Data-Driven Thermostats - For Feedback, Comfort, and Reliability

by

Milan Jain

Submitted to the Department of Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

INDRAPRASTHA INSTITUTE OF INFORMATION TECHNOLOGY DELHI

March 2019

Author ....
Department of Computer Science
March 30, 2019

Certified by ....
Dr. Amarjeet Singh
Assistant Professor, IIIT Delhi
Thesis Supervisor

Certified by ....
Dr. Vikas Chandan
Scientist, Pacific Northwest National Laboratory
Thesis Supervisor

Accepted by ....
Department of Computer Science
IIIT Delhi

Data-Driven Thermostats - For Feedback, Comfort, and Reliability

by

Milan Jain

Submitted to the Department of Computer Science
on October 10, 2019, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In buildings, air conditioning consumes a significant proportion of the aggregate electricity bill. Specifically, in residential and small-scale commercial buildings, people prefer to use room-level ACs (AC stands for Air Conditioning in this thesis) with an in-built thermostat. As thermostat settings conceptually govern AC energy consumption and user comfort; over the years, researchers spent a significant amount of time and effort in enhancing thermostats to ensure optimal usage of ACs. By analyzing occupants' behaviour, today, smart thermostats, with multiple sensory abilities, can automatically adjust the set temperature to maximize both - energy savings and user comfort.

While thermostats are smart and ubiquitous, they often rely on occupants' dynamic schedule for the automated control of the set-point temperature. For a typical home, where everyone follows a particular routine, any deviation in daily schedule often leads to user discomfort. In addition to that, smart thermostats neither consider spatial variations across the buildings, nor temporal variations, such as climate change, while changing the set temperature. Subsequently, even today, the expensive thermostats are confined to automated temperature variation, with a limited scope of boosting the energy savings and enhancing the user comfort. In this dissertation, we address these concerns and introduce Data-Driven Thermostats to make AC experience efficient and comfortable for the users.

First, we propose PACMAN that monitors room temperature to ensure tenants' participation during AC usage by providing actionable energy-feedback. Next, we recommend a Comfort-Energy Trade-off (CET) knob, realized through an optimization framework, to allow users to balance their comfort and savings without worrying about the right set temperature. Our study indicates that such a knob can reduce residents' discomfort by 23% and save 26% energy. Third, we investigate the impact of occupancy prediction errors on occupants' comfort and total energy consumption of a building. Finally, we propose *Greina* - to continuously monitor the readily available ambient information from the thermostat and timely report refrigerant leaks through the coils (or valves) of a refrigeration unit. Such leaks waste significant energy, risk occupants' health, and affect user comfort. Our methods are novel, scalable, and more effective than the state-of-the-art smart thermostats.

Thesis Supervisor: Dr. Amarjeet Singh
Title: Assistant Professor, IIIT Delhi

Thesis Supervisor: Dr. Vikas Chandan
Title: Scientist, Pacific Northwest National Laboratory

This doctoral thesis has been examined by a Thesis Committee as follows:

Dr. Amarjeet Singh .... Amarjeet Singh ....

Thesis Supervisor
Assistant Professor of Computer Science, IIIT Delhi

Dr. Vikas Chandan ....

Thesis Supervisor
Research Scientist, Pacific Northwest National Laboratory

Prof. Alex Rogers ....

Member, Thesis Committee
Professor of Computer Science, University of Oxford

Prof. Elena Gaura ....

Member, Thesis Committee
Professor of Pervasive Computing, Coventry University

Prof. Sudip Misra .... Sudip Misra ....

Member, Thesis Committee
Professor, Department of Computer Science, IIT Kharagpur

To my beloved Maa and Babuji...

इन्द्रियाणि पराण्याहुरिन्द्रियेभ्य: परं मन: |
मनसस्तु परा बुद्धिर्यो बुद्धे: परतस्तु स: || 42 ||

Translation:

The senses are superior to the body, and superior to the senses is the mind. Beyond the mind is the intellect, and even beyond the intellect is the individual consciousness.

- Bhagvat Gita

Acknowledgments

Sometimes it's the journey that teaches you a lot about your destination.

– Aubrey Drake Graham, Canadian Artist

In 2017, for a conference paper presentation, I visited the beautiful island of Maui in Hawaii, a state in the United States of America. During that visit, I, along with my girlfriend, went on a 65-mile long road trip on a two-lane highway, connecting Kahului with the town of Hana in east Maui. It is a highly recommended activity if you are travelling to the island of Maui. On the way to Hana, we encountered enchanting trails leading to waterfall pools, delightful beaches, scenic mountains, among many other breathtaking views. Once you reach Hana, you would soon realise that the roadway to Hana is not about the destination, it is about the journey. Now when I look back, I find this road-trip quite analogous to my PhD journey. The PhD was never about graduating or getting a degree, it involved learning from the ups and downs of the research, develop a problem-solving attitude, and network with the researchers across the world.

When we planned the trip to Hana, to be honest, I had no prior experience of driving a car in the USA. We drive on the left side of the road in India; however, in the USA, people drive on the right side of the way. Of course, I was low on confidence since I had a little practical knowledge about the traffic rules in the USA. All I had was the intent and faith in my driving skills. Subsequently, we decided to go on this road-trip, and I decided to stick to the basics of driving. Likewise, when I was about to start the terminal degree of my education (the PhD), I wasn't sure if I could pull this off. I wasn't sure if I am ready for this long-term commitment. I had to consider both personal and financial constraints. At that time also, all I had was the intent. I wanted to work on challenging problems, learn new stuff, and do something innovative. And I believe the reason I had this kind of attitude lied somewhere in my childhood and upbringing.

I had a privileged childhood, all thanks to my mother (Mrs Veena Jain) and my father (Mr S. K. Jain). Unlike me, my father had a rough and tough childhood because my grandfather (Mr R. D. Jain) passed away at an early age. While he was extremely talented, at many times, he had to compromise with his dreams because neither he had the resources nor the situation was favourable. He worked hard, focused on family, and made sure that his kids had to never compromise with their dreams. For us (me and my siblings), he ensured

that we understand the significance of education in our life and never underestimate our ability to achieve something we want. I still remember the way he used to wake up early in the morning and take our classes before leaving for his office. He was so efficient with his work, he would never waste a minute in his office. He would timely wrap up his work, come back in the evening, and spend time with us. For any doubt, in any subject, we used to approach him, rather than the teachers. He would explain the concept so well that there is no chance that we would forget it, anytime soon in the future. At times, when he had no upfront answers, he would spend his time reading the text, learn the concept, and explain it to us. I still wonder, if not a banker, how well of a teacher he would have been! Intentionally or unintentionally, he made sure that we always have the right mindset to tackle a problem. Now, whether you take my road trip to Hana, or beginning of my PhD, you will definitely find a glimpse of this attitude in my behaviour.

Now, as planned, we began early in the morning for this road trip. On our way, we had to pass narrow bridges, take sharp turns, and ensure a safe drive. In PhD journey as well, we face challenges almost every day. Whether due to an experiment failure, or a paper rejection, we meet our own set of highs and lows. In those times, a dose of persistence and self-motivation is often required. For me, the dose used to come from my mother. We all know that a mother's love is pure and unconditional, and I was fortunate enough to experience that. Whether I had to leave early at six in the morning or had to come late in the night, she will make sure I get a healthy home-cooked meal every time. Whether she is ill or feeling low, I have never seen her taking a day off. I guess this kind of commitment can only exist when you love what you are doing. When I used to see her commitment level, a paper rejection, or a code failure, used to look quite insignificant to me. In my PhD, this thought always boosted up my confidence at the low times, especially during the mid-PhD crisis, and eventually engraved a fight-back spirit in me. I could literally write a book on the contribution of my parents in building my character. Not only they shaped my personality, but they also supported me when I needed them the most. I know, whatever I do, I will always remain indebted to my *Maa*, *Bauji*, *Dadi*, and *Dadu* for their unconditional support, love, and character-building upbringing.

Now, at this point, I had the character and the attitude from my upbringing. Still, I needed a roadmap and the guidance to navigate the route, to both Hana and the PhD graduation. In the road trip to Hana, we needed navigation to ensure that we are on the right track and getting maximum out of the road trip. Likewise, in my life, I needed the knowledge, a sense of discriminating the right from the wrong, and awareness that I should always do what I love. I received that guidance from my teachers and the gurus. If I recall correctly, I was never interested in science until the ninth standard. Since I was so much inspired by the personality of my father, I always intended to study commerce, until I met Mrs Paramjeet Kaur.

In the ninth standard, my father received a transfer to Karnal (a city in Haryana), and now, I had to continue my ninth standard in Kendriya Vidyalaya, Karnal. In this new school, Mrs Paramjeet Kaur was our science teacher. I was always amazed by her experiment-driven approach to teaching science. In no time, science became my favourite subject. I, who was a below-average student in science, won a science quiz first time in the ninth standard. Along with Mrs Kaur, I would also like to acknowledge the immense support from my then school principal. Unfortunately, I don't remember how to spell her name correctly. Though I was an average student, she always believed that I can top the batch and score well in my board exams. Her constant encouragement significantly boosted my confidence, and I actually evolved as a topper in my school with 91.4% in the tenth grade. By the time I completed my tenth grade, commerce was no more my favourite stream, and I chose to pursue science in my high school.

Fast forward to the first semester of my masters, I met Dr Amarjeet Singh, an Assistant Professor at IIIT-Delhi, and my PhD advisor. He is one person whom I admire and respect a lot for his honest and hardworking personality. I distinctly recall our first meeting in which I explicitly showed my interest in working with him. Though I couldn't understand a single word he said (we discussed home automation), I knew I could learn a lot from this aggressive and passionate researcher. Now, when I look back, I have no regret in saying that I took the right decision. Even though he decided to go for his own startup, he never compromised on the time for his PhD students. I still wonder how he pulled this off. In the initial days, he had us (the PhD students), his startup, and a year-old baby, but somehow, even in this crazy busy schedule, he always had the time to guide us. In return, all he needed was some support from us. For instance, often, he would want us to work from his office, or other times he might want us to submit the work two weeks before the deadline. I guess that was quite reasonable of him.

Throughout my PhD, Dr Singh always stood by me, motivated me to aim higher, and more importantly, let me shape my PhD. In addition to this, he helped me in building my network. Through him only, I met Dr Vikas Chandan, who eventually became my thesis co-advisor. Vikas graduated in Mechanical Engineering from the University of Illinois at Urbana-Champaign, and he was working at IBM research at the time I met him. With Vikas, I used to get a totally different perspective because he was from the mechanical engineering background, and we were from computer science. Not just his contributions made the thesis multidisciplinary but also gave our work the right direction. With Vikas's mechanical and industrial experience, and Amarjeet's academic and entrepreneurship exposure, I believe I was fortunate to do a PhD in a diverse environment. Not only my thesis advisors but even my collaborators and committee members also played a crucial role in this long marathon of PhD.

Our collaborators and thesis evaluation committees together stood up as an essential pillar of my PhD thesis. In 2016, we first collaborated with Prof Srinivasan Keshav and Prof Catherine Rosenberg at the University of Waterloo, Canada. At that time, I worked with Dr Rachel K Kalaimani to study hybrid control strategies for heating, ventilation, and air-conditioning units (HVACs). In 2018, I visited Responsive Environment Group at MIT Media Lab to study the role of room temperature in a mediated atmosphere. In MIT, I mainly worked with Dr Nan Zhao, under the guidance of Prof Joseph A Paradiso. While MIT and the University of Waterloo offered me a different outlook of my research problem; Zenatix, IBM Research Lab, and PNNL provided me much required industrial exposure. In PNNL, I mainly worked with Dr Osman Ahmed on analysing tradeoff in cost and accuracy of the sensors for fault detection and diagnosis.

In addition to our collaborators, my thesis evaluation committees consistently guided me in the right direction and provided constructive feedback. Our core annual committee comprised of Prof Krithi Ramamritham from IIT Bombay, Prof Yuvraj Agarwal from Carnegie Mellon University, and Prof Richa Singh from IIIT-Delhi. Though the committee was a hard nut to crack, the constructive feedback from such an expert committee eventually ensured quality work from me. Likewise, the committee members of my qualifying exam, Dr Vijay Arya from IBM Research, and Prof Pushpendra Singh and Prof Pravesh Biyani from IIIT-Delhi, narrowed down my research problem. Because of the high expectations of my committee members, I always submitted, and eventually presented my work at reputed venues. In conferences, I had a habit of discussing my work in the PhD forums, which I found quite beneficial in formulating my research problem. In PerCom 2016, I won the Best PhD Forum Presentation Award. In the PhD forum, I received crucial feedback from the committee of Prof Archan Misra from SMU and Prof Nirmalya Roy from UMBC. In UbiComp 2017, a doctoral colloquium committee including Prof Anind K Dey from the University of Washington, David Dearman from Google Research, and Monica Tentori from CICESE suggested me several possible extensions of my work.

In the end, after acknowledging the feedback from committee members over the years, I eventually concluded my research work and submitted the final draft of my dissertation to the thesis evaluation committee. The thesis committee comprised of Prof Alex Rogers from the University of Oxford, Prof Elena Gaura from Coventry University, and Prof Sudip Misra from IIT Kharagpur. Again, a panel of distinguished and well-known researchers. To be honest, Prof Alex Rogers is my role model based on the kind and quality of work he does. Given his busy schedule, I was not even sure that he would accept our invitation to join my thesis committee. It was definitely an honour for me to graduate with my academic role model being my evaluator. The comments from this distinguished panel polished the final draft and compelled me to ponder a long term impact of my research. Overall, I believe this thesis is a combined effort by me, my advisors, our collaborators, and the committee members. Without the right guidance, I am quite sure that the timely completion of neither the journey to Hana nor the marathon of PhD was practical.

Moving forward, I hope you remember that I was not travelling alone; my girlfriend was accompanying me. Not only she assisted me in driving, we together cherished each moment, relished all the scenic beauties, and created memories. Dalai Lama, a spiritual guru, once said, ”a human being is a social animal”. I firmly believe in this philosophy. I think that whatever we do, wherever we are in our life, people around us make a significant impact on our thought process, lifestyle, and the way of living. In the journey of PhD graduation, I was fortunate to be accompanied by supportive siblings, encouraging and fun friends, and a delightful partner. I would begin with my sister Palak Jain and brother Kushal Jain, with whom I shared a fair share of smiles and cries. Though I was the big brother; how to smile in hard times, and how to tackle tough people, is something I have learnt from them. When tensed, we would sit for hours, or go for a stroll to chit-chat and listen to each other. For me, often, that acted as a therapy in my hard times.

Good company in a journey often makes the way seem shorter. Along with my siblings, in my journey, I was fortunate to have Mridula with me. In every sense, she is my better half, as we do compliment each other. If I am an introvert, she is an extrovert with a strong opinion about her life choices. If I prefer coffee, she is a tea lover. I am a big Bollywood fan, but she admires Hollywood. But as they say, opposite attracts! I highly appreciate the fact that we are so different because we end up experiencing the best of both the worlds. In tough times, when I am anxious, she will hold my hand and prompt me to stay stronger. If I am sad, she will find ways to cheer me up. In short, she has been my personal therapist, where an appointment was not necessary. Whether it was my PhD or roadway to Hana, if she wouldn't have supported me, I might have crashed at multiple occasions.

Last but definitely not least, I would next like to acknowledge the crucial role of my friends - the family I chose for myself. Since childhood, I feel I always had the best set of people on my side. My father had a transferable job, and we used to move one city to another, regularly. We didn't have social media in the 90s, and therefore, I lost contact with most of my childhood friends. Though I possess a very fade memory of my childhood, I could never forget Sakshi from KV Faridabad. She was a delightful girl who always protected me in the bullish environment (yes, I did get bullied!). We used to go to school together, play together, and even fight together. Good memories! I guess the next time when I shared such a strong bond of friendship was when I moved to NIT Patna for my bachelor's degree - the golden era of my life.

In my undergraduate program, I met some of the finest but lunatic people on the earth. On the first day itself, I met Sahil, Pradeep, OP, Abhishek, Mahipal, and Ramswaroop, who end up being my flatmates. We all were from different streams of engineering, but we shared a strong bond among ourselves. Though I won't deny the fact that we also had differences (like any family), no one could undermine the strength of our union. Our extended evening discussions, cooking sessions, train travels, and movie nights are some of the few memories that could never fade away from my mind. The terrific batchmates like Avinash, Niraj, Ashish, Bhuvnesh, Neetesh, and Abhishek, to name a few, were the icing on the cake. Often, during the PhD, whenever I needed financial support, they won't think twice before offering the assistance. Since my undergrad, these folks have been my extended family away from home. We all are still connected and meet regularly. Even today, in our busy schedule, we often find time to sit for a drink and talk for hours. I hope this friendship goes forever.

Soon after my B.Tech, I moved back to Delhi to pursue my master's degree at IIIT-Delhi. Since I was moving from hostels to my home, I knew things will change. I knew that I will lose the freedom of a careless bachelor, the vibrancy of my hostel life, and the liberty of breaking the rules and regulations. And that was hard! I remember the commute on my first day when I travelled from my home to my college with a bag, a tiffin box, and a water bottle in my hands. Most of the people in that subway looked tired and were sleeping, and I was wondering why are they so exhausted in the morning. Soon, I got my answers. Life in Delhi is hectic and fast-paced. You must run fast to be in this race else you will be left behind. I was commuting every day, trying to wrap up my work in time, catching up the pace of Delhi, but fell down sharply in the first semester itself. I guess that was one of the toughest times for me because all my friends were spread across India, and I was in a different environment with people in a totally different mindset. I distinctly remember that I took Mobile Computing in my first semester, and couldn't find a single person to make a group for the project. Though I chose a specialisation in Mobile Computing, I scored a five-pointer in that subject. I knew it won't be easy, and I will have to get out of this mess.

In the winters break (after the first semester), I decided to stay back and start a research project with Dr Amarjeet Singh. At that time, I bonded with Parikshit and Anil, my M.Tech colleagues, and Nipun, Manoj, Samy, Sahil, and Manaswi, who were part of the research group of the Amarjeet. That winter was a big relief! I was steadily becoming a part of the IIIT-Delhi ecosystem with the support of my new colleagues. Soon I also met Alvika, Garvita, Deepika, and Dheryta during my Masters, and now I had my own group at the IIIT-Delhi. I started enjoying my time at IIIT-Delhi. The gup-shups at the chai-time, thought-provoking whiteboard discussions, and extended lunch hours provoked me to think beyond Masters. At that time only, I chose to pursue a PhD at IIIT-Delhi.

With time, our group started expanding, and I was learning something new from everyone in this group. Now, Tanya, Anupriya, and Lokender were also part of my close friend circle. While Anil, Alvika, and Lokender taught me dedication towards the work, Dheryta taught me how to get up and fight back in adverse conditions. If I ever need motivation and pep talk, I would always turn towards Parikshit, Garvita, and Deepika for a piece of advice. Nipun, Manoj, and Samy were literally my gurus. Since Dr Amarjeet left for his startup, Nipun, Manoj, and Samy taught me the basics of research and guided me on multiple instances. The youngest in the group, Tanya and Anupriya became the most delightful

and charming members of the group. I was fortunate to be surrounded by people with such high intellect and hardworking nature.

The list of people who supported me in this whole journey is so long that it is hard to mention them all. I was indeed blessed to have colleagues like, Kuldeep, Samarth, Venki, Megha, Charul, Akanksha, Dhananjay, Sneihil, Sonia, Paridhi, Haroon, Vandana, Siddhartha, and Inderpal, to name a few. Not just friends and colleagues, I haven't seen an efficient administrative staff like IIIT-Delhi. I would like to acknowledge Mrs Priti Patel, Ms Sheetu Patel, Mr Ashutosh Brahma, Mr Vivek (left IIIT-Delhi), and Mr Kuldeep (left IIIT-Delhi) from the administrative department, Mr Prosenjeet and Ms Paridhi from the finance department, and Mr Ajay and Mr Arun Verma from the logistics department. At several occasions, the administrative staff at IIIT-Delhi went beyond the conventional way, so that we could focus and concentrate on our research. Only because of the support from friends, family, and the staff members, I could complete this long journey.

At this point, we must not forget that none of this was possible without money because research is expensive! While TCS Research and Innovation fully funded my PhD, IIIT-Delhi and IUSSTF (through BHAVAN fellowship) generously supported my visit to MIT and PNNL, respectively. In addition to this, I would also like to acknowledge the financial support from Microsoft Research, ACM India-IARCS, SIGCHI, and Xerox Research towards my international conference travels. Without the generous financial support from these agencies at different times in my PhD, none of this was feasible.

Instead, I should say that without the right attitude, proper guidance, and generous support, I would not have completed this long marathon of PhD. Now that it is over, there is only one question in mind, ”why it is over?”. But, as they say, all good things must come to an end. The only thing that stays forever is the unforgettable memories which are well framed in my mind. I hope the connections I built over these years will remain forever. I hope the knowledge I acquired in this last decade of learning will equip me to serve back for the betterment of the society. Definitely, it is not an end, this is just a new beginning. A fresh start to become an independent researcher and carry forward the legacy of novelty and innovation.

Milan Jain

Contents

List of Figures

List of Tables

Chapter 1

Introduction

If India's usage of air conditioning eventually matches the U.S. level, its energy demand for cooling would be 14 times that of the U.S.

– Michael Sivak, American Scientist Magazine (2013)

1.1 Space Cooling

Once a luxury, space cooling is now becoming an indispensable part of our lifestyle. Whether it is a home, a laboratory, a data centre, a hospital, or any other essential building, space cooling is vital in day-to-day activities of human life. Currently, 90% of people in industrialized countries like United States of America and Japan are using 500 million heating, ventilation, and air-conditioners (HVACs), consuming 5.4 TWh of energy every year, across both residential and commercial buildings. With changing climatic conditions, growing disposable incomes, we are expecting the worldwide demand for space cooling to grow from 12 TWh to 6200 TWh, by 2050 [70]. Interestingly, 70% of this demand is likely to come from 2.8 billion people living in the hottest emerging economies of the world, where currently, only 8% of people possess air-conditioners (Figure 1-1). In contrast to developed economies, people in these developing economies prefer split or room-level air-conditioning[1], especially in residential apartments and small-scale commercial buildings.

Space cooling contributes significantly to the total electricity bill and also accounts for a considerable proportion of the peak demand. Across the world, space cooling accounts

[1] In this paper, here onward, we refer split or room-level air-conditioning interchangeably to as AC or air-conditioner.

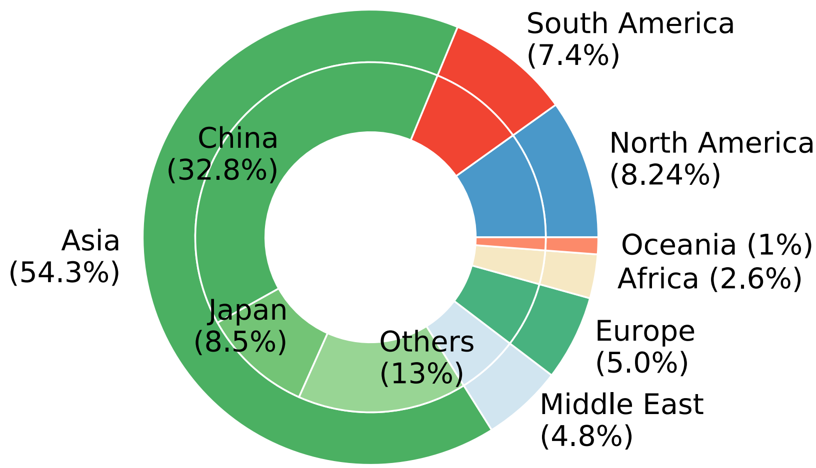


Figure 1-1: Currently, two-third ACs are found in China, Japan, and North America. In future, emerging economies with hot climate (such as India) are expected to lead the sales chart [70].

for a large share of peak demand, which puts significant stress on our power systems (Figure 1-2). In hot and humid countries, the daily peak typically occurs in the afternoon, when the outside temperature is very high. In Saudi Arabia, air conditioning accounts for a staggering 51% of total electricity demand, with demand in summer twice as high as during the winters [36]. Despite such a noteworthy contribution, occupants often possess little understanding of most suitable AC settings, the trade-off between their comfort and electricity bill, and the benefits of regular maintenance of their appliance [80]. Considering little awareness among people, lack of energy-efficient systems in homes, and estimated growth of space cooling demand, handling such a massive energy demand will soon be a severe concern globally [133].

1.1.1 Impact on Climate

Besides stressing up our grid and power systems, ACs also impact our climate by emitting greenhouse gases, directly and indirectly. Greenhouse gases (such as CO_2, water vapor, methane, among others) absorb infrared radiations which gradually increases the temperature of earth's surface and atmosphere, eventually contributing to global warming. In 2016, space cooling alone emitted 1130 million tonnes of CO_2, as computed based on the carbon footprint of electricity generation, at times of AC usage [70]. Unfortunately, considering the future demand, the carbon footprint from ACs is only going to increase, unless we explore

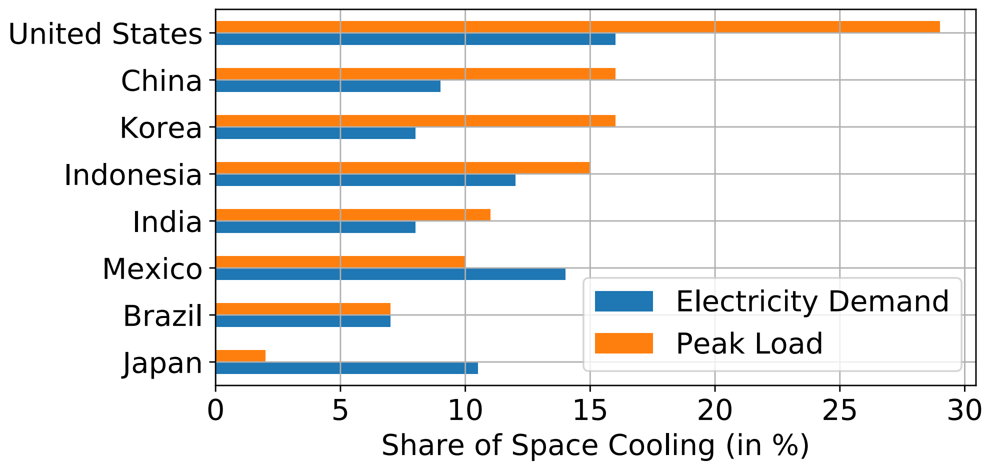


Figure 1-2: Across all the regions, space cooling contributes significantly towards both, electricity demand and the peak load [70].

energy-efficient ways to optimize both, existing and futuristic air-conditioners. Given that climate change is a reality now, it is even more critical for us to reduce both direct and indirect emissions of greenhouse gases from the ACs, substantially.

1.1.2 Government Policies

Understanding the significant impact of ACs on energy demand and environment, governments across the world imposed several policies, especially in the last few decades. For instance, in 1987, 197 UN member states signed the Montreal Protocol to discontinue the use of chlorofluorocarbons (CFCs) and hydrochlorofluorocarbons (HCFCs) as the refrigerants [70]. Since then, the agreement went under eight revisions with latest on October 2016 in Kigali, Rwanda. In Kigali amendment, the same 197 parties agreed to progressively phasedown the production and use of hydrofluorocarbons (HFCs) around the world.

Besides regulating refrigerants, different countries introduced multiple energy-efficiency standards and labeling programs to promote efficient air-conditioners. With the support from the governments of United States, India, China, Canada, and Saudi Arabia, the Clean Energy Ministerial, launched an Advanced Cooling Challenge in 2016 to encourage the development and deployment of super-efficient, smart, climate-friendly and affordable cooling technologies to improve average AC efficiency by 30% by 2030 [70]. Independently, the Government of India itself proposed replacement of inefficient ACs with energy-efficient

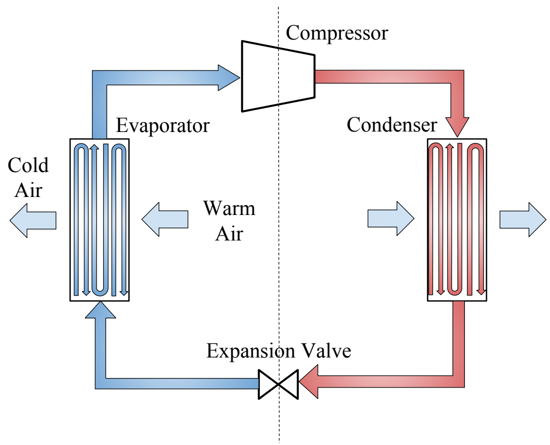


Figure 1-3: Vapor Compression Cycle: Refrigerant flows through different components of the AC to cool the space.

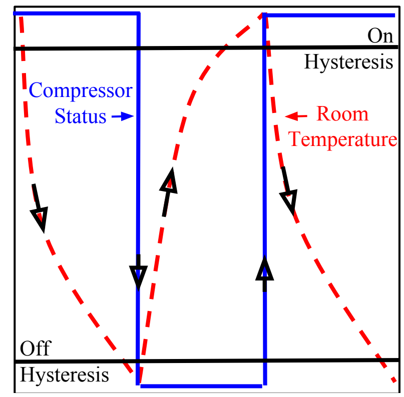


Figure 1-4: Room temperature fluctuates between *on* and *off* hysteresis to stay close to the setpoint temperature.

ACs at a subsidized cost [158]. Similarly, the French Government mandated that during extreme periods of heat, at least one room on each floor in every retirement home must be maintained at a temperature less than 25°C [70].

While there have been several efforts by different Governments across the world, both independently and together, there is a scope of improvement from technology's perspective. Occupants habitually keep their AC settings unchanged (usually at maximum), ignore current weather conditions, shy away from regular maintenance, and keep running their AC at reduced efficiency. In several instances, we noticed that tenants run their AC at maximum and use blankets instead of updating the AC settings. Often, due to irregular maintenance, tenants face problems such as ice formation, air blockage, and refrigerant leakage. With reduced efficiency, AC fails to provide the desired comfort and wastes significant energy. The widespread presence, together with the scope of improvement, make room-level air-conditioning an attractive target for energy conservation, especially in the residential and small-scale commercial buildings of emerging economies with a hot climate.

1.2 Technical Background

Typically, an AC works on the principle of the vapour compression cycle (Figure 1-3). In the beginning, fan towards the evaporator coil draws air from the room. When air passes over the evaporator coils, the low-pressure low-temperature liquid refrigerant[2] (within the coils) absorbs heat from the air, and the fan circulates cold air back into the room. Due to heat absorption, liquid refrigerant changes to high-temperature refrigerant vapours. High-temperature low-pressure refrigerant vapours enter into the compressor, where compressor increases the pressure of the refrigerant vapours for the condensation. In condensation, the aim is to eject heat from the refrigerant to the outside environment. To do so, fan towards the condenser coil extracts hot air from the outdoor which passes over the condenser coils having high-pressure and high-temperature refrigerant vapours. Due to condensation, high-temperature refrigerant vapours become low-temperature liquid refrigerant, which then enters the expansion valve. The expansion valve further reduces the pressure of low-temperature refrigerant liquid for the evaporator coils.

In this dissertation, we refer a single vapour compression cycle as *Compressor Cycle*. In any air-conditioner, compressor consumes the maximum amount of electrical energy. The AC energy consumption mainly depends on the number of compressor cycles and the duration of each compressor cycle. Interestingly, during an AC usage[3], the number of compressor cycles and the length of each cycle also govern the user comfort. Though both these parameters further rely on several contextual settings such as the building insulation, the cooling load, the age of the appliance, among many others; in a given context, the AC energy consumption is mainly a function of AC settings, as provided by the occupants on the thermostat. Since occupants specify their preference over a thermostat, the role of a thermostat during an AC usage is quite crucial.

1.2.1 Role of a Thermostat

Within an AC, the role of a thermostat is to sense the room temperature and perform specific actions to keep the room temperature close to the setpoint temperature. While actions depend on the type of controller, ACs typically have a bang-bang controller, in which, the thermostat uses *on* and *off* hysteresis as lower and upper bounds for the room temperature, respectively (as shown in Figure 1-4). When occupants turn on the AC at a set temperature, the compressor turns on, and AC starts putting cold air into the room. Due to the flow of cold air, the room temperature drops up to a certain level - *off* hysteresis,

[2] Refrigerant is a special chemical compound that switches it state from liquid to vapor, vapor to liquid back and forth for the heat transfer.

[3] *AC usage* is the time user switches on the AC unit, till she switches it off.

at which compressor turns off and allows room temperature to increase up to *on* hysteresis - the upper bound. When room temperature goes beyond the *on* hysteresis, the compressor turns on again, and AC starts cooling the space. By cycling the room temperature between *on* or *off* hysteresis, the thermostat tries to keep the average room temperature close to the setpoint temperature and maintain user comfort. On the other hand, the thermostat also decides the compressor state (*on*/*off*). Thus, a thermostat governs both user comfort and AC energy consumption; and therefore, it is crucial for optimizing AC operations.

Comprehending the noteworthiness of thermostats, lately, several studies have been focusing on developing smart and energy-efficient thermostats for the ACs. Programmable thermostats were one of the initial attempts to optimize air-conditioning across residential buildings. Beyond conventional thermostats, they offered an additional feature to set an operating schedule. However, for a regular home, where everybody follows a distinct routine, programmable thermostats proved to be an inefficient approach [94, 132]. Next in line were the smart thermostats [107]. Unlike programmable thermostats, smart thermostats learn the daily schedule of the tenants to change the thermostat temperature, automatically. Google Nest [118] is one such realization of smart thermostats for residential spaces. Google Nest offers numerous modes of operations, such as sleep mode, energy-saving mode, among many others. Akin, other commercial smart thermostats, such as Tado [152] and Sensibo [148], monitor occupants' location. When the tenants are in proximity of their residence, the thermostat will switch on the AC to pre-cool the space, and switch off the AC, as soon as tenants leave the place. Consequently, thermostats today are smart, energy-efficient, plug-n-play, and even allow remote sensing of the AC through smartphones. Given the benefits, we are anticipating the thermostat market to grow by 400% shortly [95].

Not to forget, thermostats also gather data. Smart thermostats monitor physical parameters such as room temperature, tenants' daily activities, among others, and usually upload sensed information to the cloud. Unfortunately, existing techniques use a limited amount of information and mainly focus on occupants' schedule to optimize AC operations. We must understand that the occupants' schedule is dynamic and stochastic by nature. As shown in Figure 1-5, the percentage occupancy at any hour of the day can vary significantly, depending on the sequence of events occurred that day. Since it is hard to learn and predict the occupancy accurately, energy-saving features (such as Auto-Schedule and Auto-Away) of smart thermostats (such as Google Nest) have often disappointed several customers in

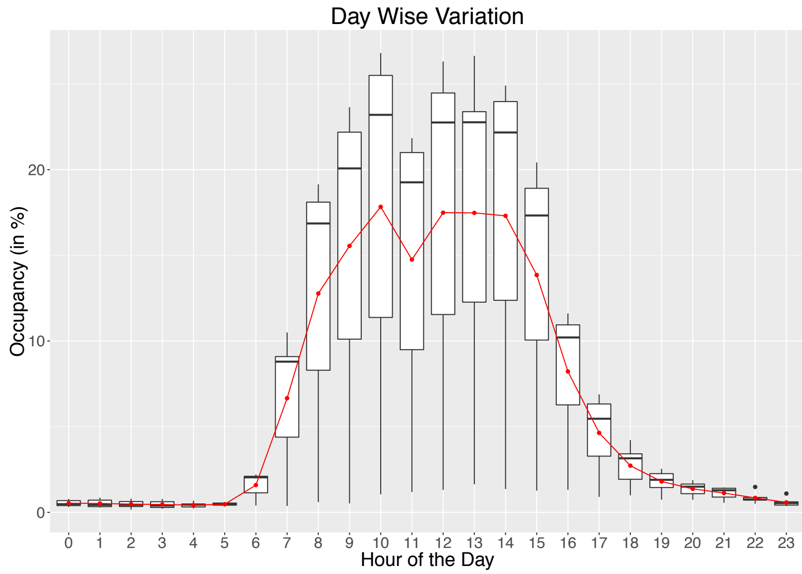


Figure 1-5: For a personal office, the plot indicates the variation in occupancy at different hour of the day. Specifically during peak hours, the variation could be quite significant.

the past. In a user study on smart thermostats, Yang et al. [166] noticed multiple instances where Google Nest learned the incorrect schedule, made erroneous assumptions about tenants' behaviour and took wrong actions. In such cases, users would typically unlearn the thermostat and try to follow an ideal schedule for the thermostat.

In addition to the learning issues, smart thermostats are also unaware of the working environment. Typically, a smart thermostat cannot compute the amount of thermal leakage happening through the room, the impact of outside weather conditions on the room temperature, and present-day efficiency of the AC. Since a smart thermostat only cares about the tenants' daily schedule, spatial variation across the buildings does not matter to them. Moreover, since thermostats do not keep track of the atmosphere, they hardly know anything about the temporal changes happening in the place. A thermostat would never know if the efficiency of an AC is depreciating with time, or thermal leakage is increasing due to poor thermal insulation of the walls. As of now, it is hard for thermostats to speculate why tenants altered the thermostat temperature on any particular day. In other words, a smart thermostat cannot link the cause with the effect.

In this dissertation, we extend *Smart Thermostats* to *Data-Driven Thermostats* a framework that primarily depends on temperature data to learn the lumped parameters of a theoretical model. By utilizing the temperature information, in this thesis, we sought to make occupancy driven thermostats, space-centric thermostats. Since we tune the model parameters using the sensor data collected from the room, the proposed framework is considerate of spatial variations across the buildings. Subsequently, *Data-Driven Thermostats* are conscious of their working environment. With time, as the framework collects more and more data, it keeps updating the model parameters to accommodate temporal changes in the building's thermal behaviour. Tuned thermal model imitates the thermal behaviour of the space and serves the application layer. Our analysis on real-world indicates that *Data-Driven Thermostat* is a genuinely low-cost, scalable, and ubiquitous way to make ACs energy-efficient, maximize user comfort, and make ACs reliable.

1.3 Research Statement

In this dissertation, we aim to explore, "How can we design a framework that can work on top of a smart thermostat and benefit from the existing infrastructure to provide actionable feedback, maintain a comfortable environment for the users, and ensure reliable functioning of the appliance?". To answer the big question, we had to answer the following questions -

1. We can save 5%-15% energy through actionable feedback [33]. How accurately can smart thermostats provide actionable energy feedback to the tenants, while using the data from its inbuilt sensors? The framework must not require additional instrumentation of the space, and it should work with minimal user attention.

2. Control approaches applying automated set-point variation can save up to 26% energy when compared with scheduled set temperature [115]. Can smart thermostats automatically change the setpoint temperature while allowing users to balance their comfort and savings.? We don't want the user to worry about the *right* set temperature. With no additional instrumentation, the thermostat should adapt to temporal changes in the environment and localize comfort for the tenants.

3. The effectiveness of predictive control strategies (for ACs) principally relies on the prediction accuracy of contextual parameters, such as weather and occupancy. Since occupancy is critical for mostly all predictive control strategies, can we quantify the impact of occupancy prediction errors on HVAC operations? Furthermore, how can we minimize the influence of prediction errors and stabilize the control?

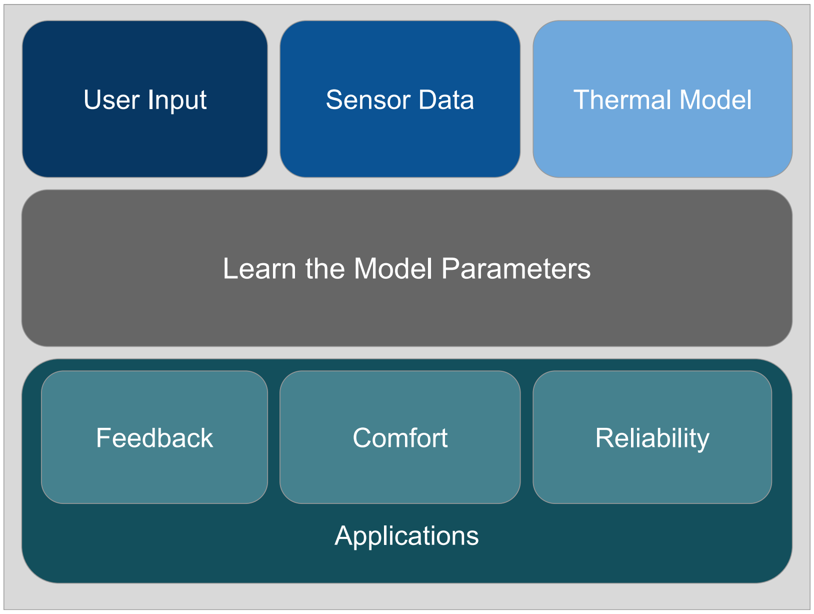


Figure 1-6: Architecture of Data-Driven Thermostats: Top layer is the input layer; Middle layer is the learning stage where framework tunes the model parameters through sensory data; Bottom layer contains applications built on top of tuned thermal model.

