Using Cameras to Improve Wi-Fi Based Indoor Positioning*

Laura Radaelli[1], Yael Moses[2], and Christian S. Jensen[3]

[1] Department of Computer Science, Aarhus University, Aarhus, Denmark
radaelli@cs.au.dk
[2] The Efi Arazi School of Computer Science, The Interdisciplinary Center, Herzliya, Israel
yael@idc.ac.il
[3] Department of Computer Science, Aalborg University, Aalborg, Denmark
csj@cs.aau.dk

Abstract. Indoor positioning systems are increasingly being deployed to enable indoor navigation and other indoor location-based services. Systems based on Wi-Fi and video cameras rely on different technologies and techniques and have so far been developed independently by different research communities; we show that integrating information provided by a video system into a Wi-Fi based system increases its maintainability and avoid drops in accuracy over time. Specifically, we consider a Wi-Fi system that uses fingerprints measurements collected in the space for positioning. We improve the system's room-level accuracy by means of automatic, video-driven collection of fingerprints. Our method is able to relate a Wi-Fi user to unidentified movements detected by cameras by exploiting the existing Wi-Fi system, thus generating fingerprints automatically. This use of video for fingerprint collection reduces the need for manual collection and allows online updating of fingerprints. Hence, increasing system accuracy. We report on an empirical study that shows that automatic fingerprinting induces only few false positives and yields a substantial accuracy improvement.

Keywords: Indoor Positioning · Wi-Fi Fingerprinting · Video Tracking

1 Introduction

Over the past decade, location-based services (LBS) have gained in prominence. LBS accounted for a revenue of USD 2.8 billion in 2010 and the expected revenue in 2015 is USD 10.3 billion [20]. However, today's location-based services target mostly outdoor users. In contrast, studies find that people spend some 87% of their time indoors [5, 10, 13], and 70% of cellular calls and 80% of data connections in the USA originated from indoors in 2013 [14]. Additionally, indoor LBS market is forecasted to grow by 40% over the period 2012–2016 [21]. Thus, time is ripe for enabling also indoor location-based services, where indoor positioning is a key enabler. Specifically, indoor positioning systems enable a range of indoor location-based services, including simple navigation services as well as more complex shopping assistants and friend finders, to name but a few.

* The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-642-55334-9_11

Key characteristics of indoor positioning systems include their maintainability and accuracy. The *maintainability* of a system captures the cost of maintaining the system, in particular ensuring that its accuracy remains acceptable. Next, *accuracy* often refers to the average positioning error in 2D or 3D with respect to the ground truth. Instead, we adopt the notion of room-level accuracy, where a position is considered correct if it belongs to the same room as the ground truth position. Thus, the exact position of a user inside a room is not relevant, but it is crucial to position the user in the correct room. Room-level accuracy enables many indoor services (e.g., silence-your-phone-in-a-meeting-room, in-store ads, and locate-a-friend) and data analyses (e.g., finding hang-outs in airports, popular shops in malls, and frequent sequences of shops visited). We consider the task of building a system with room-level accuracy that is capable of maintaining its accuracy over time.

While GPS is ineffective in indoor settings, video cameras and Wi-Fi are promising technologies for indoor positioning. Cameras can track indistinctly all people in a monitored space with high room-level accuracy, but cannot easily match a specific person with a location (if not relying on complex identification techniques). Hence, cameras alone are insufficient for positioning. On the other hand, Wi-Fi based systems can position only collaborative people that have a Wi-Fi device turned on, and can provide location of the specific user through the device, but they achieve highly different levels of room-level accuracy in different settings (reported accuracy varies from sub-meter up to 40 meters for pure Wi-Fi based systems).

There are different approaches to Wi-Fi positioning: model-based, fingerprinting, and trilateration. We focus on Wi-Fi systems that use fingerprinting. A fingerprint consists of a pair of an indoor location and a set of signal strengths of the Wi-Fi access points seen at that location. Given a set of fingerprints and the signal strengths observed by a mobile device, the system can position the device. In such a system one of the main challenges is the collection of fingerprints. When done manually by surveyors, this is time-consuming and expensive, and it needs to be repeated over time to maintain system accuracy. The need for a surveyor can be avoided, in fact a user's Wi-Fi device can constantly emit Wi-Fi signal strengths measurements, that can generate fingerprints if associated with the user's position. A classic solution is to ask the users to mark their locations on a map or select it from a list (e.g., [1, 19]). Such solutions reduce the cost, but also increase the effort by users. We propose a method that automates the collection of fingerprints, thus saving efforts by the users and increasing the fingerprint update rate to avoid decreased accuracy over time. At the core of our approach is the use of a camera to determine the room the user is in, resulting in automatic room-level fingerprinting that is transparent to the user.

Figure 1 illustrates how different information from the two sources can be integrated for reaching our goal. Two rooms are connected only to a corridor that is monitored by a camera whose field of view (FOV) is shaded gray. A user walks from the corridor to $room_2$ holding a Wi-Fi device; we can see the user's actual position and reported position (gray phones) at four times (t_1–t_4). First, the example shows that the room-level positioning of a Wi-Fi system can be inaccurate: at time t_4, the user is in $room_2$, but the reported position is in $room_1$. Second, the video system can tell that the user entered $room_2$ at time t_3 and did not leave; therefore, we can generate a fingerprint that

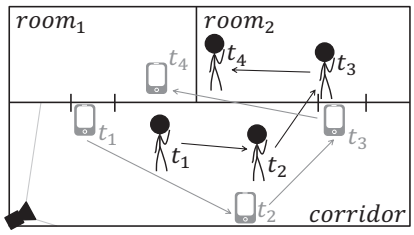


Fig. 1. Running example

associates $room_2$ with the Wi-Fi signal strengths measurements of the user received between t_3 and t_4.

To the best of our knowledge, only few works propose integration of fixed cameras with other positioning technologies, and they all investigate intra-room positioning, while we consider room-level. None of these works explore the use of cameras in fingerprinting. We provide a key building block for an automatic fingerprinting system by proposing a solution for corridor-like spaces. The proposed method can be applied to a more complex floor plan using a large set of cameras that cover the space. The basic solution proposed here is sufficient when the space is decomposed into corridor-like subspaces and the method is applied to each subspace.

