Evaluating Emergency Evacuation Events Using Building WiFi Data

Iresha Pasquel Mohottige∗, Hassan Habibi Gharakheili∗, Arun Vishwanath†, Salil S. Kanhere∗ and Vijay Sivaraman∗

∗UNSW, Sydney, Australia
Email: i.pasquelmohottige@unsw.edu.au, h.habibi@unsw.edu.au, salil.kanhere@unsw.edu.au, vijay@unsw.edu.au
†IBM Research Australia, Melbourne, Australia
Email: arvishwa@au1.ibm.com

Abstract—Building operators are required to conduct periodic drills to ensure smooth evacuations in the event of emergencies. However, quantitative evaluation of the drill for adherence to building codes is largely manual and error-prone. Further, unplanned evacuations are seldom documented, let alone evaluated. This paper explores the use of building WiFi data for quantitative evaluations of both planned and unplanned evacuation events. We collect and analyze WiFi connectivity logs spanning a period of 180 days from 14 buildings in a large University campus. For our first contribution, we isolate WiFi data for known planned evacuation drills, conduct floor-level analysis to eliminate noise associated with transient WiFi connections or persistently connected devices, and highlight the anatomy of evacuations across multiple representative buildings each with differing number of levels, exit layouts, and occupant types. Armed with a detailed understanding of the anatomy of a planned evacuation, our second contribution develops a novel method to automatically identify evacuation events from WiFi data; we use it to detect 29 unplanned evacuations, and corroborate them against documented records where available. Our third contribution deduces quantitative measures to compare planned and unplanned evacuations, in terms of evacuation speed and occupancy levels, and further quantifies the man-days of productivity loss arising from unplanned evacuation events across campus. We believe our work is the first to show that building evacuations can be evaluated systematically and accurately at scale using WiFi data, both to corroborate current manual records and to gain new insights.

Index Terms—WiFi, Connection Logs, Evacuation, Data Analytics

I. INTRODUCTION

Building safety standards mandate procedures to be in place to quickly and safely evacuate occupants in the event of emergencies such as fires. In order to periodically test these procedures, evacuation drills are enforced, at least once per year [1], [2]. The standards require that these drills measure speed of evacuation (to ensure they are within a reasonable limit, say 30 minutes) and number of occupants evacuated (to ensure drills are conducted when the building occupancy is typical). Current methods to measure these are manual and error prone, especially for multi-level buildings with many hundreds of occupants and several exit doors. Further, unplanned evacuations are rarely documented or evaluated, leading to gaps in our knowledge of whether they are conducted with similar efficiency to planned drills, and the productivity loss they cause especially since a large majority of unplanned evacuations turn out to be false alarms.

WiFi connectivity data presents an opportunity to address the gaps above. Almost everyone in the world carries one or more mobile WiFi-connected devices, and WiFi connectivity is routinely logged by campus/building IT systems without any additional cost. Monitoring the WiFi connectivity patterns can therefore indicate when individuals are entering/exiting the building and/or moving across floors, which can be immensely helpful in evaluating evacuation events. We emphasize that our objective is not to use WiFi connectivity to guide an evacuation process as it is occurring, since the false negatives (people who do not carry a WiFi connected device) and false positives (people who leave their mobile device at their desk when they step out) can significantly confound evacuation personnel. Rather, our objective is to use the WiFi connection logs for a post-mortem analysis of evacuation efficiency. While WiFi connection logs have been used to study building occupancy patterns [3], [4], we believe detecting and measuring emergency evacuations pose new questions and challenges (outlined below) which have not been addressed to-date.

The most important measures pertinent to a planned evacuation are the speed (time from sounding of alarm to last person exiting the building) and occupancy count (in relation to the typical number of occupants housed in the building) [5]. These are measured manually today by a team of fire wardens and estate management personnel, who position themselves at various exits. A WiFi count based approach can not only corroborate these measures, but also offer richer insights into the evacuation process. It can allow estimation of finer-grained occupancy, for example, for individual floors within a building, and compare it to historical occupancy profiles. It can also help profile evacuation speed as a time evolution distribution rather than a single number, and identify bottlenecks at some floors due to elevators becoming unusable. However, the use of WiFi comes with many challenges – passers-by could be mistakenly counted as occupants, occupants may leave their devices connected at their desk or conversely get disconnected while using stairways, and devices may remain connected even after evacuation if the congregation area is within WiFi range. Overcoming these challenges requires us to come up with novel approaches/methods that are significantly different to...
those developed in prior building occupancy studies that use WiFi connectivity data.

For this study we obtained WiFi connectivity logs for 14 buildings spread around a University campus, over a period of 180 days that include weekdays and weekends but exclude holiday shutdown periods (appropriate ethics approvals were obtained for this study). Using four different buildings as representative examples, we first examine the anatomy of an evacuation from a WiFi perspective, showing how it varies by building size, number of floors, layout, and use (offices versus classrooms). We then develop a novel filtering-based method to automatically detect evacuation events from the WiFi data, distinguishing it from regular people movement patterns including the start/end of lectures and lunch breaks. Next, we apply our method to detect over 25 unplanned evacuations across 14 buildings, and validate our findings against building records where available. Lastly, we quantitatively compare planned versus unplanned evacuations, and estimate the productivity loss arising from unplanned evacuations (that largely tend to be false alarms). Our work establishes a framework and methodology for large-scale, systematic, low-cost and automatic evaluation of WiFi data to identify and quantify building evacuations, providing corroborative of current manual methods as well as new insights that do not require manual effort.

The rest of the paper is organized as follows: §II highlights the anatomy of an evacuation from a WiFi perspective and §III presents our novel method to automatically detect evacuations using WiFi data. In §IV, we quantitatively evaluate evacuations, while in §V we describe the prior work. The paper concludes in §VI.

