Office Occupancy Detection based on Power Meters and BLE Beaconing

Rizky Pratama, Azkario

DOI:
10.33612/diss.147276967

Publication date:
2020

Office Occupancy Detection based on Power Meters and BLE Beaconing

Azkario Rizky Pratama

The work is supported by the University of Groningen, by the Indonesia Endowment Fund for Education (LPDP), and partially by the FIRST project under H2020-MSCA-RISE-2016 program.

ISBN: 978-94-034-2887-1 (book)
ISBN: 978-94-034-2888-8 (e-book)

Printed by ProefschriftMaken — www.proefschriftmaken.nl

This thesis was completed using the thesis LaTeX template by G. Andrea Pagani, University of Groningen.
Book cover idea: Ascariena Rafinda. *Book cover design*: M. Zaenudin.

rijksuniversiteit groningen

Office Occupancy Detection based on Power Meters and BLE Beaconing

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. C. Wijmenga
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
vrijdag 18 december 2020
om 11.00 uur

door

Azkario Rizky Pratama

geboren op 18 februari 1991
te Jogjakarta, Indonesië

Promotor:	Prof. dr. A. Lazovik Prof. dr. M. Aiello
Beoordelingscommissie:	Prof. dr. B. Jayawardhana Prof. dr. B. Koldehofe Prof. dr. H. J. Woertche

To my beloved wife and son,
Ascariena and Adnan

Contents

Acknowledgments

Pursuing a PhD degree in the Netherlands is an unforgettable and invaluable experience for me. I realize that this would not be possible without the assistance and support of many people to whom I wish to thank.

First of all, I would like to express my deep gratitude to my promoters, Prof. Marco Aiello and Prof. Alexander Lazovik. Dear Marco, thank you for your kindest support and trust. I remember the first time I met you in March 2014 when you invited me to visit the Distributed Systems group, just after we had a Skype interview. You let me explore ideas without being pessimistic. Your expertise and experience guide me to keep on track amidst the ups and downs. Thank you very much for supervising me while understanding my psychological well-being.

Dear Alexander, thank you for supervising me during my PhD trajectory. You teach me to think more critically and see things from other perspectives. We often have friendly discussions, even though sometimes you are very busy. I always feel relieved when I leave your office with new ideas or research planning approvals.

I would like to respectfully express my gratitude to the reading committee, Prof. Bayu Jayawardhana, Prof. Boris Koldehofe, and Prof. Heinrich Wörtche, for spending their time to review and giving constructive advice for my thesis.

I would like to take this chance also to express my gratitude to Dr. Lai Xu and Dr. Paul De Vrieze from Bournemouth University, UK, for initiating and organizing the EU FIRST project. I also would like to thank Prof. Yuewei Bai from Shanghai Polytechnic University, Shanghai, China, and my colleagues in Schoeneck, Germany (Stephan Boese, Norbert Eder, Michael Müller, Mahmoud Sharf, and Doortje Scherff) for their warm welcome and care during my visit. It was really great experience to take part of the EU FIRST project. I would like to extend my many thanks to Martin Sanders, for managing the project at our university.

Special thanks go to my office mates, Michel Medema, Brian Setz, and Frank Blaauw. Dear Michel, I am very grateful to you for being so kind. I remember when I got confused with Dutch taxes, health insurance, and contract extension, you helped me to find solutions and showed your care to these problems. You also quickly spend some time for translating the *samenvatting* for me. To Brian, thank you for your support in dealing with technical matters, especially regarding the group's data warehouse framework and Scala programming language that I never use before. You are always willing to help others and never doubt sharing your knowledge. Dear Frank, I am amazed at how you manage your tight schedule. I learn a lot about how to balance your life as a daddy, a husband, a researcher, and a company co-founder. Thank you also for collaborating in paper writing with me.

I also thank Michel and Panji Cahya for being my paranymphs. A fully digital defence ceremony that never is imagined before becomes a reality because of your help! Dear Panji, thank you for being my buddy in Groningen. It is good to meet a friend with the same souls to share thoughts, joys, and sadness; to become a traveling, shopping, and cooking mate; and always to be ready in a difficult time. Thank you!

To other current and former members of the Distributed Systems group; Ang Sha, Laura Fiorini, Talko Dijkhuis, Heerko Groefsema, Viktoriya Degeler, Tuan Anh Nguyen, Faris Nizamic, Ilche Georgievski, and Fatimah Alsaif, thank you for your support during my studies and for warm and friendly atmospheres. I cannot forget my first experience of outing with the group in Schiermonnikoog, even before I join as a PhD student. Furthermore, sailing experience and oberseminars are also unforgettable moments with you all. I also wish newer group members, Eren Aktas, Mostafa Hadadian, Majid Lotfian, and Saad Saleh, best of luck with your PhD study.

To other colleagues on the fifth floor of Bernoulliborg building, Estefania, Sreejita, Nicola, Ahmad, Maria, Aleke, Mohammad, Michiel, and Htet, sharing with you during our free time can relieve stress. Some of you also often pick us up for lunch in the office that reminds us to have lunch. Thank you!

To my past roommates: Bas de Bruijn. Vladyslav Tomashpolskyi, and Xun Li, I enjoy having small talks with you. Thank you also for brainstorming any confusion or sudden ideas that came to mind. The office would be extremely quiet without you.

I also would like to thank former and present supporting staffs: Ineke Schelhaas, Elina Sietsema, Desiree Hansen, Esmee Elshof, and Annette Korringa de Wit for arranging administrative matters. They are very kind and supportive in helping us in solving some bureaucratic issues.

My sincerest appreciation also goes to my colleagues in the Department of Electrical and Information Engineering, Universitas Gadjah Mada. I thank *Pak* Sarjiya, *Pak* Hanung, *Pak* Eka, and Prof. Selo for encouraging me to pursue PhD abroad and arranging any formal consents needed in the university. I will also never forget the recommendation letters of Prof. Sasongko, *Pak* Widyawan, and *Pak* Lukito that brought me to the LPDP scholarship. I would like to thank not only for the kindest letters but also for their encouragement and warmest advice during my study. I also thank Guntur, Frans, M. Faris, and *Mas* Ali for any fruitful discussions and sharing during this journey.

Also, many thanks to Indonesian friends and family in Groningen who make this city feels like home, especially with Indonesian cuisines and traditions: *Bapak ibu kos* Azka Muji - Aidina, *mas* Latif - *mba* Septi, *mas* Didik – *mba* Rosel, *mba* Nuril, *mas* Chalis - *mba* Jean, Didin - Anis, *mas* Adhyat - *mba* Nuri, Fika - Nisa, Ali - Liany, *Mas* Ali - *Mbak* Yosay, *Mas* Fajar - *Mba* Monik, *mas* Kuswanto – *umi* Fitri, *mas* Amak – *mba* Sinta, *mas* Joko - *mbak* Uchie, *mas* Riswandy, *mas* Khairul, *mas* Ega - *mba* Irma, *mba* Titis, *mba* Inna - *mas* Agung, *mas* Surya - *mba* Yassaroh, *mas* Ivan – *mba* Dita, *mas* Akbar, *mba* Retha, *mas* Krisna - *mba* Icha, *mas* Zaenal – *mba* Ayu, Umar, *mas* Ristiono - *mba* Afifah, Adityo, *mas* Lana - *mba* Arum, Bhimo, *mas* Naufal, Novita, Masyitha, Dina, *Mas* Fean, *Mas* Yoga, *mas* Tri, *mas* Fandi, *mas* Uri, *mbak* Nur, *mba* Ira, *mba* Vera D., *mba* Frita, *mba* Tia, *mas* Radit - *mbak* Nia, Reren, Ucon, Salva, *mas* Yudi – *mba* Sofa, Zaki - Nadya, Prety, Alfian, *mas* Adjie, *mba* Rosyta, *pak* Asmoro - *bu* Rini, *mba* Aryan, *mba* Isti, *mba* Era, Deka, Yovita, Rachel, Gerry, Dimas, Rio, Ayu, Reni, Marina Ika, Cancan, Afif, Rai, and Ghozi. Living in Groningen with you all brings many good memories and happy stories. I wish you success, health, and happiness.

Finally, I would like to express the heartfelt thanks to my big family, who always bring me extra energy to reach the finish line and get back home.

Dear Carien, how lucky I am to marry you in the last phase of my PhD study. We had lived 11,545 km apart for months, but you always show your care in many surprising ways. Thank you for being patient and supporting me when tired and feeling down. Also, to our dearest son, thank you for cheering us up with prayers and happiness in the last 22 weeks of pregnancy. You are the most precious gift for us. We were expecting to see you in the early of March, but the greatest God has another better plan. As we name you, Adnan, we hope you stay in the best place hereafter.

Dear *mama* Dina, *papa* Zulkifli, I realize that thankful words will never enough to express my gratitude for unconditional love and limitless support. Your blessings always accompany Carien and me to reach success in our life. I am thankful for everything you gave to us. I thank *mama* Wahyuning and *papa* Agus for the best blessings, care, and support to Carien and me. I am also thankful for your understanding during this endeavor. May your days be stuffed with health, happiness, and love.

To *eyang* Suharminah Wardiarto, thank you for your love, blessings, and support to achieve my dream. I wish you to be always as happy and healthy as today. Last but not least, to my brother, *dek* Arga, and my brother and sister in law, *mas* Ryan and *mbak* Putri, and my cutest nephew, *dek* Athar, I would like to express my gratitude for your love and taking care of our parents. I wish you blessings, happiness, and lots of luck!

Azkario Rizky Pratama
Purwokerto
November 15, 2020

Symbols

Symbols	Meaning
B	a set of beacon nodes: $B = \{b_1, \ldots, b_m\}$
β	a vector of beacon readings or beacon features
C	a set of appliance combinations: $C = \{\{laptop\}, \{PC, LCD\}, \ldots\}$
D	a set of electric loads (e.g., monitors): $D = \{d_1, d_2, \ldots, d_u\}$
ev	switching events (i.e., ON/OFF): $ev = \{ev_1, ev_2, \ldots, ev_c\}$
E	a set of electric features: $E = \{e_1, e_2, \ldots, e_h\}$
f_{ev}	event detection function
h_{occ}	occupancy state classifiers
h_{loc}	location classifiers in a beaconing system
h_{recog}	appliance classifiers in a power metering system
J	a set of electric loads; individuals: $J = \{j_1, j_2, \ldots, j_n\}$
L	a set of room locations: $L = \{l_1, \ldots, l_r\}$
M	number of deployed beacons:
m_{loc}	mapping inferred locations to occupancy states using a workspace location map
m_{mon}	mapping monitors' activation states to occupancy states using an inventory list map
N	number of people, events, electric features, or samples
$\mathcal{O}$	the ordered set of observation: $\mathcal{O} = \vec{X}_1, \ldots, \vec{X}_T$
P	BLE magnitude power (dBm); electric active power ($Watt$)
$\mathcal{P}$	a set of observation points: $\mathcal{P} = \{p_1, p_2, \ldots, p_o\}$
Q	Reactive power (VAR)
S	observation evidence from a sensory source in Dempster-Shafer Theory of Evidence

Symbols	Meaning
s	a sequence of instances
t	a time instance, $t = 1, \ldots, T$
v	a reference vector of beacons that represent a location
w	size of window width
$x_t^{j_i}$	individual j_i's power consumption
X_t	aggregate active power ($Watt$)
$\vec{X}_t$	the vector of electric features of aggregate power consumption: $\vec{X}_t = [f_1(\vec{x}_{t,e_1}), \ldots, f_h(\vec{x}_{t,e_h})]$
$y_t^{j_i}$	the binary occupancy state of individual j_i, where $y_t^{j_i} \in \{0, 1\}$
Y_t	the occupancy state of N individuals: $Y_t = y_t^{j_1}, \ldots, y_t^{j_n}$

Chapter 1

Introduction

The International Energy Agency (IAE) and the United Nations Environment Programmes Sustainable Building and Climate Initiative (UNEP-SBCI) reported that buildings are responsible for about 50-60% of the global electricity consumption [2, 3]. To respond to the finding, several efforts of reducing buildings' energy use have been made, for example, by improving buildings' thermal isolation and utilizing energy saving technologies and techniques. These attempts, however, cannot alone compensate for the increasing energy use due to *population* and *floor area growth*, the two dominant factors that rise total energy consumption both in residential and non-residential buildings. As shown in Figure 1.1, the influencing factors in residential and non-residential sectors differ in the *building use* that happens due to services, such as change in the temperature or ambient light settings. This tendency can be ascribed to the fact that non-residential building occupants are less aware of the energy consumption as they are not affected by energy bills [59]. Consequently, building consumptions and waste in non-residential buildings are higher than in households [12].

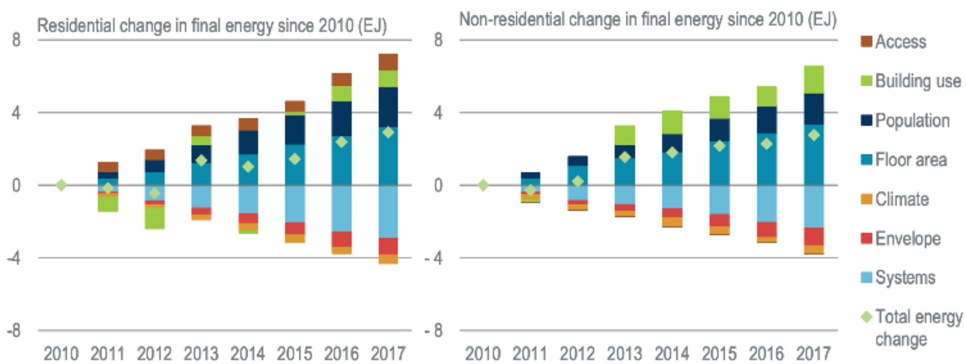

Figure 1.1: Factors that influence the building energy use, adopted from [4]

An example of occupants' inefficient behavior is the activation of power-consuming devices (e.g., lights) starting from early working time until the end of the working day (e.g., 7.00 AM until 7.00 PM), regardless of the actual occupancy. To save energy, building's energy-efficient lighting systems need to gather building context. Context may be defined as the situation of an area, information of nearby people, or properties of nearby resources [114, 5]. Let us consider a scenario as follows.

> *Suppose four employees share a room in a smart building. Ordered from the window to the innermost of the building is the space belonging to Aldo, Boy, Cecilia, and Diana. The office has central lights consisting of fluorescent lamp tubes on the ceiling. One double-tube is close to the window side (near Aldo and Boy's desks), and the other double-tube is assigned to the other side (close to Cecilia and Diana's desks). The employees have individual preferences. Aldo prefers not to use the lamps given the outdoor is clear, as sunlight provides enough illuminance to his space. Boy needs additional light, but dim light is fine with him. Thus, the utmost 75% brightness of the designated lamp block close to his space is his preference regardless of the outdoor weather. Cecilia works with a computer. Therefore, she does not need maximum brightness. However, as she is afraid of the dark, she prefers to turn all the lights in the office ON with 60% brightness when she stays in the office alone. As Diana works with documents and natural light barely reaches her space, she prefers more lights than the others. Partial lighting in the office is fine with her, given she has sufficient illuminance. Thus her preference is 90% of brightness from the nearest lamp block. The preferences are saved in a database.*

Depending on the context availability and various control mechanisms, potential energy-saving can be realized at various levels. When a building is aware of the present state of occupants in the office (e.g., obtained from PIR sensors), the building may control the lighting system. As soon as the PIR sensor detects value changes of infrared readings (i.e., due to any movements of employees), it sets the presence and triggers the lighting systems accordingly. This control is reactive and will remain active until no motion is detected for a specific time period (so-called *feedback loop*). Energy use will be lower when it turns the lighting system off when the office is vacant. However, the control is binary and does not accommodate individual preferences as the sensor cannot distinguish people present.

Finer-granularity contexts contribute to better-tuned control and energy savings with user satisfaction for the majority of occupants. For instance, Bakker, et al. report a user satisfaction of 84% for 35 participants in their experiment [35] . The context knowledge allows dimming lighting levels in a particular area, depending on occupancy. Let us say Aldo comes to the office when it is sunny. The building may delay the lighting activation until the next employee appears (e.g., as it knows Aldo does not need additional lights due to sufficient luminaries). The following person coming to the office, say, Diana, is then identified. Immediately, the building activates the partial lamp tubes near Diana's desk and adjust the luminance according to her preference. When all the occupants have arrived, the lighting systems are further adjusted based on the current condition. The control may be based on a set of predefined rules, and it will act depending on the rules and acquired contexts, such as user ID and luminance. Wozniak et al. propose predefined rules and fuzzy sets to adjust controllers according to the needs of recognized users [131]. They report up to 11.7% of energy saving on heating and lighting systems and dryers can be achieved with the proposed control.

A more complex control mechanism is automatic searching and composing the best sequence of actions based on Artificial Intelligence (AI) *planning*. AI planning is defined as an intelligent behavior in constructing strategies or action sequences to achieve some goals. In order to solve a planning problem, planners need to gather user contexts (e.g., occupant counts, identification, and activities) and the knowledge of available entities (e.g., the location of heaters and dimmable lights in a building). When it comes to a situation, such as, Aldo and Cecilia doing some activities with computers and Diana working on paper-based tasks; a planner could come to a solution of only turning ON the lamp tubes close to Diana with 90% of brightness. This decision is reasonable as Aldo does not need additional light and Cecilia does not require to turn on all the lights since she is not alone, while Diana requires more light due to paperwork. An example of planner-based indoor control can be found in [47]. The authors consider a public university restaurant with natural light coming from large windows and light fixtures that can be controlled manually or directly by the planner. They compare the manual light control to the feedback loop control based on movement sensors and a planner-based control and report an average energy saving of 71% and 89%, respectively, during a two-week observation.

From the provided scenarios, one can see how contexts are the basis for energy saving and fulfill users' expectation and needs. Numerous sensors and computational devices have been proposed to capture the context of how people live in buildings. Some of them are explicitly deployed for monitoring occupants, while others make use of the existing building infrastructure. In acquiring data, unobtrusive, off-the-shelf sensors are preferred. Sensors and devices need to work seamlessly without requiring to be worn or placed intrusively in the environment (e.g., in a way that causes a user to feel annoyed). Smartphones are a good candidate due to the proliferation of their usage. The smartphones, along with reference anchors (e.g., Wireless Fidelity (WiFi) access points and Bluetooth beacons), can support localization systems and inherently show occupancy information of building spaces. Additionally, the broad adoption of power meter technology presents a vast opportunity to reveal the context from power consumption readings. While the official numbers of actual deployment is not known, it is expected that power meter will cover at least 80% of consumers in sixteen EU member countries and will be reaching 95% on average by end of this year (2020) [42]. In the U.S., more than 50% of households have installed smart meters by the end of 2015, and it is expected that the number of smart meters will be reaching 90 million by 2020 [32]. Even further, there are independent service providers that offer power meter products to measure more granular power consumption in real-time with low-cost and quick installations. These meters also allow measuring power consumption per circuit by putting current clamps (CTs) in electrical lines.

1.1 Challenges

The acquired data from power meters and smartphones are raw power consumption and signal strength from reference anchors. Some processing activities are needed to leverage the data usefulness, such as extracting useful information and solving sensor conflicts in order to infer contextual information.

1.1.1 Need for Information Extraction

To use available data for occupancy detection in non-residential buildings, one faces the challenge of accurate high-level information extraction from the raw data. In particular, we consider two sources, power metering system and beaconing system. The former is based on power meters that measure the mixed energy consumption of several people or electrical loads, while the latter is based on Received Signal Strength (RSS) that can be exploited to infer a location or occupancy state.

Power Metering System. As a power meter is generally installed at the root of electrical distribution circuits, the recorded data is power consumption in aggregate form. It represents the total power consumption of devices being used by the occupants. To detect occupancy from power consumption, we need to detect the activation of presence-related appliances, or to mine occupancy pattern from the consumption traces. The extraction process of such information is known as *load disaggregation* or *appliance recognition* [136], that is, the process of breaking the total power readings (i.e., composite loads) into smaller components. The problem rises when in offices, homogeneous, low power consumption appliances are present. The disaggregation process is complex due to similar characteristics among appliances and oscillation or masked low-power consumptions [130]. While there is significant research in the field of load disaggregation in residential buildings, there is a dearth of research work in the office environment. To address this problem, we adopt two electrical signature forms, namely, state-transition based signatures and snapshots [77].

Beaconing System. A mobile phone can indicate indoor locations by exploiting electromagnetic signals transmitted by, for example, WiFi access points or Bluetooth beacons. The sensing is based on the observation of a user. The extracted information is thus not about occupancy (e.g., how many people present, or who is present in a space), but whether a particular occupant is in the space or where the occupant is located. Once location has been extracted, the occupancy state of the room location may be centrally inferred. Namely, if a person is located in a certain room, the occupancy state of the corresponding room is set as occupied at least by that person.

The location may be derived, for instance, from the unique combination of RSS from anchors (so-called fingerprinting technique), the nearest beacon reference [78], and the nearest neighbor classification [30]. However, the signal strength from reference nodes can vary. Different types of receivers (e.g., phones) may also deliver different measurement values, even when the mobile phones are associated with the same transmitting node (e.g., a BLE beacon) at the same distances [102]. Additionally, due to multipath propagation, the signals can be faded [25], presenting another challenge to extract location accurately in adjacent workspaces.

1.1.2 Conflicts

Several available sources may observe one common entity, but the inferences are not always correct and often present inconsistencies. The reason is that different views perceived by sensory sources may influence the observation. Additionally, the low-intrusive sensors are generally not specifically deployed for observing contexts, making the observations error-prone. Thus, the context extractions from individual sources might be inaccurate or biased. Recalling the example of occupancy context to control a lighting system, Aldo may be inferred in his workspace according to power meter readings. In contrast, BLE readings of his mobile phone may indicate that he is in the neighboring office. An option to deal with this problem is to combine the sensory readings from different modalities to generate more detailed and comprehensive measures. Alternatively, one can choose the most convincing inference between the two sensors when there are different inferred decisions.

1.2 Objectives

The objective of this thesis is to investigate simple sensing systems (i.e., power metering and beaconing systems) for occupancy detection in offices. Several research questions are addressed.

- **How is power consumption data acquired and analyzed while maintaining low-intrusiveness? How do low-intrusive power metering systems contribute to context awareness?**
- **Assuming that a power meter installed in a dedicated electric circuit of computer equipment is available, how can occupancy information be extracted? How accurate is the occupancy observation in offices based on the computer equipment activation?**
- **Assuming that a power meter with more electrical features are deployed in shared office rooms, how are active appliances recognized, and how are the present occupants distinguished? To what extent can we make use of this information for presence detection?**
- **How is beaconing localization carried out while maintaining low-intrusiveness? How precise is the occupancy inference in adjacent shared office rooms using beaconing localization?**
- **How can sensor fusion improve occupancy inference given individual sensors' benefits and faults?**

To answer these research questions, we carry out investigations empirically. The sensory sources and their corresponding programs (e.g., sensor gateways and mobile applications) should be deployed and implemented in real offices. Based on this deployment, we collect electric consumption as well as RSS data.

The data collection process is designed to be low-intrusive. We use existing mobile phones associated with users to receive beacon signals. While the phones vary and may measure inaccurate signals, we favor less calibration or training supervision. We also limit the power meter deployment. Two power meter types with different specifications are used. The power meter with only Watt measurement capability is simple to use in electric load identification. The other meter supports the measurement of more electrical variables. Using such an extra information, we identify user presence based on moving windows.

As previously mentioned, the collected data is not directly providing occupancy information, but rather information has to be extracted despite the inconsistencies and erroneous data measurements. We try several possible techniques to find out the best solution, including supervised machine learning techniques for classification, such as nearest neighbors and neural networks, and Markov models. We further fuse the sensors to improve occupancy detection in feature-level fusion and decision-level fusion schemes.

We assess the success rate of revealing occupancy detection based on several metrics, such as accuracy, F-measure, and Kappa measures.

1.3 Contributions

The research has resulted in the following, novel contributions:

1. **The identification of solution of low-intrusive power metering systems for context-aware purposes.** There is an opportunity to extend the usefulness of power meters for occupancy detection. Two reasons for this are the distributed deployment in buildings (i.e., sub-metering or circuit-level sensing) that are low-intrusive and the relationship between user presence and power consumption.

2. **A procedure for occupancy detection based on activation of low-power computer equipment (i.e., monitors)**. In the office environment, the activation of user-related appliances may indicate occupancy. Our proposal is to use power consumption changes to recognize office-related devices (in our case, computer screens). The activation/deactivation events may indicate employee occupancy. We validate the experiment in two offices.

3. **Office-related appliance recognition and fine-grain occupancy detection models based on feature rich power meters.** Office-related appliances that have small power consumption are difficult to distinguish. Meanwhile, power meters with several electrical features (e.g., measurement of reactive power and cos phi) may provide additional clues for the recognition process. We explore the sequential and non-sequential approach based on sliding windows upon power consumption readings to recognize office-related appliances and to identify user presence.

4. **A non-intrusive room-level localization system based on cosine similarity in adjacent offices.** Distinguishing a position between adjacent rooms is difficult, particularly with a non-intrusive approach (e.g., without dense fingerprinting surveys or thorough calibration processes). We propose to only sample signals in some parts of the area, followed by signal validation and classification based on cosine similarity.

5. **Decision- and feature-level fusion models to combine power metering system and beaconing system for occupancy detection.** The considered sensory systems are not perfect in detecting occupancy. We investigate sensor fusion in different levels to see how the fusion can improve the inferred occupancy.

1.4 Outline of the Thesis

Chapter 2 introduces contexts and provides a brief review of common technologies in occupancy context sensing. In particular, we look at the main goal for sensor deployment concerning the level of intrusiveness and information granularity. We then review the state-of-the-art, mainly on the localization system and power consumption monitoring, which we will focus on the rest of this thesis.

In Chapter 3, we identify the use of power metering systems for context determination. The identification includes common power meter installations and applications in buildings. We discuss the methods of information extraction from power consumption readings. This chapter serves as a background material before moving to the experimental chapters.

Chapter 4 discusses a procedure of occupancy inference experiments based on the switching state detection of computer equipment (e.g., monitors) on power consumption readings. This procedure assumes that the monitors are assigned to employees and used to support performing tasks in offices; thus, the monitor activation may reflect user presence in the workspace. We provide the experimental results for two different offices.

In Chapter 5, we analyze power consumption readings with more electrical variables. We provide an instance or a sequence of sensor readings to classifiers. Using this approach, we aim to recognize office-related appliances (e.g., LCD monitors, a CPU, laptop, and portable heater) from the aggregate power consumption and identify users in a shared office.

Chapter 6 goes further in the investigation of a beaconing-based system for occupancy detection. Specifically, we utilize mobile phones and BLE beacons to reveal occupancy in adjacent shared office rooms. As we look for low-intrusive solutions, we configure low power signal transmission on the beacons (e.g., to reduce the frequency of changing batteries) and limit training data collection (e.g., to reduce efforts to use the system). The collected training data are validated to make sure that they can represent the room location. Once completed, a classification process may be done to determine the room location of occupants. We thus compare the classification results with other low-intrusive approaches proposed in the field.

In Chapter 7, we investigate sensor fusion approaches for power metering and beaconing systems based on the level of data processing. We experiment with decision-level fusion and feature-level fusion. In decision-level fusion, the system makes temporary decisions based on sensor readings. The decisions are then combined to conclude a final inference. In feature-level fusion, feature vectors are firstly extracted from sensor readings. The combined feature vectors are then provided to classifiers.

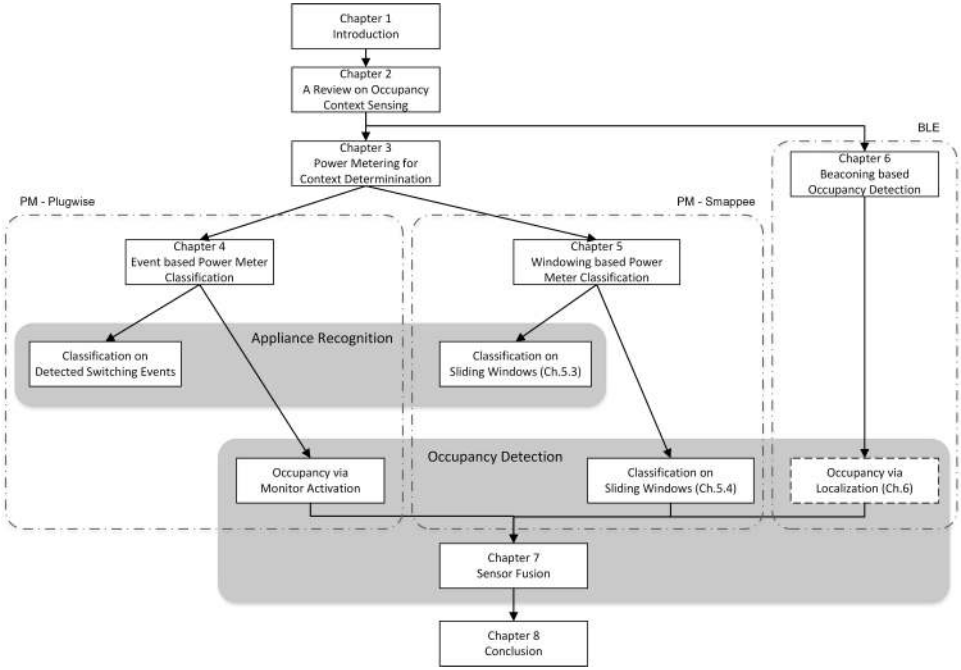

Figure 1.2: Thesis organization

Finally, we conclude our work with Chapter 8. This chapter also covers discussion on energy saving, privacy, and system portability. The schematic diagram of the thesis organization is shown in Figure 1.2. Blocks in dashed lines in this figure divide chapters based on sensing modalities, namely, Plugwise and Smappee power meters and BLE beacons. Gray-shaded boxes represent the aims of sections, particularly for recognizing appliances and detecting occupancy. The work presented in this thesis has been published in several peer-reviewed publications as shown in Table 1.1.

Table 1.1: Corresponding publications and articles produced during the study

Chapters	Venues	Citations
4	ICSOC 2017	[98]
5	APPIS 2018	[96]
	manuscript to be submitted	[94]
6	IDRBT 2017	[97]
7	Sensors, 2018	[99]
	UEMCON, 2019	[95]
	CCWC, 2017	[100]
	SMARTGREENS, 2018	[63]

Chapter 2

A Review on Ocupancy Context Sensing

Remarkable efforts have been performed to provide reliable contexts for energy-efficient buildings, including user contexts (e.g., users' location, activity), room occupancy, and energy usage (e.g., electricity consumption). This chapter presents the state-of-art in context acquisition with a particular focus on occupancy detection. Occupancy detection is defined as a process of discovering the state of living in a space.

We review scientific papers of the past decade and several less recent works which are milestones in the field. We present definitions of context in general and occupancy in particular in Section 2.1. We review sensing technologies from the way how it can extract occupancy information in Section 2.2. The intrusiveness of the sensory devices is described in Section 2.3. Finally, we provide the state of the art of the related systems from the occupancy detection perspective in Section 2.4.

2.1 Context and Occupancy

Schilit et al. were most likely the first to use the term "context" for user location, identities of nearby people, and properties of nearby resources [114]. They also introduced "context-aware" to address the ability of discovering and reacting to environment changes. A more general definition of context is given by Abowd et al., who describe context as any information that characterizes the situation of a person, place, or object [5]. Context is then useful as a foundation to provide services to a user. For example, user location contexts are needed to navigate users and show nearby shops; the activity context of the elderly is required to provide automatic assistance to improve life quality; and occupancy context is crucial to create a convenient environment by automatically adjusting lighting and air-conditioning systems, and at the same time, to reduce power consumption. In this thesis, we focus on the latter, where the context of occupancy can be improved using low-intrusive, potentially available sensory sources.

The term of *occupancy* has overlapping meanings. In its simplest meaning, it refers to a binary state of a space or so-called presence (i.e., being *vacant* or *occupied*) [8]. Other researchers mean occupancy not only as the binary present state but also as the number of people in the monitored space, e.g., [39]. These terms are defined as *occupancy detection* and *occupancy estimation* in [64, 27]. Occupancy also refers to a room location of people when the inference output is the room-level location [79, 49, 30]. Throughout this thesis, the term *occupancy* will refer to individual's present state in a particular office room. Occupancy and presence, therefore, may be used interchangeably.

2.2 Sensing Technologies

An energy-efficient building needs equipment to sense occupancy signs. Numerous sensing technologies have been proposed to do such a particular task. In this chapter, we differentiate technologies based on the purpose of their deployment, namely, conventional technologies (or explicit sensing), implicit sensing, and user-perspective sensing.

2.2.1 Explicit Sensing

A conventional way to sense occupancy in a space is by deploying a specific sensor to detect signs of occupancy, such as indoor movement. This way, the sensor (or a set of sensors) is explicitly deployed with a specific occupancy detection purpose.

Passive InfraRed sensor (PIR) is the most common sensor type in detecting movement due to its simplicity and affordable cost. PIR sensor detects occupancy by sensing infrared energy changes due to the movement of any heat radiating objects, including humans. To detect vacancy, PIR relies on a time-out period of non-detected motion. However, choosing the optimal time-out period is difficult. A small value (e.g., $15 - 20min$, or less) results in false unoccupied detection that brings disappointment to users, for example, when occupants do not significantly move during the period. On the contrary, longer time-out results in higher energy waste due to the activation of electricity devices when the space is vacant (i.e., false presences). Furthermore, this type of sensors requires a direct line of sight, which often cannot cover the whole part of the room.