The remainder of the paper is structured as follows: in Sec. 2, we formulate the problem and state our assumptions; Sec. 3 gives an overview of the assumed Wi-Fi and video systems, and presents the proposed integration method; Sec. 4 reports on experimental studies. We cover related work in Sec. 5 and conclude and provide future research directions in Sec. 6.

2 Problem Formulation

Setting: We aim to automate fingerprint collection by generating fingerprints without any active intervention by users. We consider an indoor space composed of rooms that are all accessible only via a central corridor.

A Wi-Fi positioning system is deployed in the space continuously collecting information on the signal strength measurements and location of each Wi-Fi user. In addition, a video system monitors the corridor, capturing the movements of people. It continuously detects events of a person entering or leaving a room and the associated time. We assume that a person remains in a room from the time of entering until the time of leaving the room.

Automatic Fingerprinting: A valid fingerprint for a given location can be obtained from the collected Wi-Fi signal strength measurements of a device, as long as the device remains in the same location for at least th_{fp} time units. This threshold, varies from system to system, but it is fixed at deployment time. For automatic fingerprinting, the location as well as the time interval for which the user remains in this location should be determined. In our case only the time interval a person spends within a room, *in-room-*

interval, plays a role. The interval is determined by the time of entering and leaving a room.

Assumptions: Synchronization is required when a pair of sensors monitor and report the position of the same object and do not share a physical clock. We assume that given a synchronization checkpoint when the clocks of the two sensors are aligned, the drift does not affect the synchronization within a short time frame. Fast drifting could occur when several frames are dropped in the video, but we consider very low resolution and low frame rate, hence this phenomenon is unlikely. We assume that for short videos with a synchronization checkpoint at the start (as in our experimental study), sources remain sufficiently synchronized.

A diagram of the floor is used to relate rooms with doors, and a door in the image seen by the camera is related to the appropriate room. The process of detecting doors in the image of the camera can be done either manually or automatically (by observing where do people disappear from the field of view). Furthermore a mapping between manual fingerprints in the Wi-Fi system and rooms is required to compute the room-level accuracy of a system.

As indicated above, we also assume that the space has a star topology: a central room ("corridor") has connections to n rooms that do not have connections to other rooms.

It is unrealistic to assume that all people in an indoor space are equipped with a Wi-Fi device; rather, we assume that some people may be walking in the space with no Wi-Fi positioning. We refer to people with Wi-Fi positioning as *users*. For simplicity we assume that there are no people in the space at time t_0 (i.e., at the beginning of each video monitoring session), and in our experiments we consider two people (not necessarily two users) walking in the monitored space at the same time.

3 Methods

We describe a method for integrating two different sources of information on the movements of people, namely Wi-Fi positioning and video tracking, in order to automate Wi-Fi location fingerprinting.

3.1 Wi-Fi Fingerprinting

For a Wi-Fi positioning system, a fingerprint is a pair of an indoor location and a combination of signal strengths of Wi-Fi access points visible from the location. A fingerprint is traditionally collected by a surveyor standing in a specific location with a Wi-Fi device and recording the signal strengths measured by the device for some time. The resulting fingerprints are stored in a database and are subsequently used by the system for positioning.

When a user submits the signal strength measurements "seen" by the user's device, the system returns the position in the database that corresponds to the fingerprint whose signal strength measurements are the most similar to those submitted by the user. A user's position can be computed either on the server or on the device, with each approach having different pros and cons (e.g., cost of computation, privacy). A mapping

between the coordinate system of the location and an image of the floor plan is usually available so that the user's position can be shown on the floor plan for navigation purposes.

A room-level fingerprint can be generated from signal strength measurements of a user if the in-room-interval of the user is known. In the next sections, we describe how to achieve this by using cameras.

3.2 Video Tracking

The tracking of people by using surveillance cameras is an active research topic in computer vision. We propose to use video data to assist a Wi-Fi system in collecting automatic fingerprints. Specifically, the video system can provide the crucial information about the in-room-interval of users by identifying the entrance and leaving times of a user to a given room.

Many sophisticated video tracking algorithms exist that can provide the information we need (see review in [23]), and the general method we propose can use any of these. In our implementation we use a relatively simple tracking algorithm that relies on background subtraction. Our algorithm assumes a fixed background, which is sufficient when the scene is static and not crowded as in our setting. We take an image of the empty monitored space as a fixed background and compute the pixel-wise difference with a frame to detect changes in the space. We ignore differences that are smaller than a noise threshold (more sophisticated methods are given in [4]).

After applying the basic background subtraction algorithm, we examine each frame object-wise looking for connected components, i.e. groups of pixels that form a blob and that could represent objects moving in the scene. For each frame we extract the number of objects in the scene (one in Fig. 2), the line of feet of each object (l_1 in Fig. 2), and the center of the shape of each object (c_1 in Fig. 2).

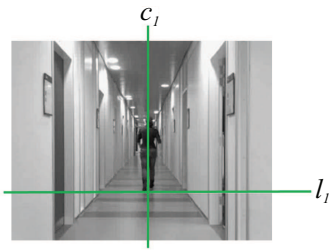


Fig. 2. Example of features extracted from video

From these features, we can compute in-room-intervals of users and people entering/leaving the monitored space. The disappearing and appearing events of persons in the scene can be identified by comparing the detected objects in successive frames. In particular, when a tracked object disappears from the field of view, it indicates that a

person either entered a room or left the corridor. When a new object enters the field of view, it may be either a person leaving a room or a person entering the corridor.

We assume that a room is associated to each door in the image seen by the camera. We determine which room is involved in the event according to the line of feet and the shape center with respect to the marked doors. If the person location does not correspond to any of the doors, we label the event as corridor leaving (or entering) event. For robustness, we consider detection in w successive frames.

Our video tracking algorithm returns a set of events, each with the timestamp of the event, its type (enter/leave), and the room involved. Fig. 3 shows the result of applying the algorithm to a video that records a person walking in the corridor and entering room 337 at time t_{x+1}, and then another person leaving the same room at time t_{y+1}; the algorithm detects a room-enter event $enter_{337,t_{x+1}}$ and a room-leave event $leave_{337,t_{y+1}}$.

