Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention

Zheng Yang, Chenshu Wu, and Yunhao Liu
School of Software and TNList
Tsinghua University
{yang, wu, yunhao}@greenorbs.com

Introduction

- Indoor localization is being more and more popular
- Most of the former approaches use RSS
- Fingerprinting-based approaches include two phases:

(1) training: building fingerprints databases of interested area

Site survey is time-consuming, labor-intensive, vulnerable to environmental dynamics, but inevitable

Google released Google Map 6.0 available only at some selected airports and shopping malls

(2) operating: query the location according to fingerprints

Introduction(cont'd)

- Propose LiFS (Locating in Fingerprint Space)
  - ❑ Considering originally separated RSS fingerprints are geographically connected by user moving paths, hence form a high dimension fingerprint space
  - ❑ Reform the floor plan of a building to the stress-free floor plan
  - ❑The spatial similarity of stress-free floor plan and fingerprint space enables fingerprints labeled with real locations

Related work

Wireless localization techniques fall into two categories:

- Fingerprinting-based techniques:

Build a fingerprinting database first, then estimate the location by mapping the measured fingerprint against the database

- Model-based techniques:

Calculate locations based on geometrical models rather than search for best-fit signatures from pre-labeled reference databases, such as prevalent log-distance path loss

Related work(cont'd)

Simultaneous Localization and Mapping:

- Standard SLAM: based on robotics and computer vision. Unreasonable for smart phone-based localization
- WiFi-SLAM: use Gaussian process latent variable models to relate RSS fingerprints and model human movements. Need a small portion of tagged fingerprints
- GraphSLAM: improve WiFi-SLAM regarding with efficiency and relying assumptions

Related work

Multidimensional Scaling:

- A set of related statistical techniques often used in data visualization
- Input: a matrix of item-item dissimilarities, output: a d-dimensional space
- Some researchers use MDS to calculate locations of wireless devices:

  Sensor nodes are capable of measuring the distances to neighboring nodes by RSS, ToA, TDoA, etc. MDS is used to assign a coordinate to each node such that the measured inter-node distances are as much preserved as possible.

Data Collection

- User participation is essential in the initial period
- Untrained users walk in a building following daily activities
- Mobile phones, carried by users, collect WiFi RSS characteristics along user movement paths
- Walking distances are measured as footsteps
- Establish the geographical relationship among fingerprints

System Architecture

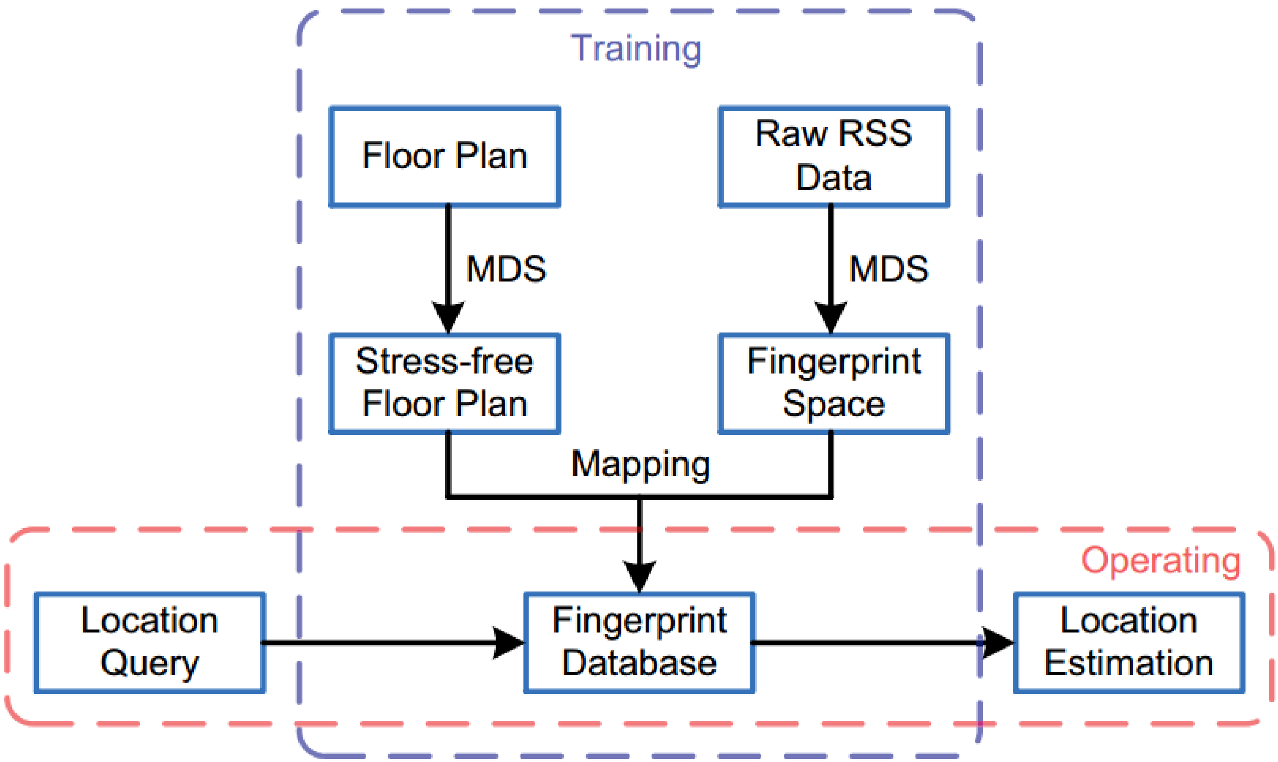

Figure 1: System architecture.

