Personalized Prediction of Indoor Comfort using Graph Convolutional Matrix Completion

Junyi Liu[1], Esha Naidu[2], Jialian Wu[1], Shira Gabriel[2], Edward Steinfeld[3], Junsong Yuan[1]
[1]Department of Computer Science, [2]Psychology, [3]Architecture, University at Buffalo

Abstract

Recent progress in environment sensing technology focuses more on measuring the physical properties of the environment, e.g., temperature and noise, but lacks the ability to understand subjective responses, or feelings about the environment, e.g., indoor comfort. Feelings depend on both environmental conditions and individual needs and preferences. Different people may feel differently in the same room experiencing the same conditions. In this work, we apply a crowdsensing based approach to predict personalized indoor comfort. We assume that similar users share similar feelings about comfort, and that indoor comfort is related to a fixed set of conditions, e.g., space, humidity, temperature. We surveyed existing users of a case study building and used their responses to learn how to predict the personal responses of new users. Technically, we apply a graph convolutional matrix completion (GC-MC) method to predict the comfort of other users, by learning the dependency between the user profiles and their ratings to a fixed set of survey questions. We collect a kitchen survey dataset of 59 questions and in total 29 users of diverse profiles.

1 Introduction

With the recent progress of Internet of Things (IoTs) and artificial intelligence (AI), “smart” systems have been employed in the indoor environment, such as smart lighting, heating, air-conditioning, and security. These smart solutions are designed to facilitate operations and improve efficiency but lack the ability to predict individual user feelings using such environments, e.g., the indoor comfort. As different users may have different feelings and experiences, it is not clear whether these smart systems benefit all inhabitants of a building, especially marginalized groups who often do not have a say in shaping policy and practices related to the control of indoor environment.

In this work, we perform personalized prediction of indoor comfort (PIC), using user ratings of a given indoor environment collected from survey data. We focus our study

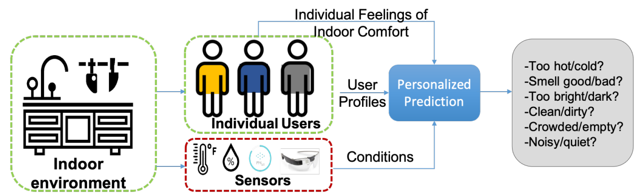

Figure 1. Illustration of personalized prediction of indoor comfort (PIC).

on the indoor comfort of a common kitchen in an office building. The goal is to predict responses of new users from existing user’s feedback about comfort. We collected data on subjective responses using an online survey questionnaire. The questionnaire includes 18 rating questions that cover different comfort aspects of the kitchen, such as temperature, humidity, and smell. The survey also requires users to input their personal profiles, such as demographics information, purpose of using the kitchen, and preferences for using the space, fixtures and appliances in the kitchen, so that we can leverage it to predict feelings of similar users. Using a prediction model and harnessing the volunteer labor of building inhabitants as “citizen scientists” will drastically reduce the cost of comfort monitoring by moving the smart sensing function from the building to the person to provide a “first person” perspective and result in portable data collection and flexible comfort prediction at low cost.

To predict comfort ratings of different users, we use a bipartite graph to represent relationships between individual users (user nodes) and their ratings (edges) of different survey questions (item nodes). To achieve our goal, we follow graph convolutional matrix completion (GC-MC) [4], a graph-based auto-encoder framework for matrix completion, and has been used to recommend films by predicting user ratings of new movies. The model combines user-item bipartite graph and nodes feature information to learn user and item embeddings. It captures the diversity of individual user profiles and question profiles, and can learn their associations, thus is suitable for predicting individual comfort feelings. Initial results show a 70% accuracy for predicting user ratings of the 18 survey questions.

2 Related Work

With the growth of the Internet of Things (IoTs), the environment data captured by sensors boost the development of machine learning for forecasting indoor environment. Several machine learning classification methods based on the radio frequency (RF) signal are explored to classify the type of indoor environments [1]. Attoue et al. [2] explored indoor and outdoor parameters to forecast the indoor temperature changes up to 2 hours with good performance on ANNs. Xu et al. [9] established a novel LSTM model modified by an error correction model to forecast indoor temperature for 5 to 30 minutes. However, they only considered the limited environmental attributes and monitorable data, ignoring user feelings about indoor environment.

