Categorization Framework and Survey of Occupancy Sensing Systems

Kjærgaard, Mikkel Baun; Sangogboye, Fisayo Caleb

Published in:
Pervasive and Mobile Computing

DOI:
10.1016/j.pmcj.2016.09.019

Publication date:
2017

Document version
Accepted manuscript

Citation for pulished version (APA):
Kjærgaard, M. B., & Sangogboye, F. C. (2017). Categorization Framework and Survey of Occupancy Sensing Systems. *Pervasive and Mobile Computing*, *38*(Part 1), 1-13. https://doi.org/10.1016/j.pmcj.2016.09.019

Categorization Framework and Survey of Occupancy Sensing Systems

Mikkel Baun Kjærgaard, Fisayo Caleb Sangogboye

{mbkj,fsan}@mmmi.sdu.dk, Center for Energy Informatics, Mærsk McKinney Møller Institute, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

Abstract

A large share of the energy usage in buildings is driven by occupancy behavior. To minimize this usage, it is important to gather accurate information about occupants' behaviour and to improve sensing systems for gathering such information. However, as research on occupancy sensing systems goes beyond *basic methods* with an increasing diversification, there is a clear need to enable adequate comparison of these systems and their properties. The systems which differ in methods and properties also lack a categorization framework for classifying different options. This article proposes a categorization framework constructed from analyzing and comparing existing sensing systems to address these needs. The classification framework is being constructed from a literature survey of 51 papers and articles presenting 46 different occupancy sensing systems. It is intended that this framework can enable developers to better benchmark and evaluate sensing system, enable organizations to identify trade-offs for adopting sensing systems and aid researchers in scoping out future research in the area.

Keywords: Occupancy Sensing Systems, Survey, Categorization Framework

1. Introduction

Improving the energy performance of buildings is crucial towards realizing a more sustainable society. One important challenge for improving the energy performance is the impact of occupancy behavior [1]. Occupancy behavior here refers to *all actions of the occupant (including presence) that affect building energy consumption* [2]. Occupancy behavior hugely influences the energy performance of both individual appliances and building-wide infrastructures. Individual appliances may include IT devices, kitchen facilities and production equipment. While building-wide infrastructures usually include lighting, heating, ventilation, cooling, IT, fire protection, security and water. Three agendas have been established towards addressing the impact of occupancy behavior: A) replace equipment and infrastructures in buildings with more efficient ones in a manner that the same occupancy behavior results in a lower energy consumption. B) engage occupants in changing their behaviors to less energy-consuming behaviors. C) improve the intelligence of infrastructures and equipment to only provide needed utilities and comfort for the actual behavior of occupants. A typical example is to control ventilation with respect to estimated or measured occupancy rates [3]. In all three cases it is important to be able to gather quantitative information about occupant's behavior as follows: in case A) to document savings in relation to occupancy behavior; in case B) to provide feedback to support behavior change; and in case C) to use occupancy behavior to optimize control.

To gather such quantitative occupancy information a wide range of occupancy sensing systems has been proposed in research and commercialized. In this article we refer to occupancy behavior sensing systems as *systems that measure, estimate, model and predict occupancy behavior based on inputs from pervasive sensing infrastructures*. Examples include systems for presence detection using PIR sensors, visual, stereo and thermal camera-based systems for counting people, and systems based on sensor-instrumented spaces to recognize the activities of individuals.

When surveying occupancy sensing systems for comparison, one has to answer many different questions. How do systems differ in types of occupancy information provided? What is the relation between the system and occupants? What is the spatial and temporal coverage; does the system allow for prediction of future occupancy situations? What types of sensor strategies are applied; are environments, objects or persons augmented? What types of methods and models are utilized? What is the resulting accuracy? These questions are important both for customers, developers and researchers who need to understand different design options and trade offs. It is believed that a categorization framework that takes cognisance of these questions can aid customers, developers and researchers to better survey, compare, and design occupancy sensing systems. This is especially important as developments in the field transit from understanding the basic mechanisms to combining different sensor strategies and modalities for providing information on complex behavioral patterns of occupants. The categorization framework can also aid researchers in scoping out future research in the area of occupancy sensing systems.

Existing surveys on occupancy sensing systems [4, 1] have so far not presented a comprehensive categorization framework for the area. Therefore, this article proposes a comprehensive categorization framework for occupancy sensing systems that is constructed based on a literature study of 51 papers and articles. The 51 papers and articles propose 46 different systems which are analyzed and grouped according to the methods and techniques utilized to form categories for this framework. The classifications of four systems are presented in detail as examples. The classifications of all the 46 systems are available from our online repository [5] and presented on the associated webpage[1]. We present an analysis of the classifications and duly highlight evaluation metrics and unexplored design options. In this article we motivate the use of occupancy sensing systems in the area of energy performance, however, there are many other application areas for such systems including safety and evacuation, building utilization and customer profiling.

2. Categorization Framework

The proposed categorization framework formulates nine categories. These categories are partly inspired by earlier works on occupancy behavior that are derived from the conducted literature study. This study was conducted by searching for key terms in relevant journals and conferences. The identified categories include and are defined as follows:

Information Type: types of occupancy information.

Occupant Relation: relation between the system and occupants.

Sensing Strategy: strategy for placement of sensors for observing occupancy behavior.

Spatial Granularity: characterization of the spatial resolution.

Temporal Granularity: characterization of the temporal resolution.

Spatial Coverage: characterization of the spatial extent.

Temporal Coverage: characterization of the temporal extent.

[1] https://github.com/mbkj/OccupancySurvey/wiki

Sensor Modality: sensor modalities for collecting data about occupancy behavior.

Methods and Models: methods and models for processing sensing data to estimate and predict occupancy information.

Earlier surveys have only considered a subset of these categories. Christensen et al. [4] discuss the three dimensions *occupancy resolution, temporal resolution and spatial resolution* mapping to the dimensions proposed in this article of *information type, spatial granularity and temporal granularity*, respectively. Nguyen et al. [1] introduce the dimensions *activities, technologies and methodologies* that maps to the dimensions proposed in this article of *information type combined with temporal coverage, sensing modality and methods and models*, respectively. Since the focus of the proposed categorization is on sensing systems, this categorization does not include design phase building construction and retrofit tools that provide occupancy modeling and simulation functionality as surveyed by Hoes et al. [6].

3. Categories of the Framework

This section presents the framework categories one by one.

Information Type. This category covers how sensing systems represent the behavior of occupants, as illustrated in Figure 1. The information type determines to a great extent what applications can be enabled by the collected information. *Presence* denotes occupation of space in a zone of interest by a human, for example, space in front of a computer [7] or rooms in a building [8] and, represented as a Boolean value [8], as a level [9] or as a count of humans [4]. *Activity* covers instant actions of humans including actions on objects in an environment, for example, typing on a keyboard [8], door passing [10], turning equipment on/off [11] or opening and closing windows and doors [2]. *Behavior* covers the longer running human processes made up by several activities. For example, working, having a meeting [8], or taking a break [12].

