

DynamicSLAM: Leveraging Human Anchors for Ubiquitous Low-Overhead Indoor Localization

Ahmed Shokry, *Student Member, IEEE*, Moustafa Elhamshary, *Member, IEEE*
and Moustafa Youssef, *Fellow, IEEE*

Abstract—We present *DynamicSLAM*: an indoor localization technique that eliminates the need for the daunting calibration step. *DynamicSLAM* is a novel Simultaneous Localization And Mapping (SLAM) framework that iteratively acquires the feature map of the environment while simultaneously localizing users relative to this map. Specifically, we employ the phone inertial sensors to keep track of the user's path. To compensate for the error accumulation due to the low-cost inertial sensors, *DynamicSLAM* leverages unique points in the environment (anchors) as observations to reduce the estimated location error. *DynamicSLAM* introduces the novel concept of mobile human anchors that are based on the encounters with other users in the environment, significantly increasing the number and ubiquity of anchors and boosting localization accuracy. We present different encounter models and show how they are incorporated in a unified probabilistic framework to reduce the ambiguity in the user location. Furthermore, we present a theoretical proof for system convergence and the human anchors ability to reset the accumulated error. Evaluation of *DynamicSLAM* using different Android phones shows that it can provide a localization accuracy with a median of 1.1m. This accuracy outperforms the state-of-the-art techniques by 55%, highlighting *DynamicSLAM* promise for ubiquitous indoor localization.

Index Terms—Unconventional localization, SLAM, indoor localization, unsupervised localization.



1 INTRODUCTION

Recent years have witnessed the advent of indoor localization systems harnessing the many untapped capabilities of phones-embedded sensors and their widespread availability. Most developed indoor localization techniques are WiFi-based, leveraging the Received Signal Strength (RSS) from WiFi access points (APs) as the metric for the location determination. To combat the noisy wireless channel, typically a WiFi fingerprint map is built that captures the signatures of the APs at different locations in the area of interest. Nonetheless, the deployment cost is prohibitive as this requires a calibration phase; which is time consuming, labour intensive and is vulnerable to environmental dynamics. To tackle this problem, a number of approaches for automating the fingerprinting process have been proposed including the use of RF propagation models [7], combining RF localization with other sensors [34], or crowdsourcing the fingerprint; where users perform the required survey process in the real-time [27], [28]. These techniques, however, suffer from low location accuracy and/or require explicit user intervention. To further address these issues, some algorithms, e.g., [15], leverage the phone inertial sensors to track the user starting from a reference point. However, the localization error quickly accumulates due to the noise of the low-cost inertial sensors. To alleviate this problem, researchers, e.g. [28], proposed to match the user path with the map information to remove any outlier locations. Nonetheless,

these indoor localization techniques require the knowledge of indoor map (i.e., wall and hallway locations).

Recently, Simultaneous Localization and Mapping (SLAM) approaches, commonly used in the robotics domain, have been introduced to provide iterative and autonomous feature maps (i.e., anchor positions and signatures; not the indoor floorplan) of previously unknown environments and localization of a subject within this environment simultaneously. For instance, ActionSLAM [12] leverages foot-mounted inertial sensors to track the user's indoor location using detected user actions as observations. However, foot-mounted sensors are not ubiquitous and the recognized activities cannot be accurately detected by standard smartphone sensors, limiting this technique scalability. To extend SLAM to standard smartphones, SemanticSLAM [1] leverages phone inertial sensors to continuously estimate the user's location and detect supervised and unsupervised anchors in the building structure (e.g., stairs, elevators, etc.) with unique patterns that surface on the phone sensors when passing through them. These anchors provide error resetting opportunities to control the inertial sensors error accumulation. Although SemanticSLAM achieves good localization accuracy, its performance is contingent on the anchor density [1], which may be limited in many scenarios when the user moves in only one floor or not passing by any elevation-change anchor; losing many opportunities for error-resetting.

In this paper, we propose *DynamicSLAM*, a novel SLAM framework that can provide accurate, robust, low-overhead, and ubiquitous indoor localization. *DynamicSLAM* is a plug-and-play indoor localization system, where no infrastructure support nor war driving are needed for the operation of the system; making it building-independent. As in tradi-

- Ahmed Shokry is with the Department of Computer Science, Alexandria University, Egypt.
E-mail: ahmed.shokry@alexu.edu.eg
- Moustafa Elhamshary and Moustafa Youssef are with Egypt-Japan University for Science and Technology (E-JUST), Alexandria, Egypt.
E-mail: moustafa.elhamshary,moustafa.youssef@ejust.edu.eg

Under submission December 26, 2018.

tional SLAM, *DynamicSLAM* tracks the user's location using dead reckoning based on the phone inertial sensors and leverages *stationary anchors* to curb the accumulated location error. However, *to further increase the anchor density and hence the localization accuracy in any environment*, *DynamicSLAM* introduces the novel concept of **human mobile anchors**. Specifically, *DynamicSLAM* leverages humans walking in the environment as anchors for error-resetting opportunities. To achieve this, *DynamicSLAM* needs to address a number of challenges, including recognizing users' encounters in a ubiquitous manner, determining their relative location, and handling the inherent error in the users' estimated locations.

We evaluate the performance of *DynamicSLAM* in a real deployment based on different Android phones. Our results show that it can achieve a median localization error of 1.1m, outperforming other state-of-the-art techniques by more than 55%. In summary, we provide the following contributions in this paper:

- **We present a novel collaborative simultaneous localization and mapping framework that leverages other users in the environment as mobile anchors to enhance the localization and mapping accuracy in environments with low building anchor density.**
- **We present different ubiquitous observation models to detect users' encounters and estimate their relative distances.**
- **We present a theoretical proof for system convergence and the human anchors ability to reset the accumulated error.**
- **We implement and evaluate our system on different Android phones and quantify its performance relative to the state-of-the-art systems.**

The rest of the paper is organized as follows: Section 2 gives a brief background about the traditional SLAM framework. Section 3 gives an overview of the proposed system main components while Section 4 gives the details of the *DynamicSLAM* probabilistic framework. We evaluate the system performance in Section 5. Finally sections 6 and 7 discuss related work and conclude the paper respectively.

2 SLAM BACKGROUND

Simultaneous Localization and Mapping (SLAM) is a computational problem originally introduced in robotics where a mobile robot is placed at an unknown location in an unknown environment and the robot continuously localizes itself within this environment, while simultaneously incrementally building a consistent *feature* map of this environment [6]. The points constituting the feature map are also called anchors [1]. In SLAM, the trajectory of the robot and the locations and signatures of all anchors are estimated online without the need for any a priori knowledge. To do that, SLAM harnesses the robot odometers to keep track of the robot motion (i.e., to estimate its pose) as well as its exteroceptive sensors for detecting anchors in the environment to build the map. In a probabilistic form, the goal of SLAM is to find the pose s^t and map Θ that maximize the joint probability density function:

$$p(s_t, \Theta | u^t, z^t, n^t) \quad (1)$$

where $u^t = \{u_1, \dots, u_t\}$ is the set of motion update measurements (displacement and heading change) capturing the control history till time t , $z^t = \{z_1, \dots, z_t\}$ is the history of anchor location observations till time t , and $n^t = \{n_1, \dots, n_t\}$ is history of anchor identifications (data associations) where n_t specifies the identity of the anchor observed at time t .

The earliest SLAM algorithm [6] is based on the extended Kalman filter (EKF). The EKF represents the robot's map and position by a high-dimensional Gaussian density over all map anchors and the robot pose. The off-diagonal elements in the covariance matrix of this multivariate Gaussian represent the correlation between errors in the robot pose and the anchors in the map. Therefore, the EKF can accommodate the natural correlation of errors in the map. To estimate the joint probability density function in Equation 1, the EKF-SLAM algorithm factors it into two different models: the state transition model (motion update model) $p(s_t | u_t, s_{t-1})$ and the observation model $p(z_t | s_t, \Theta_{n_t}, n_t)$, describing the effect of the control input and observation respectively. The motion update model is assumed to be a Markov process in which the next position depends only on the immediately preceding position and the applied control u_k , and is independent of both the observations and the map. The observation model describes the probability of making an observation z_k when the robot location and anchor locations are known. EKF-SLAM approach has several drawbacks. First, its computation complexity is quadratic in the number of anchors [32]. Second, EKF assumes a Gaussian noise for the robot motion and observation. In this case, the amount of uncertainty in sensors must be relatively small, since the linearization in EKF may introduce intolerable noise. Third, EKF represents posteriors by a Gaussian, which is a unimodal distribution that works well only in tracking problems where the posterior is focused around the true state with a small margin of uncertainty. Gaussian posteriors are a poor match for many global estimation problems in which many distinct hypotheses exist, each of which forms its own mode in the posterior. This directly impacts the data association, i.e., how to determine the identity of the detected anchors when multiple of them have a similar signature (e.g. two nearby stairs), which can lead to different maps based on the chosen association.

