SolarWalk: Smart Home Occupant Identification using Unobtrusive Indoor Photovoltaic Harvesters

Md Fazlay Rabbi Masum Billah*
University of Virginia
masum@virginia.edu

Nurani Saoda*
University of Virginia
saoda@virginia.edu

Victor Ariel Leal Sobral
University of Virginia
sobral@virginia.edu

Tushar Routh
University of Virginia
tkr2w@virginia.edu

Wenpeng Wang
University of Virginia
wangwp@virginia.edu

Bradford Campbell
University of Virginia
bradjc@virginia.edu

ABSTRACT

The key to optimal occupant comfort as well as resource utilization in a smart building is to provide personalized control over smart appliances. Additionally, with an exponentially growing Internet-of-Things (IoT), reducing the need of frequent user attention and effort involving building management to control and manage an enormous number of smart devices becomes inevitable. One crucial step to enable occupant-specific personalized spaces in smart buildings is accurate identification of different occupants. In this paper, we introduce *SolarWalk* to show that small and unobtrusive indoor photovoltaic harvesters can identify occupants in smart home scenarios. The key observations are that i) photovoltaics are commonly used as a power source for many indoor energy-harvesting devices, ii) a PV cell's output voltage is perturbed differently when different persons pass in close range, creating an unique signature voltage trace, and iii) the voltage pattern can also determine the person' walking direction. *SolarWalk* identifies occupants in a smart home by training a classifier with their shadow voltage traces. *SolarWalk* achieves an average accuracy of 88% to identify five occupants in a home and on average 77% accurate to determine whether someone entered or exited the room. *SolarWalk* enables an accurate occupant identification system that is non-invasive, ubiquitous, and does not require dedicated hardware and rigorous installation.

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**.

KEYWORDS

Photovoltaic Harvesters, Occupant Identification

ACM Reference Format:
Md Fazlay Rabbi Masum Billah, Nurani Saoda, Victor Ariel Leal Sobral, Tushar Routh, Wenpeng Wang, and Bradford Campbell. 2022. *SolarWalk*: Smart Home Occupant Identification using Unobtrusive Indoor Photovoltaic Harvesters. In *The 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '22), November 9–10, 2022, Boston, MA, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3563357.3564073

*Both authors contributed equally to this research.

BuildSys '22, November 9–10, 2022, Boston, MA, USA

ACM ISBN 978-1-4503-9890-9/22/11.
https://doi.org/10.1145/3563357.3564073

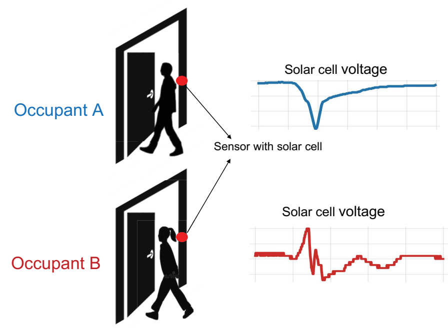


Figure 1: The photovoltaic harvester's output voltage attached to an indoor light energy-harvesting sensor fluctuates differently as different occupants of a home passes by. *SolarWalk* leverages this voltage fluctuations as an unique attribute to differentiate occupants.

1 INTRODUCTION

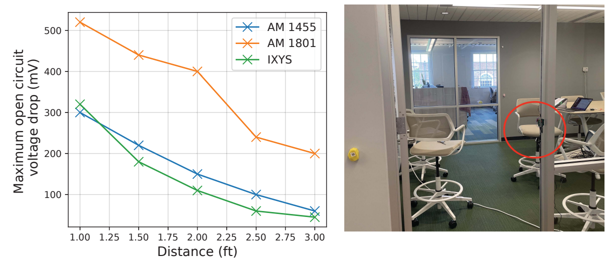
Accurate occupant identification can upgrade today's buildings from *smart* spaces to *intuitive* spaces. The knowledge of *who* is present over *someone* is present instruments indoor spaces with personalized control in applications such as HVAC, ambient lighting and many more. Such instrumentation of spaces has proven to significantly increase occupant comfort, reduce human intervention in building management, and decrease overall building energy consumption [1, 32, 40].

However, occupant identification is yet an unsolved problem due to several practical challenges. Device-based solutions that use wearable or carried devices require users to wear or carry the device most of the time [5, 22, 33]. Such approaches are intrusive and often less feasible in real-life scenario. Some solutions adopt device-free approach using camera/vision-based, audio-based techniques [2, 17, 29] to relieve the user off previous requirements, but people are likely to be reluctant to install these privacy-invasive devices. To eliminate privacy concerns, previous works explored special infrastructure/hardware-based solutions using ultrasonic, infrared, and vibration [20, 25, 37]. However, such systems often require multiple special purpose hardware, are often installation-heavy, unobtrusive, and unscalable. Some approaches leverage existing RF infrastructures like WiFi or Bluetooth [23, 41–43] to identify unique RF signatures of occupants. These approaches, however, suffer from signal degradation due to the multipath effect in indoor spaces and fail to adapt in dynamic spaces. Techniques involving on-object sensors require direct user interaction with the device and rigorous sensor installation [12, 36]. Radar-based identification systems rely on the unique signal reflection profile from human body to identify persons [10, 14, 18]. Solutions often require costly and bulky hardwares such as USRP, UWB, or mmWave radio and door-mounted systems fail to identify if multiple people walk closely.

To overcome the limitations of prior work, in this paper, we attempt to investigate how can we design an occupant identification system in a smart home scenario that is non-intrusive, ubiquitous, unobtrusive, and installation-light? To answer this question, we make two interesting observations. First, we observe that indoor photovoltaic energy-harvesting sensors are gaining tremendous traction due to their longevity and green source of energy. A growing number of these sensors are replacing battery-powered ones, pushing towards the vision of long-lived buildings [8, 9, 13]. In contrary to other harvestable sources, indoor light is more available, less variable, and is usually present if the space is occupied [11, 35], which makes the attached sensor ubiquitously deployable in most indoor spaces. Second, the output voltage fluctuations of a photovoltaic harvester when a person walks in front of it within a close range (e.g., through a door or a hallway) is a unique identifier of that person due to height, body shape, and gait differences and can be leveraged to distinguish between multiple occupants (as shown in Figure 1). Based on these observations, we introduce *SolarWalk*, a smart home occupant identification approach that adopts a photovoltaic harvester's voltage variation as a distinct feature to identify different occupants. *SolarWalk* collects the voltage variance over time from deployed sensors and feeds the voltage trace as a feature to a supervised learning-based classifier to distinguish between different occupants. Moreover, we also notice that, the voltage pattern can be an indicator of the direction of passing, e.g., entering or exiting a room, enabling *SolarWalk* to possibly determine if a room is occupied or not.

