Rethinking Comfort Profiles in Adaptive Building Energy Management Systems*

Jennifer Williams, Elnaz Shafipour, Alvin Shi, and Sebastian Stein

School of Electronics and Computer Science
University of Southampton, Southampton, England, UK
{j.williams, e.shafipour, gs1n17, s.stein}@soton.ac.uk

Abstract. As standard building occupancy schedules continue to change from static closed-door offices to dynamic open office layouts, we face new challenges for developing smart building energy management systems (BEMS) that can simultaneously *adapt* to save energy costs, while also incorporating the *comfort preferences* of the occupants. This is especially true for certain building types which by design are open layout, or partially-open layout such as schools, hospitals, and libraries. In this paper, we identify and explain three of the most critical challenges that specifically relate to incorporating feedback from building occupants into an interactive reinforcement learning algorithm. For each challenge, we propose how the challenge could be dealt with practically, within the context of our ongoing work and experimentation in this area. Overcoming these challenges opens new opportunities for artificial intelligence solutions that will place citizens in the centre and also help smart building designers move toward net-zero goals.

Keywords: building energy management · comfort profiles · reinforcement learning.

1 Introduction

Generally speaking, building energy management systems (BEMS) include both hardware components and software algorithms that control indoor climate such as heating, ventilation, and air conditioning (HVAC), indoor air temperature, indoor air quality, humidity, lighting, and certain sanitation equipment. All of these comfort-oriented components consume electricity and contribute to the overall cost for facilities to operate. Not only does energy for building operation take up nearly one third of energy consumption in the world [1], it is also estimated that the global market for BEMS and solutions based on artificial intelligence (AI) will reach (USD) $7.3 billion by 2026 [2]. However, most BEMS that are currently deployed in the real world do not have the capability to respond

* This work was supported by the UKRI Turing AI Acceleration Fellowship on Citizen-Centric AI Systems (EP/V022067/1) and by the Southampton Low Carbon Comfort Centre.

to changing occupancy schedules[1]. Most of the researched solutions to energy management involve a machine learning technique called reinforcement learning (RL [3] but these algorithms take a limited or naive view of both occupancy and comfort, such as treating occupancy as a binary problem (occupied/unoccupied) and comfort as a set room temperature for an entire building. These naive assumptions create a gap between simulation and real-world deployment.

AI researchers, engineers, and companies who work in this area envision smart buildings of the future that are adaptive, meaning that they account for natural shifts in occupancy levels and comfort needs throughout the course of the day. Further, there have been more recent efforts to include occupant *comfort preferences* into the software algorithms that manage building resources [4], rather than relying on a set temperature for an entire building or floor. Most of the energy consumed by a building is related to maintaining thermal comfort [5]. However, if a room, zone, or floor of a building has low occupancy, this can result in wasted energy and higher costs for building operators. From an industry perspective, goals for companies may include employee health, productivity, and safety, rather than just the energy saving and thermal comfort. Balancing these complex goals is very challenging to balance these goals.

Recent AI trends have studied BEMS in terms of reinforcement learning (RL) algorithms [6–9]. RL algorithms provide a means to optimize single or multiple objectives. In the work that we propose for balancing comfort preferences and energy consumption, we find that multi-objective reinforcement learning (MORL) provides the best opportunity for learning policies that optimize multiple objectives simultaneously [10]. Further, this type of problem also benefits from interactive reinforcement learning (IRL) [11] in order to incorporate feedback from building occupants when they are able to provide it. As we describe in this paper (Section 5), designing a MORL algorithm that performs meaningfully and is easy to train is a very challenging task. At the same time, previous research typically presents different models for different buildings, each with unique characteristics. Developing a one-size-fits-all solution is still not yet possible and is outside of the scope of this paper.

Designing an RL algorithm paradigm from which to work is only one step toward solving the technical challenges surrounding the problem of optimizing comfort preferences and energy consumption. In the case of comfort preferences, while there has been prior work relating to collecting [12], learning [4] and also aggregating [13] multiple inputs from people, this is far from a solved problem. In fact, a recent study [14] has highlighted more deeply that current research on comfort preferences falls short, especially because there are multiple definitions of comfort and needs may differ depending on the zone and use-case for a building.

