BOES: Building Occupancy Estimation System Using Sparse Ambient Vibration Monitoring

Shijia Pan[1*] Amelie Bonde[2] Jie Jing[3] Lin Zhang[4] Pei Zhang[1] Hae Young Noh[5]

[1]Dept. of Electrical and Computer Eng., Carnegie Mellon Univ., Pittsburgh, PA, USA, 15213;
[2]School of Computer Science, Carnegie Mellon Univ., Pittsburgh, PA, USA, 15213;
[3]School of Architecture, Tsinghua Univ., Beijing, China, 100084;
[4]Electronic Eng. Dept., Tsinghua Univ., Beijing, China, 100084;
[5]Dept. of Civil and Environmental Eng., Carnegie Mellon Univ., Pittsburgh, PA, USA, 15213

ABSTRACT

In this paper, we present a room-level building occupancy estimation system (BOES) utilizing low-resolution vibration sensors that are sparsely distributed. Many ubiquitous computing and building maintenance systems require fine-grained occupancy knowledge to enable occupant centric services and optimize space and energy utilization. The sensing infrastructure support for current occupancy estimation systems often requires multiple intrusive sensors per room, resulting in systems that are both costly to deploy and difficult to maintain. To address these shortcomings, we developed BOES. BOES utilizes sparse vibration sensors to track occupancy levels and activities. Our system has three major components. 1) It extracts features that distinguish occupant activities from noise prone ambient vibrations and detects human footsteps. 2) Using a sequence of footsteps, the system localizes and tracks individuals by observing changes in the sequences It uses this tracking information to identify when an occupant leaves or enters a room. 3) The entering and leaving room information are combined with detected individual location information to update the room-level occupancy state of the building. Through validation experiments in two different buildings, our system was able to achieve 99.55% accuracy for event detection, less than three feet average error for localization, and 85% accuracy in occupancy counting.

Keywords: building occupancy, ambient vibration, sparse sensing, localization, tracking, event detection, anomaly detection

1. INTRODUCTION

A ubiquitous building maintenance system often calls for spatio-temporal information of occupants. For example, management of a commercial building may use occupancy information to design energy saving strategies. Stores and shopping malls may utilize occupancy tracking information to estimate customer behavior patterns and then determine related advertisement strategies and up-to-date store layouts. Current occupancy estimation systems often require multiple sensors per room with line-of-sight sensing, such as arrays of cameras or infrared sensors.[1] Heavily instrumented infrastructure leads to expensive, labor-intensive deployment and maintenance, which limits its application.

To address these challenges, we introduce a Building Occupancy Estimation System (BOES) to obtain room-level occupancy information using sparse ambient vibration monitoring systems, commonly used for structural health monitoring (SHM) purposes. Although sensors in SHM systems are often sparse and noisy for occupancy estimation purposes, BOES overcomes these issues by utilizing known building layout, general walking patterns, and structural characteristics. BOES collects sensor readings, detects footstep events, and recognizes significant events based on observed walking traits (such as passing by a sensor and entering/leaving a room). Then it computes the relative locations of these events with respect to a group of sensors, whose locations are known in the building layout, to obtain the absolute location of the events and occupancy counts in the building. In order to evaluate our system performance, we deployed BOES in the hallways of two buildings, which have multiple rooms along the way, and monitored all one-dimensional traffic through the hallway and into/out of the rooms.

* shijiapan@cmu.edu; phone: 650-335-2821

Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2014,
edited by Jerome P. Lynch, Kon-Well Wang, Hoon Sohn, Proc. of SPIE Vol. 9061, 90611O
 · doi: 10.1117/12.2046510

The main contributions of this paper are:

- A sequential footstep localization and tracking algorithm using a sparse array of sensors.
- A Room-level occupancy tracking technique that leverages structural features.
- Characterization of the propagation of footstep vibration in two buildings

The remainder of this paper is structured as follows. Section 2 presents prior work in occupancy estimation and research gaps. Section 3 describes system design and detailed methodology. Section 4 shows the implementation of the platform and the deployment at different locations. This section also describes the experiments we conducted in those locations. Section 5 explains the results and performance evaluation of our system. Finally we discuss conclusions and future work in Section 6.

2. RELATED WORK

There is a large body of work revolving around occupancy detection utilizing a variety of sensors. The most common approach uses radio frequency beacons to estimate the locations of mobile devices and thus the locations of the carriers of those devices.[2–8] However, these works tend to be limited in several ways. When the range of radio is high (e.g., areas with low infrastructure requirements), the signals cannot provide accurate high resolution occupancy estimates in an indoor environment, due to multi-path and fading caused by the environment. To account for this, radio frequency based approaches generally rely on a dense infrastructure[6] or on collecting a large amount of environmental signatures.[9] Furthermore, because the approach requires the user to carry a device, it leads to inaccuracies when the device is off or not on the user.

Optical sensors such as infrared (IR) sensors, near-infrared (NIR) sensors, and cameras have also been used for occupancy tracking.[1,10–13] IR sensors are widely used by lighting control systems to detect human movements in a room, so that lights can be turned off in unoccupied rooms to save energy. NIR sensors are used for occupancy monitoring by detecting the reflection of human skin. Cameras deployed in public places can be used to track occupants' walking routes and directions, but this suffers great controversy in terms of privacy issues. Although these methods are accurate and can handle a large amount of users, they require the installation and maintenance of costly infrastructure that must be carefully placed to maintain line of sight to any potential occupants. This makes these systems impractical for room-level occupancy tracking.

Other contextual information (e.g., schedules, calendars, etc.), can also be used for occupancy estimation.[14] However, these schedules are not always available due to a variety of reasons. Even when they are available, they may be outdated, inaccurate, not provide a count of occupants, or difficult to obtain in digital form. Instead of measuring occupancy indirectly, some contextual approaches use sensors to measure them directly. Direct measurements include pressure sensors in seats or floor tiles to sense occupants.[15,16] However, like other fine-grained localization work, these also require significant infrastructure deployment and maintenance, making them inappropriate for building level occupancy tracking.

Researchers have also used seismic and acoustic sensors to explore human signatures like footstep vibration frequency characteristics and walking style.[17–20] These works focus on characterization and detection of human footsteps, where our work instead focuses on localization and tracking based on these detected footsteps. Some works also suggest methods by which localization can be performed using footsteps;[19,20] however, they only observed the possibility of using vibration sensing for localization. To the best of our knowledge, this is the first work that develops and evaluates footstep localization and occupant counting methods. Furthermore, our system monitors changes in occupancy utilizing structural information.

3. SYSTEM DESCRIPTION

BOES consists of three main modules: event detection, localization, and tracking (Figure 1). First, the sensors feed floor vibration velocity (vertical) data into the event detection module. This module estimates whether the current data contains a step. Once a *step event* (SE) is detected, it is sent to the localization module. The localization module then connects the SEs into a *sequence of step events* (SSE) and iteratively locates the

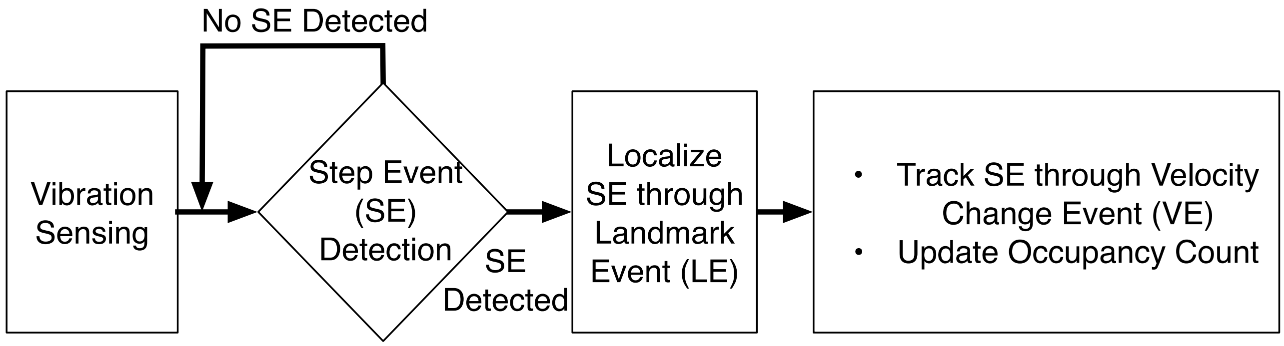

Figure 1: System overview.