Stress-free Floor Plan

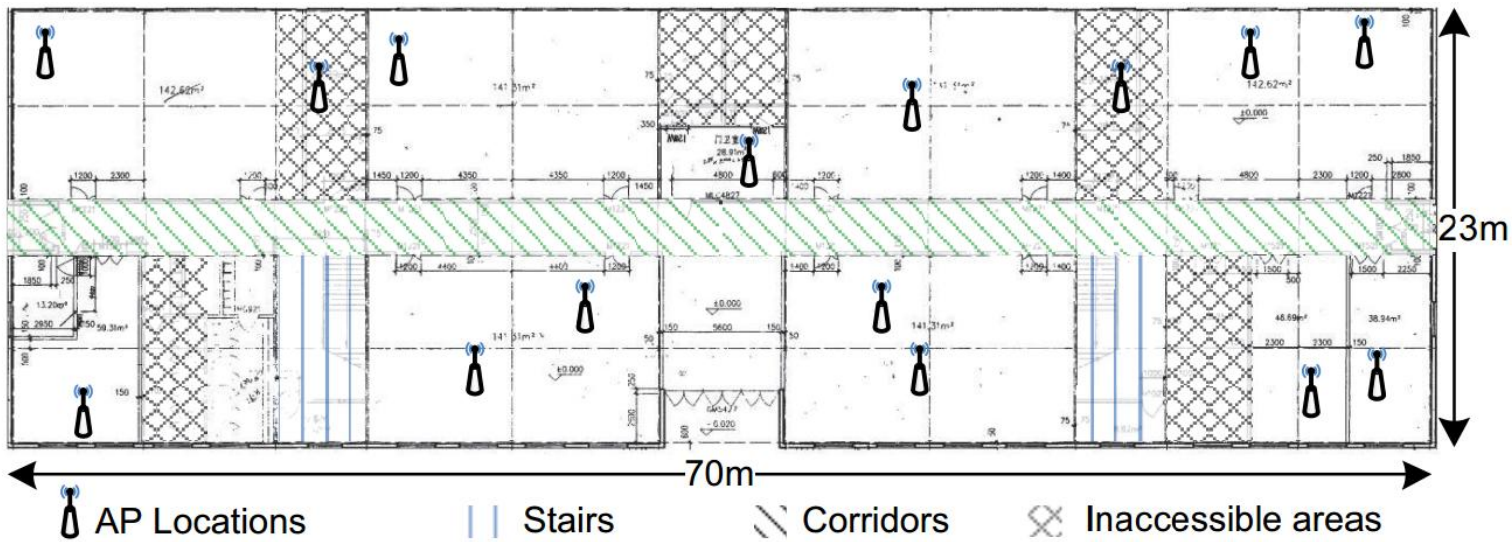

Figure 2: Floor plan of the experiment field.

Problem: The geographical distance between two locations in a floor plan does not necessarily equal to the walking distance between them due to the block of walls and other obstacles.

Stress-free Floor Plan(cont'd)

To address the distance mismatch problem, we propose the concept of stress-free floor plan:

- Sample an area at the intersecting locations of a mesh of grids in a floor plan
- The length l of a grid can be 1-3 meters
- Calculate the walking distance matrix between any two sample points.
- Use this matrix as input, MDS could output a d-dimension Euclidean space, d=2, 3

Stress-free Floor Plan(cont'd)

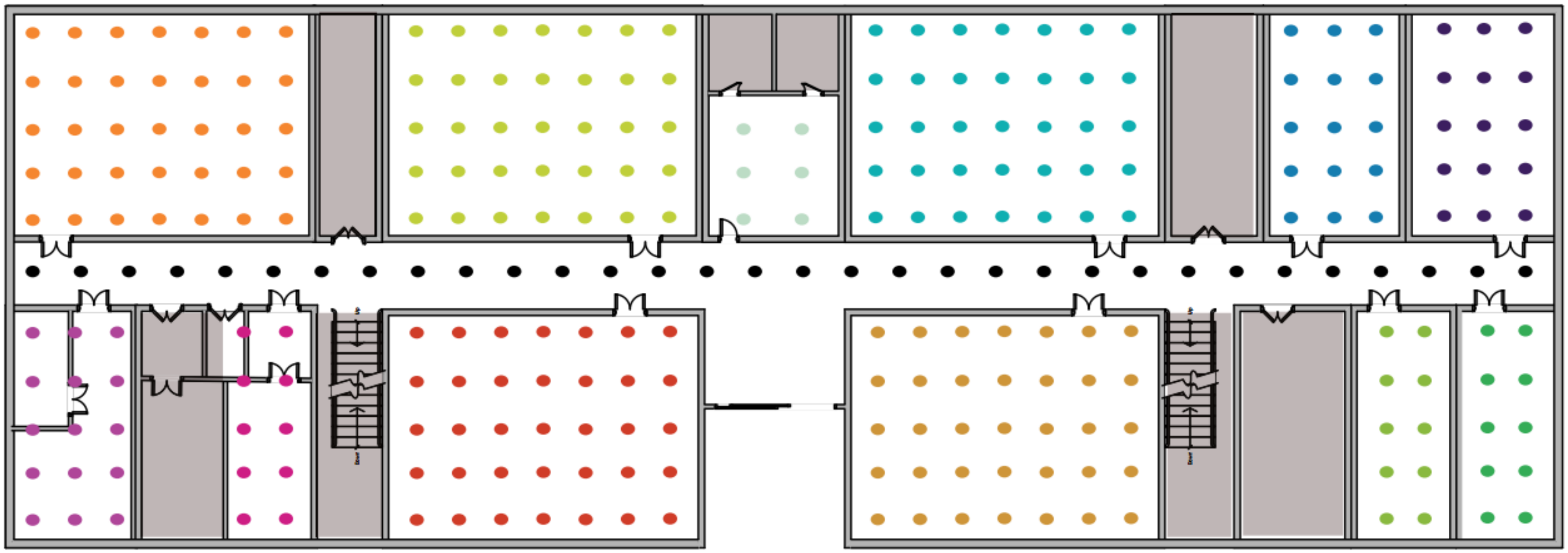

Figure 3: Floor plan with sample locations.