Energy-harvesting battery-less sensors are appealing to users and building management authorities due to their zero-maintenance and deploy-and-forget features. Typically these devices replace batteries with an energy-harvesting power supply front end. However, the potential behind re-purposing the energy harvester of these devices to function beyond just a power source is tremendous. With *SolarWalk*, an energy-harvesting PIR [7] or a door status sensor [21] is not only able to sense the presence of a person but also can determine who the event is associated with just by inspecting the ripples in the harvester voltage. By enabling this, *SolarWalk* further augments the capabilities of energy-harvesting sensors. We design *SolarWalk* to demonstrate that the jittery power supply voltage can actually be a source of context-aware data that the sensor otherwise wouldn't pay attention to.

Since *SolarWalk* is intended to be used with energy-harvesting sensors that have strict energy-budget, one challenge is to reduce the energy-overhead on the sensor while sampling and transmitting



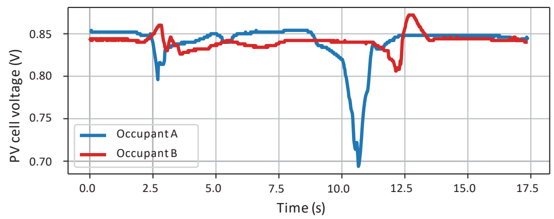
(a) **(b)**

Figure 2: a) A PV cell's open circuit voltage drops to different levels as someone walks at different distance from the solar cell's surface. b) Experimental setup with PV cell mounted on a office doorframe.

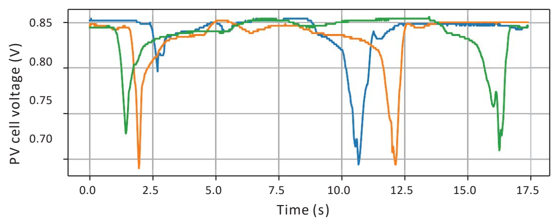
voltage data. We achieve this by making the voltage recording process event-triggered. The energy-harvesting sensors start sampling its solar cell upon detecting an external event, such as movement from an attached PIR sensor. Another challenge to design *SolarWalk*, is to be able to correctly differentiate between which voltage fluctuation actually corresponds to a walking event. Since a PV cell is just an energy-transducer without any voltage regulation, the output voltage can be noisy, as the illumination pattern of a space changes throughout the day. To overcome this challenge, we train *SolarWalk*'s classifier with historic data that contains no door events, but the steady-state output voltage of the harvester. Since the non-door event ripples are of different frequencies and amplitudes, *SolarWalk*'s classifier can differentiate between noise and event of interest.

We design and implement *SolarWalk* by collecting data from real-world deployment study. Our study involves five different occupants and we collect their shadow patterns from the solar cell mounted on two door frames of different rooms. An average US household accommodates 2.6 persons and 25.7% of US homes consist of at least two bedrooms [38, 39]. We prototype the *SolarWalk* design and evaluate the performance of the proposed solar-cell based occupant identification technique. Results show that *SolarWalk* achieves an average accuracy of 88% to identify five occupants and 77% accuracy in determining door entry or exit events. We make the following key contributions in this paper:

- We propose, *SolarWalk*, a new non-invasive technique that identifies occupants in a smart home context by inspecting the voltage variations of a photovoltaic energy-harvester as occupants walk in narrow spaces.
- *SolarWalk* augments the role of a battery-less device's energy-harvesting power supply to enable local context-awareness on the device.
- We implement the proposed design in a hardware prototype and demonstrate the feasibility of *SolarWalk* as an un-tethered, an unobtrusive occupant identification system.

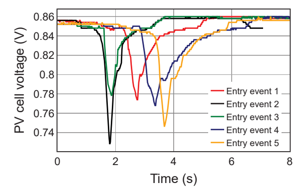


(a) Voltage fluctuations of occupant A and B are different from each other.

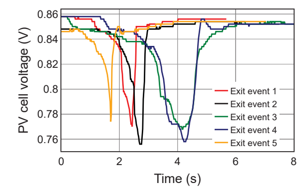


(b) Voltage fluctuations of the same occupant have similar shape.

Figure 3: This figure shows how the output voltage of the solar cell mounted on a doorframe ripples as different occupants pass through the door. The maximum voltage drop and the duration of voltage fluctuations vary differently for occupant A and B . On the other hand, these characteristics remain consistent over multiple trials by the same person.

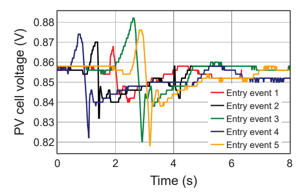


(a) Entry events of occupant A

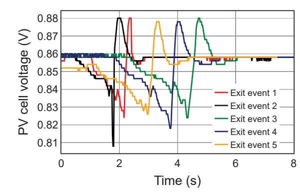


(b) Exit events of occupant A

Figure 4: Occupant A's entry and exit patterns are distinguishable. The patterns associated with the same type of event is similar. Since during entry and exit, the light is obstructed in similar but reverse direction, the entry and exit patterns tend to mirror each other.



(a) Entry events of occupant B



(b) Exit events of occupant B

Figure 5: Occupant B's entry and exit patterns are distinguishable. It is significantly different from occupant A's pattern.

2 OCCUPANT IDENTIFICATION USING PHOTOVOLTAICS

Indoor light energy-harvesting sensors typically harvest energy using one or multiple small photovolatic cells. These solar cells are usually optimized for a specific range of wavelength associated with indoor lighting conditions and the open circuit output voltage is proportional to the light intensity of its surrounding. During normal operation, the light intensity of indoor spaces changes steadily throughout the day until the light is turned off. The light intensity of the surrounding, however, undergoes a rapid change when someone passes nearby and is reflected in the output voltage of the solar cell. The maximum open circuit voltage drop induced on the solar cell decreases as the shadow of the person diminishes. Figure 2(a) shows the maximum voltage drop in three different indoor solar cells (both amorphous and monocrystalline) [15, 26] as someone walks by at different distances from the solar cell surface. Maximum voltage drop occurs when the person stands right in front of the cell completely blocking majority of the light exposure on the surface.

To better understand the characteristics of the open circuit output voltage when someone walks by within a few feet, we installed an IXYS indoor PV cell [15] on a door frame, halfway above the ground as shown in Figure 2(b). The solar cell surface is placed orthogonal to the floor. Since occupants walk in a narrow passage in spaces like doorways and hallways, such places are best suited for this study. We record the voltage traces as we enter and exit through the door. From Figure 3(a), we find that, for two different persons the voltage traces have different amplitude over time. Voltage drops as the person obscures the surface of the cell and restores itself as the person walks away. The amplitude of the ripple voltage is related to the height of the person and time length of the ripple is associated with someone's gait or walking style. However, Figure 3(b) shows that the shadow pattern is similar for the same occupant. This indicates that shadow pattern of a person as observed by a solar cell can be a characteristic feature for occupant identification.

