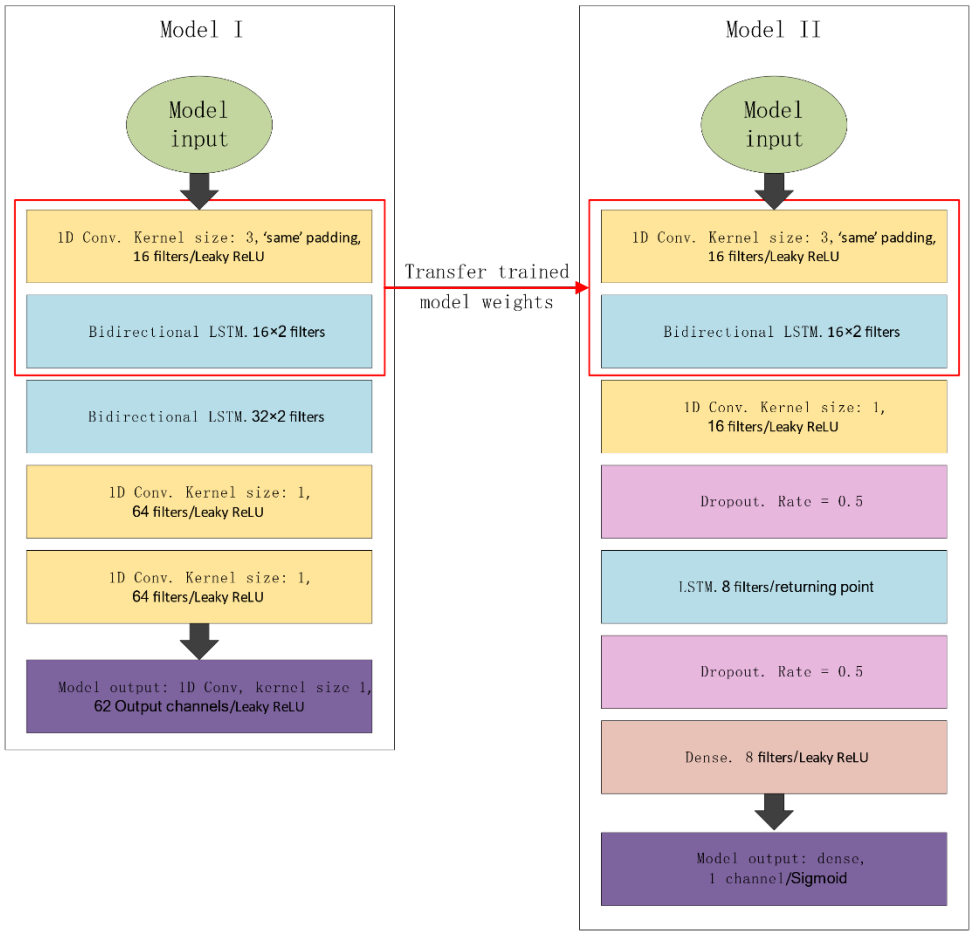

Figure 3. The illustration of transfer learning.

The weights of the proposed privacy classification model's first two layers are transferred from the NILM model and are constant (they will not change during the fitting process), thus features of the user's electricity consumption patterns will be extracted according to the pre-trained model. The second convolutional layer in the model serves for a parameter reduction purpose (its input has 32 channels while output has only 16 channels). The second LSTM layer mines the information from the extracted features. The dense layer increases the generalization ability of the model. Meanwhile, the dropout layers serve to prevent overfitting.

RESULTS

The seq2seq NILM model is the base model in the study. Its effectiveness in feature extraction is the key to the transfer learning process. Hence, the seq2seq NILM model is evaluated in the following part 3.1. The results of the privacy classification model are summarized in 3.2.

Results of The Seq2seq NILM Model

Due to the incompleteness and the low frequency of the input data, the performance of the model is limited in disaggregating the electricity use of low-power appliances (lights, Wi-Fi routers, security monitors, etc.). However, the model could achieve good accuracy in disaggregating the electricity use of high-power appliances (air conditioners, electric charging devices for vehicles, furnaces, refrigerators, etc.). Metrics for some of these disaggregation results based on the cross-validation are shown in table 1.

Table 2. Metrics of the seq2seq NILM model's performances.