Labeodan et al. evaluate occupancy detection using pressure chair sensors in an office building [71]. They modify existing chair cushions in a meeting room by embedding eight micro switches to detect state changes (i.e., closing or opening the switches based on sitting activity). Also, they use existing building space occupancy sensors, such as Carbon dioxide (CO_2) concentration, airflow rate, temperature, and humidity. The authors then compare the occupancy detection from those modalities.

Zhao et al. indicate occupancy detection based on PIRs and chair sensors in a shared office room before finally fusing them [140]. From two weeks of observation, both explicit sensing modalities detect vacant states very well, up to 99% of the times. While for the occupied event, PIR sensors installed on the ceiling can show up to 81% accuracy, while the performance degrades to 62% if they are installed on the walls. Chair sensor provides much higher performance, reaching 93.5% of the occupied states. Undoubtedly, the system cannot detect occupancy, if participants do not sit in the designated chairs.

Explicit sensing, however, requires considerable investment cost [118]. It is also limited in providing occupancy information, such as people counting, identity, and activity. The advancement of the Internet of Things with a myriad of data available encourages researchers to discover alternative approaches. Some notable strategies are discussed in the next section.

2.2.2 Implicit Sensing

Implicit occupancy sensing refers to the occupancy information extraction from existing systems (e.g., the traffic of computer networks, security card access systems, mobile and wireless communication systems) or potentially available systems for other purposes (e.g., indoor localization [30, 141], air quality controller [129], light intensity controller [59], PC's keyboard activities, webcams, or microphones [57]). A review of implicit sensing technologies is discussed in [118]. As this sensing type uses systems that are already available, the cost is relatively cheaper than the explicit sensing. However, as the sensor is not dedicated to infer occupancy, it generally requires more processing. For example, occupancy can be extracted from indoor localization systems [30, 141], speech detection [134] or speaker recognition on the recorded audio [57], or extra calibration processes with specialized equipment (i.e., Optical Particle Counters (OPCs)) [129].

Room occupancy based on location inference has been investigated using Radio-frequency Identification (RFID). Zhen et al. exploit active RFID to detect an occupant location in one of four office rooms [141]. The authors deploy seven RFID readers and split each room into three regions (in total, twelve regions are classified). They utilize Support Vector Machine (SVM) binary classifiers and use round-robin comparison to fit with the 12-class classification problem. The reported average accuracy is 93% in the classification of up to 240 RFID's signal strength vector samples per region. The occupancy information extraction from the localization system has also been investigated using Bluetooth Low Energy (BLE) beacons. For example, Conte et al. propose space occupancy classification by Bluetooth beacon received signals using machine learning approaches, namely k-Nearest Neighbor (k-NN) and decision trees [30]. The authors infer whether or not the occupant is present in a particular room based on a beaconing system.

Huang et al. propose occupancy detection using microphones and audio processing techniques [57]. Two schemes based on the number of speakers are investigated. Namely, meeting mode that involves only one speaker and party mode that includes multiple people speaking at the same time. To estimate the occupancy level, the authors propose a speaker recognition followed by summing up the number of speakers. This is possible in the meeting mode, where the speaker's voice is distinguishable, as the participants do not talk at the same time. For the party mode, the authors propose to extract the background audio energy acquired from the recorded audio. They report the accuracy of 90% for classifying up to 200 speakers in the meeting mode. For the party mode, the accuracy becomes higher when the speech measurement time is longer, up to 95% for the $25s$ measurement of up to 80 speakers.

Weekly et al. examine the correlation of particulate matter sensors, that commonly found in consumer devices (e.g., air purifiers), with human occupancy in a building observed by surveillance cameras [129]. The sensors are originally used to monitor small particles (i.e., with the size of more than $0.5, 1$, or $2.5\mu m$) for indoor air quality monitoring. The authors propose several pre-processes to extract features. It consists of filtering, variable selection, and calibration with OPCs. The authors point out that the phenomena of particles being lift off of a surface and becoming airborne when a person walks can indicate occupancy (so-called *resuspension*). A coarse sensor that only detects particles of size $\geqslant 2.5\mu m$ is sufficient. However, an accountable validation experiment is required to attest if the inferred occupancy can represent the entire room rather than only close to the camera, as in the referenced paper.

Jazizadeh and Becerik-Gerber investigate light intensity sensors for monitoring lighting systems in six rooms of a university building [59]. The aim is to estimate the energy consumption based on room light intensity. The authors detect the events of turning on/off or dimming the lights, from the lighting intensity changes. They thus correlate the events with the energy consumption of the lighting systems. This step generates useful features in supervising machine learning models. However, this work is not concerned with the prediction of occupancy states, even though the triggered events are directly related to occupant presence.

2.2.3 User-perspective Sensing

The work so far reviews context observation from the building perspective. Conversely, one can observe situations from occupant perspectives using smart devices associated with (or worn by) him (i.e., so-called wearables). From the occupant perspective, the acquired measurement is perceived in a specific, local view without necessarily knowing the context of the other participants. The sensory modalities sense only the surrounding environment and have no knowledge of the nearby instances (unless there is a communication among them).

Mashuk et al. investigate occupancy detection based on an indoor positioning system using a smart phone [81]. The idea is to estimate the location of a person as an indication of room level occupancy. The built-in mobile phone sensors include a gyroscope to detect walking orientation and an accelerometer to detect motions and count the step numbers. Furthermore, the authors utilize Bluetooth and WiFi modules for fingerprinting localization based on BLE beacons and WiFi access points installed in the environment. Given the measurements (e.g., estimated coordinates, step detection, and heading information), they perform a particle filter process to refine the estimated position. The beacons are also used as a trigger in floor-level changes. The results show that the occupancy detection cannot classify an occupied room precisely (i.e., especially between adjacent rooms) due to estimated position drifts.

Microphones in smartphones have also been explored to estimate the number of speakers involved in a conversation in room spaces under various conditions [134]. The authors propose a speech detection approach based on a lightweight clustering technique (i.e., forward clustering) to distinguish a new speaker from the previously recognized speakers. This step is then followed by counting the number of speakers. They perform experiments in various scenarios. The reported average error distance is 1.5 speakers with higher error counts when the phones are placed in the pocket of the owners.

A major advantage of user-perspective sensing is that it provides an identity connected to the phone. To preserve user privacy, however, the system is usually designed not to reveal the actual identity but to provide anonymous label instead. This method is particularly useful in personalized service automation, as there will be a signature from which phone the data is acquired. Another advantage is the ability to sense environment conditions without deployed infrastructures.

However, there are certain shortcomings associated with the use of wearables, such as the intrusiveness (in terms of user comfort), including battery drains that limit user mobility due to having to recharge, privacy threats, etc. Moreover, tests are also needed to assess reliability in terms of, for example, localization accuracy.

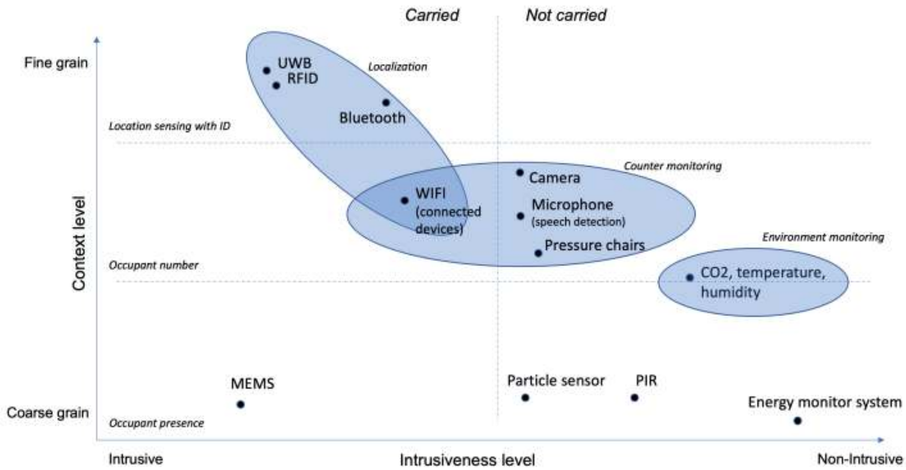

Figure 2.1: Context sensing technologies evaluated from intrusiveness and granularity dimensions, adapted from [36]

2.3 Sensing Intrusiveness

Figure 2.1 presents sensing technologies from the intrusiveness level and occupancy granularity dimensions. The term of *intrusive* has been used to refer to noticeable situations that lead to discomfort or disturbing feelings. It can be attributed to the intrusiveness of deployed devices and the discomfort perceived by the user [36]. Occupancy granularity is defined as the degree of details that can be exposed by sensory sources. It is also referred to as occupant resolution [82].

In the first dimension, the most notable separation of intrusiveness level is the requirement of carrying specific hardware to be sensing-enabled. We define an abstract partition that divides sensing systems with the requirement of taking a particular device. The half-left is the area for technologies that require a user to carry a device, while the other half is device-free sensing. The closer the position is to the origin, the more intrusive the system is. For example, RFID (e.g., [85]) is more intrusive than Bluetooth based approaches (e.g., [90]). The reason is that RFID requires a special tag or receiver, while Bluetooth requires only Bluetooth modules embedded on personal mobile devices. The energy monitoring system (i.e., power meters) has a low level of intrusiveness when it is placed out of occupants' visibility, such as in the root of the electrical energy distribution network. It, however, provides coarse-grain occupancy detection, since, it can only reveal occupancy or vacancy state of a residential building when placed in an incoming electrical line, such as

Non-intrusive Occupancy Monitoring (NIOM) [26].

In the second dimension, we divide sensor granularity into binary presence, user counting, and user identification. Environmental monitoring systems, including CO_2 concentration, temperature, and humidity, can identify the human presence and approximate the number of occupants in the space. This capability is due to their strong correlation with the number of occupants [72]. The number of occupants in an area can also be known by counting the number of connected devices [82], speakers [57], or chairs being used [71]. These are situations that can be monitored with medium-level intrusiveness. Medium intrusiveness is due to the higher number of sensor instances needed (e.g., attached on each chair). Compared to the environmental monitoring system, these are deemed to be more intrusive as they require more sensors deployed in the environment; hence, more invasive to occupants and more difficult in installation and maintenance. Finally, finer grain occupancy can be acquired through a personal localization system. RFID and Bluetooth offer personalized tracking features due to the association of RFID tags or Bluetooth signal receivers (e.g., phones) with particular occupants. While these systems depend on hardware to be carried, the adoption of Bluetooth technology in daily-used smartphones reduces the burden of carrying additional devices. The comparison of different sensor types for occupancy detection is discussed in the review by Chen et al. [27].

An ideal sensing source should be minimally obtrusive, by being able to sense environmental from afar and cover an entire environment (i.e., one sensor per room or less) [73]. Additionally, to acquire additional information (e.g., user identification (ID)), we may adopt localization systems. As will be discussed in the rest of this thesis, we will focus on the localization system and energy monitoring systems that contribute to the occupancy detection.

2.4 State of the Art

The information extraction from sensory sources covers numerous experiments with various sensory modalities. Our main concern is on localization and power metering systems, for their potentials in acquiring fine-grain occupancy with low intrusiveness.

2.4.1 Location

One advantage of the localization system is that it brings user ID, in particular, based on identifiers carried by occupants. The ID is particularly useful, for example, to control lighting or thermal systems based on occupant data, so-called occupancy-based control [91]. Previous works have shown that occupancy detection based on the localization systems with identification leads to users' comfort and energy saving, both lighting and Heating, Ventilation, and Air Conditioning (HVAC) systems [84, 144, 90, 14].

Moreno et al. propose to use very fine-grain location information (i.e., user location coordinates) for occupancy detection in a university laboratory to achieve efficiency in a heating system [84]. They deploy numerous Infrared-enabled RFID reference tags densely and require people to carry a monitor tag to be localized. The coordinate position is then estimated using neural networks, and the particle filter method is used to predict upcoming positions [85]. User comfort preferences are acquired based on user interaction through an interface. HVAC appliances are finally controlled based on occupants' identification and localization and unique adjustable comfort profiles. It is reported that the mean error localization could be lower than $1.5m$, and the energy reduction of 20% compared to a scenario without the energy management approach can be achieved.

Existing IT equipment, such as WiFi access points, may also be exploited for the same purpose. Zhou et al. achieve $1.385m$ accuracy of fine-grain localization using RSS fingerprinting (i.e., developing a database of signal strength distribution in an area) [144]. They design a mobile application to collect occupants' preferences for lamps. User preferences are accommodated when the corresponding occupant enters the zone where the lamps are located. An experiment in the eight weeks of a total of 24 weeks on the user preference-based control demonstrates up to 93% and 80% energy saving, compared to static scheduling control and PIR-based control, respectively, in the living space and four chambers. Balaji et al. investigate more coarse location information using the same sensory modality [14]. The authors propose to estimate users' locations based on the zone area of the connected access point to keep the system simple even in a large scale implementation. When a device is connected to an access point that covers the device owner's personal space, the owner is considered to be present. There is a mapping between occupants and corresponding office numbers and MAC addresses, handled by the system. About 83% accuracy is reported on the personal space occupancy detection over a ten-day experiment. HVAC system is then controlled based on the occupancy data on one experiment day. It is reported that saving 17.8% of electrical energy is achievable by controlling 55 HVAC zones (23% of total zones) in the building.

Research on the subject has been shifting to use available devices that support occupant daily activities, such as a mobile phone, not only to collect preferences but also to get location information. Coarse-grain location, such as the system based on WiFi authentication request, is non-intrusive, yet not sufficient for saving energy. At the same time, fine-grain occupancy demands significant efforts such as building a WiFi fingerprint database. Bluetooth is a potential solution to indicate room-level locations and show user preferences, as suggested by Park et al. [90], especially in places where WiFi does not cover all spaces very well. Park et al. propose LightLearn, a framework aiming at learning individual occupant preferences and environmental conditions in lighting control based on reinforcement learning. While Bluetooth makes use of existing mobile devices, the discovery of classic Bluetooth makes a nuisance on pairing new device requirements. Moreover, the authors address an individual occupancy instead of multi-user occupancy.

Distinguishing people present in a shared office is of interest because it can suggest personalized services to improve energy saving while maintaining satisfaction. Recent research has suggested that BLE advancement supports occupancy detection. However, the focus is only on a single occupant (e.g., [17, 45, 81, 30]). In this thesis, we address multi-occupant occupancy detection in shared offices. With multiple occupants, this work faces challenges such as various signals due to various handsets used by the employees as well as fast fading and multi-path propagation. These may influence the inference of multi-occupant presence, especially in adjacent rooms. More specific techniques and proposed solutions to the problems are discussed as relevant literature in Chapter 6.

2.4.2 Power Monitoring

A power monitoring system in a building may have more purposes than solely as a power measurement. As illustrated in Figure 2.2, various power meter types (i.e., centralized metering, sub-metering, and plug-based metering) have been explored to extract occupancy-related information. The farther from the origin the metering types are plotted, the finer-grained information may be collected, but the PMs become more intrusive. In the vertical axis, we see various purposes of the meters. The higher the meter position, the more generic power meter purposes. Power readings from single-point metering have been used in residential areas to reveal home occupancy status [66, 26], as shown in the top left of the figure. The detection process is non-intrusive, leveraging the existing power meter in a central panel. Yet, it only involves coarse granular detection (i.e., occupancy of a house or flat as a whole). Additionally, some efforts have been performed based on centralized metering to monitor appliances at home, without revealing additional user contexts (e.g.,[18, 19, 130]). While they extract information from coarse-grain power readings (i.e., per home), these efforts may only extract high-power electrical loads from aggregate power consumption, such as fridge and freezer, washing machines, and electric cookers.

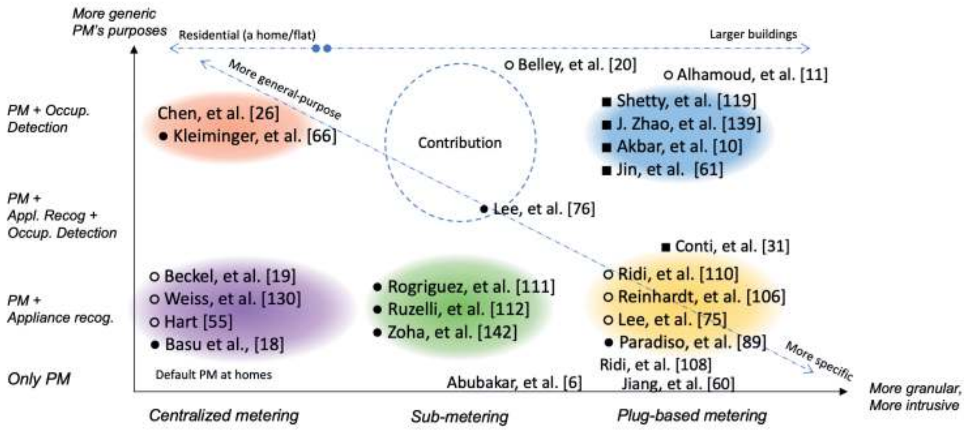

Figure 2.2: Various power meter utilization in buildings

On a larger scale, such as in large offices or other commercial buildings where many occupants live, the centralized power meter may not useful to show any information. It is unable to monitor occupancy only in a part of the buildings (e.g., on a particular floor or room) unless a specific meter is installed. The meter is defined as sub-metering or circuit-level sensing [106]. Fortunately, these meters are commercially available and relatively easy to install by clipping the meters to an incoming line of the electric board (e.g., Efergy Engage Sub-metering kit[1]). In this way, the power readings are still in aggregate forms, but with a smaller number of electric loads in a particular area.

Given the aggregate power readings, the purpose of power meters can be extended as a source of context. Some researchers have used sub-metering system for electric load identification purposes. With more granular power readings (e.g., at desk or room-level), they can extract activation of smaller power-consuming appliances, like those commonly used in the office. For example, Zoha et al. have investigated appliance recognition using plug meter per desk [142]. The authors propose Factorial Hidden Markov Models (HMM) and Generalized Likelihood Ratio to classify a combination of activated electrical loads on a desk (e.g., a PC, LCD, laptop, desk lamp, and fan). They use some combination of electrical features, including the average of real power and reactive power, power factor, and a standard deviation of real power and active power. The recognition of several combination appliances results in F-measure, ranging between 76-98% for binary state appliances

[1] https://efergy.us/engage-sub-metering-kit/

and 61-95% for multi-state appliances. Similarly, Rogriguez et al. study the identification of individual loads and the combination of them [111]. Kitchen appliances (e.g., kettle, microwave) and workstation appliances (e.g., heater fan, PC, lamp, and charger) are involved in the experiment. The authors use a high sampling power meter (i.e., $1kHz$) with two electric measures, namely, electric current and phase shifting. They generate more features derived when appliances are in transitional- and steady-state. Based on the features, active-appliance labels are then classified based on Decision Tree (DT). The identification of individual loads results in 90% accuracy for most appliances, while the recognition of the aggregate loads results in vary, ranging between 50-80%.

The recognition of appliances may indicate occupancy when the recognized appliances are those that require direct interaction (so-called user-interactive appliances [75], or usage dependent appliances [137]), for example, a computer, printer, and microwave. Lee et al. attempt to distinguish the user-interactive appliances from the others (i.e., background appliances and occupancy-reactive appliances) [75]. Their motivation is to use the recognized user-interactive appliances and the information of user presence (i.e., acquired from the other modalities) to deactivate unused power outlets for saving energy. In [31], Conti et al. have identified laptop power consumption and associated with some users. Their approach is based on plug metering per user, which provides some measurements (i.e., active and reactive power, RMS current, and phase angle). Apart from these works, other researchers generally concern with finding the activated appliances without linking this information to the occupancy, as shown in the lower part of Figure 2.2 (e.g., [110, 19, 111]). More experiments in appliance recognition with various setups and subjects, however, are needed as a proof-of-concept of benefit appliance recognition in occupancy detection.

Researchers have studied the occupancy detection in offices by mining power consumption. Yet, they mostly utilize intrusive power meter, either per appliance or per work desk, as clustered in the top right of Figure 2.2. Shetty et al. involve four participants in the experiment of individual presence states (absent/present). They employ a clustering approach of PIR sensor data and the power consumption of a PC and monitor during one-week observation [119]. Similarly, Zhao et al. deploy power meter per appliance in more varied appliances, including fans, chargers, lights, and printers [139]. The authors categorize the appliances to one of three classes (i.e., PC, lighting, and others) to infer an occupancy state (i.e., occupied with computer work, occupied with non-computer work, remote computer work, or unoccupied). They use DT, SVM, and naive Bayes classifiers to classify power readings per data point. The occupancy detection accuracies vary among the occupants and classification techniques. The best approach is DT, reaching an average accuracy of

90% and a kappa value of 69%. In addition to the present state classification, the authors also predict the number of room occupancy using regression approaches. They report a strong correlation between the prediction and ground truth, reaching 95%.

To reduce the number of deployed power meters, some researchers propose to use only one-meter per desk, representing the total power consumption per occupant. Akbar et al. utilize a power meter that measures active and reactive power, RMS voltage and current, and phase angle [10]. Several combinations of feature sets are applied using k-NN and SVM with various kernels to investigate occupancy state per desk (i.e., present, absent, and standby). It is reported that the more training data used, the more accurate the performance for all techniques. The overall accuracy reaches 93.67% based on two weeks experiment. Jin et al. utilize plug based power meter sensors at each work desk measured at a resolution of $1s$ [61]. The authors propose a Bayesian-based algorithm that does not require training labels [62]. The algorithm is based on rough estimation on working schedules followed by refining the prediction based on individual power readings. The authors compare the results with inferences from ultrasonic, acceleration, and WiFi connection. Also, they compare to threshold-based power consumption readings, the basic yet intrusive approach due to the involvement of a large number of power meters. The results show that the proposed approach is superior among threshold-based ones, and it is better than the acceleration and WiFi based inferences. It is also better compared to ultrasonic-based occupancy detection for most people.

Our work improves on the state of the art by considering sub-metering systems in an office. That is, the system measures the total consumption of occupants at room office level. In this scheme, our approach requires fewer power meters but still allows us to monitor low-power consumption devices. This work contributes to how the sub-metering benefits to occupancy detection while maintaining low intrusiveness.

Different markers in Figure 2.2 indicates some extraction techniques from power readings. Square markers (■) annotate occupancy extraction from power consumption data mining that generally uses moving windows. Circle markers (○) annotate event detection approaches for appliance recognition or electric load identification, while black-shaded circles (•) indicate recognition based on moving windows. The overview of the techniques is discussed in the next chapter.

Chapter 3

Power Metering for Context Determination

3.1 Overview

A vast number of power meters have been installed in recent years, as shown in Figure 3.1. That is, up to 20 million units are deployed annually in recent years in European countries, reaching about 165 million units in 2020 [113]. In the U.S., the number of installation by this year approaches 98 million units, and it continues to grow about 10 million units per year [7]. Smart meters have covered more than half of the U.S. households since 2017 [40].

The existence of power meters brings opportunity to improve building context

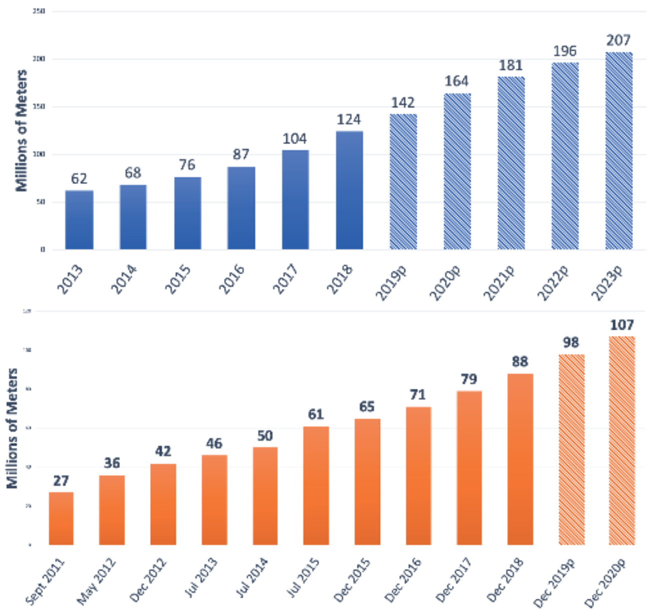

Figure 3.1: Smart meter installations in European countries (top) and the U.S. (bottom), adapted from [113, 7]

awareness. Power meters typically exist in a building and are accessible by building control systems, thus giving minimum intrusiveness level while minimizing budget allocation for adding additional sensors. Furthermore, power meters, which are based on electric wires, are considered to be robust toward signal interference, as suffered by radio signal-based sensing such as UWB, WiFi, and Bluetooth. However, this opportunity has not been fully explored by communities, mainly due to broad use cases (e.g., different office setup and appliances) and limited data availability due to restricted building access. This chapter contributes to the investigation of power metering systems for context-aware purposes. In Section 3.2, we investigate the availability of power meter and its utilization, including power meter as a measurement and monitoring device, and as a context source. In Section 3.3 and 3.4, we discuss techniques in electric load identification and data mining from power consumption readings. Finally, we summarize power metering systems as an occupancy detection source in Section 3.5, which also provides suggestions for our research. Following this chapter, we discuss experiments on power consumption readings based on event detection (Chapter 4) and sliding windows (Chapter 5).

3.2 Installation and Application

Power meters are commonly available in buildings, and they may be deployed at some locations in a building. The installation points and density influence information granularity, and thus, the power meter purposes.

3.2.1 Power meter Installations in Buildings

Power meter refers to a device that measures power consumption on the consumer side, such as in residential or commercial buildings. There are mainly two sensor installation locations. First, a centralized power meter is commonly placed at a single point sensing, usually at the incoming electrical line of a customer's building. The type of meter is usually a panel and may be equipped with a display, as shown in Figure 3.2 a). The panel power meter is relatively expensive as it has full features such as power quality analysis, high sampling rate (i.e., up to >40kHz), and complete measurement variables (e.g., current, voltage, power factor, harmonics, etc.). Such centralized sensing is seen in residential or public buildings, for example, installed by electric system operators. Using the readings, however, it is rather hard to have the consumption breakdown due to the complexity of power readings. Prior knowledge of appliances is required to break the component of consumption [130].

Second, it is also common for power meters to be deployed across the building. This is called distributed metering [106, 107] or hardware-based sub-metering

Figure 3.2: Commercial power meters: a) A panel meter from Schneider electric; b) A current clamp meter from Smappee; c) A plug-in power meter from Plugwise

[89, 22]. In this scheme, more than one meter is installed; each is responsible for a separate circuit representing a different area or different type of device. The meters in Figure 3.2 b) and c) are commonly used in the distributed metering. The former is a clamp meter, a jaw-shaped meter that works by measuring the magnetic field generated by current as it flows through a conductor. It works by clamping the jaw on a cable. The read features are not as complete as the panel power meter, and the price is relatively more affordable. The latter is a plug meter, a meter with wireless network module (e.g., based on Zigbee protocol) to communicate with other nearby plugs and to send data. Plug meter generally measures a single appliance, but it can be extended to measure a group of devices using an electric socket extension (e.g., in [61, 10]).

The more density of power meter in a building, the easier to breakdown the consumption readings as there are fewer electrical appliances involved in the measurement per meter. Thus it may be more sensitive in detecting lower power appliances [106]. Power meters spread out in a building usually have a lower sampling rate and fewer features at a more affordable price than the centralized meter. Ridi et al. propose to classify the sub-metering scheme into three sub-domains: one meter per zone, one meter per plug (i.e., each meter covers several appliances), and the most granular, one power meter per appliance [109]. The different schemes of power meter installation affect the information details and, thus, the utilization of the power readings.

3.2.2 Power meter applications

Beyond its basic function as a measurement device, a power meter may have some other purposes, such as monitoring electric loads and providing contexts of a building.

As a measurement and monitoring device. The advancement of power meters, that is, having a communication interface or protocol, has enabled information exchange between the meters and utility companies or energy suppliers. This feature allows direct and automatic billing purposes without having to survey power consumption home-by-home visually. The advanced power meter, called a smart meter, also has several sensors and control devices. These capabilities allow us to collect diagnostic information about distribution grids and home devices and to send command signals accordingly.

Another common application of power metering system is to monitor appliances or loads. The aim is to investigate power consumption behavior and usage patterns. The appliance monitoring can be done by massively deploying plug meters across a building, such as in [60]. In this way, the information is specific with much detail (i.e., per device). However, this approach is not scalable to large buildings due to high initial investments. A solution to this problem is to disaggregate composite power consumption loads. This is called Non-intrusive Load Monitoring (NILM) [55]. The idea is to decompose the total load into several component loads based on their contributions.

Other researchers focus on the recognition of activated electric loads, so-called electric load identification [38]. The motivation is to have insights into which appliance is activated so that it can be wisely controlled for remote actuation. The goal can be minimizing inefficient uses, satisfying users' preferences, or preparing certain appliances or electricity utilities to quickly react to changes in renewable energy source availability. Furthermore, it is possible to recognize possible malfunctions (e.g., [63]), so that we can mitigate problems before the appliance deteriorates. A typical approach is to deploy power meters throughout a building (i.e., submetering), and apply a supervised machine learning algorithm to classify the type of devices. Appliance recognition using centralized sensing is also possible, such as by detecting switching events on the aggregated power readings to detect appliance state changes. In this case, Weiss et al. propose to use a mobile phone to record appliance signatures for labeling the measurements in the training phase [130].

As a context source. Power consumption readings may reveal occupant information in a building, such as room occupancy, user location and activity, and identity. Lee et al. have reported that the power consumption readings may indicate room occupancy and user activity if occupants interact with user-interactive appliances, such as a coffee machine, television, and PC [75]. Any interaction with those appliances results in changes in power consumption footprints, such as higher power consumption and the presence of ripples when devices are ON. Even further, Conti et al. have identified that laptop power consumption may identify working users when power usage is acquired per individual [31].

Researchers have extracted occupants' appliance-related activities both in houses and offices. In [20], Belley et al. propose to recognize the activity of Alzheimer patients through ON/OFF event detection and activated appliances recognition from power readings. Four activity scenarios are experimented by including up to six devices operated at the same time. In [76], recognizing activated appliances is also a key to identify occupants' activity. The appliance recognition is done by exploiting the order of appliance activation using a dynamic Bayesian network. Once activated appliances are predicted, the system associates the appliances to activities based on the most frequent appliance use gathered from a social game. The authors focus on seven activities associated with appliances, such as using a computer, preparing a meal, watching TV, etc. Finally, the system shut unattended appliances down to conserve energy based on the recognized activities.

The extraction of temporal relations between consecutive activities has been proven to improve the activity detection accuracy. Alhamoud et al. investigate activity sequence patterns using the Apriori algorithm [11]. The algorithm scans the whole dataset to find all frequent activities and high dependency of two consecutive activities. It is found that *eating* activities frequently happen after *cooking* or *making coffee*. Furthermore, user location is also predicted by mining individual appliances' power consumption, such as in the kitchen, living room, work area, or outside.

3.3 Electric Load Identification

So far this thesis has focused on the power meter installation and application. The following section will describe electric load identification terminology. Electric load identification is the assignment of appliance labels to the electric loads based on unique characteristics in power meter readings. The identification includes the type or model of a load and its operational status. The meaning of this term has been broadened in recent years. *Appliance recognition* is defined as the process of recognizing the operating states of appliances from raw sensing data of electric power [76]. Likewise, *load monitoring* is defined as the process of acquiring and identifying load measurements to determine the energy consumption and status of individual loads in a system [6]. While these terms are similar in the detection of appliance states, load monitoring is generally an extension to determine the energy consumption of the known individual states. Similarly, *Load disaggregation* is defined as a separation process of single appliance power consumption from the total power readings [108, 13]. This process is usually needed to obtain appliance signatures in NILM.

In general, two approaches are common in identifying the electric loads. First, optimization is often done with an objective function, such as minimizing the residual value between real power readings and the known power consumption (e.g., as stated in an appliance manual book). An example of such method is the Least

Square Estimation [69]. Another approach is based on the recognition of electrical signatures. There are two signature forms as defined by Liang et al. [77], snapshot form and delta form. *Snapshot form* is defined as an instantaneous capture of readings at any fixed time intervals, while *delta form* is the difference of two consecutive snapshots form signatures. The former is related to a sliding window approach, while the latter is related to switching state detection. The techniques are discussed in the following.

3.3.1 Switching state Detection

Based on the premise that appliance state changes lead to energy consumption fluctuation, the switching state detection technique finds ON/OFF transition events in the power readings. The study was first carried out by George W. Hart [55]. The author initially observes the behavior of appliances, consisting of appliance active and reactive power changes when they are switching ON/OFF. He then performs cluster analysis and build appliance models. Alternatively, a user may record an appliance behavior manually to introduce the appliance to the system. Based on this knowledge, the system detects switching events, and assigns appliance labels with the best matching cluster or appliance model to the events. The changes may also be observed from electric current (e.g., [111]) or other electric variables. Figure 3.3 illustrates switching state detection on a series of power readings.

While the idea is straightforward, there are some challenges in the implementation. First, detecting events, especially low-power devices, is not easy. It is because fluctuations in the aggregate power readings can shade the power consumption per device. Furthermore, multiple consecutive events with shorter period than the sampling interval will be detected as a single event. Second, it is challenging to match ON/OFF combinations in a sequence. An ON/OFF activation state of a device might be incorrectly predicted when other events (e.g., the activation of other devices) present in between. Such an error could lead to a more complicated problem when further processes are applied on top of switching state detection. Finally, electric load identification based on switching state detection has high complexity in terms of cost and time in recording an appliances database, as the assumption of this scheme is that all possible appliance changes are introduced, and load signature database has been formed [77]. A comprehensive review of appliance load monitoring systems based on power changes is presented in [138].

