Heterogeneous transfer learning for thermal comfort modeling

Hu, Weizheng; Luo, Yong; Lu, Zongqing; Wen, Yonggang

2019

https://hdl.handle.net/10356/150158

https://doi.org/10.1145/3360322.3360843

Heterogeneous Transfer Learning for Thermal Comfort Modeling

Weizheng Hu
School of Computer Science and Engineering
Nanyang Technological University, Singapore
huwe0012@ntu.edu.sg

Yong Luo
School of Computer Science and Engineering
Nanyang Technological University, Singapore
yluo@ntu.edu.sg

Zongqing Lu
Peking University, China
Peng Cheng Lab, China
zongqing.lu@pku.edu.cn

Yonggang Wen
School of Computer Science and Engineering
Nanyang Technological University, Singapore
ygwen@ntu.edu.sg

ABSTRACT

For decades, the Predicted Mean Vote (PMV) model has been adopted to evaluate building occupants' thermal comfort. However, recent studies argue that the PMV model is inaccurate and suffers from two major issues: thermal comfort parameter inadequacy and modeling data inadequacy. To overcome these issues, in this paper, we propose a learning-based approach for thermal comfort modeling, named as Heterogeneous Transfer Learning (HTL) based Intelligent Thermal Comfort Neural Network (HTL-ITCNN). First, to address the parameter inadequacy issue, we add more relevant factors as the modeling features except for the six PMV parameters. Due to the flexibility of learning-based approaches, newly found thermal comfort parameters can be appended to extend the number of modeling features. Second, to mitigate the impact of the data inadequacy issue, we adopt the deep transfer learning techniques to train the thermal comfort model, where the model training would benefit from the transferred knowledge from the existing datasets. Due to the heterogeneity of the features among different datasets, we follow the HTL concept to conducting effective knowledge transfer among heterogeneous domains, which are the different but related datasets with varied features. To validate our solution, we conduct five-month data collection experiments and build our datasets. With the HTL-based two-stage learning paradigm, the experimental results show that the accuracy of HTL-ITCNN outperforms the PMV model by on average 73.9%. Besides, we verify the impacts of newly added features and knowledge transfer on model performance. Moreover, we demonstrate the enormous potential of personal thermal comfort modeling research.

CCS CONCEPTS

• **Computing methodologies** → **Neural networks**; **Modeling methodologies**; *Transfer learning*.

KEYWORDS

Thermal comfort modeling, pervasive sensing, wearables, deep neural network, heterogeneous transfer learning

1 INTRODUCTION

Thermal comfort is used to express the occupants' satisfaction with their surrounding thermal environment. Existing studies [2, 28] show that thermal discomfort not only affects occupants' productivity but also influences their long-term health. The current building comfort management strategy of the heating, ventilation, and air conditioning (HVAC) systems [27] are to maintain a "comfort" set-point temperature over the whole day. This fixed value generally obtains from the empirical results of canonical predicted mean vote (PMV) model [7], developed by *Fanger et al.* in 1970, which has been adopted by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) standard 55 [3].

Following Fanger's observations, some researches [11, 23] propose to build the IoT-based systems for automatic thermal comfort management by using various pervasive devices. For example, in a typical office environment, such an IoT-based system keeps monitoring the four environmental parameters (air temperature, relative humidity, air velocity and mean radiant temperature) of the PMV model via ubiquitous environmental sensors, measuring the occupants' clothing insulation via the infrared sensor, and tracking occupants' metabolic rate via wristbands. With these sensing data, the system can track occupants' thermal comfort in real-time and adjust the settings of the HVAC systems dynamically. If occupants are not satisfied with the current thermal environment, they can send their feedback to the system via their smartphones. Through the interaction with occupants, the IoT-based system can tune the control policy of the HVAC systems to achieve the optimal indoor thermal comfort management.

However, recent researches [20, 26] indicate that the PMV model, the core of those IoT-based systems, is inaccurate to evaluate human thermal comfort. With our literature survey, we summarize two major drawbacks of the PMV model: 1) thermal comfort parameter

inadequacy; and 2) modeling data inadequacy. For the former, other than the six PMV factors, there are a variety of parameters linking to the human thermal comfort, such as environmental factors: acoustic (e.g., noise [10]) and CO2 level [29]; personal information: age [16], gender [21], heart rate [24], skin temperature [13]. For the latter, compared to other data-driven research fields (e.g., Computer Vision, Nature Language Processing), collecting data from human subjects is a quite challenging and time-consuming task. The largest available open-source dataset of the community is the ASHRAE Global Thermal Comfort Database II [8], which only contains 81,846 samples collected from 49 field studies conducted between 1995 and 2016 from around the world. Even for the canonical Fanger's PMV model [7], it is developed based on a relatively small dataset which only collects the data from a thousand American and European subjects. Therefore, it is necessary to innovate the modeling methodology to develop next-generation thermal comfort models.

In this paper, we propose a learning-based approach for thermal comfort modeling. First, we take more thermal comfort parameters into consideration, such as time (i.e., the hour of a day), personal information (i.e., age, gender, weight, height, and clothing insulation) and vital signs (i.e., metabolic rate, heart rate and skin temperature). Moreover, for learning-based approaches, it is flexible to append new modeling parameters in the future. In this way, we can overcome the thermal comfort parameter inadequacy issue. Second, we adopt the transfer learning techniques to transfer the knowledge learned from multiple source domains (i.e., different but related datasets) to the target domain (i.e., our self-collected dataset) for modeling. However, after investigating a lot of available datasets from the community, we find that their feature spaces are different from ours, resulting in the difficulty of transferring the learned knowledge. To overcome this issue, we propose a Heterogeneous Transfer Learning (HTL) approach to conduct effective knowledge transfer based on deep neural network, by noticing that most of the existing datasets share several overlapped features (e.g., air temperature, relative humidity) with our datasets. In this way, although we could not solve the modeling data inadequacy issue entirely, we can mitigate its impacts at a certain level by utilizing the transferred knowledge obtained from the existing and related datasets.

To verify the effectiveness of our proposed approach, we use our developed Intelligent Thermal Comfort Management (iTCM) System to conduct the data collection experiments on our university campus for five months, and then analyze the obtained dataset for HTL-based thermal comfort modeling. Our contribution can be broken down into the following three aspects:

- We build the iTCM datasets for thermal comfort analytics. The iTCM datasets consist of one generic dataset and three personal datasets. Moreover, we will contribute to the community by releasing our datasets.
- We propose a transfer learning based approach for the thermal comfort modeling, named as HTL-based Intelligent Thermal Comfort Neural Network (HTL-ITCNN), which adopts a two-stage learning paradigm.
- We empirically show that HTL-ITCNN outperforms most of the existing methods. Also, we analyze the factors that produce performance improvement. Moreover, we present the enormous potential of personal thermal comfort models.