Appliances	MSE	MAE	EVS
Air conditioner	0.058	0.122	84.5%
Car charging device	0.107	0.090	87.8%
Furnace	0.017	0.073	83.9%
Swimming pool pump	0.105	0.171	84.8%
Refrigerator	0.002	0.034	77.6%

The mean square error (MSE) and mean absolute error (MAE) are calculated based on the hourly usage in kilowatt (kW). The explained variance score (EVS) equals to $\left(1-\frac{Var(y-\hat{y})}{Var\ y}\right)\times 100\%$. It measures the discrepancy between the model's prediction and actual data.

Table 2 shows that the features of the household's high-power appliances usage patterns can be captured even from the incomplete low-frequency records of the main grid load. Fig. 4 is a randomly selected example of the air conditioner's disaggregated value and the actual load value during a week. The similarities between the two lines in Fig. 4 indicate the effectiveness of the proposed NILM model.

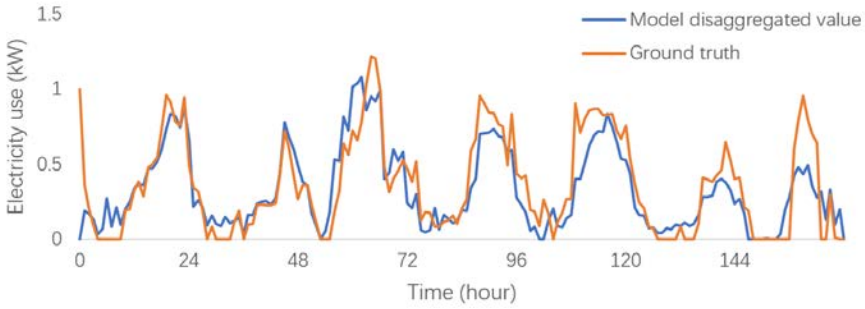

Figure 4. Disaggregated results vs. true values of the air conditioner's electricity usage.

Results of The Privacy Classification Model

Two kinds of privacies are considered in this paper: household income and household age structure. As aforementioned in 2.1, the households are divided into different groups based on their private traits (the income level and the age structure). Two computation experiments are carried out for these two kinds of private traits. By denoting the households with higher income as the positive events and households with lower income as negative events, the confusion matrix for the classification task of household income level can be carried out. Likewise, the confusion matrix for the classification task of resident's age can be calculated by denoting households with only older people (older than 65) as positive events and households with relatively younger people (include at least one family member who is younger than 65) as negative events. The receiver operating characteristic curves (ROC), lift charts, and gain charts can also be plotted base on it.

Since the dataset is imbalanced, receiver operating characteristic curves (ROC) is a good measure for model evaluation. The proposed method is compared with three other classification methods: LSTM classifier (with two LSTM layers and a dense layer), ANN (with one hidden layer), and logistic regression. Note that ANN and logistic regression do not receive sequential inputs. They take the mathematical mean of the sequence as input instead. It can be observed from Fig. 5 that mining the resident's age is more difficult than mining their income. Meanwhile, the proposed method outperforms other methods on both tasks.

The lift and gain charts are also popular visualized metrics for classification models. The lift charts and gain charts of the proposed method and the comparison methods are plotted separately in Fig. 6 and Fig.7. Both the lift charts and

gain charts show that: the proposed method is slightly better than other methods in classifying the income level, but it outperforms other methods significantly in classifying the resident's age.

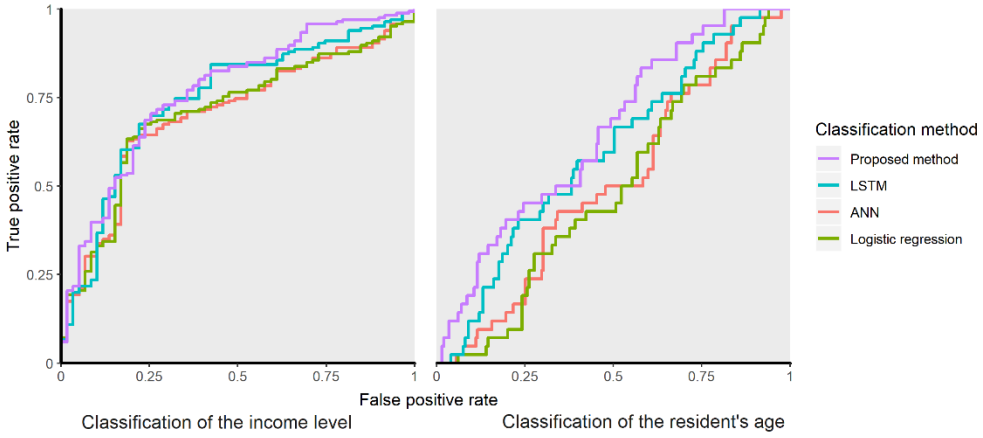

Figure 5. ROCs of the proposed method and comparison methods on the classification task of household income level and household age structure.