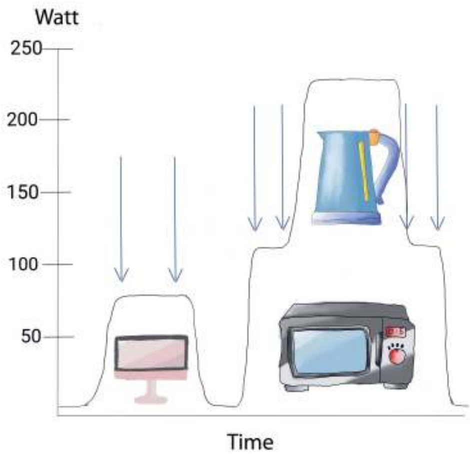

Figure 3.3: Switching state detection, adapted from [18]

3.3.2 Sliding window Approach

Another approach is to detect whether an appliance is activated during a sample duration. The implementation of this approach is generally based on sliding windows on power readings to capture appliance features [142]. The similar approach has also been done by Basu et al. [18]. The authors investigate electric load identification on centralized power meter readings in a hundred houses to identify individual loads directly from the total load readings. After clustering the houses into four categories, they employ a 10-minute overlapping sliding window and apply HMM models and nearest neighbors with various distances. It is reported that appliances with a significant amount of energy consumption could be detected. The best-recognized appliance is a water heater with 91% F-measure in three house-clusters, followed by a dish-washer, electric cooker, and an electric oven with about 60% F-measure, depending on the house-cluster. Nearest neighbor based approaches with dynamic time warping and euclidean distance seem outperform the HMM.

As this scheme does not depend on the switching state detection, the approach seems more suitable for low sampling rate power meters. However, several drawbacks appear. First, multiple appliances run simultaneously form an aggregate power consumption that is difficult to interpret. It requires classification model updates whenever a new device is introduced. Second, the size of the window width is difficult to determine. The wider window size leads to more device transitions involved in the process, while narrower sizes limit the information gathered. The best window width is often discovered from trials.

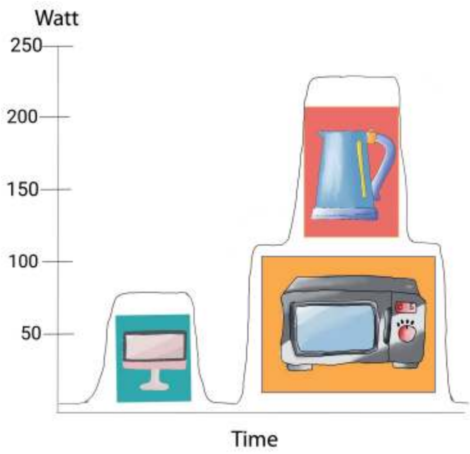

Figure 3.4: Sliding window approach, adapted from [18]

3.4 Power Consumption Data Mining

Mining on power consumption data has been done generally to detect occupancy in homes or workspaces. The existing occupancy detection in the residential sector focuses particularly on recognizing higher mean and standard deviation of power loads when the household is occupied [67, 26]. In [26], the occupancy state is detected based on thresholds, while in [67], several machine learning approaches (e.g., SVM, k-NN and HMM) are applied to classify the state of occupancy. Similar work is also done with lower resolution data (i.e., 30-minute interval and $100Wh$ increments) by firstly estimating the power consumption before applying machine learning classifiers [56].

In a cubicle office environment, binary occupancy detection has been approached by deploying power meters for each computer-related device (e.g., PCs and monitors) and a PIR sensor in each workspace [119]. Unsupervised k-means clustering is applied to interpret the absent and present state for each of four users. In [10], the setup is simpler by considering a power meter installation for each desk and letting users connect any devices to the measured power outlet. k-NN and SVM based techniques are then used to classify three occupancy states (i.e., away, present, or standby). In addition to the occupancy state classification, the prediction of room occupancy level is also done using a regression approach [139]. The authors deploy power meters to measure individual office appliances, including computers, lights, and others. Petrovic et al. infer occupancy of households and offices by mining the power consumption of WiFi routers and office appliances [93]. While it is found that the router's power consumption may indicate occupancy in a room, the detection performance is improved when additional appliance consumption data is involved in the inference. There is a relationship between the number of occupants and the increasing amount of power consumption.

3.5 Summary

Thus far, we provide an introduction to power meters in buildings and their contribution to electric load identification and occupancy detection through power consumption data mining. The chapter also discusses two forms of signatures in analyzing power consumption. The lesson learned from this chapter are summarized as follows:

1. The farther power meter deployment is from users, the less intrusive it is, but the information detail is consequently lower.

2. Power meters are commonly found in buildings. Power consumption is potentially used for appliance monitoring, and it has a relationship with contexts of users (e.g., occupancy and activity), especially the consumption of those devices that need an interaction to be activated. While approaches to recognize activated appliances have been proposed, they only identify appliances, and not many of them use this information to reveal occupancy. We aim to fill this gap by recognizing office-related appliances to indicate user presence in Chapter 4.

3. Two common schemes for analyzing power consumption exist, namely, switching state detection and sliding window approaches. Each has its drawbacks and advantages. Our aim is to ascertain which one is better for occupancy detection by experimental means. We report our findings in Chapters 4 and 5.

4. Occupancy may be gathered by mining the energy consumption of a house, workspace, device, or even a WiFi router. However, published studies on the occupancy detection based on power consumption data mining mostly rely on power meter installed either per appliance or per desk. The studies would have been less intrusive if they use fewer power meters, such as one meter per circuit or per room. We aim to fill this gap by mining power readings on the sub-metering system in Chapter 5.

Chapter 4

Event based Power Meter Classification

4.1 Overview

Power meters offer an opportunity to acquire contexts related to occupants in a non-intrusive way at a relatively low cost. To gain such benefits, power meters need to be installed sparsely and far from occupants. For example, it needs to be installed at the root of the electrical line or at the circuit-level in an electric circuit breaker rather than attached to each appliance. This way, the measured consumption comes in the aggregate forms, and the readings are affected by many devices' consumption.

The contexts may be acquired by recognizing individual appliances from the power readings. Recalling the smart building example, if the building can identify Boy's computer monitor activation from the sub-metering measurements, it becomes aware that Boy starts to work. Eventually, it may assign only 75% brightness to a specific lamp close to his workspace as he preferred. Unfortunately, the study of appliance recognition from sub-metering in the office environment is scarce as studies are mostly concerned with high-power electric appliances in residential buildings. In this chapter, we focus on typical offices, with employees working at desks mostly on their computers. In particular, we consider the electric load identification of computer monitors. The study is based on switching state detection, a method that explores appliance state-changing events by recognizing significant changes on its power consumption waveform. We propose a procedure to recognize appliances from aggregate power consumption readings inspired by Weiss et al. [130]. The differences lie in the appliances to be detected (i.e., we aim to identify low-power monitors) and the generated appliance features based only on active power. We show that the appliance activation may indicate the occupancy of employees. However, the recognition of the activated monitors from measured aggregate power consumption is not easy, especially in a large office with more identical devices.

This chapter firstly presents relevant literature on appliance recognition based on switching states in Section 4.2. It then describes the off-the-shelf power meter sensing technology and discusses the proposed procedure in Section 4.3. The experiment details and results are discussed in Section 4.4 and 4.5, respectively. Finally, Section 4.6 summarizes and concludes the experiments.

4.2 Relevant Literature

A number of studies have investigated electric load identification based on event detection. Inspired by Hart's seminal work [55], Belley et al. identify electric loads based on active power (P) and reactive power (Q) [20]. The authors experiment on three-phase power meter readings installed in a laboratory. Appliance signatures are initially extracted by detecting switching states of each appliance, followed by computing ΔP and ΔQ. They also note to which electric phase line the appliance is connected to. In the appliance identification phase, they use the collected signature database with predetermined decision rules to classify detected switching ON appliances. The rules firstly examine in which phase the event was detected. It then compares ΔP and ΔQ and, if still in an acceptable range, compares other Boolean features (e.g., whether reactive power consumption occurs continuously). The investigation on 16 appliances shows that most switching events are identified using the proposed approach, except for the fan, stereo, and refrigerator due to considerably small ΔP (each with 153, 200, and 1100W, respectively). The rest of the appliances are high power consumption such as stoves, kettles, oven, etc. Following the appliance identification, the authors investigate human activities. It is reported that four activity combination scenarios are detected with at least 97% accuracy. These scenarios include a combination of up to twelve events of six different appliances.

Weiss et al. show that home appliances may be recognized on aggregate power readings without deploying a large number of sensors [130]. They propose a single meter system to measure aggregate power consumption and recognize appliances based on events. First, the time points of switching events are detected. The event detection relies on the threshold of the absolute differences of two consecutive smoothed apparent power measurements (i.e., threshold $f_{th} = 2VA$). Several filtering methods are compared to avoid false events (i.e., due to transient behavior, not appliance switching states) as many as possible without missing the real switching events. Second, electrical parameter differences before and after the event are then computed to acquire power change information due to an activation of an appliance. Also, power levels between two consecutive events are also extracted. Finally, the differences are compared to an appliance signature database, and the event is mapped to a switching state of a device accordingly. The signature database is populated using a smartphone application that acquires occupants' feedback. The matching process relies on the nearest neighbor search in a two-dimensional space (i.e., active power and reactive power). The authors consider up to eight home appliances. Most of them consume more than $30W$, except for a console game, fluorescent lamp, and CD player that consume $15, 25,$ and $3W$, respectively. The testing runs over several hours with multiple devices running in random order. 125 of 128 switching events are identified correctly. They do not further test the detection system in real-world data, such as in daily life where occupants do some activities.

Beckel et al. and Cicchetti implement Weiss's approach with some changes (i.e., in the creation of the appliance signature database) [19, 28]. They validate the approach in their dataset, the one of six households consumption data. As a solution to missing Weiss's signature collection procedure which uses a smartphone application, Beckel et al. install up to ten plug meters per each household. They use a threshold on active power (i.e., $f_{th} = 5W$) to detect appliance activation on the plug data and extract appliance signatures from the smart meter in training sessions. The authors then test 90 days of household consumption data with an additional 15 days for training. There are nine appliances to be recognized in the chosen household, categorized as cooling appliances (a fridge and freezer), high consumption devices (a dishwasher, kettle, and stove), and others (a lamp, TV, stereo, and laptops). The results report that cooling appliances detection has nearly perfect detection with $.92$ F-measure. The detection of high consumption devices also results in almost no False Positives, but the algorithm misses many events for these appliances, resulting in low F-measure of up to .56 and .25 for dishwasher and stove, respectively. For the rest of the appliances, the F-measure are low, mostly due to low power consumption (i.e., laptop and stereo consumes $23W$ and $55.6W$), and easily be confused with switching events or variations caused by other appliances.

Event-based detection has been used in the identification of electric loads. The well-recognized appliances are usually high power consumption devices, such as kettle and stove. While the activation of these appliances may indicate contexts (e.g., presence in a kitchen), the researchers barely make use of this information. Also, difficulties arise when the involved appliances are similar, such as monitors in office environments. This similarity challenges the implementation of event-detection on small office appliances. Furthermore, while the aforementioned works use active, reactive, and apparent power measurements, most of the off-the-shelf meters measure active power only [126]; thus, the proposed approaches may not work properly.

4.3 Design

We utilize plug power meters with a basic measurement capability (i.e., active power in Watt). The meters are installed on each monitor to discover the relation of monitor usage and user presence in two offices. We investigate low-intrusive occupancy detection on the aggregation of monitors' power consumption. We propose an approach based on switching event detection and apply several machine learning techniques to recognize which monitor is activated given the aggregate consumption.

4.3.1 Sensing Technology

Plugwise Circle is a plug power meter and has a wireless communication module based on the Zigbee protocol for flexible and portable deployment. The plugs can be easily deployed on a power outlet and measure any electric appliances. Nearby connected plug meters (i.e., up to 10 meters range) will form a meshed network and communicate with each other. It measures active power (i.e., the amount of power that flows through the power meter) at $10s$ interval. Measurements from each plug meter are sent to a USB stick connected to a thin client (i.e., Raspberry Pi). The Raspberry acts as a data pooler and gateway. The measurements are then forwarded to the message queue in our distributed data warehouse.

Despite its easy deployment, measuring each device on the power outlet is expensive and considerably intrusive to users. We devise to use fewer plug meters (e.g., installed in the root of an electric circuit) to reduce costs and simplify the approach. Therefore, we use the power aggregation of plug meters to simulate measurement of an incoming line in a power circuit. Devices' switching states are then detected in the aggregate power consumption.

4.3.2 The Proposed Procedure

We propose a procedure for identifying electric low power monitor screens based on switching event detection.

Event Detection and Event Validation

Let X be the aggregate power consumption of computer monitors belonging to a set of individuals $J = \{j_1, j_2, \ldots, j_n\}$. We assign an event detection function $f_{ev}(\mathcal{O})$ to detect potential switching events on the ordered sequence of active power observation $\mathcal{O} = X_1, X_2, \ldots, X_T$.

The detection process starts with scanning power changes using a threshold of the absolute difference between consecutive power measurements ($wattThreshold = 10W$). Any events occurring within less than $60s$ (i.e., a $durationThreshold$) from the preceding event are ignored to prevent false event detection due to oscillations happening slightly after switching events. Once potential events are detected, care is taken to compute ΔX, i.e., the difference values after and before the events ($mean_{after}(ev)$ and $mean_{before}(ev)$). The average values of 30 samples before and after the event are considered in the computation. The value of ΔX needs to be higher than $10W$ to be considered as a real switching event. The precise procedure is presented as Algorithm 1.

Algorithm 1 Event detection and event validation

global variables
  $wattThreshold$, the minimum power change of consecutive measurements
  $durationThreshold$, the minimum duration of consecutive events
  $durationBetweenEvent$, the duration of two candidate events
  $validatedEvents$, the array of validated events
end global variables
procedure EVENT-DETECTION(X)
  Input: the aggregate power consumption
  Output: potential events
  for all sliding windows w in X **do**
    $range \leftarrow max(w) - min(w)$
    if $range > wattThreshold$ **then**
      if $durationBetweenEvent > durationThreshold$ **then**
        $events \leftarrow w$
  return $events$
procedure EVENT-VALIDATION($events$)
  Input: potential events
  Output: validated events
  for all event ev in $events$ **do**
    $\Delta X \leftarrow (mean_{after}(ev) - mean_{before}(ev))$ ▷ get power changes after an event occurred
    if $\Delta X > wattThreshold$ **then**
      $validatedEvents \leftarrow ev$
  return $validatedEvents$

Feature Extraction

Following the event detection, we examine all combinations of events to find start-end combination matches based on variance. That is, we calculate some parameters between the consecutive events ev_i and ev_{i+1}. Inspired by the field of dynamic systems [46] and statistics, we consider: rise time, overshoot, steady level, variable variance, and the mean of absolute difference. However, due to the limitation of the power meter in capturing positive-going transition in $10s$ interval, rise time and overshoot are less meaningful and thus not reported in this thesis (see [98] for details).

Delta X (ΔX) shows the difference value of average power before and after an event [130].

Steady level (or power level) indicates the power of a device (or set of devices) in a stable state. We use the histogram method to estimate the upper and lower levels [1]. We then assign the value that closest to the mode as the steady level.

Mean of Absolute Difference (MAD) captures the ripples during a device active period, Eq. 4.1 [62].

$$MAD = \frac{1}{N}\sum_{i=2}^{N}|X_i - X_{(i-1)}| \tag{4.1}$$

Variance measures how far a set of values are spread out from the steady level, i.e.,:

$$var = \frac{1}{N-1}\sum_{i=1}^{N}|X_i - \bar{X}|^2 \tag{4.2}$$

Finally, we compute the changes on steady level, mean of absolute difference, and variance due to event occurrences. The feature extraction process is shown in Algorithm 2.

Algorithm 2 Feature extraction

```
procedure FEATURE-EXTRACTION(X, validatedEvents)
    Input: the aggregate power consumption and validated events
    Output: the array of features
    Global variable: varThreshold, the maximum variance between two events
    for all event ev in validatedEvents do
        steadyLevel ← levels(X_{ev_i}, X_{ev_{i+1}})
        MAD ← mad(X_{ev_i}, X_{ev_{i+1}})
        variance ← var(X_{ev_i}, X_{ev_{i+1}})
        if variance < varThreshold then
            featuresArr ← {Δlevel; ΔMAD; Δvariance}    ▷ get the state-transition values after an event occurred
        return featuresArr
```

4.3.3 Techniques/Methods

Given a set of monitors $D = \{d_1, d_2, \ldots, d_u\}$, where individual $j_i \in J$ has a monitor $d_i \in D$, we assign classifier h_{recog} which labels the events detected by the event detection function f_{ev} with a monitor label d_i, formally $h_{recog} : f_{ev}(\mathcal{O}) \rightarrow d_i$. Several classification methods are possible:

k-Nearest Neighbor is one of the simplest learning techniques that works by finding the labeled samples nearest to a query and predict the class label with the highest votes [117].

Naive Bayesian is a simple probabilistic classifier that assumes features are independent given a class label [122].

Neural network is a nonlinear statistical model for regression or classification, typically represented by a network diagram [37]. This algorithm consists of neurons that are interconnected and can be arranged in various architectures.

4.3.4 Metrics

We investigate the performance of event detection and device classification to evaluate the occupancy detection based on monitor activation.

Metrics: Event Detection

Precision and sensitivity are employed to evaluate the detected events. The earlier is the rate of correct classification over all events detected by the system, while the latter is the proportion of actual events that are correctly identified over all occurred.

Metrics: Device classification

In classifying an appliance among similar appliances, the classification accuracy may be lower than expected. Thus we relax the classification to $top - n$ classification [24]. The classification is considered correct when an event is classified as one of the most likely n classes. It allows, for example, to scale down an occupancy investigation for an event of interest to a reasonable size set of suspect occupants. We use top-n classification accuracy per day with $n = 2$.

$$Accuracy = \frac{\text{correctly predicted events}}{N}, \tag{4.3}$$

where N is the total number of events being classified.

The average accuracy can then be computed as Eq. 4.4:

$$Accuracy_{avg} = \frac{1}{d_{days}} \sum_{d=1}^{d_{days}} accuracy\_day_d \tag{4.4}$$

where d_{days} is the number of observation days.

4.4 Experiments

We detect monitor activation/deactivation from the aggregate power consumption as an indication of occupancy in offices. This section describes datasets that we have collected and the experiment setup.

4.4.1 Data

To test the performance of the proposed approach, we collected the real power consumption of monitors in two different offices. The first office is an academic building located in the Zernike Campus of the University of Groningen, The Netherlands. Another office is a commercial office of a mid-size software house company located in Germany. In the second office, there are employees and internship students developing application products using laptops and external monitors.

Dataset A

In the university office, the involved monitors belonging to four graduate students are not necessarily the same due to the different periods of procurement. The data collection phases were divided into several parts. The first and second part of the experiment took place from March 13, 2017 until March 31, 2017 (for training) and from April 17, 2017 to June 22, 2017 (for testing). Another dataset was collected to study the relation of monitor activation to occupants' presence. It took place from July 19, 2017 until July 27, 2017.

Dataset B

In the commercial office, seven of ten monitors belonging to employees are of the same type (i.e., having the same brand and the same 24-inch screen size). The data collection started from April 24, 2019 until May 2, 2019 for ground truth observation, and until July 21, 2019 for training. For testing, the experiment took place from September 18, 2019 until October 15, 2019. Monitor power consumption was measured in eleven workspaces in the office. The size of the office is larger than in the first office. We include a set of volunteers $\mathcal{W} = \{W1 - W7, W10, W11, W12\}$ with a set of monitor loads $J_B = \{G1, G3, G4, G5, G7, G10, G11, G14, G16, G18\}$.

The actual presence of people, i.e., the ground truth, was taken based on the paper-based diary and human observation and improved using the measured monitors' power consumption data. If, due to some failure, we miss readings, we keep the previous valid one. This approach is common to mimic the constant consumption of simple devices, such as LCD monitors.

4.4.2 Setup

We manually observe the occupants' presence during a week observation and annotate the time. Additionally, we infer presence using a simple power consumption threshold, as in [26]. We compare the manual presence observation in the workspace (i.e., 8:00 AM to 9:00 PM) with the consumption upper threshold (i.e., $5W$). We then report the agreement to show the relation of monitor activation with occupant presence.

Prior to commencing the study, an appliance signature database is built by collecting individual power consumption during the training period. It is then followed by a feature extraction process based on switching events on the smoothed power readings. We apply the moving average technique to smooth ripples on the power consumption readings. The feature extraction from individual consumption signatures results in 252 and 363 training instances for dataset A and B, respectively. The monitors' features are then used to supervise classifier models.

After the training features are extracted, we combine the individual power consumption to obtain aggregate power consumption. Note that, this step aims to acquire aggregate power consumption, i.e., simulating a fewer number of power meters in the root of the electric circuit, as we collect data per monitor in this experiment. We build several aggregate consumptions by adding one-by-one a monitor per set. In dataset A, three power aggregations are formed, each with a combination of two to four monitors that represent up to four participants. In dataset B, eight sets are formed, each with a combination of three to ten monitors, representing up to ten participants. Tables 4.1 and 4.2 show the details of monitor combination. The aggregate power consumption is then analyzed as an input.

Once the classifier models and power consumption aggregation are ready, the final phase is to recognize the monitors on the aggregate power consumption. Similar to the signature preparation step, this process consists of standardization, event detection, and feature extraction. However, these are applied to daily aggregate power consumption, according to Algorithms 1 and 2. Finally, the outputs of this process

Table 4.1: Power consumption aggregation for testing dataset A

Set	Monitor set	#persons	#presence	#events
A1	{M4, M10}	2	12 days	82
A2	{M4, M7, M10}	3	28 days	168
A3	{M4, M7, M10, M14}	4	32 days	279

Table 4.2: Power consumption aggregation for testing dataset B

Set	Monitor set	#persons	#presence	#events
B1	{G4,G7,G11}	3	12 days	169
B2	{G4,G7,G11,G14}	4	12 days	169
B3	{G4,G7,G11,G14,G17}	5	15 days	196
B4	{G3,G4,G7, G11,G14,G17}	6	15 days	301
B5	{G3,G4,G5, G7,G11,G14,G17}	7	15 days	331
B6	{G3,G4,G5,G7,G10,G11,G14, G17}	8	15 days	398
B7	{G3,G4,G5,G7,G10,G11,G14, G17,G18}	9	15 days	414
B8	{G1,G3,G4,G5,G7,G10,G11,G14, G17,G18}	10	15 days	501

are supplied to classifiers for appliance identification.

In this work, we compare different feature combinations to see the performance: (i) the difference in power before and after an event (ΔX), (ii) steady-level, and (iii) steady-level, MAD, and Variance.

4.5 Results and Discussion

This section begins by discussing the agreement of monitor activation to occupancy. It then elaborates on the performance of switching event detection and device classification from the aggregate power consumption.

4.5.1 Occupancy via monitor activation

Prior to undertaking the appliance recognition investigation, the relation of monitor activation to occupancy states was observed to see how far the monitor activation may represent individual occupancy. Table 4.3 shows the agreement of monitors' power consumption and occupant presences over a five-minute interval. One can see that the power consumption of each monitor provides a reasonable estimation of users' presence, reaching 98% agreement. Weaker relations, reaching 83% agree-

Table 4.3: Occupancy accuracy over a 5-minute interval

Employee ID	Presence days	Agreement
Dataset A		
W1	7d	96.7949
W2	5d	89.8718
W5	4d	83.4936
Dataset B		
W1	4d	95.833
W2	2d	93.59
W3	2d	90.705
W4	4d	82.692
W5	4d	98.077
W6	6d	97.329
W7	3d	91.026
W10	2d	97.115
W11	5d	97.821
W12	5d	95.897

ment, are due to monitor activation while someone is away. For example, the monitor is waiting for a timeout to automatically put on sleep mode after plugged out from a source (e.g., a laptop).

Based on the empirical observation, it seems that the monitor activation, if recognized correctly, can reveal the present state of most of the individuals. That is, 7 of 13 participants (i.e., 54%) have only 5% error or less, and 10 of 13 participants (i.e., 77%) have up to 10% error. The problem is then how to recognize the monitor activation with a fewer number (or even a single) of power meters. We thus investigate the switching event recognition on the aggregate power consumption.

4.5.2 Event Detection Rate

Following Algorithm 1), we perform switching event detection on the aggregation of monitor consumption. The result is shown in Figure 4.1. In Dataset A, the sensitivity of event detection is higher than precision, reaching more than .90 sensitivity for the three monitor sets. In dataset B, it drops reaching about .70. It is because more events are not detected in dataset B that lower the sensitivity. This result corresponds to the nature of the dataset B, where it consists of monitors with lower power consumption. The nature of dataset B brings worse event detection performance (i.e., in terms of sensitivity) than in dataset A.

The proposed event detection results up to .80 precision for both datasets. In

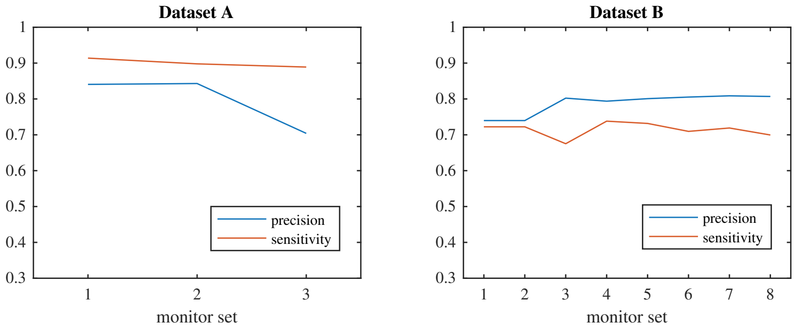

Figure 4.1: Precision and sensitivity of detected events

monitor set A3, however, the precision declines reaching .70. The reason is that the aggregate consumption ripples and sometimes drops suddenly up to $20W$ for a short time (up to 3 minutes). This behavior is affected by adding monitor M14 in dataset A. It triggers false detected events and thus reduces the precision. This result shows that the event detection is affected by the pattern of monitors' power consumption.

4.5.3 Appliance Classification

We compare several feature sets and different techniques to classify particular monitors from power consumption aggregation. The comparisons are summarized in Figure 4.2.

Feature ΔX is barely sufficient to classify monitor sets, especially in dataset B. Features steady level, MAD, and variance can improve the classification in both datasets. Other features that do not contribute positively are not shown in this thesis (we refer to [98] for details).

Given only feature steady level, *k*-NN has a performance of 77-88% in dataset A, but of only 24.3-53% in dataset B (see blue bars in the second row). The performance can be explained by the fact that in dataset A, the monitors are relatively distinguishable, considering only feature power level. In dataset B, the monitors are not distinguishable based only feature steady level due to the monitors' homogeneity. Adding features MAD and variance helps to distinguish the monitors, reaching 39-64% in dataset B (see the right graph in the third row).

Neural Networks (NN) generally perform well for considered features (either only steady level or steady level, MAD, and variance) in both datasets (see yellow bars in the second and third rows). More specifically, NN outperforms the *k*-NN using the same features (steady level, MAD, and variance), reaching 72-84% and

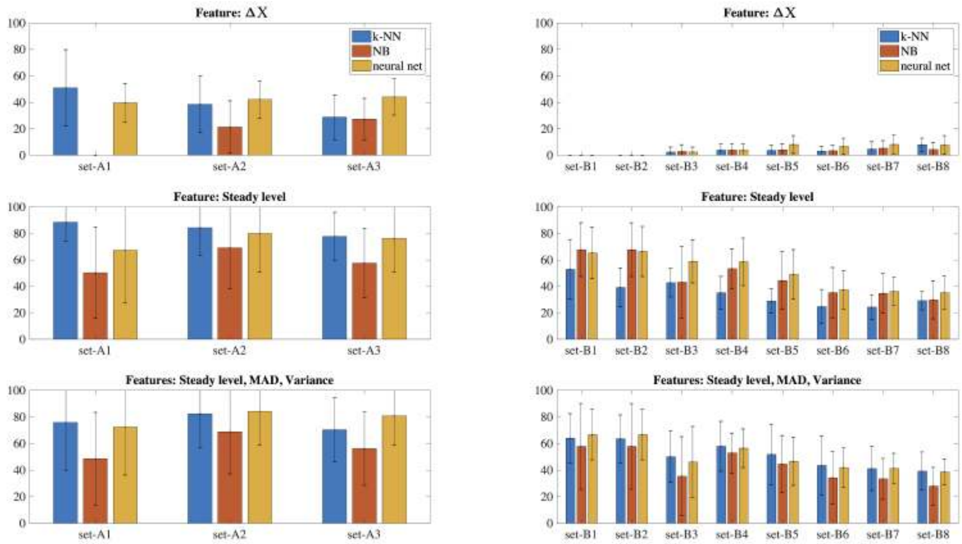

Figure 4.2: Monitor identification on the dataset A (left) and dataset B (right) using various feature combinations: the 1^{st} row: ΔX; 2^{nd} row: Steady level; 3^{rd} row: Steady level, MAD, and variance

39-67% in dataset A and B, respectively (see yellow bars in the third row). The average accuracy on dataset A, however, are still below the classification using *k*-NN with feature only steady level that reaches 77-88% (see blue bars in the left graph in the second row). In this case, *k*-NN outperforms NN when to recognize small number distinguishable appliances, such as in dataset A. *k*-NN performance drops in the higher number of non-distinguishable appliances such as in dataset B. It might be because the *k*-NN works based on the distance to the samples. In contrast, NN works by optimizing networks (backpropagation), so giving a higher chance to predict the label .

NB mostly performs worse than the other techniques, except in monitor set B1 and B2 where this technique slightly outperforms NN (see red bars in the right graph in second row). NB is a generative model that works based on the class conditional densities. On the contrary, discriminative models (e.g., *k*-NN and NN) focus on determining decision boundaries, thus perform better in classification tasks.

Finally, one can see that the performance declines as the number of monitors increases. It indicates that the more electrical appliances involved in the aggregate power consumption, the more difficult detection.

Figure 4.3 and 4.4 illustrate the proposed event detection and monitor classifica-

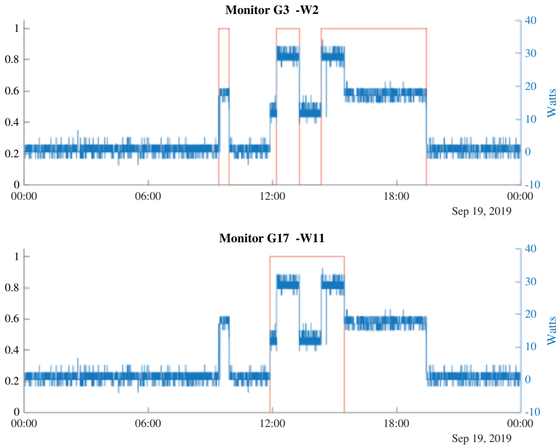

Figure 4.3: Actual monitor activation states on September 19, 2019.

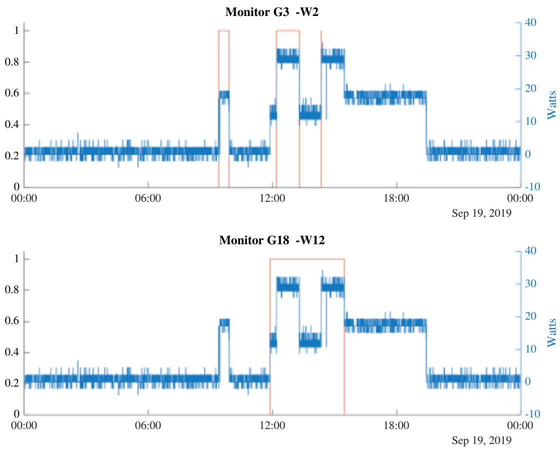

Figure 4.4: Prediction of monitor activation states (top-n classification) on September 19, 2019.

tion of monitor set B4 using NN and features steady level, MAD, and variance on September 19, 2019. There were only employee W2 and W11 who are present on the corresponding date. The blue line represents the aggregate power consumption, while the red line represents an employee's binary presence state. It can be seen in Figure 4.3 that there are six real switching states of monitor G3 belonging to employee W2 (the graph on top) and two switching states of monitor G17 belonging to employee W11 (the graph at the bottom). The proposed procedure can detect five of the six switching events correctly, but it misses the switching OFF event at around 7.00 PM, as shown in Figure 4.4. The system can detect events at noon and at 5.00 PM, however, it misclassifies the monitor G17 as monitor G18 (the graph at the bottom).

4.6 Conclusion

This study set out to explore occupancy detection in an office based on the activation of computer equipment. We propose a procedure to recognize low-power electric loads (i.e., monitors) from the aggregate power consumption (e.g., as measured in the root of electrical lines). Some features inspired by the field of dynamic systems and statistics are investigated. The meaningful features are the value changes of power in a steady level, MAD, and variance. In our procedure, events are firstly detected based on thresholds. Then, based on the events, features are extracted to be used in the classification process. This step is then followed by examining the event detection rate in terms of precision and sensitivity. We also examine classification performance per day based on various combinations of features. Finally, the average of top-n ($n = 2$) classification accuracy is provided. The present study contributes techniques to electric load identification. The proposed classification is tested in actual use cases in offices with various typical work situations. That is, one is at a university office, and another one is at a commercial office of a software company. While both offices have occupants who work in computer-related fields (e.g., researching and developing software and applications), these offices are different in terms of employees' behavior and available monitor devices.