Occupant Relation. This category denotes the way in which occupants are represented by the sensing system as listed in Figure 1. The occupant relation has a great influence on needed privacy protection measures for the data. *Anonymous* denotes that the system does not represent the identity of an observed occupant. For example, by only detecting the presence of an

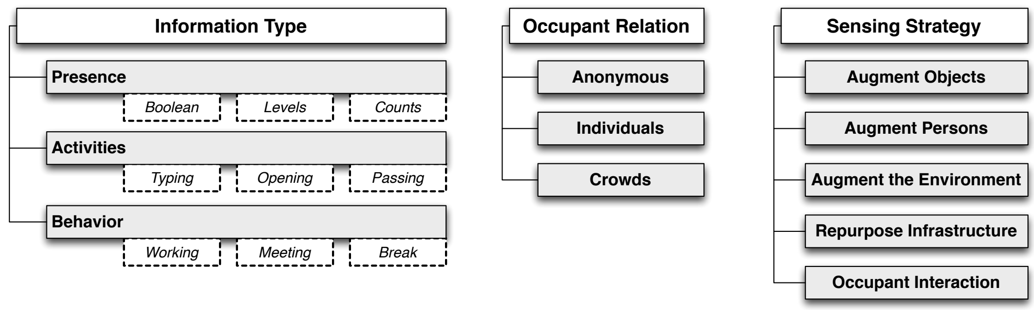

Figure 1: Information Type, Occupant Relation and Sensing Strategy

occupant [13]. However, in some situations anonymous data might be reidentified, e.g., anonymous data from a single office can with high probability be reassociated with the office owner. *Individuals* refers to the situation when the system knows the identities of the monitored humans. For example, when tracking the proximity of an employee to his work desk [4] or the time for an occupant to reach his house [14]. *Crowd* implies considering people as a crowd of humans rather than as individuals. For example to detect the behavior of crowds, such as, flocking [15]. This implies that data is aggregated making it difficult to reidentify data back to individuals.

Sensing Strategies. This category covers how sensors are applied to collect measurements about occupant behavior as listed in Figure 1. The choice of sensing strategy has a large influence on the infrastructure requirements and installation complexity. We define the following strategies:

The *augment objects* strategy proposes to observe occupant behavior by embedding sensors in the objects that occupants use in their everyday life. For instance, kitchen equipment, office equipment or other objects that people use in their everyday life [7]. As an example, with sensors in a cup you can detect if the it contains liquid or if it is tilted to be drunken from. The advantages of this strategy are that no assumptions are placed on occupants and detailed information can be collected about the use of objects. Drawbacks include that it can be complex to embed sensors in objects and that if a high number of objects have to be instrumented deployments can become costly.

The *augment persons* strategy proposes to augment persons with wearable or mobile sensor systems that enable the gathering of information about occupancy behavior [15]. The advantages of this strategy are that it enables the collection of detailed data about the physical movement of occupants and the ability to collect longitudinal data for individual occupants. Drawbacks include that occupants have to remember to carry or wear the systems and that the strategy can be costly when instrumenting a large number of occupants.

The *augment the environment* strategy proposes to place sensors in the environment of the occupants to observe their behavior [16]. The environment could here be a building or room with sensors installed to observe occupants' behaviors. The advantages of this strategy are that no assumptions are placed on occupants and detailed information can be collected about space use. Drawbacks include that placement and wiring of sensors can be difficult and that deployments can be costly when covering large areas.

The *repurpose infrastructure* strategy follows the idea to repurpose existing infrastructures of buildings to provide information about occupancy behavior. For instance, WiFi networks [17], security access cards [18] and calendar information [12]. The advantages of this strategy are that no additional assumptions are placed on occupants and deployment cost can be low due to the reuse of existing infrastructure. Drawbacks include a risk of low accuracy if the mapping from gathered data to occupancy information is complex and that robustness might be low due to the dependency on other systems not designed or maintained for high accuracy occupancy sensing.

The *occupant interaction* strategy activates occupants to provide clues about thier behavior. Occupants might be activated to provide clues about how long they will occupy a room via room controls or calendars. Another example is to use RFID [19] or other sensor modalitites to enable occupants to notify the building of their presence [20]. The strategy can be seen as an example of crowd sourcing the collection of information about occupant behavior. The difference compared to *repurpose infrastructure* is that technologies for the occupant interaction strategy are put in place to directly enable the interaction with occupants. The advantages of the strategy include the ability to collect information that sensors only can unreliable monitor and limited cost for sensor deployments. The drawbacks include that occupants have to remember to provide information and that occupants might intentionally or unintentionally misinform the system.

Spatial Granularity. This category concerns the characterization of the spatial resolution of occupancy information as illustrated in Figure 2. To characterize the spatial resolution we propose to use an already established terminology in the built environment in form of the Industry Foundation Classes (IFC) data model for buildings [21]. Following the IFC data model a *Site* is a physical area, a *Building* is a physical structure placed on a site, a *BuildingStory* is a single story of a building, a *Space* is a subpart of a story which might correspond to several rooms in the same Heating, Ventilation and Air Conditioning (HVAC) zone, a room or a subpart of a room. An *Object* is a physical element placed in a space, e.g., a computer or an appliance. A common spatial granularity considered by occupancy sensing systems are spaces in the form of rooms [13] or HVAC actuation zones [17]. Building granularity is most often considered in the case of residential homes [19] and object granularity in terms of occupant presence at individual equipment, such as, IT equipment [7]. The available spatial granularity has a large impact on what types of applications can be built, e.g., building-wide energy performance benchmarking versus room-based occupancy triggered ventilation. Aiming at a finer granularity will generally increase cost due to the deployment of a more fine grained sensing infrastructure.

Temporal Granularity. This category concerns the characterization of the temporal resolution of occupancy information as illustrated in Figure 2. *Periodic* denotes that the sampling and processing of occupancy information is executed at regular periodic intervals. Tarzia et al. [7] periodically schedule their sonar system for detection of human presence. Another option is the *event-based* scheme where an event occur when new occupancy information is available. For instance, Agarwal et. al. [13] schedule detection around door opening and closing events. A periodic sampling or processing of occupancy information places a higher demand on the sensing system in terms of network load, the volume of storage required for data storage and the processing capability required to process higher volumes of data. This bottleneck may question the sustainability of the sensing system itself. With event-driven systems a drawback is that it is more difficult to observe system failures as data is only received intermittently.