4. Often, refrigerant leaks through the coils (or valves) of an AC and slowly diminishes its cooling capacity while allowing it to be functional. Such leaks risk occupants' health, waste significant energy, and impact the room temperature. How accurately can smart thermostats timely report such leaks by only using the readily available data from its inbuilt sensors?

In this thesis, we systematically designed and studied Data-Driven Thermostats by exploring answers to all the research questions, as mentioned above.

1.4 Thesis Contributions and Outline

Our Data-Driven Thermostat comprises of three layers - (1) input layer, (2) learning layer, and (3) the application layer (as depicted in Figure 1-6). The learning layer takes sensor data (from the thermostat) and user preference (such as setpoint temperature) as the input, and regularly tunes the model (provided as the input) parameters to accommodate changes in weather conditions, user activities, and various other dynamics occurring in the room. Principally, the thermal behaviour of any space depends on the non-intuitive thermodynamics, and it is hard to monitor each factor affecting the room temperature. Therefore, in Data-Driven Thermostats, we use *Grey Box Modelling*, in which we tune the lumped parameters of a theoretical thermal model using the sensor data collected from the physical world [108, 137]. As tuning involves the actual data from a physical space, the adjusted parameters typically present an approximate thermal behaviour of the room. The system utilizes the tuned thermal model to estimate room temperature, AC energy consumption, among several other parameters for the application layer. Eventually, application layer takes the estimates to then provide - (1) actionable energy feedback, (2) adaptive and personalized comfort, and (3) corrective feedback for the appliance. Next, we discuss each application in detail.

In a study, Darby et al. [33] noticed that detailed feedback at the appliance level could result in 5%-15% savings in energy consumption. In Chapter 2, we discuss PACMAN - an energy-feedback system to non-intrusively (by only using the temperature information) predict and estimate the AC energy consumption, prior to usage and post-usage, respectively. PACMAN uses room temperature (from the thermostat) and weather data (from cloud-based weather services) to learn the parameters of a thermal model. The learned thermal model together with estimated and predicted energy profile of the appliance is beneficial for applications such as - (1) fault detection and diagnosis, (2) energy-efficient control, (3) improving disaggregation accuracy for NILM algorithms [127] (when fused with smart meter data), and (4) identifying changes in the thermal behaviour of the room (e.g., detect door/window is open when AC is running). To evaluate PACMAN, we conducted an in-situ study across seven homes in Delhi (India). In our analysis, the proposed service achieved an average accuracy of 85.3% and 83.7% with the best accuracy of 97.0% for the estimation and 93.3% for the prediction of AC energy consumption, across all homes.

If followed, actionable feedback can save significant energy; however, the realization of these energy savings mainly depends on residents' attitude; thus can benefit from automation [29]. Typically, tenants set the AC thermostat temperature (also referred to as the setpoint temperature), depending on their personal preferences. For the same setpoint temperature (in a given atmospheric condition and context), while some people might feel comfortable, others might feel hot or cold. Thus, people choose different setpoint temperature, as per their choice. However, though the set temperature is a matter of choice, it directly influences the AC energy consumption. For a given atmospheric condition, the AC energy consumption would decrease as we increase the thermostat temperature. With limited knowledge about the comfort-energy trade-off, residents prefer to set a fixed thermostat temperature (usually lowest possible) in their daily routine. In Chapter 3, we propose a Comfort-Energy Trade-off (CET) knob on smart thermostats to allow users to balance their comfort and the savings without them worrying about the *right* set temperature. The CET knob will be present on the user's smartphone so that they can report their interest (along with the setpoint temperature), that whether they want to maximizing savings, or maximizing comfort, or balance both. Our analysis based on real data, collected from a controlled experiment (across two rooms for two weeks) and an in-situ deployment (across five rooms for three months), indicates that the proposed service can reduce residents' discomfort by 23% (CET selection for maximal comfort) and save 26% energy when we set CET for maximizing savings.

Though automated control could assist users in maximizing their comfort and savings, the effectiveness of any typical predictive control strategy (like above mentioned) depends on the prediction accuracy of model inputs, such as building occupancy and outside air temperature. As prediction accuracy deteriorates, performance - in terms of occupant comfort and building energy use - degrades and may get even worse than conventional techniques. While existing studies [122] studied the influence of errors in the weather forecast, to the best of our knowledge, there exist no studies analyzing the impact of prediction errors in occupancy on user comfort and total energy consumption. With the help of a simulation framework (ThermalSim[4]), in Chapter 4, we systematically study the impact of occupancy prediction errors on HVAC operations. Besides, we model and analyze the effect of a personal environmental control system (PEC) in the presence of prediction errors. A PEC could be a fan or a heater to provide personalized thermal comfort [22]. We find that PEC, when used with a predictive control strategy, can reduce both - the variability in energy consumption and occupants' discomfort, and make predictive controllers more reliable.

When we talk about reliability, another critical aspect is, timely notifying a fault to ensure efficient functioning of the appliance. Several reasons, such as ageing, irregular maintenance, and improper handling lead to frequent breakdowns of any appliance. Some failures are sudden (such as electrical failure), others happen gradually over time, also known as slow time-varying faults. While occupants can quickly identify sudden failures, sensing the early symptoms of slow time-varying faults requires detailed attention and technical

[4]We are still working on its stable release for the community use and we hope to release it soon

training. Refrigerant leakage is one such slow time-varying mechanical fault, quite common in air-conditioners and refrigerators [23]. The puncture hole often starts as a pinhole leak and becomes bigger when goes undetected. Due to the loss of refrigerant, the compressor works with reduced efficiency and takes more time than the usual to cool the room; thus, wasting significant energy. In addition to energy wastage and user discomfort, the leakage exposes tenants to the refrigerant, which is extremely dangerous for their health [7]. Early detection of such leaks can benefit in - (1) avoiding leakage of hazardous refrigerant in the open environment, (2) reducing the energy wastage, and (3) maintaining the desired temperature. While occupants usually fail to sense the early symptoms of such leaks, current techniques to report refrigerant leakage are often unscalable. In Chapter 5, we propose *Greina* - to continuously monitor the readily available ambient information from the thermostat and timely report such leaks. We evaluate our approach on 74 outlets of a retail enterprise. Our analysis indicates that the proposed framework can report leakage a week in advance when compared to manual reporting.

The data-driven thermostat is a low-cost and scalable extension of the smart thermostat that benefits from grey-box modelling to provide energy-efficient feedback, maintain personalized comfort, and ensure timely notification to the users about a fault. The proposed framework is efficient, works out-of-the-box, and compatible with any air-conditioner with an infrared (IR) remote. Given the current scenario, smart thermostats are anyways going to replace conventional thermostats in the near future. We believe extending smart thermostats to data-driven thermostats presents a genuinely ubiquitous way to ensure energy-efficient usage of the air-conditioning.

1.5 Thesis Publications

We now enlist the major publications that contributed to this thesis.

1.5.1 Actionable Energy Feedback

1. **Milan Jain**, Amarjeet Singh, and Vikas Chandan. "Non-intrusive estimation and prediction of residential ac energy consumption." IEEE International Conference on Pervasive Computing and Communications (2016).

1.5.2 Adaptive and Personalized Comfort

1. **Milan Jain**, Amarjeet Singh, and Vikas Chandan. "Portable+: A Ubiquitous And Smart Way Towards Comfortable Energy Savings." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (2017).

1.5.3 Reliability

1. **Milan Jain**, Rachel Kalaimani, Srinivasan Keshav, and Catherine Rosenberg. "Using Personal Environmental Comfort Systems to Mitigate the Impact of Occupancy Prediction Errors on HVAC Performance." Energy Informatics Journal (2018).

2. **Milan Jain**, Mridula Gupta, Amarjeet Singh, and Vikas Chandan. "Beyond Control: Enabling Smart Thermostats for Leakage Detection." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (2019).

Chapter 2

Energy Estimation Through Thermostat

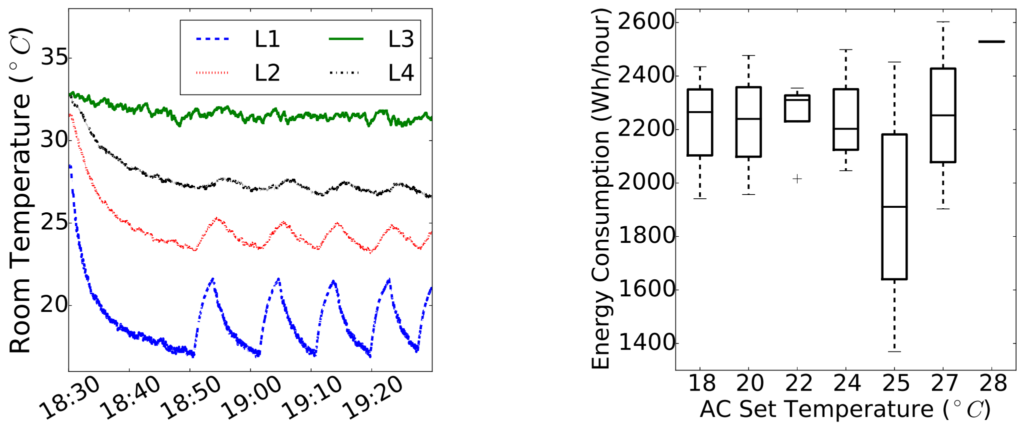
If you can't measure it, you can't improve it.

– Lord Kelvin, Lecture on "Electrical Units of Measurement" (1883)

2.1 Introduction

Occupants, who are unaware of their electrical consumption, add up to one-third to their electricity bills [119]. Among all the appliances, typically available in a home, air-conditioning (also referred to as AC in this thesis) contributes a significant proportion to the total electricity bill. However, even today, real-time and actionable feedback about AC settings, such as the impact of increasing the setpoint temperature by each degree on AC energy consumption, is broadly missing. A direct method to measur AC energy consumption is to use in-line plug monitors [20]. However, such plug monitors are typically unavailable in high current ratings, required for handling large transient current flow associated with AC switch-on events. The widespread presence, high energy consumption, and little understanding of optimal AC settings among people make ACs an attractive target for energy conservation. Though commercial smart thermostats, such as Nest [118], Tado [152], and Sensibo [148] can automatically change the thermostat temperature to optimize AC operations; they often rely on occupants' schedule and are unaware about their surroundings.

In this chapter, we discuss PACMAN, an end-to-end feedback system that leverages room temperature data for estimation and prediction of AC energy consumption. *Energy*



(a) Room Temperature

(b) Energy consumption per hour

Figure 2-1: [2-1a]Room temperature recorded at multiple locations (L1-In front of the AC, L2-Farther away from the AC, L3-In the attached bathroom with the door closed, L4-Other corner of the room). Here, the setpoint temperature is 20°C. [2-1b] The setpoint temperature is an insufficient metric to report the AC energy consumption.

estimation, post-usage, let tenants analyze the impact of chosen settings, to support better decision making eventually. On the other hand, *predictive feedback*, at the beginning of the usage, strives to influence occupants' behavior, when it matters the most. Unlike tenants' daily routine, room temperature is more stable since it chiefly relies on the physical properties of the space, such as outside weather conditions, cooling added by an AC, and heat exchange through other entities in that enclosed space. While it is hard for a thermostat to understand the rationale behind tenants' particular action, it is feasible for a thermostat to analyze temperature variations within an enclosed space. For instance, a thermostat might not understand why Alice left early on the weekend, but it can expect room temperature to be relatively higher when the outside temperature is high. Moreover, analyzing temperature variation enables a thermostat to comprehend an environment and adapt to temporal variations within the room, and spatial variations across the rooms.

However, using temperature as a proxy for AC energy consumption exposes multiple open challenges. One such hurdle is the position of the temperature sensor in the room. As shown in Figure 2-1a, the ambient temperature at any point in the room differs significantly from the temperature observed by the AC thermostat. Akin to that, the AC energy consumption should ideally increase by reducing the setpoint temperature; however, as shown in Figure 2-1b, we observed significant deviations from the expected behavior, mainly due to the impact of the thermal environment in the room and weather conditions. Since there are several other non-intuitive dynamics involved, the use of temperature sensor as a proxy for AC energy consumption is complicated. PACMAN, to address these concerns, records ambient room temperature from the room and weather information from a cloud-based weather service to develop a thermal model of the space. Since PACMAN learns the parameters of the thermal model from the data gathered through the room, we expect the thermal model to represent the space. When tenants switch on the AC, the PACMAN utilizes the thermal model to predict room temperature and project AC energy consumption at different setpoint temperatures. At the beginning of an AC usage, such predictions act as a motivation for the tenants to consider energy-efficient thermostat temperature.

Since PACMAN is purely a data-driven framework, it is practical to attach PACMAN with any smart thermostat and provide near real-time feedback to the occupants about the impact of their current settings. In such a scenario, the aim will be to potentially motivate a change in tenants' behavior to set up a slightly higher set temperature, whenever required. In this study, we evaluated PACMAN based on its accuracy in estimating and predicting the AC energy consumption. For the study, with an in-situ deployment, we collected temperature and power consumption data from seven apartments of a city, for a total duration of three months. Our analysis based on the in-situ deployment indicates that PACMAN achieved an average accuracy of 85% and 84% with the best accuracy of 97% and 93% for the estimation and prediction of AC energy consumption, respectively, across all homes. To the best of our knowledge, ours was the first work to estimate and predict the AC energy consumption non-intrusively, by only sensing the ambient temperature.

2.2 Related Work

Optimization of HVAC systems for energy conservation has been well studied in the past. Appliance-level feedback can be provided by direct monitoring (using plug loads) [161] or by disaggregating meter level data into appliance level consumption [63] using machine learning algorithms. However, using such direct or indirect monitoring systems in the context of AC usage, typically limits the feedback to actual energy consumed without getting into the details of estimation and prediction of energy and its interdependence on set temperature. Disaggregation algorithms [14, 58, 15, 89, 88], when combined with our proposed system, will only help in improving the overall accuracy of our system.

Several studies tried to make thermostats more intelligent to optimize AC energy consumption. Programmable thermostats were the initial efforts in this direction where occupants were able to give fixed schedules. However, Sachs et. al. [145] showed that households with simple thermostats saved more energy on average than programmable thermostats. Commercial reactive thermostats such as Nest [118] claimed energy savings based on its intelligent AC control that require learning a schedule before taking actions. But, one of the participants in a survey [166] reported Nest as *arrogant* due to its nature of forcing their decision on occupants based on its inaccurate learning. Similar attempts [107, 96] were made to actuate AC based on the occupancy and sleeping patterns. Scott et al. [146] proposed sensors based actuation while Rogers et al. [143] introduced a system to control heating based on calculated thermal comfort in real-time. Studies in Indian context [12, 13] showed that automated sensor based HVAC control could have false positives from sensors leading to energy wastage or false negatives leading to occupants' discomfort.

Various metrics such as Predicted Mean Vote (PMV) [45] or other similar metrics [51] were proposed in the personalized home heating domain to address occupants' comfort. However, recent study [65] showed that one type of sensor is insufficient and ineffective at capturing the larger thermal comfort picture of the room. Thus, the objective of PACMAN is to reasonably learn the thermal model of the room based on historical usages. Decisions on comfort level are left to the occupants while providing them with information, relating set temperature with energy consumption, using the learned model.

Rogers et al. [142] (closest to our work) proposed MyJuolo whereby a USB based temperature logger collects the room temperature data. Using this data, their framework learns a thermal model of the room, involving several parameters including leakage rate, heating rate, and thermal noise within the room. However, their process requires significant intervention from occupants and also a static thermal model that might give inaccurate predictions with varying operating and weather conditions. In contrast, PACMAN uses data from multiple usages, collected in real time, making it more effective for the prediction. Updating the thermal model in PACMAN, accounting for dynamic operating and weather conditions, require no effort from the occupant thus making it user-friendly as well.

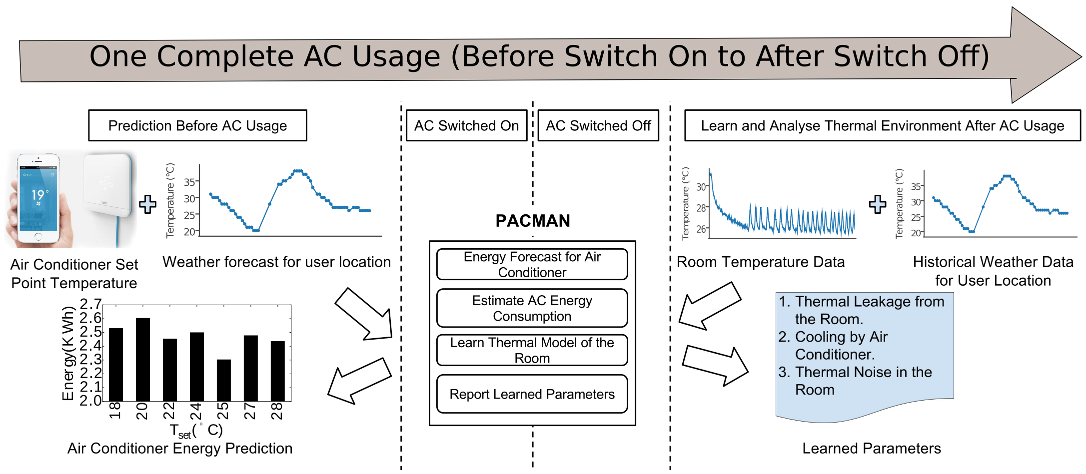


Figure 2-2: Use case of PACMAN: When the AC is switched ON, the set temperature and weather forecast is used with the learned model to predict the AC energy consumption at multiple set temperatures. When the AC is switched off, PACMAN estimates AC energy consumption during the usage. The collected room temperature data and corresponding weather conditions from recent n historical usages are used to update the thermal model.

2.3 Use Case

Figure 2-2 presents a typical usage scenario for PACMAN system. Alice has a smartphone with an application to control AC that also has access to ambient temperature from the sensor embedded in the phone. Various commercial devices [152, 148] provide such mobile applications for the users. These devices come with IR interface to communicate with the AC but lack the intelligence built into the PACMAN system. Such a combination is similar to intelligent control of central HVAC systems as performed by Google Nest [118], while also taking account of energy savings and Alice's comfort.

As Alice switches on the AC, PACMAN extracts the weather forecast, for the area surrounding her residence, from an online weather service. Weather forecast, together with the thermal model of her room (learned from the historical usage) is used to predict the AC energy consumption around its desired set temperature for a duration mentioned by her. Using the predictions, PACMAN estimates the cost Alice will pay if she runs AC at the specified set temperature, together with savings possible with increase in set temperature by each degree. Such information can assist her to decide the trade-off between cost (in terms of the electricity bill) and comfort (in terms of desired room temperature).

PACMAN as a backend service also provides room temperature readings together with weather conditions in real-time for easy accessibility of information to Alice and her better engagement. At the end of each AC usage, PACMAN estimates the AC energy consumption for the recent usage using the collected room temperature data. Estimated energy can help her understand the impact of her chosen settings and decide them appropriately over time. As PACMAN receives AC on and off instances, it can also improve the accuracy of NILM algorithms performing energy disaggregation for other appliances using meter level data. It can help her gain better insight into her electricity bills and take informative decisions (e.g. replacing an energy inefficient refrigerator).

In its final step, PACMAN learns the thermal model using the collected room temperature and weather conditions. The learned thermal model presents thermodynamic properties of the surrounding. It can be used to detect abnormal behavior of the occupants (e.g. open window) or the appliance (e.g. AC condenser working improperly).

The primary contribution of our work is to provide informative feedback on energy efficient usage of AC by developing a thermal model using easy-to-collect temperature data. Extension to other use-cases such as anomaly detection, automated control, and improved disaggregation is left for future work.

2.4 System Architecture

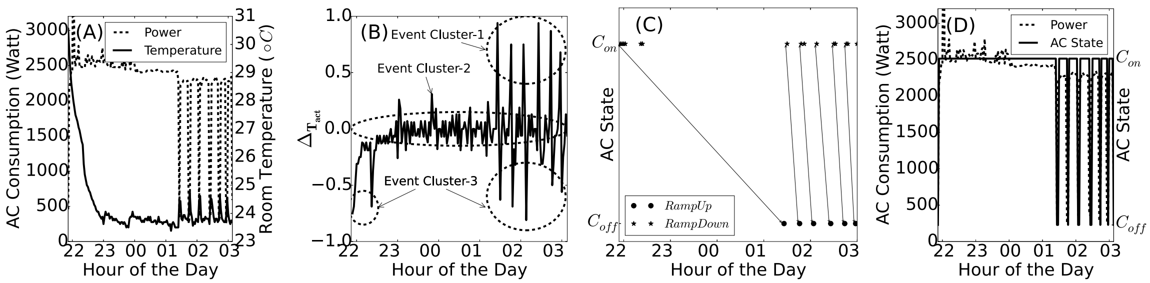
We designed PACMAN to model AC power consumption using room temperature and weather information to output actionable insights for the occupants at successive phases (from prediction to estimation) of AC usage. Keeping in mind the modularity and extensibility of the PACMAN, we partitioned it into three major components namely PACMAN-L, PACMAN-P, and PACMAN-Engine. It works as a backend service capable of supporting various commercial devices [152, 148] providing their own mobile front-ends.

PACMAN-L utilizes the set temperature, room temperature, and weather conditions, during an AC usage, to learn the following thermal parameters for the room:

1. *Cooling Rate:* The rate at which AC dissipates heat from the room.

2. *Leakage Rate:* The rate at which room gets heated because of its structure and thermal insulation. The structure includes various factors such as the number of windows, area of the room and material of the walls.

3. *Thermal Noise:* Variations in the room temperature due to unknown factors, such as



(a) Raw Usage (b) Clustering (c) Event Detection and Matching (d) Final Result

Figure 2-3: Illustration for estimation of AC state using the room temperature data going through various stages (Left to Right); (A) Sample usage from an apartment, (B) $\Delta_{\mathbf{T_{act}}}$ along with detected clusters for the usage, (C) Event detection and matching after filtering only *RampUp* (significant increase in temperature) amd *RampDown* (significant decrease in temperature) events, (D) Estimated AC state with actual power consumption of AC

open window, heat exchange with other appliances, present in the room.

PACMAN-L also estimates the AC energy consumption using the temperature data, collected at the end of each AC usage.

PACMAN-P is the prediction component providing energy consumption at various set temperatures, considering weather forecast for that duration. When the AC is switched on, PACMAN-P takes as input the set temperature, thermal parameters learned from historical usages and weather forecast to then predict the AC energy consumption prior to usage.

PACMAN-Engine is an administrative component that controls other elements of PACMAN and interacts with weather service to extract weather data for them.

2.5 System Implementation

PACMAN is implemented in Python while utilizing several libraries supporting machine learning tools (e.g. Pandas and scikit-learn). We now discuss the implementation details for each component of PACMAN, as outlined in Section 2.4.

2.5.1 PACMAN-L

As the occupant switches off the AC, PACMAN-L gathers the time annotated room and outside temperature data measured at a mismatched sampling rate for the duration ($d = t^{off} - t^{(0)}$, where $t^{(0)}, t^{off}$ are the timestamps when AC was switched on and off respectively)

Algorithm 1: Algorithm to estimate AC state

Input: $t^{(0)}$, t^{off}, $\mathbf{T_{act}}$
Output: $\mathbf{S_{est}}$

Data Initialization:
for $t \leftarrow t^{(1)}$ **to** t^{off} **do**
  $\Delta_{T_{act}}^{(t)} \leftarrow T_{act}^{(t)} - T_{act}^{(t-1)}$

`/* K-Mean clustering with 3 clusters */`
Clustering:
$[C_1, C_2, C_3] \leftarrow$ K-Mean($\Delta_{\mathbf{T_{act}}}$, 3)
$E_1 \leftarrow$ Min($[C_1, C_2, C_3]$) `//` *RampDown* `Event`
$E_2 \leftarrow$ Max($[C_1, C_2, C_3]$) `//` *RampUp* `Event`

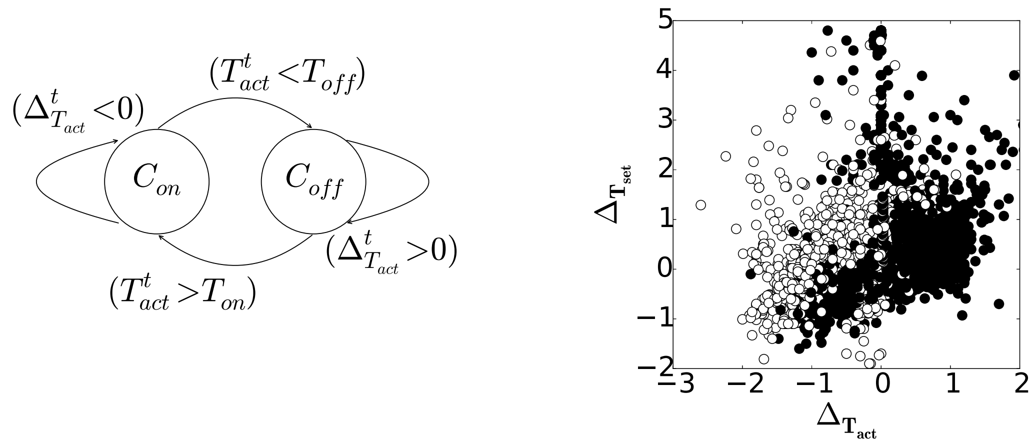
`/* Detect AC state on labeled series */`
Event Detection:
$S_{est}^{t^0} \leftarrow 1$ `// AC state at start of usage`
curr_event $\leftarrow E_1$ `// Current event`
search $\leftarrow E_2$ `// Next event to detect`
for $t \leftarrow t^{(1)}$ **to** t^{off} **do**
  if $C(\Delta_{T_{act}}^{(t)}) == search$ **then**
    curr_event $\leftarrow$ search;
    if $search == E_2$ **then** $S_{est}^{(t)} \leftarrow 0$, search $\leftarrow E_1$;
    else $S_{est}^{(t)} \leftarrow 1$, search $\leftarrow E_2$;
  else Retain previous state and event ;

AC was on at set temperature (T_{set}). To time align, this module resamples the collected data to 1 reading every 2 minutes and interpolates the missing values using linear interpolation[1].

Estimation

Figure 2-3a exhibits the correlation between the room temperature and AC power consumption. As the room achieves a temperature T_{off}, close to T_{set} (24°C), the compressor turns off resulting in an increase in the room temperature. When the room temperature increases beyond a threshold, T_{on} (this threshold varies for different AC and their set temperatures), the compressor is turned on again, and the room temperature starts decreasing. As a result, the room temperature keeps fluctuating around T_{set} resulting in multiple compressor cycles during an AC usage. Thus, we assume an AC to be a device with two-states namely (see Figure 2-4a); (1) Compressor On (C_{on}), and (2) Compressor Off (C_{off}), where transition depends upon T_{on}, T_{off}, $\mathbf{T_{act}}$, and $\Delta_{\mathbf{T_{act}}}$ defined as $\Delta_{T_{act}}^{(t)} = T_{act}^{(t)} - T_{act}^{(t-1)}$. $T_{act}^{(t)}$

[1] https://en.wikipedia.org/wiki/Linear_interpolation



(a) State transition diagram of AC

(b) AC State Classification

Figure 2-4: AC as two state diagram with possible transitions (Left). $\Delta_{\mathbf{T_{act}}}$ and $\Delta_{\mathbf{T_{set}}}$ sufficient features to classify AC state while learning and predicting (Right). White dot indicates C_{on}, while black dot indicates C_{off} state.

is the room temperature and $\Delta^{(t)}_{T_{act}}$ signifies an increase or decrease in room temperature at t_{th} time instance. We observed from our dataset that during state transition there is a sudden decrease (*RampDown*) and increase (*RampUp*) in room temperature corresponding to C_{on} and C_{off} states respectively (Figure 2-3b). Leveraging on this observation, we run K-Means[2] clustering over $\Delta_{\mathbf{T_{act}}}$, to identify *RampUp* and *RampDown* events within the usage followed by event matching (Figure 2-3c) to detect corresponding AC state at any time during the usage. PACMAN-L uses estimation algorithm (Algorithm 1) to output $\mathbf{S_{est}}$, a vector of 0 (C_{off}) and 1 (C_{on}) indicating AC state at any point during the usage.

$$Energy = P_{rated} \times \frac{(\sum_{t=t^{(0)}}^{t^{off}} S^{(t)} - 1)}{60} \times 2 \tag{2.1}$$

Equation 2.1 maps AC state ($\mathbf{S}$) to an estimate of AC energy consumption (*Energy* in KWh) using the rated power (P_{rated} in KW) mentioned by the AC manufacturer.

[2]Though we are aware that K-Means clustering is sensitive to the outliers; with prior knowledge about the number of distinct spherical clusters (k=3) and limited data size within an AC usage, we found K-Means to be most effective clustering technique for the PACMAN.

Learning

The control circuitry for a particular AC uses its set temperature to switch the compressor state. However, the room temperature (as observed at another location in a room, and used in PACMAN-L) corresponding to the compressor on or off instance, deviates from the temperature seen by AC (Figure 2-1a). To simplify this complex interdependence between T_{act} and T_{set}, we learn an AC state classifier over $\Delta_{\mathbf{T_{act}}}$ and $\Delta_{\mathbf{T_{set}}}$ (defined as $\Delta_{T_{set}}^{(t)} = T_{act}^{(t)} - T_{set})$, sufficient to classify AC state (Figure 2-4b). We use Random Forest Classifier (RDF) as AC state classifier over $\Delta_{\mathbf{T_{act}}} = \{\Delta_{T_{act}}^{t^{(1)}}, \Delta_{T_{act}}^{t^{(2)}}, \Delta_{T_{act}}^{t^{(3)}}, ...\}$ and $\Delta_{\mathbf{T_{set}}} = \{\Delta_{T_{set}}^{t^{(1)}}, \Delta_{T_{set}}^{t^{(2)}}, \Delta_{T_{set}}^{t^{(3)}}, ...\}$ using a python library [130].

Next we learn our thermal model, which is an adapted version of earlier presented model [142]. Room temperature (T_{act}) at any time instance depends upon numerous factors as discussed in previous sections. However, we explicitly considered the effect of AC cooling and external temperature for our thermal model (Equation 2.2), while accounting for all other dynamic factors as thermal noise in the room.

$$\Delta_{T_{act}}^{(t+1)} = \alpha * (T_{act}^{(t)} - T_{ext}^{(t)}) + \beta * S_{est}^{(t)} + \gamma * T_{est}^{(t)} + \epsilon \tag{2.2}$$

$T_{ext}^{(t)}$ is the external temperature and $S_{est}^{(t)}$ depicts estimated AC state, at the t_{th} instance of the AC usage. Learned parameters of the model are; (1) Leakage Rate (α), (2) Cooling Rate (β), (3) Thermal noise due to external sources (γ), and (4) Thermal noise due to internal sources (ϵ). To learn unknown parameters ($\theta = \{\alpha, \beta, \gamma, \epsilon\}$), we first compute dependent variables ($\Delta_{T_{act}}^{(t+1)}$) and independent variables $[(T_{act}^{(t)} - T_{ext}^{(t)}), S_{est}^{(t)}, T_{ext}^{(t)}]$ for previous n recent usages and then use linear regression to learn the unknown parameters.

2.5.2 PACMAN-P

PACMAN captures AC set temperature from the occupant, together with her location, to predict AC energy consumption for d_{pred}, the time duration for which occupant wants to predict. By default, PACMAN-P predicts for a duration which is average of all historical usages. For the sake of evaluation purpose we assume $d_{pred} = d$ for every usage.

Algorithm 2 presents our prediction algorithm to predict room temperature ($\mathbf{T_{pred}}$) and AC state ($\mathbf{S_{pred}}$) before using the air conditioner. Prediction algorithm uses parameter vector, $\theta = \{\alpha, \beta, \gamma, \epsilon\}$ learned over recent n historical usages, current room temperature

Algorithm 2: Algorithm to predict $\mathbf{T_{pred}}$ and $\mathbf{S_{pred}}$

Input: $T_{act}^{(0)}$, T_{set}, d_{pred}, θ, $\mathbf{T_{ext}}$
Output: $\mathbf{T_{pred}}$, $\mathbf{S_{pred}}$

`/* Current Temperature and AC State */`
Data Initialization:
$T_{pred}^{(0)} = T_{act}^{t(0)}$;
$S_{pred}^{(0)} = 1$;

`/*` Predict $\Delta_{\mathbf{T_{act}}}$ and AC state `*/`
Prediction Loop:
for $t \leftarrow t^{(0)}$ **to** t^{off} **do**
  $\Delta_{T_{act}}^{(t+1)} = \alpha * (T_{pred}^{(t)} - T_{ext}^{(t)}) + \beta * S_{pred}^{(t)} + \gamma * T_{ext}^{(t)} + \epsilon$;
  $T_{pred}^{(t+1)} = T_{pred}^{(t)} + \Delta_{T_{act}}^{(t+1)}$;
  $\Delta_{T_{set}}^{(t+1)} = T_{pred}^{(t+1)} - T_{set}$;
  $S_{pred}^{(t+1)} = RDF(\Delta_{T_{act}}^{(t+1)}, \Delta_{T_{set}}^{(t+1)})$;

($T_{act}^{t(0)}$), AC set temperature (T_{set}), prediction duration (d_{pred}) and, weather forecast ($\mathbf{T_{ext}}$) as input to first predict $\Delta_{\mathbf{T_{act}}}$. Then it uses learned AC state classifier (RDF) along with $\Delta_{\mathbf{T_{act}}}$ to predict AC state ($\mathbf{S_{pred}}$) for the entire prediction duration (d_{pred}).

2.5.3 PACMAN-Engine

PACMAN-Engine uses Wunderground[3] for weather forecast, historical weather data and real time update of outside temperature. It also manages learned models from historical usages along with meta information such as d, $T_{act}^{t(0)}$, $T_{ext}^{t(0)}$, T_{set} and various other statistics related to prediction and estimation for every AC usage.

2.6 Evaluation

We conducted an in-situ study in seven different rooms of seven homes over a period of approximately three months. During our data collection, we collected room temperature, power consumption, and external temperature data without interfering occupants' daily schedule. From each home, we collected data in different time intervals across three months. We used power consumption data as ground truth for the evaluation purpose. Table 2.1 summarizes environment of the seven rooms with their AC specifications.

[3] http://www.wunderground.com/

Experimental Setup			Air Conditioner Technical Specifications				Room Specifications		
#ID	**#Days**	**Used Set Points**	**Model**	**Tonnage**	**Rated Power (Watts)**	**Floor Area (sq.mt)**	**Window Width (mt)**	**Occupancy**	**Floor**
1	15	27, 28	O-General AXZB18GNL-W	1.5	2180	20.06	2	1-2	2/2
2	9	20, 23, 26	O-General	1.5	2180	13.2	2	2-3	2/12
3	13	25, 26, 27	Hitachi RAU518HTD	1.5	1570	20.06	2	1-2	5/12
4	41	28	Haier HW-18L2H/2013	1.5	1950	20.06	2	1-2	5/12
5	9	25	Carrier GWRAC 018DR002	1.5	1950	13.2	2	3-4	6/12
6	12	28, 29	LG L-Crescent 1.5TR 5-Star	1.5	1600	20.06	2	1-2	7/12
7	9	18, 25	Hitachi RAV518HTD	1.5	1575	20.06	2	1-2	9/12

Table 2.1: We collected data from seven rooms. Second column represents number of days we collected the data, third column shows various T_{set} used during the study. AC technical specifications are as supplied by the manufacturer. Room specifications, that impact thermal modeling, includes occupancy denoting average # of occupants in the room and floor ratio representing proportion of room's floor number to total floors in the building

2.6.1 Experimental Setup

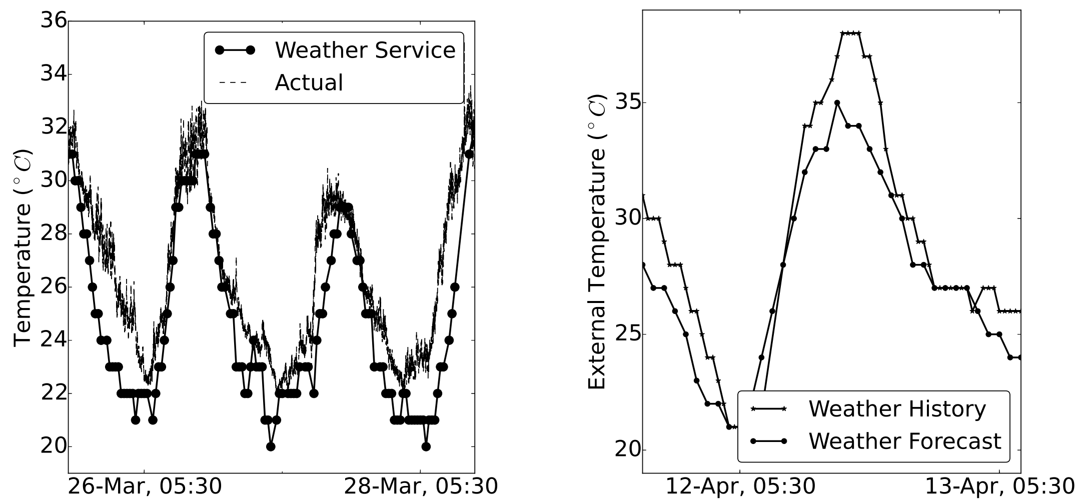
We interfaced a temperature sensor from Maxim[4] (accuracy ±0.5°C) over a single board computer[5] to record room temperature data at 1Hz. Each home had a three phase power supply with a clear knowledge of the phase supplying power to the AC. Using modbus enabled smart electricity meter, we extracted the AC power consumption from the phase level power consumption. We recorded weather conditions using Wunderground API that provided two readings of historical weather information for every hour and one reading of weather forecast every hour. Besides, we also collected usage related information such as time at which AC was switched on and switched off along with set temperature of AC for that usage from the occupants. During this in-situ deployment, we recorded average external temperature to be 28°C, with the minimum at 25°C and maximum at 33°C.

2.6.2 Evaluation Criteria

Using Equation 2.3 and 2.4, we evaluate the performance of PACMAN in estimating (E_{est}) and predicting (E_{pred}) AC energy consumption when compared with actual energy consumed (E_{act}), collected for evaluation purpose.

$$E^{error}_{est/pred} = \frac{E_{est/pred} - E_{act}}{E_{act}} \times 100 \tag{2.3}$$

[4]http://datasheets.maximintegrated.com/en/ds/DS18B20.pdf
[5]http://www.raspberrypi.org/



(a) Sensor v/s Website

(b) Historical v/s Forecast

Figure 2-5: Validation of different datasets used for the empirical evaluation of PACMAN.

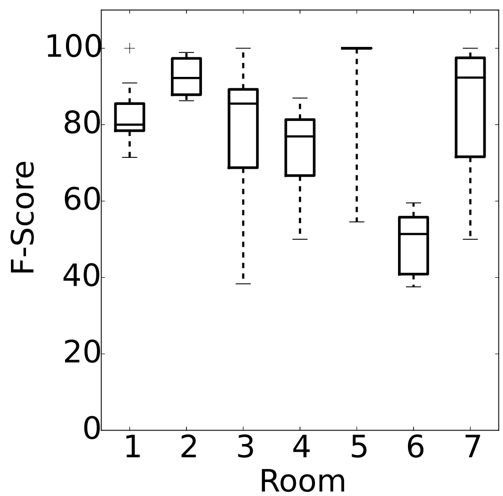
while, prediction accuracy as:

$$E^{acc}_{est/pred} = 100 - |E^{error}_{est/pred}| \tag{2.4}$$

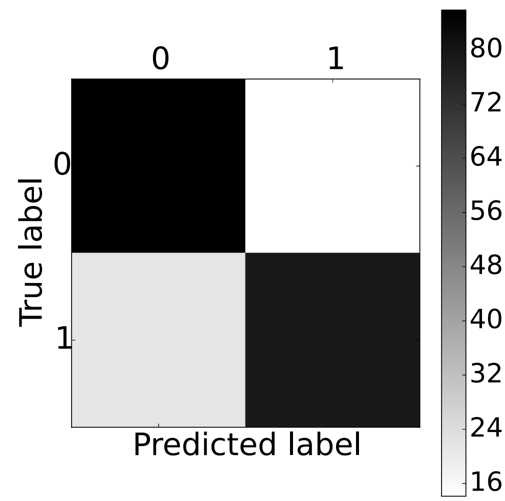
where, $E^{error}_{est/pred}$ is the error in estimating/predicting AC energy consumption and $E^{acc}_{est/pred}$ denotes estimation/prediction accuracy of AC energy consumption.