FastSLAM [23], [25], [26] is an alternative SLAM solution that mitigates the limitation of EKF-SLAM. FastSLAM combines particle filter and extended Kalman filter. The basic idea of FastSLAM is to maintain a set of particles. Each particle contains a sampled robot path and its own representation of the map where each anchor in the map is represented by its own local Gaussian instead of one big joint Gaussian as in the case of EKF-SLAM algorithm. This is based on the SLAM property where the map anchors are conditionally independent given the path. By factoring out the path (one per particle), we can simply treat each map anchor independently, thereby avoiding the costly step of maintaining the correlations between them that plagued the EKF-SLAM approach. The resulting representation requires space linear in the size of the map, and linear in the number of particles. Estimating the posterior over all paths enables

the factorization of posterior over poses and maps:

$$p(s^t, \Theta | z^t, u^t, n^t) = p(s^t | z^t, u^t, n^t) \prod_{n=1}^{N_L} p(\Theta_n | s^t, z^t, n^t) \quad (2)$$

This factorization states that the calculation of the posterior over paths and maps can be decomposed into $N_L + 1$ probabilities, where N_L is the number of anchors. FastSLAM estimates the robot path s^t using a particle filter (Localization problem). The second term (Mapping problem) can be factored into N_L separate problems, where the individual anchor location probability density function $p(\theta_n | s^t, z^t, n^t)$ is estimated using an EKF. Formally, the posterior of the m^{th} particle ($S_t^{[m]}$) contains a path $s^{t,[m]}$ and N_L local estimates of anchors $\Theta^{[m]}$. A local estimate of anchor n by particle $[m]$ is represented by anchor type $\hat{f}_{n,t}$, the mean of anchor position $\mu_{n,t}^{[m]}$, and the position covariance $\sum_{n,t}^m$.

FastSLAM has many advantages over EKF-SLAM. First, FastSLAM can be implemented in time logarithmic in the number of anchors [32], offering computational advantages over the plain EKF-SLAM. Second, FastSLAM provides significant robustness in data association as it maintains posteriors over multiple data associations, and not just by tracking only a single data association at any point in time based on the incremental maximum likelihood data association incorporated in EKF-SLAM [26]. Finally, FastSLAM can also cope with non-linear models as particle filters can handle non-linear robot motion models and is guaranteed to converge under certain assumptions [26]. Due to these advantages, we adopt FastSLAM approach in *DynamicSLAM*.

3 SYSTEM OVERVIEW

Now, we start with an overview of *DynamicSLAM* core building blocks (Figure 1), postponing the discussion on the core SLAM algorithm to the next section.

3.1 User Mobile Traces

DynamicSLAM collects time-stamped sensor measurements including accelerometer, gyroscope, and magnetometer, and the received signal strength values from the available WiFi access points and Bluetooth. The collected raw sensor measurements are preprocessed to reduce the effect of noise by applying a low-pass filter, and phone orientation changes by transforming the sensor readings from the mobile coordinate system to the world coordinate system leveraging the inertial sensors [24].

3.2 Dead reckoning- Motion Model

DynamicSLAM relies on dead-reckoning to track the user's location starting from a reference point (e.g., the last GPS fix at the building entrance [2]). Dead-reckoning leverages inertial sensors (accelerometer, magnetometer, and gyroscope) where the current user's location is estimated from the previous location, distance traveled by the user, and the direction of her motion during the traveling distance. The user direction of motion can be estimated from the magnetometer and/or the gyroscope [34], while the displacement can be obtained from the accelerometer [3].

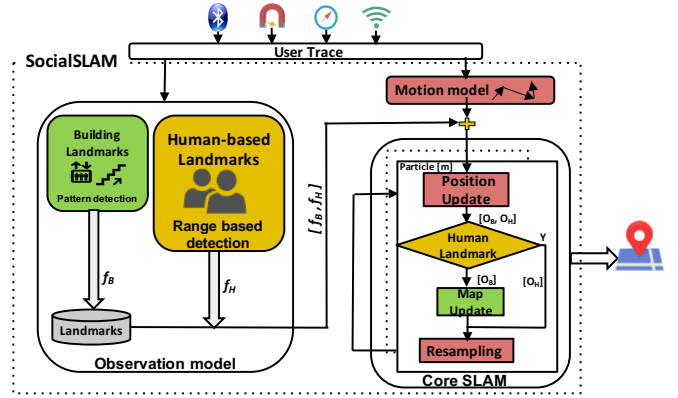


Fig. 1: *DynamicSLAM* architecture.

3.3 Anchors Detection- Observation Model

The dead-reckoning accumulated location error is unbounded [34]. Hence, it is infeasible for indoor tracking. To mitigate this error, we propose an anchor-based displacement error resetting technique. In particular, we leverage ample and unique points in the indoor environment to reset the accumulated error when the user hits one of them. These anchor points (anchors) can be estimated in a crowd-sourced manner, removing the need for any labor-intensive calibration. Traditional anchors are either physical structure/machines or organic anchors that exist in *buildings* (e.g., escalators, elevators, stairs, turns, electronic machines, etc.) [1], [34]. *DynamicSLAM* introduces a novel class of *human-based* anchors, where encountering other system users is used as dynamic mobile anchors, increasing the anchor density and making them more ubiquitously available. The details of extracting the two anchor classes and handling the associated challenges are discussed in the next section.

3.4 SLAM Algorithm

DynamicSLAM extends the FastSLAM algorithm to incorporate its motion and observation models. Specifically, when *DynamicSLAM* detects a user step, it triggers a position update where each particle updates its location belief as detailed in Section 4. When the system detects a human anchor (i.e., encountering another user), the particle location is refined by applying an extended Kalman filter on the relative users' locations and the covariance matrix representing the uncertainty in relative location estimates. The EKF helps to filter out the noise in users relative location estimation. **Two approaches** that trade-off the accuracy and efficiency are proposed in *DynamicSLAM* to fuse the two users' locations. On the other hand, if the particle detects a traditional building anchor, both the particle location and the map are updated to account for the detection of these anchors. Finally, in both cases, the particles importance weights are estimated to reflect the confidence of the location associated with each particle. The weights of particles are initialized equally and updated at each step based on the likelihood of the detected anchors (*user* and *building* anchors). Finally, a resampling step is performed using the current weights in order to refine the particles and drop those that significantly deviated from the actual user location. In the next

section, we will explain the details of the SLAM algorithm incorporated in *DynamicSLAM*.

4 THE DYNAMICSLAM SYSTEM

This section elaborates on the main components of *DynamicSLAM*: the anchor detection and the core SLAM algorithm as well as provides a proof of the system convergence. Finally, we discuss the possible ways to deploy the system.

4.1 Anchor Detection

In this section, we first briefly discuss the detection of the traditional building anchors. We then elaborate on the details of the detection algorithm for the *novel human-based anchors*.

4.1.1 Building anchors

Buildings have many unique points which present identifiable signatures on one or more cell-phone sensors. These unique patterns can be leveraged to automatically recognize these anchors. Once a user sees a anchor, *DynamicSLAM* detects its signature and leverages it to reset the dead-reckoning error. The list of physical anchors starts by the building's door entrance, detected by the loss of the GPS signal, which is used as a reference point to seed the dead-reckoning process [2]. Other indoor anchors include escalators, elevators, stairs, turns, and doors which can be detected using a finite state machine by employing the approach in [34]. Also, there may exist some anchors that are specific to some buildings. For instance, railway stations and airports have many specific machines (e.g., ticketing machines, lockers, electronic gates, etc) that can be detected using the different phone sensors [8].

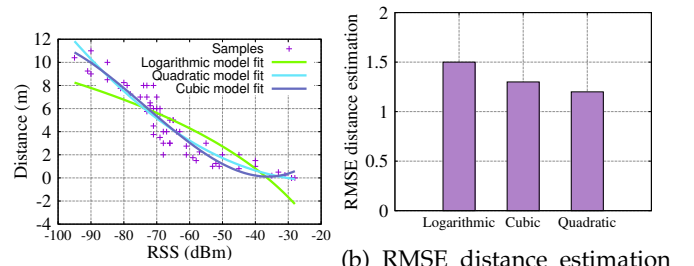
Finally, there are many organic anchors that may have a unique signature on the phone sensors without having a physical meaning to humans, e.g., areas with magnetic distortion caused by nearby electromagnetic waves. These points (organic anchors) can be detected using unsupervised learning techniques and can also be leveraged as anchors [1].

Note that since both anchor types (physical and organic) have a unique signature in the building, there is no need to differentiate between the different types of building anchors during processing. This differentiation is only relevant during the anchor extraction process.