This paper provides fundamental insights for applying the transfer learning concept to develop thermal comfort models. We combine the techniques of deep neural network and heterogeneous transfer learning for thermal comfort modeling and obtain the significant improvement of the prediction accuracy. Our findings will make contributions to the future researches in the human thermal comfort research field.

The rest of this article is organized as follows. Section 2 reviews the related work. Section 3 introduces training datasets and controlled experiments. Section 4 presents the adopted modeling features and describe the learning paradigm of HTL-based thermal comfort modeling. Section 5 evaluates the performance of our approach. The last section concludes this article.

2 RELATED WORK

Various efforts have been made to apply the machine learning techniques to the human thermal comfort modeling. Many of them focus on using traditional machine learning algorithms. *Barrios et al.* [4] proposed to use the off-the-shelf sensors to track occupants' heart rate as a proxy of the metabolic rate and adopt the linear and logistic regression approaches to develop the thermal comfort models. *Ghahramani et al.* [12] designed a dynamic Bayesian network for personal thermal comfort modeling. *Laftchiev et al.* [22] built an IoT platform for data sensing and compared the modeling performance among various classic machine learning algorithms (e.g., SVM, Kernel Ridge Regression, k-NN, Gaussian Process Regression, Quadratic Discriminant Analysis). Other studies take the Neural Network (NN) into consideration. *Weiwei et al.* [25] presented an evaluation model for personal thermal comfort based on Back Propagation (BP) NN. *Wei et al.* [32] evaluated a Deep Neural Network Learning approach to model thermal comfort.

Although many researchers have made significant contributions to improving the prediction accuracy of learning-based thermal comfort models, little effort has been paid to apply the transfer learning to this research field. The only partially related work is *Ming et al.* [19]. Authors considered that occupancy detection for buildings is crucial to improve the building energy efficiency and occupants' comfort. They proposed three frameworks for virtual occupancy sensing, and one of them is the transfer learning based method. However, different from their goal of improving the recognition accuracy of occupancy detection, our proposed HTL-ITCNN is to enhance the prediction accuracy of occupants' thermal comfort. To the best of our knowledge, none of the existing works propose to use the transfer learning techniques for thermal comfort modeling.

3 TRAINING DATASETS

In this section, we introduce the adopted datasets: ASHRAE RP-884 Dataset [6] and our collected iTCM Datasets.

3.1 ASHRAE RP-884 Dataset

The ASHRAE RP-884 dataset is one of the most widely used public datasets in the human thermal comfort research field. The objective of building this dataset is to develop the adaptive model. It comprises more than 20,000 user comfort votes collected from 52 studies across ten different climate zones around the world.

For this dataset, we extract the data based on three basic properties: 1) Selecting the votes collected from the HVAC environment; 2) Selecting the votes that contain all six PMV parameters; 3) Selecting the votes that contain occupants' actual thermal comfort feedbacks. In total, we selected 11,164 votes from the original database.

3.2 iTCM Datasets

To develop HTL-based thermal comfort models, we conduct data collection experiments to build our datasets.

3.2.1 IoT-based Thermal Comfort Management System. We adopt our previously developed intelligent Thermal Comfort Management (iTCM) system [15] for data collection. It consists of three major modules: Wireless Sensor Network (WSN), Mobile App, and Back-end Server. For the WSN, we adopt the Building-in-Briefcase (BiB) device [30] developed by engineers from the University of California, Berkeley. For the mobile App, we develop both iOS and Android versions to pair with the off-the-shelf Microsoft Band 2 [31] to track occupants' vital signs in real time. For the back-end server, we adopt an enterprise-level framework to build a reliable back-end system for data processing and storage.

3.2.2 Controlled Experiments. We conducted the experiments in a controllable chamber equipped with independent HVAC systems.

Before the start of each experiment, we set up both hardware (HVAC system, BiB device, smartphones, and Microsoft Bands) and software (mobile App and back-end server). Table 1 lists the collected various variables and the corresponding sampling intervals.

Each experiment consists of a 30-minute preparation session and a 3-hour data collection session. To obtain participants' actual comfort votes from different room temperatures, we implement each experiment session under one of the following six thermal environments: 1) 18°C-20°C; 2) 20°C-22°C; 3) 22°C-24°C; 4) 24°C-26°C; 5) 26°C-28°C; 6) 28°C-30°C. Moreover, one research student is allocated to be in charge of conducting each experiment.

- During the preparation session, first, we pass the consent forms to all participants for reading and signing. Then, we briefly introduce our experiment processes and guide them to use our developed mobile App (e.g., save personal information, set clothing settings, submit thermal comfort votes).
- During the data collection session, we only allow participants to do some light-weight tasks (e.g., read books, do homework) to simulate the general indoor environment (e.g., classroom, office). Every 10 minutes, research student will remind participants to send their thermal comfort votes. To obtain accurate feedback, we require participants to submit their votes within several seconds. Since it is difficult to ensure everyone to send the vote simultaneously, the submission interval is from 5 to 25 minutes, as listed in Table 1. Meanwhile, we also record each participant's total amount of calories burned during the submission interval to calculate the personal metabolic rate.
- After the completion of each experiment, to avoid affecting the next session and ensure the accuracy of the measured data, the research student will reset all Microsoft Bands to remove previous participants' bio-information.

Table 1: List of variable sampling interval

#	Variable	Interval
1	Actual thermal comfort votes (ASHRAE 7-point scale) with timestamp (date & time)	5-25 min
2	Burned calories (kcal)	5-25 min
3	Air temperature (°C)	2 min
4	Relative humidity (%)	2 min
5	Heart rate (bpm)	1 sec
6	Skin temperature (°C)	1 sec

Table 2: Information of modeling datasets

Dataset Name	Size	*AT Range	$^\wedge$RH Range
ASHRAE dataset	11,005	15.9-31.9°C	10.0-97.8%
iTCM generic dataset	4,293	19.6-30.6°C	37.3-83.6%
iTCM personal datasets P_1, P_2, and P_3	345 380 341	19.6-29.9°C	42.4-75.5%

*AT refers to air temperature.
$^\wedge$RH refers to relative humidity.

Following the above experimental procedures, we divide our data collection experiments into the following three stages.

- *Three-Week Preliminary Experiment:* In this stage, we recruited 30 participants. Our goal is to test and debug our iTCM system, especially the mobile App. Meanwhile, we want to collect some data for preliminary analytics.
- *Three-Month General Experiment:* In this stage, we recruited 241 participants. Our objective is to collect sufficient data to support our machine learning research for the generic thermal comfort modeling.
- *One-Month Particular Experiment:* In this stage, since there are no female applicants, we only recruited three fixed male participants. We aim to collect some data for personal thermal comfort modeling.

In total, we collect 6,257 raw votes from 274 participants, of which there are 141 males and 133 females. The male-female ratio is roughly 1:1. The age group of the participants is from 17 to 32 years old. Moreover, data privacy is under the supervision and protection of the Institutional Review Board (IRB) of our university.