2.6.3 Data Validation

Datasets used as proxy, e.g. weather data from weather service as a proxy for external temperature and historical external temperature for the weather forecast, are validated for their accuracy before using for the evaluation purpose. To validate the accuracy of external temperature data as provided by weather service, we recorded temperature in our campus for few days. For the same duration, we extracted historical weather data from the weather service. Results (Figure 2-5a) show that the historical temperature provided by the weather service is close to the actual external temperature with a mean absolute deviation (MAD) of 0.3°C. Since PACMAN takes weather forecast, from the weather service, for the prediction of AC energy consumption, the E^{acc}_{pred} depends on the accuracy of the weather forecast. We collected both the forecasted temperature and historical temperature from the website for the duration of 34 hours and observed an MAD of 0.8°C, as shown in Figure 2-5b.



(a) F-Scores



(b) Confusion Matrix

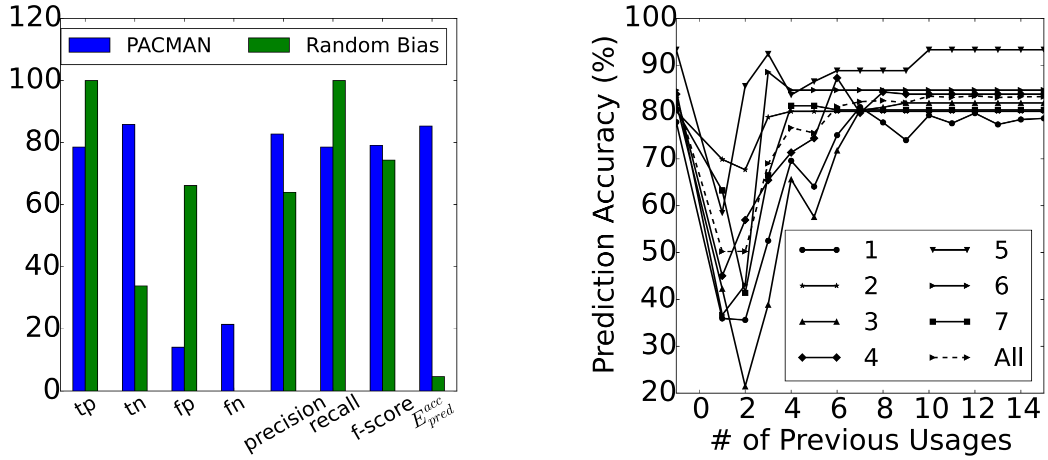
Figure 2-6: PACMAN achieved higher F-Scores across all the rooms except for the one with improper sensor positioning (e.g Room-6). It classified majority percentage of C_{on} events (80%) when averaged across all the rooms

2.6.4 Analysis of AC State Detection during Estimation

The performance of PACMAN in estimating AC energy consumption depends upon the accurate classification of AC state at any time instance during the usage. The result (Figure 2-6) show that temperature data is sufficient to estimate AC state at any instance of time with an average F-Score of 80.0% across all the rooms and the best F-score being 94.5%. We observed a poor classification accuracy for Room-6, where sensor was kept far away from the AC. This introduced a lag between the actual and predicted values leading to poor F-Scores but decent prediction and estimation accuracy. However, classification accuracy can be improved in such scenarios by keeping sensor close to the AC. We also compared our approach with a random bias approach, i.e. assuming AC state to be C_{on} always (Figure 2-7a) except for two time instances; (1) Before occupant switches on the AC, (2) After occupant switches off the AC. However, such a naive approach had comparable f-score but high number of false positives and thus poor prediction accuracy (E^{acc}_{pred}).

2.6.5 Sensitivity Analysis

PACMAN learns from historical n usages to predict for the $(n+1)^{th}$ usage. To select the value of n, we measured the average prediction accuracy while varying its value till 15 (Figure 2-7b). We observed that average prediction accuracy over all the rooms (dotted



(a) PACMAN v/s Random Bias

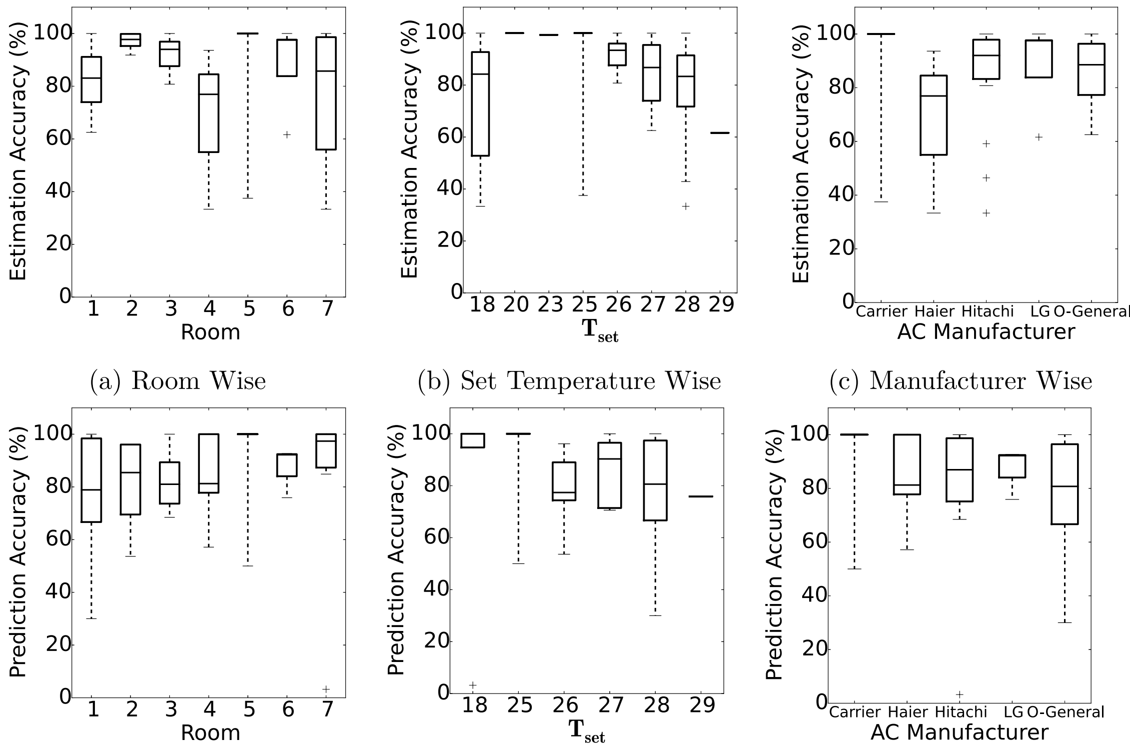
(b) Accuracy stabilizes at $n = 10$

Figure 2-7: [2-7a] Although *Random Bias* had comparable F-Scores but poor prediction accuracy. [2-7b] For $n = 10$ (recent historical usages), system achieved the peak value

line) achieves peak at 10 and show insignificant improvement beyond that. Thus, we used 10 recent historical usages to learn our thermal model to predict for the next usage.

2.6.6 Analysis of AC Energy Estimation & Prediction Accuracy

We evaluate the estimation and prediction accuracy of PACMAN for each room (with different combination of room and AC specifications), each set temperature across different homes and ACs, and across different AC manufacturers to understand the general applicability of PACMAN across diverse usages. Figure 2-8a presents the variation in estimation of AC energy consumption, using room temperature data, across all the rooms. Similarly, Figure 2-8d presents box plot for room level prediction accuracy in energy forecast of AC usage. We recorded an estimation accuracy of 85.3% (Best-97.0%) and prediction accuracy of 83.7% (Best-93.3%) when averaged over all the rooms. Further, during our experiment, occupants used AC at various set temperatures, as shown in Column 3 of Table 2.1. Figure 2-8e and 2-8b shows that neither the estimation and nor the prediction were biased to a specific set temperatures, for all the rooms. For example, since 18°C was used only for Room-7, its variation in estimation and prediction accuracy is similar. Across all the set temperatures, we observed that the average prediction accuracy was 82.4% with best accuracy of 93.2%, while the average estimation accuracy was 85.2% with best accuracy of 100.0%. For prediction, we had few set temperatures as it excludes those points that were



(a) Room Wise (b) Set Temperature Wise (c) Manufacturer Wise

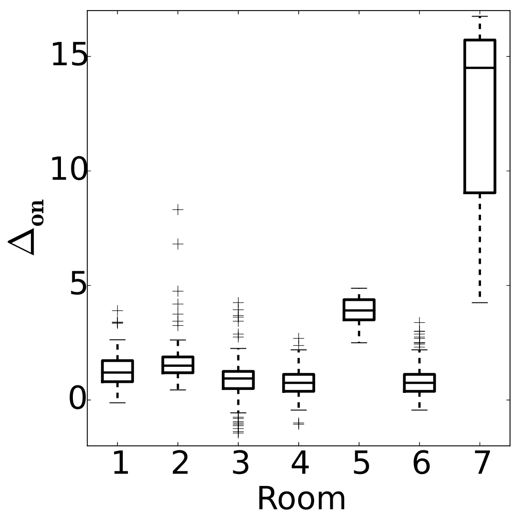
(d) Room Wise (e) Set Temperature Wise (f) Manufacturer Wise

Figure 2-8: Comparison of AC energy estimation [2-8a, 2-8b, 2-8c] and prediction [2-8d, 2-8e, 2-8f] accuracy based on various influencing parameters - room (for thermal properties), set temperature (being directly used in the prediction model) and AC manufacturer (possibly leading to different control mechanisms).

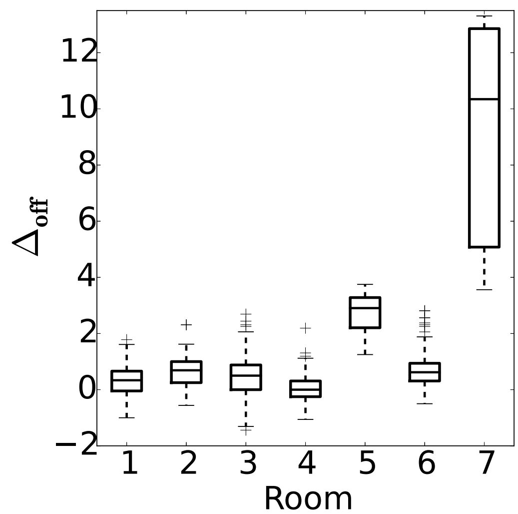
used for learning the thermal model initially. There were only a small number of usages at some of the set temperatures thus resulting in smaller variation in prediction accuracy for such temperature settings. Figure 2-8c and Figure 2-8f illustrates energy prediction and estimation accuracy, respectively, based on variations in the model of ACs used in our experimentation. We observed an average accuracy of 84.1% and 85.0% with best accuracy of 92.6% and 93.3% for the estimation and prediction of AC energy consumption, respectively.

2.7 Challenges And Discussion

PACMAN is principally designed for room level ACs and heavily depends on the room temperature. However, room temperature depends upon various factors such as sensor positioning, number of occupants and their activities, among others. In this section, we analyse the outliers in estimation and prediction accuracy of PACMAN, caused primar-



(a) Δ_{on}



(b) Δ_{off}

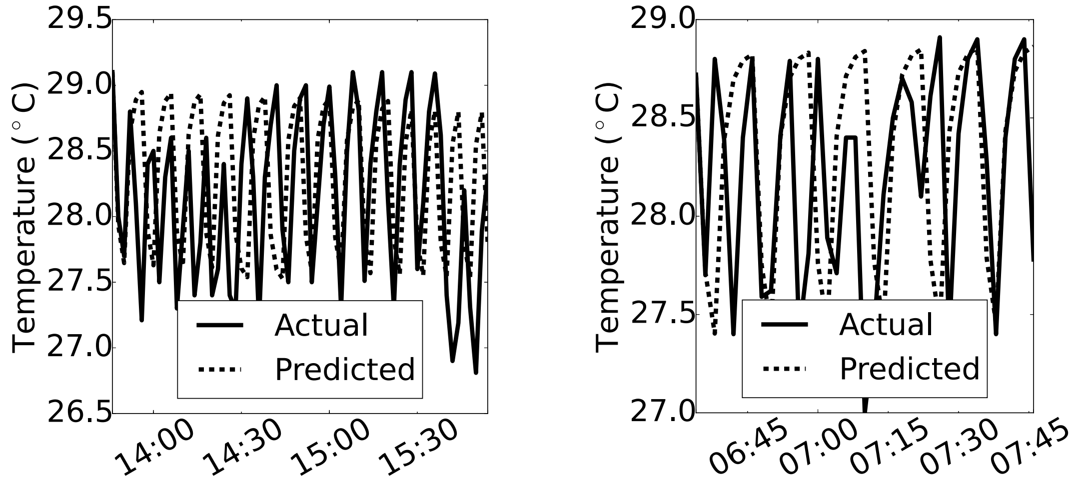
Figure 2-9: Δ_{on} [2-9a] and Δ_{off} [2-9b] varies across all the rooms during usages.

ily from imperfect applicability of the thermal model. This require further research into inclusion of multiple thermodynamic properties to enrich the current model being learned.

2.7.1 Sensor Position

In a separate controlled experiment, we installed four different temperature sensors at different locations in the room. Figure 2-1a shows that temperature seen by AC (L1) is very different than ambient temperature at other locations (L2, L3, L4). These variations impact estimation accuracy as temperature drop near the AC (major factor in estimation algorithm) is significantly larger than a location farther to AC.

Also, AC control algorithm depends upon the return air temperature that deviates from room temperature at any other point in the room. Δ_{on} is the temperature difference between room temperature and compressor on temperature, T_{on}. Δ_{off} is the difference between room temperature and compressor off temperature, T_{off}. Ideally, both T_{on} and T_{off}, should remain constant during an AC usage. However, thermal noise around the sensor results in varying Δ_{on} (Figure 2-9a) and Δ_{off} (Figure 2-9b), as also observed in our data collection. Lower was the set temperature, more were the deviations (Room-7). This makes a major impact on prediction accuracy because predicted state change might vary from actual state change of AC and this error accumulates over the time for each usage (Figure 2-10a). We deployed temperature sensor in some open space within the room while also accounting for esthetics



(a) Impact of Δ_{on} and Δ_{off} (b) Occupants' Intervention

Figure 2-10: Variation in Δ_{on} and Δ_{off} leads to misaligned predicted and recorded temperature cycles [2-10a]. Similarly, occupants' intervention distorts actual temperature cycles resulting in poor classification of AC state during predictions [2-10b].

in the room to ensure its presence is comfortable to the occupants. In the future, we plan to enrich PACMAN to automatically classify the location of temperature sensor with reference to the AC in the room and accordingly adapt the thermal model for improved accuracy.

2.7.2 Occupants' Interference

Occupants interfere in two ways, either by changing AC settings or by performing activies that impact temperature surrounding the sensor during a usage. Changing AC set temperature was observed frequently in our deployment across all the rooms resulting in poor prediction accuracy in such scenarios as PACMAN predicted assuming T_{set} to be constant during the usage. Similarly, real-time activities of occupants such as to-and-fro movements from one room to another, turning on/off other appliances (fans) also affect room temperature significantly. For example (Figure 2-10b), occupant woke up around 07:00 AM and impact of his activity is visible. However, annotating such activities for every usage require more instrumentation. Such distorted cycles impact both estimation and prediction accuracy as their algorithms being function of room temperature. However, these challenges can be addressed by incorporating context information such as occupancy information to learn a daily schedule of the occupants for the predictions.

2.7.3 Long Usages during Night

During the night, generally AC run for long durations and external temperature (initially higher) goes lower than the room temperature in mid of the night. In these scenario, where the external temperature is lower than the AC set temperature, the compressor should ideally turn off. PACMAN prediction model, accounting for variation in external temperature, will also work as per this ideal scenario. However, due to additional sources of thermal noise in the room, the compressor may still continue running, leading to poor prediction accuracy.

Variations in room temperature depend upon numerous factors such as number of occupants, human activities, the number of doors & windows together with whether they are open or closed and thermal appliances present in the room. However, annotating each factor can make the model complex along with making it cumbersome for the occupant to accurately report each of these parameters as and when they change. Therefore, while compromising on the accuracy in some of the usage instances, we simplified the thermal model with T_{set} as the only input by the occupant. It is important to note that due to these reasons, accurate prediction of AC state at any instant is a very complex task. For the use case (giving energy consumption feedback to the occupants) discussed in this chapter, even if the exact state at majority of the instances is inaccurate, it is important to achieve high level of accuracy in energy consumption prediction and estimation. Such a behavior was also observed with PACMAN whereby while the precision and recall for the prediction were inaccurate the overall energy prediction accuracy was still reasonably good.

Chapter 3

Comfortable Energy Savings

Automation applied to an inefficient operation will magnify the inefficiency.

– Bill Gates, Founder of Microsoft Corporation

3.1 Introduction

If followed, feedback can lead to significant energy savings, however, the realization of these energy savings mainly depend on the residents' attitude; thus can benefit from automation [29]. Several studies [34, 13] indicate that efficient and optimized usage of an AC can save significant energy while maintaining the user comfort. To feed in the comfort requirements, tenants typically set the AC thermostat temperature (also referred as the set temperature) which varies across people, depending on their personal preferences. For instance, the comfort of *Alice*, *Bob*, and *Eve* goes up and down for distinct values of set temperature while assuming the same effect of extrinsic factors such as weather (Figure 3-1). People similar to *Eve* prefer a higher range of thermostat temperature and others like *Alice* feel pleasant if AC operates at low temperature for a particular climate. Though set temperature is a matter of choice, it directly influences the AC energy consumption. Portrayed by the black arrow (in Figure 3-1), the AC energy consumption decreases as the user raises the thermostat temperature for a given atmospheric condition. This way, with *Alice*'s preferences, she will always end up consuming on the higher side and start believing (rather incorrectly) that savings are beyond the bound of possibilities.

People feel comfortable in a range of temperatures. *Alice* is unaware of the fact that

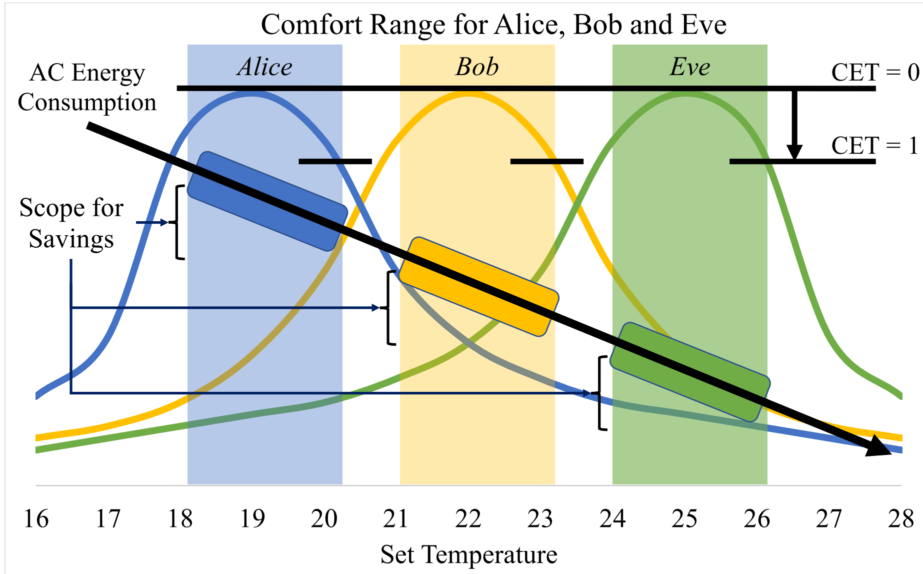


Figure 3-1: For a given weather condition, *Alice*, *Bob*, and *Eve* might feel comfortable at different set temperatures depending on their individual preferences. Even within their comfortable band (indicated by shaded region), there is a scope of varying the thermostat temperature to achieve better comfort while minimizing the AC energy consumption. Portable+ thermostat empowers the residents in deciding between attaining the peak comfort ($CET = 0$), maximizing the energy savings for $CET = 1$ (i.e. power saving mode), or maintaining a balance between both ($0 < CET < 1$).

whether she set the thermostat at $18°C$, $19°C$, or $20°C$, she will perceive similar sensations (depicted by blue shaded area in Figure 3-1) for a given environment. Now, assume that *Alice* starts the AC at $18°C$ thinking that lowering the thermostat temperature will make the room more comfortable. She hardly knows that raising the set temperature to $19°C$ can attain the peak comfort (for her) while also reducing the AC energy consumption. Moreover, moving the set temperature to $20°C$ can further increase the savings (from AC) while making her slightly (and possibly within an acceptable range) uncomfortable. With limited knowledge about the comfort-energy trade-off, residents (like *Alice*) prefer to set a fixed thermostat temperature in their daily routine. Therefore, a thermostat should report the users, *"What temperature settings will provide personal comfort and the cost efficiency?"* [150]. State-of-the-art commercially available portable thermostats [152, 148] and smart ACs [100] emulate the IR (infrared) commands of the AC remote; thus allowing them to control their AC even from their smartphone. Besides, they also monitor the user location to provide "comfort" features such as location based on/off and precooling the room as per the schedule set by the user. Although portable thermostats collect temperature data from occupied area, they neither monitor the user comfort within the room nor assist the residents in achieving their desired comfort and savings, simultaneously. A thermostat should provide occupants with a way to adjust and control the conditions; if not given, controls become ineffective, inappropriate or unusable [120]. Thus, there is a high motivation to provide users with an additional knob that allows them to specify their goal beyond the set temperature - minimize AC energy consumption (savings), maximize comfort, or maintain a balance between both. Based on the input specified by residents on the knob, the thermostat should adjust the set temperature to attain the peak comfort and energy savings, simultaneously. Control approaches applying automated set-point variation (i.e. intelligent) save up to 26% energy when compared with scheduled set temperature [115].

We build portable+ thermostat on top of the existing portable thermostat by adding a knob (varying between 0 and 1) to let the residents optimize between comfort and energy consumption. We call this knob CET - Comfort Energy Trade-off. The knob is part of the mobile application (used to remotely control the AC) where tenants mention their interest (along with the set-point temperature) - maximise savings, or maximise comfort, or balance both. Nicol et. al. [120] pointed out three contextual variables to decide if conditions are comfortable for the residents - climate, nature of the building, and time. Taking the work forward, we realize the proposed thermostat through an optimization framework that considers all three stated variables. The framework leverages a physics-based thermal model to simulate the room environment for any weather conditions. To learn the model parameters, framework records the room temperature data and adapts the generic model to imitate the thermal behaviour of the room. With time, as framework collects more and more data, it keeps updating the model parameters to accommodate temporal changes in building's thermal behaviour. For a given weather forecast, the tuned thermal model works as a simulation testbed to decide the optimal thermostat temperature for the AC, while optimizing for occupants' comfort and the energy savings, independently and jointly. In this work, we consider a lumped parameter thermal model that divides the room into multiple zones and applies for thermal balance across them (Section 3.3.2). Given the modular architecture of the framework, the thermal model (and the other components) are replaceable with their refined versions to further improve the outcome.

We analyzed the benefits of enhancing portable to portable+ thermostats towards reducing the user discomfort and improving the energy savings. For the study, we gathered temperature data from a two-week controlled experiment in two rooms (a bedroom and a living room) of a home and a three-month in-situ deployment across five rooms of five different homes. For the stated dataset, our analysis indicates that portable+ thermostats can reduce residents' discomfort by 23% when trying to attain peak comfort, and save 26% energy in power-saving mode. Pervasive adoption of portable+ thermostats can result in significant aggregate reductions at the global scale to attain sustainable use of energy.

To summarize, the major contributions of this chapter are:

1. Extending portable thermostats to a smart portable+ thermostats that equip the residents with a CET knob to specify their comfort requirements along with the set temperature.

2. Optimization framework to dynamically vary the set temperature (for the AC) as climate changes; thus, achieving the desired CET value, as input by the users on portable+ thermostat.

3. Field evaluation of extending portable to portable+ thermostats for a lumped thermal model, demonstrating simultaneous savings in energy consumption and improvement in user comfort.

3.2 Related Work

Programmable thermostats were one of the initial attempts in the direction of energy savings in central HVAC (Heating, Ventilation and Cooling) units across residential buildings. Beyond the conventional thermostats, they offered an additional feature to set HVAC operating schedule. For a (regular) home occupied by multiple people having their individual calendar to follow, programmable thermostats proved to be an inefficient approach [132, 94]. Smart thermostats [107] surpassed the shortcomings of programmable thermostats by learning residents occupancy patterns to control the HVAC. Google Nest [118] is one such realization of smart thermostats for residential spaces. Recent studies [72, 96, 134, 146] evinced the feasibility of energy-efficient control for centralized HVAC units while using reactive and predictive thermostats. However, in contrast to developed countries, room level split or window air conditioners (decentralized ACs) are common across developing countries. Given an inbuilt thermostat (having an inbuilt controller) of the window ACs, it is infeasible to use commercial smart thermostats with them.

Decentralized ACs are ubiquitous in many parts of the world; thus optimizing their energy consumption is need of the hour. Appliance-level feedback can result in 5-15% savings in power consumption [33]. Understanding the noteworthiness of the appliance-level feedback, PACMAN informed occupants about the impact of changing thermostat temperature on AC energy consumption in the Indian context [77, 78]. MyJuolo used temperature data collected from a USB logger to learn a thermal model of the room to infer the optimal thermostat temperature for the occupants [142]. Lork et. al. [106] proposed a data driven framework for AC load forecasting using advanced machine learning techniques. Previous studies [102, 103] also proposed control algorithms to optimize duty-cycles of the AC compressor. Some studies [9, 99, 101] also suggested upgrades for various AC components for energy-efficient usage of the decentralized ACs. While the feedback based approaches depend on the residents to decide their comfort, hardware upgradations (of window ACs) makes various assumptions about the human comfort during the design phase.

To further ameliorate the user experience, manufacturers introduced smart and web-based ACs having numerous modes of operations (such as sleep mode), along with the capability to control the AC from the smartphone. Similarly, state-of-the-art portable thermostats ([152], [148]) allow residents to operate their AC locally as well as remotely (from any other location) through their smartphones. When the user's smartphone location is close to home, they pre-cool/heat the room (to improve residents' comfort) and switch AC off (to save energy) as soon as occupants leave their residence. However, understanding the impact of outside weather conditions and thermostat temperature towards energy savings and occupants comfort (as studied in this chapter) provide an attractive value proposition for window AC units [34, 67, 66]. Such an analysis is largely absent in research literature that primarily focused on optimizing central HVAC for developed countries.

The specified advances (in the domain of residential conditioning) neither monitors the user experience nor quantify the effect of their energy-efficient approach on the residents. Studies [138, 52] proposed thermal imaging to monitor thermal comfort of the occupants while attempting to reduce energy footprints through air conditioning. In residential apartments, camera sensing will be a breach of the residents' privacy. McCartney et. al. [111] shows that changing the set temperature with outdoor temperature does not increase the discomfort while raising the energy savings. Following on these studies, we pro-

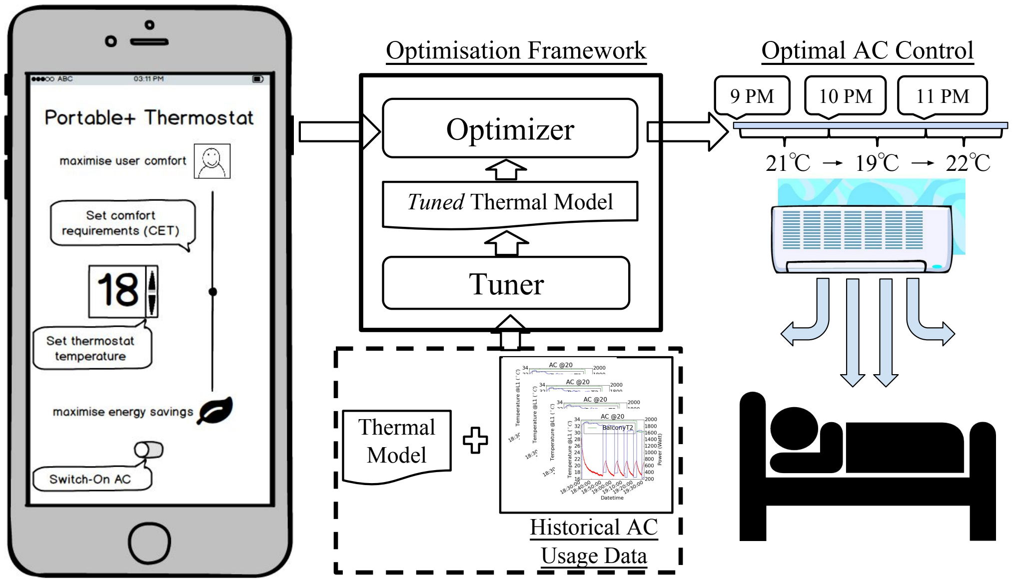


Figure 3-2: Left to Right: *Alice* mentions the thermostat temperature and her comfort requirements (CET) through her smartphone before switching on the AC. Besides the user input, the *optimizer* (of the framework) also takes a tuned thermal model generated by learning the parameters of a generic (thermal) model using the data from recent n historical AC usages from Alice's room. The set temperature, CET, and tuned thermal model act as an input for the framework to predict optimal set temperature for the *Alice* for a given weather forecast. Though the *optimizer* predicts for a fixed duration, it periodically updates the set temperature with changing atmospheric conditions.

pose portable+ thermostat with a CET knob to specify comfort requirements along with the thermostat temperature. To realize we propose an optimization framework that utilizes a theoretical thermal model to vary the set temperature while considering changing weather conditions. The framework automatically learns the model parameters (before every AC usage) using historical data to adapt the thermal model to any room while also accounting for changes in the local conditions with time.

3.3 Optimization Framework

Beyond set temperature, a portable+ thermostat provides the user with an additional knob of CET (Comfort-Energy Trade-off) realized through an optimization framework comprising of a *tuner* and an *optimizer* (Figure 3-2). Initially, *tuner* of the framework takes a lumped thermal model and tunes its parameters for a particular room environment after gathering the data from a few AC usages in that room. The process of using real-world data to tune parameters of a thermal model is called *Grey Box Modelling* and widely practised by the researchers in various domains. Next, the *optimizer* takes set temperature and CET (from the user), and the tuned thermal model (from the *tuner*) to optimally control the AC. The tuned thermal model along with current room temperature and the weather forecast empowers the *optimizer* in simulating temperature across the room. The simulated data for the thermostat temperature (set by the user) acts as a reference point to compute the comfort requirements of the user i.e. where it lies on the curve (from Figure 3-1). The *optimizer* utilizes the calculated comfort to look for an optimal thermostat temperature to achieve the desired CET value. On completion of ongoing usage (and prior to next AC usage), framework updates the model parameters while using the data gathered from the current usage and a few other recent historical AC usages.

3.3.1 Optimizer

An AC contains an inbuilt thermostat that uses *on* and *off* hysteresis to maintain the desired room temperature, set by the user on the thermostat. When temperature sensor (of the inbuilt thermostat), embedded within the AC, senses that room temperature is close to the lower threshold (*off* hysteresis), the controller (of the inbuilt thermostat) shuts down the compressor, a major power consuming component of the AC. Correspondingly, when the room temperature attains the upper threshold (*on* hysteresis), the controller again turns on the compressor. At any time instance during the AC usage, the thermostat temperature (provided by the user) is used by the AC to decide whether to keep compressor *on* or *off*. Compressor being the major power consuming component of the AC, its state (at any time instance) principally drives the AC energy consumption. Thus, to improve the user experience, the *optimizer* finds a set temperature with least impact on the user comfort while minimizing the compressor *on* duration.

When a resident turns on the AC, (s)he sets the thermostat temperature and a CET value. The *optimizer* utilizes the set temperature to simulate the room temperature and AC compressor state. However, within a room, the temperature in any region depends on numerous factors such as the distance from the AC, exposure to external weather conditions, and heat transfer from the adjacent spaces, among others. Thus, the framework leverages a theoretical thermal model to understand temperature variations within the room. As thermal conditions may vary significantly from one home to another, the *tuner* adjusts the model parameters for a particular environment using the data from its historical AC usages.

Sensation	pmv_{ll} (Lower Limit)	pmv_{ul} (Upper Limit)	pmv_{ref}
`Cold`	-	-2.5	-3.0
`Cool`	-2.5	-1.5	-2.0
`Slightly Cool`	-1.5	-0.5	-1.0
`Neutral`	-0.5	+0.5	0.0
`Slightly Warm`	+0.5	+1.5	+1.0
`Warm`	+1.5	+2.5	+2.0
`Hot`	+2.5	-	+3.0

Table 3.1: PMV ranges for each sensation level

The optimization framework, using the tuned thermal parameters and weather forecast, predicts AC compressor state (to compute *savings*) and the region-wise temperature (to calculate region-wise *discomfort*) for various thermostat temperature. While *savings* (in AC energy consumption) quantifies the impact of a particular set temperature on AC energy consumption, *discomfort* signifies the effect on occupants' comfort. The *optimizer* outputs a thermostat temperature that maximizes the weighted average of *savings* and *discomfort* for a given value of CET.

$$savings\ (\%) = \frac{E_{act} - E_{new}}{E_{act}} \times 100 \tag{3.1}$$

Energy Savings

It is the percentage change in AC energy consumption before (E_{act}) and after (E_{new}) changing the thermostat temperature (Equation 3.1).

$$discomfort = \mid pmv_{act} - pmv_{ref} \mid \tag{3.2}$$

Occupants' Discomfort

Franger's Predicted Mean Vote (*pmv*) is an ISO 7730 standard and widely used to estimate human comfort within the buildings [45]. It assigns a numerical value based on 1. ambient air temperature ($°C$), 2. mean radiant temperature ($°C$), 3. air velocity in the room (m/s), 4. relative humidity (%), 5. metabolic rate of the occupants (*met*), and 6. clothing insulation of the occupants (*clo*). On a thermal scale, *pmv* varies from `Cold` (-3) to `Hot` (+3) and consists of seven sensitivity levels for the humans (Table 3.1) each having a lower (pmv_{ll})

and an upper bound (pmv_{ul}). Comfort is subjective and differs from person-to-person. Given that people adapt themselves to a particular environment, their preferred sensation often deviates from `Neutral` [68, 138]. Therefore, the proposed framework based on the set temperature provided by the residents estimates the desired sensation of the occupants. The *optimizer* then looks for a thermostat temperature that keeps the *pmv* close to the midpoint (pmv_{ref}) of the estimated sensation level (Equation 3.2).

Though it is viable to monitor each factor which is required to compute *pmv*, the current thermal model (discussed in Section 3.3.2) of the framework computes *pmv* using temperature in the region (under consideration) as the ambient air temperature, wall temperature as the mean radiant temperature, and remaining as per the ASHRAE (The American Society of Heating, Refrigeration and Air-Conditioning Engineers) standards. We discuss *pmv* computations and value of each factor in detail in Section 3.3.3.

Comfort-Energy Trade-off

The AC settings differ across the users according to their cooling requirements. Some people prefer to optimize the AC usage to reduce their electricity bill, while others choose to cover themselves with thick blankets and set their AC thermostat to lowest possible temperature. The proposed framework allows the residents to mention their preference through CET, for any thermostat temperature.

$$score = \alpha \times \widehat{savings} + (1 - \alpha) \times (1 - \widehat{discomfort}) \tag{3.3}$$

The optimization function (Equation 3.3) maximizes *score* to get the optimal value of set temperature for the thermostat. To bring both (*savings* and *discomfort*) on the same scale ([0, 1]), $\widehat{savings}$ and $\widehat{discomfort}$ are the metric values normalized with the maximum savings and discomfort, for the particular AC usage. We subtract the later half from 1 to make the overall problem a maximization problem. α is the weight that ranges from 0 (min *discomfort*) to 1 (max *savings*) and decided based on the value set by the residents on CET on the portable+ thermostat. While finding the optimal thermostat temperature, framework ensures that the following constraints are satisfied:

1. $E_{new} \leq E_{old}$ i.e. in worst case, *optimizer* should return the actual thermostat temperature set by the users.

2. $pmv_{ll} \leq pmv_{new} \leq pmv_{ul}$ i.e. any change in thermostat temperature should keep occupants' comfort (i.e. pmv_{new}) within the same sensitivity level as of pmv_{old}.

Since occupants are free to use the AC as per their convenience, the *optimizer* considers pmv_{old} as their comfortable sensitivity level. Therefore, the second constraint ensures that user comfort (after changing the thermostat temperature) remains within the lower (pmv_{ll}) and upper (pmv_{ul}) limits of the actual sensitivity level (Table 3.1).

3.3.2 Tuner

Different parts (or regions) of a room perceive different temperature. The equation between region-wise temperature is non-linear; therefore, the framework needs a simplified model to emulate the thermal response of a building by capturing its essential behaviour. Studies show that analysis based on 'lumping' entire room into a small number of parameters can help develop smart, powerful, and energy-efficient solutions [142, 51, 52]. In contrast to extensive instrumentation of a room, embedding thermal model in the framework also ensures scalability of the approach to any environment. We now present one such thermal model to demonstrate the approach. Though in the specified model, we divide the room into three regions, Appendix A presents a generic formulation for an n-regions space.