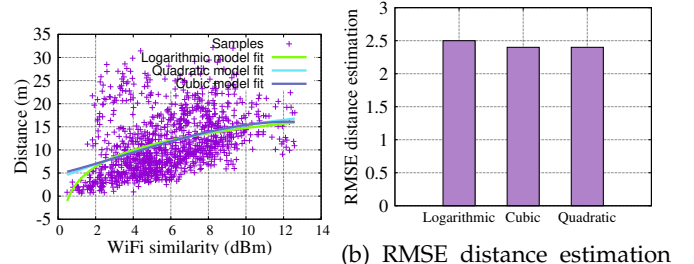
DynamicSLAM learns anchors locations through crowd-sourcing, which converges to the true anchor location over time [1], [34]. In addition, in some buildings, a map with the location of the building anchors, such as elevators and stairs, may be available. This can further speed the convergence process. Nonetheless, *DynamicSLAM* models the uncertainty of the anchor locations in its operation.

4.1.2 Human-based anchors

The density of anchors differs from one building to another. If the anchors are not dense in a given building, there are fewer opportunities for curbing the dead-reckoning accumulated error, affecting the overall localization accuracy. To alleviate this problem, *DynamicSLAM* introduces the concept of *human-based mobile anchors*, where other nearby users in the environment are used to reset the localization



(a) Fitting to raw data. (b) RMSE distance estimation for different Bluetooth models.
Fig. 2: Comparison between different Bluetooth models.



(a) Fitting to raw data. (b) RMSE distance estimation for different WiFi models.
Fig. 3: Comparison between different WiFi models.

error. The main intuition is that, when two users encounter each other, they can calculate their relative distances, and consequently, both users can refine their internal beliefs; reducing each other uncertainty in their estimated locations. *DynamicSLAM* presents *two different measurement models* to detect the users' encounter and estimate their relative distances based on the available signals and leading to different ubiquity and accuracy. These are the Bluetooth and the WiFi location models.

- 1) **The Bluetooth Model:** A user overhearing Bluetooth signals emitted from nearby users can directly infer their encounter. The Received Signal Strength (RSS) can be used to estimate the distance between the nearby users. We have conducted experiments to capture the relation between the Bluetooth RSS and the relative distance between users. We also show the quality of fitting using different fitting functions, including those based on the traditional known propagation models [4], [9], [17].

Figure 2 shows the results. The figure confirms that the quadratic fitting function leads to the best approximation. This is consistent with previous results reported in literature [9], [17].

The user encounters detection can be performed locally on the device and forwarded to the server for processing.

- 2) **The WiFi Model:** This model is based on the WiFi signals. The idea is that users in the vicinity of each other should hear similar APs with close RSS values. Therefore, to detect the encounter of two users, we check the similarity of their collected WiFi RSS measurement vectors. If the similarity is greater than a certain threshold, we detect the users' encounter.

To determine this threshold, we conducted an experiment to build a spatial physical-distance to WiFi-space distance model. Our WiFi-similarity measure takes into account the wireless channel noise that may lead to different number of APs heard in the same location at

different times. Specifically, the WiFi similarity measure is calculated as:

$$\text{Sim}_{\text{WiFi}} = \frac{||U_1 - U_2||}{n} \quad (3)$$

where n is the number of *commonly heard* access points by both users and U_i is the RSS vector of the common APs heard by user i . Note that the normalization by the number of APs makes the similarity measure independent of the number of heard APs. Moreover, one can restrict the RSS vector to the strongest N APs to reduce the noise effect. This is quantified in Section 5.

Figure 3a shows the relation between the spatial space Euclidean distance and the WiFi-space similarity measure. Figure 3b further compares the accuracy of the estimated relative users distance using different fit models. The figure shows that different models give comparable performance. We therefore use the logarithmic model as it is commonly used in the propagation literature [4]. The figure also shows that the WiFi samples are noisier than the Bluetooth samples. This is due to the higher range of WiFi signals compared to Bluetooth, which makes them more susceptible to noise. Finally, since WiFi chips cannot scan for other users' WiFi signal during their normal operation¹, WiFi-based users' encounters are performed at the *DynamicSLAM* server, where the WiFi scans of nearby APs are collected from each user and a module on the server calculates the similarity and performs the encounter detection.

4.2 Core SLAM Algorithm

DynamicSLAM is a variant of the FastSLAM [26] framework optimized to operate with anchors that can be detected based on the limited capabilities of phone-embedded sensors, as compared to robot sensors. *DynamicSLAM* uses the particle filter to enhance user's position. It relies on dead-reckoning to track the users location and **predicts** the distance (i.e. offset) between the user location and anchor location (Motion model / Prediction model). On the other hand, it relies on the Bluetooth/WiFi models to get the **estimated** distance (Observation model). Finally, *DynamicSLAM* combines both models using the particle filter equations.

DynamicSLAM core algorithm consists of three steps: location sampling (motion update), map update (observation update), and re-sampling. Without loss of generality, we assume that only a single anchor is observed at any time instance t to simplify the problem complexity. Note that, since *DynamicSLAM* needs around 150ms to get the estimated location while the typical human speed is 1.4m/s; the anchor can be assumed stationary during the estimation process. Table 1 summarizes the notations used in the paper.

4.2.1 Motion Update

It is invoked when a step is detected based on the phone inertial sensors. At this time instance, the control signal

1. WiFi scans are performed between the WiFi chip on the client and the APs installed in the environment to determine the best AP to associate to. To scan for other nearby users' WiFi signals, the device needs to be in monitoring mode, which disables the normal transmission operation.

TABLE 1: definitions and notations

Notation	Definition
\hat{u}_t	Control data at time t
\hat{l}_t	Estimated displacement at time t from the sensors
$\hat{L}_t^{[m]}$	Estimated step length for particle $[m]$ at time t
$\hat{\phi}_t$	Estimated heading change at time t
$\Phi_t^{[m]}$	Sampled heading change for the particle $[m]$ at time t
$\hat{s}_t^{[m]}$	Predicated location at time t for the particle $[m]$
s_t	User location at time t
$s_t^{a,[m]}$	User a 's location as seen by particle $[m]$ at time t
$s_{n,t}^{[m]}$	User location as seen by particle $[m]$ after refining using anchor n at time t
S^t	entire location history ($s^t = s_1, s_2, \dots$)
$\mu_{s_t}^{a,[m]}$	Mean of estimated location s_t of particle $[m]$ of user a at time t
$\Sigma_{s_t}^{a,[m]}$	Covariance matrix of estimated location s_t of particle $[m]$ of user a at time t
$\mu_{s_{n,t}}^{[m]}$	Mean of estimated location s_t of particle $[m]$ given anchor n at time t
$\Sigma_{s_{n,t}}^{[m]}$	Covariance matrix of estimated location s_t of particle $[m]$ given anchor n at time t
$z_{n,t}^a$	Estimated measurement of anchor n position at time t by user a
$\hat{z}_{n,t}^{a,[m]}$	Predicted measurement of particle $[m]$ for user a of observed anchor n at time t
$Q_{n,t}^{[m]}$	Covariance matrix of observing anchor n at time t by particle $[m]$
P_t	Covariance matrix of control data at time t
R_t	Measurement covariance at time t
$K_t^{[m]}$	Kalman gain for particle $[m]$ at time t
θ	anchor map
$\theta_{n,t}$	Location of anchor n at time t
$N_t^{[m]}$	Number of anchors for particle $[m]$ at time t
$\mu_{n,t}^{[m]}$	Mean of anchor n location at time t for particle $[m]$
$\Sigma_{n,t}^{[m]}$	Covariance matrix of anchor n at time t for particle $[m]$
$w_t^{[m]}$	Weight of particle $[m]$ at time t
f_t	anchor type (e.g. elevator)
\hat{f}_t	Detected anchor type (e.g. elevator pattern)
p_0	Likelihood of observing an anchor for the first time
p_n	Likelihood correspondence of anchor n with the observed pattern f at time t
η	Normalization factor
$J_{n,t}$	The Jacobian of measurement model for anchor n at time t
β	A threshold on the number of iterations for the iterative sampling technique

$\hat{u}_t = \langle \hat{l}_t, \hat{\phi}_t \rangle$, where \hat{l}_t and $\hat{\phi}_t$ are the estimated step length and heading from the phone inertial sensors, updates the user's location from s_{t-1} to s_t . More formally, given the control signal \hat{u}_t at time t , we draw each particle $[m]$ according to:

$$\hat{s}_t^{[m]} \sim P(s_t^{[m]} | s_{t-1}^{[m]}, \hat{u}_t) \quad (4)$$

Assuming Gaussian-distributed error for both \hat{l}_t and $\hat{\phi}_t$, the sampled displacement and heading for particle $[m]$ at time t are calculated as:

$$L_t^{[m]} \sim N(\hat{l}_t, \sigma_l), \phi_t^{[m]} \sim N(\hat{\phi}_t, \sigma_\phi) \quad (5)$$

where σ_l and σ_ϕ are the variance of the displacement and heading estimation errors respectively. Therefore, the new sampled location for a given user in Eq. 4 can be rewritten as:

$$s_t^{[m].\phi} = s_{t-1}^{[m].\phi} + \phi_t^{[m]} \quad (6)$$

$$s_t^{[m].x} = s_{t-1}^{[m].x} + L_t^{[m]} \cdot \cos(s_t^{[m].\phi}) \quad (7)$$

$$s_t^{[m].y} = s_{t-1}^{[m].y} + L_t^{[m]} \cdot \sin(s_t^{[m].\phi}) \quad (8)$$

where $s_t^{[m].\phi}$, $s_t^{[m].x}$, $s_t^{[m].y}$ are the new user heading, x, and y coordinates of the location respectively after incorporating the control input \hat{u}_t .