Before applying the collected datasets to model training, we conducted pre-processing operations. First, we find that the interval between two adjacent votes from certain participants is too short (e.g., less than 5 minutes or even less than 1 minute), which can be considered as duplicate votes. Hence, we remove these duplicate votes from the datasets. Then, we notice that the total number of +3 (hot) votes is deficient. There are only 31 "hot" votes received from 14 participants in our datasets. ASHRAE dataset also suffers from the same issue. The reason is that the upper bound temperature (i.e., 30.6°C & 29.9°C in iTCM datasets or 31.9°C in ASHRAE dataset) is still not higher enough to make participants feel hot. Thus, we consider that these "hot" votes cannot correctly reflect most of

Weizheng Hu, Yong Luo, Zongqing Lu, and Yonggang Wen

Table 3: Modeling feature list of the iTCM datasets

#	Feature	Data Source	Description
F_1	hour	Mobile App	The hour of a day (24-hour format) when participant sends the thermal comfort vote
F_2	age		Participant's age in years
F_3	gender		Participant's gender (1: male, 2: female)
F_4	weight	Participants	Participant's weight in kilograms (kg)
F_5	height		Participant's height in meters (m)
F_6	clothing insulation (CI)		Participant's clothing insulation plus the general office chair insulation
F_7	air temperature (AT)	BiB Device	Average room air temperature in degrees Celsius (°C)
F_8	relative humidity (RH)		Average room relative humidity in percentage (%)
F_9	metabolic rate (MR)		Participant's metabolic rate in metabolic equivalent (met)
F_{10}	heart rate (HR)	Microsoft Band 2	Participant's hear rate in beats per minute (bpm)
F_{11}	skin temperature (ST)		Participant's skin temperature in degrees Celsius (°C)

Table 4: Estimated clothing ensemble & insulation for male and female participants

Clothing Ensemble (Male)	Clothing Insulation	Extra Clothing	Clothing Insulation
$^\wedge$MSAS + (Short-sleeve shirt & Walking shorts) or similar	0.47		0.81
$^\wedge$MSAS + (Short-sleeve shirt & Thin Trousers) or similar	0.54	+ Long-sleeve swear shirt	0.88
$^\wedge$MSAS + (Long-sleeve shirt & Walking shorts) or similar	0.51	or similar	0.85
$^\wedge$MSAS + (Long-sleeve shirt & Thin Trousers) or similar	0.58		0.92

$^\wedge$MSAS (Men's underwear + Shoes or Sandals + Ankle socks + Standard office chair): 0.18

Clothing Ensemble (Female)	Clothing Insulation	Extra Clothing	Clothing Insulation
*WSAS + (Short-sleeve shirt & Walking shorts) or similar	0.46		0.80
*WSAS + (Short-sleeve shirt & Thin Trousers) or similar	0.53	+ Long-sleeve swear shirt	0.87
*WSAS + (Long-sleeve shirt & Walking shorts) or similar	0.50	or similar	0.84
*WSAS + (Long-sleeve shirt & Thin Trousers) or similar	0.57		0.91

*WSAS (Women's underwear + Shoes or Sandals + Ankle socks + Standard office chair): 0.17

the participants' hot sensations, and decide to remove them from both iTCM datasets and ASHRAE dataset. As a result, we will focus on the 6-point thermal comfort scale ranging from -3 to +2. Eventually, we obtain four datasets: the ASHRAE dataset, one iTCM generic dataset and three iTCM personal datasets for thermal comfort modeling, as listed in Table 2.

4 THERMAL COMFORT MODELING

In this section, we introduce the modeling features and the proposed HTL-based learning paradigm.

4.1 Feature Introduction

Table 3 lists the collected features for thermal comfort modeling. Based on the data source, we divide them into four categories.

4.1.1 Mobile App. Existing work [9] shows the linkage between time and thermal comfort. Thus, we use the mobile App to record the hour of a day when participants send thermal comfort votes.

4.1.2 Participants. We collect personal information from participants directly. Existing researches have investigated the relationship between personal information (i.e., age [16], gender [21], and weight & height [1]) and human thermal comfort. In our experiments, participants can input and save their personal data in the account profile of our mobile App. For the clothing insulation (CI), we let participants set their clothing ensemble from a list of typical dress combinations in the tropical area (e.g., Singapore) via the mobile App, as shown in Table 4. In this paper, we adopt the thermal comfort tool [14] developed by the University of California Berkeley to estimate participants' CI based on their dress combinations. Moreover, we allow participants to take on/off the extra clothing (e.g., long-sleeve swear shirt or similar), but they need to change their clothing settings subsequently.

4.1.3 BiB Devices. BiB sensors monitor the air temperature (AT) and relative humidity (RH) of the chamber, where AT and RH have been adopted by Fanger's PMV model as important environmental factors related to human thermal comfort. Since our chamber is relatively small and can only accommodate up to six participants, two BiB sensors are deployed to cover the whole chamber. We
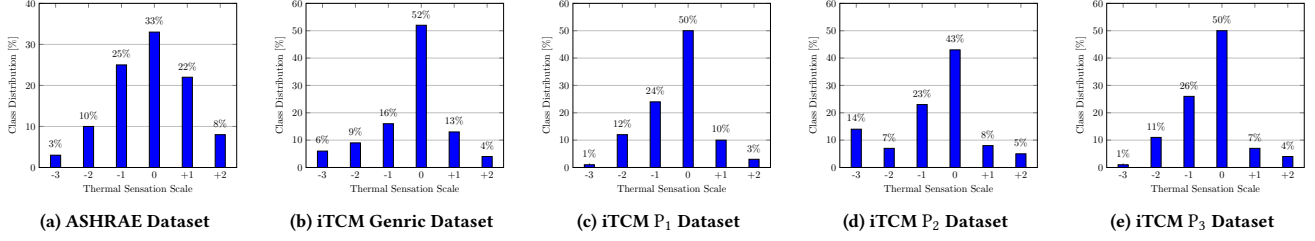
Figure 1: The class distribution of adopted datasets varies with the thermal sensation scale. We can find that all datasets have the class imbalance issue. The majority of thermal comfort votes are -1 (slightly cool), 0 (comfortable), and +1 (slightly warm).

record the average values of the AT and RH from two BiB sensors and then store them in the database.

4.1.4 Microsoft Band 2. It measures participants' vital signs, such as burned calories, heart rate (HR), and skin temperature (ST). Existing literature has shown that heart rate [24] and skin temperature [13] are relevant to human thermal comfort. In this paper, we adopt the burned calories to calculate the metabolic rate (MR), which is also one of the six PMV modeling parameters. The calculation equation [5] is presented as follows,

$$M = \frac{C * 3600}{W * T}, \qquad (1)$$

where M is the metabolic rate in met, W is the participant's weight in kg, T is the activity time interval between a participant's two adjacent votes in seconds, and C is the total number of calories burned during the duration T in kcal.