sequence by comparing the time of peak SE energy at each sensor, referred to as a *landmark event* (LE). With the detected LEs in the SSE from multiple sensors, the module analyzes the partially localized SSE and determines the travel direction to fully localize the SSE. The tracking module then searches for the *velocity change event* (VE), which is a sudden change in the *maximum floor velocity* (MV) between two or three consecutive SEs. A VE indicates that an occupant left the hallway to enter an adjoining room or left an adjoining room to enter the hallway. The tracking module combines the VEs with the localized SSEs and LEs obtained from the localization module, and updates the occupancy state of each room.

3.1 Event Detection Module

The event detection module collects digitized sensor readings from a group of sensors that monitors overlapping areas, determines whether a SE occurred, and labels the detection time of SE. In order to identify a SE, we consider all readings from a group of neighboring sensors that are within a sensing range (i.e., they can sense the same events). Due to the fast propagation speed of vibration through the floor medium and relatively low sampling rate, we assume that the readings from these sensors are synchronized.

Figure 2 shows the summary of the event detection module. The raw sensor data streams are first passed through a high pass filter to remove any offsets and low frequency noise from building vibrations. Then the system extracts a *signal energy* feature from the filtered data within a time window. The feature is defined as the sum of the squared values within the window. The window size is selected heuristically to be around one quarter of a second. This size is based on the fact that a human footstep exerts force to the floor for around two to three quarters of a second.[19] The selected window size is small enough to achieve a reasonable time resolution and to trigger multiple event detections for each SE for redundancy. At the same time, the window is large enough to prevent a high false alarm rate due to noise.

Falsely detected SEs or missed SEs may cause errors in SSE pattern recognition and result in incorrectly detected LEs and VEs. In order to have reliable SE detection, the system compares the summed energy to a noise model for anomaly detection, where an anomaly, signified as an increase in the feature value, indicates a SE.

The noise is modeled as a Gaussian distribution, whose parameters are continuously updated based on streamed non-event data. If the magnitude of the current window's feature is beyond three standard deviations (3σ) above the mean (μ) of the noise, the system considers it a potential SE. Note that the threshold of $\mu + 3\sigma$ reduces the theoretical false positive rate to below 1%. If the feature from the next window sequence also falls above $\mu + 3\sigma$ of the noise model, the system labels the potential SE as a SE.

When a SE is detected, the localization module is triggered to further characterize the event. Otherwise, the noise model is updated using the current data segment as shown in Figure 2.

Step Variations The features extracted from footsteps vary when the occupants walk differently (i.e. stomp their feet or wear different shoes). Since the system is for a building with diverse occupants, the system must be robust to these energy variations.

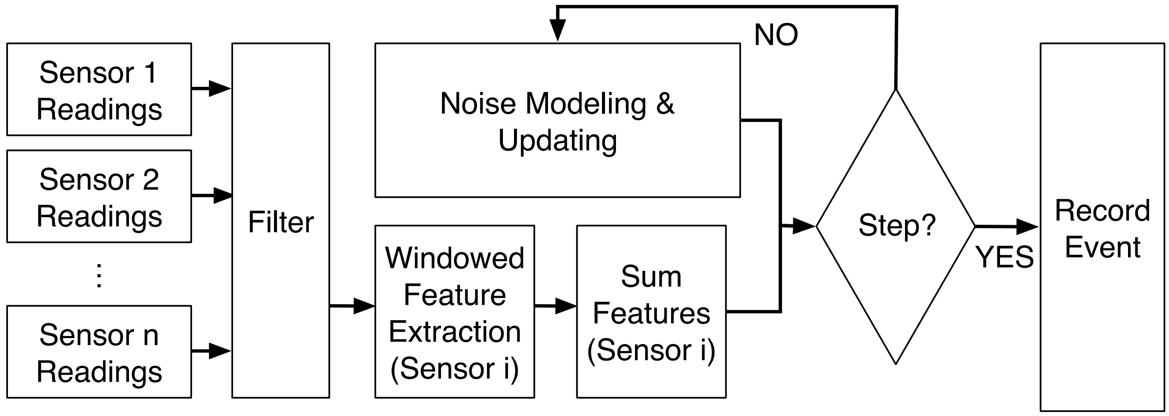

Figure 2: Event detection algorithm.

Figure 3a shows the floor velocity induced by the steps of a person wearing sneakers, and Figure 3b shows the floor velocity induced by the steps of a person wearing heels. The red dotted lines in the figures are the starting points of the detected SE windows. Comparing these two sub-figures, we observe that different kinds of shoes result in varying ranges of signal amplitude for their corresponding SEs. Our proposed detection algorithm can detect SEs in both cases, regardless of the variation.

These differences in shoes also result in different SE energies, as shown in Figure 3c and d. They represent the signal energy calculated from the SEs marked with the red dashed line shown in Figures 3a and b, respectively. We observe a clear difference between the energy levels in the SEs of these two shoe types with an average of four times difference. This suggests that the signal energy level and maximum floor velocity of SEs vary significantly between SSEs of different occupants. However, this variation can be calculated once the SSE we are analyzing contains the LE, which allows us to locate the SE.

3.2 Localization Module

The localization module combines multiple SEs into a SSE and then localizes the SSE using its LE, corresponding sensor location, and direction of travel. A LE gives the exact SE in the SSE that is closest to the corresponding sensor. Considering the one-dimensional hallway scenario, we can deduce the possible location of all the other SEs in the SSE with rotational ambiguity (i.e., we know how far each SE is from the sensor, but we do not know which side of the sensor each SE is on). The detected direction of the walk resolves that ambiguity and determines the locations of all the SEs in the SSE. In this section, we first characterize the variation in energy propagation. Then we present the sparse localization method (using LEs and walking direction) that is less sensitive to this variation.

Impulse Energy Propagation To obtain a model for the energy propagation we measured the received energy along a hallway at different distances. Figure 4 shows the relationship between the measured energy of uniform ball drop impulse events and the distance of that event from the sensor. From this figure, we can see that as the impulse's distance from the sensor increases, the energy of the event decreases with a rate of roughly $\frac{1}{distance}$.

Many existing indoor localization techniques rely on methods such as trilateration or triangulation.[21] While these approaches in theory will require three sensors with overlapping coverage to calculate the location, they only work for the overlapping coverage area of all three sensors. Moreover, our measurements (Figure 4) show a significant amount of noise that require a much higher number of sensors to reduce the noise.[21] This then severely limits the applicability of trilateration or triangulation, especially in practical scenarios where sensors are sparsely deployed. Therefore, in BOES, we present a LE-based method that enables localization in both dense (wide overlapping coverage) and sparse (little overlapping coverage) scenarios.