The experiment begins by observing the conformity of individual monitor activation to real occupancy based on individual measurements. It is shown that monitor activation has a relationship with user presence; namely, the presence of up to 77% of the total of 13 participants may be inferred with less than 10% error using monitor activation. Even further, more than half of the participants have only 5% error or less. Recognizing monitor activation on the aggregate power consumption is challenging as it thoroughly relies on the accuracy of detected events. Furthermore, it is challenging to match a switching OFF event with the previous switching ON event. The more devices involved and the lower amount of power consumed, the harder to detect switching events.

In summary, the experiment shows the relation of monitor activation and employee presence. It also demonstrates the feasibility of low-power monitor detection based on events. When the aggregate power consumption consists of different monitors, they are relatively distinguishable, reaching 80% top-n accuracy per day (i.e., in dataset A). The approach struggles and cannot distinguish the loads when the monitors are similar, as shown by the 39% top-n accuracy per day for detecting up to ten monitors in dataset B. Another strategy (e.g., moving windows) may avoid event detection problems in appliance detection, as we will explore in the next chapter. In this way, appliance signatures are taken at a particular time without detecting events in advance.

Chapter 5

Windowing based Power Meter Classification

5.1 Overview

In Chapter 4, we investigate the electric load identification for occupancy detection based on switching state detection in active power measurements. The system, however, suffers from miss detected events. It is also difficult to distinguish similar low consuming power appliances (e.g., monitors in an office). In this chapter, we explore electric load identification without event detection to avoid miss-detection problems. We apply a moving window segment-by-segment on the aggregate power consumption (i.e., sliding windows). The power consumption readings consist of some variable measurements, such as active power, reactive power, and apparent power. Every time the window slides, a classification of the power readings is performed. We analyze the readings sequentially and non-sequentially to see if historical observation brings advantages in recognizing the electric loads.

In addition to electric load identification, we also mine power consumption data for fine-grain occupancy detection using the same approach. Namely, we utilize sliding windows on the composite loads that correspond to occupants' power consumption during the occupation. There might be patterns of power consumption of each individual during his occupancy in the office to detect his presence. The patterns allow by means of low-intrusive identification to obtain fine-grain occupancy information (e.g., identifying who is sitting in a particular office). We find that the detection and recognition of up to three employees are possible based on the aggregate power readings with a sub-metering system in a shared room. The finding complements earlier studies that typically either produce coarse-grain contexts (e.g., vacancy or occupancy state of a house) or require more meters to obtain finer-grain contexts (see Figure 2.2).

The clamp-based meter technology and experiment design are described in Section 5.2. The discussion of the first experiment about electric load identification is provided in Section 5.3, followed by the second experiment to reveal occupancy in Section 5.4. Finally, lessons learned are presented in Section 5.5.

5.2 Design

We utilize power meters installed in a room (i.e., sub-metering). It is assumed that dedicated circuits of PC equipment are available in the office, and power meters are attached to them to measure the aggregate power consumption. The dedicated circuits are commonly assigned to eliminate the risk of outages and to guarantee the quality of supplied power [128, 34], or serving some offices based on their location in a building (e.g., [22]). Another assumption is that the occupants do not use a remote desktop application, hence the computer usage implies physical presence in the office. We then classify sensor readings per instance and per sequence (i.e., consecutive instances in a historical observation) for identifying office appliances and individual presence in an office.

5.2.1 Sensing Technology

We use Smappee power meters, a clamp based one that supports electricity measurement with five observed variables: active power, reactive power, apparent power, power factor, and electric current. Hence, we distinguish the notation of power readings with more features ($\vec{X}$) in this chapter. Active power is the amount of power that flows through the power meter (measured in Watt). Reactive power is the dissipated power as a result inductive or capacitive components in appliances (measured in Volt-Amps-Reactive, VAR). Apparent power is the product of the root-mean-square voltage and the root-mean-square current (measured in Volt-Amps, VA). Power factor or cosphi (in percents) represents the ratio of the active power flowing to the appliance divided by the reactive power. Electric current is the amount of electron flowing through the clamp (measured in ampere). All these variables are collected in a five-second interval.

By default, the meter supports data collection to up to 5 minutes. Data will be kept up to one month in the Smappee cloud. To acquire data more frequently (i.e., up to $5s$), we use a local REST API by connecting a thin client (i.e., Raspberry Pi) to the built-in WiFi module. We use the local API by implementing Smappee-pooler[1] module on the thin client. This module polls and forwards the measurements to the message queue (MQTT). We develop a script to consume messages coming to the queue and send them to our distributed data warehouse[2].

To study information extraction from power meters, we initially install a Smappee power meter in a work desk to investigate the aggregate power consumption. We then install the meter to measure at a larger scale (i.e., an office room), where three

[1] https://github.com/NMichas/smappee-local-mqtt
[2] Website: https://github.com/rug-ds-lab/system-core

graduate students sit during work hours. In the first case, we extract individual appliance loads from the aggregate power readings, while in the second one, we extract composite loads of several appliances belonging to occupants. We discuss the experiment cases in Sections 5.3 and 5.4, respectively. In the following, we describe common techniques to analyze the data.

5.2.2 Techniques

Given a set of individuals $J = \{j_1, j_2, \ldots, j_n\}$ and a set of electric features $E = \{e_1, e_2, \ldots, e_h\}$, the aggregate power reading at time t is

$$\vec{X}_t = [f_1(\vec{x}_{t,e_1}), f_2(\vec{x}_{t,e_2},) \ldots, f_h(\vec{x}_{t,e_h})],$$

where $\vec{x}_{t,e_k} = [x^{j_1}_{t,e_k}, \ldots, x^{j_n}_{t,e_k}]$ is the vector of feature readings e_k of individual load $j_i \in J$, and $f_k(\vec{x}_{t,e_k})$ is the aggregate function of feature e_k over all individual load $j_i \in J$. An example of such a function is the aggregation of active power $f_a(\vec{x}_{t,a}) = \sum_{i=1}^{n} x^{j_i}_{t,a}$.

We aim to identify component loads from the aggregate power readings $\vec{X}$. The loads can be electric appliances belonging to an occupant (i.e., investigated in Section 5.3) or composite loads consumed by occupants in a shared office room (i.e., investigated in Section 5.4). The latter correspond to the total power consumption of an occupant, and thus, the detection of such composite loads may indicate occupancy.

We investigate several techniques to extract information from aggregate power readings. First, in per instance classification, the classification task is performed by determining a class label of the unlabeled aggregate power consumption. Second, we perform sequence classification based on sliding windows. Finally, generative classification is investigated.

Per instance classification

This approach assumes independent data points. Thus, there is no sequential correlation in power presentations, and the classification may be performed per instance. This technique maps the aggregate power consumption ($\vec{X}_t$) into a class label using classification techniques as follows.

***k*-NN** is a classification method that infers a class label by comparing a query instance to stored training instances without constructing any classification models, so-called instance-based learning, or lazy learning method [9]. A label is assigned to each of the queries based on the majority labels of the nearest neighbors.

SVM works by generating a nonlinear decision boundary. It maximizes the margin of separation by training on a dataset. This method is generally designed to solve binary classification problems. One way to solve the multi-class problem with SVM is by one-against-one approach that combines several binary classifiers [68].

Neural Networks or feedforward neural networks is an algorithm that fits a non-linear estimator from a feature vector in a training set [37]. Generally, it is formed by input, output, and one or more (non-linear) hidden layers. Input layers consist of a set of neurons representing input features. Hidden layers consist of a set of neurons that transform input values from the previous layer with weighted linear summation followed by a non-linear activation function (e.g., hyperbolic tan function). Finally, the output layer receives values from the last hidden layer and transforms them into output values.

Sequence classification (Temporal data type)

We investigate sequence classification to study the correlation between consecutive instances in historical observation and its advantage in recognizing occupancy. The intuition is that once a person is present and actuating devices, the devices will stay activated for a long period. We provide an ordered list of power measurements into classifiers.

This technique performs classification based on non-overlapping sliding windows. Given an observation $\mathcal{O} = \vec{X}_1, \ldots, \vec{X}_T$, we define sub-sequences s_k with window length w. That is, $s_k = \vec{X}_{w(k-1)+1}, \ldots, \vec{X}_{w \cdot k}$, $1 \leqslant k \leqslant T/w$. We thus classify the sub-sequence s_k using several methods as presented next.

Sequence distance-based classification measures the similarity between a pair of sub-sequences [133]. We adapt the *k*-NN to sequential classification by appending the elements of power measurement vectors to form a list or sequence. An input sequence is assigned to a single class. The sequencing means that before a classification is being done, the full sequence needs to be prepared. The *k*-NN classifier then compares the distance between a query sequence s and train data sequences.

Similar to the per-instance classification, classification based on distance or similarity measures also works in series or temporal data type. One of distance measures is L_p-Norms, which requires two compared time series or sequences of the same length. Special case of L_p-Norms with $p = 2$ is the Euclidean distance, where

N is the number of instances:

$$dist(s, s') = \sqrt{\sum_{i=1}^{N} (s[i] - s'[i])^2} \tag{5.1}$$

Looping neural networks. We extend the traditional, feedforward neural network architecture by considering recurring events. That is, we apply Recurrent Neural Networks (RNN). RNN is a type of neural network that takes the previous output as the next input value in a sequence, as illustrated in Figure 5.1. In typical neural networks (i.e., feedforward), an input vector flows through a hidden layer resulting in an output vector without forming any cycles. While in RNN, the output from previous time step ($t - 1$) is fed to every neuron of the current time step (t). This characteristic offers advantages when used in time-series data, as the model can theoretically see the context of data readings from the previous occurrence. In our work, we design a RNN that takes a chunk (or a sub-sequence) of a whole long sequence. The chunk consists of power readings during a specific period. The network then infers a label as illustrated in Figure 5.2.

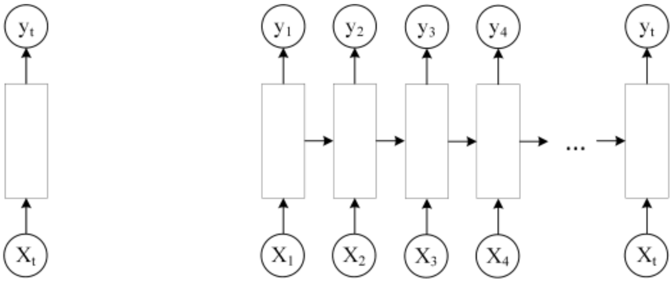

Figure 5.1: Feed forward neural network (left) and recurrent neural networks (right)

When data passes hidden layers multiple times, it might get flattened, also known as the vanishing gradient problem, and affect RNN's memory from the past readings. One solution is to utilize the Long Short-Term Memory (LSTM) architecture which can decide when to forget or keep the current input for the next output using logical gates. The comparison of standard RNN and LSTM is given in Figure 5.3.

Model-based classification (HMM)

Finally, we model the problem based on Hidden Markov Models. The reason is that the electric loads are 'hidden' as they cannot be observed directly, while we have access only to observe the aggregate power consumption. The HMM is a generative model as it relies on the class conditional densities and prior class probabilities to

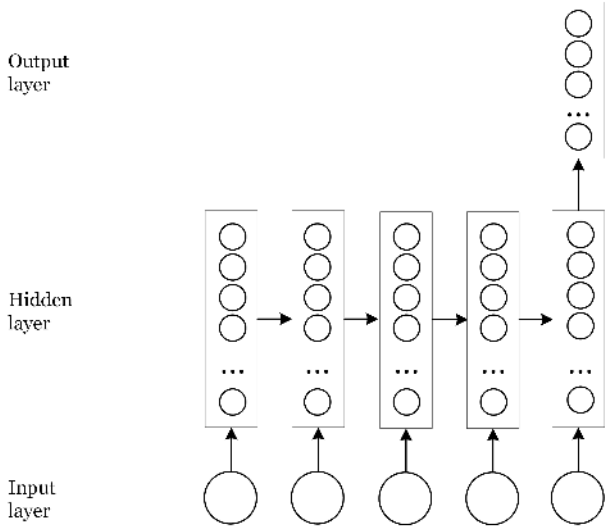

Figure 5.2: Recurrent neural networks

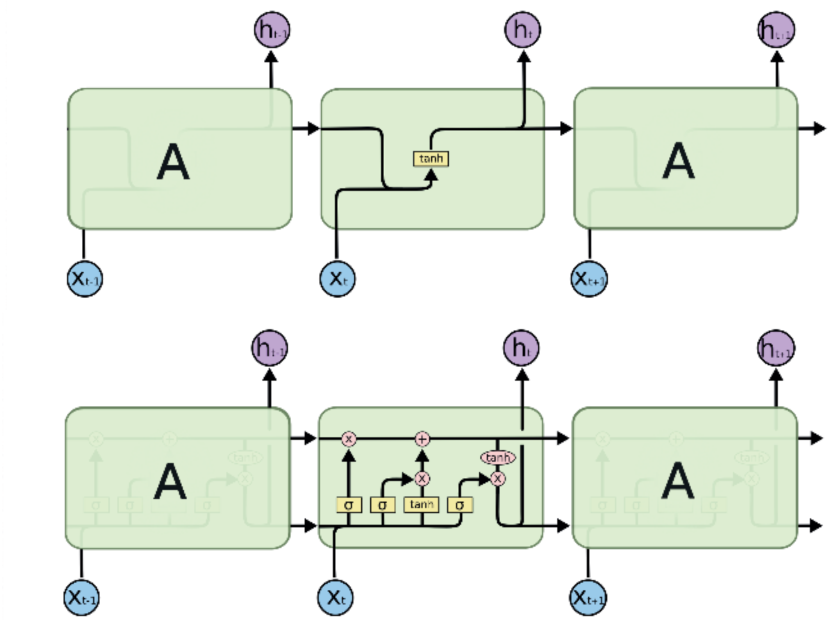

Figure 5.3: Standard RNN (top) and LSTM cells (bottom), adopted from [87]

find posterior values. An HMM model $\mathcal{M}_{j_i}$ is assumed to generate a sequence of load j_i at time $t = 1, \ldots, T$, namely, $s = x_1^{j_i}, \ldots, x_T^{j_i}$. The parameters of model

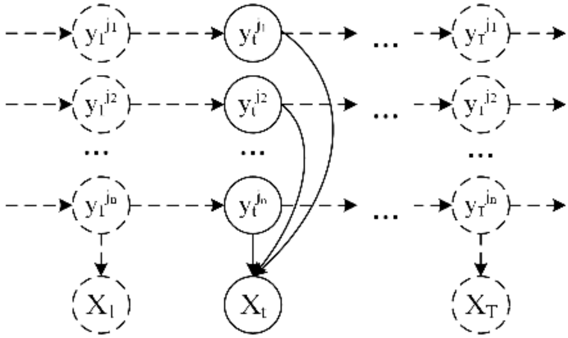

Figure 5.4: Factorial HMMs

$\mathcal{M}_{j_i}$ are learned in the training phase (e.g., using the Baum Welch algorithm). The model $\mathcal{M}_{j_i}$ then assigns a sequence of class labels with the highest likelihood to the rest of the sequence (e.g., the z-length query sequence $s' = x^{j_i}_{T+1}, \ldots, x^{j_i}_{T+z}$).

Note that in building the model, we only consider a single observed variable (i.e., active power), as considering all the observed variables to predict a single hidden variable inflates the probability space immensely, and we do not have enough training data to calculate the parameters properly.

We construct an HMM chain for each load $j_i \in J$. As we measure the total power consumption $X_t = \sum_{i=1}^{n} x_t^{j_i}$, the model is generalized to Factorial HMM [48], as illustrated in Figure 5.4. We then find the optimal sequence of hidden states by means of the Viterbi algorithm [101].

5.2.3 Metrics

In order to evaluate the classification, we provide total accuracy to indicate overall performance. We also consider Cohen's Kappa measure [29] to eliminate the accuracy bias due to imbalanced class distribution.

Accuracy defines the correct prediction of all class labels over the total number of predictions that have been made, specifically,

$$Accuracy = \frac{\text{correctly predicted windows}}{N}, \tag{5.2}$$

where N is the total number of windows being classified.

Cohen's Kappa measures the agreement between accuracy of the system to the accuracy of a random system, as shown in Equation 5.3. The total accuracy is an observational probability of agreement while the random accuracy is a hypothetical expected probability of agreement under an appropriate set of baseline constraints [29].

$$kappa = \frac{totalAccuracy - randomAccuracy(RA)}{1 - RA} \tag{5.3}$$

where RA is the sum of the products of reference likelihood and result likelihood for each class. Mathematically,

$$randomAccuracy = \frac{\sum_{c \in C} actualclass_c * predictedclass_c}{N^2}$$

5.3 Experiment-1: Office Appliance Identification

First, we experiment electric load identification with an assumption that a power meter is available in a workspace, measuring the aggregate power consumption of a user. This section describes conducted experiments. It then provides the results and discusses the findings.

5.3.1 Data

We measured power consumption in a workspace of the University of Groningen, The Netherlands, during several weeks between June 2017 and February 2018 [96]. The data collection was based on appliances that are known in active states. Specifically, we note the timestamp and the corresponding class label when a particular appliance is set off. We involved two LCD monitors, a CPU, laptop, and a portable heater. The appliances were connected to the power source through a power extension. We clamp the meter on the extension, as illustrated in Figure 5.5. In total, we collected 92,580 data points of 14 classes.

The data readings were normalized to scale the measurement to a range between 0 and 1. We applied k-fold cross-validation with $k = 5$ to assess model generalization to the dataset. Several feature sets were investigated to compare recognition performance based on various measurements. The sets consist of FS-1 (active and reactive power), FS-2 (active power, current, and power factor), and FS-3 (Active-, reactive-, and apparent-power, current, and power factor).

5.3.2 Setup

In appliance identification experiment, we assign a classifier h_{recog_2} to recognize a class label c that represents an appliance or a combination of appliances that

Figure 5.5: Office appliance identification

contribute to the total power readings $\vec{X}$. Formally, $h_{recog_2} : \vec{X} \rightarrow c, c \in C = \{\{\text{laptop}\}, \{\text{laptop and monitors}\}, \ldots\}$.

We design an RNN architecture such that we can input a sequence (e.g., 60 data points during 5 minutes) of a set of electric features E to the model. The sequences are from non-overlapping moving windows, as this performs better than overlapping windows according to our experiments [121]. For each sequence, we expect a class label to be learned or classified. We feed 20 chunks of the sequences in an iteration to reduce the number of looping needed to complete one epoch, thus, speeding up the learning process. We apply either one LSTM layer or two LSTM layers (LSTM^{L2}), stacked in the hidden layer of the network.

Each LSTM layer consists of h_t hidden states, where $h_t = 20$. The reason of the chosen architectures is that the layer size is not more important than the layer depth [52], hence we stick in the fixed size of h_t and change the level of the hidden layer. We apply Adam optimizer [65] and determine the learning rate of 0.001 to optimize a cost function during iteration. The cost function is based on cross entropy [50].

We determine the number of epochs to be completed to make sure that the model has learned sufficiently without memorizing the training data (overfitting). To do so, we implement the early stopping strategy. We evaluate the model performance on a validation set and save the best model snapshot when it outperforms the previous best winner. We terminate the training when the network does not perform better after the i-th epoch, where $i = 100$ when the maximum number of epochs is 400, and $i = 50$ when the maximum number is set to 200.

To evaluate the approach, we also consider k-NN (with $k = 7$) and SVM with linear- and polynomial-kernels (with regularization parameter $C = 1$) [21]. We choose k-NN as this is a simple yet powerful technique. The critical issue is the efficiency of the approach in the classification process as the running time is linear with the size of the data set. SVM is a widely used technique for classification due to its ability to generate nonlinear decision boundaries. Both k-NN and SVM are designed to classify on the basis of one single point of measurement, instead of giving a sequence of data as input.

5.3.3 Results and Discussion

The classification results are shown in Table 5.1. In general, the performance indicated by the accuracy and Kappa measure are comparable. The reported measures are the average of the completion of 5-folds cross-validation. The RNN with LSTM based classification delivers Kappa between 60-97%, depending on the network configuration, the number of iterations on the training phase, and the feature set taken into account. LSTM and LSTM^{L2} differ in terms of cells number. The former uses a single LSTM cell while the latter uses two. Using LSTM, we can achieve a Kappa measure of up to 90.1%, while based on LSTM^{L2}, 96.8% of the same measure can be achieved. These results can be improved up to 97.6-99.4% by increasing the number of training epochs. The higher number of epoch iterations are allowed, the better results are obtained. It is because the approach is based on iterative optimization. The approach tries to adjust model's weights along the iterations to find the optimal network configuration.

In general, the classifications using FS-3 achieve higher performance than the other sets, reaching at least 90%. SVM with the polynomial kernel is an exception. While SVM with the linear kernel can achieve Kappa measure of about 82-93%, the same classifier with the polynomial kernel (degree 3) reaches roughly 40% for feature FS-2 and FS-3. The reason is that, in SVM, the decision boundary is decided by a hyperplane that is shaped by kernel functions. The model with polynomial kernel tries to overfit the training data and fails to classify the rest of the data, while the linear kernel works quite well. The k-NN based inference results up to 99.9% Kappa measure on the all set of predictors. This result might indicate that feature combinations are distinct and non-overlapping, thus can be classified well with k-NN. As a comparison, when we put only active power, k-NN will return worse result than the one reported in this thesis.

5.4 Experiment-2: Occupancy Detection

The second experiment is the extension of the previous investigation. Given the promising results in appliance classification using k-NN and RNN LSTM for a single user, we expand the research by involving more users. We infer occupancy of three graduate students with disguised ID. To approach the real case, each user may use

Table 5.1: Cohen's Kappa measure of classifiers with different feature sets

Method	Accuracy			Kappa measure			Remarks
	FS-1	FS-2	FS-3	FS-1	FS-2	FS-3	
LSTM	.73	.629	.909	.709	.599	.901	iteration max 200 epochs
LSTM^{L2}	.779	.719	.97	.761	.696	.968	
LSTM	.819	.782	.978	.804	.765	.976	iteration max 400 epochs
LSTM^{L2}	.848	.841	.995	.836	.829	.994	
k-NN	.999	.996	.999	.999	.996	.999	$k = 7$
SVM (lin)	.84	.92	.934	.827	.914	.928	$C = 1$
SVM (poly)	.31	.444	.456	.24	.396	.408	degree=3, $C = 1$

any appliances. We do not observe the individual appliances' signatures, but we look for patterns on the occupant signatures.

5.4.1 Data

We collected power consumption in a shared room office at our university between May 31, 2018 and September 11, 2018 [94]. We used two Smappee power meters. One clamp was dedicated to measuring the total power consumption, while the other clamps were attached per user for behavior investigation, including (i) ground truth refinement, (ii) user energy profile construction, and (iii) train data labeling. In actual deployments, only one smappee power meter is needed. That is, one clamp measures total consumption in the circuit-breaker of a room, and the other two clamps alternately collect labeled training data (i.e., one class after another).

The ground truth data was manually collected in a spreadsheet document. Due to ground truth incompleteness during a long period of data collection, we generated class labels by applying a threshold (i.e., 20-Watt) to the per individual power consumption. The validation of the generated class labels was done by camera recordings from February 4, 2019 until February 8, 2019. It was shown that two of three participants had the Kappa measure of agreement of .97 between the occupancy based on the power threshold and the occupancy based on the camera observation. Another participant, unfortunately, missed the experiment at the time of camera deployment.

We apply normalization on the data readings to scale measurement values to a fixed range (i.e., between 0 and 1). We then relabel the sequential data to make sure the training process is supplied with sequences that represent a homogeneous class. We only process full-length and partially complete sequences. Full-length sequences refer to w-length consecutive instances with the same label. The partially complete sequence consists of ordered instances with the same label, but the length is slightly

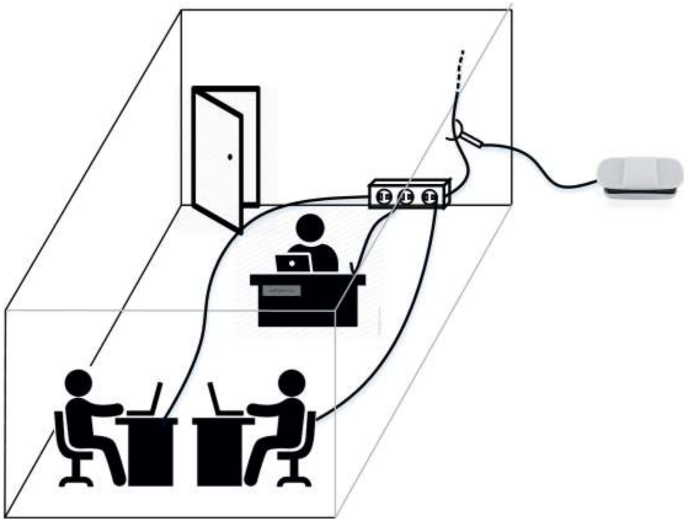

Figure 5.6: Occupancy detection in a shared office

less than w. As for the latter, we impute the last value to form a full of w- length consecutive instances. We investigate a different combination of raw measurement time series variables to discover potential patterns formed during occupancy. This includes (i) Watts, VAR, VA, current, and cosphi as proposed by Akbar et al. [10], (ii) Watts only, as the most basic measurement component in a power meter, (iii) VAR and Watts, as proposed by Hart [55], and (iv) Watts, current, and cosphi. We also add features that indicate the range of the time of day when a measured value occurred. We mark the instances as a member of the corresponding time of day. The markers that are represented using one-hot-encoding are considered as additional features to the estimators.

5.4.2 Setup

In occupancy detection experiment, classifier h_{occ} is assigned to identify a class that summarizes the presence states of all individuals in a particular room. Formally, $h_{occ} : \vec{X} \rightarrow Y$, where $Y = \left\{y^{j_1}, y^{j_2}, \ldots, y^{j_n}\right\}$ and $y^{j_i} \in \{0, 1\}$.

We divide the collected data into training, validation, and test set. We exclude 15% of the total data set for testing purposes (the hold-out test set). Two directions on the selection data portion are investigated. First, a subset of data is selected after doing randomized shuffling by preserving the proportions of the class prior probabilities. Second, data division is done based on the historical occurrence, which retains historical order and corresponds to a real-world scenario. In the first phase of the experiment, 85% of the total data set are used for exploring classification models and finding the optimal model parameters. This phase is done through 5-fold cross-validation. In the next phase, the parameters are used to retrain the model using the combination of training and validation set as the full training set. In the final phase, the trained model is used to classify the final 15% of the data to evaluate the performance.

We use similar architectures as in the previous experiment (i.e., the same Adam optimizer, and LSTM cell activation function). The number of hidden neurons in a single LSTM cell and epochs are parameters to tune in the beginning. We also tune various sequence lengths (1,2,5,10, or 20 minutes) and different features (active power only; active and reactive power; active, current, cosphi; and with the addition of time of the day every 45 and 60 minutes). We utilize our university's computer cluster (Peregrine High-Performance Computing cluster[3]) to experiment with several parameters simultaneously.

5.4.3 Results and Discussions

Based on hyperparameter optimization that works best on the training validation set, we set several parameters as following. Using the data division procedure, we find that the best k for k-NN is 11 for both sequence and non-sequence analysis. For neural networks, the best number of neurons is 30. We apply LSTM cells as activation function with Adam Optimizer. We apply 100 hidden neurons in a single cell LSTM layer and an epoch of 744. Using the saved parameters, we classify the test set and presented the results in Table 5.2.

Table 5.2: Classifier accuracy and Kappa measures on the shuffled test set

Techniques	Accuracy	Kappa
k-NN	.966	.937
k-NN^{seq}	.970	.946
NeuralNet	.941	.891
RNN LSTM	.964	.934
FHMM	.774	.622

As shown in the table, the nearest-neighbor based methods slightly achieve improvement in the sequential analysis over the per-instance classification, while the RNN achieves about 5% more of the kappa measure than the neural network. This finding suggests that one may achieve better results by looping the information from the previous input values to predict outcomes of the following instances on a se-

[3]https://www.rug.nl/society-business/centre-for-information-technology/research/services/hpc

Table 5.3: Classifier Kappa measures on the last seven days of the un-shuffled test set

Date	Max persons	Occupancy with ID		Occupant counting	
		k-NN^{seq}	RNN LSTM	*k*-NN^{seq}	RNN LSTM
2018-9-3	3	**.865**	.841	**.911**	.858
2018-9-4	3	**.877**	.801	**.905**	.816
2018-9-5	1	.850	**.876**	.855	**.876**
2018-9-6	2	.559	**.628**	.737	**.796**
2018-9-7	3	.787	**.897**	.783	**.911**
2018-9-10	3	.717	**.836**	.859	**.888**
2018-9-11	3	.854	**.952**	.855	**.960**
7-day average		.787	**.833**	.844	**.872**

quence. As for *k*-NN, the algorithm works based on the voting labels of the nearest samples to the query. Hence, the performance solely relies on sample availability and does not affect much to classification performance. FHMM is the worst performing algorithm. It might be because this algorithm does not work by finding separation boundaries among different occupancy states; rather, it finds the most probable occupancy states given the observed power consumption. Moreover, in this work, we only define two individual states (i.e., being $present$ or $absent$) without focusing on which appliances occupant use. Therefore, the power consumption range during $present$ condition is wide. The model simplification (i.e., by providing only the active power variable) might also negatively influence the results of FHMM inferences.

The results on the test set based on historical occurrence (i.e., without shuffling) in Table 5.3 show daily performances that seem to be lower than the shuffled data division (i.e., reaching 93-94% kappa measure in Table 5.2). It might be due to the variance of power consumption that we may find in practice. As we only provide the first 85% portion of the data, the last portion of the data were not sampled. In this particular case, RNN LSTM gains benefits in several days (i.e., the five of seven work days) and outperforms *k*-NN^{seq}. A possible explanation for this might be that RNN regards the output of the previous instances to predict the output of the current instance which did not occur in the training data. On the contrary, *k*-NN^{seq} does easily misclassify a class if there are very similar samples that belong to different classes. *k*-NN^{seq} outperforms when it finds many similar samples with the major label as in the first two days (i.e., September 3, 2018 and September 4, 2018).

A comparison of the two algorithms on September 10, 2018 is illustrated in Figure 5.7. The figure presents the identification of present occupants, as expressed in the seven class labels. Each class represents the combination of three occupants'

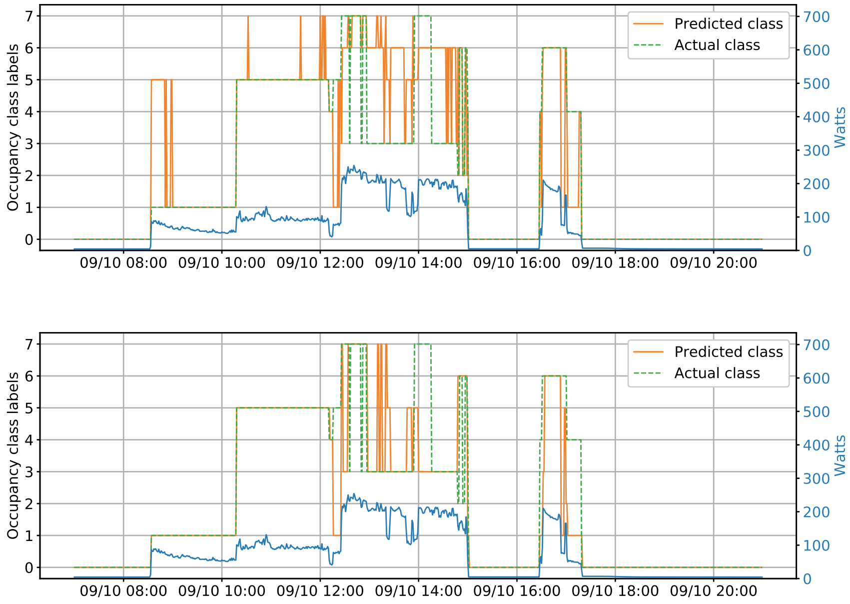

Figure 5.7: Occupancy detection with ID on September 10, 2018 using k-NN^{seq} with a Kappa of 71.7% (top) and RNN LSTM with a Kappa of 83.6% (bottom). Blue line shows the active power in $Watts$, orange line shows the labels of predicted occupancy, and dashed green line shows the labels of real occupancy

presence states. The top figure shows the prediction using k-NN^{seq}, while the bottom is the prediction using sequence RNN LSTM. It can be seen from the figures that there were some misclassified class-5 instances using k-NN^{seq} at 08.30-09.00 AM due to the power consumption going up to 80 Watts. RNN LSTM could handle it better without misclassification until 12.15 AM. While around 02.00 PM, both classifiers failed in detecting class-7. The classifier based on RNN LSTM performed better in recognizing class-3 at 02.15 PM. In the same period, the k-NN^{seq} classifier mostly misclassified as a class-6 until at 03.00 PM.

5.5 Conclusion

This study set out to investigate power consumption in sub-metering of a shared office room using sliding windows. Sequential and non-sequential approaches are explored to identify electric loads. We begin with the identification simple loads consisting of five office appliances belonging to an occupant. We then continue with the identification of composite loads that corresponds to occupants' power consumption during their presence in the office.