Moreover, the pattern for different entry and exit events are distinguishable and can be used to determine if occupant A entered or exited the room to turn on/off any device in that room. Figure 4 show output voltage fluctuations for entry and exit events. We observe the shape of the entry and exit events shape tend mirror to each other and have an opposing skewed tail, indicating a sense of direction associated with the events. As the person enters the room, they do not obscure majority of the surface area until they reach the door frame plane which is orthogonal to the solar cell surface and continues blocking the light as they move away. However, for the exit event, voltage begins to decrease earlier than the person reaches the door frame. This happens mostly due to the brighter source of the light coming from inside the room. Typically rooms are brighter than hallways because of multiple light sources. This particular voltage pattern phenomena is a good indicator to determine from which direction the person crossed the solar cell.

Motivated by these observations, in this paper, we aim to design the proposed system named *SolarWalk*, that can identify persons using tiny, non-invasive solar cells.

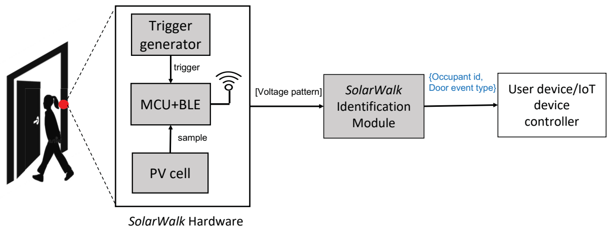


Figure 6: Overview of *SolarWalk* design.

3 SYSTEM DESIGN

3.1 Overview of *SolarWalk*

SolarWalk identifies occupants in smart homes by analyzing their associated distinct voltage patterns, reflected on a solar energy harvester as they walk in close proximity. *SolarWalk* design consists of two major components: *SolarWalk* hardware and *SolarWalk* identification module as shown in Figure 6. *SolarWalk* hardware records voltage traces from the PV cell as the event of interests occur and the identification module employs a pre-trained machine learning classifier trained from the data collected in the same physical environment.

The hardware module consists of an external trigger generator that notifies a microcontroller of a possible walking event. The microcontroller starts recording the door event until it finishes. Once the voltage trace is recorded, the MCU communicates the data over BLE to the *SolarWalk* identification module. *SolarWalk* identification module determines the identity label of who the door event is associated with and what type of door event it is (i.e., entry or exit from the room). We train the identification module with historic data containing both walk and no-walk events. During the training phase, the identification module relies on labeled voltage data with an occupant identifier and the type of event.

SolarWalk elevates the capability of an ordinary photovoltaic harvester by introducing the concept of meaningful power supply fluctuation. With *SolarWalk* we envision that, existing battery-less devices could be repurposed to do more than their usual sensing and these sensors could be crowdsourced to enable zone-specific data-fusioning.

3.2 *SolarWalk* Hardware

SolarWalk relies on the mobility of an occupant to record how someone's shadow pattern impacts the voltage generation. However, continuous sampling of solar cell voltage at the required frequency is energy-expensive, even when carefully duty-cycled. *SolarWalk* overcomes this challenge by incorporating an external trigger sensor to initiate voltage sampling. Since entering and exiting through a door in a home are not high-frequency events, the average energy-overhead can be kept significantly low.

Figure 7 shows the state machine of the software that runs on *SolarWalk* devices. The MCU waits for the trigger in low power mode with trigger enabled. Once the trigger is set ($trig = set$), the MCU starts sampling the solar cell at 50 Hz. We empirically determine the required sampling rate throughout our data collection study. The system needs to keep recording for the entire duration of the event and sets a timer ($TIM0 = t_r$) to stop sampling. The external trigger

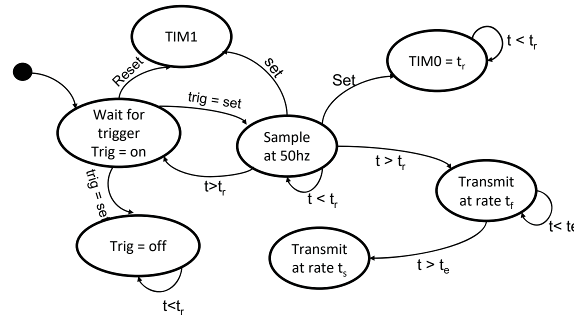


Figure 7: State machine representation of *SolarWalk* device's workflow

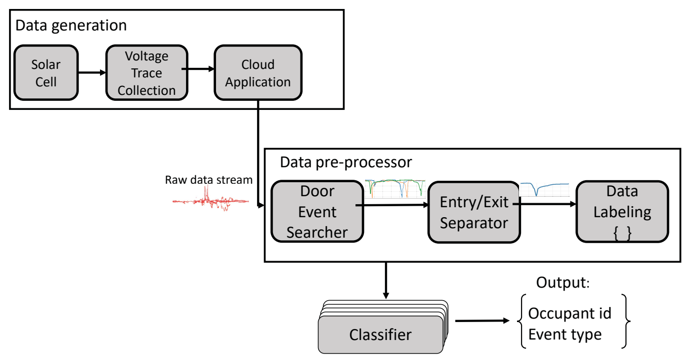


Figure 8: Block diagram of *SolarWalk* identification module

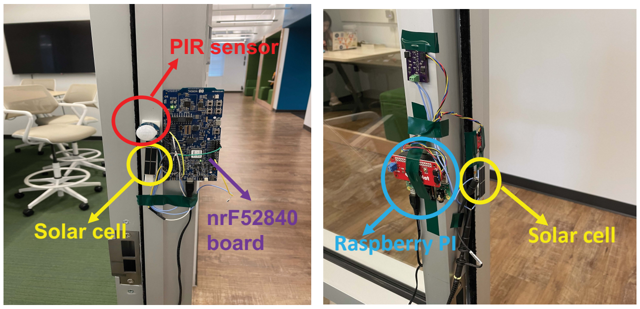
is turned off during an ongoing sampling to prevent further triggers while the event is being recorded and turned on once sampling is finished. Upon finishing sampling, the MCU transmits the data over BLE advertisements with t_f rate. The MCU also keeps track of how long it has been passed since the last trigger happened and if it is greater than t_e (TIM1), it lowers the advertisement rate to t_s to conserve more energy.

As we discussed in Section 2, the shadow pattern of a person diminishes with increasing distance from the solar cell surface. Though adoption of multiple solar cells could provide us wider range, we refrain from this design choice to make *SolarWalk* hardware unobtrusively fit in indoor spaces within a reasonable form factor. Moreover, solar energy-harvesting devices usually are extremely low-power devices and a majority of them incorporate at least one PV cell.

3.3 *SolarWalk* Identification Module

The identification module of *SolarWalk* system runs on a gateway receiver or an edge device and collects data from the hardware to perform the identification process using supervised learning techniques. We train the identification module with historic data using KNN supervised machine learning technique.