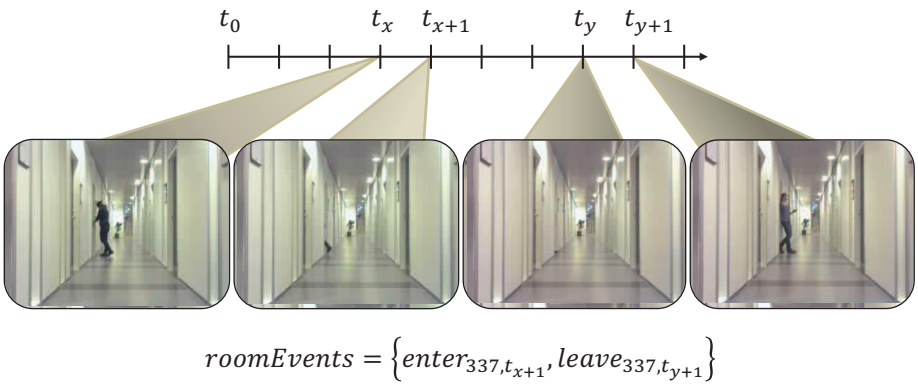


Fig. 3. Room-enter and room-leave events

We want to couple a person entering/leaving a room with a Wi-Fi device. We directly use the Wi-Fi positioning for identifying the user involved in the event. Another approach would be to use sophisticated image recognition techniques to identify the user entering/leaving a room. This is left for future research.

3.3 Integration Technique

The video system provides information on people entering and leaving rooms, and the Wi-Fi system provides signal strength measurements for all users. To generate a fingerprint, we need to match the interval a person spends in a specific room with a specific Wi-Fi user.

The proposed heuristic method aims to find proper assignments of Wi-Fi users to the in-room-intervals reported by the video system. Fingerprints are then generated using these assignments. The method is structured as follows:

1. For each room, find in-room-intervals using the video system.

2. To each in-room-interval, assign a Wi-Fi user if one exists.
3. For each user, compute when the user entered/left the corridor.

We describe each step in the following.

Find in-room-intervals: We first process the sequence of room events detected by the video system. Four different situations might occur, as shown in Fig. 4.

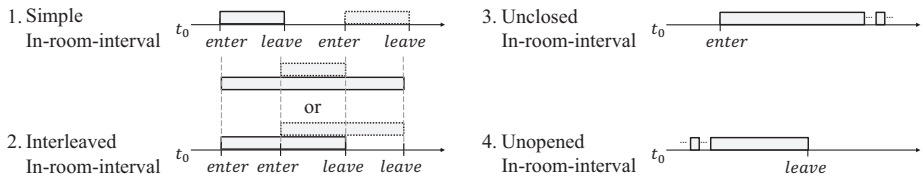


Fig. 4. In-room-intervals in a sequence of events

We can exclude unopened in-room-intervals: as we assume that at the starting time t_0 there are no people in the space, the video system must observe an $enter$ event before each $leave$ event. To contend with unclosed in-room-intervals, we set the end of the interval to the time of computation t_c (which coincides with the end of a video session in our experimental study). We consider an interleaved in-room-interval as a single in-room-interval that extends form the first $enter$ event until the last $leave$ event.

We process the events sequentially. When encountering an $enter$ event, we look for the next event. Three cases can occur. First, when the current event is the last in the sequence, it is an unclosed in-room-interval (UcI), and we set the end time of the interval to current computation time t_c. Second, when the next event is a $leave$ event, we save the interval as a simple in-room-interval (SI). Third, when the next event is an $enter$ event, we have encountered an interleaved in-room-interval (II), and we do not close the interval until it contains the same number of $enter$ and $leave$ events. In this case, ambiguity occurs. The system can choose to ignore such intervals, but we propose a heuristic to select a possible solution. The output of this step is a sequence of in-room-intervals.

Assign Wi-Fi Users: Having found in-room-intervals for each room, we now assign each interval to a Wi-Fi user. Some of the challenges in doing so are given by the overlapping intervals and the presence of people who are not Wi-Fi users.

We exploit the existing Wi-Fi positioning system to select the users that may be assigned to an in-room-interval. Candidates are all the users with a Wi-Fi position δ-close to the door of the room at both enter and leave time of a simple in-room-interval. If no user fulfills the requirements for being a candidate, we conclude that the person that entered the room is not a Wi-Fi user. If there is more than one candidate for the same interval, we simply select the first detected user, but more complex strategies could be adopted, e.g., considering the distance of candidates to the door.

The algorithm used to assign in-room-intervals to users is detailed in Algorithm 1. The algorithm goes through all in-room-intervals ordered by enter time and deals with simple and unclosed in-room-intervals (lines 8–12) and interleaved in-room-intervals in

Algorithm 1 assignIntervals()

```
1: Input: Intervals, Users, τ, δ
2: PrevIntervals ← ∅;
3: UsedEvents ← ∅;
4: for i ∈ Intervals do
5:    if i.type ∈ {SI, UcI} then
6:       U ← getUsers(i, τ, δ) \ overlap(i, PrevIntervals);
7:       if |U| > 0 then
8:          user ← pickUser(U);
9:          Assignments ← Assignments ∪ {(user, i)};
10:   if i.type ∈ {II} then
11:      iteration ← 1;
12:      while (|PrevCombo| * 2 < |i.events|) do
13:         Combinations ← computeComb(i);
14:         for l ∈ ⟨1, ..., |Users|⟩ do
15:            for c ∈ Combinations(l) do
16:               skipCombos(iteration − 1);
17:               if c.enter ∉ UsedEvents ∧ c.leave ∉ UsedEvents then
18:                  U ← c.users \ overlap(c.interval, PrevIntervals);
19:                  U ← U \ overlap(c.interval, PrevCombos);
20:                  if |U| > 0 then
21:                     c.user ← pickUser(U);
22:                     PrevCombos ← PrevCombos ∪ {c};
23:                     UsedEvents ← UsedEvents ∪ {c.enter, c.leave};
24:         iteration ← iteration + 1;
25:      Assignments ← Assignments ∪ getSimpleIntervals(PrevCombos);
26:   PrevIntervals ← PrevIntervals ∪ {i};
27: return Assignments;
```