Spatial Coverage. This category covers a characterization of the spatial coverage as illustrated in Figure 3. This category also follows the IFC data model which categorize coverage into *Site*, *Building*, *BuildingStory*, *Space* and *Object*. The coverage of an occupancy sensing system is often defined in multiples of the spatial granularity or aggregates of it, e.g., a building story is an aggregate of all spaces within that story. Another option is that

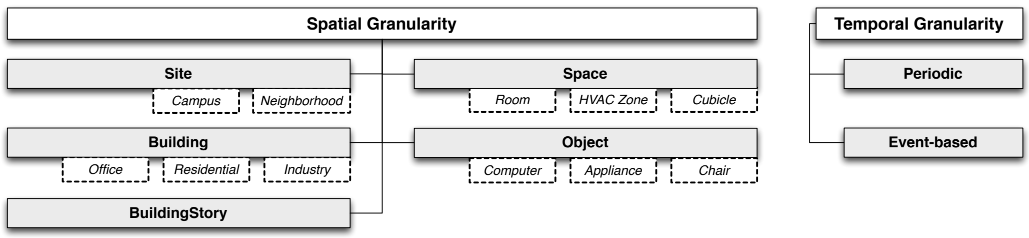

Figure 2: Spatial and Temporal Granularity

a system can extrapolate from monitoring a few spaces on a building story to all spaces of a building story. Systems often aim for a spatial coverage of a whole building, however, reported studies often conduct lab deployments that only cover a smaller set of spaces within a building [13, 7]. The coverage of systems based on world-wide tracking is extended to all relevant sites, e.g., Koehler et al. [14] use GPS tracking to predict the earliest point in time when an occupant can reach their home. The possible coverage of a occupancy sensing system mainly depends on cost. As mentioned earlier different sensing strategies provide cost effective scalability in different dimensions, e.g., number of occupants vs. coverage area. An additional factor is privacy restrictions as it might not be socially acceptable or legally allowed to monitor all areas of a building.

Temporal Coverage. This category covers a characterization of the temporal coverage as illustrated in Figure 3. *Past* denotes that the system can provide occupancy information about the past this can either be in the form of information for a particular point in time or aggregated models characterizing normal behavior [22]. *Present* denotes that the system can provide occupancy information for the current point in time [7]. Future denotes that the system is able to predict occupancy information for future points in time based on some form of method or model [19]. Systems delivering information about the *Past* or the *Future* will generally be more demanding in terms of computing power and data storage.

Sensor Modality. This category covers sensor modalities for collecting occupancy behavior data. An abundance of different sensor modalities allows for observing various aspects of occupancy behavior. Some of the sensor technologies were originally envisioned for occupancy sensing whereas others originate from efforts on related research problems. Figure 4 provides an

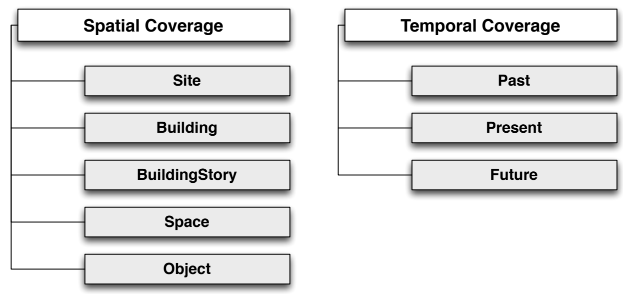

Figure 3: Spatial and Temporal Coverage

overview of the identified sensor modalities. Many systems apply or integrate several types of sensor modalities. It is beyond the scope of this article to explain the principles behind all the various sensor modalities, thus readers are for such information encouraged to survey material on the individual topics. The sensor modalities also differ in how they impact occupants' privacy as surveyed by Christin et al. [23]. To give an overview Figure 4 lists main advantages and disadvantages for each sensor modality. In our survey we identified the following sensor modalities: *Occupant data* includes data provided by the occupants themselves including social networking [12], calendar information [9] and computer networks [24]. *Force* covers the measuring of forces applied to objects and environments including pressure sensors [8], switches [25], gyroscopes and accelerometers [26] and device inputs, such as, a keyboard or a mouse [4]. *Visible light* covers the capturing of visual light including the use of video cameras [27, 28], light level sensors [11] and stereo cameras [29, 30]. *Infrared light* covers the capturing of infrared light including PIR sensors [13, 22, 31] and thermal cameras [32]. *Sound* covers the capturing of hearable sound by the use of microphones [8] and transmission setups to estimate proximity [33]. *Ultrasound* covers the use of ultrasound including sonar setups [7, 34] and transmission setups to estimate proximity [35]. *EM waves* covers radar [36] and radio-based communication setups including RFID, WiFi, Bluetooth, UWB and GPS technologies [17, 37, 38, 14, 19, 39, 4]. *Air* covers the measurement of the composition and properties of air including CO_2 [16, 40], humidity [16] and temperature [40]. *Magnetic fields* covers the measurement of magnetic fields including reed switches [13], compasses [26], access cards [18] and magnetic coils [20]. *Electricity* covers the measurement of electricity consumption including various metering setups [41]. *HVAC data* covers data from the operation of