Figure 8 shows an overview of *SolarWalk*'s identification module's training phase. The training phase consists of major blocks: data generation, data pre-processor, and classifier. The data generation block accommodates a voltage trace collection module connected to a solar cell and sends the data over a cloud application for pre-processing. To be able to differentiate between steady state



(a) ***SolarWalk*** **hardware prototype**

(b) Data collection hardware platform

Figure 9: ***SolarWalk*** **prototype implementation**

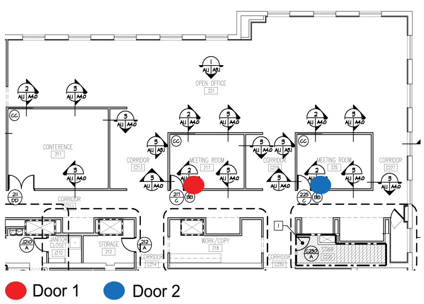


Figure 10: Floor plan showing the installed sensors on two doors of two different rooms.

voltage fluctuations and actual events of interest, the system trains the classifier with data samples containing both walk and no-walk events. The data generation block generates a robust dataset that captures the output voltage profile of the solar cell throughout different times of the day. The data pre-processor block receives a stream of data containing door entry and exit events for multiple occupants. In the pre-processing phase, the system separates door walk events from no-walks events. This process, however, is not required in the deployment phase since the identification module only receives walk events from *SolarWalk*'s hardware device. It also filters and labels entry and exit traces with user-provided label. From the time series data of door events, the pre-processor labels each occupant's entry and exit sequence. The entry and exit sequence of voltage samples are then fed into the classifier along with the occupant id label. The classifier outputs the result in terms of occupant label and the type of event.

4 IMPLEMENTATION

In this section, we discuss the implementation of *SolarWalk* hardware and the data collection platform in the training phase of *SolarWalk* identification module.

SolarWalk **hardware prototype.** We use a PIR sensor as the trigger generator to detect movement in the doorway. We incorporate a Panasonic AMN41121 [27] which can detect movement within 5m range with a 50° horizontal angle field of detection. We run *SolarWalk* software in the nRF52840 development kit [24]. The development kit accommodates a Cortex M4 processor SoC with

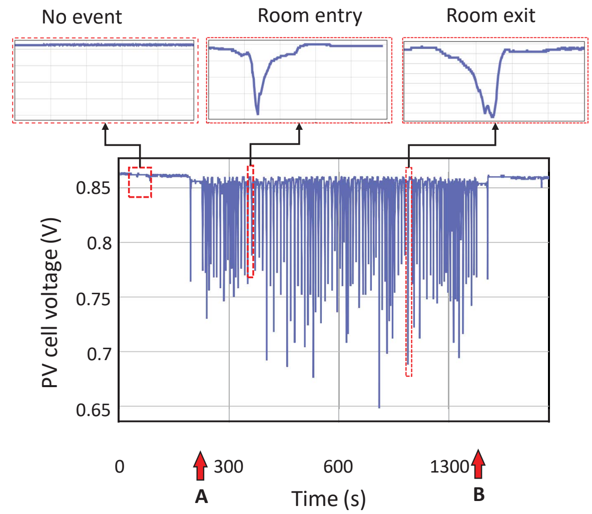


Figure 11: Data collection step of a single participant. At point A, the participant started walking. Each walk spans 10 seconds, which is either entering into the room or exiting out from the room. At point B, the participant stopped walking. In this particular trial, we collected data from 50 room entry events and 50 room exit events.

BLE 5 radio. The MCU is connected with a IXYS SMLD121H04L monocrystalline solar cell with a 22% efficiency. The solar cell is optimized to be used for both indoor and outdoor applications. The rate open circuit voltage is 2.52V with a short circuit current of 50 mA. The dimension of the solar cell is 43x14 mm. Figure 9(a) depicts the hardware prototype implementation.

At runtime, the MCU samples the solar cell at a 50hz rate using one of the internal ADC channels for $t_r = 6$ s. We set this value to capture the whole entry or exit event. We determine this value to be the maximum duration of any door event by analyzing the data collected during the training phase.

Data collection module. During the training phase, we adopt a data acquisition platform [35]consisting of a Raspberry Pi model 3A+. It connects a custom breakout board containing an ADS1015 analog to digital converter and to a Sparfun breakout board containing a VEML6030 illuminance sensor over I2C interface. This platform is configured to sample open-circuit voltage of a IXYS SMLD121H04L monocrystalline solar cell at a rate of 50 samples per second and also to record illuminance readings as a baseline for the data acquisition conditions. The ADS1015 gain stage was configured to 8, resulting in a full-scale resolution of ±4.096 Volts, and a 2 millivolt least significant bit size. The platform records and streams data using MQTT protocol to a cloud-hosted database, so we can later use the recorded data to train and evaluate our classifier models. Figure 9(b) shows the set up of the data collection module.

5 EVALUATION

To evaluate *SolarWalk*, our goal is to answer how accurately the solar cell voltage trace performs as an attribute to identify occupants in a 5-person household. We base our experiments on real-world

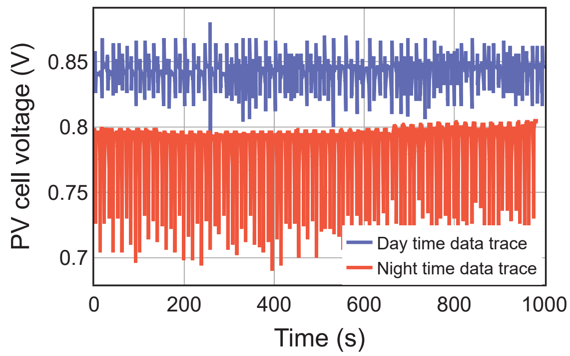


Figure 12: Voltage trace of a participant during day and night time. Open circuit voltage of solar cell changes throughout the day and can have impact on model performance. *SolarWalk* dataset includes traces from both day and night.

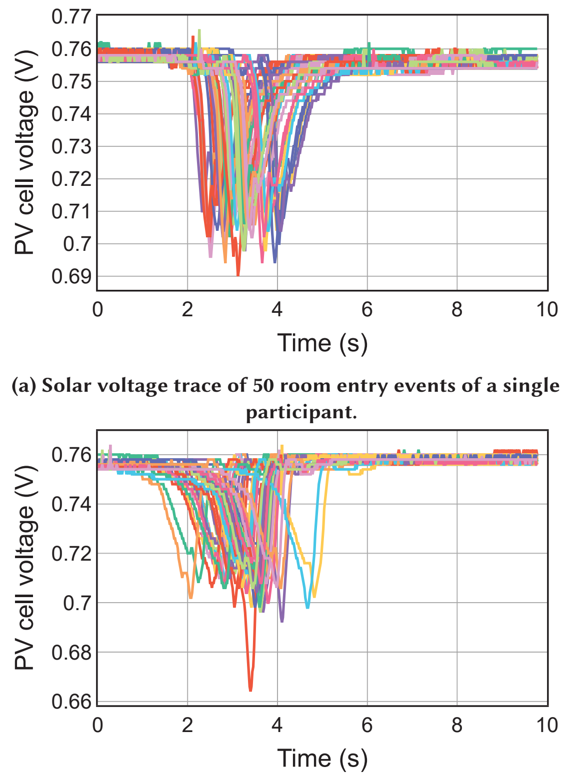
study to evaluate the performance of *SolarWalk* identification system. We explore how the identification accuracy is impacted by i) different systems parameter: the number of occupants, different classification methods, ii) environmental parameters: doors from different rooms, different times of a day, iii) physical attributes: different occupants and their heights. Another interesting feature of *SolarWalk* classifier is the ability to distinguish between two types of door events: entry and exit and we analyze how accurately the system can distinguish between these events.