a different manner (lines 13–21). First, we select the candidate users U (lines 6, 18.19). U is composed by users that are δ-close to the door of the room during a period of time of 2τ around the enter and leave times of the in-room-interval (function $\mathrm{getUsers}(\cdot)$). U excludes users already assigned to other rooms for in-room-intervals overlapping in time the current one (function $\mathrm{overlap}(\cdot)$). Second, if U contains more than one candidate, we use the function $\mathrm{pickUser}(\cdot)$ to select which candidate to assign to the in-room-interval. In our case, the function selects the first user in U. For interleaved in-room-intervals, we have to pair enter and leave events: $Combinations$ contains all the possible pairs in the interval, each coupled with the users retrieved by $\mathrm{getUsers}(\cdot)$. We iterate through all the combinations (lines 14–23), starting with the ones that have only one assigned user ($l = 1$). If no suitable combination is found during the first iteration, the next iteration skips the first combination (using function $\mathrm{skipCombos}(\cdot)$) to avoid examining the combinations in the same order as in the previous iteration. More sophisticated heuristics can be designed to address increasing number of events and ambiguities.

Assign Corridor Events: Having determined in-room-intervals for each user, we still need to decide when each one of them entered and left the corridor.

As discussed in Sec. 3.2, we assume that from the video source, we can detect events such as “entering the corridor” and “leaving the corridor.” In some cases, different options are available for pairing corridor and room events, some are shown in Fig. 5.

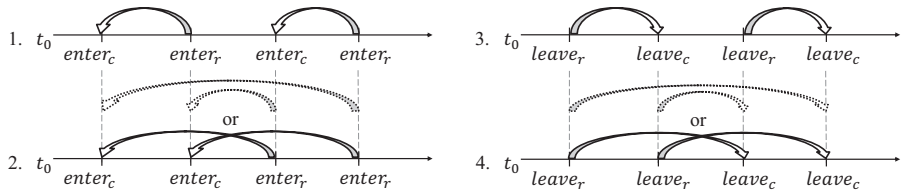


Fig. 5. Some of the possible pairings between room and corridor events

We do not have enough information to identify which pairing option is correct. Therefore, we apply the same pairing policy for all events: we assign a corridor event to the closest room event of the same type. This means that we pair a room-enter event with the closest available corridor-enter event that occurred before it, and we pair a room-leave event with the closest corridor-leave event that occurred after it. In terms of the options in Fig. 5, we choose the black straight arrows and discard the dotted ones.

4 Experimental Study

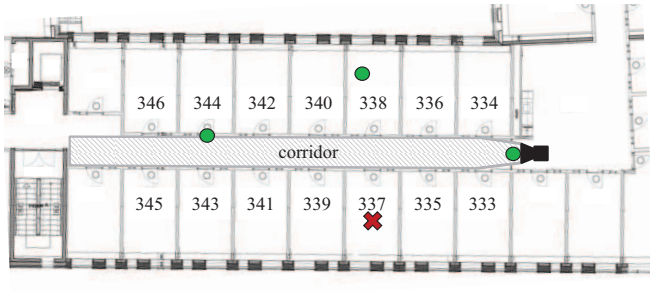
We evaluate the proposed method in a controlled environment, namely a corridor of offices in the Department of Computer Science at Aarhus University. We examine the quality of the automatic fingerprints (Sec. 4.2), study the integration method (Sec. 4.3), and consider the positioning accuracy achieved by automatic fingerprinting (Sec. 4.4).

4.1 Settings

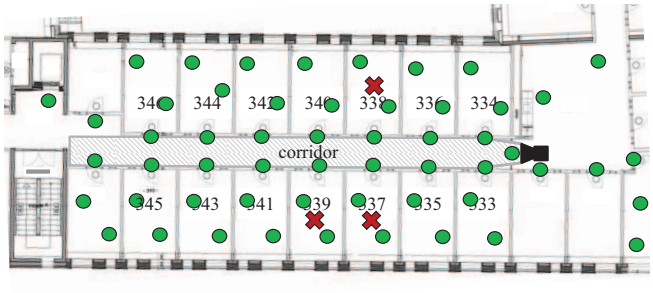
The floor plan, in Fig. 6, encompasses a corridor with 14 offices. A Dell Integrated Webcam mounted on a laptop positioned as shown in the figure is used for the monitoring. We employ a Wi-Fi positioning system [12] with two different sets of fingerprints in two different studies, and we use Samsung Galaxy S3 and Nexus 3 phones connected through a dedicated Android application.

4.2 Study 1: Fingerprint Comparison

We first compare manual and automatic fingerprints. We perform 10 surveys, during each of which we collect one manual fingerprint in the middle of room 337 (using a dedicated mobile application) and then collect an automatic fingerprint for the same room. The fingerprint location is marked as a cross in Fig. 6(a), where dots are pre-existing fingerprints. We compare manual and automatic fingerprints in two different ways. First, we compute the similarity between the manual and automatic fingerprint