Compared to the study in Chapter 4, the power meters used in this study measure more variables, including active power, reactive power, apparent power, power factor, and electric current. By considering the measurement of these variable (i.e., in feature set FS-3), we show that the appliances and their combinations can be identified up to .99 Kappa measure using *k*-NN and RNN LSTM. Next, we show up to three-person occupancy detection by recognizing the aggregate loads consumed by each individual. We find that sequence analysis gives improvements from the non-sequence analysis in neural network approaches, but not in nearest neighbors. Nonetheless, no significant benefits in using RNN LSTM as the nearest neighbors approach outperforms the other method in our experiments, reaching .946 Kappa measure. We show that the RNN LSTM might be useful when there is no sample provided to the classifier, such as in predicting typical days when power consumption patterns have not appeared at all before.

Chapter 6

Beaconing-based Occupancy Detection

6.1 Overview

The proliferation of mobile phones and wearables (e.g., smartwatches) contributes to context gathering without requiring dedicated tags or signal receivers attached to users. Those devices are equipped with communication modules that support receiving Radio Frequency (RF) signals from transmitting nodes, such as WiFi access points (APs) and BLE beacons. WiFi is more common in indoor spaces like in an office. However, as the WiFi deployment is not tailored to localization (i.e., to provide wireless Internet connection instead), the location of WiFi APs are frequently hidden and unknown, thus, unable to be used as a reference position. Furthermore, they are usually arranged to get the maximum coverage (e.g., covering the whole area with fewer APs) regardless of the installation geometry. While BLE beacons are less deployed, they are gaining popularity because people are making more use of wireless peripherals (e.g., headphones). Off-the-shelf beacons are commonly found for proximity-based services, such as in the American Museum of Natural History[1] and Amsterdam Schiphol Airport[2]. The beacons are affordable; namely, they cost as low as three euros per piece. Another advantage is its flexibility in the deployment, as it is compact with small size and battery-powered (i.e., with up to 3 years life-expectancy, depending on the configuration).

A serious weakness with beaconing-based sensory sources is the signal strength that may be erratic due to environment dynamics or hardware diversity. Let us take the scenario of a shared office as an example. When Aldo becomes the first person who comes to the office, his mobile phone receives strong signals from BLE beacons located in the office and medium signal strength from BLE in the neighboring office. As this observation matches the signals collected during the training phase, the system may correctly locate him in the office and assigns the room as occupied by Aldo. A moment later, Cecilia comes to the office and frequently moves around her

[1]https://www.amnh.org/explore/news-blogs/news-posts/bluetooth-beacons-help-navigate-museum-halls

[2]https://news.schiphol.com/amsterdam-airport-schiphol-first-airport-in-europe-with-full-beacon-coverage/

space. Aldo's phone then often receives weaker signals from beacons in the office, just similar to the signals from the neighboring office. This signal observation leads to the wrong inference of Aldo's position. It triggers all the light in the office with 60% brightness (e.g., as Cecilia's preference) since the building assumes that Aldo is not in the office, and Cecilia is alone. The automatic lighting control might even turn off the light in the office when Cecilia's phone discovers BLE signals with different strength that leads to unoccupied state inference.

In this chapter, we study beacon signals collected in some observation points using a mobile phone. We propose to use the cosine similarity approach to find vector representation in different locations in offices and to classify the location of other employees. The scenario is considered as low intrusive as it does not require employee participation to collect new training data using different phones. The empirical findings in this study provide a new understanding of how the low-intrusive approach may perform in the detection of multi occupants, especially in adjacent office rooms. This chapter begins by presenting the relevant literature and general design in Sections 6.2 and 6.3, respectively. The experiment setup and results are given in Section 6.4. Finally, we conclude the experiments in Section 6.5.

6.2 Relevant Literature

Previous research has established that localization systems can derive room occupancy. The systems localize a user by estimating a coordinate and assigning positive occupancy when the inferred position is within the area boundary of a target room, for example, a system proposed by Paek et al. [88]. The authors investigate a trilateration technique using BLE beacons for the class attendance detection system in a university. Trilateration is a process to find a location that is described in terms of fixed distances to the known points using geometry approaches. As they discover that beacon RSS readings have high variation and tend to be weaker than observation in a line-of-sight (LOS) environment, they propose geometric manipulation, mainly when there are no intersection points between beacons. They enlarge the estimated radius from the receiver to a beacon with an increment of one meter until intersection points formed. Given the intersections, the trilateration method may estimate the location. The proposed approach is evaluated in checking the attendance of four students in three classrooms. They report that no false detection occurred during the experiment. The experiment setup, however, is not particularly clear, such as whether the classrooms are adjacent, what size the rooms are, and in which part of the classrooms the students occupied. Furthermore, it is almost certain that the beacon coverage signal is not perfectly circular. There is no guarantee that the real-location can be inferred precisely by extending the radius by a uniform value (i.e., $1m$). Also, a one-meter increment seems not fit with small-medium offices.

Other localization techniques based on RF, including BLE and WiFi, are commonly based on a fingerprinting map, a pre-surveyed map of radio signal strength across the environment. A seminal study in location fingerprinting with BLE beacons is the work of Faragher and Harle [44]. The authors estimate user locations based on Maximum a posteriori (MAP) or Minimum Mean Square Error (MMSE) on the posterior distribution. The distribution is calculated from a distance between current signal measurement and the fingerprinting map. Therefore, the fingerprinting map holds a vital role. Given the BLE fingerprinting map, the positioning is even more promising than WiFi, namely, reaching the average error of up to $2.6m$ and $8.5m$ of 95% of the time, respectively. Thanks to the flexible placement of BLE, that brings advantage to better signal geometry [43]. However, the map is challenging to make. It requires site survey and manually taking readings at each point (e.g., [15]) or requires additional high-precision localization equipment to assign a true-position to the recorded BLE signals and create a signal strength map accordingly (e.g., using Active Bat [44]).

Another method exists to model the indoor radio signal propagation without building a fingerprint map, by calculating the distance between beacons and the receiver and estimating the current position [143]. The propagation formula is based on the log-distance path loss model that reflects the trends of radio propagation of wireless devices [103]. It is reported that the average error of location inference is up to $2.18m$ in semi-outdoor areas (e.g., in a corridor out of a building). The evaluation is based on a walking track with around $450m$ distance. This result is relevant since there are fewer obstacles in the investigated area, such as walls and office objects, than in the indoor office area.

Low-intrusive Approaches

Other researchers have looked at low intrusive approaches. An intuitive way to infer the current location is to take the location of the nearest beacon discovered by the signal receiver. Lin et al. demonstrate a localization system based on the strongest beacon and achieve more than 95% average accuracy in locating a person in twelve subareas [78]. The approach is applied to the smoothed signals, where the smoothing process is based on the average value of the last five samples. They select the major strongest beacon during the five consecutive timestamps.

As a proximity-based localization, iBeacon protocol describes nearby beacons into some categories (i.e., immediate, near, and far)[3]. However, the formulation

[3]https://developer.apple.com/documentation/corelocation/determining_the_proximity_to_an_ibea-

to determine proximity category and the distance approximation is not accessible. Likewise, Kyritsis et al. categorize beacon for each discovered beacon based on RSSI thresholds [70]. Two thresholds (i.e., thr_{in}, thr_{out}) are set for each beacon to determine beacon categories (i.e., strong ($RSSI > thr_{in}$); medium ($thr_{in} > RSSI > thr_{out}$); weak ($thr_{out} > RSSI > -127dBm$)). In assigning the thresholds, the authors consider the signal propagation model and room dimensions, assuming that adjacent rooms have different sizes. A calibration process is needed to relate the RSS to the known distance for a specific signal receiver (e.g., a phone). Once beacons have been assigned to a category, they predict the location of the receiver based on the beacon location with the highest category. In the case of multiple beacons have the same category (e.g., due to unpredicted environment noise), the system assigns the location based on the highest probability. That is, the probability is the normalized value of scores among the same category beacons. The score is defined as the difference of RSS measurement to the lower threshold of the category.

Various studies have assessed the efficacy of pattern recognition and machine learning on the RSS in occupancy detection. Conte et al. propose single-occupant multi-class occupancy detection using BLE [30]. The authors deploy one beacon per room in three rooms. The inference is made per instance using k-NN and DT. The experiment result is obtained based on ten-fold cross-validation with 1234 instances. They report 83% accuracy. Similarly, the same procedure is applied by Corna et al. to detect occupancy of one room [33].

Filippoupolitis propose a single user localization in ten areas using eight BLE beacons [45]. The areas are divided into two independent sectors. The authors investigate Logistic Regression, k-NN, and SVM technique on BLE signals measured using a mobile phone. They report the classification performance based on 10-cross validation. The accuracies are between 80-100% depending on the rooms.

Localization using WiFi and BLE beacon by placing one beacon per room also has been done in [123]. From beacon readings, features are extracted (i.e., max, min, mean, std. dev) followed by machine learning utilization, such as a Bayesian-based classifier. The test bed is a house with five rooms. However, as the work focus on the load disaggregation based on occupancy, the authors do not provide localization accuracy [124].

Challenges and proposed solutions

Several lines of evidence suggest that RSS of BLE beacons suffers from several disturbances. Multipath and fast fading signals during propagation appear to be significant problems in an electromagnetic-based localization, including BLE [25]. Device

con_device

heterogeneity in the market also provides various readings for the same broadcasted signal strength [102]. Furthermore, for battery-powered BLE, the beacons' battery level affects the signal transmits, up to $5dBm$ [123]. The human body may also block and weaken the RSS [143]. Interference on WiFi signals affects on dropped reception ratio to around 75% and lower RSS values (more than $10dBm$ reduction) [88]. Further, Paek et al. also observe that a phone held on hand can attenuate the received signal as much as $30dBm$. Especially in real life, these unforeseen circumstances may affect the performance of localization. Therefore, testing in real-life holds an important role.

A number of studies have addressed the problems. Castillo Kara et al. propose beacons' asymmetric transmission power [25]. The authors firstly investigate beacon signal attenuation per unit beacon. Based on this insight, they evaluate some different combinations of beacons' transmission power to improve classification models in the localization system. While they claim that some improvement exist over the homogeneous transmission power, the approach is not portable and requires much effort to investigate the transmission power setting of each BLE beacon. Furthermore, there is no way to guarantee that the best performance is the best it can achieve, unless one tries a very large number of experiments.

To deal with device heterogeneity, some researchers propose offset calibration to compensate the RSS offsets of different devices in relation to a reference device [125, 120]. While it is reported that this approach may increase the localization performance (i.e., reducing up to 80% localization error in 50% percentile), it requires calibration of each type of mobile phones used by users.

Apart from the data or program manipulation approaches, hardware-based approaches have also been proposed. Barsocchi et al. propose to use BLE beacons with multiple signal strength (i.e., $-18dBm$ and $3dBm$) to improve occupancy detection [17]. The transmitting beacons are provided and carried by a user, thus, the hardware is uniform. The receiver nodes, on the other hand, are deployed in the rooms to receive the transmitted signals. A participant walks through predetermined routes and stays about a minute in marked positions to evaluate the occupancy inference performance. The authors report up to 10% accuracy improvement compared to single signal strength transmission. However, the approach is intrusive, as it forces people to carry the transmitting beacons. In our work, we utilize existing mobile phones as receivers instead of attaching additional hardware to users. Hence, we achieve additional benefits, such as reducing intrusiveness to users in carrying special beacons and improving chances to track users as they use the phones during activities. Furthermore, this approach diminishes the building manager's responsibility to manage the battery of emitting beacon tags carried by users.

Vigneshwaran et al. equip two access doors with two BLE beacons [127]. Each beacon is set with relatively strong power ($-8dBm$) and short-interval transmission (i.e., $109ms$). The authors configure a mobile application as a receiver with a high duty cycle, that is, less than $1s$ of waiting interval. This configuration allows them to detect enter/exit movements through the observation of signal strength variation. Such a high-demand arrangement consequently impacts to the beacon lifetime (i.e., about only three months) and phones' battery life. Our work, on the other hand, aims at power saving. We set a more efficient beacon to transmit energy into $950ms$ and, most importantly, the duty cycle of $2.5s$ ON and $1.5s$ OFF to preserve phones' battery. The beacon lifetime is also longer using such a less powerful configuration (e.g., in our experience, the beacons still have more than two years of battery life after about three years of deployment).

6.3 Design

We aim to investigate the performance of multi-occupant low-intrusive room-level localization using BLE beacons. Two assumptions have been made. First, the occupants carry a mobile phone to support work activities. Second, The participants agree to keep the beacon application and Bluetooth scanning services running. They understand that the system does not recognize activities and productivity, therefore, they give consent to collect occupancy data in common spaces (i.e., offices and a social corner). At any time, they can stop the services when needed to protect their privacy.

We collect BLE beacons signals using various mobile phones. Instead of building fingerprint maps, we limit to a few hours of training data collection only in a few observation points. This simple step is to to keep the intrusiveness-level low. We ensure that the collected training data well represent each room by using a cosine distance between the reference vectors acquired in the observation points [97]. Based on the collected signals, we extract features for training purposes. We thus infer multi-occupants location in a typical week and compare to the related works to evaluate the performance of our approach.

6.3.1 Sensing Technology

We use off-the-shelf Proximity Tag from Estimote[4], which has a default battery life of two to three years. The vendor provides two Software Development Kits

[4] https://estimote.com/products/

(SDKs), namely Proximity SDK[5] and Android-Fleet-Management-SDK[6] (the latter is also known as Estimote SDK). Proximity SDK is an advanced SDK with Estimote's signal-processing technology, running on top of Estimote's proximity monitoring framework. This SDK, however, only provides information of enter or exit events to a defined area. On the other hand, Estimote SDK is a basic SDK that provides APIs for detecting beacons in two ways: *monitoring* and *ranging*. Monitoring enables to know events of entering or exiting of an area, while ranging provides more granular information. It returns a list of beacons in range, together with signal strength and estimated proximity to each of beacon. We opt to implement *ranging* beacons from Estimote SDK to collect raw RSS data and beacon categorization[7] (i.e., immediate, near, far, unknown), and process in our way, as will be discussed in the next section. Also, the SDK is the most widely used development kit and it is supported by most Estimote beacon types. Each beacon node broadcasts tiny BLE packets based on a beacon protocol. We use the most standard iBeacon protocol[8] that is fully supported by the beacon nodes we used.

6.3.2 Techniques

Prior to the localization investigation, signal references are collected to supervise classifiers. For each location $l_i \in L, L = \{l_1, l_2, \ldots, l_r\}$, we sample signals from a set of beacons $B = \{b_1, b_2, \ldots, b_m\}$ in observation points $\mathcal{P} = \{p_1, p_2, \ldots, p_o\}$. The observation point p is a position where people are mostly staying in the office rooms, such as in a chair or sofa. From each observation point $p_i \in \mathcal{P}$, we compute the median of beacon signal readings as a reference vector v that represents the location l_i.

We investigate the similarity of the reference vectors to evaluate whether such vectors are valid for representing a class of location [97]. For this, we calculate the degree of similarity between a pair of reference vectors, v_i and v_j (i.e., Eq. 6.1). We then keep the signals collected at the observation points if they have different directions with other vectors representing different classes.

$$sim(v_i, v_j) = \frac{\vec{v_i} \cdot \vec{v_j}}{|\vec{v_i}||\vec{v_j}|} \tag{6.1}$$

Having the collected training signals to represent classes properly, we continue the processing. We transform the RSS signals into magnitude power P (Eq. 6.2), and normalize the value into a range of 0 to 1 (Eq. 6.3).

[5]https://github.com/Estimote/Android-Proximity-SDK
[6]https://github.com/Estimote/Android-Fleet-Management-SDK
[7]https://developer.estimote.com/ibeacon/tutorial/part-3-ranging-beacons/
[8]https://developer.apple.com/ibeacon/

$$P = 1mW \cdot 10^{\frac{\text{RSS}\,dBm}{10}} \tag{6.2}$$

$$P' = \frac{P - min(P)}{max(P) - min(P)} \tag{6.3}$$

We extract a number of features, including mean, mode, standard deviation, and maximum value over a non-overlapping sliding window of length w. We also consider the difference between the means of the current window and the previous window as features. Furthermore, we consider Boolean features that indicate whether beacon nodes are discovered and being the strongest, among others. The feature space is summarized in Table 6.1.

Table 6.1: Feature space of a single beacon made up of N observations over a sliding window

Features	Formula
mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} P_i$
mode	$\hat{P} = argmax(P_i)_{i=1}^{N}$
std. deviation	$P_{std} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - \mu)^2}$
max	$P_{max} = max(P_i)_{i=1}^{N}$
diff	$P_{diff} = \mu_t - \mu_{(t-1)}$
isDiscovered	$1_{(\exists P_i \neq \infty, i \in N)}$, 0 otherwise
isStrongest	$1_{(max(P_{max}) \in \mathbb{R}^M)}$, 0 otherwise

Finally, we build classification models based on *k*-NN with cosine distance. A localization classifier h_{loc} assigns l^{j_i}, a location of individual j_i, to β^{j_i}, the extracted features of the RSS beacons discovered from j_i's phone. Formally, $h_{loc} : \beta^{j_i} \rightarrow l^{j_i}, l^{j_i} \in L$. The extracted feature β^{j_i} has $(n \cdot M)$ columns, where n is the number of extracted features per beacon, and M is the number of deployed beacons. We use the built models to classify any query RSS.

Baseline: Low-intrusive Approaches

We replicate the nearest beacon approach proposed by Lin et al. [78] and provide the inference results of experiments in our dataset. Furthermore, we adopt the thresholding approach proposed by Kyritsis et al. [70] with some adjustments in the calibration process. As our environment differs from the original author's setup (i.e., we have adjacent rooms with the same size), the room dimension does not significantly contribute to the threshold assignments, and thus, room inference. Therefore, we calibrate the thresholds based on the visual observation of signals strength magnitude received by different phones. The assigned thresholds are then used to infer room-level localization in our dataset.

6.3.3 Metrics

We measure overall accuracy and per class F-measure to evaluate the system in determining room-location of each participant.

Accuracy

Accuracy defines the correct prediction of all class labels over the total number of predictions that have been made. In the case of beaconing-based occupancy detection, we assign localization classifier $h_{loc} : \beta^{j_i} \rightarrow l^{j_i}$. The metric should evaluate how good the classifier is in assigning a room location $l^{j_i} \in L$ to the beacon signals of each participant j_i. The general formulation of accuracy in Equation 5.2 can be specified as:

$$Accuracy_{j_i} = \sum_{l \in L} \frac{TP_l}{TP_l + TN_l + FP_l + FN_l} \tag{6.4}$$

The True Positive or True Negative of location l (TP_l or TN_l) is the number of windows that correctly inferred as location l, or others (i.e., not in location l), respectively. False Positive or False Negative (FP_l or FN_l) is the number of instances for which location l, or not in l, are misclassified.

Since employees spend most of the work hours in their room office, we provide the correctness of whether they are in the office. Therefore, we also expose the percentage of correct inference of being present (i.e., true presence) and error inference of being not present (i.e., false absence) in the corresponding work office.

F-measure per-class

As accuracy is prone to be misleading (e.g., a high accuracy due to inferring a major class in imbalanced data), we calculate precision and recall for each room location $l \in L$. Precision in detecting location l is defined as the rate of inferred class-l that are predicted precisely. While recall in detecting location l is defined as the rate of the real "in room location l" instances that are identified correctly. Finally, we provide the harmonic mean of precision and recall with an equal weight. Precision, recall, and F-measure are defined in Equation 6.5, 6.6, and 6.7, respectively.

$$Precision_l = \frac{TP_l}{TP_l + FP_l} \tag{6.5}$$

$$Recall_l = \frac{TP_l}{TP_l + FN_l} \tag{6.6}$$

$$F - measure_l = 2 \cdot \frac{recall_l \cdot precision_l}{recall_l + precision_l} \tag{6.7}$$

6.4 Experiments

We experiment multi-occupant localization based on BLE beaconing system. This section describes datasets and the experiment setup. Finally, it discusses the localization results.

6.4.1 Data

We included five participants in carrying mobile phones loaded with BLE signal acquisition application during typical workdays on the fifth floor of the Bernoulliborg building at Zernike Campus of the University of Groningen, The Netherlands. The data collection was performed from October 15, 2018 until October 19, 2018. For training purposes, we collected additional data on the weekend on October 20, 2018. The training data covered eight observation points. At each point, the training data were collected for about 20 minutes. We opted to use occupant work desks in three room-offices and a sofa in Social Corner (SC) as observation points as these places are the most occupied position during the occupancy period.

6.4.2 Setup

We initially deployed twelve beacons across the office environment on the fifth floor of Bernoulliborg building in the Zernike Campus, the University of Groningen, The Netherlands. One of the beacons, unfortunately, was missing during the data collection period. The beacons have adjustable configuration parameters. As we focus on a low-intrusive approach, we set low power signal transmission to preserve battery life and, thus, lowering the need for beacon battery maintenance. Table 6.2 provides the overview of our low-power beaconing system, compared to other deployments.

To receive signals broadcasted by the beacons, we install application listener on the mobile phones carried by the participants. The app has capabilities of accessing the Bluetooth and Internet connection for sending the measurement to our server. The application is set to listen and wait for signals in $2.5s$ and $1.5s$, respectively. Furthermore, the application has a user interface to receive the actual room locations as ground truths. The participants have fixed workplaces. Figure 6.1 illustrates the experiment testbed in our university building.

Table 6.2: Beacon Deployment

Refs.	Transmitted signal strength	Beacon density	Broadcast interval
This thesis	$-16dBm$	2 per $30m^2$	$950ms$
[17]	-18 and $3dBm$	2 beacon per person, 1 static receiver per room ($26m^2$)	$1000ms$
[127]	$-8dBm$	1 per door	$109ms$
[44]	$0dBm$	1 per $28m^2$	$20ms$
[88]	$4dBm$	3 per classroom	$950ms$
[143]	$-12dBm$	1 per $30m^2$, semi-outdoor	$645ms$
[70]	$-12dBm$	1 per room	$350ms$
[45]	unknown	1 per $48m^2$	$125ms$
[83]	$4dBm$	unknown	unknown

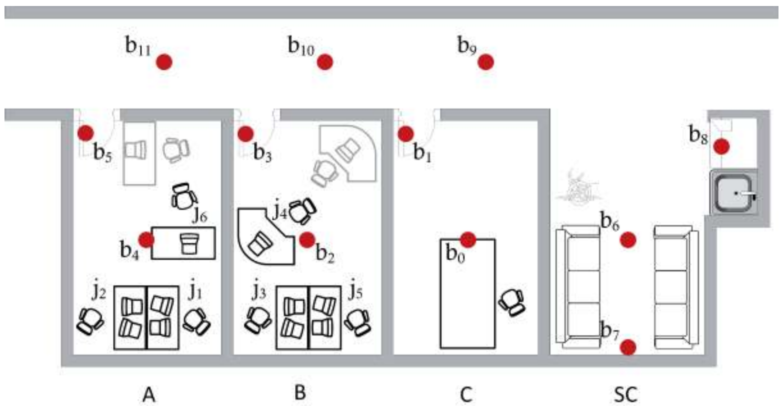

Figure 6.1: Layout of the work spaces and deployed beacons

Prior to locating people based on RSS, the received signal strengths are investigated. We collect signals from two observation points in room B at the weekend when nobody is present in the office. The first observation is from j_5's desk on signals from beacons b_3 (vertically on the access door of room B), b_2 (on the ceiling of room B), and b_0 (on the ceiling of room C). The second observation is from j_3's desk on signals transmitted by beacons b_2 (placed on the ceiling of room B) and b_4 (placed on the ceiling of room A).

In our approach, we use window width $w = 10$, which is equal to 50-second measurement. The window is non-overlapping. Such a window configuration is

Table 6.3: Calibrated thresholds for each user (dBm)

Users	b_0^{in}	b_0^{out}	b_2^{in}	b_2^{out}	b_4^{in}	b_4^{out}	b_6^{in}	b_6^{out}
j_1	-87	-93	-89	-93	-86	-98	-80	-90
j_2	-81	-88	-81	-91	-81	-90	-75	-92
j_3	-90	-95	-83	-92	-85	-90	-81	-90
j_4	-84	-89	-87	-96	-86	-92	-81	-93
j_5	-83	-93	-91	-98	-89	-97	-85	-92

based on the assumption that people will stay at least a minute in a room, for example, taking a cup of coffee at the coffee machine in Social Corner (SC). As we focus on a low-intrusive approach, the training instances are only from a single mobile phone (i.e., belong to individual j_2) to prevent the necessity of all participants to take training data.

Implementation: Low-intrusive Approaches

When comparing to the related works, care is taken to replicate the experimental setup. We consider only one beacon on ceilings per room as proposed by Lin et al. and Kyritsis et al. [78, 70]. While we have different signal sampling rates with the authors, we assign the same window size. The reason is that we consider the same number of samples in smoothing the BLE signals. For the Lin's approach, we use $w = 5$, which means that we average signal readings in about $25s$ (5 times wider than the original work). For Kyritsis's approach, we use $w = 10$ of the latest RSSI readings. We assign the thresholds thr_{in} and thr_{out} that well differentiate beacon strength categories (i.e., strong, medium, and weak). The thresholds are listed in Table 6.3.

6.4.3 Results and Discussion

This section discusses the preliminary investigation on the received signal strength in the test environment. It then elaborates the localization results based on the approaches mentioned earlier.

Preliminary RSS Observation

We investigate the signal strength collected from two observation points (i.e., at the work desks belonging to individual j_3 and j_5) located in Room B. The findings are discussed in the following.

Distance is not the only factor affecting the received signal strength. Ideally, beacon b_2 should be the strongest due to its closest distance and the same room location as the phone, and beacon b_0 should be the weakest due to its location in the neighboring room. However, there is a variability of signal strength from beacons in different places

Figure 6.2 shows BLE signals observed at the j_5's desk. The phones belonging to j_1 (the blue line), j_2 (the orange line), and j_4 (the red lines) receive the strongest signals from beacon b_3, between -75 and $-70dBm$. From beacons b_2 and b_0, the j_4's (the red line) and j_1's (the blue line) phones receive signals about $-85dBm$ and $-90dBm$, respectively.

The signals of beacon b_0, as expected, are weaker than the other two beacons, as this beacon is located in the next office with a wall separation between the rooms. Interestingly, signals from beacon b_2 are weaker than signals from b_3 for all phones, even though the distance from the phones to b_2 is closer (i.e., $2.3m$) than to b_3 ($4.9m$). It is probable that the beacons' positioning and antenna directions affects the signal strength. Both beacons b_3 and b_2 are located in Room B. That is, beacon b_3 is attached vertically on the door with the distance about $1.9m$ from the ground, while beacon b_2 is attached on the ceiling with the distance about $2.7m$ from the ground. It is apparent that placing beacons vertically on the doors or walls is more effective than on the ceiling.

Signals from beacons in neighboring offices may be discovered with a similar signal strength by some mobile phones. Figure 6.3 presents an overview of signal strength of beacons b_2 and b_4 received at the j_3's desk. One can see that the mobile phone belonging to j_4 (the red line) observes signals from beacon b_2 with signal strength between -75 and $-80dBm$, and, as expected, the same phone observes weaker signals from b_4 in the neighboring office with signal strength between -85 and $-90dBm$. Interestingly, the other mobile phones report similar signal strength for the same beacons in different locations. That is, the phone belonging to j_1 (the blue line) observes fluctuation signals between -85 and $-90dBm$ for signals from both beacons b_2 and b_4. Also, the j_2's phone (the orange line) observes about $-75dBm$, and the j_5's phone (the purple line) observes about $-85dBm$ for both beacons, even though the signal fluctuation is slightly different. A likely explanation of the similarity is the difference on sensor sensitivity .The phone belonging to j_4 is sensitive enough to distinguish beacons b_2 and b_4, while the other phones are not. The mobile phones have different receiver sensitivity and radio-frequency interference (RFI) due to noise within the devices [115]. It might be affected by the noise generated by memory interfaces, clock signals for the SD card, sensors, etc. It is likely that the mobile phone belonging to j_4 has higher sensitivity than the others.

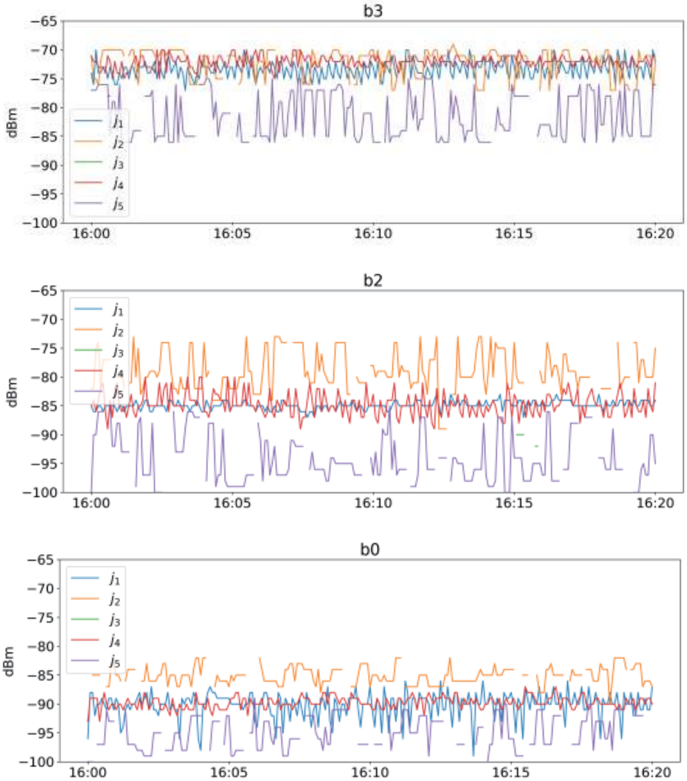

Figure 6.2: Received signal strength transmitted from three different beacons and observed from j_5's desk in Room B. Top: the observation of beacon b_3 (at the door of Room B). Middle: the observation of beacon b_2 (on the ceiling of Room B). Bottom: the observation of beacon b_0 (on the ceiling of the neighboring Room C)

The findings suggest that there are some irregular unexpected received signal strengths from the empirical experiments. To infer the location of different phones based on RSS, pre-process calibration, or complete fingerprinting surveys may be necessary. In this thesis, however, we avoid such a way to reduce intrusiveness. We use machine learning approaches to explore the pattern of the RSS instead. We also compare with other considerably low-intrusive approaches, based on the strongest beacon and thresholding approach.

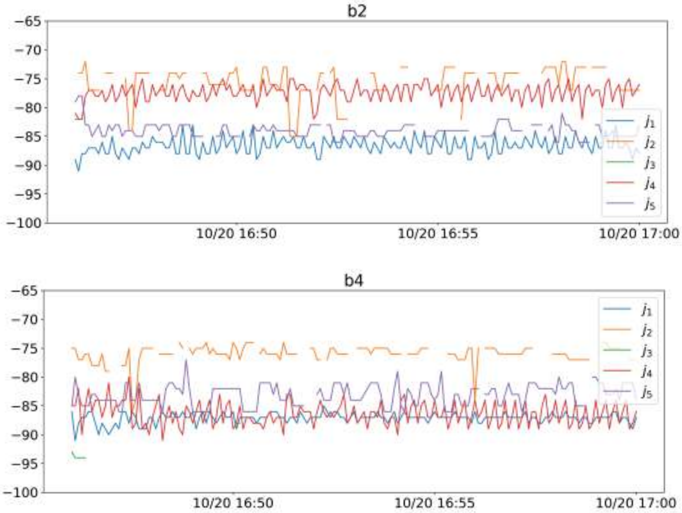

Figure 6.3: Received signal strength transmitted from two beacons on the ceilings of Room B (top) and neighboring Room A (bottom). The observation was from j_3's desk in Room B

The Classification Results and Discussions

The inference results of the proposed low-intrusive approaches are presented in Tables 6.4 and 6.5, respectively. As can be seen in the tables, utilizing the strongest beacon (i.e., Lin's approach) and the threshold-based beacons' signal-level (i.e., Kyritsis's approach) one can infer per room location with an F-measure between .82 and .91 in four of five employees in a week of observation. The Kyritsis approach slightly improves the baseline for most people except for j_1. This declined performance might be affected by the threshold assignments that do not entirely fit with the test data. The best inference of Kyritsis's approach is on individual j_5, with .906 F-measure of presence in the j_5's office (i.e., Room B) which is inferred using the baseline reaching .881 F-measure.

Table 6.6 presents room-level classification results over five participants using our proposed approach. What stands out in the table is that our approach improves the inference of majority rooms of all participants. While the approach does not improve in the inference of room C and SC for individual j_1, it significantly improves the inference in j_1's office, reaching .973 F-measure. The improvement of the self-office inference also happens to individual j_3, but the inference is still not satisfying. When we look at the self-office inference, the False Negative (FN) is higher than TP. That is, about 44.36% of the time of his presence in his office (i.e., Room B) is correct,

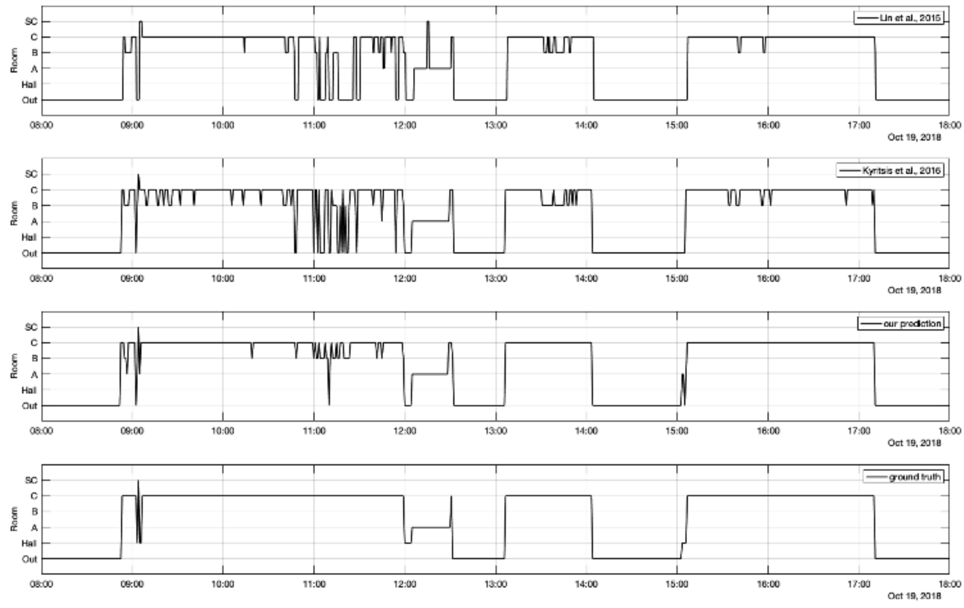

Figure 6.4: The localization Inference of j_1 based on RSS of BLE beacons

and 55.64% are misclassified as not in the office. There are many misclassifications to the neighboring Room A. Sitting near the room border may cause false inferences.