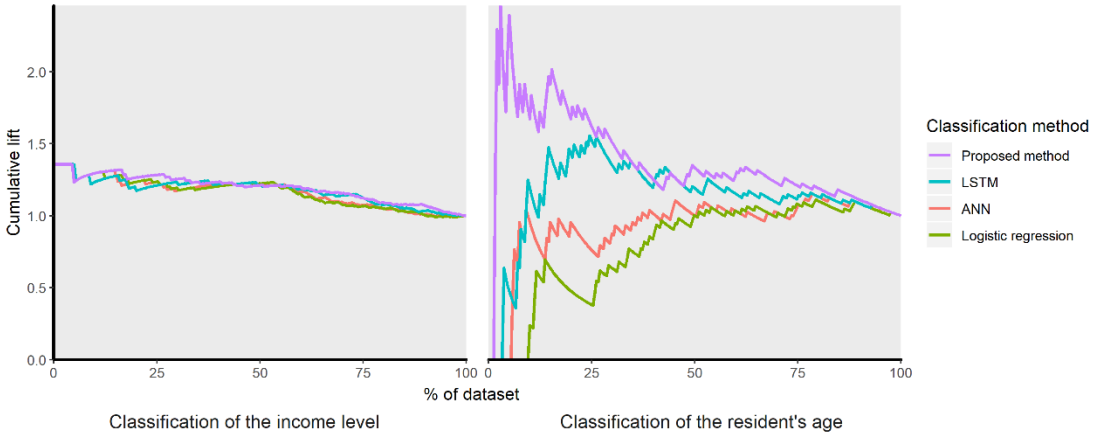

Figure 6. Lift charts of the proposed method and comparison methods on the classification task of household income level and household age structure.

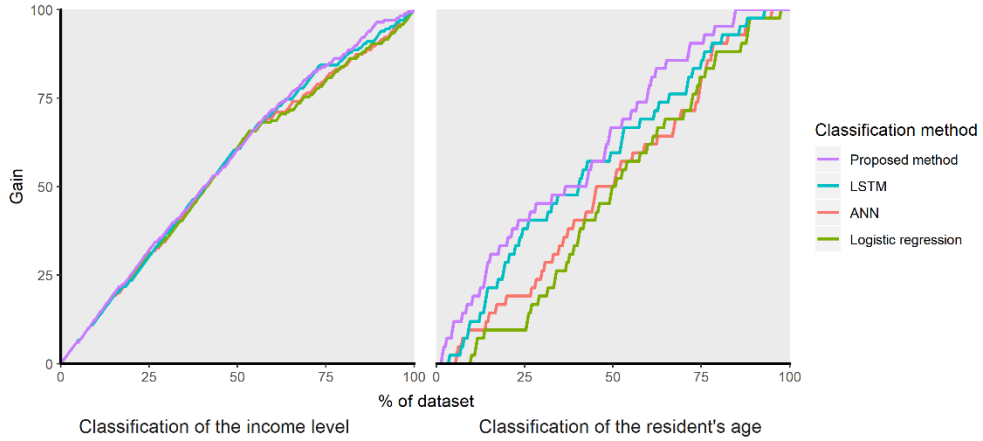

Figure 7. Gain charts of the proposed method and comparison methods on the classification task of household income level and household age structure.

Similar results are also given in table 3 and table 4. In these tables, sensitivity, specificity, positive predictive value, negative predictive value, and accuracy are calculated based on the confusion matrix as basic metrics. However, due to the imbalance of the classification task, these metrics become invalid in classifying the household age structure. Hence, the KS values and AUCs are also given as metrics to fix the problem.

It can be observed that the proposed model is the best in specificity, positive predictive value, negative predicted value, accuracy, and AUC in classifying income level (sensitivity and KS are very close among all four models). KS and AUC also indicate that the proposed model outperforms other models significantly in classifying household age structures.

It should be noticed that ANN and logistic regression can hardly be deemed as effective in classifying the household age structure (the AUCs of ANN and logistic regression are 0.51 and 0.48, while the KS values of ANN and logistic regression are 14.60 and 0.11, which are not differentiated from random guess significantly), while the proposed method performs much better.