Stress-free Floor Plan(cont'd)

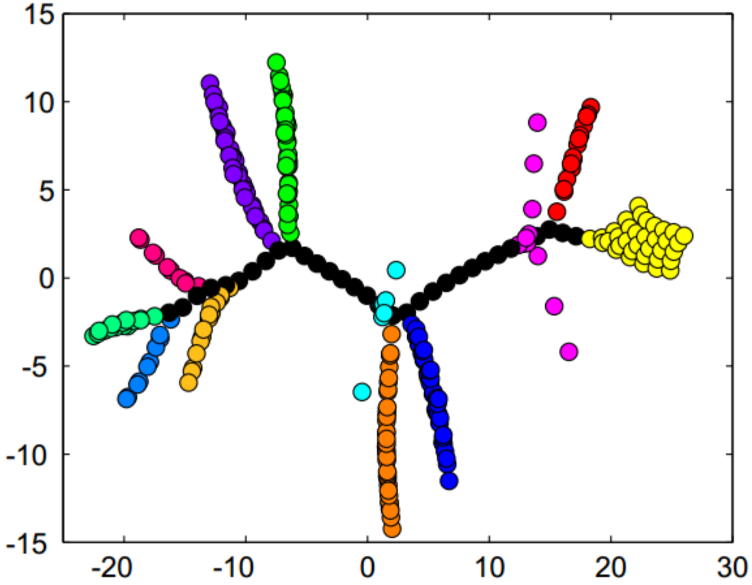

Figure 5: 2D stress-free floor plan.

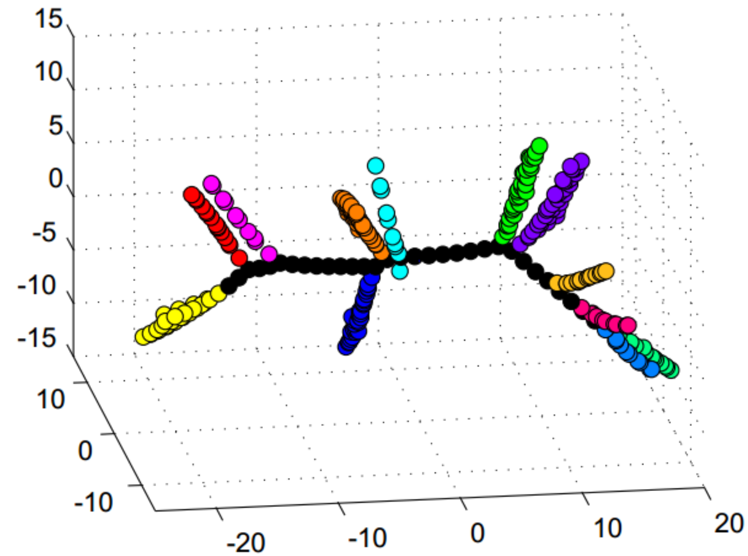

Figure 6: 3D stress-free floor plan.

Fingerprint space

- Fingerprint collection

Suppose m APs in an area A. For each location in A:

$$f = (s_1, s_2, \ldots, s_m)$$

After fingerprint collection, we have a set of fingerprints:

$$F = \{f_i, i = 1 \ldots n\} \text{ (} n \text{ is the number of records)}$$

and a distance matrix:

$$D' = [d'_{ij}]$$

Fingerprint space(cont'd)

- Pre-processing

1. Rule out fingerprints overlap

User movements are arbitrary
→ Walking paths might be intersectant
→ Fingerprints might be overlapped

Generally, for two fingerprints $f_i = (s_1, s_2, \ldots, s_m)$ and $f_j = (t_1, t_2, \ldots, t_m)$, define RSS difference (dissimilarity) δ between f_i, f_j as follows:

$$\delta_{ij} = \| f_i - f_j \|_1 = \sum_{k=1}^{m} |s_k - t_k|$$

If: $\delta_{ij} > \epsilon$ Treat them as two different points

Else: Merge them

Fingerprint space(cont'd)

2. Walking distance measurements

Traditional: calculated by integrating acceleration twice with respect to time
Presence of noise in accelerometer readings
→Error accumulates rapidly and can reach up to 100 meters

Solution: adopt the individual step counts as the metric of walking distance

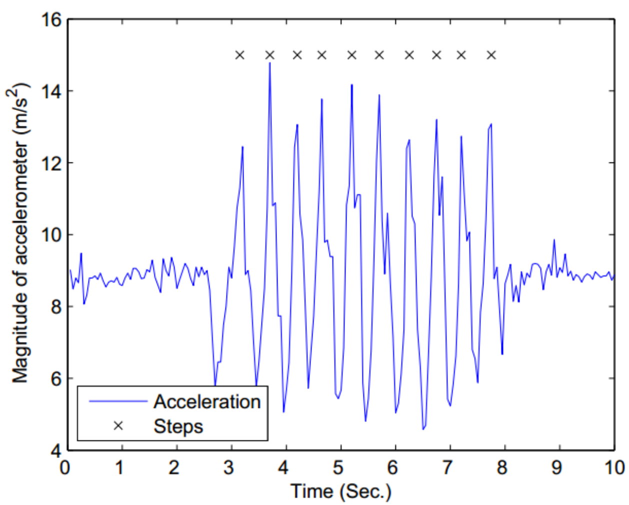

Figure 7: The acceleration pattern for 10 steps.