Figure 6.4 illustrates room-level localization using several approaches. From the figure, one can see that the localization systems often misclassified adjacent rooms. For example, starting from 09:07 until 12:00, the person is actually in Room C. However, the system sometimes infers that the person is in Room B, in turn leading to an increasing number of FN. While our proposed approach may reduce this number (i.e., it shows that our approach is more robust due to more features and matched training samples), if we look at the FN, it gives 123 inference which means 102.5 minutes or 1.7 hours of false detection on average in five workdays. This misclassification may still lead to power waste, or even worse, inconvenience or frustration for the users when it is connected to automatic control systems.

Table 6.4: Room location classification using the Lin's approach [78]

ID	Acc.	F-measure per class					Self-office (%)	
		Out.	A	B	C	SC	True Pres.	False Abs.
j_1	.95	.9870	**.91306**		.8871	.8713	84.34	15.66
j_2	.95	.9814	**.8219**				70.64	29.36
j_3	.79	.9589		**.4588**			29.89	70.11
j_4	.99	.9932		**.8879**	.9339		81.19	18.81
j_5	.97	.9922		**.8812**			79.35	20.65

Table 6.5: Room location classification using the Kyritsis's approach [70]

ID	Acc.	F-measure per class					Self-office (%)	
		Out.	A	B	C	SC	True Pres.	False Abs.
j_1	.94	.9936	**.8921**		.8333	.7083	80.85	19.15
j_2	.96	.9814	**.8608**				76.68	23.32
j_3	.80	.9626		**.4895**			32.53	67.47
j_4	.99	.9938		**.8919**	.7629		83.90	16.10
j_5	.97	.9929		**.9064**			83.57	16.43

6.5 Conclusion

The RSS based localization or occupancy detection systems suffers from disturbances, as noted already in the literature. The problems are also in line with our observations, which show that: 1) the distance between beacons and mobile phone receivers is not the only factor that affects the RSS. The beacon orientation could also effects the received signal strength; 2) receiving signals in the edge of the adjacent rooms may lead to the similar signal strength of beacons located in different rooms. Some mobile phones are not sensitive enough to sense the difference. Based on this observation, deciding a room location of adjacent rooms might be difficult due to many factors affecting the irregular RSS.

We propose to use the BLE beaconing system in a low-intrusive way. Instead of conducting dense fingerprinting surveys, we propose to only sample signals using a phone in some parts of the most occupied points and use the signals as location references. We validate the collected reference vectors based on the cosine similarity. We thus extract features from the collected training data and build nearest neighbors classification models. To test the performance, we investigate five mobile phones and compare with previous works that we consider low-intrusive (i.e., [70, 78]).

We find that the proposed low-intrusive approach performs reasonably well for four of five participants in a week investigation. It also improves the occupancy detection in self-office for whole employees over the baselines, even though the

Table 6.6: Room location classification using the proposed approach

ID	Acc.	F-measure per class					Self-office (%)	
		Out.	A	B	C	SC	True Pres.	False Abs.
j_1	.98	.9980	**.9725**		.8169	.7727	95.09	4.91
j_2	.97	.9856	**.8847**				80.68	19.32
j_3	.83	.9628		**.6126**			44.36	55.64
j_4	.99	.9938		**.9017**	.9344		83.22	16.78
j_5	.99	.9931		**.9559**	NaN		92.31	7.69

strongest beacon indication still infers better in some places with a short occupancy period. The individual whose presence is not inferred well is the one who sits near the separator wall among two adjacent rooms. This results even in 55% of misclassification. The other finding is that even in the best inference results, the false detection happens for about 100 minutes in five-day observation. These findings show that the proposed approach still needs improvement. As we will explore in the next chapter, the fusion of different modalities may provide improvements.

Chapter 7

Fusion of Power-metering and Beaconing Systems

7.1 Overview

The BLE beaconing systems often experience false negatives when neighboring offices are misclassified or BLE beacons are not detected. In this chapter, we survey sensor fusion of the two sensory sources introduced in the previous chapters, namely, the power metering system and beaconing system. Sensor fusion is defined as a technique to combine sensor readings from multiple sources to improve accuracy and reliability, or to achieve more specific inferences than could be achieved by the use of a single sensor alone [54, 140]. The fusion process, unfortunately, does not always lead to improvements. Additional data or information may even confuse the inferences based the individual sensors, especially when the additional information is incorrect or inconsistent [53]. Hence, care must be taken to properly investigate the fusion process by, for example, selecting fusion techniques that are appropriate to the available information.

Hall and Llinas further define three types of architectures based on the data manipulation level, namely, data-level fusion, feature-level fusion, and decision-level fusion [54]. In data-level fusion, the raw sensor data from commensurate sensors (i.e., sensors observing the same physical quantities, such as visual images) are combined. It is then followed by extracting feature vectors from the fused data. In the other fusion approach, feature-level fusion fusion, each sensor extracts feature vectors of the observation. The vectors are then concatenated to form a single feature vector, which in turn is input to any classification techniques. The output is a joint or fused decision based on the combined feature vectors from the sensors. Finally, in decision-level fusion fusion, each sensor performs a decision-making process based on its observational data. The temporary inference of each sensor can be combined to form a final decision.

We particularly focus on the feature-level fusion and decision-level fusion, as power metering system and beaconing system observe occupants from different perspectives. That is, the power metering system observes power consumption while the beaconing system investigates the signal strength received by phones belonging to individuals. We investigate per-user occupancy in shared offices using classification and decision-based methods. To the best of our knowledge, none of the earlier occupancy detection systems have attempted to extract multi-occupants presence from the fusion of power metering system and beaconing system.

This chapter begins by presenting the relevant literature in Section 7.2. The experimental design is described in Section 7.3. Next, it discusses the experiment on decision-level fusion and feature-level fusion in Sections 7.4 and 7.5, respectively. Finally, general conclusions are drawn in Section 7.6.

7.2 Relevant Literature

Various studies have assessed the efficacy of decision-level fusion. An example is based on a Bayesian approach [140]. Zhao et al. consider two-level occupancy detection. Namely, occupancy at room-level, which covers two private room offices, and occupancy at the work zone level, which covers three zones in a building floor. At the room level, the authors use Passive InfraRed sensors (PIRs), computer keyboards, and mice to sense the human presence in the workspaces. At the work zone level, they utilize WiFi connection and GPS position to infer either an occupant is outside of the building, in one of the observed zones, or coming/leaving to/from the building. The fusion is done using the Bayesian Belief Network (BBN), which requires prior probabilities of all possible events as input parameters. The parameters are assigned per sensor or information source and are learned with an Expectation Maximization (EM) algorithm if there is enough historical data available. Otherwise, the values are to be determined by experts or conducting a survey. It is shown that the accuracy increases from 73.4% (using PIR) and 92.8% (keyboard and mouse) to 96.7% by the BBN for a typical day. While they can achieve improvement, the testbed is intrusive in terms of attaching sensors in private offices. We address multiple occupants in shared offices with less intrusive sensors. For example, we use a power meter that measures the aggregate power consumption in the offices.

The other example is a fusion based on Dempster-Shafer Theory of Evidence (DST) that requires sensor beliefs or Probability Mass Assignment (PMA) from each sensory input. Nesa et al. propose a formula to compute PMA, assuming that the measurements follow a normal distribution [86]. The goal is to infer the occupancy of a single room occupied by two persons. They experiment with humidity-, light-, CO_2-, and temperature-sensors. As a comparison to the fusion technique, they implement single sensor inferences using decision trees, gradient boosting, and Linear Discriminant Analysis (LDA) approach. From a six-day training set, they validate the proposed approach in two testing sets; each contains two and seven days. The achieved result is promising for all sensory combinations that fused with a light sensor, that is, about 97% for classification with a decision tree, LDA, and DST. For the single sensor inference, the results reach 78% and 84% for CO_2 and temperature sensors, respectively. Their solution, however, lacks of occupant identification due to information insufficiency acquired from the applied sensors. That is, there is no information on how many persons are present and who they are.

A fusion algorithm inspired by the stigmergy of ant's pheromone release is proposed to detect occupancy of two single-occupancy offices [16]. The authors exploit motion, noise, and power meter sensors. Unlike our power meter that measures the consumption of multiple users, their current-transformer (CT) based power meters observe individual employees. This setup allows them to exploit mean and standard deviation of power meter readings that represent individual presence states, but with high intrusiveness of deployed devices. To combine with the other sensors, the authors perform optimization of two parameters from each sensor, i.e., amplitude intensity and dispersion decay. They then use an equation based on the natural exponential function to compute a sensor-specific value that needs to be summed up to see whether or not it exceeds a pre-determined threshold. An occupancy status is decided when the value is greater than the threshold.

A number of previous research into feature-level fusion has focused on environment sensors to count people numbers in an office room. Khan et al. propose to use at least 127-dimensional feature vectors in combining light, humidity, acoustic, and PIR sensors to detect occupancy [64]. Occupancy detection is in hierarchical forms. The lowest level is binary occupancy inference (i.e., occupancy/no occupancy state). One level higher is category inference (i.e., none, low, medium, high occupancy). Counting the number of occupants becomes the highest level of inference in the hierarchy. On each inference level, the system outputs inference as well as its confidence degree. The confidence is computed from the probability of the class occurred and considered as an additional feature to infer higher-level occupancy. Furthermore, they propose to combine with contextual information, such as computer activities and meeting schedules to improve decision confidence. Based on the experiment in an open office space ($92m^2$) and a small meeting room ($45m^2$) during up to two weeks, they show that embedding additional features can improve the classification accuracy up to 6% and for both *k*-NN and SVM methods.

Ekwevugbe et al. utilize environment sensors, such as air quality (i.e., CO_2 level and Volatile Organic Compounds (VOC)), acoustic, infrared cameras, and indoor climate (i.e., temperature, humidity, and illumination) to detect the number of occupants in addition to PIR [41]. They aim to predict occupancy numbers in a considerably large room (i.e., $8m$ x $13m$) occupied by up to six persons. They extract some features from the sensors and select the best describe class according to the criteria. The selected features are the first order of difference of CO_2, an average of case temperature, sounds related features, the difference of CO_2, the variance of case temperature, and the total duration of occupancy as detected by PIR sensor. Given the extracted features, they use backpropagation NN to estimate the occupancy numbers, resulting in about 70% accuracy per day. One major drawback of the environment sensor approach is the system's detection delay due to the slow mixture rate of the air [71]. Further, the detection is non-individualized, meaning that the estimation does not distinguish people, thus unable to be used in personalized-based control.

Mohebbi et al. establish information fusion between anonymous and eponymous sensors (i.e., PIR sensors and BLE) for localization systems of two occupants in a residential building [83]. There are 14 motion sensors, and 30 Estimote Sticker Beacons[1] (i.e., the light version of BLE nodes) in a space sized about $70m^2$. They propose to fuse the location estimation (so-called confidence-map) produced by each sensor modality. The confidence map is a set of values in the cartesian coordinate that representing location estimation sensed by the sensors. That is, the value will be one if any movement triggers PIR sensor or a mobile phone sees a broadcasting beacon about $1m$ distance (represented by $-70dBm$ RSS). The PIR sensor produce a single confidence map, while BLE produce a set of confidence maps (i.e., one map per-occupant). The fusion is processed by summing the confidence values across the entire two-dimensional coordinate-space using a weighted sum, resulting in the merged location estimation from several confidence maps. In this chapter, we utilize low-intrusive power metering system and beaconing system for occupancy detection system in adjacent shared rooms. We involve up to six participants to investigate fusion techniques that are appropriate to the available low-intrusive sensory systems.

7.3 Design

The power metering system and beaconing system have different perspectives on the presence sensing. The power metering system sees occupancy from a global infrastructure perspective. Namely, it discovers generic situations without paying attention to a specific individual. A concrete example is the aggregate power consumption X in shared offices. A classifier h_{occ} assigns a class label into power reading X, formally $h_{occ} : X \rightarrow Y$. Y is a class label that represents the presence of all individuals in the offices, $Y = \left\{y^{j_1}, y^{j_2}, \ldots, y^{j_n}\right\}$, where $y^{j_i} \in \{0, 1\}$ represents occupancy state of an individual j_i relative to his office space.

[1] https://estimote.com/products/

Table 7.1: Fusion of power consumption X_t and BLEs' RSS $\beta_t^{j_i}$

Instances	fused feature					Y_t
$t = 1$	X_1	$\vec{\beta}_1^{j_1}$	$\vec{\beta}_1^{j_2}$	...	$\vec{\beta}_1^{j_n}$	$j_1 j_2 j_3$
$t = 2$	X_2	$\vec{\beta}_2^{j_1}$	$\vec{\beta}_2^{j_2}$	...	$\vec{\beta}_2^{j_n}$	$j_1 j_2$
$\ldots$						
$t = T$	X_T	$\vec{\beta}_T^{j_1}$	$\vec{\beta}_T^{j_2}$	...	$\vec{\beta}_T^{j_n}$	$j_1 j_4 j_5$

On the other hand, the beaconing system observes occupancy from an individual perspective. That is, this system works by discovering broadcasted signals with respect to the individuals' position. Given a set of room locations $L = \{l_1, \ldots, l_r\}$, and a set of beacons $B = \{b_1, \ldots, b_m\}$, the signal strength of beacons received by any individual $j_i \in J$ sampled in any discrete time t, $t \subseteq \mathbb{N}$, is measured as $\beta_t^{j_i}$. A localization classifier h_{loc} assigns a class label $l_t^{j_i}$ into beacon readings $\beta_t^{j_i}$, formally, $h_{loc} : \beta_t^{j_i} \rightarrow l_t^{j_i}, l_t^{j_i} \in L$. To associate one's location l^{j_i} to occupancy y^{j_i}, we make use of a workspace location map that indicates the work space of each employee, formally, $m_{loc} : l_t^{j_i} \rightarrow y_t^{j_i}$. The $y_t^{j_i} = 1$, if l^{j_i} is equal to one's office space, and 0 otherwise. Such information is commonly maintained by building administrators, for example, to handle letters and parcels.

7.3.1 Fusion Techniques

We focus on the feature-level fusion and decision-level fusion. A notable difference between the two techniques is on the level where data fusion happens. The former method fuses data at the lower level before any decision has been made. The latter approach combines higher-level decisions inferred based on individual readings.

Feature-level fusion

Feature-level fusion is defined as an action of joining features from two sensory sources, before a decision has been made. This technique concatenates features from each sensory sources, followed by classifying the concatenated features. As the power metering system and beaconing system have different perspectives, a problem transformation needs to be performed. We transform the problem by concatenating a set of labels y^{j_i} to form a multi-label class $Y = \{y^{j_1}, y^{j_2}, \ldots, y^{j_n}\}$. This approach is known as *label combination* or *label power-set* (LC) method [105] or the Combination Method (CM) [104]. Once labels have been concatenated to represent employees in an area, beacon observation from the participants are also concatenated as illustrated in Table 7.1. Finally, conventional single-label classification methods may be used to build classification models by treating Y_t as an independent label.

Decision-level fusion

Decision-level fusion is defined as an action of combining outputs or decisions from multiple sources to form a final decision. We opt to utilize Dempster-Shafer Theory of Evidence (DST) to combine inferences obtained from individual sources [116]. The DST is an approach for dealing with uncertainty in a hypothesis based on evidential reasoning.

In DST reasoning, the system needs to infer a temporary decision and assign 'beliefs' over the possible hypothesis θ based on evidence S reported by sensory sources. Similar to probability, the sum of the degree of beliefs (also called as masses) is 1. The belief of any hypothesis θ is defined as the sum of all evidence S_k that supports hypothesis θ and the sub-hypotheses nested in θ [132], as given in Eq. 7.1.

$$Belief_i(\theta) = \sum_{S_k \subseteq \theta} m_i(S_k) \tag{7.1}$$

Given observation evidence from multiple sensors, the DST combination rule provides a mechanism to fuse probability masses of the observation of sensor-i (m_i) and sensor-j (m_j) as follows:

$$Belief(A) = m_i \oplus m_j(A) = \frac{\sum\limits_{A_k \cap A_{k'} = A} m_i(A_k) m_j(A_{k'})}{1 - K}, \tag{7.2}$$

$$\text{where } K = \sum_{A_l \cap A_{l'} = \varnothing} m_i(A_l) m_j(A_{l'})$$

Based on the belief of sensor-i and sensor-j in generating a proposition A, we can compute the combined belief of proposition A using the combination rule of Equation 7.2. This value is normalized by $1 - K$, where K indicates conflicts among the sources to be combined. For example, the $Belief(j_1 : present)$ is computed from the products of the belief that the sensory modalities identify the j_1's presence. The conflict factor K represents the disagreement of the two sensors towards proposition that j_1 is present, such as $m_i(present).m_j(absent)$ and $m_i(absent).m_j(present)$.

7.3.2 Metrics

To evaluate the performance of occupancy detection, we measure accuracy and F-measure per person. While decision-level fusion results single-label inferences (i.e., y^{j_i}), feature-level fusion outputs multi-label class Y due to the concatenation of occupants' presence states.

Single-label Inferences

In decision-level fusion, power metering, beaconing system, and the combination of them infer y^{j_i}, the occupancy state of an individual j_i, where $y^{j_i} \in \{present, absent\}$. The number of instances for which j_i's $presence/absence$ are correctly predicted count as TP_{j_i}/TN_{j_i}. Correspondingly, the number of instances for which j_i's $presence/absence$ are misclassified count as FP_{j_i}/FN_{j_i}. The accuracy and F-measure of each individual occupancy detection are then defined as follows:

- $Accuracy_{j_i} = \dfrac{TP_{j_i} + TN_{j_i}}{TP_{j_i} + TN_{j_i} + FP_{j_i} + FN_{j_i}}$
- $Precision_{j_i} = \dfrac{TP_{j_i}}{TP_{j_i} + FP_{j_i}}$
- $Recall_{j_i} = \dfrac{TP_{j_i}}{TP_{j_i} + FN_{j_i}}$
- $F - measure_{j_i} = 2 \cdot \dfrac{precision_{j_i} \cdot recall_{j_i}}{precision_{j_i} + recall_{j_i}}$

For multi-class problems, such as BLE-based location inferences, the F-measure is broken down into per room location as defined in Section 6.3.3.

Base-class Evaluation of Multi-label Inferences

Particularly in feature-level fusion, the system infers Y_t, a class labels that represents the presence state of all individuals. We break down the label into individual occupancy as proposed by Boutell et al. [23].

Let $Y_t = \{y_{t,j_1}, y_{t,j_2}, \ldots, y_{t,j_n}\}$ be the set of true labels for an instance at time t and $Y'_t = \{y'_{t,j_1}, y'_{t,j_2}, \ldots, y'_{t,j_n}\}$ be the set of predicted labels from classifier h at the same time t. The *hit* $H_t^{y_{t,j_i}} = 1$, if $y_{t,j_i} = y'_{t,j_i} = 1$, and 0 otherwise. Likewise, let the *true condition positive* $\hat{Y}_t^{y_{t,j_i}} = 1$, if $y_{t,j_i} = 1$, and 0 otherwise, and let the *predicted condition positive* $\hat{\hat{Y}}_t^{y_{t,j_i}} = 1$, if $y'_{t,j_i} = 1$, and 0 otherwise. The base-class recall and precision become:

- $Recall_{j_i} = \dfrac{\sum_t H_t^{y_{t,j_i}}}{\sum_t \hat{Y}_t^{y_{t,j_i}}}$

- $Precision_{j_i} = \dfrac{\sum_t H_t^{y_{t,j_i}}}{\sum_t \hat{\hat{Y}}_t^{y_{t,j_i}}}$

- $F - measure_{j_i} = 2 \cdot \dfrac{precision_{j_i} \cdot recall_{j_i}}{precision_{j_i} + recall_{j_i}}$

7.4 Experiment-1: Decision-level Fusion

We combine user presences inferred based on beaconing system and power metering system. In this section, we describe the collected data and experimental setup, covering inferences based on individual sensors and decision-level fusion. Finally, we discuss the results and findings.

7.4.1 Data

We collected training data for the beaconing system using a mobile phone (i.e., the phone belongs to individual j_1), from March 9, 2017 until May 2, 2017. As the collected data were unbalanced, we randomly down-sampled according to the smallest number of class instances (i.e., 22 instances). There were five room classes (i.e., three offices, one social corner, and a hallway) involved in the experiment. The model was built using k-NN with $k = 5$ using the cosine distance. We also measured power metering system for training purposes from March 13, 2017 until March 31, 2017. We tested the occupancy inference using fresh data from September 14, 2017 until October 30, 2017.

The occupancy ground truth was collected using a mobile application. Participants were asked to report their location whenever they moved to other rooms of the observation area. The application converts the room location to binary occupancy according to the workspace location map.

7.4.2 Setup

The appliance-metering system assumes that there is a plug-based power meter attached on each monitor, while the sub-metering system measures the total power consumption. The volunteers have mobile phones and monitor screens in their workspace, as listed in Table 7.2.

Presence from appliance-metering system We deploy plug meters to measure monitor screens' power consumption, as in Section 4.3. The appliance-metering based occupancy detection is provided to report how the screen activation is related

Table 7.2: List of mobile devices and monitors

ID	Phone (Android SDK version)	Monitor power rated
j_1	S6 edge+ (Android 7.0, API 24)	11.8 W and 21 W
j_2	S6 (Android 7.0, API 24)	$90W$ and $19.5W$
j_5	A5(2016) (Android 6.0.1, API 23)	$40W$ and $21W$
j_6	Nexus 5x (Android 7.1.1, API 25)	$24W$

to the occupant presences. The presence is detected using a threshold value on the plug meter readings. When the measured consumption is over a threshold, we infer the occupancy state as occupied.

Presence from sub-metering system Occupancy detection in this experiment is based on the switching state detection, and therefore relies on the event detection function f_{ev}, as defined in Section 4.3.2. Given a set of individuals $J = \{j_1, j_2, \dots, j_n\}$ and a set of monitors $D = \{d_1, d_2, \dots, d_u\}$, where individual $j_i \in J$ has at least one monitor $d_i \in D$, the classifier h_{recog} assigns monitor labels into detected switching events on the ordered sequence of power observation $\mathcal{O} = X_1, X_2, \dots, X_T$, formally, $h_{recog} : f_{ev}(\mathcal{O}) \rightarrow d_i, d_i \in D$.

Next, the detected switching state of monitor d_i is associated with individual occupancy y^{j_i} using an inventory list map of monitor devices, formally, $m_{mon} : d_i \rightarrow y^{j_i}$. The inventory list is assumed to be updated by buildings facility managers since they provide work-related equipment needed by employees. When an ON/OFF event is classified as a load d_i belonging to an employee j_i, the event indicates the start and end of presence states of an employee y^{j_i}, where $y^{j_i} \in \{0, 1\}$.

We assign belief to this sensory source by determining how closely the activated monitors agree with the real occupancy of the owner in the past (i.e., we choose the data from April 19, 2017 until May 1, 2017). The belief assignment is based on how close the hypothesis and sensor evidence in the past, specifically, similar to Lawhern et al. [74]. We count the frequency of positive/negative agreement and positive/negative disagreement between actual occupancy and predicted device activation. The actual occupancy is obtained from manual user input (or the appliance-metering system if user inputs are missing), while the prediction of activated devices is gathered from the device recognition module of the sub-metering system. The assigned beliefs are shown in Table 7.3.

Table 7.3: Probability Mass Assignment of room-level power meter, specific for participants j_1, j_2, j_5, j_6

	j_1_devices		j_2_devices		j_5_devices		j_6_devices	
Belief	ON	OFF	ON	OFF	ON	OFF	ON	OFF
Presence	71.35	23.71	98.50	38.10	83.20	4.03	68.50	13.63
Absence	28.65	76.29	1.50	61.90	16.80	95.97	31.50	86.37

Presence from beaconing system We utilize w-width overlapping moving windows to extract BLE features from the beacons' RSS, similar to Section 6.3.2. The extracted features are mean, the difference of consecutive means, mode, standard deviation, and maximum value of RSS. In addition, we put binary features to indicate which beacons are discovered and which one has the strongest received signals. See [99] for details. A localization classifier h_{loc} assigns l^{j_i}, the location of individual j_i, to β^{j_i}, the extracted BLE beacon features, formally $h_{loc} : \beta^{j_i} \rightarrow l^{j_i}$. The workspace location map is then used to infer binary occupancy y^{j_i} from the localization output l^{j_i}.

The system produces inference results and beliefs based on the evidence presented by RSS. It computes beliefs as the sum of the weight of nearest neighbors with the same label normalized by the sum of weights of k-NN. We use $k = 5$ and evenly distributed weight for the k-NN.

Presence from decision fusion Given the decision and evidence from each sensory modality, we fuse the inferences to obtain the final decision. DST is used to deal with uncertainty from each sensor. Occupancy is inferred based on 5-minute moving windows with 1-minute overlap (i.e., window width $w = 48$ instances, in 5-second sampling interval). We define work hours based on common observation, that is, from 7.00 AM to 9.00 PM. Hence, on a day observation, there are 210 time-windows during 14 work hours. The classification of the detected events and the classification of RSS are based on neural networks (see Chapter 4) and k-NN with cosine distance (see Chapter 6), respectively.

7.4.3 Results and Discussion

The presence inference results from appliance-metering, sub-metering, beaconing system, and decision-level fusion are shown in Table 7.4. As shown in the table, the appliance-metering may infer volunteer presences with .92 to .99 F-measure. This result shows that monitor activation may indicate occupancy as long as an employee uses a computer during their occupation and consistently put it on standby mode while he is away. Individual j_5 indicates this condition and resulting very high precision and recall. On the other hand, for a particular person, such as j_2, presence detection results in a lower recall than the other people, reaching .91 F-measure. It indicates a high number of false negatives due to j_2's presence without power consumption.

Table 7.4: Occupancy inference performance per individual.

ID	Source	Accuracy	Precision	Recall	F-measure
j_1	Appliance-metering	.9178	.9151	.9623	.9321
	Sub-metering	.6790	.6696	.9671	.7740
	BLE	**.8874**	.8630	.9741	**.9088**
	Fusion	.8712	.8429	.9745	.8970
j_2	Appliance-metering	.9005	.9458	.9107	.9194
	Sub-metering	.8907	.9483	.8953	.9096
	BLE	.7969	.7563	.9707	.8397
	Fusion	**.9008**	.9462	.8989	**.9149**
j_5	Appliance-metering	.9858	.9867	.9891	.9877
	Sub-metering	.7915	.7952	.9905	.8665
	BLE	.7970	.7565	.9935	.8557
	Fusion	**.8175**	.7962	.9907	**.8740**
j_6	Appliance-metering	.9341	.9472	.9765	.9578
	Sub-metering	.7924	.8107	.9640	.8692
	BLE	**.8947**	.8737	.9948	**.9286**
	Fusion	.8919	.8745	.9955	.9279

With respect to the beaconing system, the performance seems to depend on the used mobile phone. Since we use training data collected from individual j_1, the occupancy inference is better for users with mobile phones with similar RSS detection. That is, for the individuals j_1 and j_6, the inferences reach a value between .90 and .93 F-measure, while for individuals j_2 and j_5, the occupancy inferences using BLE achieve .84 F-measure. This is expected as we do not perform any calibration process to handle device heterogeneity (such as proposed in [120]).

Occupancy inference based on non-intrusive sub-metering in this experiment is based on electric load identification (i.e., switching state detection), as discussed in Chapter 4. The inference system using this modality yields .77 F-measure for presence of individual j_1, and reaches .90 F-measure for other individuals. These results are related with the appliances being used. Individual j_1 has two low consumption monitors that are difficult to detect and distinguish. The best inference using this modality is individual j_2 who has a notable consumption while he uses the monitor.

The results of this study indicate that neither sub-metering nor beaconing system performs well with the involved volunteers. The decision-level fusion shows improvement for individuals j_2 and j_5, while it does not for individuals j_1 and j_6 for whom the results using beaconing system are better. A possible explanation for this might be that the DST-based fusion process highly depends on the belief of its sensory inference. In this thesis, the prior belief of beaconing system is based on the nearest neighbor instances, while in the sub-metering system, the belief is based on the past data. The latter case, the performance highly relies on the pattern of appliance usage in the training dataset, and as such, the performance may vary considerably from test to test.

To see the fusion impact in more detail, we illustrate the decision-level fusion inference in two typical days in Figure 7.1 and 7.2. On September 20, 2017, the sub-metering inference detected monitor activation belong to j_1 only from 13:32 to 15:12, with .587 accuracy. The inference based on beaconing system provided better prediction with .919 accuracy. The decision level fusion worsened the beaconing system inferences in some time (i.e., 11:00-11:12; 11:56-12:00; 16:28-16.40), resulting accuracy .872 accuracy. It seems possible that the result was because the BLE inferences at these periods were not more confident than the vacancy inferred by the sub-metering system, resulting in final false-negative states (i.e., being vacant) for some periods. On September 23, 2017, the inference based on the beaconing system showed occupancy changes from 15:00 until 18:32. In the same period, inferences from the sub-metering system showed accurate detection with increased belief. In our results, the decision fusion performed better than each modality, reaching .962 accuracy.

7.5 Experiment-2: Feature-level Fusion

In the second experiment, we combine features extracted from the beaconing system and power metering system to infer user presences. We begin by describing the data collection procedure and experimental setup. We then continue with the results and discussion.

7.5.1 Data

We collected data from beaconing system and power metering system simultaneously for about four weeks, starting from October 1, 2018 until October 26, 2018. We consider only work hours, from 7.00 AM to 9.00 PM. We override the undiscovered beacons with a very weak signal strength value (i.e., $-120dBm$). Similar to the previous experiment, ground truth data collection relies on the volunteers' room-location reports.

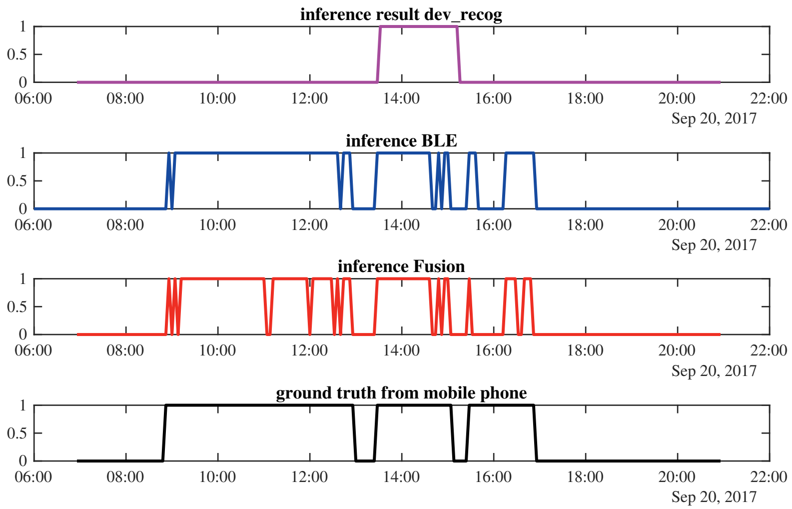

Figure 7.1: The occupancy inference of j_1 on September 20, 2017. Sensor fusion does not improve the final result.

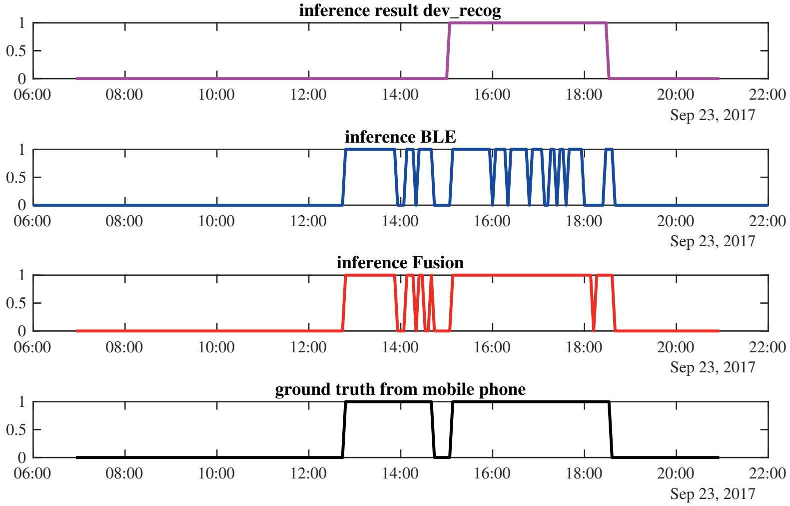

Figure 7.2: The occupancy inference of j_1 on September 23, 2017. The fusion process improves the final result

.