3 Proposed Method

3.1 Graph Convolutional Matrix Completion

Assuming a matrix has n_1 rows and n_2 columns, but only a number of its entries can be observed, which is quite smaller than the total number of entries $n_1 \times n_2$. Matrix completion is the task of filling in the missing entries of the such observed matrix. Candès et al. [5] proposed using convex optimization to approximate the rating matrix with a low-rank rating matrix, such as Netflix problem [3]. In this paper, we apply the Graph Convolutional Matrix Completion (GC-MC) [4] to achieve the personalized prediction of indoor comfort. GC-MC is a graph-based auto-encoder framework for matrix completion, which is originally used for predicting user's ratings about movies, etc.

Fig. 2 shows the method for achieving the PIC task using GC-MC approach. A bipartite graph $G = (\mathcal{W}, \mathcal{E}, \mathcal{R})$ connects the user nodes (in circle) $u_i \in \mathcal{U}$ with $i \in \{1, ..., N_u\}$ and feeling question nodes (in square) $v_j \in \mathcal{V}$ with $j \in \{1, ..., N_v\}$, where $\mathcal{U} \cup \mathcal{V} = \mathcal{W}$. N_u is the number of users while N_v is the number of questions. Each edge (u_i, r, v_i) in the bipartite graph will record the rating r of user u_i towards the question v_j, where $r \in \{1, ..., R\} = \mathcal{R}$ and $(u_i, r, v_i) \in \mathcal{E}$. In such a case, our PIC task becomes a matrix completion problem: given a matrix M of size $N_u \times N_v$, like a user-item matrix in [4], each matrix element records a rating of user u_i towards the question v_j. If we know some of the ratings from a subset of users to a subset of questions, can we predict the remaining ratings in the same matrix?

In our method, each user node is represented by the user's demographics information and preference, while each question node is characterized by the environment factors that are associated with that question. In the matrix M, there are many unobserved ratings as each user may not answer all questions, and the new users have not answered any comfort feeling questions. To predict the ratings of new users, these nodes are represented by latent embeddings learned from a message passing procedure and node attributes, then the bilinear decoder utilizes the latent embeddings to complete the user-feeling matrix and produce feeling predictions for the testing users in the bipartite graph (red line in Fig. 2).

3.2 Node Representations

User Profile: During our survey, we ask questions to identify user profile. Each user is represented as a vector which includes not only the demographics information of the user, but also the preferences of the user. We assume if two users have similar profile, they will share similar ratings regarding the feelings questions. This can help to predict the ratings of new users.

Question Profile: For comfort feeling questions, one may be related to the other, like temperature feeling is related to humidity feeling. Thus for a given user, the rating of one question may influence the rating of a related question. To learn the connections among these questions, we use multi-hot vectors to represent feeling question attributes as shown in Fig. 3, which is based on the environment factors that need to be considered when answering the corresponding feeling questions. For example, people may consider the smell and noise of a room when they plan to have a social activity. The question representation should be similar if the questions consider similar factors. For example, the attribute vector for noisy level question is the same as that for sound level from people.

As an alternative, we can leverage state-of-the-art Natural Language Processing (NLP) model to characterize each feeling question, e.g through BERT [6], a pre-trained language representation model. BERT often achieves excellent performance in fields such as question answering [7], sentiment analysis [10]. In this work, we use it to encode each question sentence to a fixed length vector [8], e.g., a 768-dimensional vector to represent each question.