Thermal Model

In a controlled experiment, we deployed temperature sensor in three distinct regions (let's say R1, R2, R3) and noted an average temperature of $23°C$, $27°C$, and $34°C$ (during the AC usage) for a set temperature of $24°C$. Motivated by Lake Thermal Stratification [32] that divides the lake into 3 different thermal zones - top zone which receives sunlight, bottom zone which is always dark and cold, and a mixing zone in between wherein water from top and bottom zone mix due to currents; we also divide the room into 3 regions - Low Impact Region (lir), Moderate Impact Region (mir), and High Impact Region (hir). Here, hir corresponds to the area in proximity of the AC, thus facing direct and the maximal impact of the cold air coming from the AC. Further, mir is the region where occupants spend their significant time and often have an indirect effect of AC cooling, while lir primarily includes the corner spaces of the room. Instead of physical boundaries, these regions are considered to be separated by a thin layer of air having negligible thermal mass. The framework

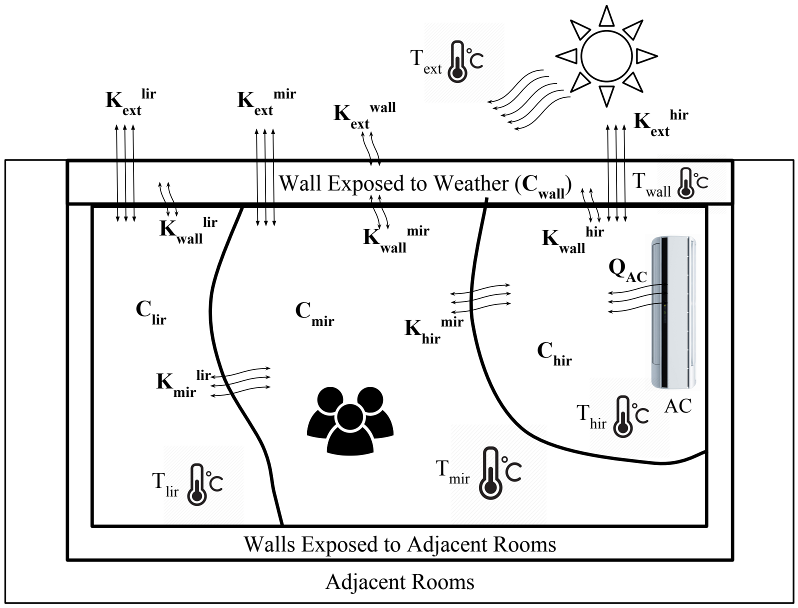


Figure 3-3: Pictorial representation of various thermal interactions considered in the thermal model of the proposed optimization framework

extends a 2^{nd} order thermal model [37] and targets to improve the user experience in *mir* for the demonstration purpose.

$$\begin{aligned}\frac{(T_{wall}^{(t+1)} - T_{wall}^{(t)}) \times C_{wall}}{\tau} &= K_{ext}^{wall} \times (T_{ext}^{(t)} - T_{wall}^{(t)}) + K_{wall}^{hir} \times (T_{hir}^{(t)} - T_{wall}^{(t)}) \\ &+ K_{wall}^{mir} \times (T_{mir}^{(t)} - T_{wall}^{(t)}) + K_{wall}^{lir} \times (T_{lir}^{(t)} - T_{wall}^{(t)})\end{aligned} \tag{3.4}$$

$$\begin{aligned}\frac{(T_{hir}^{(t+1)} - T_{hir}^{(t)}) \times C_{hir}}{\tau} &= K_{ext}^{hir} \times (T_{ext}^{(t)} - T_{hir}^{(t)}) + K_{wall}^{hir} \times (T_{wall}^{(t)} - T_{hir}^{(t)}) \\ &+ K_{hir}^{mir} \times (T_{mir}^{(t)} - T_{hir}^{(t)}) - Q_{AC} \times S_{AC} + \eta_{hir}\end{aligned} \tag{3.5}$$

$$\begin{aligned}\frac{(T_{mir}^{(t+1)} - T_{mir}^{(t)}) \times C_{mir}}{\tau} &= K_{ext}^{mir} \times (T_{ext}^{(t)} - T_{mir}^{(t)}) + K_{wall}^{mir} \times (T_{wall}^{(t)} - T_{mir}^{(t)}) \\ &+ K_{hir}^{mir} \times (T_{hir}^{(t)} - T_{mir}^{(t)}) + K_{mir}^{lir} \times (T_{lir}^{(t)} - T_{mir}^{(t)}) + \eta_{mir}\end{aligned} \tag{3.6}$$

$$\frac{(T_{lir}^{(t+1)} - T_{lir}^{(t)}) \times C_{lir}}{\tau} = K_{ext}^{lir} \times (T_{ext}^{(t)} - T_{lir}^{(t)}) + K_{wall}^{lir} \times (T_{wall}^{(t)} - T_{lir}^{(t)}) \\ + K_{mir}^{lir} \times (T_{mir}^{(t)} - T_{lir}^{(t)}) + \eta_{lir} \tag{3.7}$$

Equation 3.4 and Equation 3.5-3.7 describe the mathematical formulation for the heat lost/gained (due to indicated thermal interactions) by the wall (facing outside), and the three regions of the room (hir, mir, and lir), respectively, in a given time interval τ. During the time duration τ, we consider conductive heat transfers due to the difference in temperature of the region (under consideration) and the adjacent spaces such as weather conditions, wall (facing outside), neighbouring regions, and the AC as depicted in Figure 3-3. We assume negligible heat transfer through adjacent rooms within the home. T_{ext} takes care of the changing weather conditions and thermal noise (depicted by η_r, where $r \in \{hir, mir, lir\}$) in each region models the dynamic occupancy patterns and various other activities, such as frequent opening/closing of doors/windows in the room.

$$\theta = \{C_{wall}, C_{hir}, C_{mir}, C_{lir}, K_{wall}^{hir}, K_{wall}^{mir}, K_{wall}^{lir}, K_{ext}^{hir}, K_{ext}^{mir}, K_{ext}^{lir}, \\ K_{ext}^{wall}, K_{hir}^{mir}, K_{mir}^{lir}, \eta_{hir}, \eta_{mir}, \eta_{lir}, Q_{AC}\} \tag{3.8}$$

Table 3.2 lists all the symbols used in the thermal model. The thermal model consists of 17 parameters (Equation 3.8) where each parameter is physically significant in the real-world and should be adapted to a particular environment. Depending on the material of the wall, insulation quality, and various assumptions, it is feasible to compute theoretical values for the stated parameters, but that will make the approach unscalable across diverse type of rooms. Besides, with aging, each parameter of a thermal model can see noteworthy deviations in their values, e.g. efficiency of AC might degrade with time. Therefore, before predicting each AC usage, the *tuner* automatically adjusts each parameter (of the thermal model) based on the data gathered from a few recent historical AC usages. The *tuner* uses theoretical values from the literature [25, 55] to derive appropriate range for initializing the parameters. Initialization based on theoretical values ensures inclusion of domain knowledge while performing statistical optimization.

Symbol	Description	Unit
τ	Sampling Interval	s
r	Thermal region $\in \{hir, mir, lir\}$	–
η_r	Thermal noise in region r	–
Q_{AC}	Cooling capacity of AC	kW
$T_r^{(t)}$	Temperature in region r at time instance t	$^{\circ}C$
$T_{ext}^{(t)}$	External temperature at time instance t	$^{\circ}C$
$T_{wall}^{(t)}$	Temperature of wall (facing outside) at time instance t	$^{\circ}C$
C_r	Thermal capacity of region r	kJ/K
C_{wall}	Thermal capacity of wall (facing outside)	kJ/K
K_{wall}^r	Heat transfer coefficient between wall (facing outside) and region r	kW/K
K_{ext}^r	Heat transfer coefficient between external environment and region r	kW/K
K_{ext}^{wall}	Heat transfer coefficient between wall (facing outside) and weather	kW/K
K_{hir}^{mir}	Heat transfer coefficient between *hir* and *mir*	kW/K
K_{mir}^{lir}	Heat transfer coefficient between *mir* and *lir*	kW/K
$S_{AC}^{(t)}$	AC compressor state (*on*/*off*) at time instance t	–

Table 3.2: List of symbols used in the proposed thermal model

Tuned Thermal Model

The power of *Grey Box Modelling* lies in its ability to tune the model parameters (with time) using the sensory data. As tuning involves the actual (temperature) data collected from the room, the adjusted parameters are an approximate representation of the thermal behaviour of the room. By using the data from recent n AC usages, *tuner* keeps updating the parameters (before each AC usage) to accommodate changes in weather conditions, user activities (in the form of thermal noise), and various other dynamics of the room. The optimal parameter vector θ^* minimizes the total squared error between estimated ($T_{hir}^{(t)}$) and observed ($\bar{T}_{hir}^{(t)}$) temperature in *hir* to tune the model parameters (Equation 3.9).

$$\theta^* = \arg\min_{\theta} \sum_{t=t^{on}}^{t^{off}} (T_{hir}^{(t)} - \bar{T}_{hir}^{(t)})^2 \tag{3.9}$$

t^{on} and t^{off} are the time instances when user turned the AC on and off, respectively. The objective function is convex and quadratic with added constraints of $\{C_r, C_{wall}, K_{ext}^r, K_{wall}^r, K_{ext}^{wall}, K_{hir}^{mir}, K_{mir}^{lir}, Q_{AC}\} \geq 0,\ \forall r \in \{hir, mir, lir\}$. We solve this constrained convex optimization problem using the SNOPT solver in AMPL, a modelling language for mathematical programming. In an attempt to avoid local optimum, we initialize each parameter with twenty random instances of parameters, limits (lower and upper) of whom are derived from the domain knowledge. Empirical evaluation over the collected dataset shows that the data from $n = 4$ usages (each of length 4 hours) can tune the (discussed) thermal model with a decent accuracy.

The *optimizer* utilizes the tuned thermal model to predict region-wise temperature and AC compressor state at each time instance. While region-wise temperature computes the user comfort, the predicted compressor state (S_{AC}) calculates the AC energy consumption (Equation 3.10).

$$Energy\ (kWh) = P_r \times \frac{\sum_{t=t^{on}}^{t^{off}} S_{AC}^{(t)}}{3600} \times \tau \tag{3.10}$$

P_r is the rated power consumption of the AC, and $S_{AC}^{(t)}$ denotes the AC compressor state at any time instance t. Here, t^{on} and t^{off} are the time instances when user turned on and off the AC, respectively. $1/\tau$ denotes the sampling rate in hertz (Hz) used for the data collection.

3.3.3 Putting Things Together - A Case Study

For a random day from one of the rooms of in-situ deployment, we now illustrate the functioning of the proposed framework in finding optimal set-temperature to control the AC. For the scenario under consideration, the thermostat is present in *hir* and set at $21°C$. The sensors are collecting data every 2 minutes, thus setting $\tau = 600$ seconds to compute the AC energy consumption. Besides, we use standard ASHRAE values for other parameters to calculate *pmv* at any time instance during an AC usage. The occupants never used the fans; thus we use standard value (for a static environment) for the velocity of air i.e. 0.5 m/s. We set relative humidity to 50% as the thermal model (Section 3.3.2) only predicts the room temperature. Though it is feasible to predict humidity by using a further complex thermal model, we limit the scope of this study to temperature only and leave dynamic humidity exploration as part of our future work. We conducted both the studies during summers; thus we set clothing level to be 0.5 *clo*. Also, given that study included only bedrooms and living rooms, occupants were either sitting or sleeping during the AC usages

$\mathbf{T_{set}}$	**pmv**	*comfort*			**energy**	*savings*		*score*		
		discomfort	$\widehat{discomfort}$	$\widehat{comfort}$		*savings*	$\widehat{savings}$	$\alpha = 0.2$	0.5	0.8
21	-0.13				6.24					
22	$+0.08$	0.21	0.25	*0.75*	5.46	0.78	*0.24*	**0.65**	0.50	0.34
23	$+0.28$	0.41	0.49	*0.51*	4.68	1.56	*0.48*	0.51	**0.50**	0.49
24	$+0.49$	0.62	0.74	*0.26*	3.83	2.41	*0.74*	0.36	0.50	**0.65**
25	$+0.71$	0.84	1.00	*0.00*	2.29	3.25	*1.00*	*inv*	*inv*	*inv*

Table 3.3: Illustration of optimisation framework through an AC usage where AC is currently set at $21°C$

(also confirmed by the residents). Thus, we set metabolic rate equal to 1.0 *met*, defined for occupants sitting in the room. The value of τ and other parameters (to compute *pmv*) will be valid even for the subsequent analysis of the framework in Section 3.4.

Table 3.3 shows the analysis to find optimal set temperature for a given value of set temperature and CET (or α). When set at $21°C$, the value of $pmv = -0.13$ indicates that residents (in *mir*) will feel `Neutral` during the AC usage and consume $6.24kWh$ energy. Based on this information, the *optimizer* of the framework looks for a set temperature that maximizes the *score* and keeps the sensation level at `Neutral` (in *mir*).

Maximize Comfort ($\alpha = 0.2$)

When α (or CET) is set to 0.2, the *optimizer* changes the thermostat temperature to $22°C$ for maximizing the user comfort during the AC usage. By increasing the set temperature by $1°C$ ($21°C \rightarrow 22°C$), residents will experience better comfort as it brings the *pmv* closer to the mean value of `Neutral` sensation ($pmv = 0.0$). Moreover, running the AC at $22°C$ also reduces the AC energy consumption by 0.78 kWh (i.e. 13%) while improving the user experience.

Comfortable Energy Savings ($\alpha = 0.5$)

CET = 0.5 implies that users are ready to compromise a little on their comfort but also looking for potential energy savings for the AC. Though *scores* remain same by increasing the thermostat temperature to $22°C$, $23°C$, or $24°C$, the *optimizer* set AC at $23°C$. By doing so, energy savings increases to 25% (with a slight drop in comfort) while maintaining an equilibrium between both comfort and the savings.

Maximize Savings ($\alpha = 0.8$)

Next, we demonstrate the potential of the framework in maximizing the AC energy savings. Though setting the thermostat to $25°C$ will result in maximal savings, but it violates the second constraint of the optimization function (of the *optimizer*) i.e. pmv_{new} should remain within the same sensitivity level (`Neutral`). As $T_{set} = 25°C$ is an invalid temperature (depicted by *inv*) for the thermostat, framework operates the AC at $24°C$ to achieve the desired CET. Raising the set temperature from $21°C$ to $24°C$ increases the energy savings to 39% i.e. three times of the energy savings obtained by setting the CET = 0.2.

The framework in a portable+ thermostat dynamically decides the set temperature, as explained above, in fixed time windows by looking at the weather prediction fetched from a weather station. This will ensure that while the user provides a static reference set temperature and a personal desire to optimize between comfort and energy savings, the framework automatically keeps changing the set temperature in accordance with user inputs. Next, we assess the accuracy of our tuned thermal model in simulating the temperature (across the regions) followed by the impact analysis (of both portable and portable+ thermostats) on occupants' comfort and the AC energy savings.

3.4 Evaluation

In real world, portable+ thermostat senses temperature from a single location in the room. To emulate the actual condition in different parts of the room, proposed framework uses a lumped parameter thermal model; tuned using the temperature data gathered from the sensing point. For a comprehensive analysis, we carried out a controlled experiment in two rooms (bedroom and living room of a home) over two weeks and collected data from three different locations within the room, each representing a different thermal zone. The controlled experiments helped us in examining various hypothesis. For a broad and diverse evaluation of portable+, we conducted an in-situ deployment across five different homes over a duration of three months and monitored everyday AC usage by the occupants. For both the dataset, using Leave-p-out Cross Validation (Lpo-CV), we assess tuned parameters (of lumped thermal model) in estimating the temperature across different regions. Later, using the tuned models, we analyze the impact of portable+ thermostats on occupants' comfort and AC energy consumption. We begin this section by briefing our experimental setup used for the data collection across controlled and in-situ deployments.

3.4.1 Data Collection

For both controlled experiments and in-situ deployment, we connected a low-cost temperature sensor (Maxim DS18B20[1]) to a single-board computer (Raspberry Pi). For ground truth, we also monitored power consumption of the AC across all the rooms. We pushed the collected data to a cloud server.

Controlled Experiment

We deployed three temperature sensors in a bedroom (represented as CEBR herein) and a living room (designated as CELR herein) of a home for a duration of two weeks. While the bedroom was small in size (floor area of 13.2 m^2) having a 1.5 tonne AC, the living room was comparatively larger in size (floor area of 20.06 m^2) containing a 2 tonne AC. To decide the regions for placement of temperature sensors, we asked a few volunteers to feel the cooling from the AC at distinct points in the room. For each room, we combined volunteers' feedback with temperature measurements at those locations to deploy three temperature sensor, one in each region of the thermal model - high impact region (*hir*), moderate impact region (*mir*), and low impact region (*lir*). To monitor weather conditions, we kept one temperature sensor in the balcony of the home. During the study, we set the thermostat to different temperatures and experienced diverse weather conditions.

In-Situ Deployment

To further validate our hypothesis in real-world, we also conducted an in-situ deployment in five different homes (for three months) which consist of three bedrooms (ISR1, ISR2, ISR5) and two living rooms (ISR3, ISR4). One of the biggest challenges for in-situ deployment was to ensure occupants' comfort in daily routine. Therefore, during the data collection, we asked residents to follow their usual routine and only confirm their thermostat temperature for each AC usage. Further, to avoid inconvenience to the occupants, we monitored temperature only in *hir* (nearby AC) and AC power consumption (for ground truth). To monitor the weather conditions, we used an API of third party weather stations (Weather World Online [164]) which reports temperature as recorded by their nearest station every half an

[1] https://www.maximintegrated.com/en/products/analog/sensors-and-sensor-interface/DS18B20.html

hour. We understand that not monitoring the temperature in other regions (for tuning the thermal model) and also using a proxy (in the form of a weather server) to monitor weather conditions can develop a bias during the impact analysis of data from in-situ deployment. However, this also represents a genuinely low-cost and scalable transition to portable+ that monitors the temperature at only one location in the room. We use data from controlled experiments to understand the accuracy for estimating temperature across the regions when data from only one region is available.

$$error = \sqrt{\mathrm{E}[(T_r^{(t)} - \bar{T}_r^{(t)})^2]},\ \forall r \in \{hir, mir, lir\} \tag{3.11}$$

3.4.2 Validation of Tuned Parameters

Tuned thermal model from our optimization framework is a pre-requisite to analyze the impact of set temperature and placement of portable+ thermostat on occupants' comfort and AC energy consumption. We use root mean square error (RMSE) in estimated temperature across the different regions to evaluate the accuracy of proposed framework in tuning the parameters of the thermal model (Equation 3.11), where E[.] denotes the expectation of squared error.

Leave p-out Cross Validation (Lpo-CV)

We term an AC usage as the duration from the time user switched on the AC till they switched it off. In Leave-p-out Cross Validation (Lpo-CV), we train the model from $n - p$ AC usages and test the tuned parameters over remaining p usages. Here n is the total number of such AC usages observed during the data collection and p denotes the size of test dataset. For example, if there are 10 AC usages in a home and p equals 30%, then we train our model using first seven usages and test the tuned parameters over remaining three usages. It continues in a circular fashion until it learns the model on all possible combinations of $n - p$ consecutive AC usages. In our case, we divided the total number of usages (n) into 70% training and 30% testing dataset for all the rooms, from both the deployments.

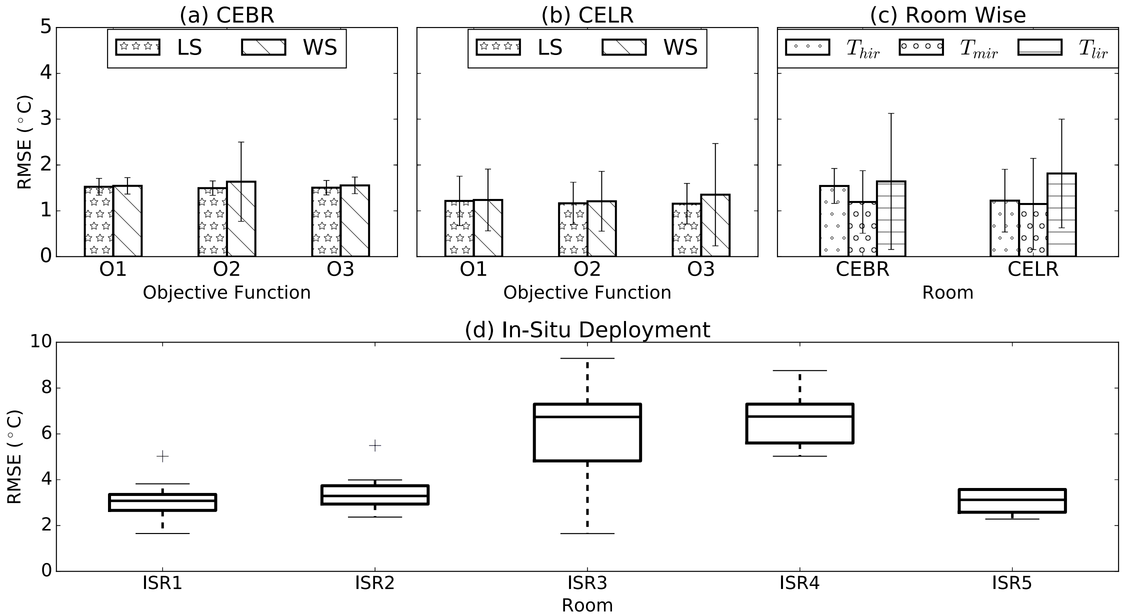


Figure 3-4: Leave p-out Cross Validation: Using data from WS instead of LS had little impact on RMSE when objective function varied from O1 to O3 but introduced uncertainty in estimating T_{hir} across both (a) Bedroom (CEBR) and (2) Living Room (CELR) from the controlled experiment. (c) Tuned thermal model (using O1 as the objective function), estimated the temperature across three regions with reasonable accuracy while using WS (to monitor weather) for both the rooms in controlled experiment. (d) Accuracy in estimating T_{hir} for both bedrooms (ISR1, ISR2, ISR5) and living rooms (ISR3, ISR4) from the in-situ deployment.

$$\begin{aligned}
&\textbf{O1:} \quad \sum_{t=t^{on}}^{t^{off}} (T_{hir}^{(t)} - \bar{T}_{hir}^{(t)})^2, \\
&\textbf{O2:} \quad \sum_{r\in\{hir,mir\}} \sum_{t=t^{on}}^{t^{off}} (T_r^{(t)} - \bar{T}_r^{(t)})^2, \\
&\textbf{O3:} \quad \sum_{r\in\{hir,mir,lir\}} \sum_{t=t^{on}}^{t^{off}} (T_r^{(t)} - \bar{T}_r^{(t)})^2
\end{aligned} \tag{3.12}$$

We begin the validation with analyzing the influence of 1. choice of objective function i.e. O1, O2, and O3 (Equation 3.12), and 2. monitoring weather conditions using a local sensor (LS) in place of an API from third party weather station (WS).

As we vary the objective function from O1 to O3, we need additional temperature information sensed across different regions. Figure 3-4a and 3-4b compare the RMSE (in estimating T_{hir}) for CEBR and CELR, respectively, across three different objective functions when using WS and LS to monitor the outside weather conditions. We observe that including more information (O1→O2→O3) in the objective function while making a little improvement in estimating T_{hir} across both the rooms, reduced standard deviation by $0.03°C$. We further noted that temperature from WS deviated (absolutely) by $6°C$ ($\sigma = 2.8°C$) from the temperature sensed by LS, resulting in some uncertainty in estimating T_{hir} (i.e. increase in standard deviation).

Next, we validate the tuned thermal model (using Lpo-CV) on the data from the controlled experiment (in estimating T_{hir}, T_{mir}, and T_{lir}) while using O1 as the objective function and weather data from World Weather Online [164]. Figure 3-4c shows the mean *error* in estimating temperature across the three regions while using the tuned thermal model. We noticed a mean RMSE of $1.54°C$ ($\sigma = 0.38°C$), $1.19°C$ ($\sigma = 0.68°C$), and $1.64°C$ ($\sigma = 1.48°C$) in estimating T_{hir}, T_{mir} and T_{lir}, respectively, in the bedroom. Even though the living room was larger than the bedroom, and connected to a corridor (possibly resulting in some thermal leakage), tuned thermal model estimated T_{hir}, T_{mir} and T_{lir} with an RMSE of $1.22°C$ ($\sigma = 0.68°C$), $1.15°C$ ($\sigma = 0.99°C$), and $1.81°C$ ($\sigma = 1.18°C$). These errors align with those presented in the existing literature [37] and show that thermal model is reasonable enough to estimate temperature across the regions, specifically for our application. While reducing objective function (to O1) and using WS (to monitor weather) brings a little uncertainty in the tuned parameters, they are critical in ensuring that the proposed approach is scalable and less intrusive for the occupants. To be closer to a scalable scenario and analyze its effectiveness, we deployed only one (temperature) sensor in *hir* and used (third-party) weather server for the validation of tuned thermal models from in-situ deployment. We observed that RMSE increased to $3.15°C$ ($\sigma = 0.77°C$) for the bedrooms (ISR1, ISR2, ISR5) and $7.4°C$ ($\sigma = 4.6°C$) for the living rooms (ISR3, ISR4), as shown in Figure 3-4d.

Where Things Went Wrong?

Primary reason for the increase in errors and standard deviation for in-situ deployment is the increase in thermal noise due to numerous uncontrolled and dynamics activities of the occupants present in the room. In Section 3.3.2 we outlined a thermal model that depicts the dynamic activities of the occupants as a noise in each region (η_r) which is constant at any

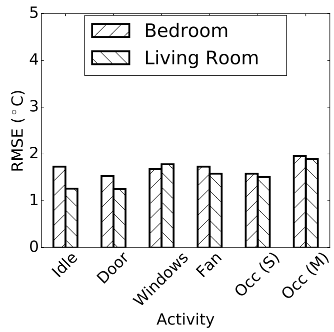


Figure 3-5: Constant noise (through various sources) results in insignificant deviation from the Idle scenario.

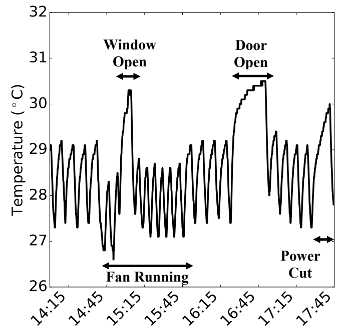


Figure 3-6: A labelled usage depicting dynamic and random noise from multiple sources at different time instances during the AC usage.

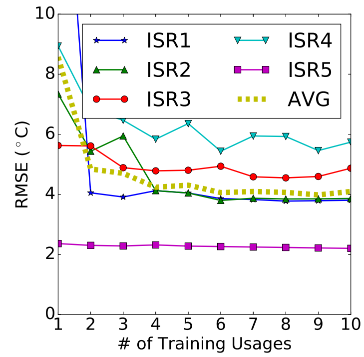


Figure 3-7: Simulation error stabilizes beyond a training dataset containing four usages.

time instance t. Thus, a source generating constant (thermal) noise will make insignificant effect on simulation accuracy, although dynamic and random activities during the AC usage can increase the simulation errors significantly. For instance, we carried out an experiment (in a controlled setting) to understand the impact of five activities in simulating the room temperature - 1. keeping doors open, 2. keeping windows open, 3. running the (ceiling) fan, 4. single occupant working, and 5. multiple occupants working (Figure 3-5). During the experiment, we carried out each activity for an hour under similar weather conditions and noted slight deviations in RMSE across the activities. However, for in-situ deployment, users were allowed to follow their daily schedule which resulted in (a few) AC usages that incorporated (random) noise from multiple of such sources. To illustrate, Figure 3-6 presents one such usage annotated with multiple activities that arose at distinct time intervals while using the AC. As the (current) model is unaware of changing fan/door/window status (on/off), it will predict assuming an average noise through these sources; thus inflating the errors. In bedroom, people used their AC mostly during the night while sleeping; thus, the influence of dynamic activities is less when compared to the living rooms. Further, living rooms also include many shared spaces such as corridor impacting the thermal leakage from the room and raising the average error during cross-validation. While a higher order thermal model can reduce these errors, but such a model will require more information; thus additional instrumentation of the home.

Parameter	**CEBR**	**CELR**	**ISR1**	**ISR2**	**ISR3**	**ISR4**	**ISR5**
Q_{AC}, kW	6	7.9	5.2	4.9	7.6	8	5.32
$C_{hir}, kJ/K$	206	475	311	516	787	862	439
$C_{mir}, kJ/K$	148	160	148	348	495	337	271
$C_{lir}, kJ/K$	90	110	50	150	275	431	112
$C_{wall}, kJ/K$	653	153	453	153	833	670	492
$K_{wall}^{hir}, kW/K$	0.34	0.18	0.22	0.12	0.83	0.35	0.16
$K_{wall}^{mir}, kW/K$	0.06	0.23	0.11	0.01	0.51	0.24	0.07
$K_{wall}^{lir}, kW/K$	0.03	0.35	0.09	0.03	0.91	0.41	0.25
$K_{ext}^{hir}, kW/K$	0.09	0.32	0.13	0.13	1.59	0.23	0.28
$K_{ext}^{mir}, kW/K$	0.24	0.13	0.02	0.02	0.84	0.44	0.07
$K_{ext}^{lir}, kW/K$	0.4	0.14	0.07	0.03	0.75	0.95	0.15
$K_{hir}^{mir}, kW/K$	0.29	1.1	0.4	0.88	0.82	1.12	0.99
$K_{mir}^{lir}, kW/K$	0.37	0.6	0.23	0.43	0.85	0.37	0.87
$K_{ext}^{wall}, kW/K$	0.15	0.15	0.17	0.09	0.72	0.36	0.16
η_{hir}	0.15	0.34	0.26	0.71	3.1	1.38	0.47
η_{mir}	0.48	0.18	0.33	0.63	3.76	3.09	0.59
η_{lir}	0.35	0.48	0.48	0.58	5.85	4.05	1.2

Table 3.4: Tuned parameters of the thermal model for each rooms from both controlled and in-situ deployments

Sensitivity Analysis

The tuner learns from historical n usages to predict and analyze $(n+1)^{th}$ usage for optimal set temperature. To select the value of n, we measured the average simulation accuracy for all the rooms (from in-situ deployment) while varying the size of training data from one to ten AC usages (Figure 3-7). We observed that simulation accuracy averaged over all the rooms (dotted line) stabilizes at four usages and shows insignificant improvement beyond that. Thus, in real-world, we used four recent historical usages to learn our thermal model to predict for the next usage. However, to set up the testbed (for Impact Analysis), we tuned the thermal model for each room using the complete dataset i.e. including all AC usages for a particular room. As we are bound to respect the privacy concerns of the volunteers, the list of learned parameters ensures that results are reproducible even though we are unable to share the datasets. Table 3.4 summarizes the learned model parameters

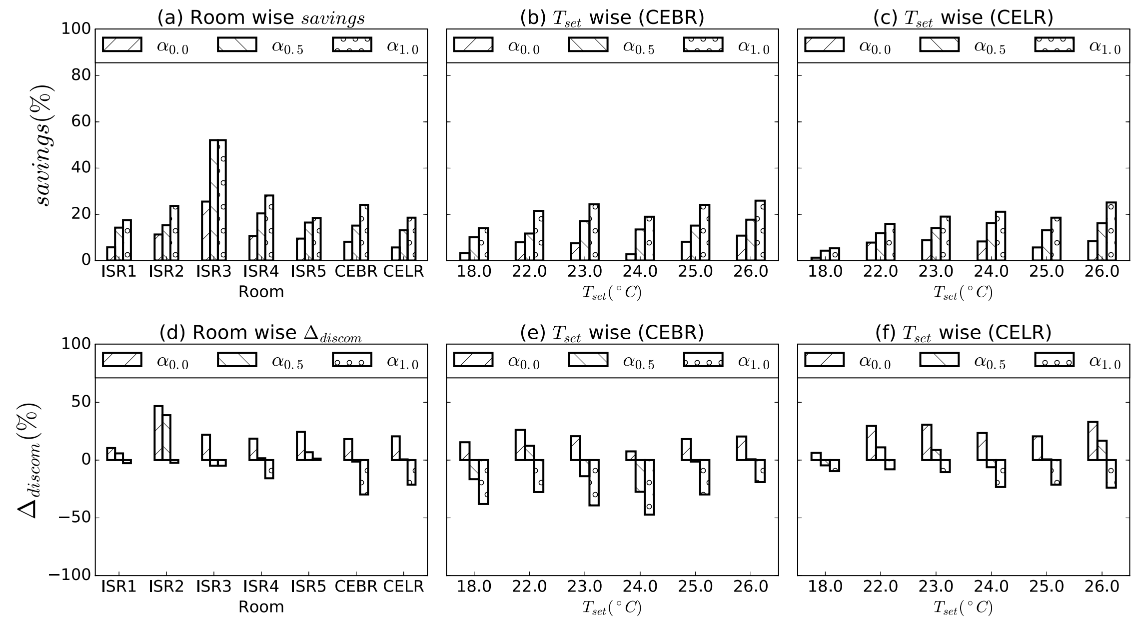


Figure 3-8: Impact Analysis: (a) Energy savings increase when we move our optimization goal from minimizing discomfort to maximize energy savings ($\alpha = 0.0 \rightarrow 1.0$) across all the rooms. The portable+ thermostat is consistent with achieving these energy savings across different set temperature for both type of rooms; (b) Bedroom and (c) Living Room from the controlled experiment. (d) When compared with portable thermostat, portable+ thermostat increases comfort (by reducing *discomfort*) across all the room. Similar to energy savings, % change in discomfort (Δ_{discom}), when moved from portable to the portable+ thermostat, has a weak correlation with T_{set} across both (e) Bedroom and (f) Living Room from the controlled experiment.

for all the rooms. Using the testbed, we next present our impact analysis on portable+ thermostats to quantify their impact on occupants' comfort and AC energy consumption, in terms of *savings* and *discomfort*.

3.4.3 Impact Analysis

Optimization framework requires tuned thermal model (of the room), external temperature, and T_{hir} at the time when AC was turned on (t^{on}). Therefore, to carry out the impact analysis, we collected historical weather data for the period when we conducted both the studies (controlled and in-situ). From our (in-situ) dataset, we observed that much (45%) of the AC usages (across all the rooms) was between 12 AM and 4 AM. Therefore, for the period of controlled study (2 weeks) and in-situ deployment (3 months), using the tuned thermal model of each room, we estimated the temperature (in *hir*, *mir*, and *lir*) between

12 AM to 4 AM while considering a diverse set of thermostat temperature. During the simulation, we set the thermostat at $25°C$ (for the controlled experiment) and frequently used set temperature (by the residents) for in-situ deployment i.e. $27°C$ (ISR1), $24°C$ (ISR2), $26°C$ (ISR3), $25°C$ (ISR4), and $21°C$ (ISR5). Based on the estimated temperature measurements (across all regions) and AC compressor state, we assess the potential of portable+ thermostats in attaining maximum savings while reducing occupants' discomfort.

$$\Delta_{discom}(\%) = \frac{discomfort_{pot} - discomfort_{pot+}}{discomfort_{pot}} \tag{3.13}$$

We use Δ_{discom} (Equation 3.13) to depict percentage change in *discomfort* when shifting from portable to a portable+ thermostat. Here, $discomfort_{pot+}$ indicates occupants' discomfort observed due to portable+ thermostats for the optimal settings, while $discomfort_{pot}$ depicts the discomfort incurred by using portable thermostats. A higher value of Δ_{discom} implies better user comfort (or reduced discomfort) if switched from portable to a portable+ thermostat to optimize AC energy consumption.

Figure 3-8a and Figure 3-8d shows the average *savings* and Δ_{discom} (when compared to the portable thermostat) across different rooms while varying the α in Equation 3.3. As we move our optimization to maximize energy savings from maximizing occupants' comfort ($\alpha = 0 \rightarrow 1$), we noted a significant increase in energy savings across all the rooms, with maximum energy savings of 50% achieved for ISR3. On the other end of the spectrum, with $\alpha = 0$, significant improvement in comfort was achieved (as expected), with the maximum change in comfort being 50% attained for ISR2. Further, in controlled experiments, we collected data at different set temperatures to understand the savings and comfort potential across different user settings. We observed that improvement in comfort (Figure 3-8e, 3-8f) along with increased energy savings (Figure 3-8b, 3-8c) is achievable across a range of set temperatures. Before moving further, we remind readers that across all the optimization (that we discuss now), we ensure that the constraints (1) $E_{new} \leq E_{old}$, and (2) $pmv_{ll} \leq pmv_{new} \leq pmv_{ul}$ are met, as discussed in the previous section.

Maximize Occupants' Comfort ($\alpha = 0.0$)

The job of the AC is to maintain comfortable room temperature for the occupants. Thus, we begin our analysis with maximizing comfort (or minimizing discomfort) for the occupants.

With $\alpha = 0.0$, our analysis shows that portable+ thermostats reduce occupants' discomfort by 23% ($\sigma = 11\%$) (Figure 3-8d) while still saving 11% ($\sigma = 7\%$) (Figure 3-8a) energy (across all the rooms), as compared to portable thermostat. These improvements over portable thermostat are given the fact that portable+ thermostat seek to attain peak (from Figure 3-1) for *Alice* which ensures minimal discomfort (in *mir*) while optimizing for AC energy consumption.

Comfortable Energy Savings ($\alpha = 0.5$)

The power of portable+ thermostat lies in its potential to attain comfortable energy savings i.e. optimize AC energy consumption while maintaining occupants' comfort. As we move α to 0.5, *savings* increased to 21% ($\sigma = 14\%$) while reducing discomfort by 7% ($\sigma = 15\%$). The high deviations (in Δ_{discom}) is given the fact that human comfort may vary significantly as per the weather conditions. Thus, with a change in climate, comfort requirements of *Alice* (curve in Figure 3-1) may shift either towards left or right. However, Alice following a daily routine will operate her AC at $18°C$ which can fall towards left on the much lower side and also close to the peak value. Depending on where user comfort resides (on the curve) for a given set temperature and weather conditions, likelihood of savings and comfort may vary significantly in magnitude. Such variations across diverse climate further advocate the necessity of portable+ thermostat for residential ACs.

Maximize Energy Savings ($\alpha = 1.0$)

Next, we move our optimization goal towards maximizing the energy savings. We observed that now portable+ thermostat is no better than a portable thermostat in reducing the occupants' discomfort while optimizing the AC energy consumption. Discomfort introduced by portable+ thermostat increased by 10% ($\sigma = 12\%$) when compared with portable thermostat, however, energy savings also increased to 26% ($\sigma = 12\%$).

To conclude, referring to *savings* for $18°C$ (from Figure 3-8b), our analysis indicates that portable+ thermostat empowers *Alice* in saving 10% energy while keeping her comfort intact ($\alpha = 0.5$) while using the set temperature from her daily routine. With slight discomfort, savings (from AC) can raise up to 14% by switching to power saving mode of the thermostat. Moreover, to do so, all she need to do is set the thermostat at any temperature and mention her requirements (peak comfort/maximal savings) at CET, and portable+

thermostat will take care of everything else. By using the portable+ thermostats, residents (like *Alice*) can attain a balance between their comfort and energy savings. Depending on their preference, they can save up to 26% energy (when set on power saving mode) and reduce their discomfort by 23% while aiming for the peak comfort. Realized savings can be significantly improved by further optimizing the control logic of portable+ thermostat through various applications, which we discuss in the next section.

3.5 Challenges And Discussion

In this work, we proposed portable+ thermostat that uses a 2^{nd} order thermal model to make AC operations energy-efficient and comfortable for the users. For the detailed thermal model, we depicted numerous sources affecting the room temperature (such as doors, windows, fans, among other) as a lumped thermal noise that ensued lower accuracy for a few usages from in-situ deployment. However, thermal model is an input to the framework, and installing motion sensors (to detect occupancy) and relay sensors (to determine door/window status) will allow the system to utilize thermal models with low errors. Having said that, the framework is open to refinements for any n-region room (Appendix A), places where residents possess inverter ACs (Appendix B), or any other scenario; therefore, model assumptions should not be confused with the limitations.

Next, though we propose portable+ thermostat, this study neither deals with the thermostat design nor its mobile application. Studies show that human-centric prototypes of such devices (and applications) can significantly influence the outcomes; thus we keep that thread open for the concerned community. Further, the optimization framework comprises of two components - (1) *optimizer*, and (2) *tuner*. The optimization function (of *optimizer*) aim to maximize the weighted average of *savings* and *discomfort*. The stated function ensured 26% savings for power-saving mode and enhanced the user experience by 23% in an attempt to attain peak comfort for the residents. Though the wide adoption across the community establishes the reason behind the choice of metrics (such as PMV), and functions (for *savings*, *discomfort*, and *score*) in the framework; they are the nut and bolts of the proposed system. For instance, in current implementation, we only consider temperature in *hir* and wall temperature to compute *pmv*; however, the formulation includes other parameters also, such as air velocity, relative humidity, metabolic rate of the occupants, and clothing insulation. While we can sense these parameters, directly or indirectly, to enrich the model, the current implementation is limited and assumes standard ASHRAE values for the other parameters. Since tightening them might boost the performance of portable+ thermostat, we encourage the community to study their variants towards enhancing the proposed framework.

The portable+ thermostat is capable enough to work for any window AC in any geography of the world, but we analyzed its effectiveness through a dataset collected from the homes of a particular area in India. Climate, users' attitude (towards energy savings), and many other factors differ significantly across the geographies. Though the shown numbers are an indication of better comfort along with notable energy savings, there can be considerable discrepancy across (and within) the countries. Therefore, a real-world implementation is critical to understand its effectiveness in achieving the desired goals.

Chapter 4

Influence of Prediction Errors

I am going to make a prediction - it could go either way.

– Ron Atkinson, Britain's Best-Known Football Pundit

4.1 Introduction

In the literature, several studies proposed different control strategies which broadly fall into two categories: *reactive* [60] and *predictive* [79, 123]. Typically, in a reactive controller, air-conditioners respond to measured occupancy in a zone. Here, occupancy is measured using motion, CO_2 sensors, or by monitoring building's WiFi infrastructure [159]. Since buildings typically take some time to respond to control inputs, better performance can be obtained using predictive control strategy where the controller selects the optimal trajectory of set points for a finite time horizon [76]. Of the predictive control techniques, perhaps the best-known approach is Model Predictive Control (MPC) [53].

In a typical MPC, a known building thermal model estimates the future system state using forecasts of model inputs, such as building occupancy and outside air temperature. However, the effectiveness of this approach depends on the accuracy of the predictions. As prediction accuracy deteriorates, MPC performance - in terms of occupant comfort and building energy use - degrades and may get even worse than conventional techniques. In recent work, Oldewurtel et al. [122] extensively studied the influence of errors in weather forecast on HVAC energy consumption and occupants' comfort and quantified the impact of mispredictions. However, the work neither addressed errors in occupancy prediction nor studied the ways to mitigate the influence of prediction errors.

In this chapter, we address this gap. While using the occupancy data from a commercial building, we study the influence of occupancy errors on MPC performance using a custom-built building simulator. We also model and analyze the impact of personal environmental control system (PEC) in the presence of prediction errors. A PEC could be an off-the-shelf desktop fan or a heater to provide individual thermal comfort [22]. We find that PEC when used with model predictive control, can reduce both - the variability in energy consumption and the occupants' discomfort.

Our contributions are as follows:

1. We present the design and development of a building thermal simulator that models conventional schedule-based, reactive occupancy-based, and predictive MPC-based HVAC controllers.

2. We extend the MPC-based control strategy proposed by Kalaimani et al. [85] and allow PEC to react between any two consecutive states of the system.