The purpose of this paper is to examine three main challenges to comfort profiles for BEMS based on very recent advances in the state of the art for reinforcement learning. This overview is important because they take into consideration a realistic and interactive use-case for a deployed BEMS, while also considering that most research and development takes place in a simulated envi-

[1] https://www.znealliance.org/acco-bems

ronment. Addressing these challenges is another step closer toward more efficient and trustworthy BEMS whose foremost purpose is to service citizen comfort and safety, while also working toward sustainability goals. We present the following three challenges along with suggested solutions based on ongoing work and experimentation:

1. Challenge 1: What kind of information needs to be collected from occupants and how often? (Section 3)
2. Challenge 2: How should occupant comfort profiles be aggregated? (Section 4)
3. Challenge 3: How can comfort profiles be incorporated into a reinforcement learning algorithm? (Section 5)

2 Background and Related Work

Previous work from [15] has examined how human behaviour changes when people are provided with a smart thermostat in their home, which may also enable them to make better economical decisions about their energy consumption with respect to price. The authors analyzed users' preferred temperature set-points at various times of the day, using different costs for electricity price-points. They found that many users were willing to reduce their electricity consumption when prices were high, even if that meant that their home temperature changed from their set-point. As the authors note, one limitation was that they used room temperature as a proxy for user comfort. They further did not take into account that comfort preferences may change throughout the day, or have shifting priorities due to user-intrinsic factors (e.g. disability, age, or short-term illness).

Recent work has introduced Gnu-RL[2] [16] which first learns from historical to pre-train policy gradients in order to reduce overall training time of the RL algorithm. This pre-training step makes the algorithm "precocial" in that it is already mature before RL training, which significantly reduces overall training time by a factor of simulation *decades* [17]. However, the Gnu-RL algorithm makes some assumptions which contribute major limitations, and which have not yet been addressed in research. It assumes that building dynamics are locally linear and also the algorithm is also missing an interactive component. As [18] point out, the linear design means that it is difficult to adapt to real-world settings that are dynamic, such as a hospital building or a college campus. Without an interactive component, it is not possible to incorporate occupant feedback into the algorithm in real-time.

Handling occupant interaction is itself a difficult problem. For example, [19] treat interaction from building occupants as evidence that they are dissatisfied based on the assumption that if people were already content and comfortable, they are not likely to provide positive feedback about their comfort. Counting the number of times that occupants submit their feedback is not ideal for large buildings, wherein the temperatures may be different depending on the zones and

[2] https://github.com/INFERLab/Gnu-RL

occupancy levels. Their RL algorithm uses deep learning to balance and optimize both comfort and energy by attempting to simultaneously minimize discomfort and minimize energy consumption. However, their approach does not attempt to represent that comfort levels for some occupants may require a priority as with disabled, elderly, or children.

3 Challenge 1: Comfort Preference Collection

Our first challenge relates to how data is collected for modelling comfort preferences. In the first instance, it is necessary to have a large dataset to train an RL algorithm. However, since comfort preferences and profiles are not part of existing datasets, this poses a challenge for modelling interaction in the algorithm training. Another facet to this challenge is that it is unknown exactly what type of information must be collected. For example, in an office building where the occupants are recurring, it may be easiest to allow the occupants to fill in a *profile*, for example using an app or other web interface. The profile information can be stored and accessed by the BEMS and RL algorithm to ensure that their office and area are comfortable when occupied. For other types of buildings such as schools or hospitals, where occupancy is changing, the BEMS would benefit from regular feedback so that it can adapt, but not so much feedback that occupants will become annoyed by answering questions.

We propose to conduct an online survey where we provide participants with an imaginary scenario that describes "their office" and ask them to enter information about their preferred temperature and ventilation set-points at various times of day. This information would allow us to create realistic *comfort profiles* that can be incorporated into the design of our ideal BEMS RL algorithm. We can further ask participants to envision the circumstances where they would become annoyed by providing real-time feedback. While this type of survey does have its limitations, it may help with overcoming a lack of explicit comfort profiles in publicly available datasets that are used alongside highly-detailed building dynamics simulations like EnergyPlus[3].