- Use local variance threshold method
- MDS tolerates measurement errors of stride lengths gracefully

$$\Delta := \begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \cdots & \delta_{1,I} \\ \delta_{2,1} & \delta_{2,2} & \cdots & \delta_{2,I} \\ \vdots & \vdots & & \vdots \\ \delta_{I,1} & \delta_{I,2} & \cdots & \delta_{I,I} \end{pmatrix}. \qquad \|x_i - x_j\| \approx \delta_{i,j}$$

$$x_1, \ldots, x_I \in \mathbb{R}^N$$

Wireless System Research Group

McMaster University ENGINEERING

Fingerprint space(cont'd)

- Fingerprint space construction
  - ❑ Operating phase of LiFS starts when the number of collected fingerprints reaches 10 times of the number of the sample locations
  - ❑ Refine the current fingerprint database uninterruptedly in operating phase
  - ❑ Value of d'[i,j] can be approximated as the length of the shortest path between f_i and f_j by passing several user path segments(Floyd-Warshall), update timely
  - ❑Use MDS map all fingerprints into a d-dimension Euclidean space

Fingerprint space(cont'd)

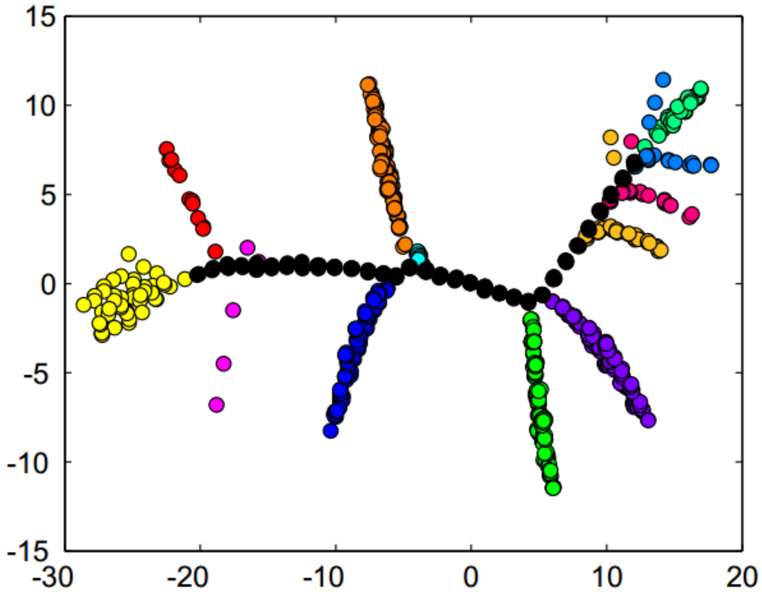

Figure 8: 2D fingerprint space.

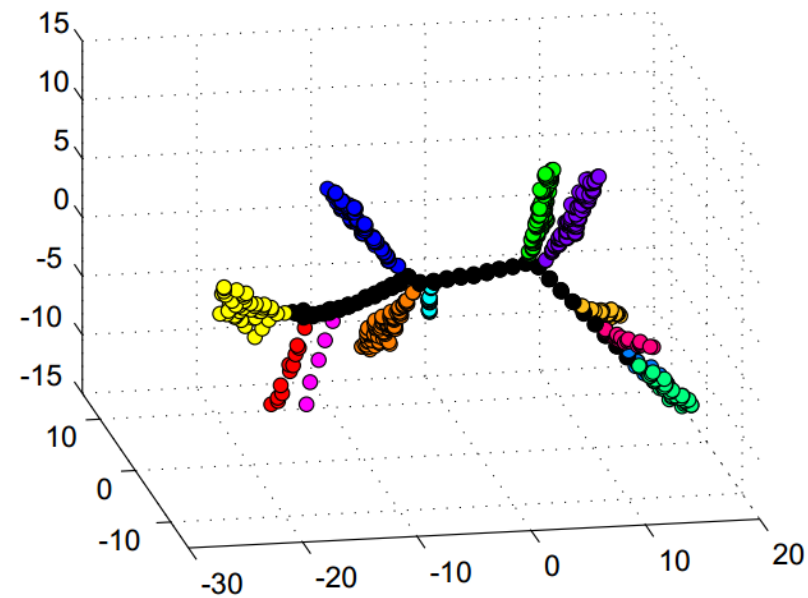

Figure 9: 3D fingerprint space.

Mapping

- Feature extraction

  1. Corridor Recognition

  ❑ fingerprints collected at corridors have a relatively large centrality values

  ❑ vertex centrality can be valued by degree, betweenness, closeness etc.

$$B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}},$$

where σ_{st} is the number of shortest paths from s to t, and $\sigma_{st}(v)$ is the number of shortest paths from s to t that pass through a vertex v.

Mapping(cont'd)

- According to the distance among fingerprints, build the Minimum Spanning Tree (MST)
- Compute the vertex betweenness for all vertices (fingerprints), and then use a watershed to distinguish the fingerprints from corridors and others areas

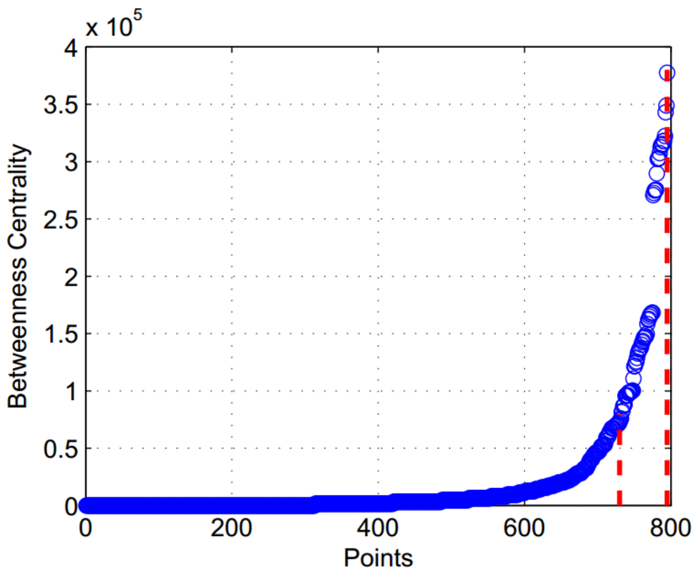

Figure 11: Betweenness distribution.