II. PLANNED EVACUATION EVENTS FOR CAMPUS BUILDINGS

The university campus in this study consists of 53 buildings spread across a 38-hectare site, serving over 60,000 students and staff. The buildings accommodate a variety of activities ranging across teaching classrooms, study spaces, research labs, social areas, administrative offices, and housing units. The campus owns and operates a rich WiFi infrastructure consisting of over 5000 access points (APs) that serve authorized university users.

A. University Campus Buildings and Datasets

We begin by considering four representative buildings on campus, and to preserve anonymity refer to them by their quadrant location on the map. These include: F21 - the largest building on campus that is primarily used for study and meeting purposes; C22 - an office building; F23 - a building which largely accommodates teaching spaces; and J17 - a multi-purpose academic building comprising offices, lecture theaters, laboratories, and study spaces. Table I shows attributes of these buildings including primary use, number of floors, area size, and number of indoor WiFi APs.

For our study, we obtained two types of datasets: (1) building evacuation drill schedule and reports provided by campus Estate Management and (2) daily WiFi sessions logs of all campus APs (provided by the campus IT department). We note that appropriate clearance (UNSW Human Research Ethics Advisory Panel approval number HC190372) was obtained from the University ethics review board for this study. We further note that we anonymize user IDs and MAC addresses contained in the WiFi logs prior to storage/analysis by applying a one-way hash function.

Evacuation dataset: Campus Estate Management provided us with: (a) schedule of planned evacuations (aka drills) across all buildings on campus over the 6-month period of this study, (b) copies of drill reports from seven planned evacuations over this period, and (c) date/time of all unplanned evacuations that were recorded for the four selected buildings listed in Table I over the study period.

Evacuation drills, typically one per year, are planned by Estate Management at the beginning of each calendar year. Each drill is attended by a team of fire wardens, some of whom walk through various floors to clear people out while others are stationed at the various exits. After the drill, a report is filed that manually estimates the timing aspects of the drill, such as time of alarm, end of evacuation, and start of reoccupation, as well as number of occupants evacuated. These reports are analyzed to check if adjustments need to be made to the evacuation procedures to speed it up, and to make qualitative judgments on whether the building was at typical occupancy level during the drill. Our study uses these manual reports as ground-truth information to validate our WiFi based analysis of evacuation drills.

WiFi dataset: The University IT department provided us with: (a) data showing the physical mapping of WiFi APs to buildings and floor levels, and (b) daily session logs across the 5249 APs for a period of 210 days from Oct-2018 to May-2019, from which we redacted data for the Christmas holiday period (since campus operations are minimal) to get 180 days of usable data. Combining the two gave us a total of around 65 million WiFi session records; Table II shows a representative snapshot. Each record contains a unique User ID (note that we have hashed this to preserve anonymity); device MAC address (also hashed to preserve anonymity); a unique AP name that clearly indicates the building name, floor level, and access point ID; time at which the device associated to/disassociated from the AP (note that this is in minutes and therefore we do not have sub-minute accuracy); and avg throughput indicating data rate during the session. For example, the top session

| Table I | ATTRIBUTES OF REPRESENTATIVE BUILDINGS. |
| Building | Primary usage | # floors | Area size (m²) | # APs |
| F21 | Study and meeting | 14 | 4254 | 303 |
| C22 | Office | 5 | 2543 | 62 |
| F23 | Teaching | 16 | 2384 | 171 |
| J17 | Multi-purpose academic | 6 | 1410 | 101 |

| Table II | SAMPLE WiFi SESSION LOGS. |
| User ID | MAC addr | AP name | Assoc. time | Disassoc. time | Thput. (Kbps) |
| 154e7e26 | b76690ac | F21_Fi_APS1 | 31/01/2019 10:40 | 31/01/2019 11:15 | 6.1 |
| 154e7e26 | 12744897 | F21_Fi_APS2 | 31/01/2019 10:55 | 31/01/2019 11:20 | 8.6 |
| b6c72a33 | d6e17807 | F21_Fi_APS12 | 31/01/2019 11:15 | 31/01/2019 11:20 | 561.8 |
in Table II belongs to user 145e7e26 who connected from device b76690ac to AP1 located in building J17 floor F1, from 10:40am to 11:15am, and used an average throughput of 0.1 Kbps over the session. The second record represents the same user but with a different device 127d4fb7 connected to AP2 located on the same floor of the building from 10:55am to 11:20am, while the third record is a different user connected to a different floor of the same building, connected for 5 minutes with an average throughput of 561.8 Kbps.

**B. Inferring Building Occupancy and Evacuations from WiFi Traces**

Evacuations cause a sudden temporal change, i.e., drop followed by a rise in building occupancy, and we expect this to be reflected in the WiFi counts, which can be estimated for a point in time by counting the total number of unique user ids connected across all APs in the building at that time [3]. However, this method may overestimate building occupants as people walking past the building may get connected to an AP in the building for a short duration and thus wrongly be counted as occupants. To get a sense of such transient connections, we plot in Fig. 1(a) the CCDF of all session durations in our log, and observe that as many as 43.2% of sessions have a duration of less than 5 minutes. Indeed, when we plotted this on a per-building basis we found that buildings that are close to heavy foot traffic areas saw as many as 55% devices being filtered. Once transient devices have been pre-filtered out, we might still have transient sessions that could either be due to the device passing by the building (“passer-by transient”), or due to it moving between floors with potential intermittent connectivity during an evacuation (“evacuation transient”). Post-filtering to remove passer-by transients during the evacuation time-window will be discussed in §II-C.