3. We quantify the impact of occupancy prediction errors on two MPC-based control strategies - with and without PEC. For analysis, we use occupancy data from forty-five volunteers over three months and simulations of a test building in both heating and cooling seasons.

4. It is important that occupancy forecast errors are realistic; thus, we propose a method to systematically introduce realistic occupancy errors into MPC predictions using real-world occupancy data.

The rest of the chapter is organized as follows. Section 4.2 discusses the literature and studies conducted in the past. In Section 4.3, we outline the control strategies studied in the chapter. In Section 4.4, we present the detailed architecture and design of the thermal simulator followed by detailed analysis in Section 4.5. In Section 4.6, we discuss several limitations and possible future directions of the study and conclude the chapter.

4.2 Related Work

4.2.1 Central HVAC Controllers

In the past, researchers have extensively studied the optimization of HVAC controllers to minimize the aggregate energy consumption and maximize user comfort [73, 77]. Agarwal et al. [3] studied aggressive duty cycling of HVAC based on occupancy patterns within the building. Lu et al. [107] proposed a smart thermostat to automate HVAC control by sensing occupancy and sleeping patterns in residential buildings. The occupancy-based control allows buildings to operate outside of comfort regimes when unoccupied, thus reducing energy usage [44]. Henceforth, several other studies also explored the use of occupancy information to optimize the HVAC energy operations [43, 11, 44, 118, 8, 72, 94, 96, 146, 165]. However, centralized HVAC controllers divide a building into thermal zones comprising of private and shared spaces. Within each zone, these control strategies maintain ASHRAE standard while assuming each zone as either occupied or unoccupied; thus, ignoring individual comfort requirements.

4.2.2 Personal Environmental Control

For personalized comfort, studies proposed to use personal environmental control systems (PECs), especially in shared spaces [22, 18, 171, 52, 51, 136]. Unlike conventional centrally-controlled HVAC system, where people share the same set point temperature [35, 74], PEC systems can meet the comfort requirements of all occupants, albeit at the cost of additional energy expenditure. Kalaimani et al. [85] merged PEC with model predictive control to further minimize the HVAC energy consumption and maximize the user comfort.

Though advanced predictive control strategies (such as MPC) have the potential to optimize HVAC operations significantly, none of the studies mentioned above quantify the influence of the prediction errors on the energy consumption of HVAC and on the occupants' comfort.

4.2.3 Error Analysis

Oldewurtel et al. [122] studied the influence of errors in weather forecast on MPC-controlled HVAC operations, and their results indicate that the quality of weather predictions highly correlates with the performance of the model predictive controller. However, the study only

Symbol	Description	Unit
$u(t)$	Supply air temperature at time t	°C
$v^{ij}(t)$	Rate of flow of supply air in room j of a VAV zone i at time t	m^3/s

Table 4.1: List of HVAC control variables

focused on prediction errors in the weather forecast and the evaluation was limited to "pure" MPC-based HVAC controller. Given that occupancy prediction is also an input to MPC, it is essential to analyze the influence of occupancy prediction errors on HVAC operations. Besides, the study [122] is limited to HVAC and does not incorporate the impact of PECs in satisfying the comfort requirements of occupants.

In this chapter, we extend the work in [122] and in [85] by first analyzing the effect of prediction errors in occupancy and later exploring the benefits of PECs in mitigating (or minimizing) the influence of prediction errors on HVAC operations. Our study indicates that predictive control strategies make HVAC operations highly unreliable. High variability has discouraged building managers to use advanced HVAC control strategies, and thus, they have continued using conventional HVAC controllers.

4.3 HVAC Control Strategies

In a typical commercial building, spaces are either private (such as offices) or shared (such as cafeteria, corridor), and a set of private and shared spaces constitutes a zone. Within each zone, there exists a VAV unit that takes air from AHU at a particular temperature ($u(t)$) and supplies it across the rooms at a specific rate ($v^{ij}(t)$) to maintain the room temperature close to the set point temperature. Here, j indicates the room number in the i^{th} zone of the building. To ensure a consistent supply of fresh air, AHU recirculates only a limited amount of used air ($r(t)$) and ejects the remaining air in the open environment.

Defined in Table 4.1, $u(t)$, $v^{ij}(t)$, and $r(t)$ are key HVAC control parameters and their values are typically decided by the control strategy. In this section, we discuss the four control strategies, implemented to analyze the influence of occupancy prediction errors on HVAC operations. The first two methods are non-predictive, and building managers widely use these strategies in commercial buildings today; when employed, the HVAC operations are independent of prediction errors. The last two are MPC-based control strategies. In the chapter, we use non-predictive control strategies as the baseline strategies for predictive control strategies when occupancy prediction is not perfect.

4.3.1 Schedule-based control

In a schedule-based control of HVAC, the building manager starts the HVAC at a fixed time in the morning and shuts it down in the evening (typically 9 AM to 6 PM). On any day, AHU supplies air at a static temperature which is chosen based on the season (summer/winter), and the set point temperature does not vary within a day. Based on ASHRAE standards[1], we set the supply air temperature ($u^{(t)}$) to 15°C for summers and 20°C for winters. For both seasons, the ratio of reuse air (r) and rate of flow of supply air ($v^{(t)}$) is constant at 0.8 and 0.236 m^3/s, respectively. The approach is naive but widely used by building managers in commercial buildings.

4.3.2 Reactive control

In reactive control strategies, VAV cools or heats the space only if people are present in the corresponding VAV zone. In the past, studies have suggested several direct and indirect HVAC control strategies to estimate occupancy; we use the occupancy data and implement the strategy proposed by Ardakanian et al. [5] for benchmarking.

4.3.3 Model Predictive Control

Model predictive control (MPC) is a recent approach for HVAC where controller can compute the room temperature over a finite time horizon [2]. Typically, a thermal model using occupancy estimates and weather forecast determines the future system state over a time horizon. In our implementation of MPC, we used Equation 4.1 as the thermal model that considers the influence of HVAC, atmospheric temperature, heating load by the occupants, and other heating or cooling loads present in the room [140]. Table 4.2 lists all symbols of

[1] American Society of Heating, Refrigeration and Air-Conditioning - a global organisation that publishes standards and guidelines related to HVAC.

Symbol	Description	Default	Unit
ρ	Density of air	1.204	kg/m^3
σ	Specific heat of air	1.003	$kJ/(kg.K)$
τ	Sampling interval	$-$	s
n_z	Total number of VAV zones in a building	$-$	$-$
n_r^i	Total number of rooms in VAV zone i	$-$	$-$
$O^{ij}(t)$	Occupancy in room j of zone i at time t	$-$	$-$
$T^{ij}(t)$	Temperature in room j of zone i at time t due to HVAC	$-$	$°C$
$T_{ex}(t)$	External temperature at time t	$-$	$°C$
C^{ij}	Thermal capacity of room j in zone i	2000	kJ/K
α_{ex}^{ij}	Heat transfer coefficient between outside and room j in zone i	0.048	$kJ/(K.s)$
Q_{ap}^{ij}	Heat load due to heating/cooling equipments in room j of zone i	0.1	kW
Q_{oc}^{ij}	Heat load due to occupant in room j of zone i	0.1	kW

Table 4.2: List of symbols used in the thermal model

the thermal model and their default values.

$$\begin{aligned}\frac{T^{ij}(t+1) - T^{ij}(t)}{\tau} \times C^{ij} = & \frac{\rho\sigma}{n_r^i} \times v^{ij}(t) \times (u(t) - T^{ij}(t)) \\ & + \alpha_{ex}^{ij} \times (T_{ex}(t) - T^{ij}(t)) \\ & + (Q_{oc}^{ij} + Q_{ap}^{ij}) \times O^{ij}(t)\end{aligned} \tag{4.1}$$

For the time horizon, the controller computes $u(t)$, $v^{ij}(t)$, and $r(t)$, by solving an optimization problem using the current state of the system, with an objective to minimize the total energy consumption (Equation 4.2).

$$\begin{aligned}Po(t) = & V(t) \times \eta_h \times (u(t) - T_{cu}(t)) + V(t) \times V(t) \times \eta_f \\ & + V(t) \times \eta_c \times (T_{mx}(t) - T_{cu}(t))\end{aligned} \tag{4.2}$$

where,

$$V(t) = \sum_{i=1}^{n_z} \sum_{j=1}^{n_r^i} v^{ij}(t) \tag{4.3}$$

$V(t)$ depicts the total air supplied across all the rooms within a building, η_h and η_c indicate the efficiency of the heating and cooling unit, respectively. $T_{cu}(t)$ and $T_{mx}(t)$ denote the temperature of air coming from the cooling and mixing unit, respectively. η_f is the efficiency of the VAV fan which is supplying air to the room. Details about the power consumed by the supply fan can be found in Rabbani et al. [136].

The optimization problem constraints the comfort index to remain within the specified bounds to ensure user comfort. In this study, we use widely used metric PMV - Predicted Mean Vote, to measure user comfort [46]. Other constraints include time-scale limitations, thermal dynamics (Equations 4.1), and constraints dictated by the system setup (such as thermal comfort and HVAC operation should remain within a desired range). In this chapter, we implemented MPC with two time-scales where the controller updates the supply air temperature every hour and the supply air volume every 10 minutes. The above time-scales are typically determined by the physical limitation of an HVAC unit. For more details about this specific formulation of the optimization problem, please refer to Kalaimani et al. [86].

4.3.4 MPC with Personal Environment Controller

Recently, Kalaimani et al. [85] proposed a hybrid HVAC controller and combined MPC with a personal environmental control system. In the study, [85] used SPOT - an off-the-shelf desktop fan/heater with local temperature sensing and a computer-controlled actuator to provide individual thermal comfort. Assuming perfect prediction of occupancy and outside temperature, the study shows that combining MPC with SPOT is effective in reducing the total energy consumption by choosing appropriate thermal setbacks during the intervals of sparse occupancy.

At the time of partial occupancy, HVAC runs at a base temperature which is slightly higher (in summers) or lower (in winters) than the desired temperature. Equation 4.1 depicts the base temperature (due to HVAC) which depends on HVAC, external weather conditions, and occupants within the space.

$$\begin{aligned}\frac{T_{hv}^{ij}(t+1) - T_{hv}^{ij}(t)}{\tau} \times C^{ij} = {} & \frac{\rho\sigma}{n_r^i} \times v^{ij}(t) \times (u^{ij}(t) - T_{hv}^{ij}(t)) \\ & + \alpha_{ex}^{ij} \times (T_{ex}(t) - T_{hv}^{ij}(t)) + Q_{ap}^{ij} \times O^{ij}(t)\end{aligned} \tag{4.4}$$

In the proposed approach, we assume that room is divided into two regions: *occupied* - the part of the room where the occupant is present; and *unoccupied* - the other part of the

Symbol	Description	Default	Unit
$T_{hv}^{ij}(t)$	Temperature in room j of zone i at time t due to HVAC	–	$^{\circ}C$
$\Delta_{oc}^{ij}(t)$	Change in temperature of occupied region of room j in zone i at time t	–	$^{\circ}C$
$T_{oc}^{ij}(t)$	Temperature in occupied region of room j in zone i at time t	–	$^{\circ}C$
$T_{un}^{ij}(t)$	Temperature in unoccupied region of room j in zone i at time t	–	$^{\circ}C$
C_{oc}^{ij}	Thermal capacity of occupied region of room j in zone i	200	kJ/K
α_{in}	Heat transfer coefficient between occupied and unoccupied regions of room j in zone i	0.1425	$kJ/(K.s)$
Q_{he}^{ij}	Heat load due to SPOT in room j of zone i	0.7	kW

Table 4.3: Notations used in the revised thermal model

room.

In the occupied region, SPOT provides the offset comfort to attain the comfort requirements of the occupant. In Equation 4.4, we show how the room temperature changes when taking into account the impact of HVAC, heat exchange with the outside, external weather conditions, and other heating/cooling loads present in the room. In Equation 4.5, we then calculate the change in temperature due to SPOT, occupant, and other heat exchanging load present in the room, followed by temperature in the occupied part of the room in Equation 4.6. On the other hand, in the unoccupied portion (Equation 4.7), both SPOT and the occupants *indirectly* influence the room temperature due to thermal coupling between the two zones, modeled by the heat transfer coefficient α_{in}.Table 4.3 lists the new notations used in the extended model.

$$\begin{aligned}\frac{\Delta_{oc}^{ij}(t+1) - \Delta_{oc}^{ij}(t)}{\tau} \times C_{oc}^{ij} = Q_{oc}^{ij} \times O^{ij}(t) + Q_{he}^{ij} \times S_{he}^{ij}(t) \\ - \alpha_{in}^{ij} \times \Delta_{oc}^{ij}(t)\end{aligned} \tag{4.5}$$

$$T_{oc}^{ij}(t+1) = T_{hv}^{ij}(t+1) + \Delta_{oc}^{ij}(t+1) \tag{4.6}$$

$$T_{un}^{ij}(t+1) = T_{hv}^{ij}(t+1) + \frac{\tau \times \alpha_{in}}{C^{ij} - C_{oc}^{ij}} \times \Delta_{oc}^{ij}(t) \tag{4.7}$$

The revised objective function has additional parameters S_f for fan and S_{he} for heater (Equation 4.8). A fan consumes negligible power, thus the objective function only considers the power consumption of SPOT's heater.

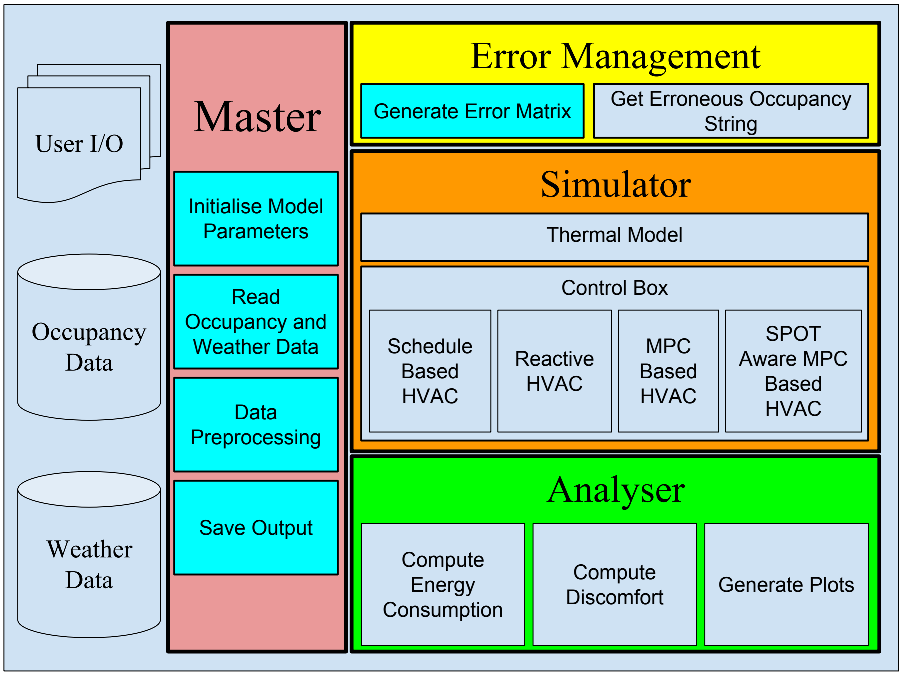


Figure 4-1: ThermalSim is a lightweight C/C++ based building simulation platform that focuses on analysing the influence of prediction errors on HVAC operations.

$$\begin{aligned} Po(t) = V(t) \times \eta_h \times (u(t) - T_{cu}(t)) + V(t) \times V(t) \times \eta_f \\ + V(t) \times \eta_c \times (T_{mx}(t) - T_{cu}(t)) + \sum_{i=1}^{n_z} \sum_{j=1}^{n_r^i} S_{he}^{ij}(t) \end{aligned} \tag{4.8}$$

The controller determines HVAC control parameters (Table 4.1) on a 10-minute timescale and in between, fan/heater (of SPOT) reacts to occupancy every 30 seconds. By doing so, SPOT assists the controller in regulating the discomfort that might arise due to mispredictions; thus ensuring both - personalized comfort and minimal influence of prediction errors on HVAC operations. Next, we discuss the simulator.

4.4 Simulator Software Architecture

To evaluate the impact of forecast errors on the different HVAC controllers, we built a thermal simulator called ThermalSim[2] [75]. ThermalSim is a lightweight C/C++ based simulation platform, whose focus is to study the influence of prediction errors on HVAC

[2] We are still working on its stable release for the community use and we hope to release it soon

```
building:{
        zones:1,                /* Number of Zones */
        rooms:1,                /* Number of Rooms */
        start:20150111T0000,            /* Start Time (Format: yyyymmddThhmm) */
        stop:20160221T2300,             /* Stop Time (Format: yyyymmddThhmm) */
        horizon:4,                      /* Horizon for simulation */
        time_step:600,                  /* Time Step */
        control:2,                      /* Control Type: 1 (No Control), 2 (Reactive), 3 (MPC) */
        ahu:{
                h_eff:0.9,              /* Heating Efficiency */
                c_eff:0.9               /* Cooling Efficiency */
        },
        room:{
                C:2000,                 /* Thermal Capacity of Room - kJ/K */
                C_SPOT:200,             /* Thermal Capacity of SPOT Region - kJ/K */
                alpha_o:0.048,          /* Heat Transfer Coefficient for Outside (kJ/K.s) */
                alpha_r:0.1425,         /* Heat Transfer Coefficient for Regions (kJ/K.s) */
                Q_l:0.1,        /* Heat Load Due to Lightening and Equipments (kW) */
                Q_h:0.1,        /* Heat Load Due to Presence of Occupants (kW) */
                Q_s:0.7,        /* Heat Load of SPOT Unit (kW) */
                fan_coef:0.094
        },
        air:{
                density:1.225,          /* Density of Air - kg/m3 */
                sp_heat:1.003           /* Specific Heat Capacity of Air - J/Kg-K*/
        },
```

Figure 4-2: Input format for ThermalSim.

operations. ThermalSim, as shown in Figure 4-1, consists of four major modules:

1. Master - to handle data I/O and preprocessing,
2. Error Management - to inject *unbiased* errors in the occupancy streams,
3. Simulator - to simulate room temperature for a given thermal model and control logic,
4. Analyser - to compute energy consumption, occupant comfort, and analyze simulated data streams.

In the current version, Simulator module incorporates AMPL [4] – an algebraic modeling language for the mathematical programming – to compute the control parameters.

4.4.1 Master Module

The Master module takes as input historical weather and occupancy data in CSV (Comma Separated Values) format, a user-generated description of the building, and simulation control parameters (Figure 4-2) including start and stop time of the simulation, parameters of the thermal model, control strategy, among others. Before executing the simulations, the Master module pre-processes the data, and after completion saves the output of simulation in the CSV format.

4.4.2 Modeling Occupancy Prediction Errors

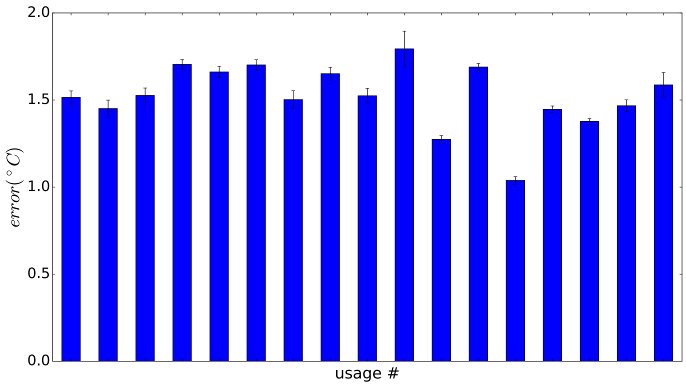
ThermalSim represents occupancy data for a day as a string of consecutive 0's (for unoccupied workspaces) and 1's (for occupied spaces). We consider only two states of occupancy because a majority of occupancy prediction algorithms use occupancy as a two-state variable. We call this string an *occupancy string*. The length of a single occupancy string depends upon the sampling rate of the occupancy data. Data sampled every ten minutes will generate an occupancy string of length 144 characters, and if the sampling rate is thirty seconds, the string will be 2880 characters long.

Error Matrix

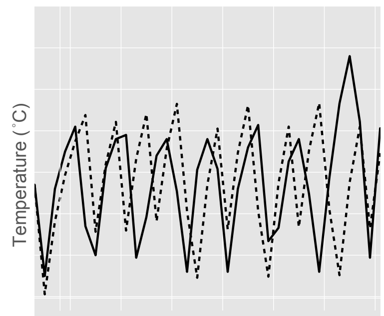
It is important that occupancy forecast errors be realistic. For example, it does not make sense to randomly flip occupancy states, since this may result in forecasting occupancy during the middle of the night, which is very unlikely. Our key insight is that a likely outcome of an errored forecast is to forecast *another valid occupancy string*, with the observation that the higher the error rate, the larger the distance, in an appropriate metric space, between the true and the errored strings.

We use the following approach: For a dataset with n occupancy strings, each cell of an *error matrix* depicts the Hamming Distance between any two occupancy strings – the number of mismatching characters [61]. To normalize, we divide value in each cell by the length of occupancy string. The *error matrix* is a symmetric matrix of size n^2 which helps in systematically injecting unbiased errors in the occupancy data.

To illustrate, consider a scenario where we want to analyze different control strategies with 10% prediction error in the occupancy data. The error management module will refer *error matrix* for an occupancy string which is closest to the day of analysis. We term the selected occupancy string as the *reference* string. The module will then look into the *error matrix* to find all those strings that have 10% error as compared to the *reference string* and randomly select one. We call the selected one an erroneous string. If the day (*reference* string) was 30% occupied, then the occupancy in the *erroneous* string may fall anywhere in between 20%-40%.



(a) ThermalSim can simulate daily room temperature with an RMSE of $1.52°C(\sigma = 0.18°C)$.



(b) Recorded v/s Simulated

Figure 4-3: The hard line indicates the actual room temperature and dotted line indicates the predicted room temperature.

4.4.3 Simulator

The simulator module takes input from the master and error management modules to simulate the room temperature. It comprises two major blocks - 1. thermal model - depicts various thermal interactions occurring within a room, and 2. control module - to compute the control parameters. In the current version, we have implemented two thermal models - 1. single region - no partition exists within a room (Equation 4.1), and 2. two regions - the occupied area is separated from the unoccupied portion by a thin layer of air (Equations 4.4-4.7). As discussed in Section 4.3, we have implemented four HVAC controllers in ThermalSim - 1. schedule-based, 2. reactive, 3. model predictive control (no SPOT device present), and 4. SPOT-aware model predictive control. In the rest of the chapter, we will use NS as an acronym for No-SPOT model predictive control and SA for SPOT-Aware MPC.

Simulator Validation

To quantify the accuracy of *ThermalSim* in simulating room temperature, from a room in residential apartment, we collected temperature data for 17 days and carried out leave-p-out cross validation with $p = 5$. In such an approach, we validate the model on p observations and use the remaining observations for training. We used a non-linear solver whose objective was to minimize the residual between predicted and actual room temperature. The simulator tunes following model parameters -

1. thermal capacity of the room (C),
2. heat transfer coefficient between outside and room (α_{ex}),
3. coefficient of heating/cooling ($\rho\sigma$)
4. heat load due to occupants (Q_{ac}), and
5. heat load due to heating/cooling appliances (Q_{ac}).

Our analysis (in Figure 4-3a) indicates that *ThermalSim* can simulate the daily room temperature with an RMSE (Root Mean Square Error) of $1.52°C(\sigma = 0.18°C)$. Figure 4-3b depicts the average (solid line) and predicted (dashed line) room temperature. Note that though the predicted room temperature follows the pattern of actual room temperature, it fails to align perfectly. Though misalignment does increase the RMSE at some time instances, we found that it has little overall impact on total energy consumption and occupants' comfort.

4.4.4 Metrics

Energy Consumption

Equation 4.9 computes the total energy consumption of a building for a day. Here, $Po(t)$ denotes the power consumption of HVAC and other heating/cooling devices, τ is the sampling rate, and n_t is the number of daily samples.

$$E = \sum_{t=0}^{n_t} Po(t) \times \frac{\tau}{3600} \tag{4.9}$$

Occupant Discomfort

ThermalSim leverages Predicted Mean Vote (PMV) [6] to estimate the comfort level of the occupants (Equation 4.10). At a given time instant t, if PMV ($P^{ij}(t)$) lies within the comfort requirements ($[P_{ll}, P_{ul}]$) of an individual then we mark the room as comfortable, else uncomfortable. $D_{\%}^{ij}$ denotes the percentage of time instances in a day when the user was uncomfortable in the room.

$$P^{ij}(t) = P1 \times T_{oc}^{ij}(t) - P2 \times v_a^{ij}(t) + P3 \times v_a^{ij}(t) \times v_a^{ij}(t) - P4 \tag{4.10}$$

$$D^{ij}(t) = max(0, P_{ll} - P^{ij}(t), P^{ij}(t) - P_{ul}) \tag{4.11}$$

$$D_{\%}^{ij} = \frac{\sum_{t=0}^{n_t}[D^{ij}(t) \neq 0]}{\sum_{t=0}^{n_t}[O^{ij}(t) = 1]} \tag{4.12}$$

Robustness

Prediction errors are stochastic in nature and their impact on energy consumption and occupant comfort depends on two factors:

Nature of the Error: If the prediction algorithm mispredicts occupancy for short time intervals (say for a minute or so), we term the prediction errors as point errors, otherwise we call them burst errors. For a particular error percentage, an erroneous occupancy string can have point errors, burst errors, or a mix of both; resulting in different values of energy consumption and occupants' discomfort for the *same* error percentage.

Timing of the Error: The occupancy prediction algorithm can make errors at any time of the day - such as during peak or non-peak time. Consider the situation where the occupancy prediction has 15% error during the peak hours and the controller assumes one of the five rooms to be occupied though it was unoccupied. In this situation there is a high chance that the HVAC might be already running during that time. Given the fact that the other four rooms are occupied, this particular prediction error will have an insignificant impact on the HVAC operations. However, during night time, the same error percentage might waste significant energy. This illustrates that the *timing* of the prediction errors has a significant impact on both comfort and energy consumption.

For a specific error percentage, depending on the nature and timing of the errors, the energy consumption and user discomfort may either increase or decrease, potentially destabilizing HVAC operations. For a specific example, consider the big circle and triangle in Figure 4-4, which depict the energy consumption and user discomfort for NS and SA controllers respectively for perfect occupancy predictions in a particular simulation scenario. For a specific error percentage, the small circles (NS) and triangles (SA) depict the energy consumption and user discomfort for fifteen different erroneous occupancy strings. We no-

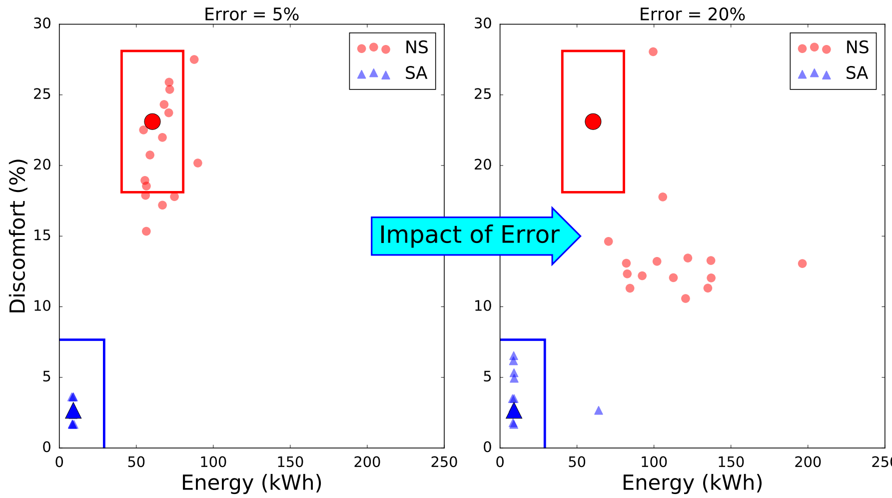


Figure 4-4: As error increases, the energy consumption and occupants' discomfort vary depending on the *nature* and the *timing* of prediction errors. 5% errors on the left and 20% on the right. Large circles/triangles indicate a perfect prediction scenario and small circles/triangles correspond to those scenarios when occupancy prediction was erroneous.

ticed that as prediction error increases from 5% (left) to 20% (right), the points indicating erroneous strings start moving away from the results obtained from perfect prediction.

Note that the circles (NS) are more scattered than the triangles (SA). In the case of NS, the system decides the control parameters such that the desired room temperature (which is the same for each room) is achieved across all the rooms. In case of a sudden change in the occupancy, NS updates the control parameters, but it takes significant time to re-attain the energy-discomfort tradeoff setpoint. In contrast, in SA, the controller knows the current state of SPOT; thus, the controller chooses a set point such that HVAC provides a certain level of comfort to the occupants and SPOT provides the necessary additional offset. SPOT, being responsive in nature, keeps the comfort level of individuals within the desired range with insignificant increase in aggregate energy consumption. Therefore, even if the error percentage increases, the energy and discomfort stays close to the perfect prediction for SA whereas NS becomes highly unstable.

To capture this phenomenon, Equation 4.13 defines a *robust* ($cs \in \{NS, SA\}$) metric which quantifies the robustness of a particular control strategy cs towards the prediction errors. It computes the number of instances that stay within the desired limits of the building manager.

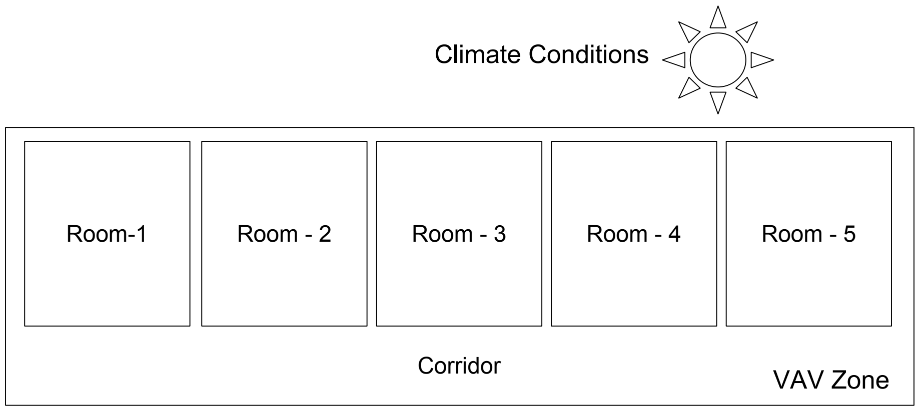


Figure 4-5: For evaluation, we considered a hypothetical building consisting of 5 rooms separated by walls.

$$robust_{cs}\ (\%) = \frac{\#\text{ of instances within limits}}{\text{total }\#\text{ of instances}} \times 100 \tag{4.13}$$

For concreteness, we use $\pm 20\ kWh$ and $\pm 5\%$ as the acceptable limits for energy consumption and occupants' discomfort, respectively, as shown by the rectangles in the figure. For the given scenario (in Figure 4-4), when the error percentage is increasing from 5% to 20%, NS is less *robust* towards the prediction error (60% $\rightarrow$ 0%), however, SA remains consistent (100% $\rightarrow$ 93%). For a predictive control strategy, a PEC system (like SPOT) mitigates the effect of prediction errors to make the HVAC operations more reliable and robust. Whenever there is an unexpected occupancy in the room, SPOT can react quickly as compared to central HVAC system which has a slower time-scale.

4.5 Evaluation

4.5.1 Test Building Description

For our evaluation, we consider a single zone in a typical building comprising of five rooms where each room is surrounded by walls on three sides and has a window exposed to weather conditions on the fourth (Figure 4-5). We assume an AHU and a VAV unit in the building. Though the structure is hypothetical, it is a typical architecture for faculty offices in Universities where thick brick walls separate the rooms. We believe that the key insights obtained for the study are well representative of more complicated building architectures. Note that ThermalSim can also deal with more complicated structures, should that be desired. When

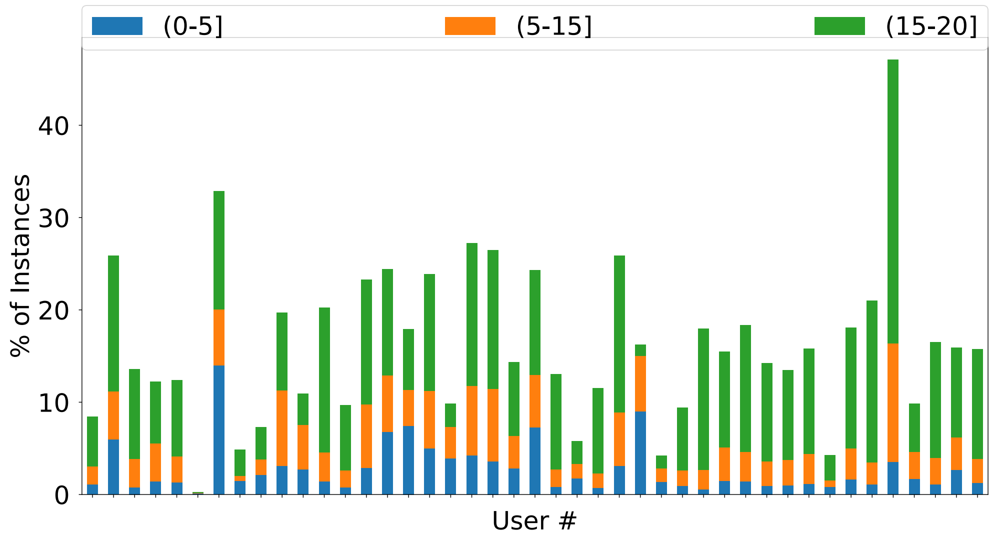


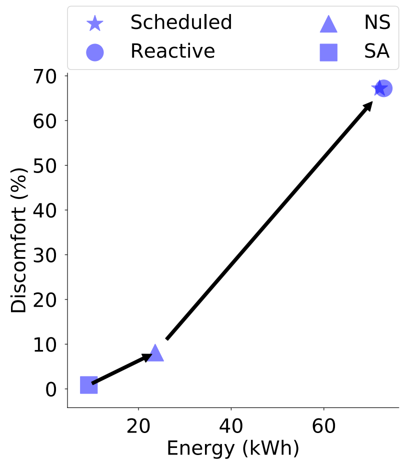
Figure 4-6: We sample occupancy every 30 seconds; in every 10-minute interval, there exist 20 measurements of occupancy. Here, the color indicates the number of 30 seconds instances in a 10-minute interval when the room was occupied. Notice that room would be marked occupied for all the three scenarios, however, the percentage of instances when the room was occupied for less than 2 minutes (in the range of (0, 5]) is relatively low.

we evaluate MPC with a reactive controller, we also consider effect of SPOT heater/fan on the room temperature. For the stated scenario, we next discuss the dataset.

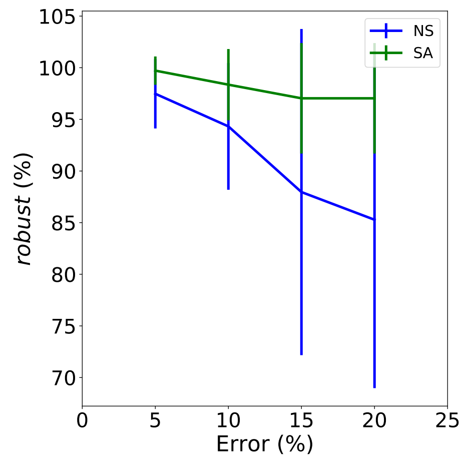
4.5.2 Dataset

ThermalSim requires real-world occupancy data to generate an *error matrix*. We leveraged an existing deployment from our university campus and gathered occupancy data (along with other information) from more than fifty volunteers – including students, faculty, and the staff members every 30 seconds for a year.

An MPC requires occupancy information in every 10-minute to compute the control parameters 24 hour time horizon; therefore, we upsample the occupancy data from 30-seconds to 10-minutes by applying the following rule - *"Mark a 10-minute interval unoccupied if all the 30-second instances indicate the room to be unoccupied, else mark the space as occupied".* However, with this strategy, even if a single instance in the 10-minute interval is occupied, the controller will mark the space as occupied for the whole duration. To understand whether such a bias is limiting or not, we analyzed the occupancy data and our analysis indicates that data has only 3% 10-minute instances where the room is occupied for



(a) Energy-discomfort plot when prediction is perfect.



(b) As error increases 5% → 20%, SA stays more robust than NS.

Figure 4-7: [Left] The arrow indicates the performance degradation, in terms of energy consumption and user comfort, when we move from predictive to non-predictive control strategies. [Right] Error bars indicate the variation in different simulated scenarios. For system to be more robust, the length of error bar should be smaller.

less than 2 minutes (Figure 4-6). Therefore, we only mark a 10-minute interval unoccupied if the room was occupied at all the 30-second instances within that interval.

4.5.3 Evaluation Setup

Our hypothesis is that the benefits of using a PEC system like SPOT along with HVAC controller mitigates the influence of prediction errors on MPC-based HVAC operation.

We validate this hypothesis assuming occupants in all the five rooms have similar comfort requirements: $[23°C, 25°C]$ in summers and $[21°C, 23°C]$ in winters. For the given setup, we compare the performance of predictive and non-predictive HVAC controllers for 25 days, both in summers and winters.

For each day, we select an occupancy string from the *error matrix* that deviates (from the current day) by the error percentage specified in the system. For instance, if we wish to introduce 10% error in the current day occupancy string, we search for another occupancy string in historical data where 288 out of 2880 instances (for a data sampled every 30 seconds) have a mismatch with the current day occupancy string. *ThermalSim* utilizes both actual and erroneous occupancy string to simulate the building (depicted in Figure 4-5) for all the four control strategies and compare their performance.

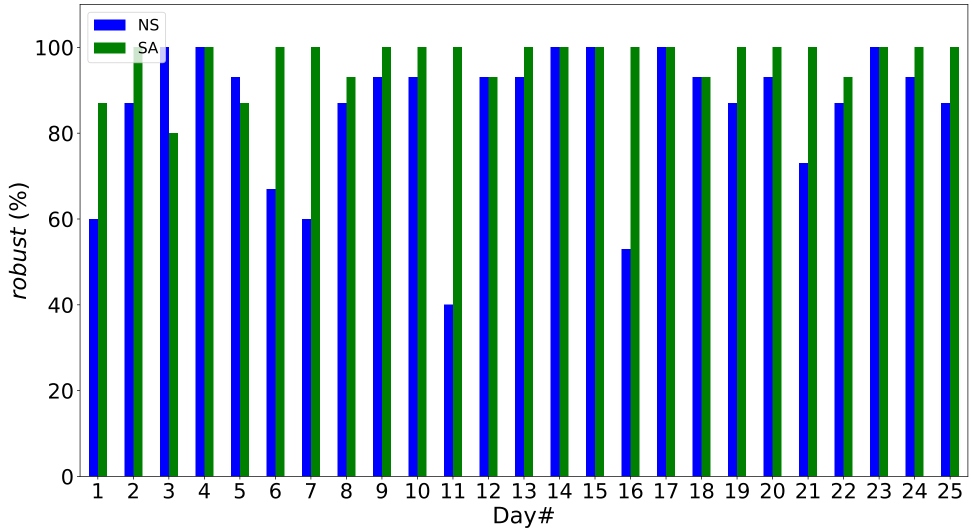
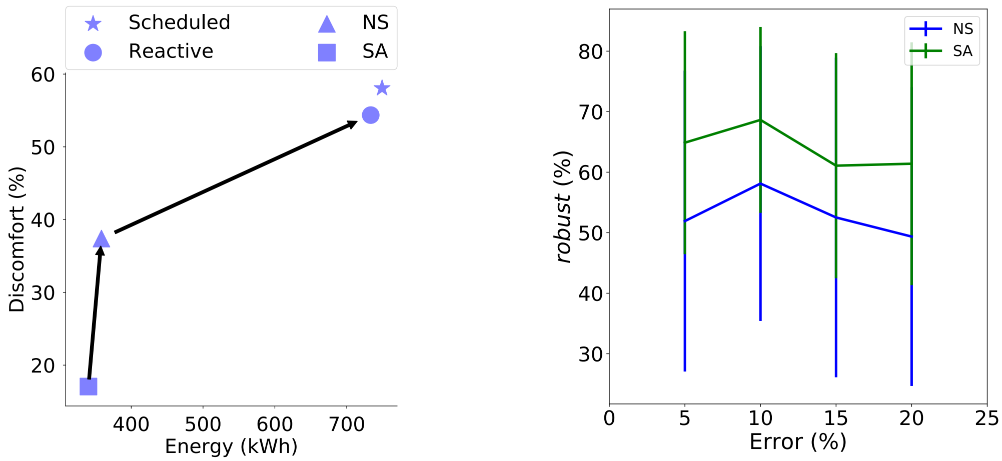


Figure 4-8: For 20% prediction error in occupancy, SA is more reliable and robust NS across all the 25 days of summer.