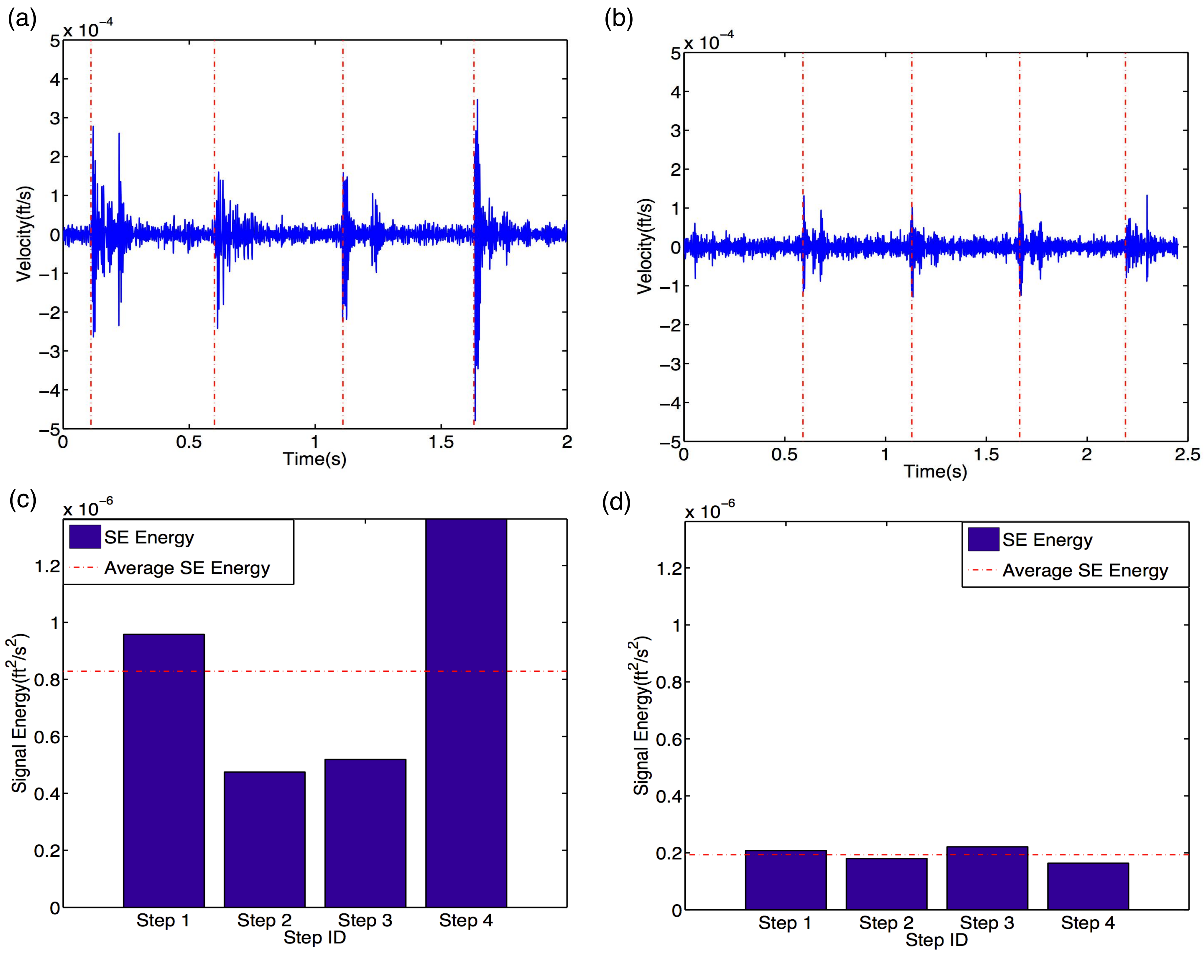

Figure 3: The signal of steps with different types of shoes varies in signal strength: (a) raw sensor readings of steps with high heels; (b) raw sensor readings of steps with sneakers; (c) energy of the SEs in (a); and (d) energy of the SEs in (b).

Landmark Event Detection Based on Figure 4, we deduce that when an occupant walks by a sensor, the energy levels of the SSE will gradually increase as the occupant approaches the sensor and decrease as the occupant moves away. The LE is defined to be the peak of the SE energies in the SSE, where the sign of the event energy change shifts. The detection of a LE only makes sense when the corresponding sensor's absolute location in the building is known. We register the deployment location of each sensor (assigned id: Sensor_i) in the system as SensorLoc_i. Once the peak is detected from the SSE of the Sensor_i, BOES marks the time and the event at the peak to be a LE for Sensor_i. We smooth the signal energy sequence of the SSE with a five-point moving average to avoid false detection of a LE due to human walking variations. The LE indicates that the location of the corresponding SE is closest to SensorLoc_i among the entire SSE. Therefore, once the LE is detected by Sensor_i, BOES can infer the absolute location of the SE with the registered location information SensorLoc_i. The LE-based location estimation may contain errors due to a mismatch between the sensor and the closest SE locations and variability in event energy. If one of the sensors has a high error rate in LE detection, the mismatch will occur only for a limited number of steps, since the locations will be updated when the LEs are detected from other sensors.

Walking Direction Estimation The walking direction eliminates the rotational ambiguity between two opposing paths in SE localization. When a SSE is mapped onto a one-dimensional hallway using LEs, we assume

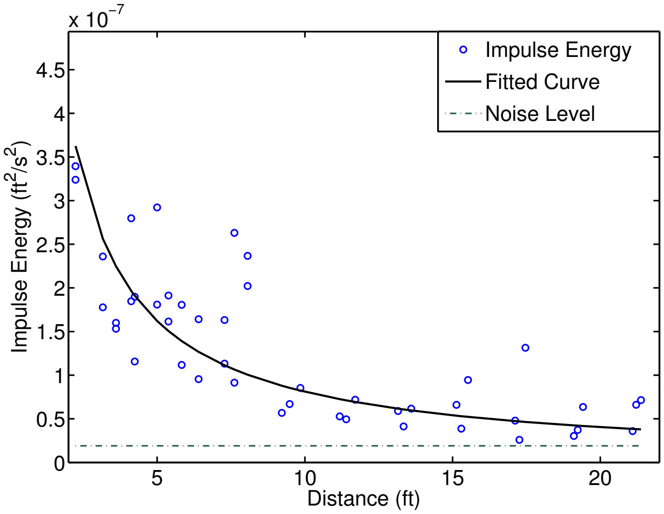

Figure 4: The relation between impulse energy and impulse-sensor distance fits the $\frac{1}{distance}$ decay. However, noise level is too high to effectively extract distance from energy.

that only the point where an occupant passes the sensor is accurately detected. In order to obtain walking direction and then localize all the other SEs, the system analyzes the LEs from multiple synchronized sensors. Since they are synchronized, the system can identify their SEs from the same footstep using the timestamp. When an occupant walks from one side of the hallway to the other side, the first sensor the occupant encounters is the one closest to the initial position. The walking direction will be consistent with the order in which the occupant passes by the sensors. This sorted list of sensors is determined by the order in which the LEs appear for the corresponding sensors.

By utilizing multiple sensors, we can also avoid false detection of LEs due to an occupant changing direction. For example, if an occupant walks from one end of the hallway towards the group of sensors but turns back before passing any of the sensors, the system will still observe a peak in event energy sequence due to the direction change. However, all the sensors will detect the peaks at the same timestamp instead of observing them sequentially, as described above for the true LE case. In this case, such event peaks can be identified as a non-LE.

In this study, we focus on scenarios where there is one occupant within the sensor coverage at a given time, so that SEs from different occupants do not mix. In the future, we plan to separate the SSEs of multiple occupants, which is out of the scope for the current work.

Once the footstep sequence is determined as well as the LE in the SSE, the system maps each SE to an absolute location on the floor plan. In order to map the previous SE to its location, we use an assumed human step length. The occupant's step length can be estimated from step frequency, which has been explored in related work.[22–24] In those works, the step length for a human is between 1.5 to 2.5 feet. Since the estimation of each step length is out of the scope of this study, we assume the human step sequence is consistent in terms of pace and step length. Therefore, we assumed a fixed step length of two feet for our system. With this step length, LEs, and walking direction, we can map the entire SSE to corresponding locations.