Another aspect to consider is that building occupants may own multiple WiFi devices, e.g., laptop, tablet, phone. During an evacuation, an occupant may carry one device (e.g., phone) on their person while exiting the building, while leaving other devices (e.g., laptop) inside. Since our WiFi data includes both User-ID and MAC-address of the device, we are able to map devices to users – Fig. 1(c) shows the histogram of daily count of WiFi devices per user in a building. We observe that 75% of users own a single device, 22.5% have 2 devices, and a negligible fraction (2.5%) have more than two devices; the average number of WiFi connected devices per-user is 1.282. For our evacuation studies, we deem a user to have evacuated if any of their devices has exited the premise.

**C. Anatomy of an Evacuation**

Evacuations are triggered by the sounding of an alarm – either automatically due to detection of an emergency event such as a fire, or manually by a person in the case of a planned drill. Upon hearing the alarm all occupants are required to leave the building by the nearest exit route, typically under the coordination of building wardens. Usually a softer monotonic alarm (during what’s termed the “pre-evacuation time”) gives people time to prepare for evacuation, and this changes to a louder fluctuating alarm to force people out – this marks the start of the actual evacuation period, which is deemed to end when the building is fully vacated.
Building occupancy profile during planned evacuations:
In this section we study the anatomy of an evacuation using known planned evacuations. We begin by considering building F21 on campus which has 303 APs, and deduce its occupancy in terms of devices and users using the WiFi logs for the day when an evacuation was planned, as shown in Fig. 2. The total count of device ids (solid blue lines) and user ids (dotted black lines) is computed every minute (real-time count is sampled at rate of 1-min), and “passer-by” devices are pre-filtered as explained earlier. The evacuation drill scheduled for 10:30am is evident as a sudden drop in occupancy – the number of connected devices drops from 2633 to 937, while user ids drop from 1897 to 881 – and this is followed by a rapid rise as the building is reoccupied. Though this is as expected, there are some interesting observations that can be made from the WiFi occupancy plot:

(a) Even after the building is vacated, 937 devices remain in the building – we expect these to be largely laptop computers, though some smart-phones might also have been left behind (which might explain why the number of residual devices is slightly higher than the residual users). (b) The WiFi occupancy plot allows us to deduce the start/end times of the evacuation – the red inset box that expands the evacuation interval in the top left part of Fig. 2 annotates the start $T_s$ when occupancy starts falling leading to the dip, and the end $T_e$ when occupancy reaches its minimum before rising again due to re-occupation – thereby allowing us to measure the speed of evacuation. (c) Though the drill was planned for 10:30am, our plot suggests it started at 10:51am and ended at 11:03am. Indeed, the drill report provided by campus Estate Management recorded a start time of 10:49am, an end time of 11:01am, and evacuation duration of 12 minutes which corroborate with the time observed in our WiFi study.

Filtering passer-by transients prior to evacuation: As noted earlier, for each building, we are able to pre-filter devices that have a daily average session duration of less than 5 minutes in that building, since they are not occupants. In addition, while evaluating each evacuation event, we need to further post-filter sessions that are passer-by transients just prior to commencement of the evacuation. Specifically, devices that are associated with the building at $T_s$, but not at $T_s - 5$ minutes, are deemed to be non-occupant transients. Of the 2633 devices connected to building F21 at time $T_s$, 511 were found to be passer-by transient devices (indeed this building has high foot-traffic around it), and the remaining 2126 devices are deemed to belong to occupants of this building when the evacuation started.

Progression of evacuation: We now undertake a deeper look into how the evacuation progresses by tracking individual device sessions between time $T_s$ and $T_e$. In order to do so we categorize the different possible connectivity patterns experienced by occupant devices into three types as illustrated in Fig. 3. Type A devices consistently remain connected to the building WiFi throughout the event (i.e., from $T_s$ to after $T_e$) – according to our data, these devices are statically connected to a single AP, and hence become residuals (e.g., laptops); type B devices disconnect from the building WiFi prior to $T_e$ (red $\times$ in figure) – they may reconnect at different floors (green $✓$ in figure) as they proceed toward exit points, but eventually disappear from the building WiFi by time $T_e$, when they either leave the building or lose WiFi coverage (such as in emergency stairways); type C devices may generate successive connections during the evacuation (similar to type B), but they manage to reconnect before $T_e$ and remain connected even after time $T_e$, possibly because they are lurking within WiFi range of the building even after evacuation. For the planned evacuation of building F21, a majority of devices, i.e., 68% are of type B, while 25% are of type A and 7% of type C.

In order to determine when a device has evacuated, we take the following approach: type A devices do not evacuate; type B devices are estimated to have evacuated at the last red “$\times$” mark; and type C devices are deemed to have evacuated at the last green “$✓$” mark. This approach sets a lower bound on evacuation time for a device, with the upper bound obviously being $T_e$. This categorization of devices gives us deeper insights into the progression of the evacuation in terms of per-floor distributions and their trajectories, as shown in Fig. 4 for building F21.

In Fig. 4(a) we see how devices are distributed across the 14 floors at the start (black bars) and end (blue bars) of the evacuation. Note that blue bars correspond to residual devices at time $T_s$. The main entrance/exit doors are located at the 2nd floor of this building. At the beginning of the evacuation, it is
seen that a large fraction (i.e., 1091 out of 2126) of occupant devices are on floors 2, 3, and 4. By the end time $T_e$, there are varying number of residual devices across the floors, ranging from 12 devices at floor 8 to 83 devices at floor 2. This gives an idea of how many temporary versus permanent occupants are on each floor – the floors with the largest drops correspond to those with the largest number of temporary occupants, e.g., students using study spaces.