How to determine the watershed?

1, the area ratio c of corridors to entire floor plan

(i.e., c =size(corridor)/size(all))

2, the large gap of betweenness values of fingerprints

Mapping(cont'd)

2. Room Recognition

❑ Observation:
Remove the fingerprints from corridors, remaining fingerprints form several clusters that are apparently spatially separated

❑ Intuition: clustering method(k-means)

K is the numbers of room in the building

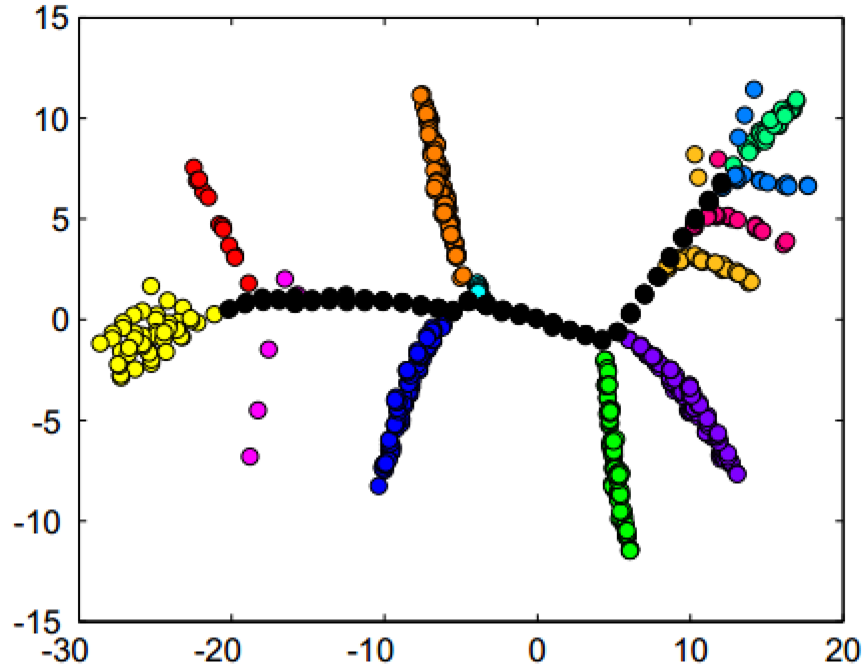

Figure 8: 2D fingerprint space.

Mapping(cont'd)

3. Reference Point Mapping

How to establish relationships stress-free floor plan and fingerprint space?

Doors are the keys, can be used as reference points!

Identify the fingerprints that are collected near doors(the points at each side of a door):

$$(\hat{f}_i, \hat{f}'_i) = \underset{f \in F_{R_i}, f' \in F_c}{\arg\min} \parallel f - f' \parallel,$$

Specifically, $\hat{f}_i$ and $\hat{f}'_i$ locate as close as possible to a door in the floor plan but in opposite sides ($\hat{f}_i$ inside the room and $\hat{f}'_i$ outside the room). Let $F_D = \{\hat{f}'_i, i = 1, 2, \ldots, k\}$

$F_D=(f_1, f_2, \ldots, f_k)$ Actually, the fingerprints in F_D can be organized in a chain in the MST

$P_D=(p_1, p_2, \ldots, p_k)$ The order of sample locations in P_D are in accord with their appearance from one side to the other side along the corridor

Mapping(cont'd)

There are two possible ways to map F_D to P_D:

$$\sigma_1 : f_i \mapsto p_i;$$

$(f_1, f_2, \ldots, f_k) \rightarrow (p_1, p_2, \ldots, p_k)$

$$\sigma_2 : f_i \mapsto p_{k-i+1}.$$

$(f_1, f_2, \ldots, f_k) \rightarrow (p_k, p_{k-1}, \ldots, p_1)$

$$l_i = \| p_{i+1} - p_i \|. \qquad l'_i = \| f_{i+1} - f_i \|.$$

The cosine similarity of:

$$(l_1, l_2, \ldots, l_{k-1})$$

$$(l'_1, l'_2, \ldots, l'_{k-1})$$

calculated by

$$\frac{l \cdot l'}{\| l \| \| l' \|}.$$

If $S_1 >= S_2$, choose σ_1
Else, choose σ_2

Mapping(cont'd)

- Space Transformation

1. Floor-level Transformation

$f_i \in F_D \qquad x_i = [x_i^1 \; x_i^2 \; \ldots \; x_i^d]^{\mathrm{T}}$

$p_i \in P_D \qquad y_i = [y_i^1 \; y_i^2 \; \ldots \; y_i^d]^{\mathrm{T}}$

$$y_i = Ax_i + B. \qquad k = |F_D|$$

We re-write the k equations as

$$H_i z = G_i,$$

where $H_i = [x_i^{\mathrm{T}} \; 1]$, $z = [A \; B]^{\mathrm{T}}$, and $G_i = y_i^{\mathrm{T}}$. Combining k equations as a matrix equation, we have

$$Hz = G,$$

The least square estimation [3] of above k equations gives

$$\bar{z} = (H^{\mathrm{T}}H)^{-1}H^{\mathrm{T}}G,$$

which minimizes $\| G - Hz \|$.