While the walking direction cannot be determined without some overlap in sensor coverage, once we have the walking direction, each SE in a SSE with a direction associated with it can be located in one-dimensional space with no additional information. This allows localization in sparse scenarios, where there is little overlapping coverage among sensors.

3.3 Tracking Module

The tracking module uses the localization results as well as the *velocity change event* (VE) to track occupants and count room-level occupancy. The VE is extracted from a SSE to detect an occupancy transfer between rooms within the monitored area. The VE marks the moments of occupancy change in each space by detecting the occupant entering another hallway/room.

When an occupant walks into a room from the hallway, we observe that the sequence of *maximum floor velocities* (MVs) of each SE experiences a sudden drop or disappears. The velocity refers to the floor velocity signals recorded by a sensor. This means the MV of the SE immediately prior to entering the room is much higher than the MV of the SE after entering. We observed a consistent pattern of sudden MV drop when an occupant enters a room from the hallway.

However, one single sudden drop of MV in a SSE may be caused by irregular walking patterns, such as stepping on the side of the heel or stepping on a hollow tile. In this case, the MV of the following SE will have a normal attenuated value compared to the SE before the single drop. In order to avoid such false detection of a VE, we developed a mechanism to verify the consistency of a sudden MV drop. We use a sliding buffer that sequentially holds the MVs for the most recent four SEs: $(MV_{t_1}, MV_{t_2}, MV_{t_3}, MV_{t_4})$, where MV_{t_4} is from the most recent step, and MV_{t_1} the least recent. We consider the case where a sudden drop of MV happens at time t_4 when MV_{t_4} is less than a half of the last two MVs (MV_{t_3} and MV_{t_2}). For example, when the system detects $MV_{t_3} \geq 2MV_{t_4}$ and $MV_{t_2} \geq 2MV_{t_4}$, it is considered as a sudden drop of MV. At this point, we mark the corresponding SE as a potential VE and wait for the next SE detection. When the next SE is detected at time t_5, the sliding buffer will update so that it now has $MV_{t_2}, MV_{t_3}, MV_{t_4}, MV_{t_5}$.

There are three possible scenarios once a potential VE is detected at time t_4:

1. The MV value increases again (i.e., $MV_{t_3} < 2MV_{t_5}$ and $MV_{t_2} < 2MV_{t_5}$), so the potential VE is a false alarm.
2. The MV value stays low (i.e., $MV_{t_3} > 2MV_{t_5}$ and $MV_{t_2} > 2MV_{t_5}$), and MV_{t_4} is confirmed to be a VE.
3. No event is detected within two seconds (considering normal human step frequency) after t_4. In this case, BOES considers the SE signal to have dropped out of detection range due to a crossing room event and marks MV_{t_4} to be a VE. If the drop is due to the occupant stopping in the hallway, the VE will be erroneously recorded. However, we assume stopping in the hallway is a short-lived event and will not affect the long-term accuracy of the occupancy count.

A "leaving a room" VE is detected the same way as an "entering a room" VE, except that we look for a sudden spike instead of a sudden drop in the MV.

As the final step, BOES updates its occupancy counts of rooms. The system sets the initial value to 0 and keeps an occupancy count for each room. Once a VE is detected, the system uses the location calculated for the associated SE to estimate the room where the VE happened. It then updates the occupancy count of that room. The occupancy count increases when an "entering room" VE is detected and decreases when a leaving room VE is detected. Because the number of occupants in each room must be a non-negative integer, the system will not record VEs that result in negative counting. In order to avoid cumulative errors, if no event has been detected for several hours, we reset the occupancy counts according to a fixed schedule that is selected based on the building usage pattern.

4. IMPLEMENTATION AND DEPLOYMENT

We implemented our system as discussed in Section 3 (also shown in Figure 1). The vibration sensing (Figure 1 first box to the left) is implemented on our sensing board. The rest of the sensing information processing modules (Figure 1) are implemented using MATLAB. We deployed our system in two types of buildings for performance evaluation. In this section, we introduce details of the implementation of the vibration sensing platform, the deployment of the system, and the design of the experiments for performance evaluation.

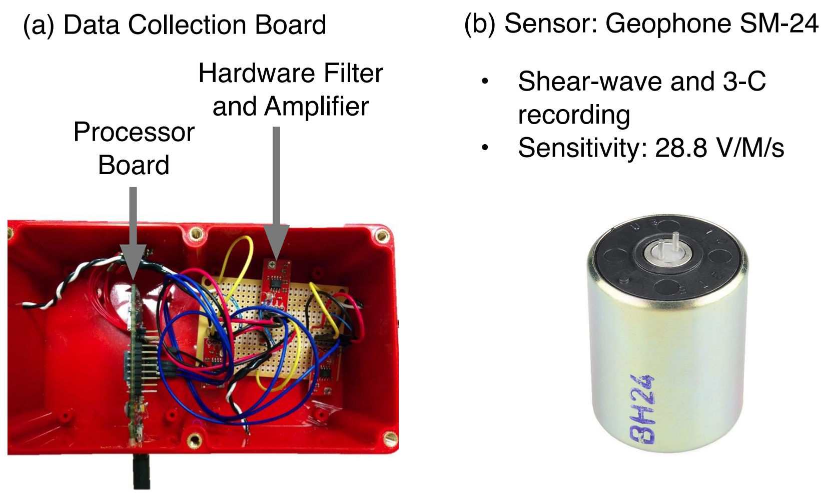

Figure 5: Hardware for data collection: (a) data collection board and (b) geophone sensor.

4.1 Vibration Sensing Platform

Figure 5 presents the data collection board and a geophone in our vibration sensing platform. In our sensing platform, multiple geophones (Figure 5b) are connected to the data collection board (Figure 5a). First, the geophone signals are fed through an amplifier and a first-order high pass filter. Then the amplified and filtered signals are read through the analog digital converter (ADC) on the data collection board to be stored and evaluated by the processor.

Geophone A geophone converts the velocity of ground movement into voltage to sense the vibration of the floor. In our system, we use a geophone SM-24 (Figure 5b) with the sensitivity of 28.8 V/m/s, a low cost analog sensor. With a properly designed amplifier it has enough sensitivity to recognize the vibration induced by human footsteps in buildings.

Amplifier and Filter Human footstep induced floor vibration has a very small amplitude with respect to the general geophone sensing range. Hence, the resulting voltage from geophones may be too small for the analog to digital converter (ADC) on the data collection board to accurately interpret (i.e., the vibration signal may be smaller than the resolution of the ADC). Therefore, BOES uses the amplifier shown in Figure 5a, which magnifies the incoming signals so that the vibration signals after digitizing have enough resolution for analysis. The amplification factor is 1000X so that the SE signals of both sneakers and high heels are amplified enough to be detected, while the strongest SE signals (from high heels that happen near the sensors) do not get clipped. Before feeding into the ADC, an anti-aliasing filter (a first order low-pass filter) is applied to the signal to filter out most of the high frequency signal. Based on existing study of the vibration and sound signatures of human footsteps in buildings,[25] the filter cutoff frequency is set at 100 Hz to fully capture the characteristic frequency band of the human footstep.

Data Collection Board The data collection board is based around the ATSAM3X8E.[26] Three geophones are connected to the data collection board. Since they are directly connected to the same board, they are easily synchronized for data collection. We use a sampling rate of 2 kHz. A sample rate that is too low will result in the alias of the unfiltered residue of the high frequency signal distorting the signal, so a high sampling rate of 2 kHz is used in order to avoid this issue.