5.1 Methodology

Experimental setup: We perform our data collection study by installing the *SolarWalk* data collection platform on two different doors of two middle rooms in our lab building. The width of both doorways is three feet. Figure 10 shows the floor plan including the installation points. Figure 9(b) depicts one of the setups. We install the device halfway above the floor on the doorframe to cover an optimal range of occupant height. The lower the position of the solar cell, the more likely the shadow of a person is going to impact the voltage. However, since solar energy-harvesting sensors usually should be placed as close as possible to the light source, we chose the midway to be the optimum point for deployment. We also deploy a working *SolarWalk* hardware on one of the doors to demonstrate the functionality and proof-of-concept implementation of the design(Figure 9(a)).

Data collection procedure: Our study involves five different occupants from different body shape in terms of height and girth. We collected 900 door entry and exit events from five participants as they walked through the door. Four participants walk a 100 times through each of the doors and one participants walk 100 times through one door. Figure 11 illustrates our data collection step of a single participant, which started at point A, and ended at point B. In this trial, we collected 50 room entry samples and 50 room exit samples. Each walk spans ten seconds.

We performed the data collection throughout different hours of the day including both day and night time to build a robust dataset,



(a) Solar voltage trace of 50 room entry events of a single participant.

(b) Solar voltage trace of 50 room exit events of a single participant.

Figure 13: Data collection step of *SolarWalk* involves each participant walking through the door every 10 seconds. However, a noticeable change in solar cell voltage pattern is observed in the first six seconds, which contains 300 voltage samples. Thus, the dimension of the input feature of our machine learning model is 1×300.

since the shadow pattern and the open circuit voltage of the solar cell is expected to change throughout the day. Figure 12 illustrates the solar trace of one participant's walk event during day time and night time. Each trace in this figure consists of 100 events (room entry or room exit) that lasts 1,000 seconds.

Data preprocessing procedure: Once we collected room entry and exit voltage traces from participants, we analyzed each trace carefully to identify the trigger point of the solar cell. Figure 13(a) illustrates 50 solar cell traces of one participant's entry event. We notice that, although each event spans for 10 seconds, a noticeable change in voltage pattern happens in the first six seconds. A similar outcome can be noticed in exit events Figure 13(b). As such, during training and testing our machine learning models, we have taken traces from the first six seconds. As our prototype collects data at 50Hz sampling rate, a single entry event or exit event contains 6×50 samples. Thus, as an input feature our ML models take 300 voltage readings.

Machine learning models: To evaluate *SolarWalk*, we implemented three supervised classifier algorithms: K-Nearest Neighbor (KNN) classifier, Random Forest, and Decision Tree. In our evaluation, the KNN classifier contains six neighbors. On the other hand,

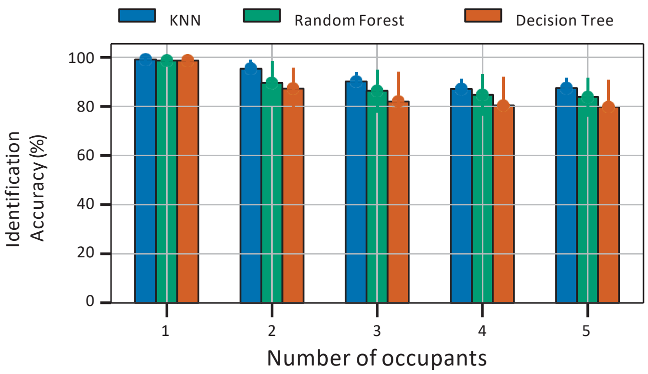


Figure 14: The figure shows that the occupant identification accuracy continues to drop as we increase the number of occupants. With five occupants *SolarWalk*'s KKN classifier achieves 88% accuracy.

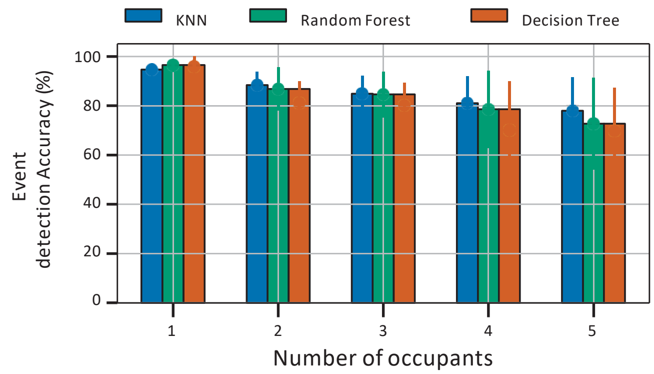


Figure 15: Here, we show how the type of event detection accuracy changes with increasing number of occupants. *SolarWalk* classifier can on average accurately identify between door entry and exit events 77% of time.

the Random Forest classifier consists of 10 trees and uses entropy as the loss function. We performed 10-fold cross-validation while training and testing each model.

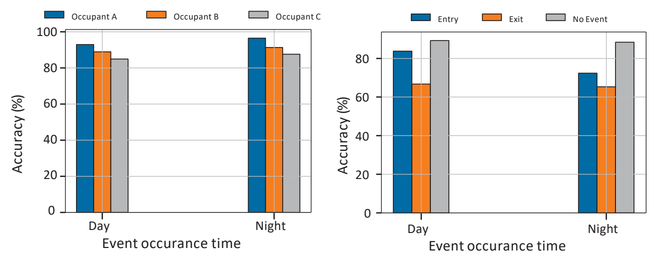
5.2 Overall System Performance

In this section, our goal is to evaluate how accurately the system can identify different occupants and distinguish between two different door activity. Results show that, our KNN-based classifier can accurately detect the identity of occupants on average 87% of the time in a 5-person home and on average 95% of the time in a 2-person small home. We also explore the performance of two other supervised learning method: decision tree [30], random forest [3] for comparison. Figure 14 the how percentage identification accuracy changes with an increasing number of occupants across different classification methods. The plot shows the distribution over 10 trials. From the result, we find that the percentage of accuracy drops from 99% for one occupant to 88% for five occupants, denoting a 13% point decrease. This represents that the solar cell shadow feature is a new accurate physical attribute for the occupant identification for homes with less than 5 people. However, as the demographic increases, the system might fail to perform acceptably and more robust learning techniques i.e., reinforcement learning is needed for high occupancy spaces such as offices or classrooms.

To determine if *SolarWalk* can differentiate between a door entry and exit event, we measure the event detection accuracy as we vary the number of occupants. Figure 15 shows that, on average *SolarWalk* can correctly differentiate between entry and exit events with a probability of .77 for five people. For two persons, it can detect events with an accuracy of 88%.