(a) Fingerprints in Study 1



(b) Fingerprints in Study 2

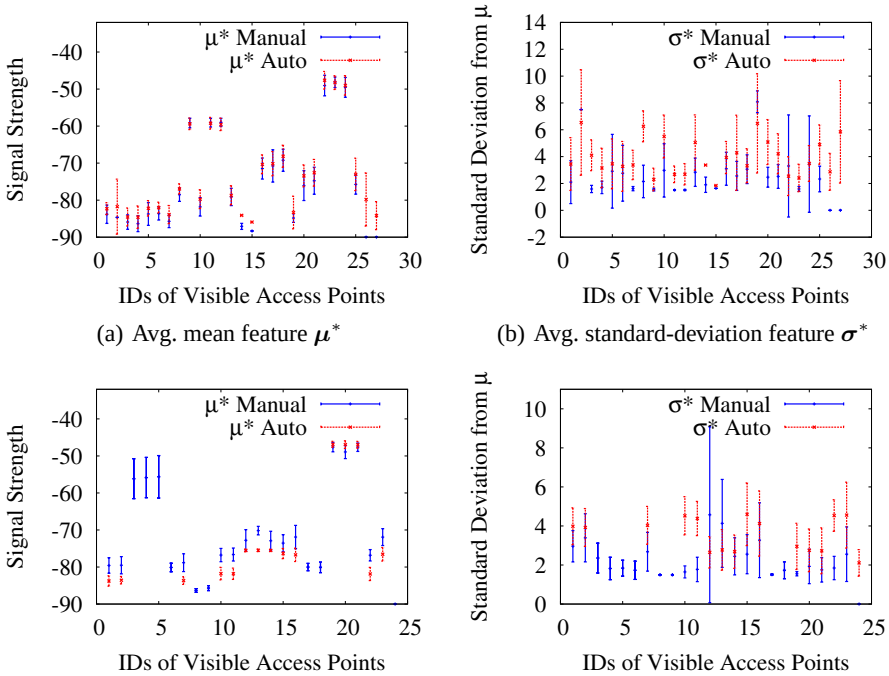
Fig. 6. Floor plan of our hallway

and between the automatic and all the other pre-existing fingerprints. Second, we compare the manual and automatic fingerprints based on two of the main features of a fingerprint.

We compute the similarity between two fingerprints by using the same algorithm used for positioning a user. The result is the same for all 10 surveys: the similarity between manual and automatic fingerprint is 1, while the similarity between the automatic fingerprint and all the other pre-existing fingerprints is near-zero.

Next, we compare the manual and automatic fingerprints based on two features: the *mean feature* $\boldsymbol{\mu}$, a vector consisting of mean values of signal strengths for each visible access point; and the *standard-deviation feature* $\boldsymbol{\sigma}$, defined as a vector of corresponding standard deviations (i.e., the ith entry is the standard deviation from the mean value for the ith visible access point). We compute $\boldsymbol{\mu}^* = \frac{1}{10}\sum_{j=1}^{10}\boldsymbol{\mu}_j$ ($\boldsymbol{\mu}_j$ is the mean vector from the jth survey) as the average over the 10 surveys; similarly, we compute $\boldsymbol{\sigma}^* = \frac{1}{10}\sum_{j=1}^{10}\boldsymbol{\sigma}_j$ as average of the standard-deviation features.

Fig. 7(a) shows $\boldsymbol{\mu}^*$ for manual (solid line) and automatic (dashed line) fingerprints. For each bar, the point in the middle represents the average value for the access point, and the height describes the standard deviation. For all access points, the average values of manual and automatic fingerprints are very close. For access points 2 and 19, the standard deviation of the automatic fingerprints is larger than for the manual ones, but in all the remaining cases, also the standard deviations are comparable.



(a) Avg. mean feature $\boldsymbol{\mu}^*$

(b) Avg. standard-deviation feature $\boldsymbol{\sigma}^*$

(c) Avg. mean feature $\boldsymbol{\mu}^*$, 2 phones

(d) Avg. standard-deviation feature $\boldsymbol{\sigma}^*$, 2 phones

Fig. 7. Manual and automatic fingerprint comparison

Fig. 7(b) shows $\boldsymbol{\sigma}^*$. The value of the average standard-deviation feature for automatic fingerprints is approximately double that of manual fingerprints. The impact of this difference depends on how relevant the standard deviation is in the similarity measure used for positioning a user and on the ratio of manual/automatic fingerprints in the system. In fact, if most of the fingerprints are actually automatic then this does not affect the system at all. A reason for the difference in standard deviation might be the fact that during manual collection, the phone is stationary, while during automatic collection, the phone might be moving in the room. This may actually be an advantage, since users being positioned are more likely to be moving around than to be stationary.

We also perform 10 surveys of the same type, but with two different phones, so that different phones are used for manual and automatic fingerprinting. The similarity results are as before, i.e., the similarity between manual and automatic fingerprints is 1, and the similarity of automatic fingerprint with the other points in the radio map is near-zero.

A comparison of the average mean feature is shown in Fig. 7(c). Fingerprints collected with the two phones are similar, but the automatic ones generally have a lower mean feature and a lower standard deviation. Results for the standard-deviation feature, in Fig. 7(d), show higher average values for one phone, but similar standard deviations.

We conclude that manual and automatic fingerprints are comparable, implying that a system that contains both types of fingerprints, or possibly only automatic ones, is feasible. When using different phones for collection, the results are less clear. Fingerprints collected with different phones at the same location are more similar than the fingerprints collected at different locations with different phones. We also find that fingerprints taken with different phones show slightly different feature values, which might reduce accuracy depending on the specifics of the Wi-Fi positioning system.

4.3 Study 2: Automatic Fingerprinting Evaluation

Evaluation metrics To evaluate the integration method, we look at true positives, false positives, and false negatives. The meanings of these when matching in-room-intervals to users are:

- true positives: correct assignments, suggesting how much manual fingerprinting work is avoided;
- false positives: assignments of an in-room-interval to an incorrect user; these assign a fingerprint to a wrong location and are very undesirable;
- false negatives: number of in-room-intervals not assigned to any user, while a user was actually in the room; these indicate how much room there is for improvement.

Experiments We record different videos of people moving in the monitored space, and we compare the results of the algorithm with the ground truth.

We perform 9 experiments. Each consists of one or two people walking in the monitored space according to a script, while the camera is recording a video and the Wi-Fi system is recording positioning data of the Wi-Fi devices they are carrying. The Wi-Fi system is initialized using fingerprints shown as dots in Fig. 6(b), where automatic fingerprint locations are shown as crosses. The experiments are sequenced according to the room involved and the complexity of the scene and script. We start with a single person walking and entering/leaving a room; we continue with two people under different conditions: walking separately or together and with or without a Wi-Fi device.

Event Detection For each experiment, we check whether the video tracking recognizes when a person enters/leaves a room and recognizes the room. All but one experiment are processed with the same parameters for background subtraction; in experiment #7 we use different values due to unusual lighting and background interference (a plant at the end of the corridor happened to be of the same color as the user's clothing).