4.2 System Deployments

We deployed the system at two different buildings, one commercial building and one residential building, and collected two types of data. First, we collected floor impulse responses caused by dropping a hockey ball from the height of three feet in a grid pattern to characterize the floor. Then floor vibrations due to human footstep

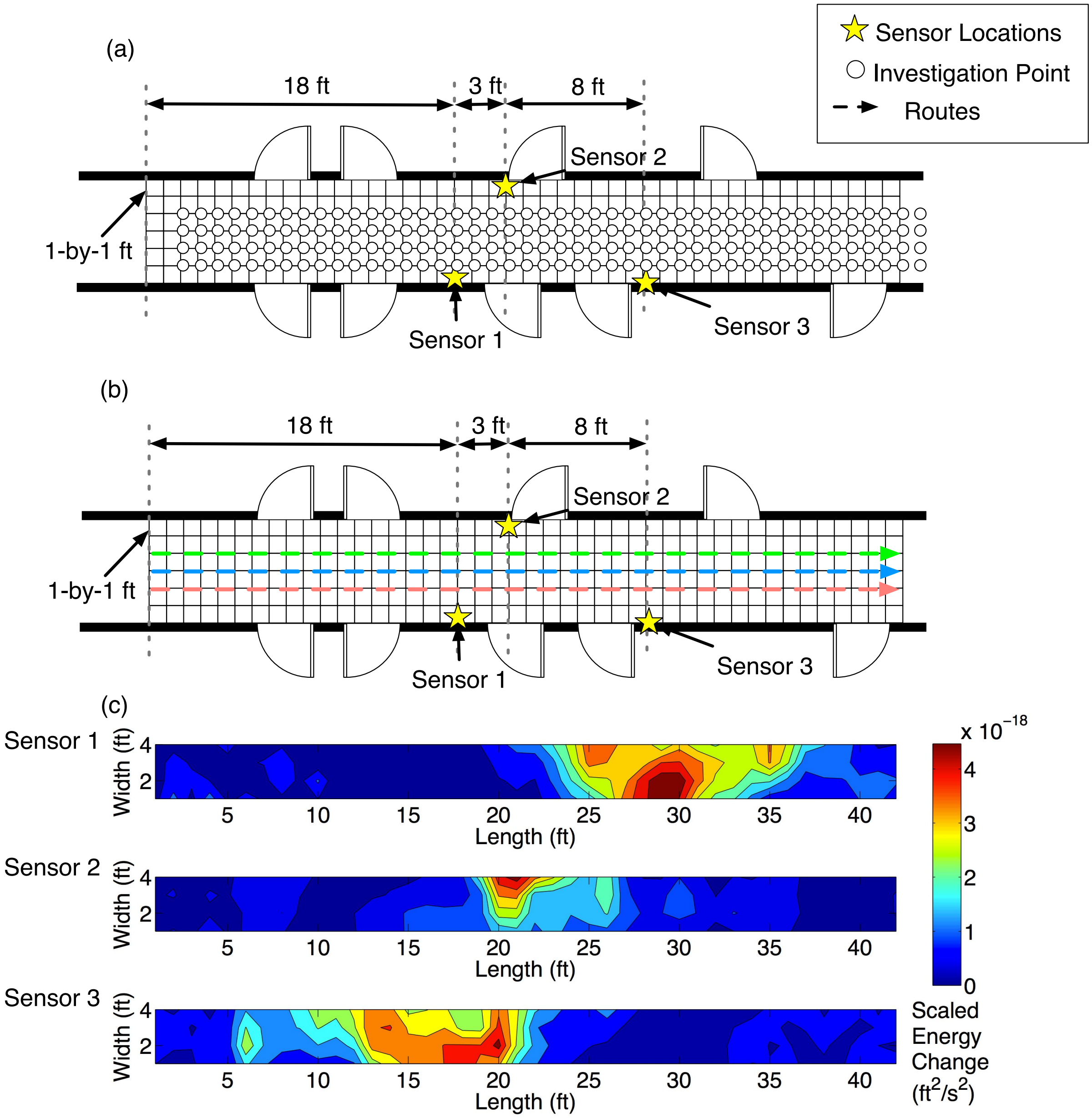

Figure 6: Hamerschlag Hall experiments: (a) shows the data collection points for the uniform impulse experiments; (b) shows the trajectories of the human footstep sequences; and (c) shows the heatmap generated by the uniform impulse experiments, and verifies our LE detection method.

sequences were collected for occupancy estimation and system validation. We define a *walking record (WR)* as a collection of SSEs that have been collected in sequential time, so the SSEs share no SEs. A WR can have any number of SSEs, including none (i.e., only noise is recorded).

4.2.1 Commercial Building: Hamerschlag Hall

We set up a vibration sensing platform for BOES system performance evaluation in a corridor in Hamerschlag Hall at Carnegie Mellon University campus in Pittsburgh. This building represents the structure of a large

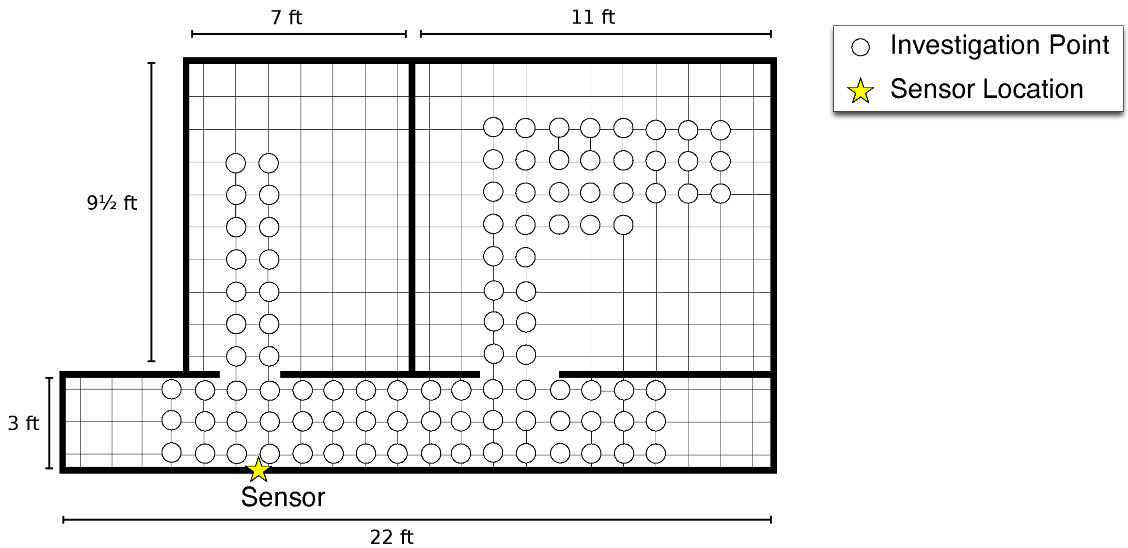

Figure 7: Townhouse experiment setup.

scaled commercial building. The floor plan and sensor deployment details are shown in Figure 6a. The space we investigated in the corridor is 6 feet by 42 feet. For sensor locations, we want to achieve the largest sensing coverage. In addition, we want the system to be particularly sensitive at critical locations, such as doors. Therefore, we placed two of the sensors along one side of the hallway about 10 feet away from each other, and placed the third one on the other side of the hallway. The sensor locations at the coordinates are marked as yellow stars in Figure 6a.

Impulse Response (Dataset 1) We collected floor impulse response data at each vertex of a one foot square grid on the floor in the hallway as shown in Figure 6a. The sensors collected multiple impulse events of dropping a hockey ball from the height of three feet. This data set is used for hallway characterization, described in Section 5.