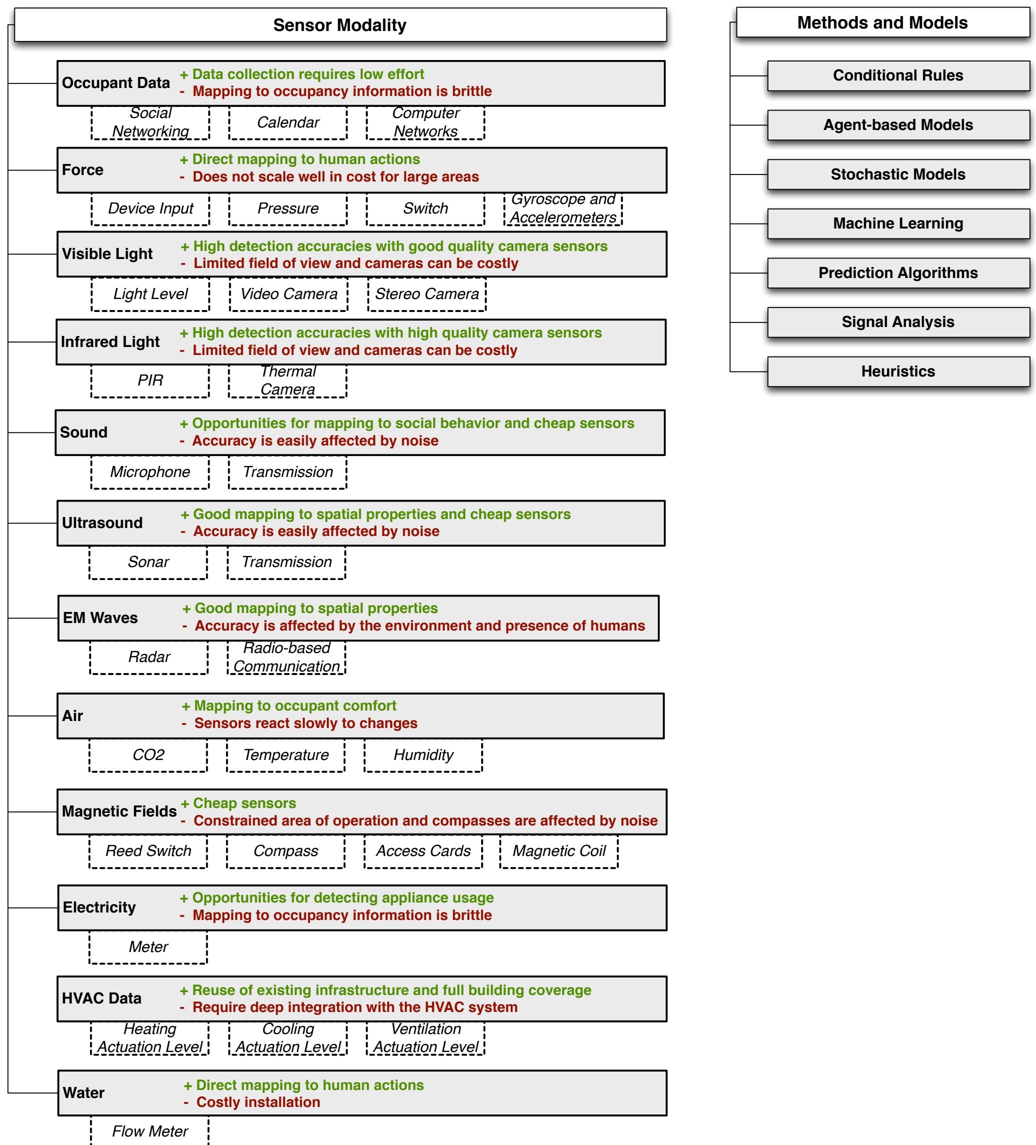

Figure 4: Sensor Modalities and Methods and Models.

HVAC systems including heating, ventilation and cooling activity levels [40], *Water* includes measurements of water consumption including flow meters [11].

Methods and Models. This category covers the methods or models that occupancy sensing systems apply to process and fuse occupant behavior data.

The goal can be to map raw sensor data to an occupancy information type, optimize accuracy by modeling spatial or temporal relations, estimate long term behavior or predict future occupancy behavior. The area of methods and models for data processing has a long history. For an introduction to the area we will refer the reader to other work. In this article we will only classify methods and models used by occupancy sensing systems into seven broad categories as illustrated in Figure 4. *Conditional rules* model the relationship between sensor input and occupancy information as conditional rules. For instance, a door opening event may indicate that the occupancy of a room changes [13]. *Agent-based models* model occupants as agents whose behaviors are defined by modeling their itinerary, path choices, and walking behavior [27]. *Stochastic models* model the probability and correlation among occupancy behavioral events and the likelihood of changes in occupants' presence [28]. *Machine learning* learns and models occupants' behavior from training data by learning the mappings between sensor inputs and occupancy levels [16]. *Prediction algorithms* enable the prediction of future states, e.g. prediction algorithms can from GPS data predict the earliest point in time an occupant can arrive back home [14]. *Signal analysis* covers signal analysis methods which includes methods for signal decomposition [40] and image processing [29]. *Heuristics* cover a broad range of simple algorithmic steps that does not fall under any of the other categories. For instance, simple thresholding on values [7].

The advantages of the strategies *Heuristics* and *Conditional rules* are that they are relatively simple to apply, however, the methods do not scale well to complex occupancy behaviors and sensing modalities with high uncertainty. *Agent-based models* and *Stochastic models* allow the modeling of more complex behaviors however, they require significant modeling efforts to arrive at accurate models. *Machine learning*, *Prediction algorithms* and *Signal analysis* are mainly data-driven techniques and their accuracy depends on the availability of good data for model construction and parameter estimation.

4. Classification Examples

To illustrate the use of the proposed categorization framework, this article presents the categorization results for four of the 46 systems covered by the literature study. The four systems have been chosen to be representative for different types of occupancy sensing systems. In the following sections we will summarize the categorizations of all the 46 systems which are also available from our online repository [5]. Table 1 lists the categorizations in a compact form for the four systems. In addition to the presented nine different categories, the table also lists infrastructure requirements, installation complexity and reported accuracy to enable readers to judge the applicability of the different systems.

Agarwal et al. [13] propose a system for presence detection of anonymous people which has a spatial granularity of spaces corresponding to rooms and event-based temporal granularity. The system covers spaces considered as individual rooms and delivers information about the present. Also, the system applies *Infrared Light-PIR* and *Magnetic Fields-Reed Switch* to augment the environment and it uses conditional rules to model the relationship between sensor input and occupancy information.

Erickson et al. [28] propose the OBSERVE system for providing occupancy counts of anonymous people. The system has a spatial granularity of spaces which is divided by sensor monitoring areas and the temporal granularity is based on periodic sampling. The system covers spaces with occupant's entries monitored by sensors and it delivers information about present and predicts future occupancy. The system applies *Visible Light-Video Camera* for occupancy counting by augmenting the environment. Also, the system applies stochastic models in the form of Markov models to predict future occupancy.

Tarzia et al. [7] propose a system for detecting the activity of known individuals. The system has a spatial granularity of the space in front of a computer and the temporal granularity is periodic. The system can cover all spaces in front of computers and can provide information about the present. The system applies *Ultrasound-Sonar* and *Force-Device Input* by augmenting the objects and it applies a few heuristics to optimize accuracy and combine sensor inputs.

Christensen et al. [4] propose a system for providing counts of anonymous persons. The system considers the spatial granularity of buildings and the temporal granularity is given by periodic sampling. The system has a spatial coverage for building and considers a temporal coverage of present. The system applies *EM Waves-Radio-based Communication (WiFi)* with a repurpose infrastructure sensing strategy. Just like the previously highlighted system, this system does not apply any structured model however, a few heuristics is utilized to optimize accuracy and combine sensor inputs.

The classification examples highlight different system options. For instance, given the goal to provide counts of occupants in a building one can compare the two options proposed by Erickson et al. [28] and Christensen et al. [4]. The two systems provide the same type of information but with different sensing strategies, sensing modalities, methods and models, and resulting accuracy. The two systems also have different implications on privacy as the system by Erickson et al. [28] involve cameras whereas the system by Christensen et al. [4] monitor the occupants' wireless devices. In regards to accuracy the former system can provide the highest accuracy but is more costly to scale especially if a high spatial granularity is required because the system need a camera per space. The later system by Christensen et al. [4] scales very well as the system repurposes an existing infrastructure but can only monitor occupants carrying wireless devices.

To utilize the proposed categorization framework for categorizing and comparing new systems, the following steps may be adopted: firstly, find classifications for compared-to existing systems in our repository [5]. Secondly, perform a classification for the new system by classifying for each of the nine categories the new system's methods and assumptions according to the subcategories. Thirdly, perform a comparison of the existing and the new systems based of the nine categories. Push the new classification to our online repository [5].