3.3 Graph Encoder

To obtain node embeddings, we use two encoders: attribute encoder and interaction encoder. The attribute encoder produces node embeddings in terms of the node attributes. For example, the attribute embedding of user i can be obtained by $a_i = \sigma(W_a x_i^a)$, where $W_a \in \mathbb{R}^{H_0 \times N_a}$ is a learnable parameter, H_0 is the output dimension of attribute encoder, and $x_i^a \in \mathbb{R}^{N_a}$ is the user profile vector that presents user i. N_a is the number of profile attributes. The interaction encoder produces node embeddings with the information of other nodes through *message passing network*. Let $\mu_{j\to i,r}$ be the edge-type specific message $\mu_{j\to i,r}$ from feeling node j to user node i, which is derived by $\mu_{j\to i,r} = \frac{1}{d_{ij}} W_r x_j$, where $W_r \in \mathbb{R}^{H_1 \times (N_u + N_v)}$ is the trainable parameter for a specific ordinal degree r. d_{ij}

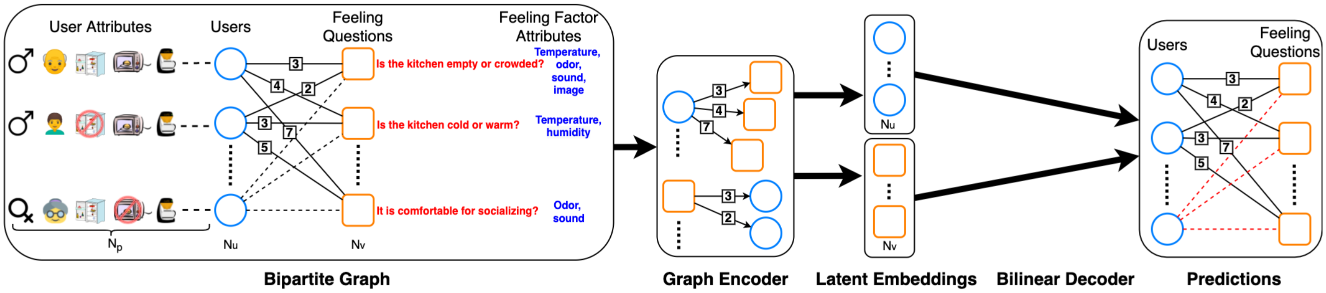

Figure 2. Personalized prediction of indoor comfort. The dotted edges in the bipartite graph indicate the unobserved feelings that we aim to predict for the incoming users during testing.

	Temperature	Humidity	Odor	Sound	Image	Others
Q1: How crowded is the room?	✓(1)	X(0)	✓(1)	✓(1)	✓(1)	X(0)
Q2: How is the temperature in the summer?	✓(1)	✓(1)	X(0)	X(0)	X(0)	X(0)
Q3: How is the humidity?	✓(1)	✓(1)	X(0)	X(0)	X(0)	X(0)
Q4: How is the ventilation?	✓(1)	✓(1)	✓(1)	X(0)	X(0)	X(0)
Q5: How is the overall noise?	X(0)	X(0)	X(0)	✓(1)	X(0)	X(0)
Q6: How is the noise from appliances?	X(0)	X(0)	X(0)	✓(1)	X(0)	X(0)
Q7: How is the sound from people in the room?	X(0)	X(0)	X(0)	✓(1)	X(0)	X(0)
Q8: How is the sound from people outside of the room?	X(0)	X(0)	X(0)	X(0)	X(0)	✓(1)
Q9: How is the smell?	X(0)	X(0)	✓(1)	X(0)	X(0)	X(0)
Q10: Is the room closest to another important place?	X(0)	X(0)	X(0)	X(0)	X(0)	✓(1)
Q11: Is the temperature comfortable?	✓(1)	✓(1)	X(0)	X(0)	X(0)	X(0)
Q12: How frequently used is this room?	X(0)	X(0)	X(0)	X(0)	✓(1)	X(0)
Q13: Is the room a good place to socialize?	X(0)	X(0)	✓(1)	✓(1)	X(0)	X(0)
Q14: Are the amenities better than in other rooms?	X(0)	X(0)	X(0)	X(0)	✓(1)	✓(1)
Q15: Is the acoustics better than the other rooms?	X(0)	X(0)	X(0)	✓(1)	X(0)	✓(1)
Q16: How is the lighting?	X(0)	X(0)	X(0)	X(0)	✓(1)	X(0)
Q17: Is the room a pleasant space?	✓(1)	✓(1)	✓(1)	✓(1)	✓(1)	✓(1)
Q18: Is the room a good place to meet people?	X(0)	X(0)	✓(1)	✓(1)	X(0)	X(0)