Human Trajectory Along the Hallway (Dataset 2) We also collected floor vibration data from human footstep sequences in the building to evaluate the localization module. Each time a person steps we manually push a button to label the data with the step time as a ground truth. In order to record true step locations, we marked predetermined step locations on the floor. Figure 6b shows the routes we designed for occupants to walk in the experiments. For each route marked in Figure 6b, we collected floor vibrations for ten footstep sequences.

Human Trajectory from Hallway to Rooms (Dataset 3) These sequences were collected from a different part of the Hamerschlag hallway in order to test the generality of our system. The set up for these tests was near the entrance to the hallway, with the sensors placed along that part of the hallway in the same configuration as in our other system deployments. We used two rooms connected to the hallway: Room 1, a carpeted room at the end the hallway, and Room 2, a stairwell to the left of Room 1.

We collected WRs of two scenarios of occupancy change that involved multiple occupants:

- In the first scenario, one occupant walks into Room 1, and then a second occupant walks into Room 2.
- In the second scenario, two occupants walk into the same room (Room 1) in sequence, then one occupant walks out of the room.

We collected 10 WRs for each scenario. Each WR of the first scenario contains two SSEs, and each WR of the second scenario contains three SSEs. Therefore, our 20 WRs together contain 50 SSEs. 40 of the SSEs record

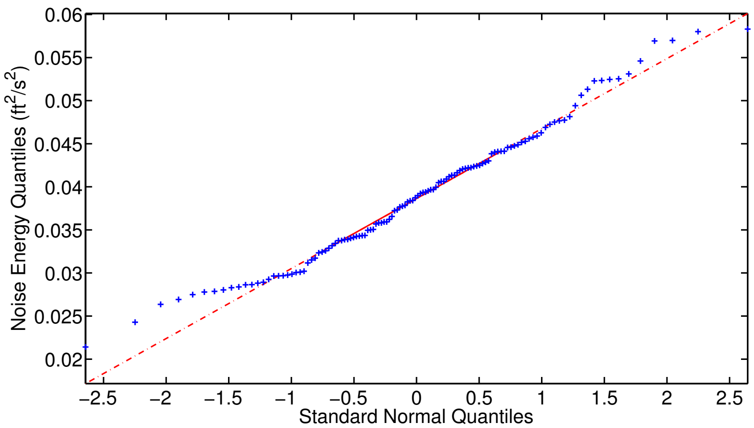

Figure 8: Q-Q plot of noise window energy sequence indicates a Gaussian distribution of the energy in noise windows.

an occupant walking into a room (20 from each scenario), and 10 of them record an occupant walking out of a room (from the second scenario).

4.2.2 Residential Building: Townhouse

We collected another set of data in a residential townhouse to validate the structural assumption underlying the VE of our tracking module. The data were collected in the hallway of the house as well as one bathroom and one kitchen, each branching off from the hallway. The floor plan of the investigated area and the sensor deployment are shown in Figure 7.

Impulse Response (Dataset 4) We collected the floor vibration due to the impulse events of dropping a hockey ball from the height of three feet on a 1 foot by 1 foot grid (Figure 7). This dataset is used to characterize the signals across multiple rooms, to analyze the extracted MV from each impulse and verify that an impulse measurement in a room adjoining the hallway will consistently show a smaller MV than an impulse measurement in the hallway. The floor plan of the house, location of the sensor deployment, and data collection points are shown in Figure 7.

5. EVALUATION

This section discusses the performance of the BOES. The three main modules (event detection, localization, and tracking) are evaluated in the following subsections. We evaluate the system based on experimental results.

5.1 Event Detection

We statistically analyzed Dataset 2 to obtain event detection accuracy. By comparing the detected SEs with the manually recorded ground truth, we computed the accuracy of the event detection algorithm for various locations of the footsteps. The number of total tested SEs is 660 (30 SSEs $\times$ 22 SEs/SSE). The accuracy was very high, about the same for all distances. The total number of correctly recognized SEs is 657. Therefore, the accuracy of step detection for Dataset 2 is 99.55%. The system also detected two events when no step event occurred. These false positives may have been caused by other activities, such as slamming doors nearby but outside our test area. The event detection accuracy did not vary much with the distance between a footstep and a sensor, within the tested area.

Noise Model Verification The event detection module assumes that the signal energy for noise is Gaussian. The Gaussianity of the noise signal energy samples was tested using the quantile-quantile (Q-Q) plot, which compares the sample quantiles with the standard normal distribution quantiles. Perfect Gaussianity results in the Q-Q plot being a straight line. Figure 8 shows that most of the samples lie on a line, which validates our

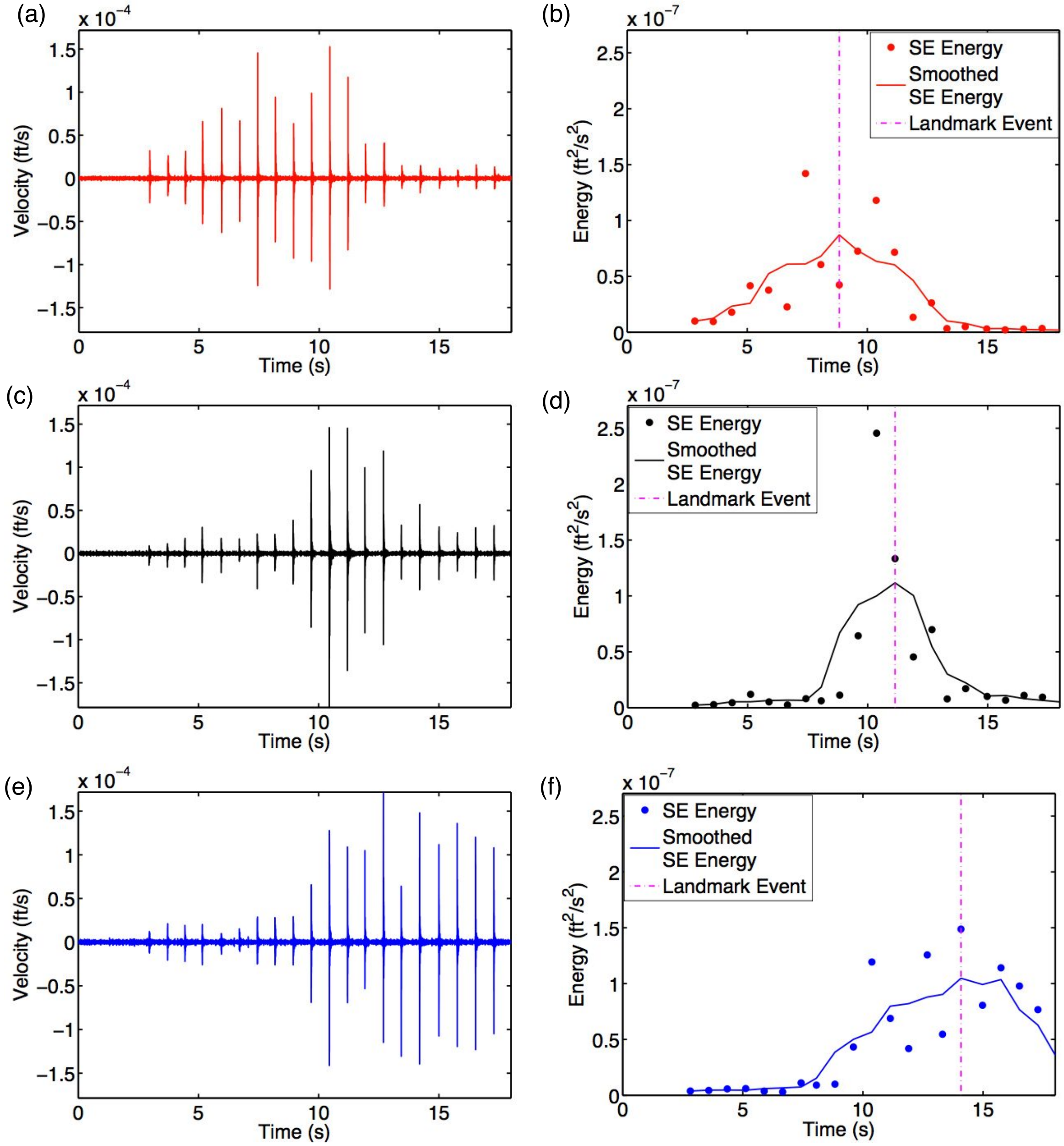

Figure 9: SSE and features sequence smoothing. The data is collected based on the deployment shown in Figure 6. (a),(c), and (e) show raw sensor readings from Sensor 1, Sensor 2, and Sensor 3, respectively. (b),(d), and (f) show the SSE extracted from the signals shown in (a),(c), and (e).

assumption. The samples fit the line much better in the center (e.g., [-1, 1]) and deviate towards the tails, which may be due to low sample size at the extremes.

