Accurate and Low-cost Indoor Location Estimation Using Kernels

Jeffrey J. Pan, James T. Kwok, Qiang Yang, Yiqiang Chen

Department of Computer Science
Hong Kong University of Science and Technology

Present in the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI-05), Edinburgh, Scotland, July 2005.

Application Background

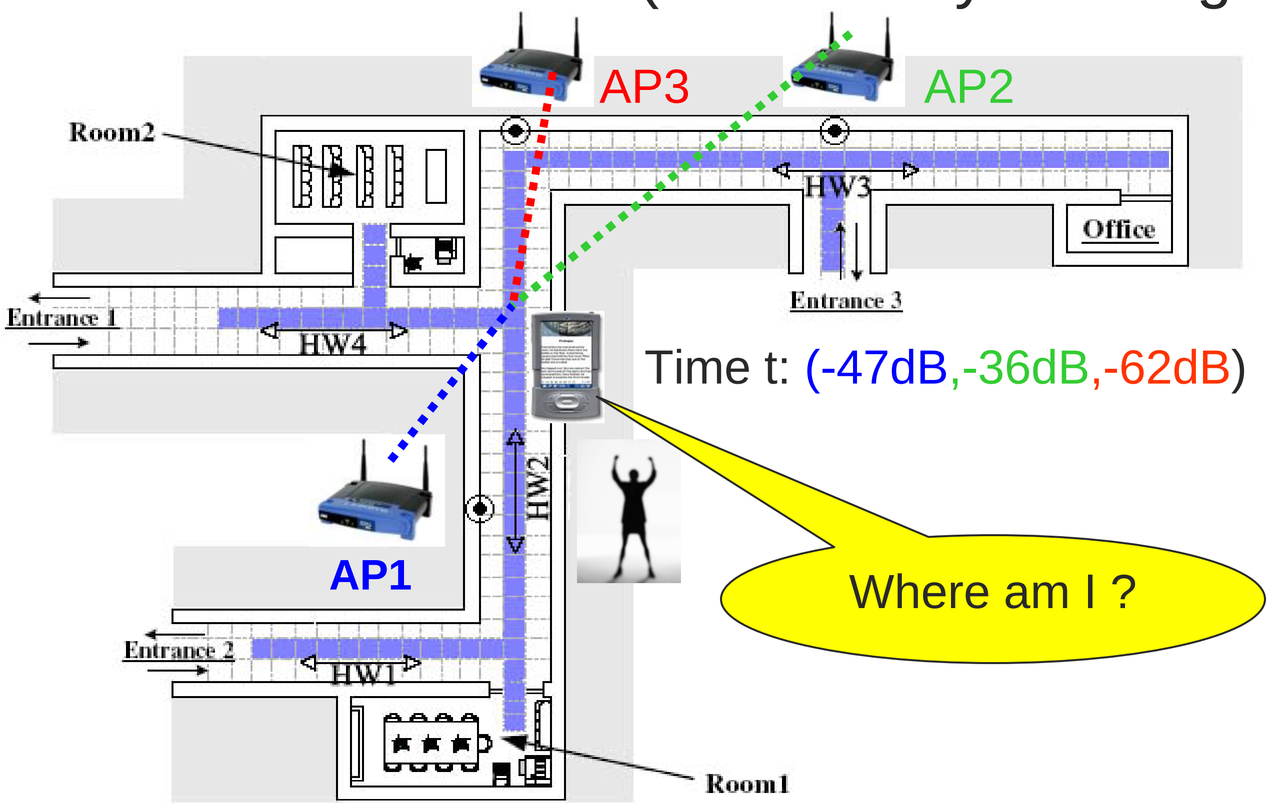
HKUST

- Positioning
  - Outdoor : Road Guiding (GPS)
  - Indoor   : Large Building (WiFi)
- Location-based Service
  - Web Content Delivery
- Behavior Analysis
  - Daily Life (L. Liao et al. AAAI-04, IJCAI-05)
  - Health Care
  - Scientific Purpose