Mapping(cont'd)

2. Room-level Transformation

A door belongs to only one room
→rooms and the fingerprints from corresponding rooms are also related

Using doors and room corners as reference points, the fingerprints and sample locations are linked determinately by the transformation matrix above discussed.

Perform the above step one room by one room and finally achieve a full mapping for all fingerprints after multiple steps of room-level transformation.

Discussion

- Global Reference Point

eg. The last reported GPS location, AP's location etc.

Not used in LiFs, but they can be integrated into LiFS, resulting in a more robust mapping solution

- Building Type

The experiment field in this paper is one floor of an academic building

This solution fits a majority of office buildings but may fail in large open environments, such as hall, atrium, gymnasium or museum, in which users' movements are difficult to characterize.

Experiments

Experiment Design

- One floor of a typical office building covering 1600m2, with the length of 70m and width of 23m, 16 office rooms,
- Totally m = 26 APs are installed
- Sample the experiment floor plan approximately every4m^2 (2m × 2m grid) and obtain 292 sample locations
- The experiment lasts five hours by 4 volunteers.
- Fingerprints are recorded every 4~5 steps during moving
- Totally 600 user traces along with 16,498 fingerprint records are collected. These traces cover most of the areas of the experimental field.

Experiments(cont'd)

Performance

$$Location\_Error = ||L(f) - L'(f)||,$$

$$Room\_Error = \frac{1}{N}\sum_{f\in F} I(R(f) \neq R'(f)),$$

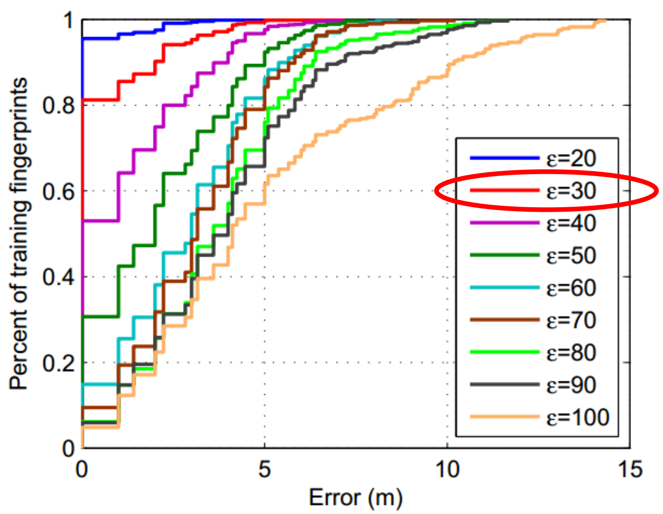

Figure 12: CDF of location error on different ϵ.

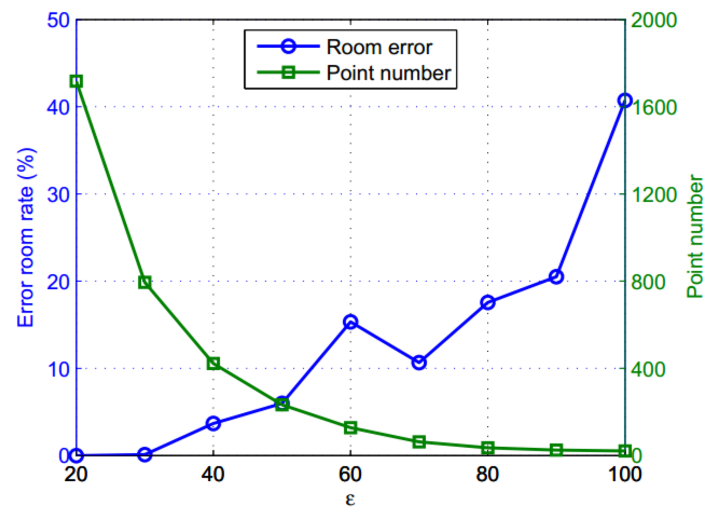

Figure 13: Room error rate vs. ϵ.

Experiments(cont'd)

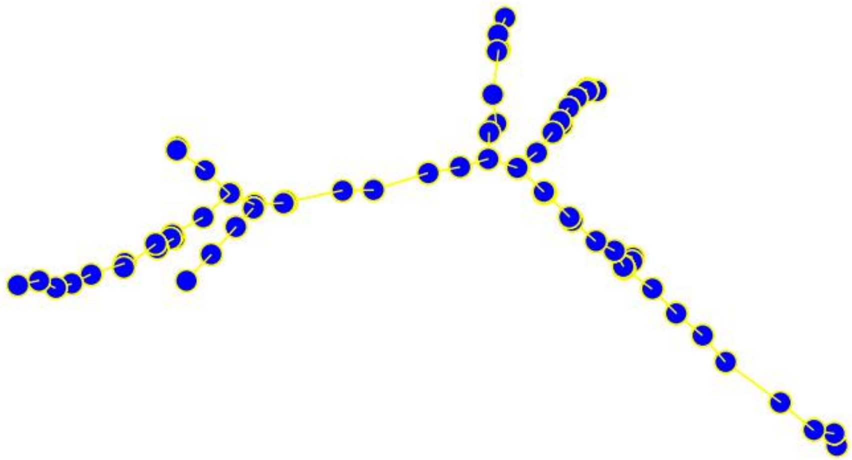

Figure 14: MST of corridor points extracted using betweenness centrality.