4 Challenge 2: Comfort Profile Aggregation

When we discuss comfort profiles, a very important element of that conversation is how to manage multiple different comfort profiles at the same time, in a BEMS. This is a very challenging research area. In fact [20] present this as a problem of learning different control policies that will allow for human behaviour to change over time, and change differently for each occupant. There is currently no suitable algorithmic approach from reinforcement learning that can manage that level of ongoing uncertainty. Towards a solution, we propose that some occupants in the building may require a weighting for their comfort profile, which would allow their preferences to take priority over other occupants. This weighting would be

[3] https://energyplus.net/

ideally related to health and safety concerns, rather than unrelated factors such as authority or financial status. It is important to consider that this type of data collection involves collecting data about occupants' behaviour and preferences, which can raise privacy concerns in the case that data is not properly anonymized or some occupants are not comfortable sharing their information. Privacy is something that all researchers dealing will need to consider.

We can further consider aggregating comfort profiles in terms of zones. For example, in an open-office layout with hot-desks, it may be preferable to offer occupants that certain zones have particular properties and allow those occupants to choose their zone accordingly.

5 Challenge 3: Incorporating Comfort into Reinforcement Learning Algorithms

In our work, we assume that there is a single-level building with different zones (e.g. an office building). Each zone has a number of occupants who may enter and their zone, or the building itself. The goal is to set the temperature of each zone to maximize the satisfaction level of the occupants. Also, to be able to understand the satisfaction level of the occupants, we will interact with them and get their feedback. As described earlier, in simulation experiments, we may utilize realistic comfort settings gathered from a survey so that our simulation reflects preferences from real people rather than randomly contrived values.

In a standard RL problem, a learning agent observes the resulting environment transitions in a number of discrete steps and learns the control policy to maximize the accumulated reward. However, in this work, we will use Interactive RL to solve this problem, which will allow us to get feedback from people in the building and adjust the building's temperature based on that feedback and other observations.

In our model, we define each zone as a tuple $z(p, t)$ where p is the number of people and t is the current temperature of that zone. Then, we define the state space as a set of zones $S = \{z_1, z_2, ..., z_n\}$. Also, the action that we take at each iteration would be a list of temperatures $a = \{t_1, t_2, ..., t_n\}$.

In this work, we will get the following information from the user.

- Specifying occupants' preferred temperature.
- Ranking different actions/schedules for a particular day.
- Occupant's zone number
- Their priority level.

Using this information we can train our initial MORL as we will have multiple zones (as well as multiple objectives) and we should set the temperature for each. Later, using interactive RL, we will require users to evaluate the system's performance qualitatively. Our expected feedback from the user would be:

- Are they present in their zone?
- Did they change their zone?

- Their responses to the thermal environment (optimal temperature set point for the zone they are currently in)
- Giving binary (positive/negative) feedback and saying in general whether they are happy with the temperature or not.

6 Discussion and Future Work

We have highlighted the current state-of-the-art for AI-based approaches to optimizing building energy and occupant comfort. While there have been many recent advances, some challenges still remain. We introduced three such challenges and described our proposed approach for each, which we are continuing to explore in our work.

Our challenges have addressed that we may assign individual weights for particular comfort profiles, and aggregate them based on this information. This adds a layer of complexity and allows us to assign higher weight to occupants with special needs, such as people with a health condition (e.g. respiratory problems who may require better air quality), disability, children and the elderly. In our future work, we will develop a new objective measure that will help us determine if an occupant should have priority. We will also use a survey to get input from people about whether they think these are fair and equal decisions. Otherwise, there is a risk of conflicts or dissatisfaction among occupants who feel that their needs are not being met.

Our proposed model will require complimentary techniques of interactive RL (IRL) as well as multi-objective RL (MORL). Such a system will be very complex to design and train, which is why we are sharing the challenges of this work at the outset. We hope that other researchers in the community who work on BEMS will share our interest in re-thinking how to best handle occupant comforts alongside energy.