4.2.2 Observation Update

At any time instance, the user may observe a human-based (O_{LH}) or building (O_{LB}) anchor. In both cases, the current user's location belief is updated. In addition, for the case of the building anchor, the feature map is also updated to refine the observed anchor location. The balance of this subsection describes the details of both cases.

Observing a Human-based anchor When user *Alice* encounters another user *Bob*, her location should be updated based on their relative distances z_t and this distance covariance matrix, i.e., the confidence in the estimated location. Note that, since each user is considered an anchor for the other user, **no map update** is invoked in this case. To update *Alice's* location based on *Bob's* encounter, we propose two different techniques that trade-off the accuracy and efficiency: One-shot sampling and iterative sampling.

1) One-shot Sampling:

To account for the relative distance measurement error between the two users, we incorporate the EKF to update the user's location belief based on the observation. The EKF captures the deviation of the predicted measurement, $\hat{z}_{b,t}^{a,[m]}$, (based on the estimated users' locations) from the estimated measurement, $z_{b,t}^a$ (based on the two observation models described in Section 4.1.2), taking into account the confidence in the different quantities

Specifically, *Alice's* **predicted** distance measurement, i.e., the location offset between *Alice's* (a) and *Bob's* (b) locations is calculated as:

$$\hat{z}_{b,t}^{a,[m]} = \|\hat{s}_t^{a,[m].(x,y)} - s_t^{b,[m].(x,y)}\| \quad (9)$$

The mean and covariance matrix of *Alice's* location distribution after incorporating *Bob's* encounter is:

$$\mu_{s_{b,t}^a}^{a,[m]} = \hat{s}_t^{a,[m]} + K_t^{[m]}(z_{b,t}^a - \hat{z}_{b,t}^{a,[m]}) \quad (10)$$

$$\Sigma_{s_{b,t}^a}^{a,[m]} = [J_{s_{b,b}}^T (Q_{b,t}^{[m]})^{-1} J_{s_{b,b}}^T + P_t^{-1}]^{-1} \quad (11)$$

Where $J_{s_{b,b}}^T$ is the transpose of the Jacobian of the measurement model with respect to *Bob's* location $\mu_b = s_t^{b,[m].(x,y)}$, P_t is the covariance matrix of the control data at time t , $Q_{b,t}^{[m]}$ is the human-based anchor observation covariance matrix given by:

$$Q_{b,t}^{[m]} = J_{b,t} \Sigma_{b,t-1}^{[m]} J_{b,t}^T + R_t \quad (12)$$

and R_t is the measurement covariance matrix. Finally, $K_t^{[m]}$ is the Kalman gain given by:

$$K_t^{[m]} = \Sigma_{s_{b,t}^a}^{a,[m]} (J_{s_{b,b}}^a)^T (Q_{b,t}^{[m]})^{-1} \quad (13)$$

Finally, the new location of a particle $[m]$ of *Alice* after incorporating *Bob's* encounter is sampled from the following distribution:

$$s_{b,t}^{a,[m]} \sim N(\mu_{s_{b,t}^a}^{a,[m]}, \Sigma_{s_{b,t}^a}^{a,[m]}) \quad (14)$$

Intuitively, when a user encounters another user, her location is updated based on her previous state, the confidence of the two users' locations, and the innovation

Algorithm 1 Refine_Particle_Position ($\hat{s}_t^{[m]}$, $[m]$, n)

Input:

$\hat{s}_t^{[m]}$: Predicated location at time t for the particle $[m]$.
 $[m]$: Particle index.
 n : Anchor index.

- 1: $\hat{z}_{n,t}^{[m]} = \|\mu_{n,t-1}^{[m]} - \hat{s}_t^{[m]}\|$ ▷ predict measurement
 - 2: $K_t^{[m]} = \Sigma_{s_{n,t}^{[m]}} J_{s,n}^T (Q_{n,t}^{[m]})^{-1}$ ▷ Kalman gain
 - 3: $\Sigma_{s_{n,t}^{[m]}} = [J_{s,n}^T Q_{n,t}^{[m]-1} J_{s,n} + P_t^{-1}]^{-1}$
 - 4: ▷ covariance of proposal distribution
 - 5: $\mu_{s_{n,t}^{[m]}} = \hat{s}_t^{[m]} + K_t^{[m]} (z_{n,t} - \hat{z}_{n,t}^{[m]})$
 - 6: ▷ mean of proposal distribution
 - 7: $s_{n,t}^{[m]} \sim N(\mu_{s_{n,t}^{[m]}}, \Sigma_{s_{n,t}^{[m]}})$ ▷ sample pose
-

Algorithm 2 DynamicSLAM (z_t, u_t, S_{t-1})

Input:

S_{t-1} : System state at time $t-1$.
 u_t : Control data at time t .
 z_t : Measurement observation at time t .

- 1: **for** $m = 1$ to M **do** ▷ particle m
 - 2: Retrieve $\langle \langle s_{t-1}^{[m]}, w_{t-1}^{[m]} \rangle, \langle \mu_{1,t-1}^{[m]}, \Sigma_{1,t-1}^{[m]}, f_{1,t-1}^{[m]} \rangle, \dots, \langle \mu_{N,t-1}^{[m]}, \Sigma_{N,t-1}^{[m]}, f_{N,t-1}^{[m]} \rangle \rangle$ from S_{t-1}
 - 3: ▷ get particle's old state
 - 4: $\hat{s}_t^{[m]} \sim P(s_t | s_{t-1}^{[m]}, \hat{u}_t)$ ▷ motion model- predict pose
 - 5: **if** Human anchor h is observed **then**
 - 6: ▷ observation model
 - 7: $Refine_Particle_Position(\hat{s}_t^{[m]}, [m], h)$
 - 8: ▷ refine user's pose
 - 9: **else** ▷ building anchor
 - 10: **for** $n = 1$ to $N_{t-1}^{[m]}$ **do** ▷ calculate sampling distribution
 - 11: $Refine_Particle_Position(\hat{s}_t^{[m]}, [m], n)$
 - 12: ▷ refine user's pose
 - 13: $p_n = |2\pi Q_{n,t}|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\hat{z}_{n,t}^T Q_{n,t}^{-1} \hat{z}_{n,t})\} p(f_t | \hat{f}_t)$
 - 14: ▷ association likelihood
 - 15: **end for**
 - 16: $p_{1+N_{t-1}^{[m]}} = \eta p o p(f_t | \hat{f}_t)$ ▷ likelihood of new anchor
 - 17: $\hat{n} = \underset{\{n \in 1, \dots, 1+N_{t-1}^{[m]}\}}{\text{argmax}} p_n$ ▷ maximum likelihood
 - 18: $N_t^{[m]} = \max\{N_{t-1}^{[m]}, \hat{n}\}$ ▷ new number of anchors
 - 19: $Map_Update(N_{t-1}^{[m]}, \hat{n}, [m])$ ▷ map update
 - 20: $w_t^{[m]} = w_{t-1}^{[m]} p_n^{[m]}$ ▷ importance weight
 - 21: **end if**
 - 22: **end for**
 - 23: **loop** M times
 - 24: Draw with replacement a random index m with probability $\propto w_t^{[m]}$ ▷ resampling
 - 25: **end loop**
 - 26: add $\langle \langle s_t^{[m]}, w_t^{[m]} \rangle, \langle \mu_{1,t}^{[m]}, \Sigma_{1,t}^{[m]}, f_{1,t}^{[m]} \rangle, \dots, \langle \mu_{N_t,t}^{[m]}, \Sigma_{N_t,t}^{[m]}, f_{N_t,t}^{[m]} \rangle \rangle$ to S_t
-

(the difference between the estimated and predicted relative distance). Algorithm 1 summarizes the process of refining the particle position.