Fig. 4(b) helps us understand the floors from which disconnections happen; specifically, for each floor, the number of type A/B/C devices last seen on that floor is depicted. Considering floor 2 as an example, we see 83 residual devices which do not leave the building (i.e., type A shown by blue bar), 332 devices which last-disconnected from an AP at floor 2 (i.e., type B shown by red bar), and 57 devices which last-reconnected to an AP at floor 2 (i.e., type C shown by light-blue bar). It is not surprising to note that a majority of devices are deemed to evacuate at floors 2, 3, and 4, since the building has exits on levels 2. However, some devices are deemed to have exited at upper floors, probably due to poor or no-existent WiFi coverage in emergency stairways. Another observation is that the majority of type C devices are seen at floor 2, highlighting those users who walk out but stay within range of the building WiFi.

Lastly, we show in Fig. 4(c) the transition of individual devices (types B and C only) across floors between two points in time, i.e., start of evacuation and the time at which device type is declared. The x-axis denotes the timeline, the left y-axis indicates the floor at which a device is located at $T_s$ while the right y-axis indicates the floor at which the device is last seen during the evacuation. Each line represents the transition of a device – thicker lines indicate a group of devices. Note that this building has five stairways, two interiors with WiFi coverage and three exteriors (emergency) with poor/no WiFi coverage. We can make a number of observations: (1) devices are moving from upper floors to lower floors, with concentration at two lower floors 1 and 2 where main exits are located in this building; (2) focusing on devices that start from upper floors (e.g., floor 14), their last-seen floor varies, ranging from floor 8 (exit via emergency stairways) to floor 1 (used interior stairways and exited via main door); (3) a majority of devices that start from floors 2, 3, and 4 evacuated in less than 5 minutes as shown by the three thickest lines, while devices starting from upper floors (e.g., floors 12 to 14) need longer time (10-12 minutes).

Having understood the anatomy of a planned evacuation event in the largest building (F21) of the campus, we now look at the planned evacuation events in three other representative buildings on our campus (i.e., C22, J17, and F23) to appreciate the generality of our methodology.

**Planned Evacuations in three representative buildings:**

For each of the three buildings, we plot in Fig. 5 the building occupancy profile (device-based) for the day on which the evacuation drill (highlighted by the red box) is scheduled. Building C22 is a medium-sized building that accommodates about 600 people, as shown in Fig. 5(a). It is located next to one of the main campus gates at which students and staff typically access public transportation to/from city center, and hence one can observe several WiFi connections from outside the building, i.e., users who wait for buses on the street. Though this results in a heavily fluctuating occupancy profile during working hours, the dip caused by a planned evacuation in the afternoon, i.e., at about 2:30pm is clearly visible. Fig. 5(b) shows the occupancy profile of a fairly large academic building J17. The planned drill at around 11:30am is seen as a sharp dip (i.e., evacuation) followed by a sharp rise (i.e., full re-occupation). It is interesting to note that there is a significant dip followed by a rise between 5:00pm and 6:00pm, which is due to staff leaving work, and postgraduate students simultaneously arriving to attend evening classes that commence at 6:00pm – the distinction in profile will become significant when we automatically detect unplanned evacuations in the next section. Lastly, the teaching-focused building F23 in Fig. 5(c) shows fluctuations (of up to 250 devices) in occupancy profile at hourly boundaries (i.e., 1pm, 2pm, 3pm) since lectures begin/end at those boundaries resulting in students moving in and out of the building. Observe that the fluctuations in occupancy caused by these events are nearly half of that caused by the evacuation drill at 10am.

Table III summarizes devices count at $T_s$ (before and after post-filtering short-term passers-by), and distribution of static (type-A), disconnected (type-B) and reconnected (type-
validate our method using ground-truth data of six unplanned evacuations that have occurred in the four representative buildings. Form this table, we observe that a significant fraction of devices (267 of 738) in building F23 are passer-by transients, and are thus excluded; C22 has the highest number of static (type-A) devices highlighting the behavior of occupants in permanent offices who tend to leave devices inside the building during evacuation; devices that disconnect from building WiFi (type-B) dominate in all the four buildings with a varying share (49% to 68%); and about 10% of devices tend to stay in close proximity of the building after evacuation (type-C).

Having understood the detailed anatomy of planned evacuations in four representative buildings, we next develop a method to automatically detect evacuation events, especially unplanned evacuations that may go unrecorded.

### III. AUTOMATIC DETECTION OF EVACUATIONS

In the previous section, we analyzed evacuation drills for which we had well-organized ground-truth data. In this section, we focus on emergency evacuations that occur unexpectedly. Consequently, the ground truth data is often missing or incomplete. As explained earlier (§II-A), the campus Estate Management provided us with documented records of unplanned evacuations that have occurred in the four representative buildings where available.

The motivation of this section is to develop a method to automatically detect unplanned evacuation events based on the knowledge gained from the 14 planned evacuations. We validate our method using ground-truth data of six unplanned evacuations, and use it to detect 29 unplanned evacuations across campus.

#### A. Detection Method

We in §II-B showed that building occupancy profile displays a sudden drop followed by a steep rise during evacuations. At evacuations, sudden change in building occupancy introduce a high-frequency component reflecting movements taking place at timescales between 15-min and 45-min. This is evidenced by majority of planned evacuations – for example, building F21 (shown in Fig. 2) was evacuated in 12 minutes and re-occupied in 23 minutes. Consolidating data from the 14 planned evaluations, the evacuation duration lasts between 7-14 minutes and the time for re-occupancy ranges from 9-23 minutes. This implies that an evacuation completes between 16 and 37 minutes. Thus we choose 15 and 45 as lower and upper bounds of the timescale for evacuations. It is also important to note that there could be other higher frequency components corresponding to occupancy changes at timescales of less than 15-min (due to movements at boundaries of classes and/or seminars). Other typical movements within a building contribute to lower frequency components of the occupancy signal. Therefore, our objective is to isolate evacuation events by applying a band-pass filter to the daily occupancy signal \( X(t) \) of buildings. Given the above timescales of interest our band-pass filter would need cut-off frequencies at 15 and 15 minutes.