Table 3. The metrics for the proposed method and comparison methods in classifying income level.

	Proposed model	**LSTM**	**ANN**	**Logistic regression**
Sensitivity	0.97	0.99	1	1
Specificity	0.19	0.02	0	0
Positive predictive value	0.77	0.74	0.74	0.74
Negative predictive value	0.69	0.50	-	-
Accuracy	76.44%	73.78%	73.78%	73.78%
KS	0.46	0.46	0.46	0.47
AUC	0.77	0.75	0.71	0.70

Sensitivity, specificity, positive predictive value, negative predictive value, and accuracy are calculated based on the confusion matrix. The negative predictive values for ANN and logistic regression do not exist because all predicted events of these two methods are positive events based on the imbalanced training set.

Table 4. The metrics for the proposed method and comparison methods in classifying household age structures.

	Proposed model	**LSTM**	**ANN**	**Logistic regression**
Sensitivity	0	0	0	0
Specificity	1	1	1	1
Positive predictive value	-	-	-	-
Negative predictive value	0.83	0.83	0.83	0.83
Accuracy	82.57%	82.57%	82.57%	82.57%
KS	0.26	0.17	0.15	0.11
AUC	0.65	0.60	0.51	0.48

Because of the highly imbalanced training set, all models' predictions are always positive events. Thus, sensitivity, specificity, positive predictive value, negative predictive value, and accuracy of the four models are indifferent. In this case, the KS (Kolmogorov-Smirnov) values and AUCs are better metrics for model evaluation.

CONCLUSION

According to the above results, the conclusion can be drawn as follows. First, the proposed transfer base model, the seq2seq NILM model, can capture the electricity load patterns of high-power appliances with incomplete low-frequency load data, while the load data are widely collected by a simple smart electric meter around the world. Second, all four models are valid in classifying the household income level (the proposed transfer learning model and the LSTM are based on the time series low-frequency load data, while the ANN and logistic regression are based on the averaged load data). The proposed method outperforms the comparison methods according to most of the metrics

(including the accuracy, the AUC, the lift chart, gain chart, and the ROC). Third, only the proposed method and LSTM are effective in classifying the household age structure, while the averaged-load-data-based ANN and logistic regression are not significantly better than random guess (the AUC of these two methods here are only 0.51 and 0.48, and the KS is only 0.15 and 0.11). In this case, the proposed method performs better than the comparisons substantially.

Some limitations of the study are listed here. First, although electricity consumption patterns of high-power appliances can be captured by the seq2seq NILM model based on low-frequency data, the use of low-power appliances, which may also reflect people's private traits, are not collectible. Second, it is hard to explain how the features of electricity consumption patterns help in inferring the privacies in detail according to the deep learning model. Third, the transfer learning-based proposed model outperforms other models significantly in probability predicting (revealed by ROC chart, lift chart, and gain chart), but it still needs to be fine-tuned to improve the classification accuracy, especially for the imbalanced dataset.

DISCUSSION

According to the study, ANN and logistic regression, which takes the averaged electricity load data as inputs, are valid in classifying the household income level but fail in classifying household age structure. It implies that the average electricity use is related to household income, while it is not significantly related to the resident's age. However, the proposed method is proved to be effective in both tasks. It shows that the time series of electricity load records are more informative than the averaged electricity load, and the proposed method is capable of validating the information. Moreover, the proposed model uses the pre-trained CNN and LSTM layer transferred from a seq2seq NILM model as a feature extractor; and, it outperforms the comparison LSTM model. Thus, it can be inferred that resident's electricity consumption patterns are closely related to their private traits since the feature extractors pre-trained with the NILM model (which reveals resident's electricity consumption patterns based on their main electricity load) can boost the model performance in classifying household income level and household age structure.

The results also show that resident's electricity consumption patterns can be captured by the proposed seq2seq NILM method based on the widely collected low-frequency electricity load records. Then the consumption patterns can be used in inferring people's privacies effectively with the proposed transfer learning method. It means that even if the data was well secured and was free from illegal attacks, the energy providers could still use the information to make contracts that are harmful to specific groups of energy consumers. For example, people of different ages and with different incomes vary in energy price sensitivity (Nesbakken, 1999) and bargaining power (Calvi, 2020). Older people are more price-sensitive but have weaker bargaining power; people with higher income are more price-sensitive to energy prices than those with lower income. Hence, energy providers can use personalized pricing strategy to maximize their profits when they infer the related privacies. Meanwhile, some consumers (especially the socially vulnerable groups like older people and impoverished people) will receive a higher energy price accordingly, which leads to the inequalities in society.