2) Iterative Sampling:

In this technique, the one-shot sampling described in the previous section is repeated on the same pair of users until the particle position converges. To determine convergence, the algorithm stops when the difference between two successive location updates for particle

Algorithm 3 Map_Update ($N_{t-1}^{[m]}$, \hat{n} , $[m]$)**Input:**

$N_{t-1}^{[m]}$: Number of anchors seen by particle $[m]$ at time $t-1$.
 \hat{n} : anchor index with the highest probability.
 $[m]$: Particle index.

```

1: if  $\hat{n} = 1 + N_{t-1}^{[m]}$  then           ▷ New anchor detected
2:    $\mu_{\hat{n},t}^{[m]} = s_t^{[m]}$              ▷ initialize its mean
3:    $\Sigma_{\hat{n},t}^{[m]} = R_t$              ▷ initialize its covariance
4: else
5:    $K_t^{[m]} = \Sigma_{\hat{n},t-1}^{[m]} (Q_{\hat{n},t,t}^{[m]})^{-1}$    ▷ Kalman gain
6:    $\mu_{\hat{n},t}^{[m]} = \mu_{\hat{n},t-1}^{[m]} - K_t^{[m]} \hat{z}_{\hat{n},t}^{[m]}$    ▷ update mean
7:    $\Sigma_{\hat{n},t}^{[m]} = (I - K_t^{[m]}) \Sigma_{\hat{n},t-1}^{[m]}$    ▷ update covariance
8: end if

```

$[m]$ is less than a certain threshold β .

The iterative sampling approach has the advantage of increased accuracy at the expense of increased computations (as we quantify in Section 5).

Observing a Building Anchor: Similar to the human-based anchors, encountering a building-based anchor, e.g., an elevator, updates the user's position using the EKF. In addition, the system needs to (I) resolve the uncertainty in the type of detected anchors, i.e., stairs vs elevators vs other types, due to the inherent noise in the phone sensors and (II) Update the building anchor location.

(I) Data Association (Detecting building anchor type) To account for this, we compute the conditional probability of actually observing each anchor type (f) when the detected pattern is of type \hat{f} . This uncertainty is a function of the sensors noise and is independent from the building. It is typically described by a confusion matrix, where each entry indicates the probability $p(f|\hat{f})$ of the anchor (f) given pattern \hat{f} . Given this confusion matrix and the anchor location uncertainty Q_n for each anchor n in the particle's map, the likelihood of correspondence with the observed pattern is calculated as [1]

$$p_n^{[m]} = \eta |2\pi Q_{n,t}^{[m]}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} (\hat{z}_{n,t}^{[m]})^T Q_{n,t}^{[m]-1} \hat{z}_{n,t}^{[m]}\right\} \cdot p(f_t|\hat{f}_t) \quad (15)$$

Where η is a normalization factor. This likelihood takes into account the distance between the current user's location and the building anchor, the location uncertainty of both, and the confusion between the observed and actual anchor types (f_t vs \hat{f}_t).

We assume that the probability of observing a newly unseen anchor given the detected pattern is calculated as:

$$p_n^{[m]} = \eta p_0 \cdot p(f_t|\hat{f}_t) \quad (16)$$

where p_0 is a constant determined empirically. (II) *Updating anchor Location (Map update)* The anchor with the highest probability (\hat{n}_t) is selected as the observed anchor and its location is updated using the standard EKF formulas. (Algorithm 3).

4.2.3 Resampling

The last step assigns a weight to each particle to reflect the confidence of its location. The particles' weights are initialized uniformly and updated at each step based on the

likelihood of the detected anchors, i.e., the anchor \hat{n}_t with the highest probability, as:

$$w_t^{[m]} = w_{t-1}^{[m]} \cdot p_{\hat{n}_t} \quad (17)$$

These weights are normalized to add to one. Then, the location and the map of the particle having the maximum weight are selected as the current estimate. Finally, the algorithm draws with replacement M new particles to avoid degeneration. Algorithm 2 summarizes the full *DynamicSLAM* algorithm.

4.3 DynamicSLAM Convergence

In this section, we argue about the system convergence as well as the effectiveness of human anchors ability to reset the accumulated error. Here convergence refers to reducing the error in the users' locations to zero. Assume that user x encounters an anchor n (human or building anchor). We define the error in user x location $\Psi_t^{x,[m]}$ as follows:

$$\Psi_t^{x,[m]} = \mu_{s_t}^{x,[m]} - s_t^x \quad (18)$$

Note that the ground-truth user location s_t^x is used in the previous equation only to prove *DynamicSLAM* convergence and is not needed during the actual system operation. From equations 6-8 and 9, $\hat{s}_t^{x,[m]} = s_{t-1}^{x,[m]} + u_t$, $\hat{z}_t^{x,[m]} = \mu_{n,t-1}^{[m]} - s_{t-1}^{x,[m]} - u_t$. Also, by setting $J_n = I$, $J_s = -I$, and $Q_{n,t}^{[m]} = R_t + \Sigma_{n,t-1}^{[m]}$, the mean and covariance of Eq. 10 and 11 of the proposal distribution are defined as a follows

$$\mu_{s_{n,t}}^{x,[m]} = -\Sigma_{s_{n,t}}^{x,[m]} (R_t + \Sigma_{n,t-1}^{[m]})^{-1} \cdot (z_{n,t}^a - \mu_{n,t-1}^{[m]} + s_{t-1}^{x,[m]} + u_t) + s_{t-1}^{x,[m]} + u_t \quad (19)$$

$$\Sigma_{s_{n,t}}^{x,[m]} = [(R_t + \Sigma_{n,t-1}^{[m]})^{-1} + P_t^{-1}]^{-1} \quad (20)$$

First, we study the human anchor ability to reset the accumulated error.

Lemma 1. If user a with error $\Psi_t^{a,[m]}$ encounters another user b with error $\Psi_t^{b,[m]}$, where $\Psi_t^{a,[m]} < \Psi_t^{b,[m]}$, then $\Psi_t^{b,[m]}$ will shrink in expectation. Conversely, if $\Psi_t^{a,[m]} > \Psi_t^{b,[m]}$, then the later may increase but will not exceed $\Psi_t^{a,[m]}$ in expectation.

Proof The expected error of user a at time t is given by:

$$E[\Psi_t^{a,[m]}] = E[\mu_{s_t}^{a,[m]} - s_t^a] = E[\mu_{s_t}^{a,[m]}] - E[s_t^a] \quad (21)$$

We can get the first term from the sampling distribution (Eq 19) which depends on user location, anchor location, estimated measurements, and the uncertainty in anchor and user locations. The second term can be obtained from the motion model (Eq 7 and 8). Therefore:

$$E[\Psi_t^{a,[m]}] = -\Sigma_{s_b,t}^{a,[m]} (R_t + \Sigma_{b,t-1}^{[m]})^{-1} \cdot (E[z_{b,t}^a] - \mu_{b,t-1}^{[m]} + s_{t-1}^{a,[m]} + u_t) + \Psi_{t-1}^{a,[m]} \quad (22)$$

For linear SLAM, $E[z_{b,t}^a] = s_t^{b,[m]} - E[s_t^{a,[m]}] = s_t^{b,[m]} - u_t - s_{t-1}^{a,[m]}$. Hence,

$$E[z_{b,t}^a] - \mu_{b,t-1}^{[m]} + s_{t-1}^{a,[m]} + u_t = s_t^{b,[m]} - u_t - s_{t-1}^{a,[m]} - \mu_{b,t-1}^{[m]} + s_{t-1}^{a,[m]} + u_t = \Psi_{t-1}^{a,[m]} - \Psi_{t-1}^{b,[m]} \quad (23)$$

By substitution from equations 20 and 23 in 22,

$$E[\Psi_t^{a,[m]}] = \Psi_{t-1}^{a,[m]} + [I + (R_t + \Sigma_{b,t-1}^{[m]})P_t^{-1}]^{-1} \cdot (\Psi_{t-1}^{b,[m]} - \Psi_{t-1}^{a,[m]}) \quad (24)$$

As R_t , $\Sigma_{b,t-1}^{[m]}$, and P_t^{-1} matrices are positive semidefinite, the value of $[I + (R_t + \Sigma_{b,t-1}^{[m]})P_t^{-1}]^{-1}$ is a contraction matrix. Hence, $E[\Psi_t^{a,[m]}]$ depends on the difference $\Psi_{t-1}^{b,[m]} - \Psi_{t-1}^{a,[m]}$. If $\Psi_{t-1}^{b,[m]} < \Psi_{t-1}^{a,[m]}$, then user a location error will be reduced in expectation. If $\Psi_{t-1}^{b,[m]} > \Psi_{t-1}^{a,[m]}$, then the $E[\Psi_t^{a,[m]}]$ will increase by a certain amount depending on the difference, which makes the upper bound for the expected error, $E[\Psi_t^{a,[m]}]$, $\Psi_{t-1}^{b,[m]}$. Now, we study the effect of a user that encounters a building anchor with known location.

Lemma 2. If a user a encounters a building anchor with known location n^* , then the error in her location will shrink in expectation.

Proof The building anchor with known location has zero error $\Omega_t^{n^*,[m]} = 0$ and $\Sigma_{n^*,t}^{[m]} = 0$. By substitution in Eq. 24

$$E[\Psi_t^{a,[m]}] = \Psi_{t-1}^{a,[m]} - [I + R_t \cdot P_t^{-1}]^{-1} \cdot \Psi_{t-1}^{a,[m]} \quad (25)$$

The following theorem proves our system convergence.

Theorem 3. *DynamicSLAM* converges even for $M=1$ particles if there is at least one building anchor whose location is known in advance.