##### Choice of Band-Pass Filter:

In signal processing, for a band-pass filter there exists a wide range of options such as Butterworth, Bessel, Chebychev, Elliptic, or Savitsky-Golay (SG). Typically a filter with certain properties is employed depending upon the use-case and practical considerations such as computational complexity or real-time processing. Given
our frequency band of interest, we need a filter with the following properties: (a) maximally flat pass-band (preserving), (b) nearly zero stop-band (attenuation), and (c) sharp roll-off at cut off frequencies (transition). We show, in Fig. 6, the frequency response \( |H(f)| \) of five popular filters. We observe that Chebychev (solid green) and SG (dashed pink) filters do not meet the attenuation requirement due to significant ripples in their stop-band regions. Bessel (dashed pink) and Elliptic (dotted blue) filters fail due to undesirable behavior (fluctuating) in their pass-band. Butterworth filter (solid black) appears to satisfy all three conditions, though its transition is not the sharpest among these options. Therefore, we choose to use Butterworth filter which sufficiently satisfies our three specifications stated above.

**Energy of Filtered Signal:** By applying our band-pass filter to the building occupancy signal, we preserve the information of potential evacuation event. Note that a sudden change in occupancy signal (due to evacuation) would yield high energy. We, therefore, quantify the energy (computing root-mean-square energy) of the filtered signal to determine if it contains an evacuation event or not. Let \( X_{BP}(t) \) be the resulting band-pass component of \( X(t) \) and \( N \), the size of the moving window. The root-mean-square energy (RMSE) is computed by:

\[
RMSE(t) = \sqrt{\frac{1}{N} \sum_{k=-N/2}^{k=N/2} X_{BP}(t + kT)^2}
\]

where \( T \) is the sampling rate which is a minute (as explained in §II-C). We have tuned the size of moving window to maximize the RMSE for planned evacuation events and found that the window size 5 gives the best results. In Fig. 7, we illustrate the process of isolating evacuation event from a building occupancy signal. It can be seen that the time-trace of RMSE, in Fig. 7(c), displays a “sharp” peak corresponding to evacuation planned before noon in building F21. We note that the height of this peak (arising from an evacuation) varies across buildings.

**Setting Threshold for Energy:** In order to identify a sharp peak (corresponding to evacuation events) in RMSE signal we compute the ratio of peak-to-mean (PMR) – for example, the PMR of the planned evacuation shown in Fig. 7(c) is

\[22.71 = \frac{421.7}{18.5} \]

A high PMR value results from high energy of occupancy change (due to evacuation) compared to average energy of minor changes that occur in a building over the course of a day. Note that, as per the earlier discussion, we employ a band-pass filter to remove the low and high frequency components. By analyzing the PMR for the 14 planned evaluations, we found that a threshold value of 18 seems suitable for detecting unplanned evacuations. Our selection of threshold value \( \theta = 18 \) is validated by the detection of six ground-truth unplanned evacuations (detailed in §III-B).

**Estimating \( T_s \) and \( T_e \):** Let us now go further and identify the start and end times \( (T_s \) and \( T_e) \) of the detected evacuation. We begin by time \( T^* \) at which the peak of RMSE is observed. Given \( T^* \), we search backwards and forwards on the corresponding occupancy signal \( X(t) \) to determine time \( T_s \) (when occupancy starts falling) and \( T_e \) (when occupancy is flattened) for the detected evacuation.

**Summary:** Given a building daily occupancy signal \( X(t) \), we first obtain the band-pass filtered signal \( X_{BP}(t) \), and next construct the RMSE from the filtered signal. Then, we compute the PMR value and check it against the threshold \( \theta = 18 \). The PMR \( \geq \theta \) detects an evacuation for which we can deduce \( T_s \) and \( T_e \). In next subsection §III-B, a detailed evaluation of our method is presented.

**B. Detecting Unplanned Evacuations**

We now evaluate the efficacy of our method by applying it to data from the four representative buildings. Table IV shows the number of unplanned evacuations in each of these buildings. The second column is our ground-truth of unplanned evacuations obtained from campus Estate Management – numbers in square brackets indicate events that occurred during atypical time of building occupancy (i.e., late night after 10pm or early in the morning before 5am) which is out of the scope of

<table>
<thead>
<tr>
<th>Building</th>
<th>Number of unplanned evacuation events</th>
<th>Ground-truth</th>
<th>Auto detected</th>
<th>Auto detected with ground-truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>F21</td>
<td>2[+]3</td>
<td>6</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>C22</td>
<td>2[+]3</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>F23</td>
<td>2[+]1</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>J17</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
At time 2:50pm on 4-Mar-2019.

At time 2:56pm on 25-Mar-2019.

At time 2:36pm on 30-Apr-2019.

Figure 8. Auto-detected emergency evacuations in building F21 (without ground-truth records).

At time 1:51pm on 19-Feb-2019.

At time 1:52pm on 4-Mar-2019.

At time 2:56pm on 25-Mar-2019.

At time 3:02pm on 29-Apr-2019.

Figure 9. Emergency evacuations in building C22: (a) ground-truth recorded, but not auto-detected, and (b, c, d) auto-detected without ground-truth records.

Figure 10. Auto-detected emergency evacuations in building F23: (a) incorrect record by EM, and (b) auto-detected without ground-truth.