5.2 Localization

The goal of this section is to assess whether the LE-based localization method can universally detect LEs and accurately localize SEs from detected SSEs. We used the data from Dataset 1 and Dataset 2 for this evaluation.

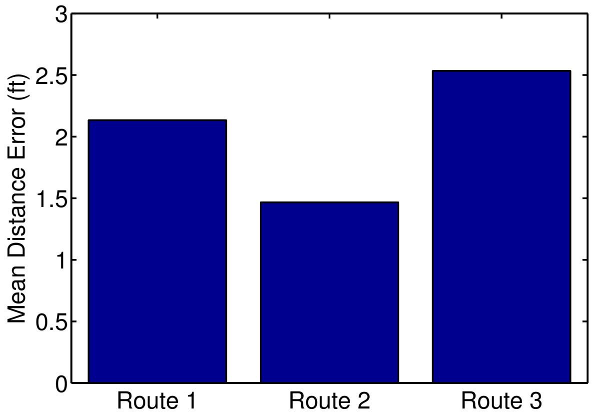

Figure 10: Accuracy of the detection of LEs of the detected footsteps when taking different routes in the hallway.

Impulse Response In order to verify whether our observations of SE energy change (i.e., closer events have higher energy) are universal, we analyzed the characteristics of floor vibration using data from impulses at different locations on the floor. Figure 6c presents heatmaps that were generated from data collected by different sensors at different locations across the floor (Dataset 1). For all three sensors, the heatmaps consistently show that the signal due to a closer impulse has a higher energy than the signal due to a further away impulse. This supports the universality of our observations on the event energy change with distance.

Human Trajectory LE Detection After verifying the characteristics of the floor with the impulse response test, we investigated the localization performance with human trajectory data (Dataset 2). Figures 9b, d, and f present the SEs detected from the data collected from Sensors 1, 2, and 3, respectively, which are marked as stars in Figure 6. The SEs are marked with solid dots, whose horizontal coordinates represent the times when the SEs occurred while the vertical coordinates indicate the signal energies for SEs. After detecting the SEs, the system smooths the sequences of SE energies, plotted as solid lines in Figures 9b, d, and f. The peak of the smoothed energy curve is defined as the LE, which is marked with a dashed vertical line in Figures 9b, d, and f. The system smooths the energy sequence to reduce the effects of noise.

Walking Direction Estimation Based on the detection results from the localization, we can deduce the walking direction of the SSEs. The order of the detected LEs indicates the direction of the footstep sequence. By comparing the order of LEs in Figures 9b, d and f, we observe that the LE from Sensor 1 appears first followed by Sensor 2 and finally Sensor 3. This indicates that the walking direction is from Sensor 1 towards Sensor 3, which matches the ground truth. We tested the walking direction estimation with the LEs detected using the 30 routes in Dataset 2 and observed 100% accuracy.

Human Trajectory Localization We quantified localization accuracy for our system by calculating the average LE location error. The estimated LE locations from Dataset 2 are compared with true locations.

Figure 10 shows that the average error of the LE localization is within three feet. In addition, the error did not vary much for different routes on the hallway. Potential sources of the localization error include:

1. Sensor location and the LE location do not coincide. For example, the sensor location happens to be between two SE locations. This is a system error inherent to the current method.
2. Inconsistency in an occupant's walking pattern may cause a variation in the SEs, which results in an erroneous LE detection.

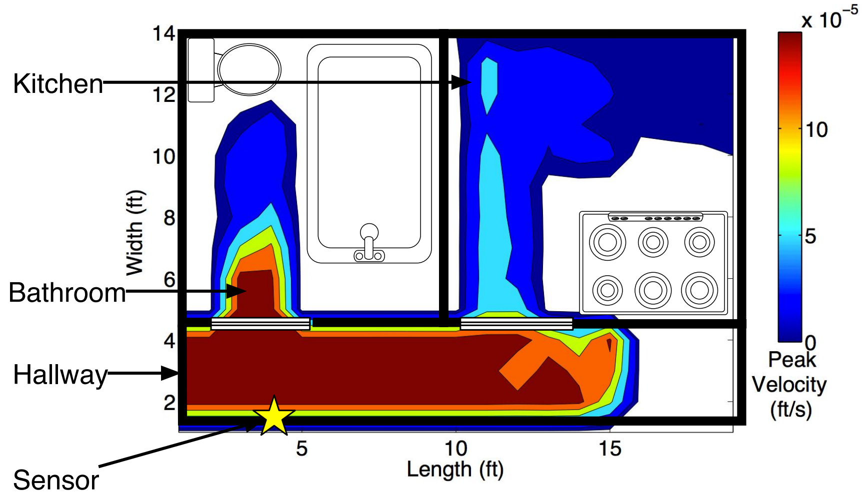

Figure 11: Impulse responses at a townhouse: the MVs detected from the data of the sensor indicate an observable sudden drop in amplitude of MVs in different rooms.

5.3 Tracking

Our tracking mechanism is based on the observation that when an occupant passes from one hallway/room to another hallway/room in a building structure, the maximum floor velocity (MV) within a SE suddenly changes (i.e., drops or rises).

Impulse Response In order to verify this, we conducted an impulse response experiment in a residential building, where we dropped a ball to apply impulse loads on the vertices of a square grid in each room (Dataset 4), as shown in Figure. 7. We investigated the hallway and two different rooms: a bathroom and a kitchen.

Figure 11 shows a heatmap of the MVs of impulses at different locations. We observe a significant amplitude drop when comparing the MVs in the rooms to the MVs in the hallway. In each room, we can see that this drop occurs no more than two feet from the entrance. This supports our assertion that room entrances mark a noticeable change in MVs, which our tracking model can use to detect an occupant passing between a hallway and a room.

Human Trajectory Tracking After verifying the sudden changes in MVs due to entering/leaving a space in the residential building, we further investigated the human footstep sequences of entering a room from the hallway at the Hamerschlag Hall. The goal of this evaluation was to estimate the VE detection accuracy. Figure 12 shows an example of VE detection. We ran our algorithm through the 20 WRs in Dataset 3, which contain 40 SSEs of entering a room and 10 SSEs of leaving a room. For the entering room routes, our system successfully detected 38 VEs (95% true positive). For the leaving room routes, the successful detection was 8 out of 10 (80% true positive). We also ran our algorithm through the 10 SSEs from Dataset 2, where the occupant passed by the group of sensors. The system falsely detected 3 entering events (30% false positive). This may be due to structural changes in the hallway that has a similar effect to the hallway-room structure connection. In addition, we observed a higher detection accuracy for entering a room compared to leaving. This may be due to an unintentional walking behavior change when leaving a room. It calls for different thresholds for the MV detection for leaving cases.