Figure 3. Multi-hot representations of indoor comfort survey questions according to the environment factors the question is associated with.

is a normalization constant $\sqrt{|\mathcal{N}_i|}$ or $\sqrt{|\mathcal{N}_i||\mathcal{N}_j|}$ with $\mathcal{N}_i$ is the set of first-hop neighbours of user i. Then we aggregate all edge-type messages $r \in R$ to produce the interaction embedding:

$$h_i = \sigma\left[\text{concat}\Big(\sum_{j\in\mathcal{N}_{i,1}} \mu_{j\to i,r}, ..., \sum_{j\in\mathcal{N}_{i,R}} \mu_{j\to i,R}\Big)\right], \quad (1)$$

where concat$(\cdot)$ is concatenation operation of vectors, $\sigma(\cdot)$ is an activation function such as ReLU$(\cdot)$, and $h_i \in \mathbb{R}^{H_2}$.

Lastly we aggregate attribute embedding and interaction embedding to generate the final user node embedding as:

$$s_i = \sigma[W \cdot \text{concat}(h_i + a_i) + b] \quad (2)$$

where $W \in \mathbb{R}^{H_3\times(H_0+H_2)}$ includes learnable parameters and b is a bias. The node embedding of feeling question t_j is encoded in the same way. Both $s_i \in \mathbb{R}^{H_3}$ and $t_j \in \mathbb{R}^{H_3}$ are vectors of dimension H_3; in our experiments we set $H_0 \in \{10, 20, 30, 40, 50\}$, $H_1 = 50$, $H_2 = 250$, $H_3 = 75$.

3.4 Bilinear Decoder

Following [4], the bilinear decoder takes the node embeddings as input and complete the input matrix M to M^*. The feeling predictions for the incoming users, i.e., the predicted edge (red line in Fig. 2), can be obtained from M^*. In order to reconstruct the matrix M^*, the bilinear decoder treats each one of the five ratings as a separate class. The decoder uses softmax function to generate a probability distribution for each rating as follows:

$$p(M^*_{ij} = r) = \frac{e^{s_i^T Q_r t_j}}{\sum_{k\in R} e^{s_i^T Q_k t_j}}, \quad (3)$$

where Q_r with shape of $H_3 \times H_3$ is a learnable parameter matrix. The final predicted degree of feeling is calculated by following:

$$M^*_{ij} = \mathbb{E}_{p(M^*_{ij}=r)}[r] = \sum_{r\in R} r\, p(M^*_{ij} = r). \quad (4)$$

3.5 Loss Function

The loss is calculated by the negative log likelihood as:

$$Loss = -\sum_{i,j;O_{ij}=1}\sum_{r\in R} y_{ij} \log p(M^*_{ij} = r), \quad (5)$$

where $y_{ij} = 1$ if $r = M_{ij}$ and zero otherwise. $O \in \{0, 1\}$ indicate whether the user i is incoming user or not.

4 Experiments

Dataset. To perform our experiments, we perform a user study through a public kitchen in an office building. The survey data mainly contain two types of questions: 1) user profiling questions that includes user demographics information and his/her preferences; these are multiple choices questions and user can choose one of them that best reflects the user profile or the preference; 2) user's rating about the indoor comfort (examples in Fig. 3). The ratings have 5 levels, with 3 the neutral rating. We survey the indoor comfort feelings from 29 users with 59 questions, including 41 questions for user profile, e.g., demographics information and user preference, and 18 questions for how they feel about the kitchen environment. Our task is to predict the new user ratings of these 18 questions.

Results. In Table 1, we use accuracy and approximate accuracy (Approx accuracy) (as shown in Eq. (6) & (7)) to quantitatively evaluate matrix completion. In Eq. (6) & (7), n is the total number of predictions, y_i is the ground truth user feeling, and y_p is the prediction of user feeling. Moreover, we test different proportion [20%, 40%, 50%, 60%, 80%] of entries in user-question matrix M as the training data, and the rest entries as the testing data. Following the evaluations of matrix completion, we first evaluate the results of the GC-MC model with BERT representation and multi-hot representation. As shown in Table 1, the GC-MC model can make most of correct predictions, and the model with multi-hot representation performs better than with BERT representation when using a small amount of training dataset. Both representations can make over 90% approximate accuracy, that means the model can predict similar or close feelings to users.