The aim of this study is not to oppose the use of the IoT-based smart grid. It should be acknowledged that the IoT-based smart grid has many advantages for both energy users and energy providers, and for academic research and environmental protection. Some examples of these advantages have been raised in the introduction part. The study is demonstrating that data mining, as an essential scheme for prediction and decision-making, can lead to privacy risks based on the data collected by even some of the most common IoT-devices. Therefore, to make full use of IoT data without privacy violation, more studies should be done on IoT-data management, data process regulation, and energy pricing regulation.

In future studies, it would be meaningful to use electricity load records of higher quality to capture the features of the low-power appliances' usage pattern; and make better inferences about people's different types of privacies to prove the importance of privacy protection in IoT. More importantly, the relationship between people's traits (education level, attitude to environment protection, income, and age) and their electricity consumption patterns of different appliances can be studied using the techniques proposed in this paper. It would be essential to theories of energy consumption behaviors.

REFERENCES

Ammar, M., Russello, G., Crispo, B., 2018. Internet of Things: A survey on the security of IoT frameworks. Journal of Information Security and Applications 38, 8-27.

Calvi, R., 2020. "Why are older women missing in India? The age profile of bargaining power and poverty," Journal of Political Economy 128, 2453-2501.

D'Incecco, M., Squartini, S., Zhong, M., 2020. Transfer Learning for Non-Intrusive Load Monitoring. IEEE Transactions on Smart Grid 11, 1419-1429.

Gharavi, H., Ghafurian, R., 2011. Smart grid: The electric energy system of the future. IEEE.

Hou, J., Qu, L., Shi, W., 2019. A survey on internet of things security from data perspectives. Computer Networks 148, 295-306.

Kerai, M., 2020. Smart Meter Statistics in Great Britain: Quarterly Report to end December 2019 Department for Business, Energy and Industrial Strategy (U.K.), p. 5.

Li, X., Lim, M.K., Ni, D., Zhong, B., Xiao, Z., Hao, H., 2020. Sustainability or continuous damage: A behavior study of prosumers' electricity consumption after installing household distributed energy resources. Journal of Cleaner Production 264.

Liu, Y., Wang, X., You, W., 2019. Non-Intrusive Load Monitoring by Voltage–Current Trajectory Enabled Transfer Learning. IEEE Transactions on Smart Grid 10, 5609-5619.

Mocanu, D.C., Mocanu, E., Nguyen, P.H., Gibescu, M., Liotta, A., 2016. Big IoT data mining for real-time energy disaggregation in buildings, 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 003765-003769.

Mustafa, M.A., Cleemput, S., Aly, A., Abidin, A.J., 2019. A secure and privacy-preserving protocol for smart metering operational data collection. IEEE Transactions on Smart Grid 10, 6481-6490.

Nalmpantis, C., Vrakas, D., 2018. Machine learning approaches for non-intrusive load monitoring: from qualitative to quantitative comparation. Artificial Intelligence Review 52, 217-243.

Nesbakken, R., 1999. Price sensitivity of residential energy consumption in Norway. Energy economics 21, 493-515.

Pan, S.J., Yang, Q., 2009. A survey on transfer learning. IEEE Transactions on knowledge and data engineering 22, 1345-1359.

Pecan Street Inc., 2019. Dataport.

Porambage, P., Ylianttila, M., Schmitt, C., Kumar, P., Gurtov, A., Vasilakos, A.V., 2016. The Quest for Privacy in the Internet of Things. IEEE Cloud Computing 3, 36-45.

U.S. Energy Infomation Administration, 2019. Annual Electric Power Industry Report, Form EIA-861 detailed data files.

Wang, X., Zha, X., Ni, W., Liu, R.P., Guo, Y.J., Niu, X., Zheng, K., 2019. Survey on blockchain for Internet of Things. Computer Communications 136, 10-29.

Welikala, S., Dinesh, C., Ekanayake, M.P.B., Godaliyadda, R.I., Ekanayake, J.J.I.T.o.S.G., 2017. Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting. 10, 448-461.

Yang, Y., Wu, L., Yin, G., Li, L., Zhao, H., 2017. A Survey on Security and Privacy Issues in Internet-of-Things. IEEE Internet of Things Journal 4, 1250-1258.