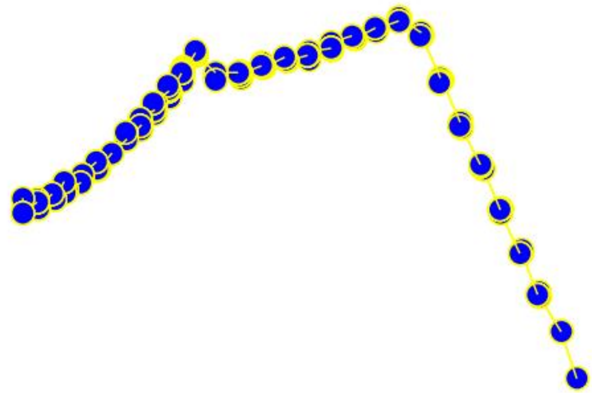

Figure 15: MST of corridor points.

Experiments(cont'd)

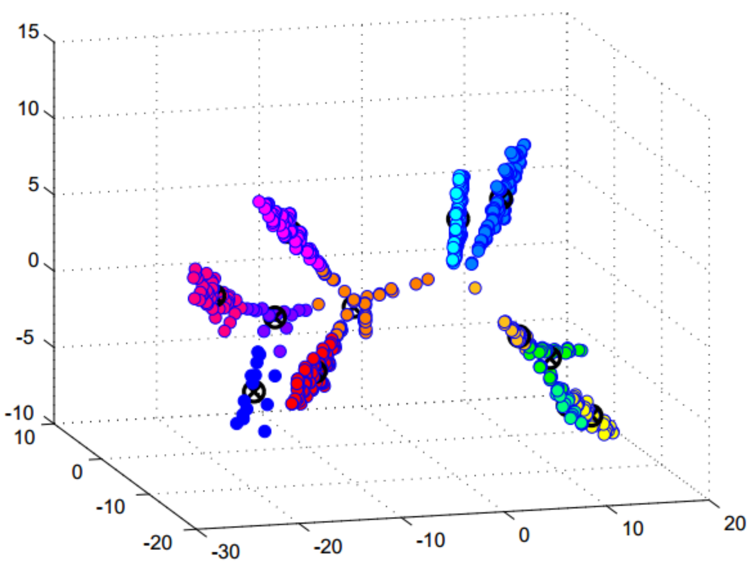

Figure 16: Clustering results by K-Means.

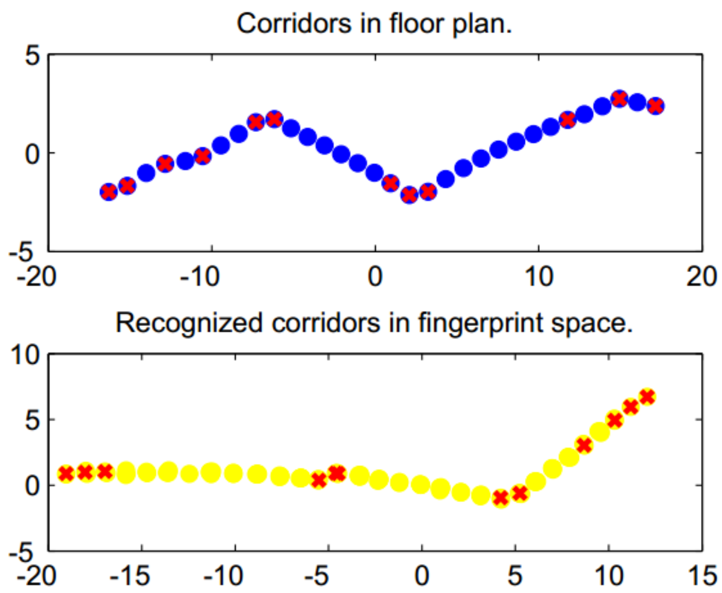

Figure 17: Floor plan corridor vs. Recognized corridor. Points marked with 'X' are reference points.