References

1. T. L. Vu, N. T. Le, Y. M. Jang, *et al.*, "An overview of internet of energy (IoE) based building energy management system," in *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 852–855, 2018.
2. I. Global Industry Analysts, "Global building energy management systems (BEMS) market to reach $7.3 billion by 2026," Mar 2022.
3. R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction.* MIT press, 2018.
4. Y. Xu, S. Chen, M. Javed, N. Li, and Z. Gan, "A multi-occupants' comfort-driven and energy-efficient control strategy of VAV system based on learned thermal comfort profiles," *Science and Technology for the Built Environment*, vol. 24, no. 10, pp. 1141–1149, 2018.
5. L. Yang, H. Yan, and J. C. Lam, "Thermal comfort and building energy consumption implications–a review," *Applied energy*, vol. 115, pp. 164–173, 2014.

6. E. Mocanu, D. C. Mocanu, P. H. Nguyen, A. Liotta, M. E. Webber, M. Gibescu, and J. G. Slootweg, "On-line building energy optimization using deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3698–3708, 2018.
7. Y. Pan and L. Zhang, "Roles of artificial intelligence in construction engineering and management: A critical review and future trends," *Automation in Construction*, vol. 122, p. 103517, 2021.
8. F. Wang, L. Zhou, H. Ren, X. Liu, S. Talari, M. Shafie-khah, and J. P. Catalao, "Multi-objective optimization model of source–load–storage synergetic dispatch for a building energy management system based on TOU price demand response," *IEEE Transactions on Industry Applications*, vol. 54, no. 2, pp. 1017–1028, 2017.
9. Z. Wang and T. Hong, "Reinforcement learning for building controls: The opportunities and challenges," *Applied Energy*, vol. 269, p. 115036, 2020.
10. K. Van Moffaert and A. Nowé, "Multi-objective reinforcement learning using sets of pareto dominating policies," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 3483–3512, 2014.
11. A. L. Thomaz, G. Hoffman, and C. Breazeal, "Real-time interactive reinforcement learning for robots," in *AAAI 2005 workshop on human comprehensible machine learning*, vol. 3, 2005.
12. F. Jazizadeh, A. Ghahramani, B. Becerik-Gerber, T. Kichkaylo, and M. Orosz, "Human-building interaction framework for personalized thermal comfort-driven systems in office buildings," *Journal of Computing in Civil Engineering*, vol. 28, no. 1, pp. 2–16, 2014.
13. K. Aduda, W. Zeiler, and G. Boxem, "Smart Grid-BEMS: the art of optimizing the connection between comfort demand and energy supply," in *2013 Fourth International Conference on Intelligent Systems Design and Engineering Applications*, pp. 565–569, 2013.
14. J. Williams, B. Lellouch, S. Stein, C. Vanderwel, and S. Gauthier, "Low-carbon comfort management for smart buildings," in *2022 IEEE International Smart Cities Conference (ISC2)*, pp. 1–5, 2022.
15. "Save money or feel cozy?: A field experiment evaluation of a smart thermostat that learns heating preferences," in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, vol. 16, International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2017.
16. B. Chen, Z. Cai, and M. Bergés, "Gnu-RL: A precocial reinforcement learning solution for building HVAC control using a differentiable MPC policy," in *Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, pp. 316–325, 2019.
17. Z. Zhang and K. P. Lam, "Practical implementation and evaluation of deep reinforcement learning control for a radiant heating system," in *Proceedings of the 5th Conference on Systems for Built Environments*, pp. 148–157, 2018.
18. Y. Lei, S. Zhan, E. Ono, Y. Peng, Z. Zhang, T. Hasama, and A. Chong, "A practical deep reinforcement learning framework for multivariate occupant-centric control in buildings," *Applied Energy*, vol. 324, p. 119742, 2022.
19. L. Scarcello and C. Mastroianni, "Cognitive systems for energy efficiency and thermal comfort in smart buildings," in *IoT Edge Solutions for Cognitive Buildings*, pp. 329–345, Springer, 2022.
20. H. Zhang, D. Wu, and B. Boulet, "A review of recent advances on reinforcement learning for smart home energy management," in *2020 IEEE Electric Power and Energy Conference (EPEC)*, pp. 1–6, 2020.