Proof Assuming that the building contains at least one of building anchor with known location, some of the system users will eventually observe the building anchors, which according to Lemma 2 will lead to reducing their error. Similarly, when two users encounter each other, the one with the lower accuracy will enhance his estimate based on the other user (from Lemma 1). Therefore, the error of the different users in *DynamicSLAM* will always reduce, leading to convergence.

4.4 Discussion

DynamicSLAM can have different deployment modes. In this paper, we assumed that the system operations are performed on a centralized server. However, the system can be completely implemented on devices that communicate with each others in a peer-to-peer manner. The devices in this case will broadcast their own estimated locations, confidence in the estimated locations, and RSS heard from other devices. Note that, unlike the traditional collaborative SLAM-based techniques, *DynamicSLAM* does not need to exchange particles local maps as the users' locations are updated based on their local maps only.

The peer-to-peer implementation is more private. It can solve the latency and connection problems due to the centralized process. However, making all computations on smart-phones may be an overhead due to the limited computational power on smart-phones. In both cases, cryptographic approaches can be used to prevent malicious users from introducing errors in the system, e.g. by broadcasting fake or wrong locations.

TABLE 2: Default parameter values

Parameter	Range	Default value
Measurement Model	Bluetooth, WiFi	WiFi
Location update technique	One-shot sampling, Iterative sampling	Iterative sampling
Number of particles	1 to 100	75
User inverse density (m^2 /user)	3 to 3000	30
RSS threshold (Bluetooth model)	-70 to -95	-90
WiFi threshold (WiFi model)	2 to 14	8
Number of APs (WiFi model)	1 to 16	5
Iterative sampling threshold β	0 to 1	0.1

5 EVALUATION

In this section, we describe the testbed and the data collection methodology followed by evaluating the performance of the system components under different parameter settings. We end the section by comparing *DynamicSLAM* performance against other systems and quantifying its energy footprint.

5.1 Testbed & Data Collection Methodology

We deployed *DynamicSLAM* in the engineering building in our university campus. The building area is $3000m^2$ containing offices, labs, meeting rooms as well as corridors. The building has 30 building anchors (physical and organic). The data is collected by 23 users using different Android phones (e.g., LG Nexus 5, Samsung Galaxy Note 3, Galaxy 4, Galaxy Tab, among others). The data is then processed in a centralized server. Each user walked around in a random independent path in the building for at least 250m passing by some building anchors (10 on average) as well as encountering other users (7 on average). If a user encounters another group of users, the system selects the user with the highest confidence.

5.2 Effect of Changing *DynamicSLAM* Parameters

In this section, we study the effect of the different parameters on the system performance including the user encounter detection method, user density in the building, sampling method, and number of particles. Table 2 shows the default parameters values used throughout the evaluation section.

5.2.1 Users' encounter observation models

Bluetooth model Figure 4 shows the accuracy of the estimated relative users' distance (in terms of Root Mean Square Error) at different cut-off Bluetooth RSSs. The figure shows that the higher the cut-off threshold, i.e. the stronger the Bluetooth signal, the lower the relative distance error. However, requiring a strong Bluetooth signal reduces the users' encounter opportunities due to the lower range. For example, setting the cut-off RSS threshold to -70dBm, the encounter detection range is around 4m (From Figure 2), and the RMSE of relative distance is around 1m.

WiFi Model Similarly, decreasing the similarity threshold for the WiFi signal leads to better accuracy at the expense of a shorter range (Figure 5). We use a WiFi threshold that is equal to 8dBm in our experiments. Figure 6 further shows the effect of controlling the number of APs used in the similarity calculation. The figure shows that increasing the number of the strongest APs used in similarity calculations leads to higher accuracy until we reach an optimal point at five APs. This can be explained by noting that a few number of APs leads to ambiguity in the user location while a large

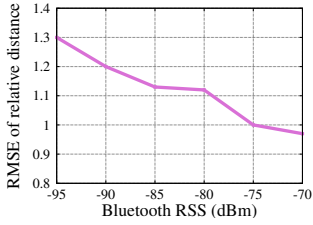


Fig. 4: Relative distance error versus different Bluetooth RSSs.

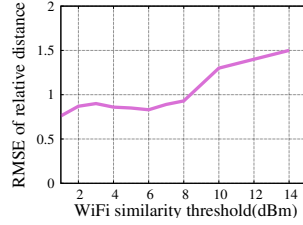


Fig. 5: Relative distance error versus WiFi similarity threshold.

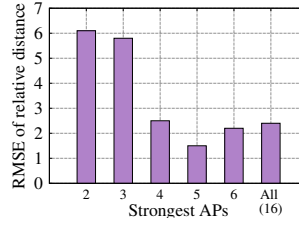


Fig. 6: Relative distance error using different number of WiFi APs.

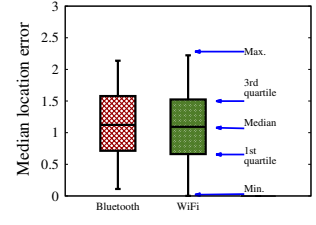
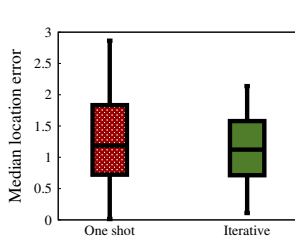
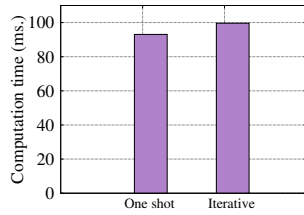


Fig. 7: Effect of relative distance observation models on the location error.

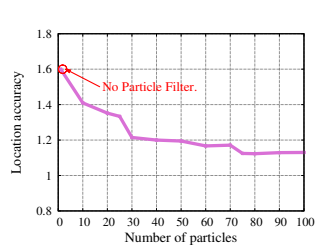


(a) Localization error.

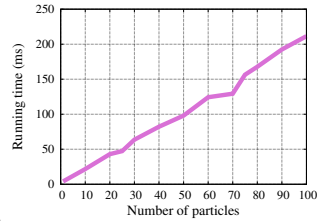


(b) Computation time per estimate.

Fig. 8: Comparison of system performance using one-shot sampling vs iterative sampling.



(a) Localization error.



(b) Computation time per estimate.

Fig. 9: Effect of number of particles on system performance.

number of APs adds noise to the APs vector, both lead to reduced accuracy.

Models Comparison Figure 7 compares the box-plot of relative distance error estimation by the two encounter observation models. Evident from the figure, the Bluetooth and WiFi models have comparable performance. This is an interesting result as one should expect that Bluetooth should lead to better accuracy due to its shorter range. However, WiFi has more than one AP in a scan, compared to only one device for Bluetooth. This compensates for the noise due to the higher range of WiFi. Finally, we note that the two observation models are independent and can be used together to provide both higher range and better accuracy.

5.2.2 Sampling Method

Figure 8a shows the box-plot of the two sampling techniques used to update the user’s location: one-shot and iterative sampling. The figure confirms that iterative sampling outperforms one-shot sampling in the median as well as the different quantile errors. This comes at the expense of increased computation time as shown in Figure 8b.

5.2.3 Number of Particles

Figure 9a shows the effect of increasing the number of particles on performance. The figure shows that as more particles

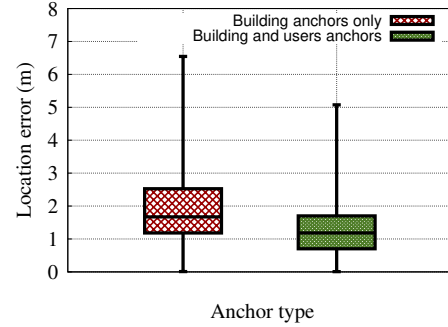


Fig. 10: Effect of using building anchors only and using users anchors on localization accuracy.

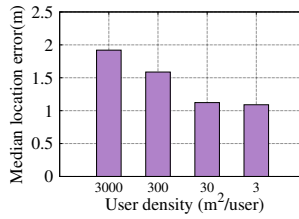


Fig. 11: Effect of (inverse) user density on localization accuracy.

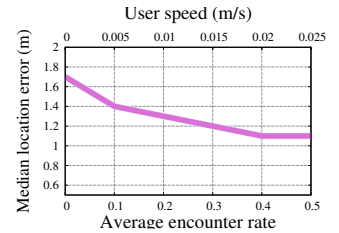


Fig. 12: Effect of average encounter rate on localization accuracy.

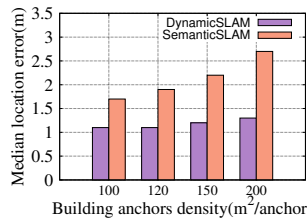


Fig. 13: Effect of (inverse) building anchor density on localization accuracy.