To mitigate any bias in the selection of erroneous occupancy strings, *ThermalSim* evaluates fifteen different erroneous occupancy patterns for each day and error percentage. Furthermore, a separate analysis for each of the two seasons provides better understanding of the influence of seasonal variations.

4.5.4 Insights

In the jurisdiction corresponding to our temperature data set, i.e., Southwestern Ontario, we find that for all control strategies, the HVAC system consumes less energy in summers when compared to winters (see Figures 4-7a, 4-9a). In our setting, the outside temperature in summers is only a few degrees higher than the desired room temperature and so the HVAC has to put in less effort to achieve the desired comfort. On the other hand, in winters, the HVAC energy consumption is significantly higher because the outside temperature is quite cold. In winters, all control strategies attempt to maintain a room temperature in the range of $21°C$ and $23°C$ which is much higher than the outside temperature (approx. $-10°C$). Consequently, HVAC has to expend more energy in winters than in summers to attain the desirable comfort conditions in the occupied zones.



(a) Energy-discomfort plot when prediction is perfect.

(b) Even with slow heater, SA is better or comparable than NS.

Figure 4-9: [Left] The arrow indicates the performance degradation, in terms of energy consumption and user comfort, when we move from predictive to non-predictive control strategies. [Right] Error bars indicate the variation in different simulated scenarios. For system to be more robust, the length of error bar should be smaller.

User Experience

The schedule-based and reactive controllers can make occupants uncomfortable and yet consume significant energy, even with perfect prediction. When set to follow a fixed schedule, HVAC supplies air at a constant flow and temperature, and does not consider occupants' schedules or daily temperature changes. For pictorial representation, we use energy-discomfort plot where x-axis denotes the daily energy consumption of the building and y-axis represents the total discomfort for the users. Consequently, with a *schedule-based* control strategy, user experience lies in the top-right corner of the energy-discomfort plot with maximum energy consumption along with notable discomfort for both the seasons (see Figures 4-7a, 4-9a). On the other hand, a reactive controller with occupancy information is marginally better or equivalent to the schedule-based controller. Model predictive control (with no SPOT) shows significant improvement in minimizing both energy consumption and occupants' discomfort. Given the weather forecast and occupancy prediction, MPC keeps updating the temperature and volume of supply air at regular time intervals.

As central HVAC unit cannot cater to the dynamic schedule of the occupants, discomfort in NS is slightly higher than the hybrid control strategy that integrates SPOT with

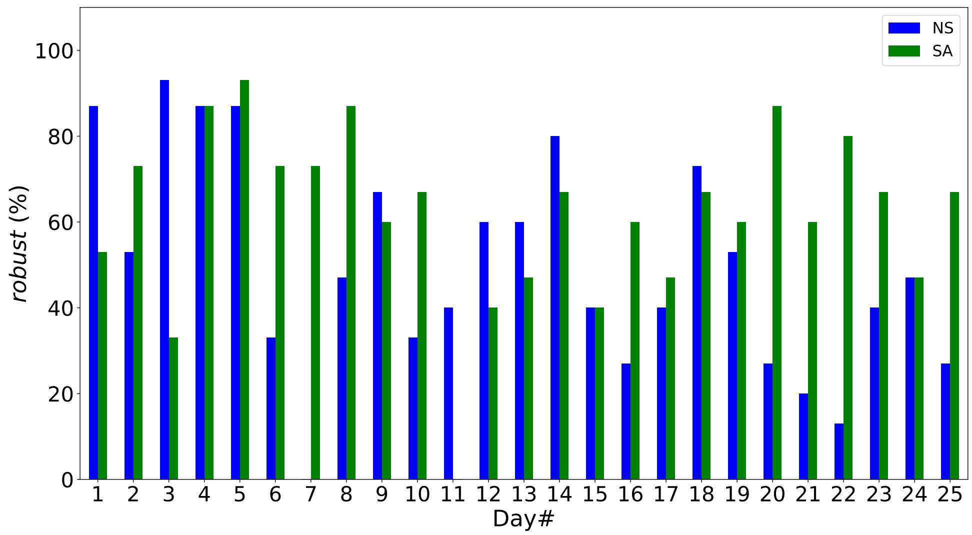


Figure 4-10: For error percentage as high as 20%, note that SA has less deviation in HVAC operations than NS.

MPC to satisfy the comfort requirements of each individual in the building. In SA, the central HVAC system is aware of the SPOT system, therefore, the controller choses the set point temperature such that HVAC can provide minimal comfort, and SPOT can offset the individual comfort requirements. This results in additional savings in energy when there is partial occupancy is in line with the results from previous study by Rachel et al. [85]. Next, we observed that the discomfort is negligible for summer as opposed to winter. The fan assists the occupant in quickly achieving her desired comfort level as opposed to a heater which takes comparatively more time to increase the temperature to provide the offset. In conclusion, irrespective of the season, both SA and NS strategies improve comfort and energy compared to schedule-based and reactive, with SA outperforming NS.

Error Analysis

When occupancy predictions are erroneous, depending upon the *nature* and *timing* of errors, energy consumption and occupants' discomfort vary, hence HVAC operations become highly variable. When analysed over 25 days each for 15 different occupancy patterns, we find that the SA control strategy is more robust than NS even with a high error percentage. As the prediction error increase from 5% to 20%, the performance of NS drops while SA performance remains quite consistent (Figure 4-7b). For 20% prediction error, SA ($\sigma = 5\%$)

is 12% more robust than NS ($\sigma = 16\%$).

Next, Figure 4-8 shows that the SA is consistently robust across all the days as compared to NS for 20% prediction error in the occupancy. Highly unreliable HVAC operations lead to significant variations in the energy consumption and the user comfort. Though the fan makes insignificant impact on the room temperature, it quickly achieves the desired comfort level by providing a cooling perception to the user. Thus, the fan is very helpful in dealing with the unexpected changes in the occupancy of the room while NS alone fails to do so.

We noticed that a fan is more effective and quicker than a heater in mitigating the effect of prediction errors on both energy consumption and discomfort. When a room gets occupied, a heater slowly increases the room temperature to achieve the comfort requirements of the occupants. This results in few intervals of discomfort for the user. This effect is visible in Figures 4-9b, 4-10. For 20% prediction error, we noticed that the SA is now even less robust than NS on few days due to the slow reaction of the SPOT heater. However, the average performance of SA is still better or comparable than NS.

4.6 Challenges And Discussion

In this chapter, we analysed the influence of prediction errors in occupancy on the HVAC operations while leveraging a custom-built building simulator - *ThermalSim*. In this section, we summarize our results, discuss various limitations of the study followed by research questions which are open for the community.

Our insights include the following: First, our dataset indicates that aggregate energy consumption is higher in winters than in summers. Second, integrating a PEC like SPOT with a predictive HVAC controller is definitely better or comparable than a pure MPC based approach. Third, for SA controller, fast reactive device (such as fan) is 20% better than the heater, in terms of occupants discomfort. Finally, NS typically fails to satisfy the comfort requirements on any day.

Our work suffers from two main limitations. First, while the thermal model (of *ThermalSim*) considered the effect of numerous sources (such as weather, occupancy) affecting the room temperature, there still exist various other factors (such as humidity) which are critical for such analysis. We plan to explore such factors and enrich the data for a deeper analysis in future.

Second, we carried out the study through a dataset collected from a particular part of the world. Climate, users' attitude (towards energy savings), and many other factors differ significantly across the geographies. Though the results indicate that SA is more robust than NS, there can be considerable discrepancy across (and within) the countries. A real-world implementation of the technology is critical to understand its effectiveness in achieving the desired goals.

We find that mitigating the effect of prediction errors possess considerable potential in optimising the HVAC operations with predictive controllers. While model predictive control (MPC) is one of the most promising state of the art HVAC control strategies, its performance is limited by the accuracy of the weather and occupancy predictions. Therefore, we designed a custom-built building simulator – *ThermalSim* – to analyse the influence of prediction errors on HVAC operations. We also proposed a method to introduce realistic errors in occupancy for the analysis. Our initial analysis indicates that prediction error (in occupancy) of 20% can make the HVAC operations highly unstable in terms of both energy consumption and occupants' comfort. Recent literature shows that it is feasible to use a personal thermal comfort system – SPOT – along with predictive strategy to ensure personalised comfort in personal and shared spaces. We observed that while SPOT is effective in attaining better personalised comfort, it also strengthens the predictive strategies by mitigating the influence of predictions errors on energy consumption and occupants' comfort because it works at a finer time-scale than the MPC-based HVAC. Employing a personal thermal comfort system, such as SPOT, we stay in the acceptable region 95% of the times as oppose to 83% of the times even for the prediction errors as high as 20%, in the occupancy; thus, motivating a reliable control strategy across the commercial buildings.

Chapter 5

Leakage Detection

Watch the little things, a small leak can sink a great ship.

– Benjamin Franklin, Founding Father of the United States

5.1 Introduction

While energy savings and user comfort are essential, fault detection is of paramount importance. Refrigerant leakage is a common mechanical fault in compressor based appliances which are primarily used for air-conditioning and refrigeration purposes [23]. The puncture hole often starts as a pinhole leak and becomes bigger when goes undetected. Due to the loss of refrigerant, compressor (a component within the RU) works with reduced efficiency and takes more time (than the usual) to cool the room, thus, wasting significant energy. In addition to that, the leakage exposes tenants to refrigerant which is extremely dangerous for their health [7]. The consequences of refrigerant leakage are even worse for retail outlets who set up cold-rooms to preserve perishable food items (usually at 5°C-8°C), and the stored product goes bad due to improper cooling by the RU during the leakage. Early detection of such leaks can benefit the retail enterprises in - 1. increasing their profits by reducing the energy wastage, 2. avoiding leakage of hazardous refrigerant in the open environment, and 3. maintaining the product quality [40, 39, 49, 34].

Unfortunately, store managers usually lack necessary skills to sense the leakage at an early stage, and undetected leaks widen with time as RU remains functional during the leakage. At breakdown point, RU stops cooling followed by the complete shutdown of outlet for few days. Figure 5-1 depicts one such scenario where store manager reported the

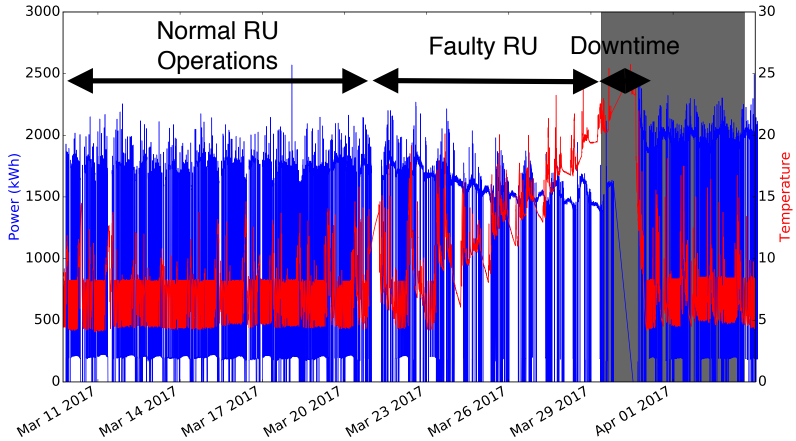
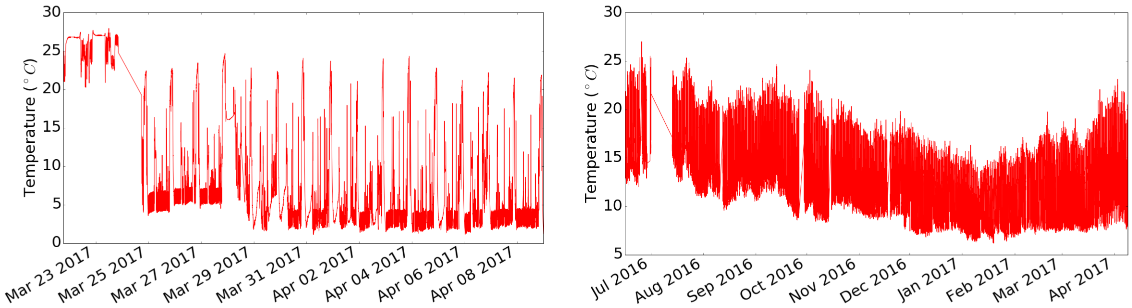


Figure 5-1: Gas leakage started on March 21, but manager kept using the RU for a week and complained on March 29 when it stopped cooling. Delay in reporting refrigeration leak led to complete shutdown of store operations for 2-3 days.

problem when RU broke, and the store went out of operations. In addition to repairing cost, the owner also dealt with the staling of stored products and business loss owing to the downtime. Current techniques for leakage detection are typically categorized as direct and indirect methods [169]. Direct methods, such as Halide leak detector, require an expert to use the detector for on-spot leakage detection with certain precautions [21, 81, 109]. On the other hand, indirect techniques need specialized sensors to monitor temperature, pressure, and mass flow rate at multiple points inside the appliance [47, 139, 129]. As pressure and mass flow sensors are more expensive than temperature sensor, several studies focused on limiting the number of sensors for leakage detection [48, 169]. However, even with the limited sensors, Yoo et al. [169] required temperature sensors at five points inside the system - 1. air inlet of indoor unit, 2. air inlet of outdoor unit, 3. evaporator midpoint, 4. condenser midpoint, and 5. compressor discharge point. Such an extensive sensor installation (that too within the appliance) often limits indirect techniques to laboratory setup and considered as an expensive and unscalable solution to the problem.

In this chapter, we reinforce smart thermostats with *Greina*[1] - a framework that observes deviations in the measured temperature from the estimated temperature to detect refrigerant leakage. The estimates come from a lumped parameter thermal model whose parameters are tuned to a particular room environment by the ambient information sensed by the thermostat. Smart thermostats are 1. plug-n-play - no need of technician for instal-

[1] *Greina* is an *Icelandic* word that means *to identify*.



(a) During daytime store manager frequently visits the cold-room which leads to high temperature in working hours.

(b) In winters (December 2016 - March 2017), room temperature stays in a lower range as compared to summers.

Figure 5-2: Room temperature depends on the activities of store manager and seasonal environmental changes.

lation, 2. allow remote sensing of RU, 3. never intervene in the daily routine of the store managers, and 4. anticipated to increase by 400% in next couple of years [95]. Given the benefits, studies in the past explored smart thermostats for user feedback [78], occupancy detection [92], energy-efficient control [107], among many other application. However, to the best of our knowledge, no one has studied smart thermostats for leakage detection. We believe enhancing smart thermostats for leakage detection presents a genuinely low-cost and scalable solution for the problem.

Despite all the benefits, one must also note that the overall problem is non-trivial and exhibits multiple challenges. Room temperature is sensitive to the actions of outlet manager (Figure 5-2a) and outside weather conditions (Figure 5-2b), and both these parameters differ significantly across the outlets. Even within a store, the daily routine of the manager, climatic conditions, and fitness level of an appliance change considerably with time. To consider both temporal and spatial variations, *Greina* utilizes a first-order thermal model which simulates the room temperature while considering the influence of outlet manager and climate outside the room. *Greina* collects weather data through a cloud-based weather service, and monitors user activity through a door sensor. With time, as framework receives more and more *clean* data, it keeps updating the model parameters to accommodate temporal changes in building's thermal behaviour. However, *clean* training data - when no fault exists within the RU - is itself a challenge with most of the existing techniques for leakage detection. To ensure clean data during the learning phase, *Greina* does not consider those days which were identified as leaking while monitoring the RU. In the initial stages, when sufficient *clean* data is unavailable to tune the model parameters, *Greina* benefits from transfer learning and uses the parameters of a contextually ‘similar’ outlet. Both, *clean* data and transfer learning ensure that the proposed framework is ready-to-work from the day of installation. Finally, *Greina* compares the estimates from the tuned thermal model with sensor measurements and when the actual temperature is sufficiently above the estimate, the system raises a leakage flag.

Of course, room temperature can be higher than the estimates even when RU is working fine (false positive), and vice-a-versa, the framework might confuse the initial symptoms of leakage with the noise generated from manual interventions (false negative). For instance, in Figure 5-1, the room temperature is in the range of 8°C-10°C in the initial stages of the refrigerant leak, akin to temperature variations due to tenants’ activities in Figure 5-2a. If misclassification is a false negative (misinformed that RU is fine), then repairing will get delayed until the outlet manager identifies the leakage. However, if misclassification is a false positive (misinformed that RU broke), then the company might end up paying a significant amount to their maintenance contractor for the unnecessary visits. While such visits are immoderate, they are annoying for the store manager and disruptive to the daily operations. Henceforth, *Greina* employs CUSUM (Cumulative Sum Control) technique to ensure that user gets notified only when the system is confident about the leakage.

We evaluate *Greina* using the data collected from 74 outlets of a retail enterprise for one year. For ground truth verification, we use fault logs as reported by the servicing company after addressing the fault. Results indicate that *Greina* is comparable or better than manual reporting. The proposed framework can reduce the average reporting delay by 5-6 days, when compared to manual reporting. In other words, even if outlet managers could use the leakage detectors (such as Halide leak detector), they would have delayed the repairing by almost a week. We couldn’t compare *Greina* with indirect methods (of leakage detection) because proposed techniques are often limited by the need of specialized sensors to monitor temperature at multiple points within the RU.

Given the energy-saving features and other advancements, smart thermostats are anyways going to replace traditional thermostats very soon. As of now, thermostats can monitor room temperature, tenants’ daily activities [2], and even upload the sensed information to

[2] Though, in this study, we used door sensor to monitor human activities, Appendix C discusses an extended thermal model which replaces door sensor with the motion sensor of the smart thermostat.

the cloud. Location-based detailed weather information is readily available from several cloud-based climate monitoring applications. By combining the two pieces of information, we can use *Greina* for leakage detection without instrumenting the appliance. In addition to being scalable, the proposed approach is reliable enough to maintain food quality and generate minimal false alarms. Finally, *Greina* requires minimal (or no) intervention from the store manager and ensures non-interruptive working hours. Currently, in collaboration with an energy-analytics based venture, we are working on the deployment of proposed thermostats across all the 74 outlets which were considered as part of this study.

5.2 Related Work

Anomaly detection is an active research area with an extensive literature in numerous domains and applications [26, 93]. One such problem is fault detection and diagnosis (FDD) where the objective is to first identify a malfunctioned appliance followed by a *root-cause* analysis to diagnose the problem [54, 71]. Usually, the partial or complete failure of the hardware (present in the equipment) brings down the whole system. Though the literature on fault detection in mechanical units is considerably extensive, we will limit the scope of this survey only to compressor-based appliances.

5.2.1 Fault Detection Frameworks

Typically, fault detection frameworks for compressor-based appliances are designed either for Refrigeration Units [19, 156], or Heating, Ventilation, and Air Conditioning Units (HVAC) [163, 23, 38, 170]. Irrespective of type of appliance, most fault detection engines are designed on the principal - simulate a baseline and compare with the measured signal (mostly energy) to mark any deviations due to the fault [41, 112, 116, 50, 125, 27]. O'Neil et al. [124] simulated the energy consumption through a building simulation framework, EnergyPlus [31], and compared with current energy consumption to identify acceptable performance. Mavromatidis et al. [110] suggested an artificial neural network based model to provide energy baseline, and use the baseline to detect faults in supermarket refrigerators. Li et al. [105] studied correlation between electrical signals and the common faults for roof-top air conditioners by conducting a series of experiments. Srinivasan et al. [149] proposed an energy model to pick anomalies in the power signature of supermarket refrigerators.

Power consumption data provides better visibility of leakage than the room temperature because it is less sensitive to environmental noise, such as occupants' activity. Even technicians prefer to analyze electrical component such as input/output voltage and supply current for fault detection. Moreover, with the advent of advanced techniques to estimate usage and operating conditions of the appliance through NILM (Non-Intrusive Load Monitoring) [16, 17] and EMI (Electromagnetic Interference) signatures [28, 59, 147, 58], the use of energy signal for fault detection is worthwhile. However, one must note that though power signal is useful for fault detection, it won't allow us to monitor the consequences. When fault occurs, power signal cannot tell if the appliance is maintaining suitable temperature for the occupants (if conditioning the air), or the stored products (when using for refrigeration).

Studies also monitored different other parameters to develop a black box model for fault detection [167, 168, 62]. Keres et al. [90] monitored compressor frequency to identify fault condition in the refrigerant unit. Porter et al. [135] collected actual operating parameters from a set of microsystem sensors installed throughout the refrigeration and compared with ideal conditions of the system for fault detection. Payne et al. [129, 128] proposed self training of a fault free model, and used the model to detect fault through a data-clustering approach while observing temperature at nine points within an air-conditioner. Similarly, Kulkarni et al. [97] designed a random forest-based binary classifier to detect issues in refrigeration cases while using temperature signal and defrost state from supermarket refrigeration cases. While most of the above mentioned studies focused only on detecting 'abrupt' faults, none of them looked into the problem of refrigerant leakage.

5.2.2 Leakage Detection Frameworks

Specifically for leakage detection, researchers designed systems primarily from the perspective of mechanical engineering, typically categorized as direct and indirect methods.

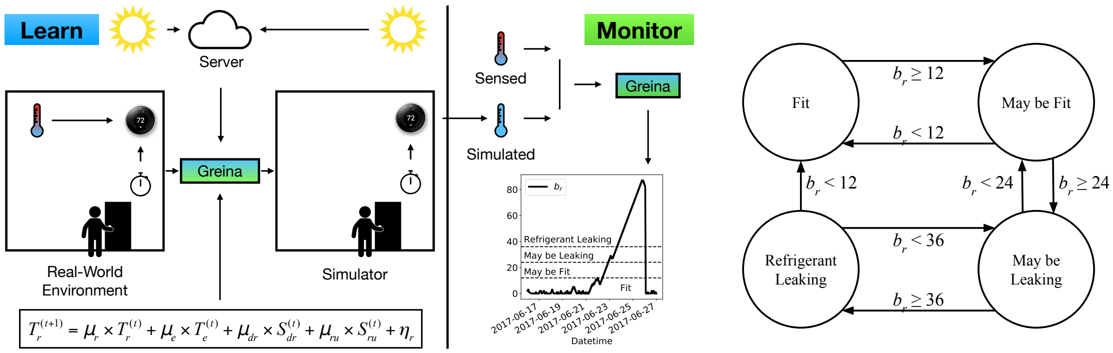
Direct Methods

In direct methods, typically a technician visits the site and uses a leakage detector [21, 109] to confirm the leak. Hailey detector [81] is one such widely used leak detector in which flame changes the color when refrigerant is present in the environment. Parekh [126] proposed composition of fluorescent, alkyl substituted perylene-dye, and a polyhalogenated hydrocarbon refrigerant to visually detect refrigerant leakage. Though these detectors are accurate, one must use them with certain precautions. For instance, we cannot use Hailey detector with hydrocarbon refrigerants. Moreover, the effectiveness of direct methods primarily relies on manager's ability in timely detecting the leakage.

Indirect Methods

Alternatively, in indirect methods, framework monitors multiple parameters at different points within the appliance for leakage detection [82, 153, 141, 84, 42, 113, 114]. Taylor et al. [154] designed a neural network to analyze data from multiple monitoring alarms to predict refrigerant leakage. Rossi et al. [144] proposed a statistical rule-based leakage detection technique which could detect 5% loss of refrigerant through extensive instrumentation. Breuker et al. [24] studied nine temperature and one relative humidity parameters to characterize *soft* faults (such as loss of refrigerant) and their impact on the operations of roof top air conditioners. As majority of leakage detection research [104, 117, 57] assumes steady-state, Kim et al. [91] proposed a methodology for developing a steady-state detector for any generic vapour-compression system. Steady-state in itself is a misleading term because parameters of a vapour-compression system are dynamic in real-world [91]. Therefore, the study recommended inclusion of all leakage detection features (for best results) which makes overall approach expensive due to extensive instrumentation of the appliance.

To minimize the number of sensors and sensing points, Fisera et al. [47, 48] developed an energy consumption model by monitoring nine parameters (including relative humidity and carbon-dioxide level) to distinguish anomalous and degradation events. In a patented technology, Suzuki et al. [151] compared theoretical heat dissipation of the condenser with actual temperature difference (in the condenser) for the leakage detection. In the same direction, Yoo et al. [169] monitored temperature difference between inlet air and midpoint of heat exchanger to detect the refrigerant charge level. Even though number of sensing points are less, invasive sensing often limits the evaluation of stated approaches to a controlled environment. Consequently, the efficacy relies on the validity of assumptions from controlled environment to a real-world scenario.



(a) *Greina* is a two-step process. The proposed framework first learns the typical temperature profile of the room followed by monitoring the RU for refrigerant leakage.

(b) For each room r, *Greina* maintains a bucket variable (b_r) to label the RU at each hour.

Figure 5-3: [Left] *Greina* takes ambient information from the smart thermostat and weather conditions from a third-party cloud-based weather server to tune the parameters of a lumped thermal model. Tuned thermal model simulates the room temperature for leakage detection. To incorporate temporal changes in the environment, *Greina* regularly updates the model parameters. [Right] From data, we observed that if the temperature within a room is consistently beyond the estimates for 36 hours, then there are high chances of refrigerant leakage in the RU.

5.3 Approach

In *Greina*, every outlet goes through a two-step process - 1. learn the *normal* behaviour, and 2. monitor the refrigeration unit for leakage, as shown in Figure 5-3.

5.3.1 Learn Normal Behaviour

In a typical room r, the change in temperature, in a given time interval τ (in seconds), primarily depends on the heat transferred by the weather outside the room, the heat added by the open door, and the heat extracted by the refrigeration unit (Equation 5.1).

$$\frac{(T_r^{(t+1)} - T_r^{(t)}) \times C_r}{\tau} = K_e^r \times (T_e^{(t)} - T_r^{(t)}) + (Q_{dr} \times S_{d_r}^{(t)}) + (Q_{ru} \times S_{ru_r}^{(t)}) + \eta_r \tag{5.1}$$

Here, $T_r^{(t)}$ and $T_e^{(t)}$ denote the average room and external temperature in the last time interval τ (between t and $t-1$), respectively. C_r is the thermal capacity of the room, and K_e^r is the heat transfer coefficient between room and external weather conditions. We derived the model from Bacher et al. [10].

When supply in the front runs out, the manager opens the room to get the fresh stock and $S_{d_r}^{(t)}$ denotes the state of door (open/close) at time t. Given that RU is a two-state appliance - 1. compressor on, and 2. compressor off, where compressor is the major power consuming component of RU, $S_{ru_r}^{(t)}$ is the state of RU at time t. Correspondingly, Q_{dr} and Q_{ru} denote the amount of heat added through the door, and heat extracted by the refrigeration unit, respectively. η_r is the thermal noise introduced by the random events.

$$\begin{aligned} \mu_r = 1 - \frac{K_e^r * \tau}{C_r}, \quad \mu_e = \frac{K_e^r * \tau}{C_r}, \quad \mu_{dr} = \frac{Q_{dr} * \tau}{C_r}, \\ \mu_{ru} = \frac{Q_{ru} * \tau}{C_r}, \quad \eta_r' = \frac{\eta_r * \tau}{C_r} \end{aligned} \tag{5.2}$$

In state space representation, we can lump the parameters (as shown in Equation 5.2) and rewrite the thermal model as Equation 5.3. *Greina* learns the model parameters ($\theta =< \mu_r, \mu_e, \mu_{dr}, \mu_{ru}, \eta_r' >$) through linear regression while using the data streams of room temperature ($\mathbf{T_r}$), external temperature ($\mathbf{T_e}$), door status ($\mathbf{S_{d_r}}$), and RU status ($\mathbf{S_{ru_r}}$). Tuning the parameters of a physics-based thermal model by using the real-world data is called *Grey Box Modelling* and widely practised by the researchers in several domains [37]. As tuning involves sensor data, the model with adjusted parameters represents an approximate thermal behaviour of the room.

$$T_r^{(t+1)} = \mu_r \times T_r^{(t)} + \mu_e \times T_e^{(t)} + \mu_{dr} \times S_{d_r}^{(t)} + \mu_{ru} \times S_{ru_r}^{(t)} + \eta_r' \tag{5.3}$$

Greina tunes model parameters every month through Stochastic Gradient Descent [131]. The online learning ensures that model adapts to recent changes occurring in the environment, without forgetting the existing knowledge. To ensure *clean* data during the learning phase, *Greina* does not consider those days which were identified as leaking while monitoring the RU. For instance, if *Greina* labelled refrigeration unit as *Refrigerant Leaking* for five days in the last month, the learning module won't consider data from those five days to tune the model parameters. Rationale behind this is that if RU was unable to keep room temperature within the limits for more than 36 hours consecutively, it was not an ideal situation and there was a problem. During those five days, the problem could have been Refrigerant Leakage (true positive), or an ongoing maintenance (false positive); either way, those five days are abnormal for the outlet.

Initially, we specified two key features of *Greina* - 1. it uses readily available information

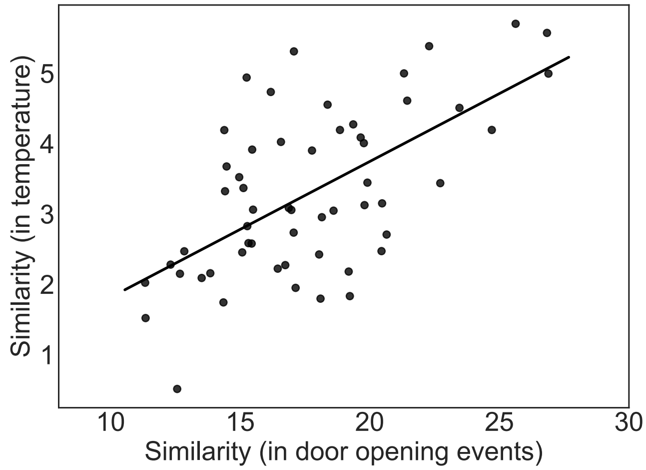


Figure 5-4: If the managers of two rooms have 'similar' routine, then there are good chances that the hourly temperature profile of those rooms will also be 'similar'.

from the smart thermostat, and 2. it can start monitoring the RU without waiting for the sensor data for a long time. Though the proposed framework can collect room temperature ($\mathbf{T_r}$) and door status ($\mathbf{S_{d_r}}$) information from the thermostat, and external temperature ($\mathbf{T_e}$) from any third-party weather server, the refrigeration state ($\mathbf{S_{ru_r}}$), which typically requires appliance-level monitoring is unavailable to the framework. In addition to that, waiting for *clean* data for a long time in the initial stages of deployment can significantly delay the monitoring process. *Greina* leverages domain advancements to fulfill these requirements.

State of Refrigeration Unit

Jain et. al. [78], proposed a classification based algorithm to determine compressor state from the room temperature. The intuition was to first identify compressor on and off events from the temperature data followed by event sequencing to estimate the state vector for RU. When compressor turns on, the room temperature quickly goes down due to the addition of cold air, and when compressor turns off, the room temperature shoots up due to thermal leakage. The estimation algorithm (EA) uses *k-means* to segregate sudden increase and sudden decrease in room temperature (in τ time) from normal variations. On the segregated events, the algorithm applies sequencing and recreates the compressor cycles of the refrigeration unit. We employ estimation algorithm to represent RU state as a function of room temperature ($\mathbf{S_{ru_r}} = EA(\mathbf{T_r})$) and learn the model parameters. One must note

Algorithm 3: Ranking Algorithm

```
/* D_r^h indicates the median of total number of door opening instances
   occurred in the h^th hour for the r^th room across all the days.
   h ∈ [1,24]                                                          */
```

Input: $\mathbf{D_r}$ where r is the ID of new room

Output: Parameters $< \mu_r, \mu_e, \mu_{dr}, \mu_{ru}, \eta'_r >$ of most 'similar' outlet (room r^*)

```
/* Score indicates the `similarity' in the door opening patterns of
   two rooms across the day.                                           */
```

Compute Score:

$S_i = \|\mathbf{D_i} - \mathbf{D_r}\|^2$, $\forall i \in$ {available outlets};

```
/* Select the room number with minimum score                           */
```

Select 'Similar' Outlet:

$r^* = \underset{i}{\operatorname{argmin}}(\mathbf{S})$

that *Greina* estimates state vector only when RU is working fine because learning requires *clean* data.

Parameter Initialization

When a store comes under monitoring, there exists no data to learn the model parameters for the particular outlet. To begin monitoring, we explored correlation between the 'similarity' in manager's daily routine and the hourly temperature profile across the day (Figure 5-4). We designed a ranking algorithm (Algorithm 3) that sorts all the available stores based on the 'similarity' in the daily routine of their respective managers. The framework selects the most 'similar' store and uses its parameters to monitor the refrigeration unit for leakage detection until sufficient clean data (from the new outlet) is available to update the model parameters. We use $l2$-norm to measure the similarity. We do not use room temperature to measure 'similarity' because it is possible that RU is faulty at the time of installation. When there exists no 'similar' outlet, the monitoring module (of *Greina*) uses 10°C as a default threshold [3].

5.3.2 Monitor for Leakage Detection

At the end of each hour, the tuned thermal model estimates typical temperature profile of the room ($\mathbf{T_r}$), for the framework to look for refrigerant leakage. However, it is infeasible to

[3] The refrigeration units are supposed to maintain a temperature range of 5°C-8°C in the cold-rooms. Typically, temperature remains higher (than the limits) during the working hours, and within the limits during the non-working hours. The 10°C is the median value for both working and non-working hours.

estimate the temperature profile which perfectly aligns with the actual temperature profile, and the reason lies in the control strategy embedded within the internal thermostat of a typical refrigeration unit. The inbuilt thermostat of an RU uses *on* hysteresis as an upper threshold and *off* hysteresis as a lower threshold. When room temperature goes beyond the *on* hysteresis, the thermostat turns on the compressor, and the room temperature starts decreasing. Subsequently, when the room temperature reaches the *off* hysteresis, the thermostat turns off the compressor and allows room temperature to increase up to *on* hysteresis. However, in real-world, the temperature readings from inbuilt thermostat of RU differs significantly from the deployed sensor measurements. As it is difficult for the external temperature sensor to match the temperature readings as sensed by the inbuilt sensor, it is also hard to accurately predict room temperature at every time instance. If we observe deviations in the raw temperature profiles ($\mathbf{T_r}$ and $\tilde{\mathbf{T}}_{\mathbf{r}}$) for anomaly detection, the misalignments will lead to false conclusions [4].

Instead, *Greina* analyses hourly mean temperature to mark if the RU in a particular hour is anomalous or not. The reason is that the cycling helps the thermostat in achieving a temperature which is the average of *on* and *off* hysteresis. For instance, in the current scenario, outlets use 5°C as *off* hysteresis and 8°C as *on* hysteresis to maintain an average temperature of 6.5°C in the room. Though the temperature at any time instance may misalign, the average temperature in that hour should remain close to 6.5°C. Thus, Equation 5.4 first computes mean actual (T_r^h) and estimated ($\tilde{T}_r^h$) temperature followed by Equation 5.5 to determine the decision boundary for leakage detection in hour h.

$$T_r^h = \frac{\sum_h \mathbf{T_r}}{(3600/\tau)}, \quad \tilde{T}_r^h = \frac{\sum_h \tilde{\mathbf{T}}_{\mathbf{r}}}{(3600/\tau)} \tag{5.4}$$

$$\hat{T}_r^h = \tilde{T}_r^h + \sigma_r^h \tag{5.5}$$

Here, $\tilde{T}_r^h$ is the estimated mean temperature, σ_r^h is the standard deviation in $\mathbf{T_r}$ in an hour h, and $\hat{T}_r^h$ denotes the estimated decision boundary for the particular hour h and room r. If the sensed mean temperature (T_r^h) is beyond the estimate ($\hat{T}_r^h$), then we mark that specific hour anomalous (Algorithm 4). Furthermore, to gain confidence that deviations are due to leakage, *Greina* applies CUSUM (Cumulative Sum) control strategy [162] and main-

[4] We discuss misalignment in detail in Section 5.5 - Discussions

Algorithm 4: Update Bucket

```
/* Sensed and Estimated Room Temperature                    */
Input: T_r^h, T̂_r^h, h_mon, an_lock, b_r
Output: b_r
/* Update Bucket Value                                      */
if T_r^h > T̂_r^h then
    b_r = b_r + 1;
    an_lock = h_mon; // Reset the lock
else if T_r^h ≤ T̂_r^h then
    an_lock = an_lock − 1;
    if an_lock == 0 then
        b_r = 0; // Reset the bucket
    else
        b_r = b_r − 1;
```

tains a bucket variable (b_r) to monitor consecutive such anomalous instances. Whenever room temperature goes beyond the decision boundary, the monitoring module increments the b_r value by one (Algorithm 4). For every consecutive hour, when room temperature is within the estimated limits, b_r decreases by one unit. If room temperature stays below the decision boundary for h_{mon} hours, *Greina* assumes RU is working fine and resets the bucket. In case of missing information, b_r value remains unchanged.

Label the Refrigeration Unit

In the final step, *Greina* labels the refrigeration unit based on its bucket value (b_r) in the particular hour h. Our analysis indicates that if the room temperature is beyond the estimated decision boundary for consecutively 36 hours, then there are maximum chances of refrigerant leakage. Though we learned the current transition thresholds from the data, we can always adjust these settings based on user requirements.

5.4 Evaluation

For the study, we deployed a customized thermostat and power meter across 74 outlets of a retail enterprise. From July 2016 until June 2017, thermostats collected room temperature and door information from the cold-rooms, and energy meters gathered appliance level power consumption. Every cold room is 9 ft wide, 9 ft long, and 8 ft high. With thick walls and doors, these highly insulated cold rooms can maintain storage temperature anywhere between −2°C and 5°C. We keep thermostats close to the blower fan (of the RU), because that allows *Greina* to focus more on the output of refrigeration unit and less on the thermal noises in the environment. We also monitored the weather conditions (in the region) through an API of a cloud-based weather service [160]. All the outlets were located in a city where outside temperature is usually between 21°C and 37°C, around the year. The retail enterprise maintains a log of calls from the store managers regarding the complaint in their cold-rooms. We used the fault logs for ground truth verification. During the study period, repair person identified 42 cases of refrigerant leakage across 39 outlets, in addition to several other defects. In remaining 35 stores, neither manager nor repair person mentioned any instance of refrigerant leakage in the logs during the study period. In one instance, the outlet manager called a local repair person instead of reporting the fault to the authorised maintenance contractor.

5.4.1 Evaluation Metrics

We evaluate *Greina* primarily on two aspects - 1. model accuracy in estimating the decision boundary ($\widehat{\mathbf{T_r}}$), and 2. minimising the delay in reporting refrigerant leakage.

Modelling Error

For each hour, we compute mean absolute deviation (denoted by e_h) in measured ($\mathbf{T_r}$) and estimated temperature ($\tilde{\mathbf{T}}_\mathbf{r}$) to quantify the accuracy of model parameters in simulating the room temperature in an hour h (Equation 5.6).

$$e_h = \frac{\sum_h |\tilde{T}_r^{(t)} - T_r^{(t)}|}{(3600/\tau)} \tag{5.6}$$

$$rd_m = dt_m - dt_s \quad rd_g = dt_g - dt_s \tag{5.7}$$

Delay in Reporting the Leakage

We labelled all the leakage instances by start date (dt_s) - when symptoms became visible in the data, and end date (dt_e) - when repair person repaired the leak. Reporting delay is the number of days between the start and the leakage reporting dates. Equation 5.6 computes reporting delay for store manager (rd_m) and *Greina* (rd_g). Here, dt_m and dt_g are the dates

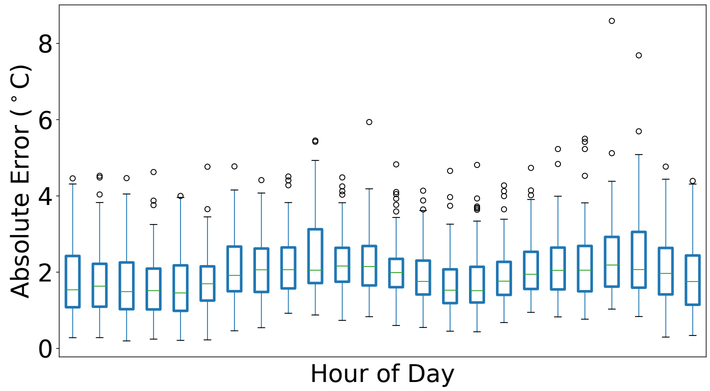


Figure 5-5: During working hours, the outlet managers are much more noisy than the non-working hours.

when the manager and *Greina* reported the leaks, respectively.