5. Analysis of Classified Systems

This section analyses the classifications of the 46 systems to characterize existing work. The following method was used for the analysis. Firstly, each system was classified according to the classification framework. Secondly, the classifications were encoded as strings in the JSON data format. Finally, we developed python scripts that analyse the classifications and produce tables and graphs summarizing the classifications. The system classifications and scripts are available in our online repository [5].

5.1. Summarizing Classifications

A first outcome of the analysis is statistics computed for the categories. The article includes statistics for the categories sensor modality and methods and models. Figure 5 shows statistics for how many systems use each of the sensor modalities. The following observations can be made from this figure:

[2]Please note that as some systems are covered in several papers there are more papers listed per category than systems in the pie chart.

	Agarwal et al. [13]	**Erickson et al. [28]**	**Tarzia et al. [7]**	**Christensen et al. [4]**
Information Type	Presence-Boolean	Presence-Counts	Activity	Presence-Counts
Occupant Relation	Anonymous	Anonymous	Individuals	Anonymous
Spatial Granularity	Space (Room)	Space (Sensor view)	Space (In front of computer)	Buildings
Temporal Granularity	Event-based	Periodic	Periodic	Periodic
Spatial Coverage	Space (Room)	Space (Room)	Space (In front of computer)	Buildings
Temporal Coverage	Present	Present, Future	Present	Present
Sensor Modality	Infrared Light-PIR, Magnetic Fields-Reed Switch	Visible Light-Video Camera	Ultrasound-Sonar, Force-Device Input	EM Waves-Radio-based Communication (WiFi)
Sensing Strategy	Augment the environment	Augment the environment	Augment the objects	Repurpose infrastructure
Methods and Models	Conditional Rules	Stochastic Models (Markov Model)	Heuristics	Heuristics
Infrastructure Requirements	Wireless sensor nodes and base stations	Wireless sensor nodes and base stations	Repurpose existing computer speakers and microphones	Repurpose existing wireless infrastructure
Installation Complexity	Install infrastructure and configuration equipment	Install infrastructure and configuration equipment	Install software	Install software
Accuracy	Deployed in a subset of rooms on a single building story. No quantified accuracy. Simulation shows that the detected occupancy information can help reduce energy consumption in the building with 10-15%	Deployed in ten rooms on a single building story. The system shows good accuracy results (close to zero) for prediction quantified using the Jensen-Shannon divergence. Simulation shows that the model-based prediction can help save 42%	Deployed in a single room with data collected for twenty participants. The system provides good accuracies for separating if an occupant is absent (96%) or passively engaged (98%)	Deploy components in different setups in two buildings. Reports on the accuracy of the different sensing options, e.g., accuracy for PIR (91%), PC-activity (91%) and DCHP to detect presence and accuracy for using WiFi to detect arriving (75%) or leaving (90%) a cubicle.

Table 1: Case studies of occupancy sensing systems

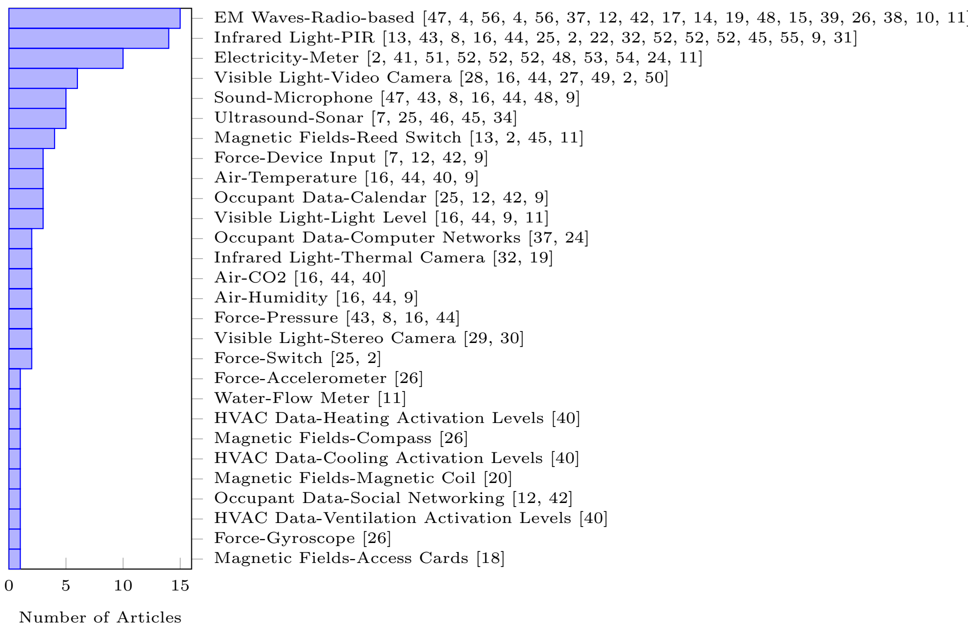

Figure 5: Summarizing the categorizations for the dimension sensor modalities[2].

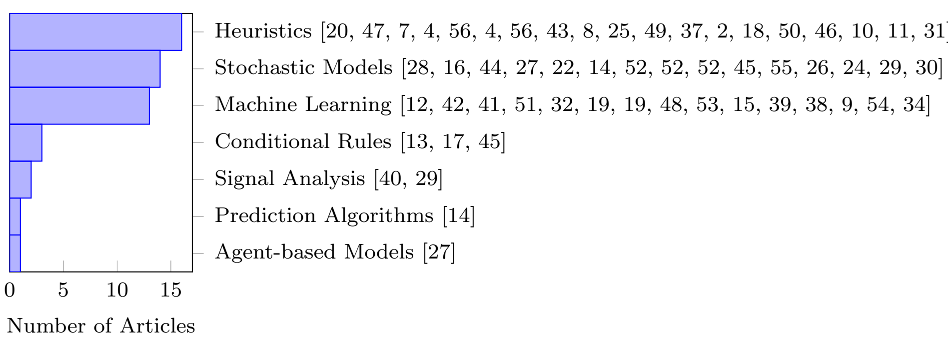

Figure 6: Summarizing the categorizations for the dimension methods and models[2].

i) *EM Waves-Radio-based Communication* is the most frequently used sensor modality (15 cases) followed by *Infrared Light-PIR* (14 cases) and *Electricity-Meter* (10 cases); ii) The used sensor modalities span a very broad range of technologies and physical measurement types; iii) nine out of the 46 systems combine several sensor modalities. In a similar manner Figure 6 presents the statistics for how many systems use the different methods and models. From the figure it can be observed that most systems use heuristics (16 cases) followed by stochastic models (14 cases) and machine learning (13 cases).