Figure 11. Auto-detected emergency evacuations in building J17 (without ground-truth records).

campus Estate Management does not have a record. There are prominent dips in all four plots supporting our truly detected evacuations. In C22, we automatically detect 4 emergency evacuations of which only one is captured by the ground-truth dataset. The second ground-truth emergency event is missed by our method. To further investigate these discrepancies, we plot in Fig. 9 the occupancy profile of building C22 on four days, i.e., the day with recorded emergency but not detected as well as three detected ones but not recorded. Fig. 9(a) shows the building daily occupancy on 22-Jan-2019 for which an emergency is recorded at time 5:05pm (highlighted by vertical red dashed line). Even though there is a presence of a sudden dip at the recorded time it is not followed by a re-occupation, and hence is not detected by our method. C22 houses administrative spaces of which staff typically would start leaving work to go home at 5pm. Therefore, we can assume C22 building occupants decided to finish their work at this end-of-day emergency thus not reoccupying the building (no sharp peak in energy is observed). Moving to Figures 9(b), 9(c) and 9(d), a visible dip in occupancy in each of these three plots (afternoon events highlighted by red boxes) suggest that these are likely evacuation events.

Considering building F23, it appears that 1 out of the 3 detected emergency evacuations corroborate with ground truth Estate Management records. There are 2 ground truth emergency evacuation records with Estate Management and our method does not detect the other record of emergency on 27-Feb-2019 at 2:32pm. We in Fig. 10 plot the building occupancy to analyze the discrepancies. The building occupancy on 27-Feb-2019 is shown by dotted black lines in Fig. 10(a) – we see that the occupancy profile seems normal with no significant dip at 2:32pm. Instead, we detect an emergency next day (28-Feb-2019) at about same time (2:42pm). In the building occupancy for this detected day shown by blue lines in the same Fig. 10(a)
we can visually confirm a sudden dip followed by a steep rise occurring in the afternoon. There has been an error in the date recorded for the evacuation. This clearly highlights the possible human errors in manual Estate Management records. The other evacuation detected for F23 from our method with no Estate Management record is shown in Fig. 10(b) and as highlighted by red box there is a likely evacuation. Lastly, for building J17, automatic evacuation detection by our method corroborates the ground-truth evacuation while detecting 2 additional emergency evacuations. The two unaccounted evacuations are fairly obvious in the building occupancy profile for 28-Feb-2019 as shown in Fig. 11(a) and for 4-Mar-2019 as shown in Fig. 11(b).

In summary, it is seen that our method misses one (out of a total of 7) unplanned evacuation in ground-truth Estate Management records available for the four representative buildings. Additionally, for these four buildings we detect 10 more evacuations which are not recorded by campus EM.

Auto-detection across 14 buildings: We extended our automatic detection to all 14 buildings across our university campus. In total, we detect 29 emergency events in 14 buildings (the four representative buildings included). We detect 4 unplanned emergencies in F25, 3 unplanned emergencies each in H20 and G14, and 2 unplanned emergencies in C20. Our method did not detect any emergency evacuations in the remaining five buildings K17, F12, K15, M15 and J18.

In the next section, we deduce quantitative measures to compare planned and unplanned evacuations detected by the method we developed in this section, and then quantify the loss of productivity caused by evacuation events across our campus.

IV. EVALUATING EVACUATION EVENTS

Having understood the anatomy of a planned evacuation event in §II, and developing a method to automatically detect the evacuation events in campus buildings in §III, we now quantitatively evaluate planned and unplanned evacuation events. We mainly focus on three metrics: speed of evacuation and count of evacuees, and additionally for planned evacuations we quantify typicality of occupancy level. Then compare the planned evacuations with unplanned evacuations. Lastly, we compute the loss of productivity arising from interruptions to day-to-day work during evacuations.

A. Planned Evacuations

We evaluate planned evacuations using three metrics: (a) speed of evacuation indicates how quickly an evacuation is executed – this metric is expressed as the time difference between $T_s$ and $T_e$; (b) evacuee count indicates the total population of the evacuees – it is measured by change of occupancy from $T_s$ to $T_e$; and (c) typicality of building occupancy for an evacuation – this is the percentile level at which the building is occupied at start of evacuation compared to its historic occupancy count.

**Speed of Evacuation:** There is no standard specification for a minimum evacuation time, but it is recommended to evacuate a building within 30 minutes of an emergency, since building materials can resist a fire up to 30 minutes [6]. We note that minimizing the time for evacuation is key for safety of building occupants, which is the reason drills are conducted.

The drill report dataset from campus Estate Management records the time (to minute resolution) for the start and end of evacuation, which lets us corroborate the recorded evacuation speed against the one we obtain from the WiFi occupancy profile (i.e., $T_s- T_e$). Table V compares our method against the drill report for the four representative buildings. The match is nearly perfect, being identical in three of the four buildings (F21, C22, and F23), differing by a minute for building J17. This gives us confidence that evacuation speed automatically measured from WiFi occupancy is as precise as manual measurements. Focusing on the speed of individual devices, we plot in Fig. 12 the CCDF of devices evacuation time in our representative buildings. It is seen that the speed of devices displays a minor variation (less than two minutes) in three buildings C22, J17, and F23. In F21 (the largest building), however, we see a 3-minute gap in the CCDF. This is because devices that start from higher floors are likely to use emergency stairways, and hence disappear from WiFi logs for a few minutes till their last-seen time.