Problem Description

HKUST

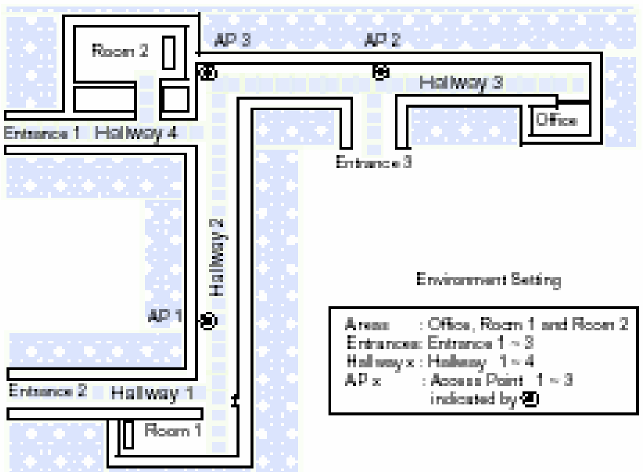
- A user with a mobile device walks in an **indoor** wireless environment (Covered by WiFi signals)



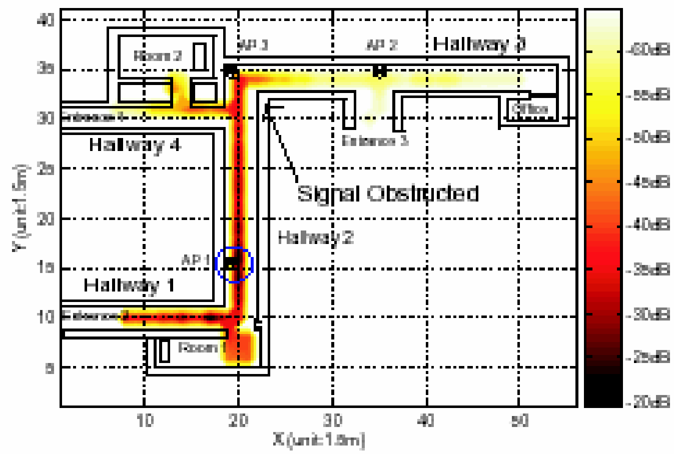
Off-the-Shelf Hardware, Only Signal Strength

Noisy Propagation Channel at 2.4G

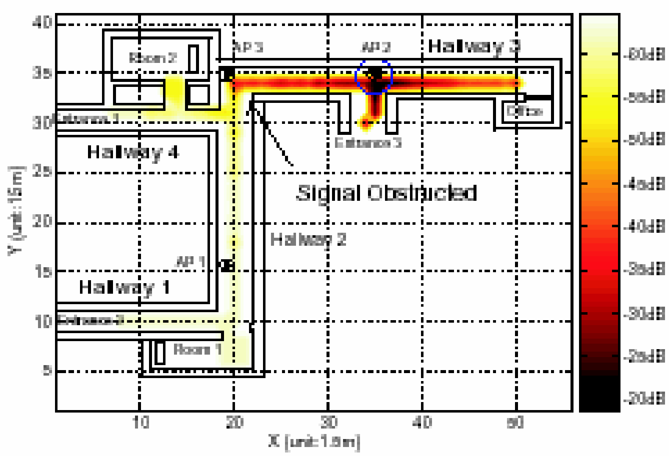
HKUST



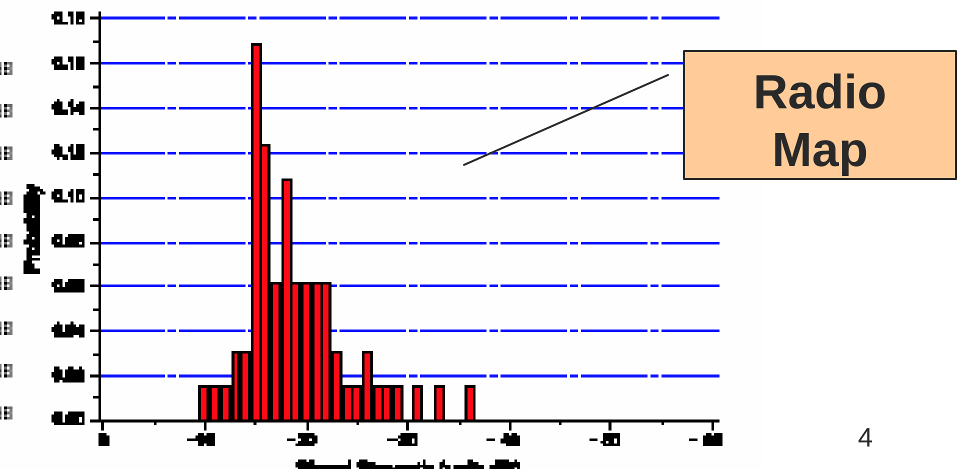
(a) Layout of the experimental test-bed.