5.2. Accuracy

A second analysis compares occupancy sensing systems in terms of information type and accuracy. Accuracy of occupancy information is a central evaluation criteria for occupancy sensing systems. The four classification examples illustrated that systems are often evaluated using different accuracy metrics. In particular different metrics are applied for different kinds of occupancy information. In our survey we collected information on the metrics used in the evaluation for each of the surveyed systems and the resulting accuracy. For most systems we report an average accuracy claimed for the system in question. Average is here over different conditions or trials. For other systems we report a range if the system under different conditions exhibited different performance levels, e.g., depending on amount of occupants, building type and internal system parameters. As the systems were evaluated under different conditions it is error prone to do a very fine-grained analysis of the results. However, the listed results give some indications to which systems provide the best accuracies.

Table 2 lists the classification results for the different accuracy metrics and occupancy information types. For the articles that quantify the performance using several metrics we list the most common metric to enable comparison.

Most articles view their problem as a classification problem. The number of classes differs based on the information type computed as highlighted in the table with *(C: X)* where X is the number of classes. If $X = N$ it means that a system has an open set of classes. For classification problems, accuracy is defined as the number of correctly classified instances. For classification problems, an extensive literature exists on different accuracy metrics that are more or less robust to class-imbalance in the test data. Robust accuracy metrics include precision and recall, and the F-measure. Some papers also use an accuracy ratio as accuracy metric. Numeric occupancy information, such as, occupant counts can be viewed as regression problems and the accuracy quantified using associated accuracy metrics. Accuracy metrics include average error as the average difference between ground truth and the produced estimates and the Root Mean Squared Error (RMSE) calculated as the root of the average squared difference.

From Table 2 we can make several observations: i) within each category there is a huge span in reported accuracies both among the individual systems and for the systems that study the effect of different conditions and parameters; ii) For presence-count, authors have approached the problem both as a classification problem and a regression problem; iii) it is a concern that most

	Presence-Boolean	Presence-Count	Presence-Track	Activities	Behavior
Accuracy	*75%-90%* [4, 56] *78%* [19] *80%* [54] *82%* [52] *84%* [19] *86%* [17] *87%* [53] *88%* [12, 42] *92%* [14] *98%* [10]	*64%-93%* [49] *65-96* [9] *66%* [52] *82%* [16, 44] *83%* [29] *84%* [39] *84%* [50] *88%* [40] *90%* [45]		*87%* (C:N) [11] *98%* [10]	*75%* (C:3) [46] *92%* (C:4) [41, 51] *95%* (C:2) [52] *95%* (C:5) [43, 8] *96%* (C:2) [7]
Precision and Recall				*75%-78%* (C:N) [48] *91%-94%* (C:N) [47]	
F-measure					*0.87* (C:N) [15]
Accuracy Ratio		*0.4-1.1* [4, 56]			
Error		*2.2* [34]	*2.5-8.5 m* [26]		
RMSE		*0.35* [32] *0.78* [30] *5.7* [27] *21.7* [31]			
Indicative Graphs	- [13] - [2] - [22] - [20] - [55] - [37] - [24] - [25]	- [2] - [37] - [18] - [38] - [25]	- [37]	- [2]	- [38]

Table 2: Evaluation results for different accuracy metrics and occupancy information types. For classification problems the number of classes is given by *(C:X)* where X is the number of classes. If $X = N$ it means that a system has an open set of classes.

systems are evaluated with accuracy as metric because this measure is prone to class-imbalances, e.g., a system can easily achieve a high accuracy if there is seldom any occupants and the system is good at detecting non-presence. Therefore we encourage authors in future work to apply precision, recall and the F-measure as more robust metrics.

5.3. Comparisons Among Categories

A third type of analysis considers the classification of systems with regards to several categories. This type of analysis can highlight if systems only explore a subset of the possible design choices, e.g., do systems consider all combinations or just a few.

To analyze *sensing strategies* in comparison to *occupant relation*, Table 3 categorizes the surveyed occupancy sensing systems with regards to these two dimensions. The following observations can be made from Table 3: i)

	Anonymous	Individuals	Crowds
Augment the Environment	[13, 32, 2, 22, 55, 28, 27, 34, 45, 53, 50, 29, 9, 52, 52, 52, 30, 19, 49, 16, 44, 25, 41, 51]	[43, 8, 48, 25, 41, 51]	
Augment Persons		[17, 4, 56, 39, 37, 14, 11, 10, 48, 26]	[15]
Augment Objects	[46, 9]	[12, 42, 43, 8, 11, 7]	
Repurpose Infrastructure	[4, 56, 40, 18, 9, 31, 54, 38, 24]	[17, 4, 56, 12, 42, 37, 47]	
Occupant Interaction		[20, 19]	

Table 3: Categorization of sensing strategy versus occupant relation.

	Presence	Activities	Behavior
Past	[22, 38, 24]		[38]
Present	[13, 17, 32, 2, 2, 22, 20, 4, 56, 4, 56, 39, 55, 40, 28, 12, 42, 37, 18, 34, 45, 53, 50, 29, 9, 31, 54, 14, 52, 52, 43, 8, 10, 38, 30, 24, 19, 19, 49, 26, 16, 44, 25]	[2, 11, 10, 48, 47]	[46, 15, 52, 43, 8, 38, 7, 41, 51]
Future	[22, 55, 28, 27, 14, 19, 19]		[41, 51]

Table 4: Categorization of occupancy information versus temporal coverage.

augmentations of the environment tend to provide an anonymous relation with people whereas systems augmenting persons tend to provide the systems with an individual relation; ii) augmenting objects as presented in 6 out of 8 cases, provide the systems with an individual relation as object ownership is known. The repurpose infrastructure strategy has resulted in systems of both categories; iii) so far most work have considered augmenting the environment (30 out of 46 cases) whereas only little work has considered the crowd relation (1 out of 46).

For comparing the system classifications for *occupancy information* in relation to *temporal coverage*, Table 4 lists classification results for these two dimensions. The following observations can be made from Table 4: i) most systems provide presence information for the present (34 of 46 cases); ii) more work focuses on behavior than activities (10 versus 4 cases); iii) less work focuses on aggregating past data and information or predicting the future in particular for occupants' activities and behaviors.

To analyze the categorizations for *spatial granularity* in relation to *temporal granularity*, Table 5 presents categorization results for these two dimensions. The following observations can be made from the table: i) most systems consider spatial granularity of space (42 of 46 cases) and they more often utilize the periodic temporal granularity than its event-based counterpart; ii) some systems consider building or object spatial granularity and in

	Site	Building	BuildingStory	Space	Object
Periodic		[2, 4, 56, 31, 54, 14, 30, 24]		[2, 22, 20, 4, 56, 39, 40, 28, 27, 12, 42, 37, 34, 45, 46, 53, 29, 9, 31, 15, 52, 52, 52, 43, 8, 30, 7, 47, 26, 16, 44, 25, 41, 51]	[11, 48]
Event-based		[2, 18, 19]		[13, 17, 32, 2, 55, 18, 50, 52, 52, 52, 10, 38, 19, 49]	

Table 5: Categorization of spatial granularity versus temporal granularity.

both cases they usually favor the periodic systems over event-based ones. iii) None of the classified systems has a building-story or site spatial granularity.