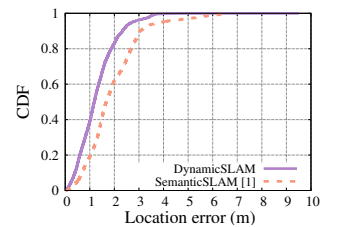


Fig. 14: Comparison between DynamicSLAM and SemanticSLAM.

are used to represent the user’s location, the localization accuracy increases. This highlights that the particle filter does capture the dynamics of the system. The figure also shows that as more particles are used to represent the user’s location, the localization accuracy increases. Nonetheless, the computation time of the system increases linearly with the increase in the number of particles as shown in Figure 9b. To find a comprise of the computation time- accuracy tradeoff, *DynamicSLAM* uses only 75 particles.

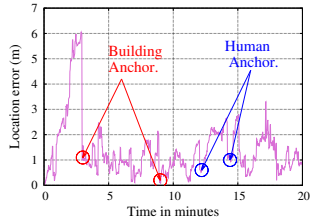


Fig. 15: An example of accuracy over time.

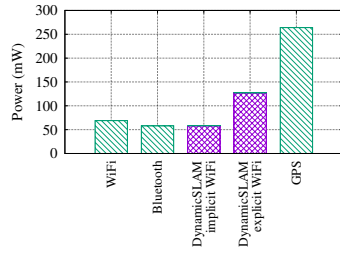


Fig. 16: Power consumption.

TABLE 3: Summary of the location accuracy percentiles of *DynamicSLAM* vs *SemanticSLAM*.

Technique	25 th percentile	50 th percentile	75 th percentile	90 th percentile	Maximum
<i>DynamicSLAM</i>	0.7	1.1	1.7	2.3	5.1
<i>SemanticSLAM</i> [1]	1.2 (-170%)	1.7 (-54%)	2.6 (-52%)	3.0 (-30%)	6.6 (-29%)

5.2.4 User Density

Figure 10 shows the effect of using the traditional building anchors only and using the new dynamic anchors on localization accuracy. Figure 11 shows the effect of the number of meters² per user (*inverse* of the user density) on the median localization error. Figure 12 further shows the effect of users' encounter rate on the localization accuracy. As more users move around, there is a higher probability for users' encounters, leading to more opportunities of individual user location calibration, and hence better localization accuracy. The secondary x-axis shows the user speed to achieve a certain number of encounters (linear relation) under the assumption that other users are uniformly and independently distributed in the area [21].

Note that, regardless of the user speed, since *DynamicSLAM* needs around 150ms to get the estimated location while the typical human speed is 1.4m/s [20]; the user can be assumed stationary during the estimation process.

5.3 Comparison With Other Systems

In this section, we compare the localization accuracy of *DynamicSLAM* against the *SemanticSLAM* technique that employs the SLAM approach based on the standard phone sensors *using only traditional building anchors*. Figure 13 shows the comparison between our proposed *DynamicSLAM* system and the *SemanticSLAM* system under different building anchors densities. The figure shows that reducing the building anchor densities affects the *SemanticSLAM* system accuracy while *DynamicSLAM* maintains its accuracy. This is due to the fact that *DynamicSLAM* leverages the users in the environment as anchors. Hence, its performance is less affected by the building anchors density.

Figure 14 also shows the CDF of the localization accuracy for both systems and Table 3 summarizes the results. *DynamicSLAM* achieves a median distance error around 1.1m, which is better than *SemanticSLAM* by 55%. Moreover, *DynamicSLAM* reduces the worst case error by 29% and enhances all other quantiles. This is due to the new human-based mobile anchors that provide ample opportunities for error resetting as illustrated in Fig 15.

5.4 Energy Footprint

DynamicSLAM harnesses inertial sensors, WiFi, and Bluetooth for its operation. The inertial sensors are always running to detect the phone orientation are used in a variety of mobile applications. Therefore, using them consumes virtually zero extra energy. Similarly, many users turn-on WiFi indoors to connect the Internet. Thus, as many sensors are activated by default for other purposes, this curbs the overall power consumption of *DynamicSLAM*. Moreover, to further enhance *DynamicSLAM* energy footprint, we adopt an adaptive sensing strategy where the sensors sampling rate and their activation are based on the nearby anchor density. For example, if a certain area with high anchor density, *DynamicSLAM* can decrease the sensor sampling rate as the location error will be frequently calibrated by these anchors. Additionally, Bluetooth and WiFi can be turned off if the server predicts that the probability of different users' encounter is small. Figure 16 compares the power consumption (calculated using the PowerTutor [36]) of *DynamicSLAM* against the GPS (at 2Hz) (used as a baseline).

6 RELATED WORK

6.1 Indoor Localization

Mobile phone localization has been well-studied in literature [15], [27], [29], [30], [31], [34]. The most ubiquitous indoor localization techniques are either WiFi-based or dead-reckoning based. WiFi-based techniques [4], [19], require calibration to create a prior "RF map" for the building to counter the wireless channel noise. This calibration process is time consuming, tedious, and requires periodic updates. This led to the emergence of new calibration-free techniques [7], [27]. These techniques, however, suffer from low location accuracy or requiring explicit user intervention. In dead-reckoning based system [15], [34], the inertial sensors on mobile phones are leveraged to dead-reckon the user starting from a reference point. Nevertheless, dead-reckoning error quickly accumulates with time; leading to a complete deviation from the actual path. Therefore, many techniques have been proposed to reset the dead-reckoning error including snapping to environment anchor points, such as elevators and stairs [1], [34] and matching with the map information [28]. *DynamicSLAM* combines dead-reckoning as its motion model with building anchors and its novel human-based anchors, as observation models, that are ubiquitously in any building gives it an ample opportunities to reset the dead-reckoning error.

6.2 Encounter-based Localization

Recent work has explored more social aspects of indoor localization named encounters-based approaches [16], [22]. The basic idea behind these approaches is to leverage the useful proximity information between users to enhance the localization accuracy. SocialLoc [16] utilizes users' meeting or missing information to correct localization errors. A user may intersects with another user, which is defined as an encounter. If the user does not meet another specific user during the localization process, such an event is defined as non-encounter. SocialLoc utilizes these encounter and non-encounter events to cooperatively correct the localization

errors. SocialSpring [22] is a general system that sits on top of an indoor positioning system. It leverages the pairwise distance measurements between the different users in the environment to refine position estimates.

DynamicSLAM, on the other hand, combines building anchors with human-based anchors (users' encounter) in a collaborative SLAM framework to enhance the localization accuracy.

6.3 Simultaneous Localization and Mapping

SLAM [6] is a well-known technique in the robotics domain where robot motion and observation models (e.g., odometers, laser, ultrasound) are cost-prohibitive for smartphones. Recently, a number of SLAM-based indoor localization systems have been proposed for tracking human users. For instance, *WiFi-SLAM* [11] builds a WiFi map based on a Gaussian Process Latent Variable Model; which allows for detecting the latent-space locations of unlabeled signal strength data. *ActionSLAM* [12] leverages a dedicated foot-mounted inertial sensor to track the user's motion while using actions uniquely related to certain locations as anchors for the user's location correction. Recently, *SemanticSLAM* [1], generalizes the SLAM concept to work with smartphones without using any external hardware. It harnesses smartphone inertial sensors to dead-reckon the user's location and leverages anchor points (elevator, escalator, stairs), which can be uniquely detected by phone sensors for the accumulated error resetting. These systems, however, either require special hardware or are dependent on a limited set of anchors that are not ubiquitously available in all buildings. *DynamicSLAM*, on the other hand, combines building anchors with the newly introduced human-based anchors that are ubiquitously in any environment.

6.4 Multi-Robot SLAM

Multi-Robot SLAM is an extension for SLAM that can benefit from having multiple robots in many places to carry out diverse tasks at the same time [5], [10], [13], [14], [18], [33], [35], [37]. The typical ultimate target of Multi-Robot SLAM systems is to construct the joint (common) map of an unknown environment, i.e. a virtual map of the anchors the robots see. For this, they use the *traditional/static building anchors* to reset the robots' accumulated error. For instance the work in [13], uses robots' specific sensors (laser range-finders) to detect robot pairs encounter. Then, by determining their relative position using laser sensors, the system combines the observations from both robots into the common map. Hence, it *indirectly* refines robots' locations by obtaining a more accurate common map.

DynamicSLAM, on the other hand, does enhance the environment map (traditional anchors) using a similar approach to multi-robot SLAM. In addition, we also leverage the other users (other robots in multi-robot SLAM) as a new type of anchors. Specifically, we leverage other users (robots) as a temporary anchor to *directly* refine users' positions and reset accumulated error. This leads to increased anchor density and hence better localization accuracy as quantified in the evaluation section.

7 CONCLUSIONS

We presented *DynamicSLAM*: an accurate, low-overhead, and collaborative simultaneous localization and mapping system. *DynamicSLAM* tracks the user's location using dead-reckoning and leverages physical anchors existing in buildings as well a novel human-based dynamic anchors to reset the accumulated localization error in a unified probabilistic framework. A number of methods to detect users' encounters and to handle their relative distance measurement error have also been proposed. Furthermore, we presented a theoretical proof of system convergence as well as the human anchors ability to reset the accumulated error.