Fig. 8(a) shows the number of events detected by the algorithm with respect to the real events.

We have no false negatives, i.e., we do not miss any events. We have only few false positives, meaning that we detect as an event something that is not an event. Most of these are due to the configuration of the scene: people leaving the FOV are classified as people entering the last room in the corridor. Or they are due to the simple background subtraction algorithm we use: people wearing clothing that is the same color as a part of the background. We can conclude that we achieve good results for our setting and our experiments. However, in a more general setting, a more sophisticated background subtraction algorithm must be considered. Only experiment #9 was affected by false

Actual value	Algorithm outcome: Event, enter	Event, leave	Not event, enter	Not event, leave
total	25	12	0	0
Event	13	9	0	0
Not event	12	3	-	-

(a) Room enter and leave event detection

Actual value	Algorithm outcome: x in i	x not in i
total	14	3
x in i	12	1
x not in i	2	2

(b) User x to room i match

Actual value	Algorithm outcome: User x, enter	User x, leave	Not User x, enter	Not User x, leave
total	16	11	0	1
User x	16	11	0	1
Not User x	0	0	-	-

(c) User x to corridor event match

Fig. 8. Confusion matrices

positives. Here, the person leaving the corridor is classified as entering room 340, creating an incorrect fingerprint. All other errors are "dismissed," either because there is no corresponding enter/leave event or because no Wi-Fi traces correspond to the event.

User Assignment To generate an automatic fingerprint, we need to assign a Wi-Fi user to an in-room-interval; assignment results from our algorithm are shown in Fig. 8(b).

We observe two false positives. The false positive in experiment #3 corresponds to an empty in-room-interval (duration 0) derived from a faulty event detection, and the false positive in experiment #9, as already mentioned, gives a real error, since a fingerprint is generated for a wrong assignment of an in-room-interval to a user. We have only one false negative: in experiment #7, an in-room-interval is not assigned to $user1$, but to an individual without a Wi-Fi device. This does not lead to a wrong fingerprint.

We can conclude that the assignment of in-room-intervals to users in our method performs well; all types of in-room-intervals (simple, unclosed, and interleaved) are detected correctly, and users are correctly assigned in all cases but two (one with no consequences, and one yielding a wrong fingerprint).

Corridor Event Assignment We want to check whether the algorithm can correctly recognize when a user enters/leaves the corridor. Corridor events are not used in our experiments to generate fingerprints due to the configuration of the corridor; in fact, it is very long, and we cannot assign all the measurements collected along the whole corridor to a single point in the middle of it. In different setting, these could be used to generate automatic fingerprints.

The results in Fig. 8(c) show that we have only one error in the corridor assignments (experiment #9), which is due to a faulty event detection that leads to a faulty user assignment. We conclude that assignment of corridor events in our method works well in the experiments.

4.4 Study 2: Room-Level Accuracy Evaluation

Evaluation metrics Our method targets room-level positioning. Knowing the floor plan of the building, we can map each Wi-Fi location to a room. Hence, we can define the *Wi-Fi room location* $r_{u,t}$ as the room where user u is at time t according to the Wi-Fi system. We use R to denote a set of Wi-Fi room locations.

An error occurs when a position differs from the ground truth. For a user u at time t, $r_{u,t}^{true}$ denote the ground truth room location. We consider a function $\mathrm{roomMatch}(\cdot)$ that counts the number of matches with the ground truth of a set of Wi-Fi room locations R:

$$\mathrm{roomMatch}(R) = |R_{\text{match}}|, \text{ where } R_{\text{match}} = \{r_{u,t} \in R | r_{u,t} = r_{u,t}^{true}\} \ . \qquad (1)$$

We also count matches in an alternative way based on the overall state and not on a single user. Then a match occurs at a specific time only if all users are correctly positioned at that time:

$$\mathrm{timeMatch}(R) = |T_{\text{match}}|, \text{ where } T_{\text{match}} = \{t | \forall u \ (r_{u,t} = r_{u,t}^{true})\} \ . \qquad (2)$$

When R contains locations of a single user only, $\mathrm{roomMatch}(R) = \mathrm{timeMatch}(R)$.

Experiments We want to check whether uploading a new automatic fingerprint improves the room-level accuracy. Therefore, before each experiment described in the previous section, we upload the automatic fingerprints generated in the previous experiment, so that we can compare the accuracy before and after experiments.

For each experiment, we measure the room-level accuracy of the Wi-Fi system for the period of time during which at least one user is inside the monitored space. Results are shown in Fig. 9, where the number of room matches for each user is computed using the $\mathrm{roomMatch}(\cdot)$ function (Equation 1) and the number of matches for both users is computed using the $\mathrm{timeMatch}(\cdot)$ function (Equation 2). We expect the accuracy of

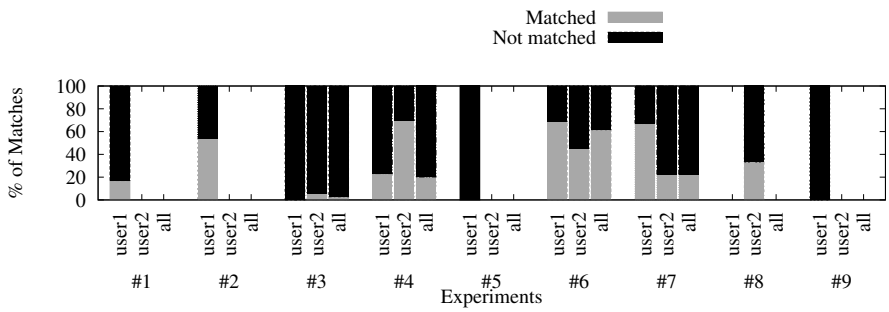


Fig. 9. Room-level accuracy evaluation

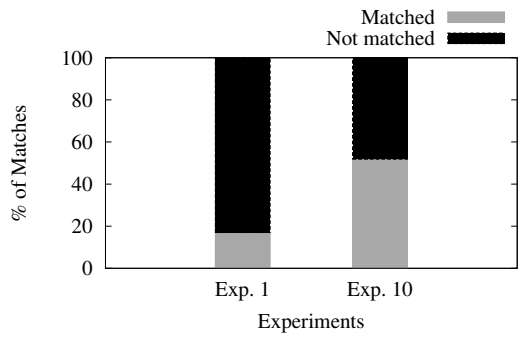
the system to increase from one experiment to the next one due to new and more up-to-date fingerprint uploaded. From the results we see that this is true to some extent. Notice that for each pair of experiments involving the same room (e.g., experiments #1 and #2), the accuracy is about three times higher in the experiment performed after uploading an automatic fingerprint. Therefore, uploading a new fingerprint with current measurements improves the positioning accuracy for that location.