5.3 Environmental Effect

Since a photovoltaic's energy conversion efficiency is dependent on a number of factors including the spectrum of exposed light and illuminance of the surface, its open circuit voltage varies throughout different indoor spaces and hours of the day. Therefore, the shadow pattern of a person is different in multiple doors. However, it should still preserve characteristics to be distinguishable from another person. In this section, we explore how *SolarWalk* performs during the day vs night and the performance among two doors.

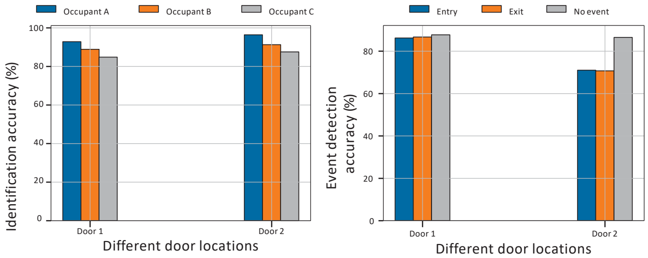


(a) Occupant identification accuracy

(b) Type of event detection accuracy

Figure 16: These plots show the effect of different times in a day on the system's accuracy. Since the steady state voltage of the solar cell undergoes variation due to different illuminance levels throughout the day, the voltage pattern's DC component shifts. Yet, system performance stays similar with a slightly higher accuracy for night events.

In Figure 16 we show the impact of different hours of the days has on the identification accuracy of three participants and event detection accuracy. For all three occupants, we evaluate the results from the data collected during two different time durations of the day. The day time voltage readings are collected within 12 pm to 4 pm and night time readings are collected after 8 pm. From Figure 16(a) we find that, for all three participants, the identification accuracy at night time is higher than day time by 3.5%, 2.3%, and 2.7%. This slight difference happens due to indoor spaces getting illuminated by natural night during day time. Therefore, someone's shadow makes the surface of the solar cell less illuminated during night than day and results in a larger voltage drop. We verify this observation from the attached illuminance sensor in the data collection module. The event detection accuracy stays similar for the exit and no event scenario as shown in Figure 16(b), but entry event accuracy drops by 11.4%. This could happen, since while entering through the door, as opposed to, exiting to the hallway, the brighter illuminance of the room light plays a role to distort the shape of the pattern. To summarize, event detection accuracy is affected more by the brightness variation throughout the day than occupant identification.



(a) (b)

Figure 17: a) *SolarWalk*'s identification accuracy remains similar for multiple occupants over two deployment locations. b) Event detection accuracy achieved at different doors. The accuracy of Door 2 is at least 15% lower for entry and exit events than Door 1. For the direction of movement, location seems to play an important role.

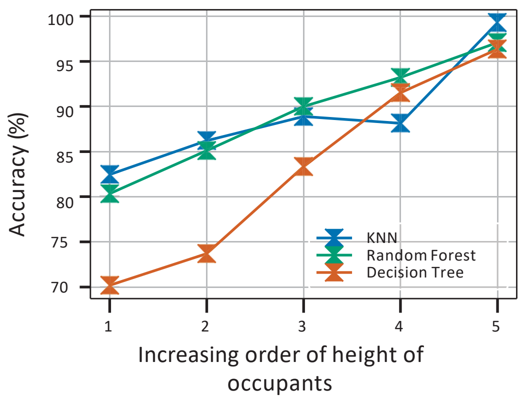


Figure 18: This figure plots the identification accuracy of the model with increasing occupant height. Occupant's height plays as an important factor for the system's identification accuracy. The taller height produces more distinguishable shadow pattern.

Figure 17(a) shows if the classifier performs in a similar manner in terms of occupant identification accuracy on two different doors. The identification accuracy for different occupants stays within a difference of 3.6% between the doors. However, the accuracy of individual occupants drops from overall single occupant identification accuracy because we only consider data from one location for this scenario. Figure 17(b) shows the event accuracy. We find that Door 1 achieves higher accuracy than Door 2 for all event types, which denotes that deployment location matters more for event detection type. Since the pattern of entry and exit events tend to change more depending on the position of source of light.

5.4 Sensitivity to Physical Attributes

A individual's body shape features plays role in the shadow formation [28, 34]. In order to understand how *SolarWalk* performs across individual occupants, we analyze the system's occupant identification accuracy and event detection accuracy for each participants.

Figure 18 shows individual occupant's identification accuracy in the increasing order of height. Identification accuracy improves with increasing height, which is expected. The maximum and minimum identification accuracy is 99% and 82% respectively. Figure 19 shows event detection accuracy for different occupants. We find that occupant E achieves a maximum of 100% and 85.7% accuracy for entry and exit event detection respectively. All of the participants achieve an entry accuracy of more than 80%, however, exit events see less accuracy. This matches our findings from previous sections.

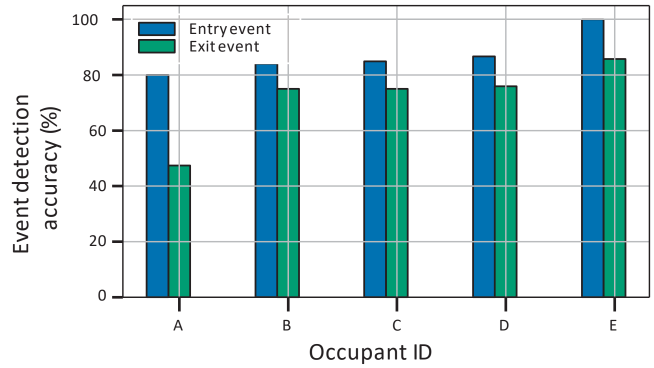


Figure 19: Event detection accuracy of an individual participant. Entry events are likely to be detected more accurately than exit events.

6 RELATED WORK

We categorize the related works into two major directions: augmenting energy-harvesting power sources and occupant identification in smart home context. We highlight some of the prior works that share similar motivations as *SolarWalk*.

6.1 Harvesters as Sensors

Instrumenting the power sources of energy-harvesting devices to generate more expressive and meaningful data have been explored in several prior works. Monjolo presents an energy-harvesting power meter where the rate of harvesting energy by power source is leveraged to calculate how much current is consumed by the attached load. It demonstrates a case of redefining the harvester to produce more expressive data. While Monjolo exploits the recharge rate of an intermittent energy-harvesting node to infer energy-metering information, we demonstrate that even noisy solar cell voltage data can have context-rich data. In [4] authors infers the occupancy status of a room by simply inspecting if the node was able to harvest, wake up, and transmit a beacon. Since the node was able to harvest signifies, the room was lit and someone could be present. These works attempt to infer intuitive information exploiting the energy-harvester that comes as an inevitable design choice of self-powered systems.

6.2 Occupant Identification

Scalable and cost-efficient unobtrusive occupancy detection setups for smart buildings have been attempted previously in several works. Various types of sensors have been exploited for satisfying the myriad requirements of identification at diversified smart environments, which include RF-based [5, 6, 19], Ultrasonic-based [31] techniques or also collecting information from on-object sensors [12].