5.4.2 Model Validation

Across all the outlets, our analysis on clean data (when RU was fit) indicates that model can simulate the room temperature with a mean absolute error of 2°C (with $\sigma = 0.9$°C), as shown in (Figure 5-5). Estimation error is primarily due to sudden thermal noise by the random events - such as leaving the door open, refilling the food products at a higher temperature. For the same reason, the error is usually higher during the operational hours, as also evident from the bumps. Erroneous estimations might mislead *Greina* that room temperature is above the estimated temperature, but adding standard deviation in Equation 5.5 minimises such instances.

5.4.3 Results and Analysis

In the study period, we noted 42 instances in the logs where either the store manager or the repair person mentioned the keywords - *gas*[5] and *leakage*. In 4 cases, though the records had the *gas* keyword, no leakage occurred in the refrigeration unit at that time, as validated by the store manager and the data. Furthermore, refrigerant leakage shows physical symptoms, such as water dripping and ice formation, which are visible before any leakage pattern in the temperature data. In 3 such instances, outlet managers quickly intimated the maintenance contractor to fix the refrigeration unit, and *Greina* had no data to analyse the symptoms for

[5]local people often use the word gas to refer refrigerant

leakage detection. In remaining 35 instances, the proposed framework correctly identified 25 leaks and failed to detect leakage in the remaining scenarios. While few of them were genuine system failures, others happened due to ungovernable circumstances -

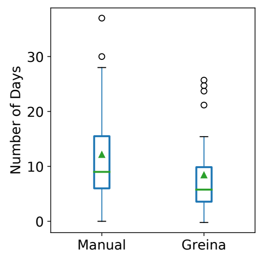
1. **Too Early to Detect:** During the learning phase, *Greina* borrows the model parameters from a 'similar' outlet. While *Greina* successfully identified six leaks when knowledge transferred from one store to another, it failed twice.

2. **Improper Learning:** Outlet managers generate significant thermal noise in the room through their dynamic and random activities. While the model can deal with such noise with a substantial amount of data, *Greina* incorrectly learns the model parameters when the data is insufficient. Moreover, as the occupants' behaviour differs significantly across the outlets, there exists no definite way to compute - *How much data will suffice to train the model correctly?*

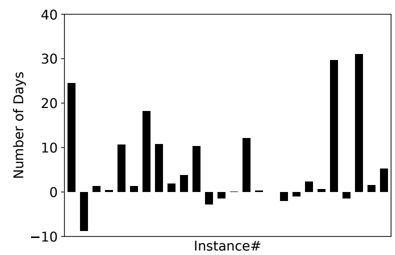
3. **Noisy Sensor:** In two cases, we noticed that a faulty sensor was providing incorrect temperature readings which resulted in false negative. However, we believe that these issues are solvable with better governance around the deployment.

4. **Low on Refrigerant:** Quite often, a refrigeration unit is actually low on refrigerant (due to heavy usage), and there exists no leakage. In such scenarios, though temperature remains in a higher range as RU is not running at the full capacity, but differs from the temperature patterns as in the case of refrigerant leakage. Therefore, when RU was low on refrigerant in 3 such instances, the technician mentioned *gas top up* in the log; however, *Greina* failed to find the symptoms of refrigerant leakage.

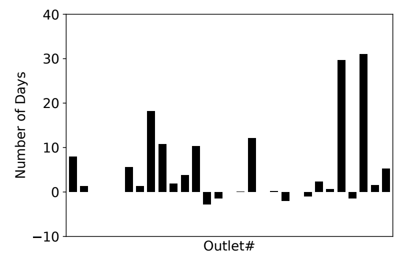
Furthermore, *Greina* reported only 6 instances of false positives - marked RU as leaking while it was working fine. In four out of six cases, either store manager or technician shut down the refrigeration unit for construction and the repairing work. As the shutdown was unexpected, *Greina* confused the rise in room temperature with leakage and marked the refrigeration unit as leaking. In remaining two instances, RU broke abruptly due to an electrical fault and stayed down until a technician came to fix it. Consequently, room temperature went high and *Greina* raised a refrigerant leakage flag. As the company, maintenance contractor, and the outlet managers are usually aware of these situations; we believe these flags were harmless. In addition to this, *Greina* also pinpointed a case of



(a) *Greina* performed better or comparable to manual reporting in leakage detection.



(b) In 25 instances both *Greina* and manual reported correctly identified the fault. In 19 out of 25, *Greina* detected the leakage before store manager.



(c) Due to delay in fixing the fault and backlog complaints, *Greina* detected the fault after manual logging in six instances.

Figure 5-6: Our analysis indicates, a simple yet powerful framework, *Greina* reduced the average leakage reporting delay by 5-6 days, significant enough to avoid energy wastage and maintain food quality by timely repairing the refrigeration unit.

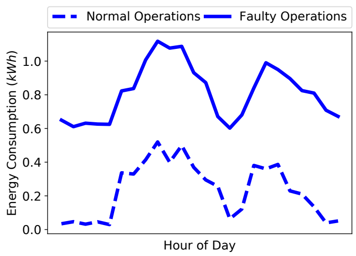
refrigerant leakage where store manager called an unauthorised technician to repair the RU and didn't notify the enterprise. Though such cases are rare, they are a serious concern for the company.

Beyond Accuracy

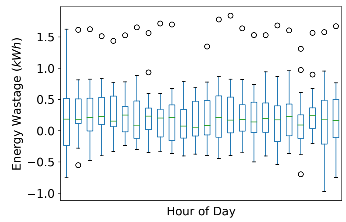
Though accuracy is essential, minimization of reporting delay is a more significant concern for the stakeholders because the delay is directly proportional to energy wastage, health hazards, and product wastage. In our analysis, we noted an average reporting delay of seven days for *Greina*, while managers had a mean reporting delay of 12 days (Figure 5-6a). In 19 out of 25 instances, *Greina* detected leakage before the store manager (Figure 5-6). In eight cases, the difference was as high as 10-30 days. The early detection of leakage exhibits several benefits.

Minimize Energy Wastage

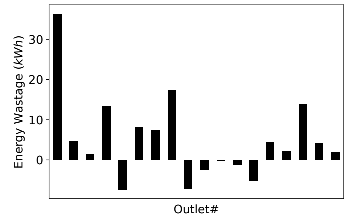
In Section 5.1, we presented a scenario (Figure 5-1) when refrigeration unit had a gas leakage, and outlet manager kept using the RU for more than a week. While such situations are common across all the stores, we noticed significant energy wastage in doing so. Figure 5-7a compares hourly energy consumption during the normal operations with faulty operations, when RU has a leakage. There are two essential takeaways - 1. energy consumption increases significantly during the working hours (almost $4x$), and 2. energy consumption is



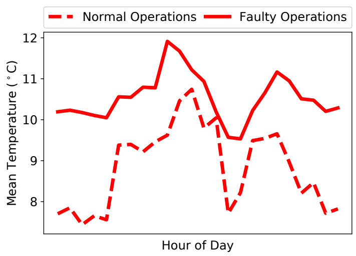
(a) Leakage increases the power consumption of RU by $4x$ and $2x$ during non-working and working hours, respectively.



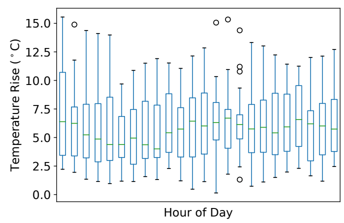
(b) Within reporting delay period (d_m^g), outlets wasted significant energy in both working and non-working hours of the day.



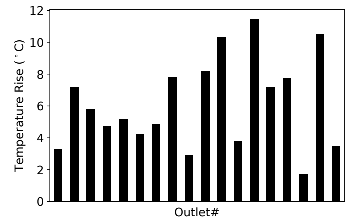
(c) The company could have saved 5-6 kWh energy every day during d_m^g period which is twice the daily consumption.



(d) When compared with normal operations, the temperature increased by 2°C-5°C across the day, during refrigerant leakage.



(e) The increase in temperature is evident across all the outlets; thus, risking perishable items every day during the reporting delay.



(f) Early detection of refrigerant leakage by *Greina* had kept the rooms 5°C-6°C colder during the reporting delay period (d_m^g).

Figure 5-7: Beyond being accurate in detecting the gas leakages, *Greina* can also save energy and keep the room 5°C-6°C colder, every day during the d_m^g period - the number of days between manual reporting and leakage detection by *Greina*.

very high when refrigeration unit is leaking (around $7x$ during the non-working hours).

$$d_m^g = dt_g - dt_m \tag{5.8}$$

While the activities (or routine) of store managers are hard to change, *Greina* seems robust enough to minimize the energy wastage due to gas leakage. Figure 5-7b compares the hourly energy wasted in the d_m^g period - the number of days between the reporting dates of manager and *Greina* (Equation 5.8). Our analysis indicates that if central maintenance team had taken the recommendations from *Greina*, they might have saved 5-10 kWh energy every day (when RU was faulty) which is twice the typical daily power consumption of RU (Figure 5-7c). The negative wastage depicted those scenarios when RU stopped working, and the outlet ended up consuming less energy in comparison to normal operations.

Minimizing Risk to Product Quality

Next, we observed that room temperature remains significantly high during the gas leakage (Figure 5-7d). High temperature risks the product quality and impacts the store operations. The average room temperature increased by 2°C-3°C in both working and non-working hours. When we analyzed across all the 19 instances where *Greina* reported before the store manager, we noted a median increase of 6°C during the d_t^g period for each hour of the day, across different outlets (Figure 5-7e). If maintenance team had taken the recommendations from *Greina*, they could have kept the rooms 5°C-6°C colder every day (when RU was faulty), by timely repairing the RU (Figure 5-7f).

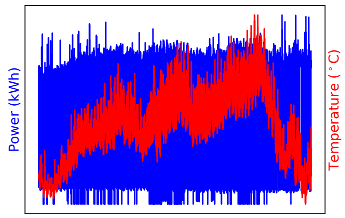
Even though store managers have a benefit of observing the physical symptoms of leakage, *Greina* identified 19 out of 25 leakages before the store manager. Though in two instances, *Greina* couldn't detect the leakage before manager; we also noted that in remaining three cases, it occurred due to the negligence of the maintenance contractor. While outlet manager timely reported the leakage to the maintenance team, the maintenance contractor didn't take any action for several days. Consequently, the symptoms became visible and *Greina* detected the leakage even though store manager repeatedly complained about the leakage.

Beyond Leakage

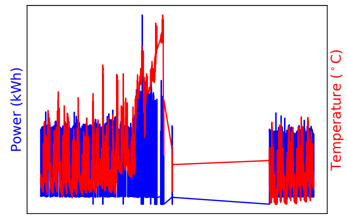
Though we designed *Greina* for refrigerant leakage, we observed that the proposed framework also identified other common time-varying faults.

Ice Formation

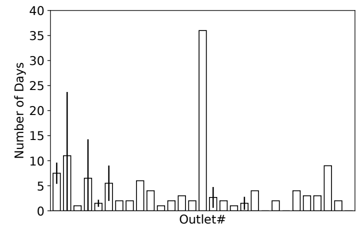
Dirty air filters, defective evaporator or condenser fans, lack of refrigerant often results in the formation of ice in the refrigeration unit. The ice obstructs the path of cold air and the refrigeration unit works at a reduced efficiency. Though the symptoms of ice formation in temperature data may depend on the fault, Figure 5-8a depicts one scenario where the increase in temperature is akin to refrigerant leakage (Figure 5-1). If the ice forms due to heavy usage or a dirty filter, store managers usually clean the filters, however, if the ice forms due to a fault, ice keeps on forming even after cleaning and manager needs to call the technician to repair the RU. *Greina* identified 18 such instances where ice formed due



(a) When ice forms due to other reasons (such as dirty filters), the peak power consumption remains consistent.



(b) Before failure, motor within the RU draws high current which increases the peak power consumption of RU by $1.5x$.



(c) Even though outlet managers timely report the leakage, the technicians introduce a significant delay in repairing the RU.

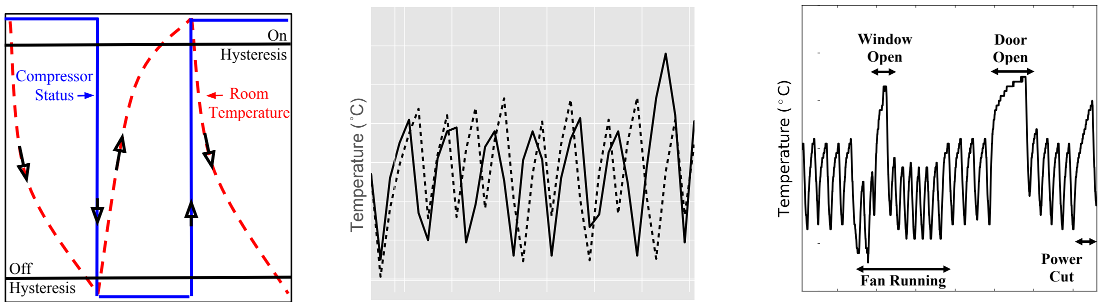
Figure 5-8: [Left and Middle] In addition to leakage, *Greina* also reported 28 instances of ice formation and motor failure during the study. [Right] *Greina* can be extended to monitor the efficiency of maintenance contractor in repairing a reported fault.

to dirty air filter or defective fans.

Condenser Motor Failure

The job of the condenser is to cool the high-pressure refrigerant gas received from the compressor by moving outside air across the condenser coils. Through condensation, the high-pressure, high-temperature refrigerant gas changes to low-temperature liquid refrigerant. However, due to wear and tear, and high temperature during the summers, the condenser fan motor usually fails and RU stops cooling the room. In an attempt to cool the room during the motor failure, RU starts drawing more current which increases the peak power consumption of RU and heats up the system further (Figure 5-8b). Eventually, RU breaks down, and manager calls a technician to repair the motor. *Greina* detected ten such instances of motor failure and raised the alarm 1-2 days before complete shutdown of the RU.

To conclude, our analysis on a reasonably rich dataset collected from actual field deployments, indicates that *Greina* possesses the power to *timely* identify the refrigerant leakage instances. Beyond being accurate, on an average, *Greina* also proved to reduce the reporting delay by five to six days. The improvement in reporting delay can minimize the energy wastage and maintain desired temperature for the stored items. Interestingly, to achieve these benefits, the company needs to only upgrade their traditional thermostat to a smart thermostat with *Greina* running in the cloud and leveraging the data collected from the thermostat. Our study indicates that with smart thermostats, the simple yet powerful



(a) For a compressor cycle, the room temperature stays within *on* and *off* hysteresis.

(b) While hard line depicts recorded temperature signal, dashed line represents estimated temperature signal.

(c) Stochastic user activities (especially in residential apartments) require time-varying noise in the thermal model.

Figure 5-9: [Left and Middle] Inability to accurately estimate room temperature close to *on* and *off* hysteresis leads to misalignment in estimated and recorded temperature signals. [Right] Current model is unaware of dynamic user activities and assumes constant thermal noise in the room. It is another major source of modelling error, especially in residential apartments.

framework, *Greina* is easily scalable.

5.5 Challenges and Discussion

In this chapter, we designed *Greina* for leakage detection by only sensing the ambient information from the smart thermostats. In collaboration with an energy-analytic firm, we are currently focused on the deployment of *Greina* across all the 74 outlets for rigorous evaluation of the system in real-world. In this section, we discuss three possible dimensions to extend the proposed framework.

5.5.1 Modelling Error

We rely on our thermal model to first learn a decision boundary for *Greina* to monitor the RU for any leakage on a daily basis. Though current model estimates with a mean absolute error of 2°C (with $\sigma = 0.9$°C), the reader should also be mindful of the fact that the model is a replaceable module of the whole framework. In Appendix C, we present one such extension of current thermal model which can estimate room temperature with an RMSE of 2.85°C (with $\sigma = 1.3$°C) even in a highly noisy environment of residential apartments. On further analysis, we noticed that following are the primary sources of error.

Misalignment

When we turn on the AC (also applicable on RU), the compressor which is the major power consuming component of the AC also turns on, and starts putting cold air into the room. Due to the flow of cold air, room temperature drops up to a certain level - *off* hysteresis (as depicted in Figure 5-9a). At this point, compressor turns off and allow room temperature to increase up to a certain level - *on* hysteresis. As temperature reaches the *on* hysteresis, compressor turns on again and starts cooling the space.

The control algorithm of AC decides compressor state based on the return air temperature which is measured inside the AC, near the filters. Let's call it T_{rat}. On the other hand, we install temperature sensor just outside the AC, near the fan. Now, let's say T_r represents room temperature as measured by the sensor and $\tilde{T}_r$ depicts the estimated room temperature. As per the control logic, the controller will turn off the compressor when $T_{rat} \geq T_{off}$. However, given the complex non-linear thermodynamics, it is impractical to precisely estimate T_{rat} or T_r at any time instance; thus, $T_{rat} - T_{off}$ and $T_{rat} - T_{on}$. As a result, $T_{rat} - T_{off} \neq \tilde{T}_r - T_{off}$ and $T_{rat} - T_{on} \neq \tilde{T}_r - T_{on}$. Basically, though the estimated room temperature follows the pattern, it fails to align perfectly with the measured room temperature. As depicted in Figure 5-9b also, the actual (hard black line) and estimated (dashed black line) temperature signals are misaligned. To minimize the effect of error due to misalignment, we consider hourly mean temperature and add standard deviation to be more conservative while monitor the RU for refrigerant leakage.

Constant Thermal Noise

Another major reason of error is the thermal noise due to numerous dynamics activities of the occupants. Especially in a residential apartment, multiple activities occur at distinct time intervals during the AC usage (as shown in Figure 5-9c). However, in current implementation (Equation 5.1 and Equation C.1-C.4), we assume thermal noise is constant at any time. As the (current) model is unaware of such dynamic activities, it will compute room temperature while assuming constant noise from these sources; thus inflating the error. While constant noise makes little impact on estimating the room temperature in cold rooms (space is highly insulated and activities are relatively less), it makes significant impact on the temperature estimation for the apartments. One way to handle noise due to dynamic activities is to first identify the activity (based on temperature, occupancy, and time of the day), and then compute thermal noise for that particular time instance. Having said that, we repeat that the model is a replaceable module of the whole framework; thus, community is encouraged to explore other such variants of the current model to enhance *Greina*'s performance.

5.5.2 False Negatives & Positives

We discussed multiple scenarios when *Greina* generated both, false negatives and false positives. After scrutinizing those situations, we noted that we can avoid many of those instances through user-friendly and interactive interfaces (or advanced notification systems). Basically, by involving users, we can empower *Greina* with a two-way communication with the outlet managers for accurate estimation of the refrigerant leakage. In several instances, managers spotted unusual events, such as high temperature or ice formation. If they could have notified *Greina* through a device, *Greina* would have used the information to detect the leak, even early. Similarly, we discussed multiple cases where store manager timely reported the leak but the repair person (or the maintenance contractor) only refilled the gas instead of fixing the leakage. Consequently, refrigeration unit again went down in a couple of weeks and company had to suffer from the business loss, in addition to usual consequences of refrigeration leakage. If *Greina* could validate existence of anomalous pattern even after the corrective action, the maintenance company could reassess the appliance. Not only the two-way communication would minimize such instances, it will also develop a sense of trust for the system.

Having said that, one must also remember that user attention is costly. When users are involved, the line between a useful system and an annoying system is typically very thin. User input is useful only if the system is designed while considering -

1. How frequently user should input the information?,
2. How much time does it take the user to input the information?,
3. How frequently system should notify the users?,
4. Is the interface intuitive and user-friendly?

For the same reason, the new frontiers opening in the domain of smart and ambient notification and attention management systems, the design of web and mobile application, and many such factors can significantly influence the outcome of this study. In future, we plan to implement the proposed framework (along with the user feedback) to critically evaluate the effectiveness of *Greina* in achieving the desired goals at a much larger scale.

5.5.3 Beyond Refrigeration

Beauty of *Greina* lies in its modular and systematic architecture, especially the replaceable thermal model, and ability to work on top of smart thermostats. While *Greina* can benefit from readily available information from the thermostat (temperature, occupancy), it can be adapted to diverse environments by tweaking the proposed thermal model. In Appendix C, we discuss one such extension to a noisy home environment. Poor thermal insulation and time-varying stochastic activities of people are the primary two basis of thermal noise in the residential apartments. However, in a separate analysis, we found that a non-linear thermal model can estimate room temperature with an RMSE of 2.85°C (with $\sigma = 1.3$°C).

Though it is feasible to extend *Greina* for the home environment, the real challenge lies in evaluating the efficacy of proposed framework in such an environment. In residential apartments, tenants call local technicians, and typically no fault logs are available. Ground truth data collection (at a large scale) demands enormous support from tenants and the technicians in sharing the information whenever there is a fault. For the same reason, several studies in the past have either evaluated their approach in theory, or in a controlled environment. We believe a comprehensive evaluation of *Greina* across residential apartments can bring up additional intriguing insights about both, the environment and the system.

Chapter 6

Conclusion and Future Work

Intelligence is the ability to adapt to change.

– Prof. Stephen Hawking

While rising temperature, across the globe, is making air-conditioning a necessity rather than a luxury; growing income and reducing prices of hardware components are making them affordable, especially in emerging hot economies like India. As of now, only 5%-8% of people in India use air conditioners. However, given the current market trend, meeting the electricity demand (especially during the peak demand) will become even more challenging in the future. To fulfill this growing demand in the future, we must ensure energy-efficient usage of the appliance. Since thermostat plays a crucial role in deciding user comfort and AC energy consumption, in this dissertation, we proposed Data-Driven Thermostats - an extension of smart thermostats.

One must remember that the field of optimizing air-conditioning experience through thermostats is decades old. Since the advent of programmable thermostats in 1885, thermostats today are smart, remotely accessible, better in design, and able to suggest optimal settings by learning occupants' schedule. Google Nest [118] is one such realization of smart thermostats for residential spaces. While thermostats are smart, they still rely on the dynamic and stochastic schedule of the occupants to optimize AC usage. To learn the daily schedule, smart thermostats need that tenants follow a particular schedule during the training period. Often, thermostat learns an irregular schedule and any deviation from the learned schedule leads to discomfort. In a study, Yang et al. [166] showed that the energy saving features of the Nest, such as Auto-Schedule and Auto-Away, were neither effective in saving energy, nor in maintaining the user comfort.

On the contrary, Data-Driven Thermostat fundamentally relies on temperature signal, which is relatively stable than occupants' schedule. The proposed framework takes sensory data from the thermostat to tune the lumped parameters of a thermal model, which eventually serves the application layer. With time, as the framework collects more and more data, it keeps updating the model parameters to accommodate temporal changes in the building's thermal behavior. The proposed framework is compatible with existing infrastructure (scalable), require no additional instrumentation (low-cost), and need minimal user intervention (user-friendly). Our analysis, based on real-world data, indicates that Data-Driven Thermostats can provide actionable energy-feedback, dynamically vary the set point temperature for personalized and localized comfort, and timely identify leakage to ensure reliable functioning of the appliance. In this chapter, we first conclude our thoughts on the proposed framework and later present our outlook on the future of thermostats.

6.1 Feedback

In several parts of the world, air-conditioning is still a luxury item, and the reason is, in any typical home, AC contributes a significant proportion towards the monthly electricity bill. Since the information is typically available at the end of the month, it is hard for the occupants to reason, *"Why?"*. Often, appliance-level plug monitors are unavailable in high current ratings, required to handle large transient current flow associated with AC switching on events. Even if such a plug monitor exists, either it is expensive, or need technical expertise for installation, or could not communicate over the network for real-time analysis. Therefore, we started exploring indirect ways of appliance-level monitoring and noticed a strong correlation between temperature and power signal.

We conducted an in-situ deployment in seven different apartments and asked the tenants to follow their daily routine. In every home, we installed a smart meter to monitor AC power consumption for the ground truth verification and a temperature sensor in the master bedroom for data modeling. In the evening, we used to conduct a brief survey for ground truth verification, in which, we would ask simple questions like, *"At what times did they use the AC?"*, *"What was the setpoint temperature during each usage?"*, *"How many people were present during the usage?"*. Based on the insights from the data, we designed PACMAN - a feedback system that only senses the room temperature to estimate and predict AC energy consumption for the users. Given that the desired information is readily available through smart thermostats, our research testifies that PACMAN is a low-cost and scalable extension to smart thermostats for providing actionable energy feedback. Our retrospective evaluation indicates that, on average, PACMAN can estimate and predict AC energy consumption with an accuracy of 84% and 85%, respectively.

6.2 Comfort

> "Ahmm, setting a thermostat temperature is quite challenging task. I come home, feeling very hot and start the AC at 23... then in some time it becomes very cold and I turn it off... and after some time, turn it on again (may be at a higher temperature). You never know what to set AC on.... therefore, during night I fix my AC at 24-25 and use blanket... at some point I feel cold and take the blanket, after few minutes I start feeling hot and remove the blanket. Though, it's annoying but that's the best way I have figured out till now..." - Participant

User comfort is subjective; it differs from one person to another. Under particular weather conditions, user comfort might depend upon context, environment, and user activities. With this thought in mind, we designed PACMAN to estimate and predict AC energy consumption at different setpoint temperatures; so that, the tenant could decide *right* thermostat temperature for themselves. However, weather condition changes with time, and a feedback-based system would require tenants to regularly update the thermostat temperature - highly inconvenient for the occupants. For instance, during night time, when occupants are sleeping, they might not bother to change the thermostat temperature; instead set AC to a lower temperature and take a blanket when required.

Ideally, a thermostat should automatically set a temperature that would suffice user comfort and also save energy. However, in doing so, thermostats should not be limited to thermostat temperature, instead provide users an option to specify their goal - minimize AC energy consumption, or maximize user comfort, or maintain a balance between them. Eventually, based on user preference, the thermostat should change the setpoint temperature to achieve the desired trade-off between comfort and energy. To address this concern, we added a Comfort-Energy Trade-off (CET) knob on the mobile application of smart thermostats, where residents can specify their intention (max. comfort/savings), along with the thermostat temperature. Next, we designed and developed an optimization framework that automatically tunes the parameters of a generic thermal model to a particular room environment. Subsequently, the framework takes tuned thermal model along with the user input (set temperature and CET value) and current weather conditions to dynamically vary the AC thermostat temperature.

Our study indicates that the proposed framework is practical with data from a single location in the room (somewhere close to the AC); thus, possess great potential to scale up with existing infrastructure. Portable+ thermostat, empowered with the estimates of temperature across different regions, can provide region-specific comfort when compared to smart thermostats. Our evaluation based on the data collected from a controlled experiment (spanning two rooms for over two weeks) and an in-situ deployment testified its efficacy in reducing occupants' discomfort by 23% (when maximizing the user experience), and maximizing energy savings by 26%, during the power-saving mode. Overall, we found that dynamically varying the setpoint temperature to maintain comfort energy tradeoff (based on user preference) is feasible (through any smart thermostat), user-friendly (require minimal or no manual intervention), and effective (as evaluated based on real-world data).

6.3 Reliability

Unlike air-conditioning in residential apartments, people often prefer centralized heating, ventilation, and air-conditioning (HVAC) units in commercial buildings. Typically, in an HVAC, there is a chiller unit (or a furnace if we are heating) to supply cold air/water to Air Handling Units (AHUs). The role of an AHU is to maintain the desired temperature in a zone, often consisting of private (such as offices) and shared regions (labs, lobby). In such shared spaces, it is impractical to shut down chillers and AHUs when occupants are not present in their individual spaces. Instead, building managers run HVAC on a setback schedule, in which, HVAC supplies less cold air to the room, which is not in use. Now, to generate an optimal setback schedule to run the HVAC, predictive control strategies for commercial HVAC, often require an estimate of occupancy across the building. Therefore, in such cases, the effectiveness of the control strategy relies on the accuracy of occupancy predictions. High variability in the outcome due to prediction errors, makes predictive control strategy unreliable for commercial buildings. Thus, in the next part, we analyzed the influence of occupancy prediction errors on users' comfort and total energy consumption.

To study the influence of occupancy prediction errors, we first designed a custom-built building simulator (ThermalSim) and systematically introduced realistic errors in the occupancy data. In our analysis, we explored that occupancy prediction error, as high as 20%, can make a model predictive control (MPC) based strategy highly unreliable, as measured on an energy-comfort scale. Next, we noticed that a personal thermal comfort system (like SPOT) could strengthen the predictive strategy by mitigating the influence of predictions errors on energy consumption and occupants' comfort. Our studies shows that it is feasible to use a personal thermal comfort system – SPOT – along with the predictive control strategy to ensure personalized comfort in personal and shared spaces. As SPOT works at a finer time-scale than the MPC-based HVAC, it can even cater to the dynamic schedule of the occupants, and satisfy the comfort requirements of every individual, while mitigating the influence of prediction errors; thus, making predictive controllers reliable even for the commercial HVACs.

Beyond energy-efficient control, another critical aspect of energy-efficient usage of an appliance is its reliable functioning. Usually, aging, irregular maintenance, and improper handling lead to frequent breakdowns of the appliance. We, understanding the benefits of fault detection through smart thermostats, next developed an unsupervised self-learning framework *Greina* , that senses ambient information from the smart thermostat for refrigerant leakage detection. Often, due to the loss of refrigerant, the compressor works with reduced efficiency and takes more time than the usual to cool the room; thus, wasting significant energy. In addition to that, the leakage exposes tenants to the refrigerant, which is extremely dangerous for their health [7]. The proposed technique employs Grey-Box Modelling to estimate a decision boundary and later uses the estimates for leakage monitoring. The performance evaluation of *Greina* on data from 74 stores (from a region in India) indicates that the simple yet powerful framework can reduce the reporting delay by a week with best of around 20-30 days in few instances. During these days, the retail enterprise could have saved twice the energy, AC consumes on a typical day. Moreover, by timely repairing the air-conditioning unit, they could have kept the rooms 5°C-10°C colder every day, when refrigerant was leaking through the AC.

6.4 Outlook

Looking forward, we envision smart thermostats to go beyond their functional limitations and play a crucial role in keeping the occupants comfortable, relaxed, and productive. While this dissertation identified and addressed a few fundamental challenges in developing such a thermostat, there are several exciting dimensions of smart thermostats which are yet to explore. In this section, we discuss two such horizons.

6.4.1 Cognitive Thermostats

It is a well-studied fact that the ambient room temperature effects office productivity [64], social relations [69], a students' ability to learn [83], and sleep [121], among many other psychological and physiological parameters [30, 155]. In a study, Lang et al. [98] showed that workers in a cold room not only make more errors but also give low throughput. In another study, Perez et al. [83] demonstrated how hard it is for the students to focus when the room temperature is too hot or cold. If we carefully observe these studies, we will notice that to achieve optimal output; the cognitive mindset of the tenants must adapt to that environment. Basically, while students might need a learning-centric environment in schools, people at home might prefer a calm and relaxing atmosphere. The thermostats, which can assist occupants in achieving the desired cognitive state, we define them as *Cognitive Thermostats*.

Here, the research question is, *"Given an environment, can thermostats vary the set-point temperature such that tenants can achieve the cognitive mindset, desired for that environment?"*.

Only from the perspective of a thermostat, the problem is hard. As of now, thermostats collect limited information; which is inadequate for the thermostat to comprehend the environment, and psychological and physiological parameters of the occupants. However, with the advent of extensive sensor networks as part of smart homes, each sensing unit is now capable of providing a partial view of the atmosphere. For instance, with limited accuracy, we can capture cognitive load on a person, by monitoring changes in the person's forehead and nose temperature using a commercial thermal camera [1]. Using multiple such sensing units, we believe that it is feasible to get better visibility of the environment by combining the partial view from these sensing units.

In collaboration with MIT Media Lab, we first conducted a preliminary study to quantify the accuracy of the thermal camera in estimating cognitive load in a temperature variant atmosphere. In an office room, we placed an infrared camera along with a visible camera and asked volunteers to complete a few tasks in a given time. During the study, we asked the volunteers to wear a Zephyr BioHarness for ground truth data collection. From our initial analysis, we noticed multiple challenges in estimating cognitive load from temperature changes on the face and the nose in a noisy environment. As of now, we are still working on system development. Hopefully, once the system is ready, we will carry out a large-scale user study to get conclusive evidence for the hypothesis, as above-mentioned.

6.4.2 FDD for Commercial HVAC

Beyond feedback and comfort, another essential aspect of energy-efficient usage of an AC is its fault-free functioning. As discussed in Chapter 5, faulty appliances waste significant energy, compromise user comfort, and in some cases, they can even be hazardous for the occupants. Therefore, it is imperative to implement automated fault detection and diagnosis to monitor the health of the appliance regularly. In this dissertation, we implemented one such fault detection technique to report instances of refrigerant leakage. While there exist multiple leakage detection techniques in the literature, we ensured that our approach utilizes only that information, which is readily available from a smart thermostat. By doing so, we ensured scalability, ubiquity, and adaptability for the proposed framework.

However, if we look at centralized HVAC in commercial buildings, it is impractical to report faults by merely monitoring the room temperature. Since an HVAC unit comprises of numerous components, the faults may arise at room-level, zone-level, or even at the building-level, and it is hard to pinpoint a fault without extensive information about the system. Fortunately, building managers do monitor multiple parameters at different points in the HVAC to maintain the desired temperature in every room.

Now, the question is, *"With limited additional sensing, can building managers use existing sensors to monitor the health of an HVAC unit?"*.

In collaboration with Optimisation and Control Group at PNNL[1], we started exploring an exciting domain of fault detection and diagnosis for HVAC units in commercial buildings; but with the help of existing sensor network. With the help of Adaptive Control Team at

[1] Pacific Northwest National Laboratory, Richland, WA, USA

PNNL, we first simulated the HVAC for a large office building, using EnergyPlus and Dymola frameworks. The building consisted of three floors, with each floor divided into five zones and each zone comprising of both private and shared spaces. In the simulation framework, we monitored 192 parameters, which included temperature, pressure, air-flow, and many other physical attributes at different points in the HVAC unit. In the dataset, there were three faults for the analysis purpose, offset in outdoor temperature sensor, offset in supply air temperature sensor, and cooling coil valve stuck at a particular position.

In our preliminary study, we specifically focused on the problem of cooling coil valve stuck. Currently, we are studying the trade-off between cost, accrued in maintaining the sensor network, and benefits, in terms of accuracy and time in detecting the issue. Our initial investigation indicates that it is practical to identify valve-stuck issue using an ensemble machine learning techniques, by only monitoring ambient temperature at AHU and the zone-level. As of now, we intend to incorporate multiple such faults in our analysis and study different machine learning and deep learning techniques in accurately generating a health report for the appliance, every day.

6.4.3 Complex Environments

Beyond residential and commercial spaces, we noticed one another intriguing application scenario, which is a big room with multi-split air conditioners. Quite often, especially in developing economies, small-scale commercial offices, banks, retail outlets use multiple AC units over a centralized air-conditioning system. Such buildings are generally limited to a large hall on a single floor, where workspace comprises of multiple cubicles. Typically, there is no central authority to control these ACs and tenants use a remote to set the thermostat temperature, for each AC. For such a scenario, there are several open-ended research questions for the community to explore. Since installed ACs typically run on their full capacities without accounting for cooling impact from neighboring ACs serving the same zone, such commercial premises possess tremendous potential for energy savings. Karmakar et al. [87] proposed one such scheduling algorithm to coordinate the compressor cycles of multiple ACs to reduce peak demand load. Therefore, extending smart thermostats for multi-AC premises could be highly beneficial from both the perspectives, energy efficiency, and user comfort.

Beyond personal benefits, a smart thermostat could even be beneficial for Smart Grids, an essential component of the Smart Cities project. Since air-conditioners are one of the most power-consuming appliances in almost every type of building, coordinated scheduling of ACs, at city-level, could benefit grids in managing the peak demand load. Moreover, by applying advanced control strategies, such as model predictive control (MPC), smart grids can even suggest optimal setpoint temperature for different hours of the day. It is even feasible to provide an incentive to the users, for following the recommended setpoint schedule of the ACs. If allowed, the tenants could automatically vary the setpoint temperature to optimize AC usage and ensure user comfort.

While benefits are endless, we still need to explore them and apply them in the real world. For starters, this dissertation introduces the concept of Data-Driven Thermostats and testifies its efficacy through data collected from multiple field deployments. Furthermore, our research validates that Data-Driven Thermostats are practical, compatible (with both old and new generation ACs), scalable, and quite effective. Henceforth, I believe, in the future, further extending the Data-Driven Thermostats will be crucial in altering the way we currently interact with our air-conditioning units and eventually make ACs highly efficient and comfortable.

Appendix A

Generic n-region Implementation of Thermal Model

In Section 3, we presented a thermal model for a room divided into three regions (hir, mir, lir) based on the cooling impact of AC. The shown model is derived from a generic n-region formulation of a thermal model adapted from [157, 56] (Equation A.1-A.4).

$$\frac{(T_{wall}^{(t+1)} - T_{wall}^{(t)}) \times C_{wall}}{\tau} = K_{ext}^{wall} \times (T_{ext}^{(t)} - T_{wall}^{(t)}) + \sum_{r=1}^{r_l} K_{wall}^{r} \times (T_r^{(t)} - T_{wall}^{(t)}) \tag{A.1}$$

$$\begin{aligned} \frac{(T_1^{(t+1)} - T_1^{(t)}) \times C_1}{\tau} =& K_{ext}^{1} \times (T_{ext}^{(t)} - T_1^{(t)}) + K_{wall}^{1} \times (T_{wall}^{(t)} - T_1^{(t)}) \\ &+ K_1^2 \times (T_2^{(t)} - T_1^{(t)}) + \eta_1 - Q_{AC} \times S_{AC} \end{aligned} \tag{A.2}$$

$$\begin{aligned} \frac{(T_r^{(t+1)} - T_r^{(t)}) \times C_r}{\tau} =& K_{ext}^{r} \times (T_{ext}^{(t)} - T_r^{(t)}) + K_{wall}^{r} \times (T_{wall}^{(t)} - T_r^{(t)}) \\ &+ K_{r-1}^{r} \times (T_{r-1}^{(t)} - T_r^{(t)}) + K_{mir}^{r+1} \times (T_{r+1}^{(t)} - T_r^{(t)}) \\ &+ \eta_r, \forall r \in [2, r_l - 1] \end{aligned} \tag{A.3}$$

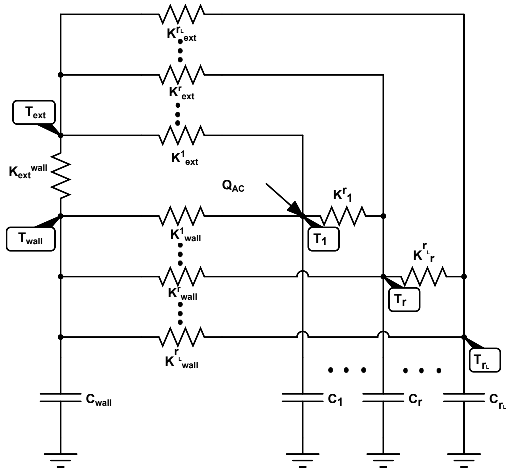


Figure A-1: Resistive-Capacitive Network for the Generic Implementation

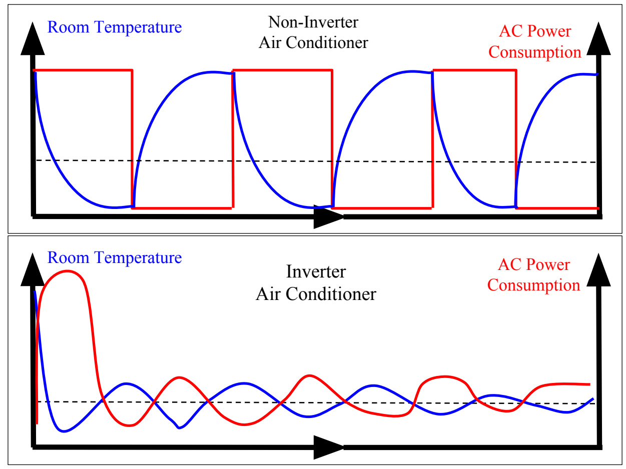


Figure A-2: Pictorial representation of room temperature and AC power consumption for Non-Inverter (top) and Inverter (bottom) AC

$$\begin{aligned}\frac{(T_{r_l}^{(t+1)} - T_{r_l}^{(t)}) \times C_{r_l}}{\tau} =& K_{ext}^{r_l} \times (T_{ext}^{(t)} - T_{r_l}^{(t)}) + K_{wall}^{r_l} \times (T_{wall}^{(t)} - T_{r_l}^{(t)}) \\ &+ K_{r_l-1}^{r_l} \times (T_{r_l-1}^{(t)} - T_{r_l}^{(t)}) + \eta_{r_l}\end{aligned} \tag{A.4}$$

For generic implementation, $r \in [1, r_l]$ where $r = 1$ implies the region closest to the AC, and r_l denotes the farthest region from the AC. τ is the time constant and $T^{(t)}$, C, and η indicate the temperature (at any time instance t), thermal capacity, and thermal noise, respectively, for the location depicted by their subscript. K is the heat transfer coefficient between the regions (including wall and external conditions) shown by their subscript and the superscript.

$$|\theta| = 5n + 2 \tag{A.5}$$

Figure A-1 presents the Resistive-Capacitive network of the generic implementation where the thermal capacity of a particular area is shown as a capacitor, and resistor indicates the heat transfer between the two regions (including the wall and external conditions). The size of the set of parameters (for the thermal model) is directly proportional to the number of regions (n) in a room (Equation A.5). Though the large size of parameters makes the optimisation problem complex and slow, increasing number of regions will only need a refined thermal model for the framework.