In summary, we collect the above-introduced thermal comfort features via the equipment we have during our data collection experiments. Moreover, we admit that there are another two factors (i.e., mean radiant temperature and air velocity) adopted by Fanger's PMV model for thermal comfort modeling. However, we do not have the related sensors to measure them in the current research stage. In the future, we will purchase the required sensors and take them into consideration for human thermal comfort research.

4.2 Feature Correlation Analysis

We adopt the widely used Pearson correlation statistic to measure the degree of the relationship between each feature and the actual thermal comfort index, as shown in Table 5.

In the table, we list the absolute value of the calculated correlation coefficient in descending order. Meanwhile, we also calculate the related p-value for each feature. First, we can find that ST is the only feature that has a strong association, and AT is the only feature that has a medium association. Moreover, the rest of the features have week associations. Among them, the p-values of MR, weight, and gender are larger than the general cutoff value 0.05, which indicates weak evidence against the null hypothesis. Here, we have to admit that most of the feature we collected only have week associations to the actual thermal comfort index in the Exploratory Data Analysis (EDA) process. However, we still decide to adopt all features for modeling. The main reason is that our feature space is tiny (only 11 features), these features of weak association may not be able to play important roles in model training, but

Table 5: Pearson correlation coefficient r analysis

#	Feature	★Coefficient r	p-value
F_{11}	skin temperature (ST)	0.532	
F_7	air temperature (AT)	0.485	
F_{10}	heart rate (HR)	0.248	
F_6	clothing insulation (CI)	-0.236	<0.05
F_8	relative humidity (RH)	0.113	
F_2	age	-0.059	
F_1	hour	0.058	
F_5	height	-0.055	
F_9	metabolic rate (MR)	0.017	0.271
F_4	weight	0.016	0.281
F_3	gender	0.014	0.353

★Coefficient r has the strong association if $0.5 <= |r| <= 1$, the medium association if $0.3 <= |r| < 0.5$, and the weak association if $0 <= |r| < 0.3$.

the developed models will still benefit from them. In subsection 5.2.2 of the performance evaluation, we analyze the performance improvement with the different set of features.

4.3 Heterogeneous Transfer Learning (HTL) Based Thermal Comfort Modeling

We propose to leverage the HTL-based approach to develop thermal comfort models.

4.3.1 Class Weight Mechanism. We aim to develop the learning-based models to classify participants' thermal comfort votes based on the ASHRAE 6-point thermal sensation scale (warm = +2, slightly warm = +1, comfortable = 0, slightly cool = -1, cool = -2, cold = -3). The existing classic thermal comfort models (e.g., PMV model, adaptive model) are regression models. However, the truth is that people generally cannot distinguish between two thermal comfort indexes (e.g., 0 v.s. 0.5, 1.0 v.s. 1.3). Compared to these numbers, people prefer to express their real feelings by using thermal comfort categories (e.g., comfortable, cool, warm). Therefore, in this paper, we consider the thermal comfort modeling as a classification problem, but still, our datasets and proposed modeling methodology can be used to develop regression thermal comfort models.
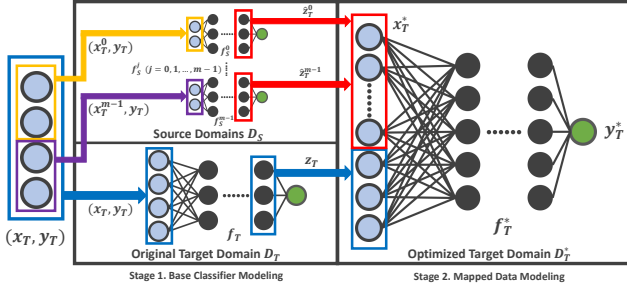
Figure 2: The schematic diagram of the two-stage HTL-based learning paradigm. In the first stage, we use each source-domain dataset to develop the related pre-trained base classifier via the Deep Neural Network (DNN) technique, respectively. In the second stage, we obtain the knowledge transferred features of target-domain dataset via the previously developed base classifiers and then input them into a new DNN for training the HTL-based classifier.

To develop a fair classifier which can effectively classify each label, we need to address the class imbalance issue, as discussed in existing researches [17, 18]. For example, in a binary classification problem, 90% of labels are 1, and only 10% of labels are 0. Without considering the class imbalance issue, the derived classifier will prefer to predict label 1 to get high prediction accuracy. However, such a classifier is not meaningful. As such, we plot the label distribution of the ASHRAE dataset and our iTCM datasets for analysis, as shown in Figure 1. We can see that the label distribution of all datasets is not balanced. More than half of the received votes are -1 (slightly cool), 0 (comfortable) or +1 (slightly warm), where the major reason is that the experimental thermal environments of all datasets are still within participants' acceptable comfort zone. Therefore, we aim to reduce the impact of class imbalance issue by assigning the weight to each label with the following equation,

$$w_i = \lambda_i \frac{p_{max}}{p_i}, \tag{2}$$

where w_i is the weight assigned to label i, p_{max} is the distribution of a label which has the maximum amount of samples, and p_i is the distribution of label i. Here, $i = 0, 1, 2, ..., n - 1$ denotes that there are n labels. λ_i is the factor used for fine-tuning the weight of each label. Using the same binary classification problem as the example, if $\lambda_0 = 0.8$ and $\lambda_1 = 1.2$, then, the weight for label 0 is $(90 \div 10) \times 0.8 = 7.2$ and the weight for label 1 is $(90 \div 90) \times 1.2 = 1.2$. In this way, we can assign the higher weights to those labels with the lower distribution. Moreover, during the model training, the cost function will consider learning more knowledge from those labels with the smaller sample size, rather than mainly focusing on the labels with the bigger sample size.

4.3.2 Learning Paradigm. In this paper, we propose the Heterogeneous Transfer Learning (HTL) based Intelligent Thermal Comfort Neural Network (HTL-ITCNN) for thermal comfort modeling via a two-stage learning paradigm, as shown in Figure 2. Our objective is to use the dataset of each source domain to develop the

related pre-trained classification model via the Deep Neural Network (DNN) technique, and then use it to map the features of the original target domain which are overlapped with that of itself to the high-dimension space. In this way, the mapped features not only hold its original information but also contain the transferred knowledge from the related source domain. To simplify the problem, we assume that the feature space of each source domain is a subset of that of the original target domain.

- *Stage 1. Base Classifier Modeling:* In this stage, we develop the pre-trained classification models for all source domains and the original target domain, as the base classifiers.

  First, as plotted in Figure 2, we use D_S to denote the multiple source domains, D_T to denote the original target domain, and D_T^* to denote the optimized target domain. Also, we suppose that there are m source domains and $D_S = (D_S^0, D_S^1, ..., D_S^{m-1})$. Next, we develop the base classifier f_S^j by using the dataset of the related source domain D_S^j, where $j = 0, 1, 2, ..., m - 1$. Meanwhile, we also use the dataset of original target domain D_T to train the related base classifier f_T. Finally, we can obtain $m + 1$ pre-trained base classifiers for feature mapping in the second stage.