6. Future Work

This literature study identified several similarities and differences between the studied systems by clearly visualizing the design options and the choices behind identified systems to be understood by both experts and newbies in this field. The classification results highlight a number of unexplored design options that could be relevant for future works to consider: (i) Explore if an occupancy sensing system can augment persons in an anonymous manner maybe by considering persons as individuals of a crowd; (ii) Develop methods for occupancy sensing systems to predict future activities and behaviors; and (iii) How systems with a space granularity can be adapted to provide information at other requested spatial granularities.

For experimental evaluation of occupancy sensing systems the categorization framework can also be used to highlight the evaluated system's assumptions and methods. This can take the form of a classification for the nine categories that thereby explicitly categorize the evaluated system. In addition to design options, the surveyed systems also differ with regards to methods and metrics used for evaluation. For instance, evaluation methods include simulation, emulation and deployment [57] and evaluation metrics include error and accuracy metrics, such as, the percentage of correct classifications, percentage deviation and Jensen-Shannon divergence. Also, some of the systems are evaluated with regards to ground truth information whereas others are only evaluated within the context of energy efficiency applications. The variation in evaluation methods is a challenge for documenting improvements over existing methods. Therefore there is a need for more future work on comparative studies on public data sets. Due to privacy constraints associated with the data collection for occupant behavior, it might be a bottleneck to get allowance to share such datasets publicly to support comparative studies.

7. Conclusion

In this article we have proposed a categorization framework for occupancy sensing systems. The proposed categorization framework has been constructed from a literature survey of 51 papers and articles presenting 46 different occupancy sensing systems available from [5]. The categorization framework is based on the nine categories namely: *information type*, *occupant relation*, *sensing strategy*, *spatial granularity*, *temporal granularity*, *spatial coverage*, *temporal coverage*, *sensing modality*, and *methods and models*. It is believed that this categorization framework will aid developers to better benchmark and evaluate sensing system, enable consumers identify trade-offs for adopting these sensing systems and could also aid researcher in scoping out future research in the area of occupancy sensing systems.

Acknowledgment

This work is supported by the Innovation Fund Denmark for the project COORDICY (4106-00003B).

[1] T. A. Nguyen, M. Aiello, Energy intelligent buildings based on user activity: A survey, Energy and Buildings 56 (0) (2013) 244 – 257.

[2] A. Caucheteux, A. Es Sabar, V. Boucher, Occupancy measurement in building: A litterature review, application on an energy efficiency research demonstrated building, International Journal of Metrology and Quality Engineering 4 (2013) 135–144.

[3] T. Lawrence, J. Braun, Calibrated simulation for retrofit evaluation of demand-controlled ventilation in small commercial buildings, ASHRAE Trans (2007) 227–240.

[4] K. Christensen, R. Melfi, B. Nordman, B. Rosenblum, R. Viera, Using existing network infrastructure to estimate building occupancy and control plugged-in devices in user workspaces, Int. J. Commun. Netw. Distrib. Syst. 12 (1) (2014) 4–29.

[5] Classifications and wiki: https://github.com/mbkj/occupancysurvey.

[6] P. Hoes, J. Hensen, M. Loomans, B. de Vries, D. Bourgeois, User behavior in whole building simulation, Energy and Buildings 41 (3) (2009) 295 – 302.

[7] S. P. Tarzia, R. P. Dick, P. A. Dinda, G. Memik, Sonar-based measurement of user presence and attention, in: UbiComp '09, 2009, pp. 89–92.

[8] I. Georgievski, T. A. Nguyen, M. Aiello, Combining activity recognition and AI planning for energy-saving offices, in: IEEE UIC, 2013, pp. 238–245.

[9] A. Khan, J. Nicholson, S. Mellor, D. Jackson, K. Ladha, C. Ladha, J. Hand, J. Clarke, P. Olivier, T. Plötz, Occupancy monitoring using environmental & context sensors and a hierarchical analysis framework, in: ACM BuildSys 2014, 2014, pp. 90–99.

[10] J. Ranjan, Y. Yao, K. Whitehouse, An RF doormat for tracking people's room locations, in: ACM UbiComp '13, 2013, pp. 797–800.

[11] J. Ranjan, E. Griffiths, K. Whitehouse, Discerning electrical and water usage by individuals in homes, in: BuildSys '14, ACM, 2014, pp. 20–29.

[12] S. K. Ghai, L. V. Thanayankizil, D. P. Seetharam, D. Chakraborty, Occupancy detection in commercial buildings using opportunistic context sources, in: IEEE PerCom Workshops, 2012, pp. 463–466.

[13] Y. Agarwal, B. Balaji, R. E. Gupta, J. Lyles, M. Wei, T. Weng, Occupancy-driven energy management for smart building automation, in: BuildSys'10, 2010, pp. 1–6.

[14] C. Koehler, B. D. Ziebart, J. Mankoff, A. K. Dey, Therml: occupancy prediction for thermostat control, in: UbiComp '13, 2013, pp. 103–112.

[15] M. B. Kjærgaard, M. Wirz, D. Roggen, G. Tröster, Detecting pedestrian flocks by fusion of multi-modal sensors in mobile phones, in: Ubicomp '12, 2012, pp. 240–249.

[16] B. Dong, K. P. Lam, Building energy and comfort management through occupant behaviour pattern detection based on a large-scale environmental sensor network, Journal of Building Performance Simulation 4 (4) (2011) 359–369.

[17] B. Balaji, J. Xu, A. Nwokafor, R. Gupta, Y. Agarwal, Sentinel: occupancy based HVAC actuation using existing wifi infrastructure within commercial buildings, in: SenSys'13, 2013, p. 17.

[18] S. Hay, A. Rice, The case for apportionment, in: BuildSys '09, 2009, pp. 13–18.

[19] J. Scott, A. J. B. Brush, J. Krumm, B. Meyers, M. Hazas, S. Hodges, N. Villar, Preheat: controlling home heating using occupancy prediction, in: UbiComp 2011, 2011, pp. 281–290.

[20] Y. Cheng, K. Chen, B. Zhang, C. M. Liang, X. Jiang, F. Zhao, Accurate real-time occupant energy-footprinting in commercial buildings, in: BuildSys '12, 2012, pp. 115–122.

[21] Industry foundation classes - data model: http://www.buildingsmart.org.

[22] W.-K. Chang, T. Hong, Statistical analysis and modeling of occupancy patterns in open-plan offices using measured lighting-switch data, Building Simulation.

[23] D. Christin, A. Reinhardt, S. S. Kanhere, M. Hollick, A survey on privacy in mobile participatory sensing applications, Journal of Systems and Software 84 (11) (2011) 1928–1946.

[24] A. Schoofs, D. T. Delaney, G. M. P. O'Hare, A. G. Ruzzelli, COPOLAN: non-invasive occupancy profiling for preliminary assessment of HVAC fixed timing strategies, in: ACM BuildSys 2011, 2011, pp. 25–30.