**Count of Evacuees:** In Table VI we show the number of people evacuated in drills of representative buildings, comparing our method (two last columns) with manual drill
The evacuee count reported in the drill dataset is roughly estimated by several fire wardens stationed at various exits. We compute the count of evacuees in two ways: one is by taking the difference of raw building occupancy (users count) at times $T_s$ and $T_e$ without post-filtering (third column), and the other one is by counting total users of types B and C after post-filtering passer-by transients (the last column) as we discussed in §II. Note that users’ type is obtained from the type of their associated device which is last declared between $T_s$ and $T_e$ – in other words, we map devices to users. We note that the difference count of users (i.e., # diff. users) is not very accurate because: (a) it does not account users with multiple devices (a user who indeed evacuated but has a device left inside the building, is not counted as an evacuee!), (b) it uses the raw occupancy data (unfiltered) containing passer-by transients. Therefore, we observe that count of B&C users is larger than count of user difference in three buildings F21, C22, and J17, except in F23 which has a significant number of passer-by transients as explained in Table III.

Typicality of Evacuation: Regulations for building fire safety [7] recommend that fire drills be conducted with “appropriate number” of people inside a building. This is indeed merely a qualitative measure. We, instead, are able to quantify this measure using the metric of typicality. This quantification can help specify measurable guidelines for scheduling drills. Fig. 13 shows the CCDF of building occupancy during the period of our dataset. We mark (red ×) building occupancy at start of its planned evacuation – note that occupancy count is obtained from the data without post-filtering passer-by transients. We can see that typicality is fairly high in percentile (seemingly appropriate) across four buildings with F21 at 89%, C22 at 93%, J17 at 95%, and F23 at 89%.

Evaluating 14 Planned Evacuations: We now evaluate the planned evacuations for all the 14 buildings of our study. Our results show that: (a) building evacuation speed varies between 6 to 20 minutes with an average of 11 minutes. We found that it takes longer to evacuate larger buildings, except in residential buildings. We think this is because occupants of campus dorms tend to resist evacuation until they are forced out, especially when they know it is an evacuation drill; as a result evacuation events are always slower (takes longer time) in residential buildings, both large or small size; (b) count of evacuee varies by building size, but more than 50% of occupants (at $T_s$) evacuated across all buildings – a significant fraction of single-device owners leave their device behind as they expect to come back to their desks soon; and (c) typicality ranges between 36% and 99% with an average of 85% – except two residential buildings (36% and 47%), others seem relatively typical.

B. Comparing Evacuations: Drills versus Emergencies

We now analyze emergency evacuations and check how do they differ from drills using two metrics, namely speed and evacuee count. The typicality metric is of interest only to drills (i.e., planned evacuations). Though we analyze data for 14 buildings, we only detail results for 4 representative buildings.

Among the remaining 10 buildings there were no emergencies in 5 and for the rest, we provide a summary of the results.

Speed of Evacuation: We found that emergencies are generally slower than their corresponding building drill, except in two events: Four of the emergency evacuations in F21 took longer to evacuate than the drill spending 14, 17, 20 and 23 minutes whereas the drill was completed in 12 minutes. Other two emergencies spend 12 minutes each, similar to the drill; in J17, the three emergency evacuations were on average half the speed of the drill which took 6 minutes; in C22, four emergencies on average were 65% slower than the drill (9, 11, 13 and 15 minutes for emergencies versus 7 minutes for drill); in F23, three emergencies on average were 3 times slower than the drill. We think drills are carried-out faster than emergencies because for planned events the team of fire wardens are well-organized and prepared so that they are able to force people out for a faster evacuation.

For the five buildings (H20, G14, F25, B16 and C20) where we detected emergency evacuations, the emergency evacuations were slower than the planned evacuations.

Count of evacuees: Emergencies can occur at any time in a building (at low or high occupancy), and hence count of evacuees may not be very insightful when comparing emergencies with drills. We, therefore, compute the ratio of evacuees count to the building occupancy at start of evacuation.

Our results in F21 show that evacuee ratio was 64% for the drill, and it varies from 35% to 70% for the six emergencies. This is not surprising since the occupancy of this building is highly variable depending on teaching and non-teaching periods. For all three emergency evacuations that took place in teaching space F23, evacuation ratios (55%, 61% and 64%) are found to be higher than the evacuation ratio (51%) of the drill. For the office building C22, it is found about 50% evacuee ratio during drill, and average of 45% for emergencies. We do not see a major variation of this ratio for planned and emergencies in C22 because the composition of the building occupants (only staff) persists across the year. Lastly, the highest evacuee ratios were found in academic building J17 (69% for drill and 68%, 74% and 84% for the 3 emergency evacuations).
For the rest of five buildings with detected emergencies, G14, F25, B16 and C20 showed higher evacuee ratio during emergencies than drills. For H20, all three emergency evacuations displayed on average 12% lower evacuee ratio than the drill. Overall, we note that the evacuee ratio for planned and unplanned evacuations for various buildings can vary depending on factors such as type and composition of occupants in the building, and the nature of emergency.

C. Loss of Productivity

Evacuations (both planned and emergency) cause interruption to normal activity of building occupants. While people safety outweighs financial losses, repeated evacuations especially when caused by false alarms can incur heavy cost due to loss of business hours. Therefore, it is recommended to minimize asset and revenue loss when planning for drills [1].

We quantify the loss of productivity (human-hours) by taking the product of two metrics, namely loss of business hours and number of re-occupants. Loss of business hours is computed by adding time to evacuate and time to reoccupy. It is important to note that we do not consider productivity loss for those evacuee who do not return to the building, since we assume they are engaged in other useful activities (instead of waiting). The number of re-occupants can be computed by simply taking the change of building occupancy level during re-occupation process.

We note that for drills, the building is reoccupied fairly quickly after it is completed. However, for emergencies it can take longer time to first identify and mitigate the cause of emergency alarm, before occupants are allowed to enter the building. According to our analysis, reoccupation process can take up to twice the amount of time needed for evacuation. We, therefore, estimate an upper-bound value $T_e + 2(T_e - T_s)$ for the reoccupation time ($T_r$, i.e., time at which building is back to its normal operation). Recall that $T_e - T_s$ is time required to evacuate. We can obtain the exact reoccupation time if the building occupancy at $T_e$ is equal to occupancy at $T_s$. Note that we use the upper-bound estimate if the exact $T_r$ is not obtained.