(b) Signal distribution from AP 1.



(c) Signal distribution from AP 2.



(d) Signal distribution at a fixed location.

Learning-based Location Estimation

HKUST

- Two phases: offline Training and online Localization
- **Offline phase** – collect samples to build a mapping function F from signal space S to location space L

Loc.	Time	(AP1,AP2,AP3)
(1,0)	1s	(-60,-50,-40) dB
(2,0)	2s	(-62,-48,-35) dB
…..	….	(… , … , …)dB
(9,5)	9s	(-50,-35,-42) dB

Training… → Mapping function F

- **Online phase** – given a new signal s , estimate the most likely location l from F
  - s = (-60,-49,-36)dB , compute F(s) as the estimated location

Outline

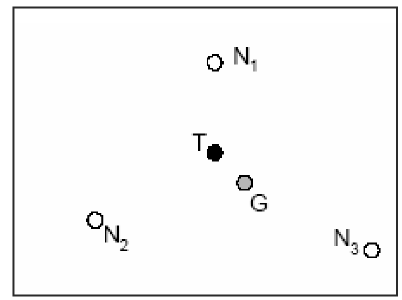
HKUST

- Introduction to Location Estimation
  - Application Background
  - Problem Description
  - Noisy Characteristics of Propagation Channel
  - Basic Framework for Location Estimation
- Related Work
  - Microsoft Research's RADAR (INFOCOM'2000)
  - University of Maryland's Horus (PerCom'2003)
- Motivation of Our Approach
- The LE-KCCA Algorithm
  - Kernel Canonical Correlation Analysis (KCCA)
  - Choices of Kernels
- Experimental Setup and Result
- Strength and Weakness
- Future Work

Related Works

HKUST

- Microsoft Research's RADAR [P. Bahl et al. INFOCOM2000]
  - K-Nearest-Neighbor Method
  - Offline - for each location, compute the **signal mean**
  - Online – estimate location with KNN and triangulation



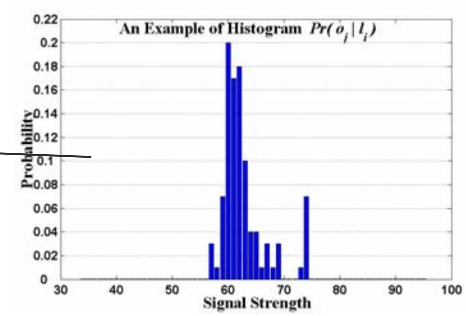
- Strength
  - Small number of samples could estimate the **signal mean** well
- Weakness
  - Accuracy is relatively low
  - Reason – The K nearest neighbors retrieved in the signal space may not necessarily the K nearest neighbors in the location space

Related Works (Cont')

HKUST

- **University of Maryland's Horus [M. Youssef et al. ,2003]**
  - Maximum Likelihood Estimation (MLE)
  - Offline - for each location, build the **Radio Map** of each AP
  - Online - apply Bayes' rule for estimation
- **Strength**
  - Accuracy is high
- **Weakness**
  - Need relatively large number of samples
  - Reason – More samples are needed for establishing an accurate **Radio Map** rather than a **signal mean**