Evaluation of *DynamicSLAM* in a typical university building using different Android phones shows that it can achieve a median distance error of 1.1m, which outperforms other state-of-the-art indoor localization techniques by approximately 55%, highlighting its promise for ubiquitous indoor localization.

REFERENCES

- [1] H. Abdelnasser, R. Mohamed, A. Elgohary, M. F. Alzantot, H. Wang, S. Sen, R. R. Choudhury, and M. Youssef. Semanticslam: Using environment landmarks for unsupervised indoor localization. *IEEE Transactions on Mobile Computing*, 15(7):1770–1782, 2016.
- [2] M. Alzantot and M. Youssef. Crowdinside: automatic construction of indoor floorplans. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, pages 99–108. ACM, 2012.
- [3] M. Alzantot and M. Youssef. Uptime: Ubiquitous pedestrian tracking using mobile phones. In *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, pages 3204–3209. IEEE, 2012.
- [4] P. Bahl and V. N. Padmanabhan. Radar: An in-building rf-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 775–784. Ieee, 2000.
- [5] A. Birk and S. Carpin. Merging occupancy grid maps from multiple robots. *Proceedings of the IEEE*, 94(7):1384–1397, 2006.
- [6] H. Durrant-Whyte and T. Bailey. Simultaneous localization and mapping: part i. *IEEE robotics & automation magazine*, 13(2):99–110, 2006.
- [7] R. Elbakly and M. Youssef. A robust zero-calibration rf-based localization system for realistic environments. In *Sensing, Communication, and Networking (SECON), 2016 13th Annual IEEE International Conference on*, pages 1–9. IEEE, 2016.
- [8] M. Elhamshary, M. Youssef, A. Uchiyama, H. Yamaguchi, and T. Higashino. Transitlabel: A crowd-sensing system for automatic labeling of transit stations semantics. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, pages 193–206. ACM, 2016.
- [9] S. Feldmann, K. Kyamakya, A. Zapater, and Z. Lue. An indoor bluetooth-based positioning system: Concept, implementation and experimental evaluation. In *International Conference on Wireless Networks*, volume 272, 2003.
- [10] J. W. Fenwick, P. M. Newman, and J. J. Leonard. Cooperative concurrent mapping and localization. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*, volume 2, pages 1810–1817. IEEE, 2002.
- [11] B. D. Ferris, D. Fox, and N. Lawrence. Wifi-slam using gaussian process latent variable models. 2007.
- [12] M. Hardegger, D. Roggen, S. Mazilu, and G. Tröster. Actionslam: Using location-related actions as landmarks in pedestrian slam. In *Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on*, pages 1–10. IEEE, 2012.
- [13] A. Howard. Multi-robot simultaneous localization and mapping using particle filters. *The International Journal of Robotics Research*, 25(12):1243–1256, 2006.
- [14] A. Howard, G. S. Sukhatme, and M. J. Mataric. Multi-robot mapping using manifold representations. In *ICRA*, volume 4, pages 4198–4203. Citeseer, 2004.

- [15] Y. Jin, H.-S. Toh, W.-S. Soh, and W.-C. Wong. A robust dead-reckoning pedestrian tracking system with low cost sensors. In *Pervasive Computing and Communications (PerCom), 2011 IEEE International Conference on*, pages 222–230. IEEE, 2011.
- [16] J. Jun, Y. Gu, L. Cheng, B. Lu, J. Sun, T. Zhu, and J. Niu. Social-loc: Improving indoor localization with social sensing. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, page 14. ACM, 2013.
- [17] J. Jung, D. Kang, and C. Bae. Distance estimation of smart device using bluetooth. *Personal Computing Platform Research Team*, pages 13–8, 2013.
- [18] J. Ko, B. Stewart, D. Fox, K. Konolige, and B. Limketkai. A practical, decision-theoretic approach to multi-robot mapping and exploration. In *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453)*, volume 4, pages 3232–3238. IEEE, 2003.
- [19] A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, et al. Place lab: Device positioning using radio beacons in the wild. In *International Conference on Pervasive Computing*, pages 116–133. Springer, 2005.
- [20] R. V. Levine and A. Norenzayan. The pace of life in 31 countries. *Journal of cross-cultural psychology*, 30(2):178–205, 1999.
- [21] B. Liu, P. Brass, O. Dousse, P. Nain, and D. Towsley. Mobility improves coverage of sensor networks. In *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, pages 300–308. ACM, 2005.
- [22] P. Martin, Y. Shoukry, P. Swaminathan, R. W. Ouyang, and M. Srivastava. Social spring: encounter-based path refinement for indoor tracking systems. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, pages 156–159. ACM, 2014.
- [23] M. Michael, T. Sebastian, K. Daphne, and W. Ben. Fastslam 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI)*, volume 2, 2003.
- [24] N. Mohssen, R. Momtaz, H. Aly, and M. Youssef. It's the human that matters: accurate user orientation estimation for mobile computing applications. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, pages 70–79. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014.
- [25] M. Montemerlo and S. Thrun. *FastSLAM: A scalable method for the simultaneous localization and mapping problem in robotics*, volume 27. Springer, 2007.
- [26] M. Montemerlo, S. Thrun, D. Koller, B. Wegbreit, et al. Fastslam: A factored solution to the simultaneous localization and mapping problem. *Aaai/iaai*, 593598, 2002.
- [27] 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 *Proceedings of the 8th international conference on Mobile systems, applications, and services*, pages 271–284. ACM, 2010.
- [28] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen. Zee: Zero-effort crowdsourcing for indoor localization. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 293–304. ACM, 2012.
- [29] H. Rizk, M. Torki, and M. Youssef. Cellindeep: Robust and accurate cellular-based indoor localization via deep learning. *IEEE Sensors Journal*, 19(6):2305–2312, 2018.
- [30] A. Shokry, M. Elhamshary, and M. Youssef. The tale of two localization technologies: Enabling accurate low-overhead wifi-based localization for low-end phones. In *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, page 42. ACM, 2017.
- [31] A. Shokry, M. Torki, and M. Youssef. Deeploc: A ubiquitous accurate and low-overhead outdoor cellular localization system. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 339–348. ACM, 2018.
- [32] S. Thrun, W. Burgard, and D. Fox. *Probabilistic robotics*. MIT press, 2005.
- [33] S. Thrun and Y. Liu. Multi-robot slam with sparse extended information filers. In *Robotics Research. The Eleventh International Symposium*, pages 254–266. Springer, 2005.
- [34] H. Wang et al. No need to war-drive: Unsupervised indoor localization. In *Mobisys*. ACM, 2012.
- [35] S. B. Williams, G. Dissanayake, and H. Durrant-Whyte. Towards multi-vehicle simultaneous localisation and mapping. In *Proceedings 2002 IEEE international conference on robotics and automation (Cat. No. 02CH37292)*, volume 3, pages 2743–2748. IEEE, 2002.
- [36] Z. Yang. Powertutor—a power monitor for android-based mobile platforms. *EECS, University of Michigan*, retrieved September, 2:19, 2012.
- [37] X. S. Zhou and S. I. Roumeliotis. Multi-robot slam with unknown initial correspondence: The robot rendezvous case. In *2006 IEEE/RSJ international conference on intelligent robots and systems*, pages 1785–1792. IEEE, 2006.



Ahmed Shokry is with the Wireless Research Center at Alexandria University, Egypt. He received his B.Sc. in Computer and Systems Engineering from Alexandria University in 2013. He also received his M.Sc. in Computer Science and Engineering from the same university in 2019. His research interests include location determination systems, theoretical and experimental algorithms.



Moustafa Elhamshary is a late assistant professor at Tanta University, Egypt. Previously, he was a research fellow at Graduate School of Information Science and Technology, Osaka University, Japan. His research interests include location determination technologies, mobile sensing, human activity recognition and location-based social networks



Moustafa Youssef is a professor at Alexandria University and founder & director of the Wireless Research Center of Excellence, Egypt. His research interests include mobile wireless networks, mobile computing, location determination technologies, pervasive computing, and network security. He has tens of issued and pending patents. He is the Lead Guest Editor of the upcoming IEEE Computer Special Issue on Transformative Technologies, an Associate Editor for IEEE TMC and ACM TSAS, an Area Editor of Elsevier PMC, as well as served on the organizing and technical committees of numerous prestigious conferences. He is the recipient of the 2003 University of Maryland Invention of the Year award, the 2010 TWAS-AAS-Microsoft Award for Young Scientists, the 2013 and 2014 COMESA Innovation Award, the 2013 ACM SIGSpatial GIS Conference Best Paper Award, the 2017 Egyptian State Award, multiple Google Research Awards, among many others. He is an ACM Distinguished Speaker, an ACM Distinguished Scientist, and an IEEE Fellow.