The occupancy count estimation results from the combination of location estimation and occupancy change estimation. After each detection of VE, we updated the occupancy count for each room.

We analyzed 20 WRs from Dataset 3 in order to test the accuracy of occupancy counting. A walking record detection is considered successful when all the VEs in the record are detected at the correct SEs. The footstep sequence analysis of the 20 walking records resulted in 17 correct occupancy estimation, achieving 85% true positives for occupancy counting.

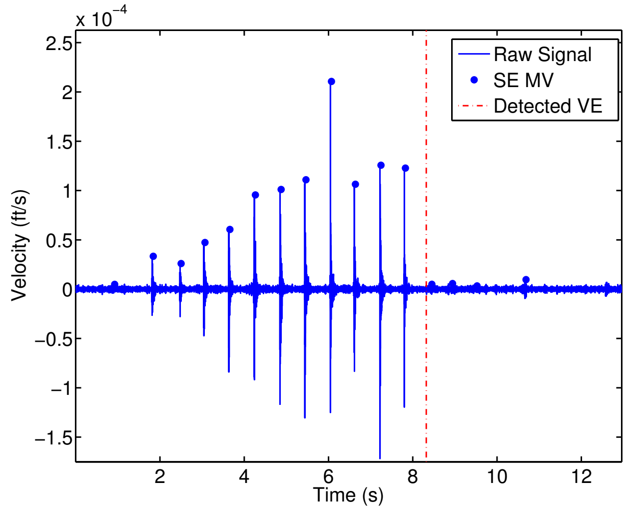

Figure 12: Observation of a sudden drop in signals when an occupant enters another room.

6. CONCLUSION

This paper introduces BOES, a building occupancy estimation system that utilizes sparse ambient vibration sensors. The system is designed to detect, localize, and track footsteps of an occupant, for the purpose of room-level occupancy estimation. The system first detects footsteps using an anomaly detection technique that identifies signal energy changes from the ambient noise model. It then localizes footsteps based on signal energy variations and known sensor locations. Finally, structural features and a sequence of localized footstep events are used to track and count occupants. Through validation in two different buildings, BOES achieved 99.55% event detection accuracy, located the footsteps within three feet, and obtained an occupancy count accuracy of 85%.

Although the estimation of the number of occupants moving as a group is not addressed in this paper, it is an interesting topic of exploration in the future. We plan to address the problem with a hierarchical approach: first investigate the level of noise in the signal to estimate the population level, and then to explore signal separation methods to estimate the number of separate sources of the signal and separate each source into a footstep sequence.

ACKNOWLEDGMENTS

This research was supported by the National Science foundation under awards CNS-1149611, funding from Intel, and the Nick G. Vlahakis Graduate Fellowship.

REFERENCES

[1] V. L. Erickson, S. Achleitner, and A. E. Cerpa, "Poem: power-efficient occupancy-based energy management system," in *Proceedings of the 12th international conference on Information processing in sensor networks*, pp. 203–216, ACM, 2013.

[2] G. Fierro, O. Rehmane, A. Krioukov, and D. Culler, "Zone-level occupancy counting with existing infrastructure," in *Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, pp. 205–206, ACM, 2012.

[3] D. Li, B. Balaji, Y. Jiang, and K. Singh, "A wi-fi based occupancy sensing approach to smart energy in commercial office buildings," in *Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, pp. 197–198, ACM, 2012.

[4] K. Gubi, "Roughmaps: Indoor positioning using existing infrastructure and symbolic maps," *Bericht, CHAI Group, School of IT, The University of Sydney* , 2010.

[5] R. Melfi, B. Rosenblum, B. Nordman, and K. Christensen, “Measuring building occupancy using existing network infrastructure,” in *Green Computing Conference and Workshops (IGCC), 2011 International*, pp. 1–8, IEEE, 2011.

[6] A. Bekkelien, M. Deriaz, and S. Marchand-Maillet, “Bluetooth indoor positioning,” Master’s thesis, University of Geneva, 2012.

[7] M. S. Bargh and R. de Groote, “Indoor localization based on response rate of bluetooth inquiries,” in *Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*, pp. 49–54, ACM, 2008.

[8] S. Feldmann, K. Kyamakya, A. Zapater, and Z. Lue, “An indoor bluetooth-based positioning system: Concept, implementation and experimental evaluation.,” in *International Conference on Wireless Networks*, pp. 109–113, 2003.

[9] P. Bolliger, “Redpin-adaptive, zero-configuration indoor localization through user collaboration,” in *Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*, pp. 55–60, ACM, 2008.

[10] A. Kamthe, L. Jiang, M. Dudys, and A. Cerpa, “Scopes: Smart cameras object position estimation system,” in *Wireless Sensor Networks*, pp. 279–295, Springer, 2009.

[11] W. Kahl and R. Settanni, “Lighting control system with infrared occupancy detector,” 1987.

[12] S. B. Gentry, J. F. Mazur, and B. K. Blackburn, “Method and apparatus for detecting an out of position occupant,” 1994.

[13] I. Pavlidis, P. F. Symosek, and B. S. Fritz, “Near-ir human detector,” 2002.

[14] S. K. Ghai, L. V. Thanayankizil, D. P. Seetharam, and D. Chakraborty, “Occupancy detection in commercial buildings using opportunistic context sources,” in *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*, pp. 463–466, IEEE, 2012.

[15] P. Orlewski, L. Federspiel, M. Cuddihy, M. Rao, and S. Fuks, “Advanced occupant detection system: Detection of human vital signs by seat-embedded ferroelectric film sensors and by vibration analysis,” in *National Highway Traffic Safety Administration, http://www-nrd.nhtsa.dot.gov/pdf/ESV/esv22/Session%203%20Written.pdf* , 2011.

[16] A. Nunes, M. Piedade, and R. Neves, “Cost effective immersive room with pressure sensing floor,” in *CD Proc. IMEKO Intl Conf. Cultivating Metrological Knowledge*, pp. 27–30, 2007.

[17] J. Sabatier and A. Ekimov, “A review of human signatures in urban environments using seismic and acoustic methods,” in *Technologies for Homeland Security, 2008 IEEE Conference on*, pp. 215–220, IEEE, 2008.

[18] A. Leon and A. David, “Microphone and geophone data analysis for noise characterization and seismic signal enhancement,” Master’s thesis, University of Calgary, 2009.

[19] H. Lee, J. W. Park, and A. S. Helal, “Estimation of indoor physical activity level based on footstep vibration signal measured by mems accelerometer for personal health care under smart home environments,” *Control and Instrumentation* , 2009.

[20] J. Hamilton, J. B.S., M. Kasarda, and P. Tarazaga, “Characterization of human motion through floor vibration,” *IMAC XXXII* , 2014.

[21] H. Liu, H. Darabi, P. Banerjee, and J. Liu, “Survey of wireless indoor positioning techniques and systems,” *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* , pp. 1067–1080, 2007.

[22] D. Grieve and R. J. Gear, “The relationships between length of stride, step frequency, time of swing and speed of walking for children and adults,” *Ergonomics* , pp. 379–399, 1966.

[23] V. Zatsiorky, S. Werner, and M. Kaimin, “Basic kinematics of walking: step length and step frequency: a review,” *Journal of sports medicine and physical fitness* , pp. 109–134, 1994.

[24] F. Danion, E. Varraine, M. Bonnard, and J. Pailhous, “Stride variability in human gait: the effect of stride frequency and stride length,” *Gait & Posture* , pp. 69–77, 2003.

[25] A. Ekimov and J. M. Sabatier, “Vibration and sound signatures of human footsteps in buildings,” *The Journal of the Acoustical Society of America* , p. 762, 2006.

[26] “At91sam arm-based flash mcu datasheet,” in *http://www.atmel.com*, Atmel Corporation, 2012.