[25] E. Patti, A. Acquaviva, M. Jahn, F. Pramudianto, R. Tomasi, D. Rabourdin, J. Virgone, E. Macii, et al., event-driven user-centric middleware for energy-efficient buildings and public spaces, Systems Journal, IEEE 99 (2014) 1–10.

[26] T. Toftkjær, M. B. Kjærgaard, The impact of sensor errors and building structures on particle filter-based inertial positioning, Pervasive and Mobile Computing 8 (5) (2012) 764–776.

[27] V. L. Erickson, Y. Lin, A. Kamthe, R. Brahme, A. Surana, A. E. Cerpa, M. D. Sohn, S. Narayanan, Energy efficient building environment control

strategies using real-time occupancy measurements, in: BuildSys '09, 2009, pp. 19–24.

[28] V. Erickson, M. A. Carreira-Perpinan, A. E. Cerpa, OBSERVE: Occupancy-based system for efficient reduction of HVAC energy, in: IPSN 2011, 2011, pp. 258–269.

[29] P. Kelly, N. E. O'Connor, A. F. Smeaton, Robust pedestrian detection and tracking in crowded scenes, Image and Vision Computing 27 (10) (2009) 1445 – 1458.

[30] F. C. Sangoboye, M. B. Kjærgaard, Plcount: A probabilistic fusion algorithm for accurately estimating occupancy from 3d camera counts, in: BuildSys 2016, ACM, 2016.

[31] M. Kjærgaard, A. Johansen, F. Sangogboye, E. Holmegaard, Occure: an occupancy reasoning platform for occupancy-driven applications, in: CBSE, ACM, 2016.

[32] A. Beltran, V. L. Erickson, A. E. Cerpa, Thermosense: Occupancy thermal based sensing for hvac control, in: BuildSys'13, 2013, pp. 11:1–11:8.

[33] Z. Zhang, D. Chu, X. Chen, T. Moscibroda, Swordfight: enabling a new class of phone-to-phone action games on commodity phones, in: MobiSys'12, 2012, pp. 1–14.

[34] J. He, A. Arora, A regression-based radar-mote system for people counting, in: IEEE PerCom 2014, 2014, pp. 95–102.

[35] N. B. Priyantha, A. Chakraborty, H. Balakrishnan, The cricket location-support system, in: MobiCom '00, 2000, pp. 32–43.

[36] Q. Pu, S. Gupta, S. Gollakota, S. Patel, Whole-home gesture recognition using wireless signals, in: MobiCom'13, 2013, pp. 27–38.

[37] R. K. Harle, A. Hopper, The potential for location-aware power management, in: UbiComp 2008, 2008, pp. 302–311.

[38] A. J. R. Ruiz, H. Blunck, T. S. Prentow, A. Stisen, M. B. Kjærgaard, Analysis methods for extracting knowledge from large-scale wifi monitoring to inform building facility planning, in: IEEE PerCom 2014, 2014, pp. 130–138.

[39] G. Conte, M. De Marchi, A. A. Nacci, V. Rana, D. Sciuto, Bluesentinel: A first approach using ibeacon for an energy efficient occupancy detection system, in: BuildSys '14, ACM, 2014, pp. 11–19.

[40] A. Ebadat, G. Bottegal, D. Varagnolo, B. Wahlberg, K. H. Johansson, Estimation of building occupancy levels through environmental signals deconvolution, in: BuildSys 2013, 2013, pp. 1–8.

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

[42] L. V. Thanayankizil, S. K. Ghai, D. Chakraborty, D. P. Seetharam, Softgreen: Towards energy management of green office buildings with soft sensors., in: COMSNETS, IEEE, 2012, pp. 1–6.

[43] T. A. Nguyen, M. Aiello, Beyond indoor presence monitoring with simple sensors, in: PECCS 2012, 2012, pp. 5–14.

[44] B. Dong, B. Andrews, K. P. Lam, M. Hynck, R. Zhang, Y.-S. Chiou, D. Benitez, An information technology enabled sustainability test-bed (itest) for occupancy detection through an environmental sensing network, Energy and Buildings 42 (7) (2010) 1038 – 1046.

[45] T. W. Hnat, E. Griffiths, R. Dawson, K. Whitehouse, Doorjamb: Unobtrusive room-level tracking of people in homes using doorway sensors, in: SenSys '12, ACM, 2012, pp. 309–322.

[46] P. Jaramillo, O. Amft, Improving energy efficiency through activity-aware control of office appliances using proximity sensing - A real-life study, in: IEEE SmartE Workshop, 2013, pp. 664–669.

[47] S. Thakur, M. Saha, A. Singh, Y. Agarwal, Wattshare: detailed energy apportionment in shared living spaces within commercial buildings, in: BuildSys 2014, 2014, pp. 30–39.

[48] M. Saha, S. Thakur, A. Singh, Y. Agarwal, Energylens: combining smartphones with electricity meter for accurate activity detection and user annotation, in: e-Energy '14, 2014, pp. 289–300.

[49] T. Teixeira, A. Savvides, Lightweight people counting and localizing in indoor spaces using camera sensor nodes, in: ACM/IEEE ICDSC 2007, 2007, pp. 36–43.

[50] A. Kamthe, L. Jiang, M. Dudys, A. Cerpa, Scopes: Smart cameras object position estimation system, in: EWSN '09, Springer-Verlag, 2009, pp. 279–295.

[51] J. Zhao, R. Yun, B. Lasternas, H. Wang, K. Lam, A. Aziz, V. Loftness, Occupant behavior and schedule prediction based on office appliance energy consumption data mining, in: CISBAT 2013, 2013, pp. 549–554.

[52] M. Milenkovic, O. Amft, An opportunistic activity-sensing approach to save energy in office buildings, in: e-Energy '13, ACM, 2013, pp. 247–258.

[53] M. Jin, R. Jia, Z. Kang, I. C. Konstantakopoulos, C. J. Spanos, Presencesense: Zero-training algorithm for individual presence detection based on power monitoring, in: BuildSys '14, ACM, 2014, pp. 1–10.

[54] W. Kleiminger, C. Beckel, T. Staake, S. Santini, Occupancy detection from electricity consumption data, in: ACM BuildSys 2013, 2013, pp. 1–8.

[55] C. Duarte, K. V. D. Wymelenberg, C. Rieger, Revealing occupancy patterns in an office building through the use of occupancy sensor data, Energy and Buildings 67 (0) (2013) 587 – 595.

[56] R. Melfi, B. Rosenblum, B. Nordman, K. Christensen, Measuring building occupancy using existing network infrastructure, in: IGCC 2012, 2011, pp. 1–8.

[57] M. B. Kjærgaard, Studying sensing-based systems: Scaling to human crowds in the real world, IEEE Internet Computing 17 (5) (2013) 80–84.