Table 7.5: List of mobile devices and monitors

ID	Phone (Android SDK version)	Monitor power rated
j_1	S6 edge+ (Android 7.0, API 24)	11.8 W and 21 W
j_2	LG Nexus 5x (Android 7.1.1, API 25)	20.6 W and 24 W
j_3	A5(2016) (Android 6.0.1, API 23)	34.8 W
j_4	Xperia XZ (Android 8.0.0, API 26)	64 W
j_5	Galaxy S3 (Android 4.3, API 18)	14 W

7.5.2 Setup

Depending on the portrayal of context observation, a classification can infer single-label or multi-label classes. When the provided context comes from an individual's perspective, such as the beaconing system, the inference has a single-label. In this case, a localization classifier h_{loc} assigns a label $l_t^{j_i}$ to $\beta_t^{j_i}$, a vector of M-dimensional beacons sampled by individual j_i at time t, formally, $h_{loc} : \beta_t^{j_i} \to l_t^{j_i}$. When the supplied information is in a broader or more general perspective, such as when more than one individual are involved in the measurement, the inference is on multiple labels. Multi-label classification deals with a set of labels $y \subseteq Y$, where Y is a set of disjoint labels with $|Y| \geqslant 1$. From the sub-metering system, a class label Y, which is a set of presence states of all individuals $j_i \in J$, is assigned to the total power consumption X_t. Formally, $h_{occ} : X_t \to Y_t, Y_t = \{y_t^{j_1}, y_t^{j_2}, \ldots, y_t^{j_n}\}$. From the beaconing system, a set of RSS vectors discovered by n-person may be concatenated that represents the presence state of all individuals j_i, where $i = 1, \ldots, n$. Thus classification can be drawn as $h_{occ_2} : \beta_t'' \to Y_t$, where β_t'' is a vector with $(n \cdot M)$ columns.

Feature-level fusion is done by concatenating the features, including the RSS observation over all participants (i.e., β''), and the aggregate power consumption X. This step is to fit with the observation perspective of X, which monitors the overall consumption of the users. Subsequently, the classifier h_{occ_3} assigns the label to the concatenated features, formally $h_{occ_3} : [\beta_t'', X_t] \to Y_t$, where $[\beta_t'', X_t]$ represents concatenated features at time sampled at time t.

We involve five volunteers, as listed in Table 7.5. Decision Tree (DT) is utilized as the features grow with the number of users. This method is particularly useful in efficiently classifying high dimensional feature sets by growing classification trees [58]. The trees may also provide alternative splitting nodes when some beacon nodes are not reachable (e.g., out of coverage).

The evaluation is on fresh test data, shuffled with the proportion of 85:15 for training and testing set. We thus tune parameters using Scikit-learn Randomized-SearchCV on the training data portion [92].

Table 7.6: Multi-class classification $h_{loc} : \beta^{j_i} \rightarrow l^{j_i}$. Bold font marks individuals' offices.

ID	Overall accuracy	F-measure per class -1	0	A	B	C	SC
j_1	.941	.946	.600	**.948**		.759	.681
j_2	.846	.902		**.646**			0
j_3	.826	.877			**.706**		
j_4	.985	.993	.831		**.955**	.899	.680
j_5	.990	.994			**.975**		

We also compare with Factorial HMM and Combinatorial Optimization (CO) with no classification improvement (i.e., see [95] for details). Thus we only report the inference based on DT in this thesis.

7.5.3 Results and Discussion

In this experiment, we have classifiers h_{loc} and h_{occ} that classify beacon and power meter readings. The location classifier h_{loc} assigns exactly one location label of a set of possible rooms in an office to a vector of signal strength readings. Table 7.6 shows the room-level localization based on a decision tree classifier on the RSS data. As can be seen from the table, the location inference of individual j_2 and j_3 yields .65 and .71 F-measure in their offices, respectively. This result is attributed to a high number of False Negatives (FN). In this case, it is due to missing beacon measurements occurred in a short period during occupancy. A likely explanation is that a human error (e.g., forgetting to start the application or accidentally stopping the measurement) or system failure (e.g., application crash or operating system service interruption). For the other three volunteers (i.e., j_1, j_4, j_5), their presence in their offices (either room A or B) can be inferred with more than .95 F-measure.

The location inference of individual j_1 and j_4 in SC results in a comparable performance, reaching .68 F-measure. Interestingly, the inference of the same individuals in room C is with a notable difference, reaching .759 and .90, respectively. A possible explanation for these results may be the environmental factors, such as physical room condition, signal blockage, or environment noise, giving various impact on the inference of room C.

The classifier h_{occ} assigns a label that represents a combination of individuals' occupancy states. Table 7.7 shows occupancy inference based on the beaconing system, sub-metering, and feature-level fusion, evaluated per individual j_i. The individual occupancy results confirm the previous discussion of single-label inferences. That is, the presence detection of j_1, j_4, j_5 in their office based on the beaconing system reaches .96-.98 F-measure, while the detection of j_2, j_3 yields worse results, reaching .71-.79 F-measure. The detection recalls for both individuals reach .72 and .59, respectively. The False Negatives in the beaconing system inference cause these results. Using the other system, i.e., sub-metering, the presence detection of individual j_4 in his office results .76 F-measure. This result is almost certainly due to his appearance in the office without having monitor consuming power energy. Undoubtedly, the sub-metering system cannot discover the occupancy of a person who does not leave power consumption fingerprints.

Comparing the results, one can see that the classification on the fused features outperforms the inference based on beaconing and sub-metering systems. The improvements happen for all the five volunteers, reaching .99 F-measure in the detection of particular volunteers. The results are likely due to the benefits of taking information from incomplete information from multiple sensor inputs. While the power metering fails in detecting j_4 due to no power consumption measured, the beaconing system also fails to detect j_2 and j_3 due to False Negatives (e.g., misclassification or missing beacon measurements). The feature-level fusion improves the inference by apportioning the shortcomings of each sensor input.

Table 7.7: Occupancy inference of sub-metering ($h_{occ} : X \to Y$); BLE ($h_{occ_2} : \beta'' \to Y$); and feature-level fusion ($h_{occ_3} : [\beta'', X] \to Y$), the multi-label Y is evaluated per base class j_i

ID	Source	Precision	Recall	F-measure
j_1	Sub-metering	.9911	.9728	.9819
	BLE	.9843	.9377	.9604
	Fusion	.9883	.9889	**.9886**
j_2	Sub-metering	.9696	.9377	.9534
	BLE	.8963	.7155	.7958
	Fusion	.9742	.9508	**.9623**
j_3	Sub-metering	.9595	.8416	.8967
	BLE	.9016	.5928	.7153
	Fusion	.9675	.8809	**.9222**
j_4	Sub-metering	.8428	.6957	.7622
	BLE	.9809	.9483	.9643
	Fusion	.9661	.97	**.9681**
j_5	Sub-metering	.9267	.9496	.938
	BLE	.989	.9647	.9767
	Fusion	.9906	.9889	**.9897**

7.6 Conclusion

This chapter sets out to examine the fusion of power metering system and beaconing system in detecting occupancy of users in adjacent shared offices. We investigate decision-level fusion and feature-level fusion and compare them to the inferences based on the individual systems. The former allows the individual systems to infer a local decision before performing a final inference; the latter yields an output once features of the systems have been combined.

The decision-level fusion combines the electric load identification based on sub-metering system and the location inference based on beaconing system. The presented experiment shows that the decision-level fusion results depend on the inferred occupancy of both systems. When the inferences of the systems are contradicting, the one with stronger belief will have more impact to decide the final inference. Assigning the proper beliefs, however, is a challenge. In this thesis, the belief of sub-metering system is set based on the agreement between monitor activation (i.e., detected by the appliance-metering system) and the actual presence in the past. Thus, the sensor beliefs may vary even for the same person and the same sensor, depending on the chosen portion of the training data. As a consequence, the final inference may also vary. Note that in this thesis, the results of this fusion scheme are not always improved compared to the inferences based on its individual sensory sources.

Feature-level fusion concatenates features from each sensory system. The most prominent finding from the analysis is that the results are always better than the individual sensory systems. The improvements present in the occupancy inferences of all individuals. In this fusion scheme, we show that the occupancy may still be detected when an employee does not consume electricity or when beaconing system is faulty, by relying on the other sensory system. A limitation of this approach is that it is unable to learn labels that do not occur in the training data. Furthermore, the label combination of multi-occupancy states makes a very sensitive classification and requires accurate ground truth. It is because small mistakes in labeling data can shift the present state of the other persons.

In summary, this chapter confirms that no single sensing modality outperforms in all conditions for all users. In the sub-metering system, False Negative (FN) inferences appear when the system misclassifies the active appliances. Hence, this modality performs well only when the appliance signatures are distinct. FNs also appear when people are present without consuming any consumption as the power metering system does not sense any evidence of presence. In the beaconing system, FNs happen when the inferred location is not the actual room where a person lies. Missing beacon data also affects FNs. This modality provides good results when a phone for testing matches with the phone used for collection training data. The fusion schemes show different remarks. On the one hand, decision-level fusion does not always improve the inference of the single modality because the fusion only considers local inferences that might be misclassified. It also disregards the raw data. On the other hand, feature-level fusion always improves F-measure, but the label combination approach makes classification models rigid and very sensitive to wrong labels. Feature-level fusion might be a good option when it involves a limited number of occupants, as it is relatively easier to observe and maintain label correlations for training purposes. Otherwise, one may want to consider decision-level fusion and concentrate on improving the accuracy of single sensor modality and finding the way how to acquire proper belief assignments.

Chapter 8

Conclusion

The aim of this thesis is to investigate the use of simple power metering and beaconing systems for non-intrusive occupancy detection in offices. These systems are based on off-the-shelf hardware that have been increasingly adopted to measure electrical energy use and to enable proximity-based services. They have different, complementary characteristics. Power metering system captures electricity consumption centrally, and may show the occupancy at the aggregate level (unless a plug-meter attached to each device, which is more intrusive). This characteristic makes this modality low-intrusive, but with limited information. On the other hand, beaconing system is a modality that observes signal strength in the environment from the perspective of a user. It provides identity (that should be kept to protect the actual ID) to support preference-based controls. Nonetheless, as this sensing is based on signals, it faces some problems such as the requirement to collect signal references (known as fingerprinting surveys) and is susceptible to environmental disturbances.

While a number of approaches using these simple sensors have been proposed to reveal occupancy, some gaps are still present: 1) There is a lack of study on extending power meter usage, particularly sub-metering, for occupancy detection in non-residential buildings (see Figure 2.2). As discussed in Chapter 3, occupancy from power metering system is revealed either from detecting appliances or power consumption data mining. A number of approaches have been proposed for electric load identification, but not many of them are extended to reveal occupancy. Furthermore, mining electric power consumption data have been done to extract occupancy. However, most of them use plug-based metering, either per appliance or per desk, which is intrusive. 2) Localization has been researched extensively, but most works have not concerned on the beaconing-based low-intrusive approach to recognize multi-occupant presence in adjacent shared offices; 3) To the best of our knowledge, none of the earlier fusion systems have attempted to improve the precision of multi-occupants presence detection by combining the power metering system and beaconing system. We performed empirical investigations to answer our original research questions.

8.1 Answers to the Research Questions

How is power consumption data acquired and analyzed while maintaining low-intrusiveness? How do low-intrusive power metering systems contribute to context awareness?

Acquiring low-intrusive power consumption data for context recognition can be done by deploying power meters per room. This way is less intrusive than deploying a power meter for every appliance (i.e., plug-based metering), while still giving more information than installing a power meter placed at a single point (i.e., centralized metering).

Literature review on the field shows that the power meter may determine contexts in two ways, namely, electric load identification and power consumption data mining. First, electric load identification may benefit to occupancy detection through the detection of the appliances that are related to occupancy (e.g., require physical interaction to be activated, such as computers). Second, mining data on power consumption benefits to reveal occupancy if there is a pattern indicating a situation (e.g., the presence of certain individuals related to their appliances being on).

Assuming that a power meter installed in a dedicated electric circuit of computer equipment is available, how can occupancy information be extracted? How accurate is the occupancy observation in offices based on the computer equipment activation?

One way to extract occupancy from power meter readings is by identifying operating electrical loads, for example, by detecting switching events (i.e., the action of activating or deactivating a device) followed by discovering which device turned ON/OFF. Based on the empirical observation, computer monitor activation can reveal the presence of 7 of 13 participants with only 5% error or less, and 10 of them have less than 10% error. We propose a procedure that computes the state-transition signatures on each detected event and identifies electric loads using classification techniques. We thus test the actual use cases in two offices: one is an office in the academic building (i.e., dataset A); another one is the commercial office of a software house company (i.e., dataset B). We find that the load identification performance relies on the precision of event detection. The more devices involved and the lower amount of monitor power consumption, the harder it is to detect the switching events and, thus, to identify loads accurately. The proposed procedure reaches 80% top-n accuracy per day in dataset A with up to four different monitors. It struggles in dataset B, which has more similar monitors (i.e., seven of ten monitors having the same brand and size), reaching 39% top-n accuracy per day.

Assuming that a power meter with more electrical variables are deployed in shared room offices, how are active appliances recognized, and how are the present occupants distinguished? To what extent can we make use of this information for presence detection?

The office appliances and the present occupants can be seen as electric loads (or composite loads) that contribute to the power meter readings when they consume electricity. Instead of detecting the appliances switching ON/OFF or the occupants activating/deactivating electric loads, the alternative way to extract information is to find patterns in the power readings. The pattern may exist, for example, when a specific appliance is operating or when a person present in the space regularly uses certain devices. The sliding window approach is applied to read power consumption sequentially. We use various sizes of windows and accordingly mine patterns on the power consumption with different electrical variables using nearest neighbors, neural networks, and Markov-based approaches.

The initial experiment shows that this approach can recognize five office-related appliances and their combinations with .99 Kappa measure using k-NN and LSTM. The extended experiment on the present occupant recognition shows .93-.94 Kappa measure, and it reaches an average of .833 Kappa measure per day when the power consumption patterns have not appeared in the training phase. This result is based on the presence detection and recognition of three employees in a shared office. In general, the nearest neighbors based approach performs better than the other methods considered in the experiments.

How is beaconing localization carried out while maintaining low-intrusiveness? How precise is the occupancy inference in adjacent shared office rooms using beaconing localization?

We design the low-intrusive beaconing localization by limiting the training data collected in the most visited spaces. The collection process should involve as few occupants as possible to keep the intrusiveness level low (i.e., only one in our experiment). The collected training data needs to represent different location labels to build a room-level localization. We utilize cosine similarity to measure the difference among training references and to classify based on the nearest neighbor technique. The performance of this system achieves more than .885 F-measure for detecting the presence of four of five participants in offices during a week of investigation. One participant is detected with only .61 F-measure due to his workspace position near to the separator between the adjacent rooms. Compared to the other low-intrusive works (i.e., [78, 70]), this proposal improves the inference of whole participants in places where people stay for long occupancy periods, even though the participant for whom the inference is worst is still misclassified 55% of the times.

How can sensor fusion improve occupancy inference given individual sensors' benefits and faults?

The occupancy inference based on switching state detection in the sub-metering system, presented in Chapter 4, performs worse on the detection of individuals who work with low consumption devices or do not work consistently with electrical appliances during their presence in the office. The occupancy inference based on the beaconing system, in Chapter 6, relies on the quality of training data. It tends to perform better when the phone used for localization has similar signals to the collected training data (i.e., note that signal measurements by various phones may vary even in the same environment). Also, this modality suffers from a high number of False Negatives, mainly due to missing beacon measurements that can occur in a short period during occupancy, and environmental factors (e.g., noise and obstacles).

To overcome these shortcomings, we consider sensor fusions. We work on two sensor fusion schemes, namely, decision-level fusion and feature-level fusion. The Decision-level fusion combines the inferred local decisions of individual sensory modalities. In our experiment, decision-level fusion based on Dempster-Shafer Theory of Evidence (DST) depends on the beliefs to its sensory modalities. Depending on the belief assignments, the fusion may or may not improve the final results. That is, it improves for two of four participants (i.e., j_2 and j_5), reaching .91 and .87 F-measure, respectively. However, the fusion for the other two participants is slightly lower than the inferences based only on the beaconing system, reaching .897 and .928 F-measure.

The other fusion scheme, feature-level fusion, combines the data at the lower level before any decision has been made. In our experiment, the feature-level fusion improves the inference of all participants in terms of F-measure, reaching up to .99 in the detection of two of five volunteers. This improvement shows that this fusion scheme benefit from multiple sensory modalities when one or some of them are incomplete or inaccessible.

8.2 Discussion on Energy Saving

Determining the occupancy of a room does not directly solve the problem of energy saving, but it is an essential building block to achieve such energy saving. In fact, energy-consuming appliances (e.g., air conditioners, heaters, and lights) can be controlled efficiently if the presence can be precisely determined. In non-residential buildings, we can mainly save energy by turning off unused devices in unoccupied spaces, or, even further, adjusting the electrical devices regarding the preferences or activities of present occupants. Recalling the example mentioned in Chapter 1, the office's energy saving can be more than 50% when the lighting system recognizes Aldo, Cecilia, and Diana presence in the office and activates half of the available lamp tubes in the space with 90% of brightness. The amount of saved energy may be affected by factors such as multi-occupancy detection supports, systems' precision in inferring occupancy, granular lighting control based on occupant identification, accommodation of user preferences, and reactive or predictive capabilities.

The **multi-occupancy** detection system enables lighting control based on spatial and temporal information of present occupants. The system may adjust the corresponding luminaries depending on occupancy states. A multi-occupancy system based on PIR sensors and RFID tags to recognize occupants coming to or leaving from workspaces has been proposed by Manzoor et al. [80]. The granular lighting control based on seating placements and occupants' presence achieves an energy saving of 13% in a one-day observation (i.e., 13-hour duration), compared to reactive controlling based on PIR sensors only. The **precise occupancy inference** also influences the amount of saved energy. The more accurate the system inference, the more savings can be achieved. More recently, Zou et al. [144] propose WinLight, a WiFi-based occupancy detection, and compare to PIR sensors. The occupancy detection based on PIR sensors achieves 76.91% accuracy in their experiment. Based on this inference, the average weekly consumption on lightings is 29.04 kWh. WinLight improves the accuracy of occupancy detection, reaching 99% accuracy, and saves up to 51% of energy based on its multi-occupancy inference, consuming only 14.24 kWh. This accurate multi-occupancy inference reduces up to 82.83% energy usage compared to the static scheduling, which consumes 82.94 kWh. The energy-saving may be further increased by dimming the light based on the occupants' location and preference, reaching the consumption of 5.73 kWh.

Yeh et al. accommodate reactive strategy to control personal lights, appliances, and HVAC systems [135]. The authors use individual tags to identify users and users' proximity to the active tags. Further, they also collect users' temperature preferences. The information is used to control desk lights, adjust the electric currents through wires, and set the air conditioner temperature. It is reported that 16.5-46.9% energy saving can be achieved depending on the number of people, compared to a baseline without controlling intervention. While it is generally conceded that **user preferences** improve the overall comfort levels, the energy-saving may vary depending on the building or occupancy types and user preferences. A solution to increase the saving might be to set constraints according to comfort standards. For example, as the comfort standard of working with a computer is 300-500 lux, the control system should not trigger additional lights only to satisfy user preferences when this luminous condition has been fulfilled.

Reactive strategies provide a simple way to control lighting. To control HVAC systems, however, this strategy may be less effective due to the slow response of thermal dynamics. The combination of reactive and **predictive gives a higher chance of power-saving**. According to a simulation-based study by Goyal et al., the additional savings are limited, though, if the building needs to maintain the airflow rate during unoccupied times (as per ASHRAE[1] standards) [51]. In the study, the authors compare a feedback control (i.e., based on the occupancy measurement) and model predictive controls (i.e., based on the occupancy prediction) in achieving energy saving over an HVAC system in a medium-sized office (1-5 people). The results show that the energy-saving of 42-59% over the baseline (i.e., no occupancy information provided) can be achieved using the feedback on occupancy measurement. Interestingly, an accurate occupancy prediction in 24-hour ahead only adds energy-saving of 1-13% compared to the control based on occupancy measurements, which may not worth the additional computation complexity in the prediction.

Note that the indicated energy savings are not aimed to directly compare the achieved energy savings among the systems. The performances are merely an indication, and they may vary depending on the system and experiment setups, occupancy patterns, building layouts, etc. Nonetheless, as the above studies show, multi-occupancy support, user identification, and precise inference in detection systems, as we did on this thesis, are key factors to improve energy savings.

8.3 Discussion on Privacy

The work in this thesis uses power meters and mobile phones as well as BLE beacons to acquire occupancy contexts. Power meters are deployed at the office level while the mobile phones are assumed to be available per user (e.g., provided by the employer to support work-related tasks). Compared to the camera-based surveillance and sound or noise-based sensing systems, our approach is less intrusive in terms of privacy exposure. There is no audio or video recorded for detection purposes. Furthermore, the occupants have a full control to limit localization access, as they can anytime switch the Bluetooth module on or off. While both camera and beaconing systems enable occupants' tracking, thus may be harmful to privacy, our approach does not aim at tracking movement and recognizing activities (e.g., either working or browsing entertaining sites, like social media). Current and historical data are not associated with individual productivity and may be removed to prevent misuses. We acknowledge that, however, a proper data protection mechanism is needed to protect the collected data from any harmful actions. Furthermore, clear purpose-statements need to be clearly stated, and participants must give a consent.

[1] American Society of Heating, Refrigerating and Air-Conditioning Engineers

The more details the building knows (e.g., who is present), the better the building may adjust its service to meet personal demands (thus improving user satisfaction and save energy). The personalized services and privacy invasion risks, however, are a trade-off. Personalized services require detailed contexts to understand the situation better and act accordingly. The context acquisition may be regarded as something harmful or unpleasant for a particular person, as this act puts them in the risk of data breaches. It is left to the users to decide whether they are fine with the context acquisition for better-tuned services and energy saving, or they prefer to keep their context hidden with non adaptive environments. The analogy is with better-matched search results returned by Google[2] search engine, which makes use of historical searches and cookies, compared to DuckDuckGo[3], which does not store any identification and only sets a cookie for saving site settings.

8.4 Discussion on Portability

We envision that the systems proposed in this thesis could be applied to other office buildings, independent of building's topology and usage patterns. Especially for the power metering system, the extension depends on the number of occupants in the office (i.e., measured by a sub-metering power meter). We have validated our approach for up to three occupants in a small/medium-sized shared office. We expect that it may be expanded up to five or six people, but will suffer in larger offices (i.e., more than ten people), since the amount of power consumption of each user will be more likely to be similar and overlapping, thus more difficult to distinguish accurately. As indicated in the experiments, the performance will degrade when similar electrical loads belonging to different occupants are involved in the measurement. This result is the consequence of utilizing low-intrusive sub-metering systems where the measured consumption is the aggregate of multi occupants. When the loads change (e.g., when a new employee joins or the existing one resigns), one also needs to accommodate the changes and retrain the classification models.

Most approaches proposed in this thesis perform supervised learning. Thus, the process of collecting appliances' electrical signatures and ground truth (i.e., the actual occupancy states given some power meter readings) holds a vital role in the classifier training process. To collect the signatures and occupancy states, one may build an interface that allows participants to give input on the actual states, as performed by Ruzelli et al. [112]. According to the input states, the interface triggers the system to build electric load profiles and save them to a database for further

[2] https://www.google.com/
[3] https://duckduckgo.com/

recognition.

Compared to the power metering system, the beaconing system is more flexible and compact. It is because the localization inference is made per user, so different sets of users will not affect each other significantly. Even further, our approach is designed to be low intrusive, that is, without requiring complete fingerprinting surveys to the whole area of buildings, making it more manageable to execute. Though, according to our experiments, the inference may be poor in deciding the room locations when performed in the border of two adjacent rooms.

To improve the inference performance from both sensory modalities, we perform fusions at the feature-level and decision-level. While it seems that the feature-level fusion is promising in accuracy improvement, this fusion scheme requires accurate label-features mapping. That is, this approach maps all possible combinations of sensor readings to class labels. Thus, the number of labels will grow exponentially, i.e., 2^n, as the number of loads or people increases. This mapping seems reasonable in small or medium-sized offices with up to five occupants, as we prove in the experiments. In a larger office, it seems more difficult to collect the training data of the whole presence combinations. The decision-level fusion, on the other hand, seems to provide better implementation flexibility in different office buildings, as the labeling process is done per sensor modality before any fusion processes. However, as the fusion process relies on the beliefs of each modality's inferences, it requires the collection of more historical data that represents each sensory modality's actual behavior.

To conclude, the proposed approaches have acceptable portability in small to medium-sized shared offices. Some conditions are needed to fulfill to use these systems, such as having phones and dedicated applications, regularly using electrical appliances during presences, and having an inventory list map and a workspace location map to relate the recognized appliances and inferred locations to individual occupancy states.

8.5 Future Directions

The issues explored in this thesis are open to further investigation. We outline some of many interesting, possible future directions. First, it would be interesting to involve more participants in the occupancy detection study. The studies of occupancy detection generally consider a limited number of occupants or, if not, a single occupant. One of the reasons is difficulties in collecting the non-intrusive ground truth. The ground truth collection process by visual occupancy monitoring requires immense efforts on a large scale. Consequently, accurate deployment plans and good support from participants are needed (e.g., the actual presence monitoring using cameras requires clear purpose-statements and consent from employees).

Next, this thesis focused on the occupancy context inference in multi-occupant shared offices. While the presented accuracy seems promising with low-intrusive sensory modalities, and previous studies show high opportunities in energy-saving based on an accurate system (e.g., [144]), this thesis has not implemented occupant-centric controls. The integration of the low-intrusive context-aware system and context-based building control is required to see the actual energy savings and occupants' feedback.

This thesis mostly utilize supervised classification techniques that need direct instructions to teach classifiers. Other methods that allow dynamic adjustment of controls based on the recent data might worth investigating, such as Reinforcement Learning. This approach may automate the occupancy inference but may have a slow convergence rate as it needs to explore all possible class labels for various inputs.

Finally, it is necessary to convince people that occupancy detection using the non-intrusive sensors is harmless to privacy (i.e., given proper data protection) and that the acquired information contributes to power consumption savings without sacrificing comfort and safety. This step requires intensive information campaigns for occupants.

Summary

Energy consumption for both residential and non-residential buildings is significant and has been increasing regularly. Floor area expansion and building use intensification are factors that raise energy demands, not to mention the population growth. For non-residential building, asking the user to be directly involved in energy saving can be challenging as occupants (e.g., employees) are less aware of and affected by high energy bills compared to their domestic situation. Employees are less careful when leaving empty office spaces heated and illuminated, resulting in unnecessary energy consumption. This thesis focuses on finding solutions for solving energy waste in non-residential buildings by automatically detecting presence, thus enabling energy saving automation.

To reduce energy consumption due to unnecessary use, precise and detailed user contexts play an important role. User contexts (e.g., occupancy and activity of users) provide grounds to buildings' control and energy management systems for efficient lighting and HVAC actuation. We explore sensing systems that indicate occupancy. Namely, we extract occupancy from power consumption (i.e., power metering or sub-metering systems) and proximity location (i.e., mobile phones with beaconing systems).

Power metering systems may reveal occupant contexts once electric loads can be identified. Appliance signatures are essential in this identification process. The signatures can be related to switching states (i.e., indicated by power consumption changes) and electrical traces sampled during a particular period (e.g., through sliding windows). We initially identify the relationship between employees' presence and computer monitor use in experiments in actual offices. We show that the presence of 10 out of 13 participants is related to their monitor consumption with less than 10% error, and for seven of them the error is smaller than 5%. The low-power appliances (i.e., monitors), however, are challenging to be identified based

on switching state on aggregate power readings, especially when the electric loads are similar. To this end, the accuracy measure is relaxed to top-n accuracy with $n = 2$, which means that the classification is considered correct when an event is classified as belonging to one of the most likely two classes. We show that it reaches 80% top-n accuracy per day in a small size office (i.e., up to three users). It performs much worse in a larger office (i.e., ten users with more homogeneous monitors) due to difficulties in distinguishing monitors with similar power consumption and matching pairs of ON/OFF switching events, reaching only 39% top-n accuracy. Following this experiment, we investigate electrical traces in a sub-metering system using sliding windows. We use Cohen's Kappa, a measure of the agreement between the observational accuracy and hypothetical expected accuracy, to avoid bias due to imbalanced class distribution. In the beginning, we notice that aggregate power consumption of appliances on a small scale (e.g., belonging to a single employee) can be identified almost perfectly. Through another experiment which involves three employees in a shared office with random loads, we show that the employee presence can be distinguished, reaching .93 Kappa measure.

Mobile phones and off-the-shelf Bluetooth Low Energy (BLE) beaconing system are used to reveal occupancy. Based on those devices, room occupancy context may be identified, for instance, based on the nearest or strongest transmitting beacon, matching received signal strength to fingerprints, or classifying a room based on received signals. One benefit of this modality is that it uses available phones that support location inferences without requiring users to carry additional hardware. Mobile phones also allow individual preference to be collected through a user interface (e.g., tolerable temperature ranges and preferred light intensity in particular conditions). Thus, it supports actuation control not only based on location but also occupant identification and preferences. However, a system involving multi-occupants faces device heterogeneity problems, not to mention other issues due to signal propagation. We observe that multiple devices receive irregular unexpected signal strength that may cause problems in distinguishing adjacent rooms. In general, additional steps (e.g., calibration or complete fingerprint surveys) are needed to improve inference precision. This step, however, requires significant efforts to be performed. We propose to use limited training data in a few observation points to lower the setup effort. Collecting good training data is a key to achieve acceptable performance; that is, training data observed in some places must be able to describe its room locations. For this purpose, we compare the direction of training data based on cosine distance, resulting in validated training data. We then extract features and apply the nearest neighbor approach to infer the room location of each occupant. We show that four of five employees are detected with at least .88 F-measure. Another participant is detected with only .61 F-measure due to his workspace location closed to a separator of adjacent rooms. Our approach outperforms other related low-intrusive methods except for short occupancy periods. In this specific case, the indication of the nearest beacons results in more accurate inference.

To conclude the work, we consider fusion approaches of the power metering and beaconing systems to improve inference precision. We study fusions at decision-level and feature-level. The former allows sub-systems to infer local decisions and combines the outputs to form a final decision. The latter yields only a decision after sensor readings have been combined. The approaches are tested in an actual office environment populated by researchers and software developers.

The general theme of this thesis is to show how low-intrusive sensing modalities may provide precise, detailed occupancy in small-medium shared offices. We seek how far we can extract information from the available systems. We finally discuss potential energy saving, user privacy, and portability, to provide insight into how the proposed occupancy detection systems may impact building use and control.

Samenvatting

Het energieverbruik van zowel woningen als utiliteitsgebouwen is aanzienlijk en neemt voortdurend toe. Uitbreiding en intensiever gebruik van gebouwen, evenals een groeiende bevolking, leiden tot een groeiende vraag naar energie. Voor utiliteitsgebouwen is het een uitdaging om de gebruikers te betrekken bij het besparen van energie, aangezien hoge energiekosten voor hen minder inzichtelijk zijn en geen directe gevolgen hebben zoals voor de eigen woonruimte het geval is. Werknemers zijn veelal minder oplettend als het gaat om het nodeloos verwarmen en verlichten van ongebruikte kantoren, wat resulteert in onnodig energieverbruik. Dit proefschrift is gericht op het vinden van oplossingen om energieverspilling in utiliteitsgebouwen te verminderen door het automatisch detecteren van aanwezigheid, waardoor het mogelijk wordt om energiebesparende maatregelen te automatiseren.

Om het energieverbruik te verlagen en verspilling tegen te gaan is het belangrijk om over nauwkeurige en gedetailleerde informatie betreft de bezetting van een gebouw en de activiteiten van de gebruikers te beschikken, waarmee systemen de verlichting en klimaatregeling in het gebouw efficiënt kunnen beheren. We onderzoeken systemen waarmee de bezetting van een gebouw gedetecteerd kan worden, gebaseerd op het energieverbruik, gemeten op verschillende niveaus, en positiebepaling met behulp van mobiele telefoons in combinatie met Bluetooth beacons.

Aan de hand van het gemeten energieverbruik is het mogelijk om verschillende apparaten te herkennen, waarmee vervolgens de aanwezigheid en activiteiten van gebruikers achterhaald kunnen worden. Hiervoor zijn de verbruikskenmerken van de apparaten belangrijk, die zowel gebaseerd kunnen zijn op veranderingen in het verbruik alsook op het verbruik gemeten over een bepaalde periode. In eerste instantie tonen we het verband aan tussen de aanwezigheid van werknemers in een kantoor en het gebruik van het computerscherm. Voor 10 van de 13 werknemers kan de aanwezigheid achterhaald worden met een foutmarge van minder dan 10%,

en voor 7 van deze werknemers is de foutmarge minder dan 5%. Apparaten met een laag verbruik, zoals computerschermen, zijn echter lastig te herkennen op basis van veranderingen in het verbruik wanneer dit verbruik over meerdere apparaten gemeten wordt, voornamelijk als deze apparaten sterke gelijkenissen vertonen. Om deze reden wordt de nauwkeurigheid gemeten met het flexibeler top-n voor n = 2, waarmee een classificatie als correct wordt beschouwd als deze tot de twee meest waarschijnlijke behoort. We tonen aan dat hiermee een nauwkeurigheid van 80% per dag wordt behaald voor een klein kantoor bestaande uit maximaal drie mensen. Voor een groter kantoor, met tien gebruikers en veelal vergelijkbare schermen, is de nauwkeurigheid met 39% beduidend lager, voornamelijk doordat het lastig is de verschillende monitoren en aan en uit signalen te onderscheiden. Vervolgens onderzoeken we verbruiksmetingen over meerdere apparaten met behulp van sliding windows op basis van Cohen's Kappa, een criterium die robuuster is voor categorieën van verschillende groottes doordat deze de overeenstemming tussen de waargenomen nauwkeurigheid en de verwachte nauwkeurigheid meet. Op kleine schaal, wanneer de apparaten maar van één werknemer zijn, kunnen deze bijna foutloos worden herkend. Een ander experiment voor een kantoor dat door drie mensen wordt gedeeld, waarvoor een Kappa waarde van .93 behaald wordt, toont aan dat de aanwezigheid van de individuen achterhaald kan worden.