- *Stage 2. Feature Mapping & HTL-Based Modeling:* In this stage, we first use the developed $m + 1$ base classifiers in the first stage to obtain the mapped features, and then feed them into a DNN to develop the HTL-Based thermal comfort model.

  First, we use (x_T, y_T) to denote the data-label pair of the original target domain D_T, where x_T denotes the data space with full features, and y_T denotes the label space. Then, based on the feature space of each source domain D_S^j, we split x_T into m subsets and $x_T = (x_T^0, x_T^1, ..., x_T^{m-1})$. Each subset can be presented as the data-label pair (x_T^j, y_T), where $j = 0, 1, 2, ..., m-1$. After that, we can obtain the mapped features $\hat{z}_T^j$ of each subset (x_T^j, y_T) with the following function,

$$\hat{z}_T^j = f_S^j(x_T^j), \quad \forall j, \tag{3}$$

where $z_T^j \in z_T$ and $\hat{z}_T = (\hat{z}_T^0, \hat{z}_T^1, ..., \hat{z}_T^{m-1})$ is the set of mapped features. Moreover, f_S^j is the base classifier of the source domain D_S^j. $D_S^j \in f_S$ and $f_S = (f_S^0, f_S^1, ..., f_S^{m-1})$ is the set of base classifiers of all source domains. Next, we obtain the mapped features z_T of set (x_T, y_T) via the base classifier f_T of the original target domain D_T as follows,

$$z_T = f_T(x_T). \tag{4}$$

Eventually, we concatenate $\hat{z}_T$ and z_T as the input x_T^* and then feed x_T^* into a new DNN to obtain the final classifier f_T^* of the optimized target domain D_T^* as follows,

$$y_T^* = f_T^*(x_T^*), \tag{5}$$

where f_T^* is the HTL-based classifier and y_T^* is the optimal prediction results.

Furthermore, to achieve the optimal multi-classification performance for all classifiers in both stages, we tune the weights $w = (w_0, w_1, ..., w_{n-1})$ for n labels by adjusting the related weight factors $\lambda = (\lambda_0, \lambda_1, ..., \lambda_{n-1})$ during the training process via Eq. (2).

5 PERFORMANCE EVALUATION

In this section, we evaluate the performance of HTL-ITCNN.

5.1 Experiment Settings

In this paper, we assume that there is only one source domain (ASHRAE dataset) which means $m = 1$ and one target domain (iTCM dataset). After analyzing the modeling features of both datasets, we find that there are four overlapped features: AT, RH, MR, and CI. In this case, HTL-ITCNN consists of three DNN-based classifiers: 1) One source-domain base classifier; 2) One original target-domain base classifier; and 3) One optimized target-domain classifier. For the detailed DNN settings of all three classifiers, we select the tanh function as the activation function and adopt the categorical cross-entropy as the cost function. In the training process, we use the batch normalization technique for normalizing the input layer as well as the hidden layer, encode the labels with one-hot encoder, and train the models via the Adam optimizer with the learning rate 0.001. Besides, to mitigate the over-fitting issue, we also adopt the L2 regularization technique and set the L2 factor as 0.0001. Moreover, we use the fixed random seed for dataset shuffling and model training. The rest of the settings are listed below,

- The source-domain base classifier f_S^0 uses the fully-connected structure with two hidden layers. Each hidden layer has 32 neurons, respectively. The batch size is 128.
- The original target-domain base classifier f_T uses the fully-connected structure with two hidden layers. Each hidden layer has 256 neurons, respectively. The batch size is 64.
- The optimized target-domain classifier f_T^* uses the fully-connected structure with two hidden layers. Each hidden layer has 512 neurons, respectively. The batch size is 64.

In addition, we initialize the weights of all neural networks by using the truncated normal distribution with zero mean and standard deviation of $\frac{1}{\sqrt{Z}}$, where Z is the number of inputs to a neuron layer, and initialize the biases of all neural networks as zero. Also, we select the accuracy, macro-F1 score, and Matthews correlation coefficient (MCC) as the evaluation metrics. Among them, macro-F1 score and MCC are sensitive to rare labels. Higher scores indicate that the classifier has better classification performance on those labels with the lower distribution.

Furthermore, as mentioned in section 3, we only consider the 6-point thermal sensation (from -3 to +2) Each instance of the collected data consists of various features and the occupant's actual vote as the ground truth. To train the HTL-based thermal comfort models, we split all datasets into training and testing sets by using the 10-fold cross-validation (CV) approach. For the ASHRAE dataset, since it is independent of our iTCM datasets, we only select the best model derived via the 10-fold CV as the source-domain base classifier f_S^0. For the iTCM dataset, we use it for training both original target-domain base classifier f_T and optimized target-domain classifier f_T^*. Different from the training process of the ASHRAE dataset, we save the ten models derived via the 10-fold CV and then use each of them to train the optimized target-domain classifier f_T^* in each fold, respectively.

For the comparison baseline, we select the predicted mean vote (PMV) model. It predicts the average comfort vote from a group of occupants based on four environmental factors (air temperature,

Table 6: Multi-classification performance analysis

Source domain			
Algorithm	Accuracy (%)	Macro-F1 (%)	MCC (%)
PMV	**32.68**	17.08	6.07
Base Classifier f_S^0	32.36	**28.63**	**15.02**

Target domain			
Algorithm	Accuracy (%)	Macro-F1 (%)	MCC (%)
PMV	36.27	22.65	9.71
Base Classifier f_T	62.15	51.97	44.15
MLP	62.19	49.56	42.32
HTL-ITCNN	**63.08**	**53.6**	**45.50**

relative humidity, mean radiant temperature, and air velocity) and two personal factors (clothing insulation and metabolic rate). Since the prediction results of the PMV model are non-integer, we round its outputs to integers for classification performance comparison,

$$y_{pmv}^* = \begin{cases} +3, & \text{if } y_{pmv} \geq +3, \\ -3, & \text{if } y_{pmv} \leq -3, \\ Round(y_{pmv}), & \text{otherwise,} \end{cases} \quad (6)$$

where y_{pmv} is original PMV score and y_{pmv}^* is the 7-point thermal comfort scale ranging from -3 to +3, where the ascending integers denote cold, cool, slightly cool, neutral, slightly warm, warm and hot, respectively. Moreover, since we do not have related sensors to measure the mean radiant temperature (MRT) and air velocity (AV), we assume that the MRT is equal to the AT and the AV is equal to 0.1 m/s to calculate the PMV score.

5.2 Generic Thermal Comfort Model

We evaluate the performance of the generic thermal comfort model trained by the iTCM Generic Dataset.