Appendix B

Thermal Model for Inverter AC

In Section 3.3.2, we discussed a thermal model with a non-inverter AC. Such an AC depends upon return air temperature (as measured by internal thermostat) and set temperature (provided by the user) to decide the compressor state (*on*/*off*). The compressor of a non-inverter AC has a fixed speed motor that can either run at full speed, or stop altogether. Thus, to maintain the desired temperature, the compressor motor of a non-inverter AC runs at full speed until return air temperature reaches the *off* hysteresis. Subsequently, controller completely shuts down the AC compressor to allow rise in temperature, until *on* hysteresis is attained. Due to temperature cycles, average room temperature remains close to the thermostat temperature set by the user. Therefore, for a fixed speed compressor, the shown thermal model denoted the AC state through a binary variable S_{AC} which can either be *on* (i.e. $S_{AC} = 1$) or *off* (i.e. $S_{AC} = 0$).

On the other hand, an inverter AC has a variable speed motor that can work even at a low RPM (rotations per minute); also known as Variable Frequency Drive (VFD). Initially, the compressor motor of an inverter AC runs at full speed to achieve the desired temperature. When controller senses that return air temperature is close to the set temperature, in place of complete shutdown, the motor rotates at a minimum speed required to maintain the set temperature. Thus, inverter ACs offer significant energy savings over non-inverter AC while maintaining the user comfort. However, to extend the proposed thermal model for an inverter AC, S_{AC} is updated to take any value between 0 and 1. Since the framework is designed to find optimal thermostat temperature for the AC, it can work for inverter ACs.

Appendix C

Extension to Home Environment

A residential apartment differs from a cold room in two major aspects:

1. Poor Insulation - Typically, thermal insulation in a residential apartment is substandard as compared to a cold room. Often, heat leaks through walls, gaps around doors and windows, and multiple such sources.

2. Noisy Occupants - In cold room, loss of cooling only happens (given superior insulation) when manager opens the door for cleaning or shifting the goods. In contrast, significant amount of cooling is consumed by the occupants in a residential apartment.

Thus, we need a high order thermal model to capture the non-linearity and estimate the room temperature in a residential apartment. Equation C.1-C.4 depict one such thermal model derived from Bacher et al. [10], also discussed in Chapter 3.

$$\frac{(T_w^{(t+1)} - T_w^{(t)}) \times C_w}{\tau} = K_e^w \times (T_e^{(t)} - T_w^{(t)}) + \sum_{r=1}^{r_l} K_w^r \times (T_r^{(t)} - T_w^{(t)}) \tag{C.1}$$

$$\begin{aligned}\frac{(T_{r_1}^{(t+1)} - T_{r_1}^{(t)}) \times C_{r_1}}{\tau} =& K_e^{r_1} \times (T_e^{(t)} - T_{r_1}^{(t)}) + K_w^{r_1} \times (T_w^{(t)} - T_{r_1}^{(t)}) \\ &+ K_{r_1}^{r_2} \times (T_{r_2}^{(t)} - T_{r_1}^{(t)}) + Q_{ac} \times S_{ac} + \eta_{r_1}\end{aligned} \tag{C.2}$$

Symbol	Description	Unit
τ	Sampling Interval	s
r	Thermal region $\in \{r_1, r_2, r_3\}$	$-$
η_r	Thermal noise in region r	$-$
Q_{oc}	Cooling load due to occupants and their activities	$-$
Q_{ac}	Cooling capacity of AC	kW
$T_r^{(t)}$	Temperature in region r at time instance t	$^{\circ}C$
$T_e^{(t)}$	External temperature at time instance t	$^{\circ}C$
$T_w^{(t)}$	Temperature of wall (facing outside) at time instance t	$^{\circ}C$
C_r	Thermal capacity of region r	kJ/K
C_w	Thermal capacity of wall (facing outside)	kJ/K
K_w^r	Heat transfer coefficient between wall (facing outside) and region r	kW/K
K_e^r	Heat transfer coefficient between external environment and region r	kW/K
K_e^w	Heat transfer coefficient between wall (facing outside) and weather	kW/K
K_{r1}^{r2}	Heat transfer coefficient between $r1$ and $r2$	kW/K
K_{r2}^{r3}	Heat transfer coefficient between $r2$ and $r3$	kW/K
$S_{ac}^{(t)}$	AC compressor state (*on*/*off*) at time instance t	$-$
$S_{oc}^{(t)}$	State of occupants at time instance t	$-$

Table C.1: List of symbols used in the proposed thermal model

$$\begin{aligned}\frac{(T_{r_2}^{(t+1)} - T_{r_2}^{(t)}) \times C_{r_2}}{\tau} =& K_e^{r_2} \times (T_e^{(t)} - T_{r_2}^{(t)}) + K_w^{r_2} \times (T_w^{(t)} - T_{r_2}^{(t)}) \\ &+ K_{r_1}^{r_2} \times (T_{r_1}^{(t)} - T_{r_2}^{(t)}) + K_{r_2}^{r_3} \times (T_{r_3}^{(t)} - T_{r_2}^{(t)}) \\ &+ Q_{oc} \times S_{oc} + \eta_{r_2}\end{aligned} \tag{C.3}$$

$$\begin{aligned}\frac{(T_{r_3}^{(t+1)} - T_{r_3}^{(t)}) \times C_{r_3}}{\tau} =& K_e^{r_3} \times (T_e^{(t)} - T_{r_3}^{(t)}) + K_w^{r_3} \times (T_w^{(t)} - T_{r_3}^{(t)}) \\ &+ K_{r_2}^{r_3} \times (T_{r_3}^{(t)} - T_2^{(t)}) + \eta_{r_3}\end{aligned} \tag{C.4}$$

Here, the first region (r_1) is the area in proximity of the AC, thus facing direct and the maximal impact of cold air coming from the AC. The second region, r_2 is the region where occupants stay, and the region receives indirect cooling from r_1 region where AC is present. r_3 indicates corner spaces in the room. The thermal model in Equation 5.1 is a special case of above mentioned model where whole room is considered as single region and manager's activities are monitored through a door sensor. Table C.1 describes all the notations used in the extended thermal model.

The idea is to logically divide the room into multiple thermal regions and capture heat transfer at region level. In the model, we assume that thermostat is installed closer to the AC and each region is considered to be separated by a thin layer of air having negligible thermal mass. To evaluate the model, we installed smart thermostat in the bedrooms of five residential apartments, and collected temperature and occupancy data for a month. In parallel, we gathered weather information from a cloud based weather service. Leave p-out cross validation (with $p = 10$) indicates that even in such a noisy environment, the extended model can estimate room temperature with a mean RMSE of 2.85°C (with $\sigma = 1.3$°C).

$$\begin{aligned}\theta = \{C_w, C_{r_1}, C_{r_2}, C_{r_3}, K_w^{r_1}, K_w^{r_2}, K_w^{r_3}, K_e^{r_1}, K_e^{r_2}, K_e^{r_3}, K_e^{w}, K_{r_1}^{r_2}, K_{r_2}^{r_3}\\ , \eta_{r_1}, \eta_{r_2}, \eta_{r_3}, Q_{ac}, Q_{oc}\}\end{aligned} \tag{C.5}$$

One must note that we only used temperature in r_1 ($\mathbf{T_{r_1}}$), outside temperature ($\mathbf{T_e}$), and occupancy information ($\mathbf{S_{oc}}$) to learn the parameter set θ (Equation C.5). While temperature and motion data can certainly be captured from a smart thermostat, climatic conditions are readily available from cloud based weather services. Once *Greina* learns the model parameters, it will simulate temperature in r_1 and compare with actual temperature data for leakage detection. Details for monitoring stage remain same as specified in Section 5.3. We believe, an interchangeable thermal model empowers *Greina* to be genuinely pervasive and ubiquitous. Depending upon the design requirements, researchers can explore different variants of the thermal model for diverse scenarios.

Bibliography

[1] Yomna Abdelrahman, Eduardo Velloso, Tilman Dingler, Albrecht Schmidt, and Frank Vetere. Cognitive heat: exploring the usage of thermal imaging to unobtrusively estimate cognitive load. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):33, 2017.

[2] Abdul Afram and Farrokh Janabi-Sharifi. Theory and applications of hvac control systems–a review of model predictive control (mpc). *Building and Environment*, 2014.

[3] Yuvraj Agarwal, Bharathan Balaji, Seemanta Dutta, Rajesh K Gupta, and Thomas Weng. Duty-cycling buildings aggressively: The next frontier in hvac control. In *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, 2011.

[4] AMPL. AMPL, 2019. [Online; accessed 11-January-2019].

[5] Omid Ardakanian, Arka Bhattacharya, and David Culler. Non-intrusive techniques for establishing occupancy related energy savings in commercial buildings. In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*, 2016.

[6] ASHRAE. ASHRAE, 2016. [Online; accessed 11-January-2019].

[7] Kriengkrai Assawamartbunlue and Michael J Brandemuehl. Refrigerant leakage detection and diagnosis for a distributed refrigeration system. *HVAC&R Research*, 12:389–405, 2006.

[8] Anil Aswani, Neal Master, Jay Taneja, David Culler, and Claire Tomlin. Reducing transient and steady state electricity consumption in hvac using learning-based model-predictive control. *Proceedings of the IEEE*, 2012.

[9] Anil Aswani, Neal Master, Jay Taneja, Valton Smith, Andrew Krioukov, David Culler, and Claire Tomlin. Identifying models of hvac systems using semiparametric regression. In *American Control Conference (ACC), 2012*, 2012.

[10] Peder Bacher and Henrik Madsen. Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings*, 43:1511–1522, 2011.

[11] Bharathan Balaji, Jian Xu, Anthony Nwokafor, Rajesh Gupta, and Yuvraj Agarwal. Sentinel: occupancy based hvac actuation using existing wifi infrastructure within commercial buildings. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, 2013.

[12] Nipun Batra, Pandarasamy Arjunan, Amarjeet Singh, and Pushpendra Singh. Experiences with occupancy based building management systems. In *Intelligent Sensors, Sensor Networks and Information Processing, 2013 IEEE Eighth International Conference on*, 2013.

[13] Nipun Batra, Manoj Gulati, Amarjeet Singh, and Mani B Srivastava. It's different: Insights into home energy consumption in india. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, 2013.

[14] Nipun Batra, Jack Kelly, Oliver Parson, Haimonti Dutta, William Knottenbelt, Alex Rogers, Amarjeet Singh, and Mani Srivastava. Nilmtk: an open source toolkit for non-intrusive load monitoring. In *Proceedings of the 5th international conference on Future energy systems*, 2014.

[15] Nipun Batra, Amarjeet Singh, and Kamin Whitehouse. If you measure it, can you improve it? exploring the value of energy disaggregation. In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, 2015.

[16] Nipun Batra, Amarjeet Singh, and Kamin Whitehouse. Gemello: Creating a detailed energy breakdown from just the monthly electricity bill. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 431–440, New York, NY, USA, 2016. ACM.

[17] Nipun Batra, Hongning Wang, Amarjeet Singh, and Kamin Whitehouse. Matrix factorisation for scalable energy breakdown. In *AAAI Conference on Artificial Intelligence*, pages 4467–4473, Palo Alto, CA, USA, 2017. AAAI Press.

[18] Fred S Bauman, Thomas G Carter, Anne V Baughman, and Edward A Arens. Field study of the impact of a desktop task/ambient conditioning system in office buildings. *Ashrae Transactions*, 1998.

[19] Alireza Behfar, David Yuill, and Yuebin Yu. Automated fault detection and diagnosis methods for supermarket equipment (rp-1615). *Science and Technology for the Built Environment*, 23:1253–1266, 2017.

[20] Belkin. Wemo WiFi Smart Light Switch, 2019. [Online; accessed 11-January-2019].

[21] Florian Bender, Aleksandr Skrypnik, Achim Voigt, Joachim Marcoll, and Michael Rapp. Selective detection of hfc and hcfc refrigerants using a surface acoustic wave sensor system. *Analytical Chemistry*, 75:5262–5266, 2003.

[22] Gail Brager, Hui Zhang, and Edward Arens. Evolving opportunities for providing thermal comfort. *Building Research & Information*, 2015.

[23] Michael R Brambley, Philip Haves, Sean C McDonald, Paul Torcellini, D Hansen, DR Holmberg, and KW Roth. Advanced sensors and controls for building applications: Market assessment and potential r&d pathways. Technical report, EERE Publication and Product Library, Washington, DC (United States), 2005.

[24] Mark S Breuker and James E Braun. Common faults and their impacts for rooftop air conditioners. *HVAC&R Research*, 4:303–318, 1998.

[25] Vikas Chandan and Andrew G Alleyne. Optimal control architecture selection for thermal control of buildings. In *American Control Conference (ACC), 2011*, 2011.

[26] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41:15, 2009.

[27] Bin Chen and James E Braun. Simple rule-based methods for fault detection and diagnostics applied to packaged air conditioners/discussion. *ASHRAE Transactions*, 107:847, 2001.

[28] Ke-Yu Chen, Sidhant Gupta, Eric C Larson, and Shwetak Patel. Dose: Detecting user-driven operating states of electronic devices from a single sensing point. In *Pervasive Computing and Communications (PerCom), 2015 IEEE International Conference on*, pages 46–54, Piscataway, NJ, USA, 2015. IEEE.

[29] Victor L Chen, Magali A Delmas, and William J Kaiser. Real-time, appliance-level electricity use feedback system: How to engage users? *Energy and buildings*, 70:455–462, 2014.

[30] Hwan-Hee Choi, Jeroen JG Van Merriënboer, and Fred Paas. Effects of the physical environment on cognitive load and learning: towards a new model of cognitive load. *Educational Psychology Review*, 26(2):225–244, 2014.

[31] Drury B Crawley, Linda K Lawrie, Curtis O Pedersen, and Frederick C Winkelmann. Energy plus: energy simulation program. *ASHRAE journal*, 42:49–56, 2000.

[32] Jonas MK Dake and Donald RF Harleman. Thermal stratification in lakes: analytical and laboratory studies. *Water Resources Research*, 1969.

[33] Sarah Darby et al. The effectiveness of feedback on energy consumption. *A Review for DEFRA of the Literature on Metering, Billing and direct Displays*, 486(2006):26, 2006.

[34] Lucas W Davis and Paul J Gertler. Contribution of air conditioning adoption to future energy use under global warming. *Proceedings of the National Academy of Sciences*, page 201423558, 2015.

[35] RJ Dear, T Akimoto, EA Arens, G Brager, Christhina Candido, KWD Cheong, B Li, N Nishihara, SC Sekhar, S Tanabe, et al. Progress in thermal comfort research over the last twenty years. *Indoor air*, 2013.

[36] Ayhan Demirbas, Ayman A Hashem, and Ahmed A Bakhsh. The cost analysis of electric power generation in saudi arabia. *Energy Sources, Part B: Economics, Planning, and Policy*, 12(6):591–596, 2017.

[37] T Dewson, B Day, and AD Irving. Least squares parameter estimation of a reduced order thermal model of an experimental building. *Building and Environment*, 1993.

[38] Temperature-Only Refrigeration Cycle Diagnostics. Advanced automated hvac fault detection and diagnostics commercialization program. *Contract*, 500:03–030, 2005.

[39] Bing Dong, Mikhail Gorbounov, Shui Yuan, Tiejun Wu, Abhishek Srivastav, Trevor Bailey, and Zheng Oâ Neill. Integrated energy performance modeling for a retail store building. In *Building simulation*, volume 6, pages 283–295, Salmon Tower Building, NY, USA, 2013. Springer.

[40] Tom Downey and John Proctor. What can 13,000 air conditioners tell us? *the Proceedings of the 2002 ACEEE Summer Study on Energy Efficiency in Buildings*, 1:53–67, 2002.

[41] Zhimin Du, Xinqiao Jin, and Lizhou Wu. Fault detection and diagnosis based on improved pca with jaa method in vav systems. *Building and Environment*, 42:3221–3232, 2007.

[42] Serge Dube. Refrigerant leak-detection systems, April 30 2009. US Patent App. 12/256,504.

[43] Varick L Erickson, Stefan Achleitner, and Alberto E Cerpa. Poem: Power-efficient occupancy-based energy management system. In *Proceedings of the 12th international conference on Information processing in sensor networks*, 2013.

[44] Varick L Erickson, Miguel Á Carreira-Perpiñán, and Alberto E Cerpa. Observe: Occupancy-based system for efficient reduction of hvac energy. In *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, 2011.

[45] Poul O Fanger et al. Thermal comfort. analysis and applications in environmental engineering. *Thermal comfort. Analysis and applications in environmental engineering.*, 1970.

[46] Povl Ove Fanger. Assessment of man's thermal comfort in practice. *Occupational and Environmental Medicine*, 1973.

[47] Radek Fisera and Martin Hrncar. System and method for performance monitoring of commercial refrigeration, June 28 2012. US Patent App. 12/976,088.

[48] Radek Fisera and Petr Stluka. Performance monitoring of the refrigeration system with minimum set of sensors. *World Academy of Science, Engineering and Technology*, 6:396–401, 2012.

[49] Christina Francis, Graeme Maidment, and Gareth Davies. An investigation of refrigerant leakage in commercial refrigeration. *International Journal of Refrigeration*, 74:12–21, 2017.

[50] Tanuja Ganu, Dwi Rahayu, Deva P Seetharam, Rajesh Kunnath, Ashok Pon Kumar, Vijay Arya, Saiful A Husain, and Shivkumar Kalyanaraman. Socketwatch: an autonomous appliance monitoring system. In *Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on*, pages 38–43, Piscataway, NJ, USA, 2014. IEEE.

[51] Peter Xiang Gao and Srinivasan Keshav. Optimal personal comfort management using spot+. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, 2013.

[52] Peter Xiang Gao and Srinivasan Keshav. Spot: a smart personalized office thermal control system. In *Proceedings of the fourth international conference on Future energy systems*, 2013.

[53] Carlos E Garcia, David M Prett, and Manfred Morari. Model predictive control: theory and practice—a survey. *Automatica*, 1989.

[54] Janos Gertler. *Fault detection and diagnosis*. Springer, Salmon Tower Building, NY, USA, 2013.

[55] MM Gouda, S Danaher, and CP Underwood. Building thermal model reduction using nonlinear constrained optimization. *Building and environment*, 2002.

[56] MM Gouda, Sean Danaher, and CP Underwood. Low-order model for the simulation of a building and its heating system. *Building Services Engineering Research and Technology*, 2000.

[57] IN Grace, Datta Datta, and SA Tassou. Sensitivity of refrigeration system performance to charge levels and parameters for on-line leak detection. *Applied Thermal Engineering*, 25:557–566, 2005.

[58] Manoj Gulati, Shobha Sundar Ram, and Amarjeet Singh. An in depth study into using emi signatures for appliance identification. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, pages 70–79. ACM, 2014.

[59] Sidhant Gupta, Matthew S. Reynolds, and Shwetak N. Patel. Electrisense: Single-point sensing using emi for electrical event detection and classification in the home. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, pages 139–148, New York, NY, USA, 2010. ACM.

[60] D Gyalistras, A Fischlin, M Morari, C Jones, F Oldewurtel, A Parisio, F Ullmann, C Sagerschnig, and AG Gruner. Use of weather and occupancy forecasts for optimal building climate control. *Technical Report*, 2010.

[61] Richard W Hamming. Error detecting and error correcting codes. *Bell Labs Technical Journal*, 1950.

[62] Hua Han, Zhikun Cao, Bo Gu, and Neng Ren. Pca-svm-based automated fault detection and diagnosis (afdd) for vapor-compression refrigeration systems. *HVAC&R Research*, 16:295–313, 2010.

[63] George William Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 1992.

[64] Alan Hedge, Wafa Sakr, and Anshu Agarwal. Thermal effects on office productivity. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 49, pages 823–827. SAGE Publications Sage CA: Los Angeles, CA, 2005.

[65] Chuan-Che Jeff Huang, Rayoung Yang, and Mark W Newman. The potential and challenges of inferring thermal comfort at home using commodity sensors. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2015.

[66] MA Humphreys and JF Nicol. An adaptive guideline for uk office temperatures. *Standards for thermal comfort*, 1995.

[67] Michael Humphreys. Outdoor temperatures and comfort indoors. *Batiment International, Building Research and Practice*, 1978.

[68] Michael A Humphreys and Mary Hancock. Do people like to feel 'neutral'?: Exploring the variation of the desired thermal sensation on the ashrae scale. *Energy and Buildings*, 2007.

[69] Hans IJzerman and Gün R Semin. The thermometer of social relations: Mapping social proximity on temperature. *Psychological science*, 20(10):1214–1220, 2009.

[70] International Energy Agency. The Future of Cooling, 2019. [Online; accessed 20-January-2019].

[71] Rolf Isermann. *Fault-Diagnosis Applications: Model-Based Condition Monitoring Actuators, Drives, Machinery, Plants, Sensors, and Fault-tolerant Systems*. Springer Publishing Company, Incorporated, Salmon Tower Building, NY, USA, 1st edition, 2011.

[72] Srinivasan Iyengar, Sandeep Kalra, Anushree Ghosh, David Irwin, Prashant Shenoy, and Benjamin Marlin. iprogram: Inferring smart schedules for dumb thermostats. In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, 2015.

[73] Milan Jain. Data driven feedback for optimized and efficient usage of decentralized air conditioners. In *Pervasive Computing and Communication Workshops (PerCom Workshops), 2016 IEEE International Conference on*, 2016.

[74] Milan Jain. Decision support system for room level air conditioners. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pages 350–354. ACM, 2017.

[75] Milan Jain. Thermalsim: A thermal simulator for error analysis. *arXiv preprint arXiv:1708.04524*, 2017.

[76] Milan Jain. Energy-efficient thermostats for room-level air conditioning. *arXiv preprint arXiv:1805.06284*, 2018.

[77] Milan Jain and Amarjeet Singh. Pacman: predicting ac consumption minimizing aggregate energy consumption. *DSpace at IIIT-Delhi*, 2014.

[78] Milan Jain, Amarjeet Singh, and Vikas Chandan. Non-intrusive estimation and prediction of residential ac energy consumption. In *2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2016.

[79] Milan Jain, Amarjeet Singh, and Vikas Chandan. Portable+: A ubiquitous and smart way towards comfortable energy savings. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2017.

[80] Mohit Jain, Deepika Chhabra, Jennifer Mankoff, and Amarjeet Singh. Energy usage attitudes of urban india. In *ICT4S*, 2014.

[81] Edward A Jeffers, R Philip McLeroy, and Jose L Sacerio. Halogen gas leak detector, 1984. US Patent 4,488,118.

[82] Jin-Ho Jeong. Refrigerant leakage sensing system and method, March 17 2005. US Patent App. 10/938,647.

[83] Josean Perez, Julio Montano and Jose Perez. Room temperature and its impact on student test scores, 2019. [Online; accessed 29-July-2019].

[84] Prasad S Kadle and Mahmoud Ghodbane. Refrigerant leak detection system, April 15 2014. US Patent 8,695,404.

[85] Rachel Kalaimani, Milan Jain, Srinivasan Keshav, and Catherine Rosenberg. On the interaction between personal comfort systems and centralized hvac systems in office buildings. *Advances in Building Energy Research*, 2018.

[86] Rachel K Kalaimani, Srinivasan Keshav, and Catherine Rosenberg. Multiple time-scale model predictive control for thermal comfort in buildings. In *Proceedings of the Seventh International Conference on Future Energy Systems Poster Sessions*, page 11. ACM, 2016.

[87] Gopinath Karmakar, Ashutosh Kabra, and Krithi Ramamritham. Coordinated scheduling of thermostatically controlled real-time systems under peak power constraint. In *Real-Time and Embedded Technology and Applications Symposium (RTAS), 2013 IEEE 19th*, pages 33–42. IEEE, 2013.

[88] Jack Kelly, Nipun Batra, Oliver Parson, Haimonti Dutta, William Knottenbelt, Alex Rogers, Amarjeet Singh, and Mani Srivastava. Nilmtk v0. 2: a non-intrusive load monitoring toolkit for large scale data sets: demo abstract. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, 2014.

[89] Jack Kelly and William Knottenbelt. Neural nilm: Deep neural networks applied to energy disaggregation. In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, 2015.

[90] Stephen L Keres, Alberto Regio Gomes, and Andrew D Litch. Fault detection and diagnosis for refrigerator from compressor sensor, December 6 2016. US Patent 9,513,043.

[91] Minsung Kim, Seok Ho Yoon, Piotr A Domanski, and W Vance Payne. Design of a steady-state detector for fault detection and diagnosis of a residential air conditioner. *International journal of refrigeration*, 31:790–799, 2008.

[92] Sun Ho Kim, Hyeun Jun Moon, and Young Ran Yoon. Improved occupancy detection accuracy using pir and door sensors for a smart thermostat. *Building Simulation 2017*, 15:2753–2758, 2017.

[93] Woohyun Kim and Srinivas Katipamula. A review of fault detection and diagnostics methods for building systems. *Science and Technology for the Built Environment*, 24:3–21, 2018.

[94] Wilhelm Kleiminger, Silvia Santini, and Friedemann Mattern. Smart heating control with occupancy prediction: how much can one save? In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, pages 947–954. ACM, 2014.

[95] Knud Lasse Lueth. Smart thermostat market for the period 2015 to 2021, 2016. [Online; accessed 12-November-2018].

[96] Christian Koehler, Brian D Ziebart, Jennifer Mankoff, and Anind K Dey. Therml: occupancy prediction for thermostat control. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, 2013.

[97] Kedar Kulkarni, Umamaheswari Devi, Amith Sirighee, Jagabondhu Hazra, and Praveen Rao. Predictive maintenance for supermarket refrigeration systems using only case temperature data. In *2018 Annual American Control Conference (ACC)*, pages 4640–4645, Piscataway, NJ, USA, 2018. IEEE.

[98] Susan Lang. Study links warm offices to fewer typing errors and higher productivity. *Cornell Chronicle*, 2004.

[99] Ky Le, T Tran-Quoc, Jean-Claude Sabonnadière, Ch Kieny, and Nouredine Hadjsaid. Peak load reduction by using air-conditioning regulators. In *Electrotechnical Conference, 2008. MELECON 2008. The 14th IEEE Mediterranean*, 2008.

[100] LG. Smart AC, 2019. [Online; accessed 11-January-2019].

[101] Bin Li and Andrew G Alleyne. A full dynamic model of a hvac vapor compression cycle interacting with a dynamic environment. In *American Control Conference, 2009. ACC'09.*, 2009.

[102] Bin Li and Andrew G Alleyne. Optimal on-off control of an air conditioning and refrigeration system. In *American Control Conference (ACC), 2010*, 2010.

[103] Bin Li, Richard Otten, Vikas Chandan, William F Mohs, Jeff Berge, and Andrew G Alleyne. Optimal on–off control of refrigerated transport systems. *Control Engineering Practice*, 2010.

[104] Haorong Li. *A decoupling-based unified fault detection and diagnosis approach for packaged air conditioners*. PhD thesis, Purdue University, 2004.

[105] Yunhua Li, Mingsheng Liu, Josephine Lau, and Bei Zhang. Experimental study on electrical signatures of common faults for packaged dx rooftop units. *Energy and Buildings*, 77:401–415, 2014.

[106] Clement Lork, Batchu Rajasekhar, Chau Yuen, and Naran M Pindoriya. How many watts: A data driven approach to aggregated residential air-conditioning load forecasting. In *Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on*, 2017.

[107] Jiakang Lu, Tamim Sookoor, Vijay Srinivasan, Ge Gao, Brian Holben, John Stankovic, Eric Field, and Kamin Whitehouse. The smart thermostat: using occupancy sensors to save energy in homes. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, pages 211–224. ACM, 2010.

[108] Henrik Madsen et. al. Estimation of continuous-time models for the heat dynamics of a building. *Energy and Buildings*, 1995.

[109] Dennis Martell and Jan Krcma. Refrigerant gas leak detector, September 27 1994. US Patent 5,351,037.

[110] Georgios Mavromatidis, Salvador Acha, and Nilay Shah. Diagnostic tools of energy performance for supermarkets using artificial neural network algorithms. *Energy and Buildings*, 62:304–314, 2013.

[111] Kathryn J McCartney and J Fergus Nicol. Developing an adaptive control algorithm for europe. *Energy and buildings*, 2002.

[112] Yu Meng, Margaret H Dunham, F Marco Marchetti, and Jie Huang. Rare event detection in a spatiotemporal environment. In *Granular Computing, 2006 IEEE International Conference on*, pages 629–634, Piscataway, NJ, USA, 2006. IEEE.

[113] Gordon R Morrow. Refrigerant leak detector system, October 4 1994. US Patent 5,351,500.

[114] Gordon R Morrow and Coy M White. Refrigerant leak detector system, June 11 1996. US Patent 5,524,445.

[115] Florian Nägele, Thomas Kasper, and Bastien Girod. Turning up the heat on obsolete thermostats: A simulation-based comparison of intelligent control approaches for residential heating systems. *Renewable and Sustainable Energy Reviews*, 75:1254–1268, 2017.

[116] Balakrishnan Narayanaswamy, Bharathan Balaji, Rajesh Gupta, and Yuvraj Agarwal. Data driven investigation of faults in hvac systems with model, cluster and compare (mcc). In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, pages 50–59, New York, NY, USA, 2014. ACM.

[117] J Navarro-Esbri, Eb Torrella, and R Cabello. A vapour compression chiller fault detection technique based on adaptative algorithms. application to on-line refrigerant leakage detection. *International Journal of Refrigeration*, 29:716–723, 2006.

[118] Nest. Nest Learning Thermostat | Programs Itself Then Pays for Itself, 2018. [Online; accessed 30-November-2018].

[119] Tuan Anh Nguyen and Marco Aiello. Energy intelligent buildings based on user activity: A survey. *Energy and buildings*, 56:244–257, 2013.

[120] J Fergus Nicol and Michael A Humphreys. Adaptive thermal comfort and sustainable thermal standards for buildings. *Energy and buildings*, 2002.

[121] Kazue Okamoto-Mizuno and Koh Mizuno. Effects of thermal environment on sleep and circadian rhythm. *Journal of physiological anthropology*, 31(1):14, 2012.

[122] Frauke Oldewurtel. *Stochastic model predictive control for energy efficient building climate control*. PhD thesis, Georgia Institute of Technology, 2011.

[123] Frauke Oldewurtel, Alessandra Parisio, Colin N Jones, Manfred Morari, Dimitrios Gyalistras, Markus Gwerder, Vanessa Stauch, Beat Lehmann, and Katharina Wirth. Energy efficient building climate control using stochastic model predictive control and weather predictions. In *American control conference (ACC), 2010*, 2010.

[124] Zheng O'Neill, Xiufeng Pang, Madhusudana Shashanka, Philip Haves, and Trevor Bailey. Model-based real-time whole building energy performance monitoring and diagnostics. *Journal of Building Performance Simulation*, 7:83–99, 2014.

[125] Kartik Palani, Nabeel Nasir, Vivek Chil Prakash, Amandeep Chugh, Rohit Gupta, and Krithi Ramamritham. Putting smart meters to work: Beyond the usual. In *Proceedings of the 5th International Conference on Future Energy Systems*, pages 237–238, New York, NY, USA, 2014. ACM.

[126] Manher Parekh. Method for detecting leakage in a refrigeration system, September 22 1992. US Patent 5,149,453.

[127] Nilavra Pathak, Md Abdullah Al Hafiz Khan, and Nirmalya Roy. Acoustic based appliance state identifications for fine-grained energy analytics. In *Pervasive Computing and Communications (PerCom), 2015 IEEE International Conference on*, 2015.

[128] W Vance Payne, Jaehyeok Heo, and Piotr A Domanski. A data-clustering technique for fault detection and diagnostics in field-assembled air conditioners. *International Journal of Air-Conditioning and Refrigeration*, 26:1850015, 2018.

[129] William V Payne, Piotr A Domanski, Jaehyeok Heo, and Zhimin Du. Self-training of a fault-free model for residential air conditioner fault detection and diagnostics. Technical report, NIST, USA, 2015.

[130] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 2011.

[131] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12:2825–2830, 2011.

[132] Therese Peffer, Marco Pritoni, Alan Meier, Cecilia Aragon, and Daniel Perry. How people use thermostats in homes: A review. *Building and Environment*, 46(12):2529–2541, 2011.

[133] Luis Pérez-Lombard, José Ortiz, and Christine Pout. A review on buildings energy consumption information. *Energy and buildings*, 40(3):394–398, 2008.

[134] Devika Pisharoty, Rayoung Yang, Mark W Newman, and Kamin Whitehouse. Thermocoach: Reducing home energy consumption with personalized thermostat recommendations. In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, 2015.

[135] Michael Ramey Porter, Joseph James Rozsnaki, and Osman Ahmed. Refrigeration system fault detection and diagnosis using distributed microsystems, March 27 2008. US Patent App. 11/786,033.

[136] Alimohammad Rabbani and S Keshav. The spot* personal thermal comfort system. In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*, 2016.

[137] Peter Radecki and Brandon Hencey. Online building thermal parameter estimation via unscented kalman filtering. In *American Control Conference (ACC), 2012*, 2012.

[138] Juhi Ranjan and James Scott. Thermalsense: determining dynamic thermal comfort preferences using thermographic imaging. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016.

[139] Neng Ren, Jun Liang, Bo Gu, and Hua Han. Fault diagnosis strategy for incompletely described samples and its application to refrigeration system. *Mechanical Systems and Signal Processing*, 22:436–450, 2008.

[140] P Riederer, D Marchio, JC Visier, A Husaunndee, and R Lahrech. Room thermal modelling adapted to the test of hvac control systems. *Building and Environment*, 2002.

[141] Charlie M Rinehart. Refrigerant leak detection system, August 10 2004. US Patent 6,772,598.

[142] Alex Rogers, Siddhartha Ghosh, Reuben Wilcock, and Nicholas R Jennings. A scalable low-cost solution to provide personalised home heating advice to households. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, 2013.

[143] Alex Rogers, Sasan Maleki, Siddhartha Ghosh, and Nicholas R Jennings. Adaptive home heating control through gaussian process prediction and mathematical programming. In *Second International Workshop on Agent Technology for Energy Systems*, 2011.

[144] Todd M Rossi and James E Braun. A statistical, rule-based fault detection and diagnostic method for vapor compression air conditioners. *Hvac&R Research*, 3(1):19–37, 1997.

[145] H Sachs. Programmable thermostats. *American Council for an Energy Efficient Economy Technical Report*, 2004.

[146] James Scott, AJ Bernheim Brush, John Krumm, Brian Meyers, Michael Hazas, Stephen Hodges, and Nicolas Villar. Preheat: controlling home heating using occupancy prediction. In *Proceedings of the 13th international conference on Ubiquitous computing*, 2011.

[147] Sense. The sense home energy monitor, 2018. [Online; accessed 12-November-2018].

[148] Sensibo. Sensibo: Smart Air Conditioner | Control Your AC With Your Phone, 2018. [Online; accessed 30-November-2018].

[149] Shravan Srinivasan, Arunchandar Vasan, Venkatesh Sarangan, and Anand Sivasubramaniam. Bugs in the freezer: Detecting faults in supermarket refrigeration systems using energy signals. In *Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems*, pages 101–110, New York, NY, USA, 2015. ACM.

[150] Stewart Unsdorfer. Recommended Thermostat Settings for Home Comfort, 2019. [Online; accessed 11-January-2019].

[151] Takahisa Suzuki, Tatsuo Tsunooka, Toshihiro Tahara, and Tetsuji Nobuta. Vapor compression type refrigeration apparatus including leak detection and method for detecting refrigerant leaks, January 27 2004. US Patent 6,681,582.

[152] Tado. Smart AC Control for you home, 2018. [Online; accessed 30-November-2018].

[153] SA Tassou and IN Grace. Fault diagnosis and refrigerant leak detection in vapour compression refrigeration systems. *International Journal of Refrigeration*, 28:680–688, 2005.

[154] Dan W Taylor and David W Corne. Refrigerant leak prediction in supermarkets using evolved neural networks. In *Recent Advances In Simulated Evolution And Learning*. World Scientific, Hackensack, NJ, USA, 2004.

[155] Lee Taylor, Samuel L Watkins, Hannah Marshall, Ben J Dascombe, and Josh Foster. The impact of different environmental conditions on cognitive function: a focused review. *Frontiers in physiology*, 6:372, 2016.

[156] Claus Thybo and Roozbeh Izadi-Zamanabadi. Development of fault detection and diagnosis schemes for industrial refrigeration systems-lessons learned. In *Control Applications, 2004. Proceedings of the 2004 IEEE International Conference on*, volume 2, pages 1248–1253, Piscataway, NJ, USA, 2004. IEEE.

[157] A Tindale. Third-order lumped-parameter simulation method. *Building Services Engineering Research and Technology*, 1993.

[158] ToI. How cheap power, inefficient ACs are pushing Delhi to the brink, 2019. [Online; accessed 02-March-2019].

[159] Amee Trivedi, Jeremy Gummeson, David Irwin, Deepak Ganesan, and Prashant Shenoy. ischedule: Campus-scale hvac scheduling via mobile wifi monitoring. In *Proceedings of the Eighth International Conference on Future Energy Systems*, 2017.

[160] TWC Product and Technology. Wunderground, 2014. [Online; accessed 12-November-2018].

[161] Thomas Weng, Bharathan Balaji, Seemanta Dutta, Rajesh Gupta, and Yuvraj Agarwal. Managing plug-loads for demand response within buildings. In *Proceedings of the Third ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, 2011.

[162] Wikipedia contributors. Cusum — Wikipedia, the free encyclopedia, 2017. [Online; accessed 12-November-2018].

[163] Dick Wirz. *Commercial refrigeration for air conditioning technicians*. Cengage Learning, Boston, USA, 2017.

[164] WWO. Weather World Online, 2019. [Online; accessed 11-January-2019].

[165] Rayoung Yang and Mark W Newman. Living with an intelligent thermostat: advanced control for heating and cooling systems. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2012.

[166] Rayoung Yang and Mark W Newman. Learning from a learning thermostat: lessons for intelligent systems for the home. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pages 93–102. ACM, 2013.

[167] Zhenyu Yang, Karsten B Rasmussen, Anh T Kieu, and Roozbeh Izadi-Zamanabadi. Fault detection and isolation for a supermarket refrigeration system–part one: Kalman-filter-based methods. *IFAC Proceedings Volumes*, 44:13233–13238, 2011.

[168] Zhenyu Yang, Karsten B Rasmussen, Anh T Kieu, and Roozbeh Izadi-Zamanabadi. Fault detection and isolation for a supermarket refrigeration system–part two: Unknown-input-observer method and its extension. *IFAC Proceedings Volumes*, 44:4238–4243, 2011.

[169] Jin Woo Yoo, Sung Bin Hong, and Min Soo Kim. Refrigerant leakage detection in an eev installed residential air conditioner with limited sensor installations. *International Journal of Refrigeration*, 78:157–165, 2017.

[170] Yuebin Yu, Denchai Woradechjumroen, and Daihong Yu. A review of fault detection and diagnosis methodologies on air-handling units. *Energy and Buildings*, 82:550–562, 2014.

[171] Hui Zhang, Edward Arens, DongEun Kim, Elena Buchberger, Fred Bauman, and Charlie Huizenga. Comfort, perceived air quality, and work performance in a low-power task–ambient conditioning system. *Building and Environment*, 2010.