Mobiele telefoons en Bluetooth Low Energy (BLE) beacons worden gebruikt voor het bepalen van aanwezigheid, bijvoorbeeld op basis van de dichtstbijzijnde beacon of de beacon met het sterkste signaal, door de signaalsterktes te vergelijken met fingerprints, of door de ruimte te bepalen aan de hand van de ontvangen signalen. Een voordeel van deze methode is dat het mogelijk is om de locatie te bepalen zonder dat het voor gebruikers noodzakelijk is hiervoor extra apparatuur bij zich te dragen. Via mobiele telefoons kunnen ook andere voorkeuren worden vergaard, zoals bijvoorbeeld de gewenste temperatuur en lichtintensiteit in bepaalde situaties. Het is dus niet alleen mogelijk om processen in het gebouw te regelen op basis van locatie maar ook op basis van identiteit en voorkeuren. Zo'n systeem moet echter met verschillende type apparaten werken, en heeft ook te maken met problemen zoals de manier waarop signalen zich verspreiden. Zo observeren we bijvoorbeeld sterke afwijkingen in de gemeten signaalsterkte tussen verschillende telefoons, wat problemen kan geven bij het onderscheid maken tussen aangrenzende ruimtes. Om de nauwkeurigheid te verbeteren zijn aanvullende stappen vereist, zoals kalibratie of het in kaart brengen van de signalen door middel van fingerprint surveys. Dergelijke stappen zijn echter zeer ingewikkeld en tijdrovend. Wij stellen voor om een beperkte hoeveelheid informatie te verzamelen door op een klein aantal plaatsen metingen te verrichten, wat het opzetten van het systeem aanzienlijk eenvoudiger maakt. Het is daarbij belangrijk dat deze locatiegegevens de verschillende ruimtes nauwgezet in kaart brengen. Hiervoor onderzoeken we locatiedata op basis van Cosine Distance, waar verschillende kenmerken uitgehaald worden om vervolgens de locatie van elke gebruiker te benaderen met de Nearest Neighbor methode. We tonen aan dat vier van de vijf werknemers gelokaliseerd kunnen worden met een F-measure van tenminste .88. Voor de andere persoon is de F-measure maar .61, voornamelijk omdat de werkplek dicht bij een aangrenzende ruimte gelegen is. Onze aanpak presteert beter dan andere methoden met een vergelijkbare impact op de gebruikers, behalve voor korte perioden van aanwezigheid. Voor dergelijke gevallen is het bepalen van de locatie op basis van de dichtstbijzijnde beacon nauwkeuriger.

Als laatste bekijken we twee manieren om de informatie van het energieverbruik en de Bluetooth beacons te combineren om de nauwkeurigheid te verbeteren. In het ene geval worden de resultaten van de afzonderlijke systemen gecombineerd om tot een uiteindelijk resultaat te komen. In het andere geval is het resultaat rechtstreeks gebaseerd op de samengevoegde sensordata. Beide methoden worden getest in een kantoor waar verschillende onderzoekers en softwareontwikkelaars werken.

Het doel van dit proefschrift is om aan te tonen dat niet-ingrijpende meetsystemen in staat zijn om precieze en gedetailleerde informatie te geven over aanwezigheid in kleine en middelgrote kantoren. We onderzoeken in welke mate informatie van deze systemen verkregen kan worden, en bespreken tenslotte mogelijkheden tot energiebesparing, privacy van gebruikers en portability, om inzicht te geven in de manier waarop het voorgestelde systeem het gebruik en beheer van gebouwen kan veranderen.

Bibliography

[1] IEEE Standard for Transitions, Pulses, and Related Waveforms - Redline. *IEEE Std 181-2011 (Revision of IEEE Std 181-2003) - Redline*, pages 1–71, 2011.

[2] *Technology Roadmap Energy Efficient Building Envelopes*. International Energy Agency, 2013.

[3] *United Nations Environment Programmes Sustainable Building and Climate Initiative*. International Energy Agency, 2013.

[4] *2018 Global Status Report - Towards a zero-emission, efficient and resilient buildings and construction sector*. United Nations Environment Programme, 2018.

[5] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles. Towards a better understanding of context and context-awareness. In *Handheld and Ubiquitous Computing*, pages 304–307, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg.

[6] I. Abubakar, S. Khalid, M. Mustafa, H. Shareef, and M. Mustapha. Application of load monitoring in appliances' energy management – a review. *Renewable and Sustainable Energy Reviews*, 67:235 – 245, 2017.

[7] M. S. Adam Cooper. Electric company smart meter deployments: Foundation for a smart grid (2019 update). `https://www.edisonfoundation.net/iei/publications/Documents/IEI_Smart%20Meter%20Report_2019_FINAL.pdf`, 2019. (Accessed on 05/05/2020).

[8] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, and T. Weng. Occupancy-driven energy management for smart building automation. In *Proc. of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, BuildSys '10, pages 1–6, 2010.

[9] C. C. Aggarwal. *Data Classification: Algorithms and Applications*. Chapman & Hall/CRC, 1st edition, 2014.

[10] A. Akbar, M. Nati, F. Carrez, and K. Moessner. Contextual occupancy detection for smart office by pattern recognition of electricity consumption data. In *2015 IEEE International Conference on Communications (ICC)*, pages 561–566, 2015.

[11] A. Alhamoud, P. Xu, F. Englert, A. Reinhardt, P. Scholl, D. Boehnstedt, and R. Steinmetz. Extracting human behavior patterns from appliance-level power consumption data. In *Wireless Sensor Networks*, pages 52–67, Cham, 2015. Springer International Publishing.

[12] A. Allouhi, Y. E. Fouih, T. Kousksou, A. Jamil, Y. Zeraouli, and Y. Mourad. Energy consumption and efficiency in buildings: current status and future trends. *Journal of Cleaner Production*, 109:118 – 130, 2015. Special Issue: Toward a Regenerative Sustainability Paradigm for the Built Environment: from vision to reality.

[13] K. C. Armel, A. Gupta, G. Shrimali, and A. Albert. Is disaggregation the holy grail of energy efficiency? the case of electricity. *Energy Policy*, 52:213 – 234, 2013. Special Section: Transition Pathways to a Low Carbon Economy.

[14] B. Balaji, J. Xu, A. Nwokafor, R. Gupta, and Y. 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*, SenSys '13, pages 17:1–17:14, New York, NY, USA, 2013. ACM.

[15] P. Baronti, P. Barsocchi, S. Chessa, F. Mavilia, and F. Palumbo. Indoor bluetooth low energy dataset for localization, tracking, occupancy, and social interaction. *Sensors*, 18(12), 2018.

[16] P. Barsocchi, A. Crivello, M. Girolami, F. Mavilia, and E. Ferro. Are you in or out? monitoring the human behavior through an occupancy strategy. In *2016 IEEE Symposium on Computers and Communication (ISCC)*, pages 159–162, 2016.

[17] P. Barsocchi, A. Crivello, M. Girolami, F. Mavilia, and F. Palumbo. Occupancy detection by multi-power bluetooth low energy beaconing. In *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–6, 2017.

[18] K. Basu, V. Debusschere, A. Douzal-Chouakria, and S. Bacha. Time series distance-based methods for non-intrusive load monitoring in residential buildings. *Energy and Buildings*, 96:109 – 117, 2015.

[19] C. Beckel, W. Kleiminger, R. Cicchetti, T. Staake, and S. Santini. The eco data set and the performance of non-intrusive load monitoring algorithms. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, BuildSys '14, pages 80–89, New York, NY, USA, 2014. ACM.

[20] C. Belley, S. Gaboury, B. Bouchard, and A. Bouzouane. An efficient and inexpensive method for activity recognition within a smart home based on load signatures of appliances. *Pervasive and Mobile Computing*, 12:58 – 78, 2014.

[21] A. Ben-Hur and J. Weston. *A User's Guide to Support Vector Machines*, pages 223–239. Humana Press, Totowa, NJ, 2010.

[22] M. Berges, E. Goldman, H. S. Matthews, and L. Soibelman. Training load monitoring algorithms on highly sub-metered home electricity consumption data. *Tsinghua Science and Technology*, 13(S1):406–411, 2008.

[23] M. R. Boutell, J. Luo, X. Shen, and C. M. Brown. Learning multi-label scene classification. *Pattern Recognition*, 37(9):1757 – 1771, 2004.

[24] A. Caliskan, F. Yamaguchi, E. Dauber, R. Harang, K. Rieck, R. Greenstadt, and A. Narayanan. When coding style survives compilation: De-anonymizing programmers from executable binaries. *Proceedings 2018 Network and Distributed System Security Symposium*, 2018.

[25] M. Castillo-Cara, J. Lovon-Melgarejo, G. Bravo-Rocca, L. Orozco-Barbosa, and I. Garcia-Varea. An empirical study of the transmission power setting for bluetooth-based indoor localization mechanisms. *Sensors*, 17(6), 2017.

[26] D. Chen, S. Barker, A. Subbaswamy, D. Irwin, and P. Shenoy. Non-intrusive occupancy monitoring using smart meters. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, BuildSys'13, pages 9:1–9:8, New York, NY, USA, 2013. ACM.

[27] Z. Chen, C. Jiang, and L. Xie. Building occupancy estimation and detection: A review. *Energy and Buildings*, 169:260 – 270, 2018.

[28] R. Cicchetti. Nilm-eval: Disaggregation of real-world electricity consumption data. *Master's thesis, ETH Zurich*, 2014.

[29] J. Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46, 1960.

[30] G. Conte, M. De Marchi, A. A. Nacci, V. Rana, and D. Sciuto. Bluesentinel: A first approach using ibeacon for an energy efficient occupancy detection system. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, BuildSys '14, pages 11–19, New York, NY, USA, 2014. ACM.

[31] M. Conti, M. Nati, E. Rotundo, and R. Spolaor. Mind the plug! laptop-user recognition through power consumption. In *Proceedings of the 2nd ACM International Workshop on IoT Privacy, Trust, and Security*, IoTPTS '16, pages 37–44, New York, NY, USA, 2016. ACM.

[32] A. Cooper. Electric company smart meter deployments: Foundation for a smart grid. `http://www.edisonfoundation.net/iei/publications/Documents/Final%20Electric%20Company%20Smart%20Meter%20Deployments-%20Foundation%20for%20A%20Smart%20Energy%20Grid.pdf`, 2016. (Accessed on 09/03/2018).

[33] A. Corna, L. Fontana, A. A. Nacci, and D. Sciuto. Occupancy detection via ibeacon on android devices for smart building management. In *2015 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 629–632, 2015.

[34] D. F. Cowan. Security and confidentiality on laboratory computer systems. In *Informatics for the clinical laboratory*, pages 59–86. Springer, 2005.

[35] C. de Bakker, M. Aarts, H. Kort, and A. Rosemann. The feasibility of highly granular lighting control in open-plan offices: Exploring the comfort and energy saving potential. *Building and Environment*, 142:427 – 438, 2018.

[36] A. De Paola, M. Ortolani, G. Lo Re, G. Anastasi, and S. K. Das. Intelligent management systems for energy efficiency in buildings: A survey. *ACM Comput. Surv.*, 47(1):13:1–13:38, 2014.

[37] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan. *Neural Network Design*. Martin Hagan, USA, 2nd edition, 2014.

[38] Y. Du, L. Du, B. Lu, R. Harley, and T. Habetler. A review of identification and monitoring methods for electric loads in commercial and residential buildings. In *2010 IEEE Energy Conversion Congress and Exposition*, pages 4527–4533, 2010.

[39] A. Ebadat, G. Bottegal, D. Varagnolo, B. Wahlberg, and K. H. Johansson. Regularized deconvolution-based approaches for estimating room occupancies. *IEEE Transactions on Automation Science and Engineering*, 12(4):1157–1168, 2015.

[40] EIA. Frequently Asked Questions - How many smart meters are installed in the United States, and who has them? `https://www.eia.gov/tools/faqs/faq.php?id=108&t=3`, 2018. (Accessed on 29/06/2020).

[41] T. Ekwevugbe, N. Brown, V. Pakka, and D. Fan. Improved occupancy monitoring in non-domestic buildings. *Sustainable Cities and Society*, 30:97 – 107, 2017.

[42] EuropeanComission. Cost-benefit analyses & state of play of smart metering deployment in the eu-27. `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52014SC0189`, 2014. (Accessed on 09/03/2018).

[43] R. Faragher and R. Harle. An analysis of the accuracy of bluetooth low energy for indoor positioning applications. In *Proceedings of the 27th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2014)*, volume 812, pages 201–210, 2014.

[44] R. Faragher and R. Harle. Location fingerprinting with bluetooth low energy beacons. *IEEE Journal on Selected Areas in Communications*, 33(11):2418–2428, 2015.

[45] A. Filippoupolitis, W. Oliff, and G. Loukas. Bluetooth low energy based occupancy detection for emergency management. In *2016 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS)*, pages 31–38, 2016.

[46] G. F. Franklin, J. D. Powell, and A. Emami-Naeini. *Feedback Control of Dynamic Systems*. Pearson, 7th edition, 2014.

[47] I. Georgievski, T. A. Nguyen, F. Nizamic, B. Setz, A. Lazovik, and M. Aiello. Planning meets activity recognition: Service coordination for intelligent buildings. *Pervasive and Mobile Computing*, 38:110 – 139, 2017.

[48] Z. Ghahramani and M. I. Jordan. Factorial hidden markov models. In *Advances in Neural Information Processing Systems*, pages 472–478, 1996.

[49] S. K. Ghai, L. V. Thanayankizil, D. P. Seetharam, and D. Chakraborty. Occupancy detection in commercial buildings using opportunistic context sources. In *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, pages 463–466, 2012.

[50] P. Golik, P. Doetsch, and H. Ney. Cross-entropy vs. squared error training: a theoretical and experimental comparison. In *Interspeech*, pages 1756–1760, 2013.

[51] S. Goyal, H. A. Ingley, and P. Barooah. Occupancy-based zone-climate control for energy-efficient buildings: Complexity vs. performance. *Applied Energy*, 106:209 – 221, 2013.

[52] A. Graves, A. Mohamed, and G. Hinton. Speech recognition with deep recurrent neural networks. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 6645–6649, 2013.

[53] D. L. Hall and A. K. Garga. Pitfalls in data fusion (and how to avoid them). In *Proceedings of the Second International Conference on Information Fusion (Fusion'99)*, volume 1, pages 429–436, 1999.

[54] D. L. Hall and J. Llinas. An introduction to multisensor data fusion. *Proceedings of the IEEE*, 85(1):6–23, 1997.

[55] G. W. Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992.

[56] S. Hattori and Y. Shinohara. Actual consumption estimation algorithm for occupancy detection using low resolution smart meter data. In *Proceedings of the 6th International Conference on Sensor Networks - Volume 1: SENSORNETS,*, pages 39–48. INSTICC, SciTePress, 2017.

[57] Q. Huang, Z. Ge, and C. Lu. Occupancy estimation in smart buildings using audio-processing techniques, 2016.

[58] G. James, D. Witten, T. Hastie, and R. Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[59] F. Jazizadeh and B. Becerik-Gerber. *A Novel Method for Non Intrusive Load Monitoring of Lighting Systems in Commercial Buildings*, pages 523–530. 2012.

[60] X. Jiang, S. Dawson-Haggerty, P. Dutta, and D. Culler. Design and implementation of a high-fidelity ac metering network. In *2009 International Conference on Information Processing in Sensor Networks*, pages 253–264, 2009.

[61] M. Jin, R. Jia, Z. Kang, I. C. Konstantakopoulos, and C. J. Spanos. Presencesense: Zero-training algorithm for individual presence detection based on power monitoring. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, BuildSys '14, pages 1–10, New York, NY, USA, 2014. ACM.

[62] M. Jin, R. Jia, and C. J. Spanos. Virtual occupancy sensing: Using smart meters to indicate your presence. *IEEE Transactions on Mobile Computing*, 16(11):3264–3277, 2017.

[63] M. Kalksma., B. Setz., A. R. Pratama., I. Georgievski., and M. Aiello. Mining sequential patterns for appliance usage prediction. In *Proceedings of the 7th International Conference on Smart Cities and Green ICT Systems - Volume 1: SMARTGREENS,*, pages 23–33. INSTICC, SciTePress, 2018.

[64] A. Khan, J. Nicholson, S. Mellor, D. Jackson, K. Ladha, C. Ladha, J. Hand, J. Clarke, P. Olivier, and T. Plötz. Occupancy monitoring using environmental & context sensors

and a hierarchical analysis framework. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, BuildSys '14, pages 90–99, New York, NY, USA, 2014. ACM.

[65] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization, 2014.

[66] W. Kleiminger, C. Beckel, and S. Santini. Household occupancy monitoring using electricity meters. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 975–986, New York, NY, USA, 2015. ACM.

[67] W. Kleiminger, C. Beckel, T. Staake, and S. Santini. Occupancy detection from electricity consumption data. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, BuildSys'13, pages 10:1–10:8, New York, NY, USA, 2013. ACM.

[68] S. Knerr, L. Personnaz, and G. Dreyfus. Single-layer learning revisited: a stepwise procedure for building and training a neural network. In *Neurocomputing*, pages 41–50. Springer Berlin Heidelberg, 1990.

[69] Kosuke Suzuki, Shinkichi Inagaki, Tatsuya Suzuki, Hisahide Nakamura, and Koichi Ito. Nonintrusive appliance load monitoring based on integer programming. In *2008 SICE Annual Conference*, pages 2742–2747, 2008.

[70] A. I. Kyritsis, P. Kostopoulos, M. Deriaz, and D. Konstantas. A BLE-based probabilistic room-level localization method. In *2016 International Conference on Localization and GNSS (ICL-GNSS)*, pages 1–6, 2016.

[71] T. Labeodan, W. Zeiler, G. Boxem, and Y. Zhao. Occupancy measurement in commercial office buildings for demand-driven control applications—a survey and detection system evaluation. *Energy and Buildings*, 93:303 – 314, 2015.

[72] K. P. Lam, M. Hoynck, R. Zhang, B. Andrews, Y.-S. Chiou, B. Dong, D. Benitez, et al. Information-theoretic environmental features selection for occupancy detection in open offices. In *Eleventh International IBPSA Conference*, pages 27–30. Citeseer, 2009.

[73] G. Laput, Y. Zhang, and C. Harrison. Synthetic sensors: Towards general-purpose sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 3986–3999, New York, NY, USA, 2017. ACM.

[74] V. Lawhern, W. D. Hairston, and K. Robbins. Detect: A matlab toolbox for event detection and identification in time series, with applications to artifact detection in eeg signals. *PloS one*, 8(4):e62944, 2013.

[75] S. Lee, G. Ryu, Y. Chon, R. Ha, and H. Cha. Automatic standby power management using usage profiling and prediction. *IEEE Transactions on Human-Machine Systems*, 43(6):535–546, 2013.

[76] S.-C. Lee, G.-Y. Lin, W.-R. Jih, and J. Y.-J. Hsu. Appliance recognition and unattended appliance detection for energy conservation. In *Proceedings of the 5th AAAI Conference on Plan, Activity, and Intent Recognition*, AAAIWS'10-05, pages 37–44. AAAI Press, 2010.

[77] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng. Load signature study - part i: Basic concept, structure, and methodology. *IEEE Transactions on Power Delivery*, 25(2):551–560, 2010.

[78] X. Lin, T. Ho, C. Fang, Z. Yen, B. Yang, and F. Lai. A mobile indoor positioning system based on ibeacon technology. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4970–4973, 2015.

[79] X. Lu, H. Wen, H. Zou, H. Jiang, L. Xie, and N. Trigoni. Robust occupancy inference with commodity wifi. In *2016 IEEE 12th Int. Conf. on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 1–8, 2016.

[80] F. Manzoor, D. Linton, and M. Loughlin. Occupancy monitoring using passive rfid technology for efficient building lighting control. In *2012 Fourth International EURASIP Workshop on RFID Technology*, pages 83–88, 2012.

[81] M. S. Mashuk, J. Pinchin, P. Siebers, and T. Moore. A smart phone based multi-floor indoor positioning system for occupancy detection. In *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 216–227, 2018.

[82] R. Melfi, B. Rosenblum, B. Nordman, and K. Christensen. Measuring building occupancy using existing network infrastructure. In *2011 International Green Computing Conference and Workshops*, pages 1–8, 2011.

[83] P. Mohebbi, E. Stroulia, and I. Nikolaidis. Sensor-data fusion for multi-person indoor location estimation. *Sensors*, 17(10), 2017.

[84] M. V. Moreno, M. A. Zamora, and A. F. Skarmeta. User-centric smart buildings for energy sustainable smart cities. *Transactions on Emerging Telecommunications Technologies*, 25(1):41–55, 2014.

[85] M. Moreno-Cano, M. Zamora-Izquierdo, J. Santa, and A. F. Skarmeta. An indoor localization system based on artificial neural networks and particle filters applied to intelligent buildings. *Neurocomputing*, 122:116 – 125, 2013. Advances in cognitive and ubiquitous computing.

[86] N. Nesa and I. Banerjee. Iot-based sensor data fusion for occupancy sensing using dempster–shafer evidence theory for smart buildings. *IEEE Internet of Things Journal*, 4(5):1563–1570, 2017.

[87] C. Olah. Understanding lstm networks. `https://colah.github.io/posts/2015-08-Understanding-LSTMs/`, 2015. (Accessed on 05/05/2020).

[88] J. Paek, J. Ko, and H. Shin. A measurement study of BLE ibeacon and geometric adjustment scheme for indoor location-based mobile applications. *Mobile Information Systems*, 2016, 2016.

[89] F. Paradiso, F. Paganelli, A. Luchetta, D. Giuli, and P. Castrogiovanni. Ann-based appliance recognition from low-frequency energy monitoring data. In *2013 IEEE 14th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, pages 1–6, 2013.

[90] J. Y. Park, T. Dougherty, H. Fritz, and Z. Nagy. Lightlearn: An adaptive and occupant centered controller for lighting based on reinforcement learning. *Building and Environment*, 147:397 – 414, 2019.

[91] J. Y. Park, M. M. Ouf, B. Gunay, Y. Peng, W. O'Brien, M. B. Kjærgaard, and Z. Nagy. A critical review of field implementations of occupant-centric building controls. *Building and Environment*, 165:106351, 2019.

[92] 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*, 12:2825–2830, 2011.

[93] T. Petrovic, K. Echigo, and H. Morikawa. Detecting presence from a WiFi router's electric power consumption by machine learning. *IEEE Access*, 6:9679–9689, 2018.

[94] A. R. Pratama, F. J. Blaauw, A. Lazovik, and M. Aiello. Office low-intrusive occupancy detection based on power consumption. in press.

[95] A. R. Pratama, A. Lazovik, and M. Aiello. Office multi-occupancy detection using BLE beacons and power meters. In *2019 IEEE 10th Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, pages 0440–0448, 2019.

[96] A. R. Pratama, F. J. Simanjuntak, A. Lazovik, and M. Aiello. Low-power appliance recognition using recurrent neural networks. In *Applications of Intelligent Systems*, volume 310, pages 239–250, 2018.

[97] A. R. Pratama, Widyawan, A. Lazovik, and M. Aiello. Indoor self-localization via bluetooth low energy beacons. *IDRBT JOURNAL OF IJBT*, 1(1):1–15, 2017.

[98] A. R. Pratama, Widyawan, A. Lazovik, and M. Aiello. Power-based device recognition for occupancy detection. In *Service-Oriented Computing – International Conference on Service Oriented Computing (ICSOC) 2017 Workshops*, pages 174–187, 2018.

[99] A. R. Pratama, W. Widyawan, A. Lazovik, and M. Aiello. Multi-user low intrusive occupancy detection. *Sensors*, 18(3), 2018.

[100] G. D. Putra, A. R. Pratama, A. Lazovik, and M. Aiello. Comparison of energy consumption in Wi-Fi and bluetooth communication in a smart building. In *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 1–6, 2017.

[101] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proc. of the IEEE*, 77(2):257–286, 1989.

[102] M. Radhakrishnan, A. Misra, R. K. Balan, and Y. Lee. Smartphones and BLE services: Empirical insights. In *2015 IEEE 12th International Conference on Mobile Ad Hoc and Sensor Systems*, pages 226–234, 2015.

[103] T. S. Rappaport et al. *Wireless communications: principles and practice*, volume 2. 1996.

[104] J. Read, B. Pfahringer, and G. Holmes. Multi-label classification using ensembles of pruned sets. In *2008 Eighth IEEE Int. Conf. on Data Mining*, pages 995–1000, 2008.

[105] J. Read, B. Pfahringer, G. Holmes, and E. Frank. Classifier chains for multi-label classification. In *Machine Learning and Knowledge Discovery in Databases*, pages 254–269, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[106] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz. On the accuracy of appliance identification based on distributed load metering data. In *2012 Sustainable Internet and ICT for Sustainability (SustainIT)*, pages 1–9, 2012.

[107] A. Reinhardt, D. Burkhardt, M. Zaheer, and R. Steinmetz. Electric appliance classification based on distributed high resolution current sensing. In *37th Annual IEEE Conference on Local Computer Networks - Workshops*, pages 999–1005, 2012.

[108] A. Ridi, C. Gisler, and J. Hennebert. ACS-F2 — a new database of appliance consumption signatures. In *2014 6th International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, pages 145–150, 2014.

[109] A. Ridi, C. Gisler, and J. Hennebert. A survey on intrusive load monitoring for appliance recognition. In *2014 22nd International Conference on Pattern Recognition*, pages 3702–3707, 2014.

[110] A. Ridi, C. Gisler, and J. Hennebert. User interaction event detection in the context of appliance monitoring. In *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pages 323–328, 2015.

[111] A. Rogriguez, S. T. Smith, A. Kiff, and B. Potter. Small power load disaggregation in office buildings based on electrical signature classification. In *2016 IEEE International Energy Conference (ENERGYCON)*, pages 1–6, 2016.

[112] A. G. Ruzzelli, C. Nicolas, A. Schoofs, and G. M. P. O'Hare. Real-time recognition and profiling of appliances through a single electricity sensor. In *2010 7th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pages 1–9, 2010.

[113] T. Ryberg. Smart metering in europe, m2m research series 2018. `http://www.berginsight.com/ReportPDF/ProductSheet/bi-sm14-ps.pdf`, 2018. (Accessed on 05/05/2020).

[114] B. Schilit, N. Adams, and R. Want. Context-aware computing applications. In *1994 First Workshop on Mobile Computing Systems and Applications*, pages 85–90, 1994.

[115] A. C. Scogna, H. Shim, J. Yu, C. Oh, S. Cheon, N. Oh, and D. Kim. Rfi and receiver sensitivity analysis in mobile electronic devices. In *DesignCon*, volume 7, pages 1–6, 2017.

[116] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, 1976.

[117] G. Shakhnarovich, T. Darrell, and P. Indyk. *Nearest-Neighbor Methods in Learning and Vision: Theory and Practice (Neural Information Processing)*. The MIT Press, 2006.

[118] W. Shen, G. Newsham, and B. Gunay. Leveraging existing occupancy-related data for optimal control of commercial office buildings: A review. *Advanced Engineering Informatics*, 33:230 – 242, 2017.

[119] S. S. Shetty, H. D. Chinh, M. Gupta, and S. K. Panda. User presence estimation in multi-occupancy rooms using plug-load meters and pir sensors. In *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pages 1–6, 2017.

[120] A. Shokry, M. Elhamshary, and M. Youssef. The tale of two localization technologies: Enabling accurate low-overhead wifi-based localization for low-end phones. In *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '17, New York, NY, USA, 2017. Association for Computing Machinery.

[121] F. J. Simanjuntak. Deep learning approach for electric appliances recognition. `http://ugm.id/fsimanjuntak`, 2017.

[122] J. F. Trevor Hastie, Robert Tibshirani. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2nd edition, 2009.

[123] A. S. Uttama Nambi, A. Reyes Lua, and V. R. Prasad. Loced: Location-aware energy disaggregation framework. Technical report, 2015.

[124] A. S. Uttama Nambi, A. Reyes Lua, and V. R. Prasad. Loced: Location-aware energy disaggregation framework. In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, BuildSys '15, page 45–54, New York, NY, USA, 2015. Association for Computing Machinery.

[125] T. Vaupel, J. Seitz, F. Kiefer, S. Haimerl, and J. Thielecke. Wi-fi positioning: System considerations and device calibration. In *2010 International Conference on Indoor Positioning and Indoor Navigation*, pages 1–7, 2010.

[126] E. Viciana, A. Alcayde, F. G. Montoya, R. Baños, F. M. Arrabal-Campos, A. Zapata-Sierra, and F. Manzano-Agugliaro. Openzmeter: An efficient low-cost energy smart meter and power quality analyzer. *Sustainability*, 10(11), 2018.

[127] S. Vigneshwaran, S. Sen, A. Misra, S. Chakraborti, and R. K. Balan. Using infrastructure-provided context filters for efficient fine-grained activity sensing. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 87–94, 2015.

[128] R. Walawalker, V. Iyer, and M. Murthy. Utility interface and power quality: The flip side. In *Frontiers of Power Conf.*, pages VII–VII. Engineering Energy Laboratory, Oklahoma State University; 1998, 2002.

[129] K. Weekly, D. Rim, L. Zhang, A. M. Bayen, W. W. Nazaroff, and C. J. Spanos. Low-cost coarse airborne particulate matter sensing for indoor occupancy detection. In *2013 IEEE International Conference on Automation Science and Engineering (CASE)*, pages 32–37, 2013.

[130] M. Weiss, A. Helfenstein, F. Mattern, and T. Staake. Leveraging smart meter data to recognize home appliances. In *2012 IEEE International Conference on Pervasive Computing and Communications*, pages 190–197, 2012.

[131] M. Woźniak and D. Połap. Intelligent home systems for ubiquitous user support by using neural networks and rule-based approach. *IEEE Transactions on Industrial Informatics*, 16(4):2651–2658, 2020.

[132] H. Wu. *Sensor data fusion for context-aware computing using dempster-shafer theory*. PhD thesis, Carnegie Mellon University, the Robotics Institute, 2003.

[133] Z. Xing, J. Pei, and E. Keogh. A brief survey on sequence classification. *SIGKDD Explor. Newsl.*, 12(1):40–48, 2010.

[134] C. Xu, S. Li, G. Liu, Y. Zhang, E. Miluzzo, Y.-F. Chen, J. Li, and B. Firner. Crowd++: Unsupervised speaker count with smartphones. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '13, pages 43–52, New York, NY, USA, 2013. ACM.

[135] L.-W. Yeh, Y.-C. Wang, and Y.-C. Tseng. ipower: an energy conservation system for intelligent buildings by wireless sensor networks. *International Journal of Sensor Networks*, 5(1):1–10, 2009.

[136] L. Yu, H. Li, X. Feng, and J. Duan. Nonintrusive appliance load monitoring for smart homes: recent advances and future issues. *IEEE Instrumentation Measurement Magazine*, 19(3):56–62, 2016.

[137] A. A. Zaidi, F. Kupzog, T. Zia, and P. Palensky. Load recognition for automated demand response in microgrids. In *IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society*, pages 2442–2447, 2010.

[138] M. Zeifman and K. Roth. Nonintrusive appliance load monitoring: Review and outlook. *IEEE Transactions on Consumer Electronics*, 57(1):76–84, 2011.

[139] J. Zhao, B. Lasternas, K. P. Lam, R. Yun, and V. Loftness. Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining. *Energy and Buildings*, 82:341 – 355, 2014.

[140] Y. Zhao, W. Zeiler, G. Boxem, and T. Labeodan. Virtual occupancy sensors for real-time occupancy information in buildings. *Building and Environment*, 93:9 – 20, 2015.

[141] Z. Zhen, Q. Jia, C. Song, and X. Guan. An indoor localization algorithm for lighting control using rfid. In *2008 IEEE Energy 2030 Conference*, pages 1–6, 2008.

[142] A. Zoha, A. Gluhak, M. Nati, and M. A. Imran. Low-power appliance monitoring using factorial hidden markov models. In *2013 IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pages 527–532, 2013.

[143] H. Zou, H. Jiang, Y. Luo, J. Zhu, X. Lu, and L. Xie. Bluedetect: An ibeacon-enabled scheme for accurate and energy-efficient indoor-outdoor detection and seamless location-based service. *Sensors*, 16(2), 2016.

[144] H. Zou, Y. Zhou, H. Jiang, S.-C. Chien, L. Xie, and C. J. Spanos. Winlight: A wifi-based occupancy-driven lighting control system for smart building. *Energy and Buildings*, 158:924 – 938, 2018.