Radio Map



Motivation of Our Approach

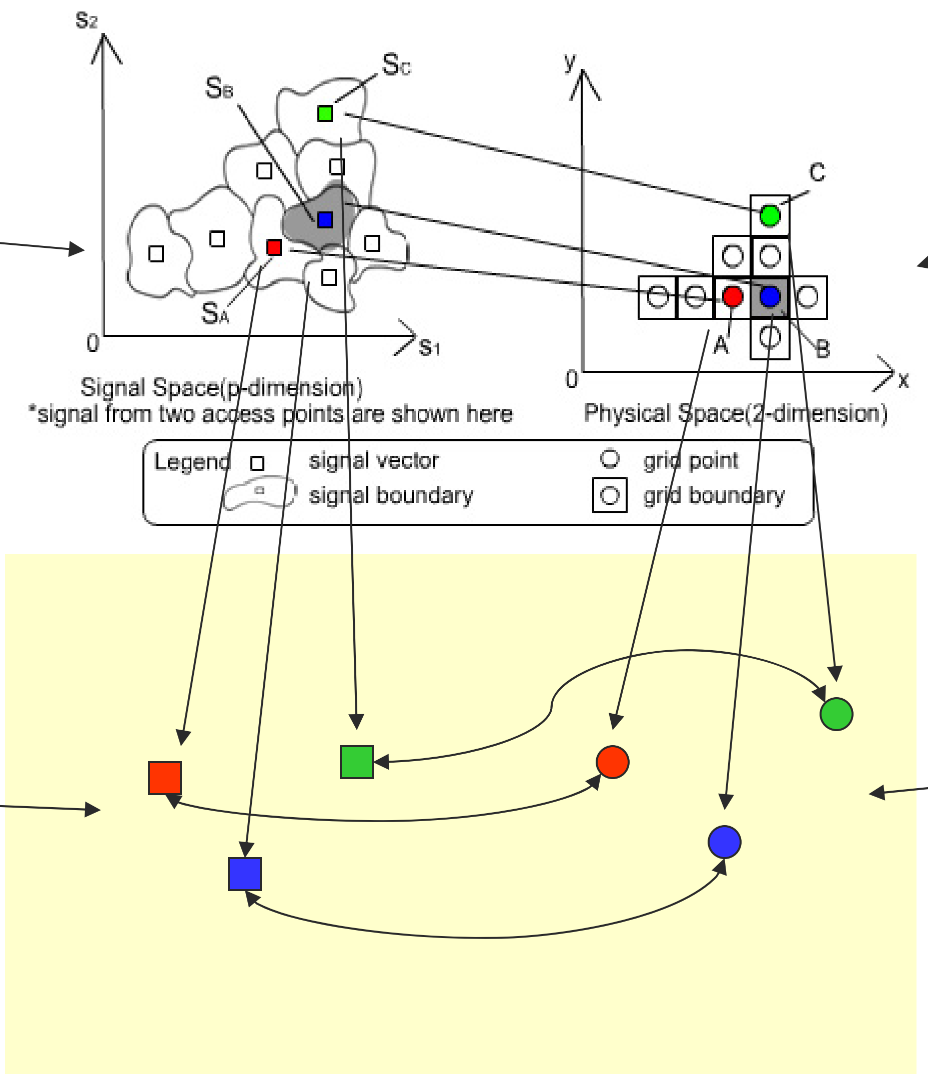
HKUST

- Observation (Motivated by RADAR)
  - Similar signals may not be nearby locations
  - Dissimilar signals may not be far away
- Idea
  - Maximize the similarity correlation between signal and location spaces under feature transformation
- Goal
  - Accuracy as high as possible (Horus)
  - Calibration Effort as low as possible (RADAR)

Motivation of Our Approach (Cont')

Original Signal Space

Original Location Space



Feature Signal Space

Feature Location Space

(Kernel) CCA

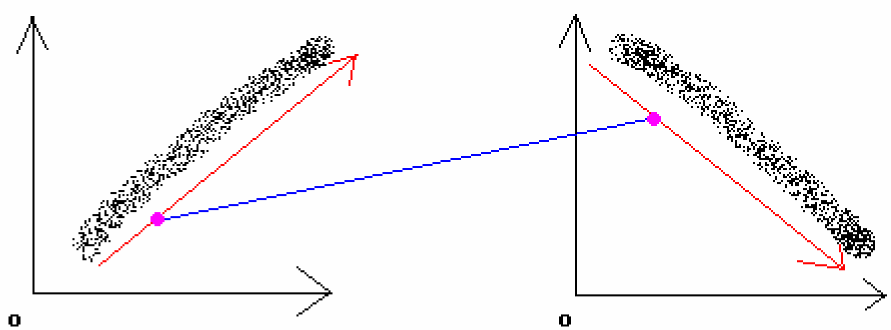
HKUST

- **C**anonical **C**orrelation **A**nalysis (CCA)
  - [H. Hotelling, 1936]
  - Two data set X and Y
  - Two linear Canonical Vectors **Wx Wy**
  - Maximize the correlation of projections