In Table VII, we show the loss of productivity for all evacuations (drills and emergencies) in our representative buildings computed in “MAN-DAYS”. It is important to note that total loss of productivity would be proportional to building size and frequency of events. For planned evacuations, the largest building F21, results in the largest loss and the smallest building C22 results in the smallest loss. It is also evident that several emergencies in F21 results in significant losses.

<table>
<thead>
<tr>
<th>Building</th>
<th>productivity loss (drill)</th>
<th>productivity loss (emergencies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F21</td>
<td>68.5</td>
<td>161.5</td>
</tr>
<tr>
<td>C22</td>
<td>6.9</td>
<td>16.5</td>
</tr>
<tr>
<td>F23</td>
<td>7.3</td>
<td>39.7</td>
</tr>
<tr>
<td>J17</td>
<td>21</td>
<td>49.8</td>
</tr>
</tbody>
</table>

Table VII
LOSS OF PRODUCTIVITY (HUMAN-DAYS).

Estimating occupancy, specifically based on WiFi activity, has received considerable attention in the literature. They can broadly be classified as follows.

Occupancy estimation based on WiFi signal strength: Using WiFi received signal strength between a pair of stationary transmitter/receiver antennas, [8] developed a framework to count the total number of people walking in an area. The experimental results showed that the proposed approach has good accuracy. Other techniques that use only WiFi received signals to track motion and detect occupancy behind walls and inside buildings are proposed in [9]–[11]. The advantage of these approaches is that they do not rely on users carrying their WiFi enabled devices, such as smartphones or laptops, to determine occupancy.

Occupancy inference using WiFi connection information: A practical system for accurate occupancy estimation based on commodity WiFi infrastructure is developed in [12]. The work captures the MAC addresses of users’ WiFi devices as they communicate with the APs, which is analyzed to estimate occupancy. By analyzing WiFi signatures from mobile phones carried by users, ARIEL [13] presents a system that is able to identify the room a mobile phone (or user) is in with over 95% accuracy. Linear regression and support vector machine based techniques are applied to the received signal strength at WiFi APs to estimate occupancy in [14]. Using a combination of number of WiFi devices, electrical energy demand and water consumption, [3] showed 48% improvement in accuracy for occupancy estimation compared to using only WiFi data. Similarly, [15] propose mechanisms to infer occupancy in a building by combining information from WiFi APs, badge access and other auxiliary data sources such as calendar schedules and instant messaging clients. While all of the above works make important contributions, they do not apply occupancy estimates to tackle broader problems, such as reducing building energy consumption or evaluating building evacuation in an emergency.

WiFi for energy management in buildings: There is an emerging body of work that specifically uses WiFi-based occupancy information to reduce the energy consumption of buildings, in particular, of that consumed by its heating, ventilation and air conditioning (HVAC) systems. We provide a brief overview of some closely related work next. The work in [16] showed the potential of using existing IT infrastructure in buildings, including WiFi APs, to lower the building energy demand. Occupancy is implicitly obtained by tracking MAC and IP addresses at these APs, which in turn can be used to direct HVAC and lighting to only the occupied zones, thus saving energy. A practical system, which infers occupancy using WiFi and uses it to control the HVAC of a commercial building leading to 18% reduction in energy consumption was demonstrated by Sentinel in [17]. Learning the spatial occupancy patterns enabled by mobile WiFi connection logs and using that information to drive HVAC scheduling is shown to reduce the energy consumption of a number of
buildings spanning a large campus by over 30% in [4]. None of the above studies evaluate the efficacy of using WiFi to understand the nuances associated with emergency evacuation in buildings. Our work fills this important gap in the literature.

Building Evacuation: We now briefly review related work on building evacuation, which are largely based on modeling and simulations, but has not studied the role that WiFi can play in this domain. Network flow based modeling and optimization techniques are developed in [18]–[21] to study various problems associated with building evacuation such as how long it takes for occupants to be evacuated, where are bottlenecks likely to occur and so on. While the treatment is comprehensive, applying the techniques to different types of building is onerous and time consuming, since it requires detailed knowledge of the building layout, floor plans, emergency exits and stairway information, which may even not be readily available from facility managers. Geographic Information System (GIS), agent-based and IoT-based data driven techniques are being considered to evaluate the efficacy of evacuation drills, but the research is still in its infancy [27]. A summary of modeling research in the context of evacuation in high-rise buildings is available in [28].

VI. CONCLUSION

Evaluating building evacuations are largely manual, cumbersome, and error-prone. In this paper, we developed a systematic method to automatically detect and evaluate evacuation events using building WiFi trace data. Collection of campus-wide WiFi data over a period of six months, allowed us to examine the anatomy of building evacuations across multiple representative buildings with different number of levels, exit layouts and occupant types. Having understood the evacuation anatomy, we then developed a novel filtering-based method to automatically detect evacuation events from WiFi data. Using our method we detected 29 unplanned evacuations and evaluated the efficacy of our automatic detection by corroborating them against documented records where available. Lastly, we deduced measures namely speed, number of evacuees and typicality to quantitatively compare planned and unplanned evacuations and further estimated the productivity loss arising from evacuation events across our campus. Our findings showed that during emergencies occupants evacuate building slower compared to drills while number of evacuees vary depending on factors such as type and composition of occupants of the building and the nature of the emergency. Our work has shown the advantage of using building WiFi data for systematic and accurate evaluation of evacuations at scale, compared to current manual methods.

REFERENCES