On the other hand, we observe a drop every two experiments (every time a new room is visited); one possible reason for this behavior is that measurements are more similar to new fingerprints collected in other location than to old fingerprints collected

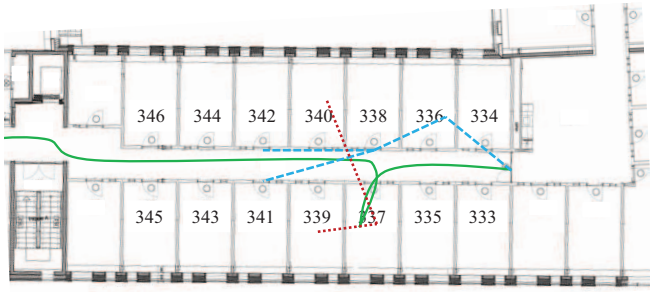
at the ground truth location. This calls for a method to “retire” fingerprints when they are too old, or at least include the “age” of a fingerprint in the computation of a user’s position.

In experiments #8 and #9, users are asked to walk in the corridor without entering any room. The resulting low accuracy can be due to the fact that no new fingerprints are uploaded for the corridor during any of the experiments.

We perform a last experiment (experiment #10), where the path taken by the user is the same as in experiment #1; the user walks in the corridor, enters room 337, stays there for 1 minute, and then leaves the room and the corridor. A comparison of the room-level positioning accuracy in the two experiments is shown in Fig. 10(a).



(a) Room-level accuracy comparison



(b) Trace comparison

Fig. 10. Experiments #1 and #10

The accuracy in the last experiment is three times higher than the accuracy in the first one, and three automatic fingerprints have been uploaded for room 337 between the two experiments.

Fig. 10(b) shows the traces in the two experiments compared with the ground truth. The solid line represents the ground truth, the dashed line experiment #1, and the dotted line experiment #10. We see that not uploading new fingerprints for the corridor reduces the accuracy of positioning in it, since the position of the user is assigned only to rooms

in experiment #10. Moreover, the wrong fingerprint generated in experiment #9, which assigns a user to room 340 when the user is actually in the corridor, affects the positioning in experiment #10 (when the user is in the corridor, the position is reported as room 340). We also find that the positioning accuracy in rooms for which new fingerprints are uploaded increases, and the user is correctly positioned in room 337 most of the time.

5 Related Work

Different technologies can provide positioning in indoor spaces [16]. We utilize two different technologies: Wi-Fi and video cameras.

Wi-Fi based positioning has been studied widely, and a wide range of systems have been proposed. Liu et al. [15] present a survey of approaches to wireless positioning. There are three main different approaches: model-based, fingerprinting, and trilateration. In a model-based and trilateration positioning the location of access points is known and positioning is done using propagation models or lateration methods [2]. Our work is focused on the systems that use Wi-Fi fingerprinting. Kjærgaard [12] presents a taxonomy of fingerprinting techniques. A known problem of fingerprinting-based systems is that they call for the collection of fingerprints, which is costly in terms of time and resources. Thus, a number of recent studies investigate the use of robots [11], techniques that enable users to do fingerprinting [19], and techniques that enable user feedback [7] for reducing the cost of fingerprint collection. We tackle the problem of fingerprint collection by exploiting cameras to allow automatic fingerprinting, thus eliminating or reducing the need for active user involvement.

In the realm of video-based systems, different techniques have been investigated for the tracking of objects in a space monitored by cameras. Some techniques use a single camera, and some use multiple cameras with or without overlapping fields of view; other techniques aim to recognize the shapes of objects; and yet some techniques focus on positioning. Yilmaz et al. [23] contribute a comprehensive survey. We employ a simple background subtraction technique that is sufficient in our experimental setting. More advanced techniques are likely to be needed in more general settings, especially when execution time is an issue [3, 6].

We propose a method of improving a Wi-Fi based positioning system by using cameras. Only few works consider the integration of cameras with Wi-Fi based systems, most of which consider phone cameras and not fixed cameras such as surveillance cameras [8, 17]. These other approaches rely on a phone camera for the main positioning and use Wi-Fi positioning to prune the space and improve the efficiency of video matching algorithms. Our approach is the opposite, as we use Wi-Fi as the main positioning technology.

Van den Berghe et al. [22] consider fixed cameras in a proposal for fusing camera and Wi-Fi sensor data in order to provide intra-room positioning. They consider a fingerprinting-based Wi-Fi positioning system and a camera installed on the ceiling of a room. Their proposal detects moving objects in video, and it projects a found object onto the floor plan using a Gaussian model. A particle filter “combines” the results of camera detection with the Wi-Fi positioning in real-time. Accuracy findings in terms of 2D error inside the room of the resulting system are reported. When only one person is

in the room, the accuracy of the underlying Wi-Fi system alone is improved by 0.5–2 meters, but in case of many people walking in the room, the improvement is insignificant (less than 0.5 meters). Results are not given with respect to room-level accuracy.

The use of external cameras has also been explored in connection with RFID localization systems. Works in this direction vary with respect to RFID localization algorithms and video processing algorithms [9, 18]. Most of the proposed systems are evaluated in controlled environments; we use the same general approach to evaluate our system.

6 Conclusions and Future Work

We propose a method that enables automatic collection of fingerprints for a Wi-Fi based indoor positioning system in corridor-like spaces, thus providing an important building block towards a full-fledged organic positioning system. The method integrates two different technologies, namely Wi-Fi positioning and video cameras. We target room-level accuracy and study this level of accuracy in a real positioning system when introducing automatic fingerprinting.