Utilizing cellular frequencies, BlueSentinel [5] has been proposed that can detect the number of users in a room and track them inside the building. Including iBeacon's location information and KNN as the classifier, BlueSentinel achieves accuracy near 83%. A similar methodology was deployed on a large scale in [6], encompassing diversified surroundings like office buildings and dormitories on a university campus. However, electricity and water consumption information was added to WiFi data over 4 weeks duration with occupancy varying from 0 to 550. Mean absolute percentage error exhibits that incorporation of multi-modal data to estimate the occupancy escalates detection accuracy. As height and weight combination is a unique feature for personalizing, non-intrusive occupant identification has been proposed by utilizing those features in [19]. This system takes into account 7 distinct features of a human being (including hand weight distance, bouncing pattern during walking etc.) for identification. After evaluating in multiple test beds, it has been demonstrated this system can detect a person with accuracy varying from 90%-100%. In [31], authors present MODES, which utilizes thermal and vibration information with an accuracy of *73%* and *84%* in high and low occupancy scenarios. However, all these techniques require heavy infrastructure, multiple device installations or carried devices which fails to achieve scalability.

In another branch of work, researchers have focused on utilizing on-object sensors to infer occupancy. For example, SenseTribute [12] collects personalized features from different on-object sensors such as accelerometers and gyroscopes installed on domestic utility products (refrigerator, towel dispenser etc.) to classify occupants. Since, different occupants interact with an object in different manners such as the pattern of knocking on a door, or opening a fridge, the vibration data collected from the attached sensors can be a unique personal attribute. SenseTribute achieves an identification accuracy of 74% and 96% for known and unknown training labels. MotionSync [16] proposes an approach to determine personalized energy consumption by occupants by finding the correlation between motion data from users' wearables and appliances. It classifies the appliances in five categories based on their interfaces to learn the interaction between user-appliance. We share similar motivations of these work to eliminate the need for infrastructure-heavy methods, rather exploit already existing sensors and augment them with richer capabilities.

7 DISCUSSIONS

In this paper, we demonstrate how solar energy-harvesting sensors can utilize their power source to collect context-rich information such as the identity of occupants interacting with the sensor just by sampling their raw harvester output voltage. However, in this section, we highlight some remaining challenges and potential research directions.

Larger demographic and deployment conditions. Since *SolarWalk* relies on the shadow feature of a person, if two people in a home have similar body shape, the system might fail to distinguish them. A natural following step to build upon our initial results and make our technique more robust is to collect data representing a wider range of demographic and deployment spaces. For instance, recruiting participants with different and similar body sizes and likeness and with different gait patterns could support a more comprehensive evaluation of the proposed approach, since it would be based on a larger demographic. Though, we argue that *SolarWalk* enables accurate person identification in an average-size smart home, a possible dimension to explore is deploying the system at locations with a wider range of illumination conditions as to cover a larger set of possible real-world scenarios. For instance, evaluating the system with different combinations of natural and artificial light sources could better capture specific deployment conditions and increase the deployability of the system more than on just doors.

Potential applications. As we demonstrated in this work, a single and brief voltage time series recording from a solar cell can potentially carry enough information to classify occupants with reasonable accuracy, what can be even more valuable for smart building applications is to create "dynamic" spaces according to the occupant's preferences and needs. We can imagine *SolarWalk* to incorporate multiple solar-powered sensors present in a space and the spatio-temporal correlation between these sensors' solar cell voltage readings can provide even richer information, not only allowing better occupant classification but also enabling other potential applications such as activity recognition, or monitor the zones inside a shopping complex or museum to analyze which items get more attention or track complex usage patterns of smart building spaces. These envisioned applications also come with a number of challenges, for instance, deciding optimal strategies to process and exchange information between energy-constrained battery-less sensors, and the development of machine learning models that provide best results given the sensors' limited energy and computation capability.

Incorporating new occupants. Currently, *SolarWalk* does not incorporate any policy to handle data from unknown users. However, a realistic scenario would be able to update the model if the occupant situation in a home changes over time. In this case, online learning-based techniques such as reinforcement learning can be adopted to increase the robustness of the system.

Data labeling. An important challenge to using the supervised learning technique as the ones used in our work is the need to label the data with ground truths. While controlled experiments can be used to collect labeled data, they are time-consuming and needed to be repeated for each set of conditions (e.g if the illumination source changes). One possible alternative approach to collect labeled data is to use an user's interaction with another smart device in the room. For instance, if an occupant walks into a room and their cellphone connects to a voice assistant device, the system can use the logged id to label the data previously collected by the *SolarWalk* device to the respective occupant.

8 CONCLUSION

Future smart buildings will be a lot more personalized, greener, and full a of large network of nearly-invisible devices. To enable such a vision, one crucial step is to design systems that are aware of their contextual cues, yet simple, unobtrusive, and, installation-friendly. As a forward step, in this paper, we introduce *SolarWalk* to enable occupant-specific personalized control by sensing the voltage perturbance of photovoltaic energy-harvester. *SolarWalk* not only demonstrates a novel and accurate non-invasive, infrastructure-free occupant identification system, but also introduces the concept of empowering the power sources of battery-less energy-harvesting applications with meaningful contextual data. We believe innovation flourishes more rapidly when systems build on existing resources, which would otherwise just be wasted. We envisage this work would enable more interesting applications in the field of smart building research.

9 ACKNOWLEDGEMENTS

We thank the anonymous reviewers for their valuable insights and helpful feedback on improving this paper. This work is supported in part by the National Science Foundation (NSF) under grants CNS-1823325 and CNS-2144940, and the Strategic Investment Fund at the University of Virginia.