5.2.1 Multi-classification Performance Analysis. With the class weight (CW) mechanism proposed in Section 4.3.1, we analyze the multi-classification performance of all classifiers for the source domain and the target domain, respectively.

- *Source Domain:* As shown in Table 6, the accuracy of the PMV model is slightly higher than the source-domain base classifier f_S^0. However, for another two metrics, f_S^0 is around twice as good as those of the PMV model. Also, from Figure 3, we can see that the PMV model prefers to predict the label 0 rather than other labels. In contrast, the f_S^0 can achieve average 31.0% accuracy on every class.
- *Target Domain:* Table 6 shows that the performance of all DNN-based approaches outperforms the PMV model significantly. The accuracy of Multi-Layer Perceptron (MLP) is almost equal to the original target-domain base classifier f_T, but for another two metrics, f_T performs better than the MLP. HTL-ITCNN achieves the best performance on all three metrics. Besides, Figure 4 shows that the PMV model still cannot classify each label of the target-domain dataset effectively.

The MLP already achieves excellent multi-classification performance, but it has a high probability of misjudging label +2. In contrast, HTL-ITCNN has average 53.8% accuracy on every class. The prediction accuracy of the label +2 is the lowest, and even though it also achieves 43% accuracy.

In conclusion, we prove that the PMV model is inaccurate to predict participants' thermal comfort sensations. In contrast, all DNN-based approaches have better multi-classification performance than the PMV model. Especially with our proposed class weight mechanism, HTL-ITCNN achieves the best performance and can classify labels effectively.

5.2.2 Performance Improvement Analysis. The performance improvement of HTL-ITCNN is boosted by two key factors: 1) the growing number of features; and 2) the transferred knowledge from the source domain. Since both factors are tightly linking to the modeling features, we define the following four feature sets to analyze the performance improvement.

- s_1: It contains four PMV factors: CI (F_6), AT (F_7), RH (F_8), and MR (F_9).
- s_2: It contains s_1 plus two vital signs factors: HR (F_{10}) and ST (F_{11}).
- s_3: It contains s_2 plus one time factor: hour (F_1).
- s_4: It contains s_3 plus four personal factors: age (F_2), gender (F_3), height (F_4), and weight (F_5).

After that, for comparison of the performance improvement, we use the same hyper-parameter settings to train the models based on each feature set. The results are listed in Table 7.

With the growing number of features from s_1 to s_4, the performance of the original target-domain base classifier f_T and HTL-ITCNN is increasing. Although we do not verify all features one by one, we still show that the newly added features have positive impacts on the thermal comfort modeling. Especially the personal factors, they can significantly improve the performance of HTL-ITCNN on all metrics. In contrast, the improvement boosted by transferred knowledge from the source domain is decreasing with the growing number of features, as shown in Figure 5. It is interesting that when adopting the same features for modeling, the transferred knowledge from the source domain has the highest compatibility on the target domain. With the growing number of features in the target domain, the positive impacts of the transferred knowledge from the source domain are diluted by newly added modeling features. The insight we can obtain is that Homogeneous Transfer Learning can achieve significant performance improvements via the better effectiveness of knowledge transfer than Heterogeneous Transfer Learning.

In summary, for modeling features, seeking highly-related new thermal comfort factors can observably improve the model prediction accuracy. For knowledge transfer, although the performance improvements are not impressive, we show the potential to apply the transfer learning principles to the thermal comfort research.

5.2.3 Performance Comparison. We compare the performance of HTL-ITCNN with some classic machine learning algorithms. The results are presented in Table 8.

First, we can find that the PMV model is just slightly better than the Random Guess. Then, with the increment of modeling

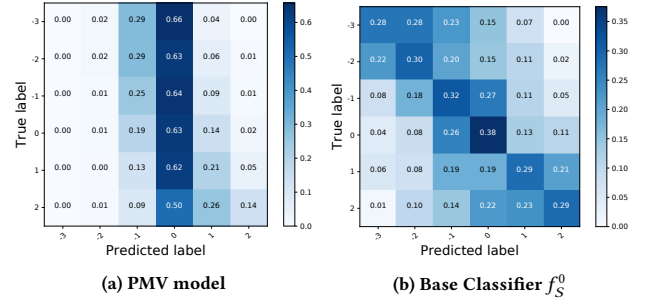
(a) PMV model

(b) Base Classifier f_S^0

Figure 3: Confusion matrix diagrams of the source domain. We can see that compared to the PMV model, the DNN-based classifier can better distinguish among all labels.

(a) PMV model

(b) MLP
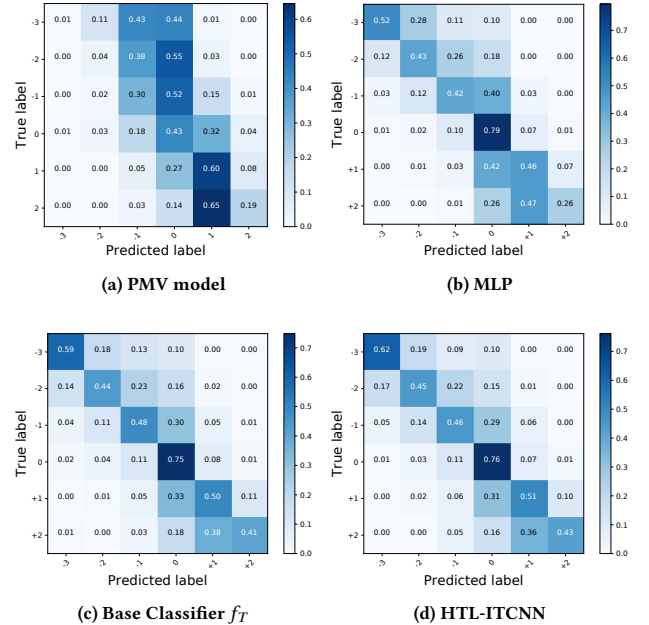
(c) Base Classifier f_T

(d) HTL-ITCNN

Figure 4: Confusion matrix diagrams of the target domain. We can find that compared to the PMV model, HTL-ITCNN has the remarkable advantages of prediction performance.

features, all learning-based approaches outperform the PMV model. Among the machine learning algorithms, the Naive Bayes has the lowest performance. But it still outperforms the PMV model by 28.5%, 33.2%, and 76.3% on the accuracy, macro-F1, and MCC, respectively. Compared to the Naive Bayes, Linear Support Vector Machine (SVM) has higher accuracy but has lower macro-F1 and MCC. RBF SVM has much better performance than the Linear SVM by 10.3%, 88.8%, and 98.6% on the accuracy, macro-F1, and MCC, respectively. The overall performance of the Decision Tree is better than Radial Basis Function (RBF) SVM, especially the macro-F1 and MCC. The k-Nearest Neighbors (k-NN) algorithm has quite good performance, whose accuracy is slightly higher than the MLP, but