Our experimental study indicates that automatic fingerprints are comparable with fingerprints that are collected manually. In an evaluation of the effect of uploading new automatic fingerprints to the positioning system, we find that the accuracy increases for the rooms that have been newly fingerprinted, but also decreases for rooms with old fingerprints. Hence, we conclude that a system that allows new fingerprints to be uploaded also needs a method for discarding obsolete fingerprints. In experiments where we upload new fingerprints for rooms, but not for the corridor, the overall trace might not show that a user actually passed through the corridor before entering a room.

A natural next step is to utilize the proposed method in a solution that covers an entire floor or building. Our method can be applied independently to different monitored subspaces, and by exploiting topological connections between the subspaces, it may be possible to further improve the in-room-interval detection capability. A key challenge is to extend the automatic fingerprinting to non-corridor-like spaces (e.g., rooms with multiple doors), where one might need to recognize a person entering and leaving a room through different doors. A possible solution is to rely on vision re-identification techniques.

An interesting direction for future work is to study an online version of the method that would make it possible to turn off the phone or to reduce scanning interval time to save battery. In fact, in addition to positioning accuracy, battery consumption is an important problem. Under the assumption that the user needs room-level positioning, we may use cameras that detect when a user enters and leaves a room to turn the Wi-Fi device off and on.

Acknowledgments The authors thank Thor S. Prentow for setting up and running the Wi-Fi positioning system. This research was supported in part by the Geocrowd Initial Training Network funded by the European Commission as an FP7 Peoples Marie Curie Action under grant agreement number 264994. Laura Radaelli did part of this work while visiting IDC Herzliya.

References

1. A. Barry, B. Fisher, and M. L. Chang. A long-duration study of user-trained 802.11 localization. In *Proc. MELT*, pp. 197–212, 2009.
2. S. Bell, W. R. Jung, and V. Krishnakumar. WiFi-based enhanced positioning systems: accuracy through mapping, calibration, and classification. In *Proc. ISA*, pp. 3–9, 2010.
3. M. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. Van Gool. Online multiperson tracking-by-detection from a single, uncalibrated camera. *IEEE TPAMI*, 33(9):1820–1833, 2011.
4. M. Cristani, M. Farenzena, D. Bloisi, and V. Murino. Background Subtraction for Automated Multisensor Surveillance: A Comprehensive Review. *EURASIP J. Adv. Signal Process*, Article 43, 24 pages, 2010.
5. W. H. Dörre. Time-activity-patterns of some selected small groups as a basis for exposure estimation: A methodological study. *J. Expo. Anal. Environ. Epidemiol.*, 7:471–491, 1997.
6. A. Elgammal. Background subtraction: Theory and practice. In *Wide Area Surveillance*, volume 6 of *Augmented Vision and Reality*, pp. 1–21. 2014.
7. T. Gallagher, B. Li, A. Dempster, and C. Rizos. Database updating through user feedback in fingerprint-based Wi-Fi location systems. In *Proc. UPINLBS*, pp. 1–8, 2010.
8. H. Hile and G. Borriello. Positioning and orientation in indoor environments using camera phones. *IEEE Computer Graphics and Applications*, 28(4):32–39, 2008.
9. A. Isasi, S. Rodriguez, J. L. D. Armentia, and A. Villodas. Location, tracking and identification with RFID and vision data fusion. In *Proc. RFID Sys Tech*, pp. 1–6, 2010.
10. C. S. Jensen, K.-J. Li, and S.Winter. ISA 2010 workshop report: the other 87%: A report on the second international workshop on indoor spatial awareness (San Jose, California - November 2, 2010). *SIGSPATIAL Special*, 3(1):10–12, 2011.
11. K.-H. Kim, A. W. Min, and K. G. Shin. Sybot: an adaptive and mobile spectrum survey system for WiFi networks. In *Proc. MOBICOM*, pp. 293–304, 2010.
12. M. B. Kjærgaard. Indoor location fingerprinting with heterogeneous clients. *Pervasive and Mobile Computing*, 7(1):31–43, 2011.
13. N. E. Klepeis, W. C. Nelson, W. R. Ott, J. P. Robinson, A. M. Tsang, P. Switzer, J. V. Behar, S. C. Hern, and W. H. Engelmann. The national human activity pattern survey (NHAPS): A resource for assessing exposure to environmental pollutants. *J. Expo. Anal. Environ. Epidemiol.*, 11(3):231–252, 2001.
14. T. E. Lacroix. Indoor LBS market report. Sep. 2013.
15. H. Liu, H. Darabi, P. Banerjee, and J. Liu. Survey of wireless indoor positioning techniques and systems. *IEEE Trans. Syst., Man, Cybern., C*, 37(6):1067–1080, 2007.
16. R. Mautz. Indoor Positioning Technologies. *Geodätisch-geophysikalische Arbeiten in der Schweiz*, 2012.
17. H. Morimitsu, R. Pimentel, M. Hashimoto, R. Cesar, and R. Hirata. Wi-Fi and keygraphs for localization with cell phones. In *Proc. ICCV Workshops*, pp. 92–99, 2011.
18. T. Nick, S. Cordes, J. Gotze, and W. John. Camera-assisted localization of passive RFID labels. In *Proc. IPIN*, pp. 1–8, 2012.
19. J.-g. Park, B. Charrow, D. Curtis, J. Battat, E. Minkov, J. Hicks, S. Teller, and J. Ledlie. Growing an organic indoor location system. In *Proc. MobiSys*, pp. 271–284, 2010.
20. J. t. Sythoff and J. Morrison. Location-based services market forecast, 2011–2015. *Pyramid Research*, May 2011.
21. Technavio. Global indoor LBS market 2012–2016. Aug. 2013.
22. S. Van den Berghe, M. Weyn, V. Spruyt, and A. Ledda. Fusing camera and Wi-Fi sensors for opportunistic localization. In *Proc. UBICOMM*, pp. 169–174, 2011.
23. A. Yilmaz, O. Javed, and M. Shah. Object tracking: a survey. *ACM Comput. Surv.*, 38(4), 2006.