REFERENCES

[1] Bharathan Balaji, Jian Xu, Anthony Nwokafor, Rajesh Gupta, and Yuvraj Agarwal. 2013. Sentinel: occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*. 1–14.
[2] Kevin W Bowyer, Karen Hollingsworth, and Patrick J Flynn. 2008. Image understanding for iris biometrics: A survey. *Computer vision and image understanding* 110, 2 (2008), 281–307.
[3] Leo Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32.
[4] Bradford Campbell and Prabal Dutta. 2014. An energy-harvesting sensor architecture and toolkit for building monitoring and event detection. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*. 100–109.
[5] Giorgio Conte, Massimo De Marchi, Alessandro Antonio Nacci, Vincenzo Rana, Donatella Sciuto, et al. 2014. BlueSentinel: a first approach using iBeacon for an energy efficient occupancy detection system.. In *BuildSys@ SenSys*. Citeseer, 11–19.
[6] Aveek K Das, Parth H Pathak, Josiah Jee, Chen-Nee Chuah, and Prasant Mohapatra. 2017. Non-intrusive multi-modal estimation of building occupancy. In *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*. 1–14.
[7] Desk Occupancy Sensor . [n.d.]. EnOcean. https://www.enocean.com/en/product/occupancy-sensor-sub-desk-eosd/.
[8] Enocean Self-powered IoT. [n.d.]. . https://www.enocean.com/en/technology/energy-harvesting/.
[9] Everactive. [n.d.]. . https://everactive.com/.
[10] Chao Feng, Jie Xiong, Liqiong Chang, Fuwei Wang, Ju Wang, and Dingyi Fang. 2021. RF-Identity: Non-Intrusive Person Identification Based on Commodity RFID Devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (2021), 1–23.
[11] Maria Gorlatova, Aya Wallwater, and Gil Zussman. 2012. Networking low-power energy harvesting devices: Measurements and algorithms. *IEEE Transactions on Mobile Computing* 12, 9 (2012), 1853–1865.
[12] Jun Han, Shijia Pan, Manal Kumar Sinha, Hae Young Noh, Pei Zhang, and Patrick Tague. 2017. Sensetribute: smart home occupant identification via fusion across on-object sensing devices. In *Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments*. 1–10.
[13] Josiah Hester and Jacob Sorber. 2017. The future of sensing is batteryless, intermittent, and awesome. In *Proceedings of the 15th ACM conference on embedded network sensor systems*. 1–6.
[14] Chen-Yu Hsu, Rumen Hristov, Guang-He Lee, Mingmin Zhao, and Dina Katabi. 2019. Enabling identification and behavioral sensing in homes using radio reflections. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.
[15] IXOLAR High Efficiency SolarMD. 2016. IXYS. https://ixapps.ixys.com/DataSheet/SLMD121H04L_Nov16.pdf.
[16] Josiah Jee, Aveek K Das, Parth H Pathak, and Prasant Mohapatra. 2016. Motionsync: personal energy analytics through motion tags and wearable sensing. In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*. 65–74.
[17] Kaustubh Kalgaonkar and Bhiksha Raj. 2007. Acoustic doppler sonar for gait recogination. In *2007 IEEE Conference on Advanced Video and Signal Based Surveillance*. IEEE, 27–32.
[18] Avinash Kalyanaraman, Dezhi Hong, Elahe Soltanaghaei, and Kamin Whitehouse. 2017. Forma track: tracking people based on body shape. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–21.
[19] Nacer Khalil, Driss Benhaddou, Omprakash Gnawali, and Jaspal Subhlok. 2016. Nonintrusive occupant identification by sensing body shape and movement. In *Proceedings of the 3rd ACM international conference on systems for energy-efficient built environments*. 1–10.
[20] Nacer Khalil, Driss Benhaddou, Omprakash Gnawali, and Jaspal Subhlok. 2017. Sonicdoor: scaling person identification with ultrasonic sensors by novel modeling of shape, behavior and walking patterns. In *Proceedings of the 4th ACM International Conference on Systems for Energy-Efficient Built Environments*. 1–10.
[21] Magnet Contact Sensor. [n.d.]. EnOcean. https://www.enocean.com/en/product/magnet-contact-sensor-emcs-oem/.
[22] Jani Mantyjarvi, Mikko Lindholm, Elena Vildjiounaite, S-M Makela, and HA Ailisto. 2005. Identifying users of portable devices from gait pattern with accelerometers. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, Vol. 2. IEEE, ii–973.
[23] Ghassem Mokhtari, Qing Zhang, Ghavameddin Nourbakhsh, Stephen Ball, and Mohanraj Karunanithi. 2017. BLUESOUND: A new resident identification sensor—Using ultrasound array and BLE technology for smart home platform. *IEEE Sensors Journal* 17, 5 (2017), 1503–1512.
[24] Nordic Semiconductor. 2019. nrF52840 Dev Kit. https://www.nordicsemi.com/Products/Development-hardware/nrf52840-dk.
[25] Shijia Pan, Ningning Wang, Yuqiu Qian, Irem Velibeyoglu, Hae Young Noh, and Pei Zhang. 2015. Indoor person identification through footstep induced structural vibration. In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications*. 81–86.
[26] Panasonic Amorphous Solar cell. [n.d.]. . https://www.endrich.com/sixcms/media.php/2308/Catalogue20Sanyo20Amorton.490768.pdf.
[27] PIR motion sensor. [n.d.]. Panasonic Electronics. https://www.mouser.com/datasheet/2/315/panasonic_amn1_2_4-1196943.pdf.
[28] KC Prabin. 2010. Theory of Shadow. *Himalayan Physics* 1 (2010), 101–105.
[29] Ruijie Quan, Xuanyi Dong, Yu Wu, Linchao Zhu, and Yi Yang. 2019. Auto-reid: Searching for a part-aware convnet for person re-identification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 3750–3759.
[30] J. Ross Quinlan. 1986. Induction of decision trees. *Machine learning* 1, 1 (1986), 81–106.
[31] Hamid Rajabi, Zhizhang Hu, Xianzhong Ding, Shijia Pan, Wan Du, and Alberto Cerpa. 2022. MODES: multi-sensor occupancy data-driven estimation system for smart buildings. In *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*. 228–239.
[32] Juhi Ranjan, Erin Griffiths, and Kamin Whitehouse. 2014. Discerning electrical and water usage by individuals in homes. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*. 20–29.
[33] Juhi Ranjan, Yu Yao, and Kamin Whitehouse. 2013. An RF doormat for tracking people's room locations. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. 797–800.
[34] Makoto Shinzaki, Yumi Iwashita, Ryo Kurazume, and Koichi Ogawara. 2015. Gait-based person identification method using shadow biometrics for robustness to changes in the walking direction. In *2015 IEEE Winter Conference on Applications of Computer Vision*. IEEE, 670–677.
[35] Victor Ariel Leal Sobral, John Lach, Jonathan L Goodall, and Bradford Campbell. 2021. Thermal Energy Harvesting Profiles in Residential Settings. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*. 520–523.
[36] Elahe Soltanaghaei and Kamin Whitehouse. 2016. Walksense: Classifying home occupancy states using walkway sensing. In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*. 167–176.
[37] Vijay Srinivasan, John Stankovic, and Kamin Whitehouse. 2010. Using height sensors for biometric identification in multi-resident homes. In *International conference on pervasive computing*. Springer, 337–354.
[38] Statista. 2020. Statista. https://www.statista.com/statistics/206393/distribution-of-housing-units-in-the-us-by-number-of-bedrooms/.
[39] United States Census Bureau. 2016-2020. United States Census Bureau. https://www.census.gov/quickfacts/fact/table/US/HCN010217.
[40] Wenpeng Wang, Jianyu Su, Zackary Hicks, and Bradford Campbell. 2020. The Standby Energy of Smart Devices: Problems, Progress, & Potential. In *2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI)*. IEEE, 164–175.
[41] Yunze Zeng, Parth H Pathak, and Prasant Mohapatra. 2016. WiWho: WiFi-based person identification in smart spaces. In *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 1–12.
[42] Jin Zhang, Bo Wei, Wen Hu, and Salil S Kanhere. 2016. Wifi-id: Human identification using wifi signal. In *2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE, 75–82.
[43] Jin Zhang, Bo Wei, Wen Hu, Salii S Kanhere, and Ariel Tan. 2016. Human identification using WiFi signal. In *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. IEEE, 1–2.