Table 7: Prediction result comparison with different features

#	Classifiers	Accuracy(%)	Macro-F1(%)	MCC(%)
s_1	Base Classifier f_S^0	32.36	28.63	15.02
s_1	Base Classifier f_T	44.89	36.14	25.07
s_2		50.17	41.28	29.87
s_3		54.60	44.63	34.42
s_4		**62.15**	**51.97**	**44.15**
s_1	HTL-ITCNN	53.39	42.91	31.97
s_2		54.30	43.63	32.96
s_3		58.09	47.47	38.15
s_4		**63.08**	**53.06**	**45.50**

Table 8: Prediction result comparison among algorithms

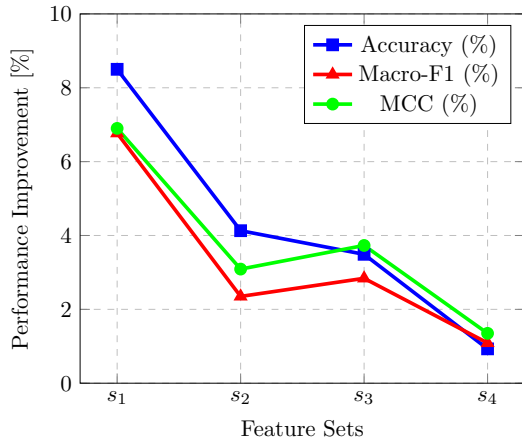| Algorithm	Accuracy(%)	Macro-F1(%)	MCC(%)
Random Guess	33.57	16.96	0.03
PMV	36.27	22.65	9.71
Naive Bayes	46.61	30.18	17.12
Linear SVM	53.58	18.62	15.48
RBF SVM	59.10	35.18	30.75
Decision Tree	60.01	49.07	40.76
MLP	62.19	49.56	42.32
k-NN	62.31	49.17	42.30
HTL-ITCNN	**63.08**	**53.06**	**45.50**
Random Forest	66.32	53.90	47.79

Figure 5: The performance improvement achieved by transferred knowledge with different feature sets. We can see that with the incremental number of non-overlapped features, the performance improvement achieved by knowledge transfer of overlapped features would decrease.

Table 9: Prediction results on three iTCM personal datasets

#	Algorithm	Accuracy(%)	Macro-F1(%)	MCC(%)
P_1	PMV	41.16	31.69	22.58
	HTL-ITCNN	36.84	20.88	8.72
	HTL-ITCNN-P_1	**62.85**	**49.07**	**45.06**
P_2	PMV	19.47	19.29	1.26
	HTL-ITCNN	35.32	26.31	11.07
	HTL-ITCNN-P_2	**65.26**	**54.99**	**53.43**
P_3	PMV	30.50	20.93	10.10
	HTL-ITCNN	35.87	23.08	8.20
	HTL-ITCNN-P_3	**64.20**	**46.84**	**47.26**

its macro-F1 and MCC scores are slightly lower than the MLP. It is surprising that k-NN achieves such outstanding performance. After analysis, the reason is that k-NN cannot handle the class imbalanced dataset which results in high prediction accuracy. For our proposed HTL-ITCNN, it outperforms the PMV model by 73.9%, 134.3% and 368.6% on the accuracy, macro-F1, and MCC, respectively. Moreover, we notice that compared to MLP and k-NN, HTL-ITCNN does not achieve remarkable improvements on the accuracy, but it significantly outperforms them on macro-F1 and MCC scores which presents its advantages of the high prediction accuracy on those labels with the lower distribution. However, HTL-ITCNN still fails to beat the Random Forest, which is recognized by academia as the best algorithm on small datasets (e.g., iTCM datasets).

Overall, machine learning techniques present significant advantages in thermal comfort modeling. Although our proposed HTL-ITCNN does not beat the Random Forest on small datasets, we will verify the performance of HTL-ITCNN on the large datasets to compare it with Random Forest in our future work.

5.3 Personal Thermal Comfort Model

To better fulfill the individual's thermal comfort demand, some researches [12, 22] propose to build personal thermal comfort models for individuals rather than a group of people. As such, we also develop personal HTL-ITCNN for performance evaluation.

As shown in Table 9, it is not surprising that all participants' personal thermal comfort models have the best accuracy. However, the prediction results of the PMV model and the generic HTL-ITCNN on different personal datasets are interesting.

- For the first participant P_1, the PMV model performs better than the generic HTL-ITCNN on all metrics. This case shows that the PMV model has better compatibility on this person than our approach, which confirms the value of PMV model.
- For the second participant P_2, the overall performance of the PMV model is quite bad. However, our generic HTL-ITCNN performs much better than the PMV model on all metrics. This case indicates that our approach has better compatibility on this person than the PMV model.
- For the third participant P_3, the PMV model achieves the acceptable performance, but our generic HTL-ITCNN outperforms the PMV model, especially on the accuracy and macro-F1 metrics. Combining with the previous two participants' cases, our approach has a stable prediction performance.

Overall, the performance of the generic HTL-ITCNN (average 36.01% accuracy) is better than the PMV model (average 30.38% accuracy). However, it still cannot predict every participant's thermal comfort accurately. In contrast, personal HTL-ITCNN demonstrates its outstanding prediction performance (average 64.10% accuracy). Therefore, we indicate that personal thermal comfort modeling, as one of the critical future research directions, has great potential.

6 CONCLUSIONS

Intending to enhance the prediction accuracy of the thermal comfort, we proposed an Heterogeneous Transfer Learning (HTL) based approach for thermal comfort modeling. We built our datasets and designed HTL-ITCNN. The experimental results show that the growing number of features and the transferred knowledge from the source domain have significant positive impacts on the performance improvement of HTL-ITCNN. Moreover, our HTL-ITCNN outperforms the PMV model and most of the machine learning algorithms. Furthermore, we also verify the performance of personal HTL-ITCNN and demonstrate the tremendous potential of the research direction on personal thermal comfort modeling.

ACKNOWLEDGMENTS

REFERENCES

[1] A Aggelakoudis and M Athanasiou. 2005. Thermal comfort study of occupants in University of Patras. In *International conference on passive and low energy cooling for the built environment*. Heliotopos Conferences, Santorini, Greece, 263–268.
[2] Takashi Akimoto, Shin-ichi Tanabe, Takashi Yanai, and Masato Sasaki. 2010. Thermal comfort and productivity-Evaluation of workplace environment in a task conditioned office. *Building and environment* 45, 1 (2010), 45–50.
[3] ASHRAE. 2017. ANSI/ASHRAE Standard 55. https://www.ashrae.org/technical-resources/bookstore/standard-55-thermal-environmental-conditions-for-human-occupancy. [Accessed on 18-Sep-2019].
[4] Liliana Barrios and Wilhelm Kleiminger. 2017. The Comfstat-automatically sensing thermal comfort for smart thermostats. In *Pervasive Computing and Communications (PerCom), 2017 IEEE International Conference on*. IEEE, IEEE, New York, NY, USA, 257–266.
[5] Christianne de Faria Coelho-Ravagnani, Flávia Carolina Lemos Melo, Fabricio CP Ravagnani, Franz Homero Paganini Burini, and Roberto C Burini. 2013. Estimation of metabolic equivalent (MET) of an exercise protocol based on indirect calorimetry. *Revista Brasileira de Medicina do Esporte* 19, 2 (2013), 134–138.
[6] Richard De Dear and Gail Schiller Brager. 1998. Developing an adaptive model of thermal comfort and preference. *ASHRAE Transactions* 104 (1998), 145–167.
[7] Poul O Fanger et al. 1970. *Thermal comfort. Analysis and applications in environmental engineering.* Copenhagen: Danish Technical Press., Randbølvej, 2720 København, Denmark. 244 pages.