In Table 2, we test the influences of using different number of users for training. We evaluate the results of random choice and the GC-MC model with different representations for our task. The random choice means we randomly choose the option of each feeling question as the predicted feelings for testing users. We choose different proportion, i.e., 50% or 80% of users for training and the remaining users are treated as new users for testing. The result indicates that the GC-MC model can make most of correct predictions for new users and meet the feelings of most people.

$$\text{Accuracy} = \frac{\sum_{i=1}^{n} \mathbb{1}(|y_i - y_p| == 0)}{n} \tag{6}$$

$$\text{Approx Accuracy} = \frac{\sum_{i=1}^{n} \mathbb{1}(|y_i - y_p| <= 1)}{n} \tag{7}$$

Table 1. Comparative study of different size of training data

REPR	Training Pct.	Accuracy	Approx Accuracy
Multi-hot Ques.	20%	69.30%	93.76%
Multi-hot Ques.	40%	72.29%	93.95%
Multi-hot Ques.	50%	75.48%	94.64%
Multi-hot Ques.	60%	73.56%	96.15%
Multi-hot Ques.	80%	78.10%	97.14%
BERT Ques.	20%	68.82%	93.76%
BERT Ques.	40%	71.97%	93.95%
BERT Ques.	50%	73.18%	94.64%
BERT Ques.	60%	75%	96.63%
BERT Ques.	80%	79.05%	97.14%

Table 2. Comparative study of different size of training users

REPR	Training Pct.	Accuracy	Approx Accuracy
Random Choice	-	21.07%	60.54%
Multi-hot Ques.	50%	72.99%	94.83%
Multi-hot Ques.	80%	74.52%	94.83%
BERT Ques.	50%	73.56%	94.83%
BERT Ques.	80%	74.33%	94.83%

5 Conclusion

In this paper, we propose to apply Graph Convolutional Matrix Completion for personalized prediction of indoor comfort. We collect a user survey data of a specific kitchen environment, then apply the GC-MC to encode the representations of both users and the feeling questions to predict individual feelings of new users. We show a reasonable result on predicting feedback about indoor comfortability.

Acknowledgement. This work is supported in part by National Science Foundation Grant CNS1951952.

References

[1] M. I. AlHajri, N. T. Ali, and R. M. Shubair. Classification of indoor environments for iot applications: A machine learning approach. *IEEE Antennas and Wireless Propagation Letters*, 17(12):2164–2168, 2018.

[2] N. Attoue, I. Shahrour, and R. Younes. Smart building: Use of the artificial neural network approach for indoor temperature forecasting. *Energies*, 11(2):395, 2018.

[3] J. Bennett and S. Lanning. The netflix prize. In *Proceedings of KDD cup and workshop*, volume 2007, page 35. New York, NY, USA., 2007.

[4] R. Berg, T. N. Kipf, and M. Welling. Graph convolutional matrix completion. *arXiv preprint arXiv:1706.02263*, 2017.

[5] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational mathematics*, 9(6):717–772, 2009.

[6] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[7] C. Qu, L. Yang, M. Qiu, W. B. Croft, Y. Zhang, and M. Iyyer. Bert with history answer embedding for conversational question answering. In *Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval*, pages 1133–1136, 2019.

[8] H. Xiao. bert-as-service. https://github.com/hanxiao/bert-as-service, 2018.

[9] C. Xu, H. Chen, J. Wang, Y. Guo, and Y. Yuan. Improving prediction performance for indoor temperature in public buildings based on a novel deep learning method. *Building and Environment*, 148:128–135, 2019.

[10] H. Xu, B. Liu, L. Shu, and P. S. Yu. Bert post-training for review reading comprehension and aspect-based sentiment analysis. *arXiv preprint arXiv:1904.02232*, 2019.