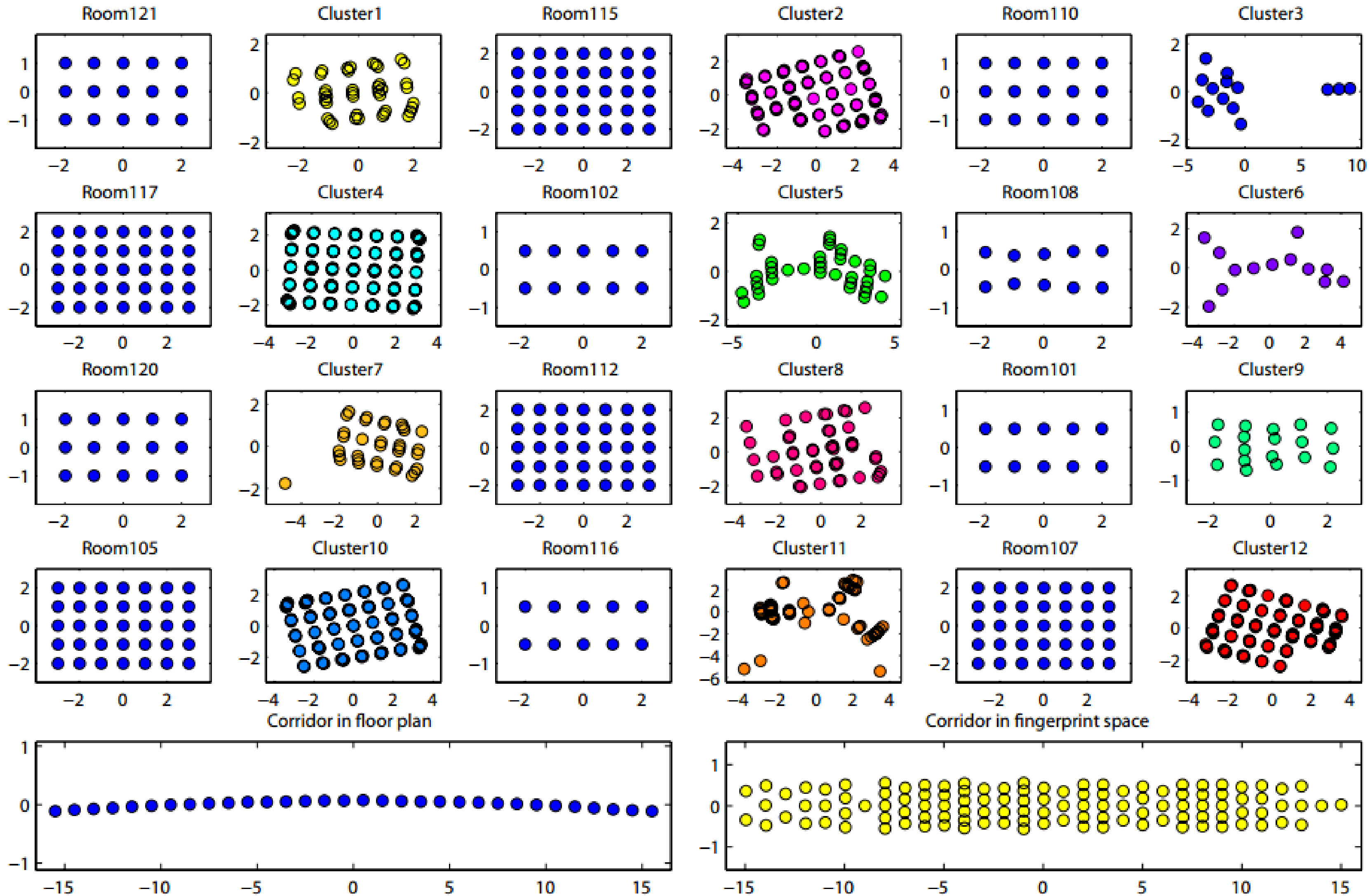

Figure 18: Fingerprints clusters vs. floor plan rooms.

Experiments(cont'd)

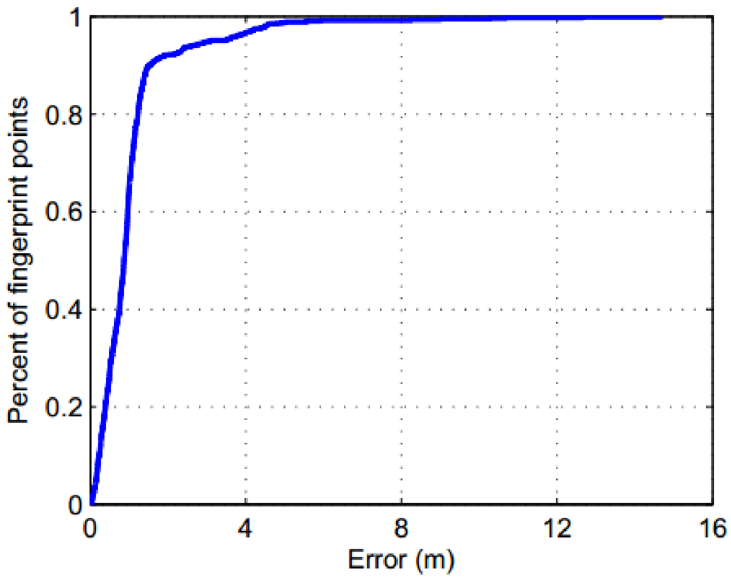

Figure 19: CDF of mapping error.

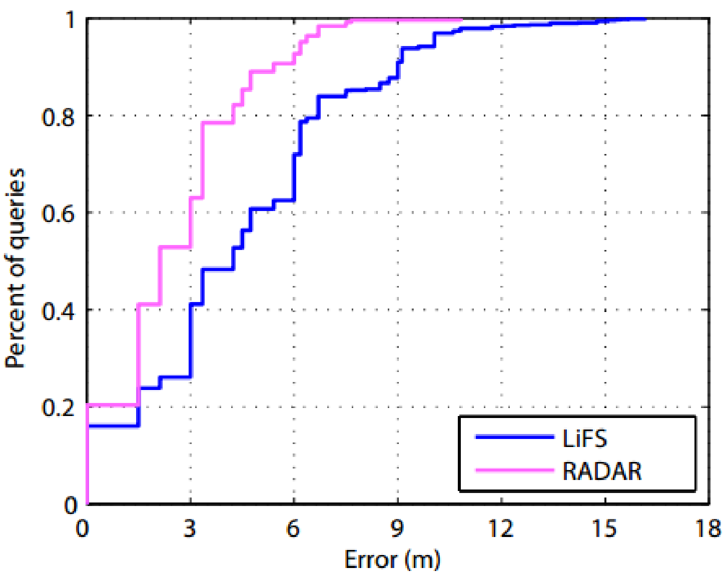

Figure 20: CDF of localization error.

Conclusion

1, Create fingerprint space in which fingerprints are distributed according to their mutual distances in real world.

2, Design and implement LiFS, an indoor localization system based on off-the-shelf WiFi infrastructure and mobile phones.

3, The preliminary experiment results show that LiFS achieves low human cost, rapid system deployment, and competitive location accuracy.

4, Sets up a novel perspective to cut off human intervention of indoor localization approaches.

My points of view

- Incomplete data. This paper simply assume the RSS equals to 0 when the corresponding AP cannot be detected.
- Depends on parsing the fingerprints from corridors too much, for the buildings with out corridors such as the mall or airport, it won't work well
- The cluster method simply choose K-means, maybe some other density-based clustering method
- Need prior information, such as the floor plan and numbers of room

...

Thanks

Q&A