[8] Veronika Földváry, Toby Cheung, Hui Zhang, Richard de Dear, Thomas Parkinson, Edward Arens, Chungyoon Chun, Stefano Schiavon, Maohui Luo, Gail Brager, et al. 2018. Development of the ASHRAE Global Thermal Comfort Database II. *Building and Environment* 142 (2018), 502–512.
[9] Jr. Frederick H. Rohles. 1979. The Effect of Time of Day and Time of Year on Thermal Comfort. *Proceedings of the Human Factors Society Annual Meeting* 23, 1 (1979), 129–132.
[10] Monika Frontczak and Pawel Wargocki. 2011. Literature survey on how different factors influence human comfort in indoor environments. *Building and environment* 46, 4 (2011), 922–937.
[11] Peter Xiang Gao and Srinivasan Keshav. 2013. SPOT: a smart personalized office thermal control system. In *Proceedings of the fourth international conference on Future energy systems*. ACM, ACM, New York, NY, USA, 237–246.
[12] Ali Ghahramani, Chao Tang, Zheng Yang, and Burcin Becerik-Gerber. 2015. A study of time-dependent variations in personal thermal comfort via a dynamic bayesian network. In *Sustainable human–building ecosystems*. ASCE, Reston, VA USA, 99–107.
[13] Peter Höppe. 2002. Different aspects of assessing indoor and outdoor thermal comfort. *Energy and buildings* 34, 6 (2002), 661–665.
[14] Tyler Hoyt, Stefano Schiavon, Alberto Piccioli, Dustin Moon, and Kyle Steinfeld. 2013. Center for the Built Environment(CBE) thermal comfort tool. http://comfort.cbe.berkeley.edu/. [Accessed on 18-Sep-2019].
[15] Weizheng Hu, Yonggang Wen, Kyle Guan, Guangyu Jin, and King Jet Tseng. 2018. iTCM: Toward Learning-Based Thermal Comfort Modeling via Pervasive Sensing for Smart Buildings. *IEEE Internet of Things Journal* 5, 5 (2018), 4164–4177.
[16] Madhavi Indraganti and Kavita Daryani Rao. 2010. Effect of age, gender, economic group and tenure on thermal comfort: a field study in residential buildings in hot and dry climate with seasonal variations. *Energy and buildings* 42, 3 (2010), 273–281.
[17] Nathalie Japkowicz. 2000. The class imbalance problem: Significance and strategies. In *In Proceedings of the 2000 International Conference on Artificial Intelligence*. CSREA Press, Las Vegas, Nevada, USA, 111–117.
[18] Nathalie Japkowicz and Shaju Stephen. 2002. The class imbalance problem: A systematic study. *Intelligent data analysis* 6, 5 (2002), 429–449.
[19] Ming Jin, Ruoxi Jia, and Costas J Spanos. 2017. Virtual occupancy sensing: Using smart meters to indicate your presence. *IEEE Transactions on Mobile Computing* 16, 11 (2017), 3264–3277.
[20] Wooyoung Jung and Farrokh Jazizadeh. 2017. Towards Integration of Doppler Radar Sensors into Personalized Thermoregulation-based Control of HVAC. In *Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments (BuildSys '17)*. ACM, New York, NY, USA, Article 21, 4 pages.
[21] Sami Karjalainen. 2007. Gender differences in thermal comfort and use of thermostats in everyday thermal environments. *Building and environment* 42, 4 (2007), 1594–1603.
[22] Emil Laftchiev and Daniel Nikovski. 2016. An IoT system to estimate personal thermal comfort. In *Internet of Things (WF-IoT), 2016 IEEE 3rd World Forum on*. IEEE, New York, NY, USA, 672–677.
[23] Abraham Hang-yat Lam, Yi Yuan, and Dan Wang. 2014. An Occupant-participatory Approach for Thermal Comfort Enhancement and Energy Conservation in Buildings. In *Proceedings of the 5th International Conference on Future Energy Systems (e-Energy '14)*. ACM, New York, NY, USA, 133–143.
[24] Weiwei Liu, Zhiwei Lian, and Yuanmou Liu. 2008. Heart rate variability at different thermal comfort levels. *European journal of applied physiology* 103, 3 (2008), 361–366.
[25] Weiwei Liu, Zhiwei Lian, and Bo Zhao. 2007. A neural network evaluation model for individual thermal comfort. *Energy and Buildings* 39, 10 (2007), 1115–1122.
[26] Antonio Martinez-Molina, Paola Boarin, Isabel Tort-Ausina, and José-Luis Vivancos. 2017. Post-occupancy evaluation of a historic primary school in Spain: Comparing PMV, TSV and PD for teachers' and pupils' thermal comfort. *Building and Environment* 117 (2017), 248–259.
[27] Faye C McQuiston and Jerald D Parker. 1982. *Heating, ventilating, and air conditioning: analysis and design.* John Wiley and Sons., Inc., New York, NY.
[28] Jonn Are Myhren and Sture Holmberg. 2008. Flow patterns and thermal comfort in a room with panel, floor and wall heating. *Energy and Buildings* 40, 4 (2008), 524–536.
[29] OA Seppänen, WJ Fisk, and MJ Mendell. 1999. Association of ventilation rates and CO2 concentrations with health andother responses in commercial and institutional buildings. *Indoor air* 9, 4 (1999), 226–252.
[30] Kevin Weekly, Ming Jin, Han Zou, Christopher Hsu, Alexandre Bayen, and Costas Spanos. 2014. Building-in-Briefcase (BiB). *arXiv preprint arXiv:1409.1660* arXiv:1409.1660, - (2014), –. http://arxiv.org/abs/1409.1660
[31] Wikipedia. 2015. Microsoft Band 2. https://en.wikipedia.org/wiki/Microsoft_Band_2. [Accessed on 18-Sep-2019].
[32] Wei Zhang, Weizheng Hu, and Yonggang Wen. 2018. Thermal Comfort Modeling for Smart Buildings: A Fine-Grained Deep Learning Approach. *IEEE Internet of Things Journal* 6, 2 (2018), 2540–2549.