$$S_{x,\mathbf{w}_x} = (\langle \mathbf{w}_x, \mathbf{x}_1 \rangle, \ldots, \langle \mathbf{w}_x, \mathbf{x}_n \rangle)$$

$$S_{y,\mathbf{w}_y} = (\langle \mathbf{w}_y, \mathbf{y}_1 \rangle, \ldots, \langle \mathbf{w}_y, \mathbf{y}_n \rangle)$$

$$\begin{aligned} \rho &= \max_{\mathbf{w}_x, \mathbf{w}_y} corr(S_x \mathbf{w}_x, S_y \mathbf{w}_y) \\ &= \max_{\mathbf{w}_x, \mathbf{w}_y} \frac{\langle S_x \mathbf{w}_x, S_y \mathbf{w}_y \rangle}{\|S_x \mathbf{w}_x\| \|S_y \mathbf{w}_y\|} \end{aligned}$$



- Kernel **CCA**
  - [D.R Hardoon, S. Szedmak, and J. Shawe-Taylor, 2004]
  - Two non-linear Canonical Vectors

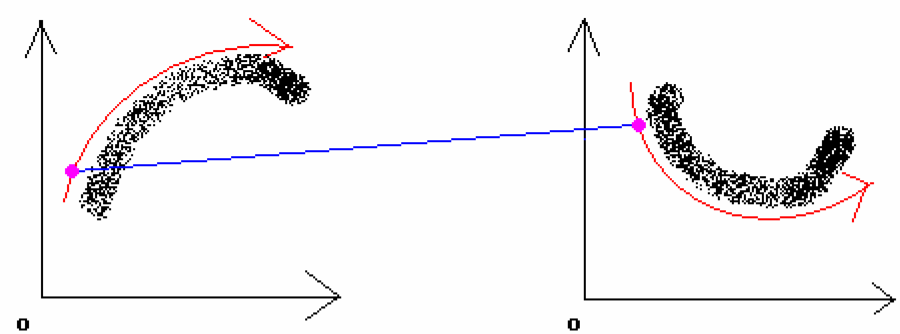
$$\mathbf{w}_x = X\alpha \qquad \mathbf{w}_y = Y\beta$$

  - K is the kernel

$$\phi : \mathbf{x} \mapsto \phi(\mathbf{x})$$

$$\kappa(x, z) = \langle \phi(x), \phi(z) \rangle$$

$$\rho = \max_{\alpha, \beta} \frac{\alpha' K_x K_y \beta}{\sqrt{\alpha' K_x^2 \alpha \cdot \beta' K_y^2 \beta}}.$$



LE-KCCA

HKUST

- **Offline phase**
  - Signal strengths are collected at various grid locations.
  - KCCA is used to learn the mapping between signal and location spaces.
    - λ_i's and α_i's are obtained from the generalized eigen-problem
    - κ is a regularization term

$$(\mathbf{K_x} + \kappa\mathbf{I})^{-1}\mathbf{K_y}(\mathbf{K_y} + \kappa\mathbf{I})^{-1}\mathbf{K_x}\alpha = \lambda^2\alpha,$$

  - For each training pair ($\mathbf{s}_i$, l_i), its projections

$$P(\mathbf{s}_i) = [P_1(\mathbf{s}_i), P_2(\mathbf{s}_i), \ldots, P_T(\mathbf{s}_i)]'$$

on the T canonical vectors are obtained from

$$P_{\mathbf{x}}(\tilde{\mathbf{x}}) = \phi_x(\tilde{\mathbf{x}})'\mathbf{w}_{\phi_x(\mathbf{x})} = \mathbf{k}'_{\tilde{\mathbf{x}}}\alpha,$$

LE-KCCA (Cont')

HKUST

- **Online phase**
  - Assume the location of a new signal strength vector is **s**
  - Again, use

    $$P_{\mathbf{x}}(\tilde{\mathbf{x}}) = \phi_x(\tilde{\mathbf{x}})' \mathbf{w}_{\phi_x(\mathbf{x})} = \mathbf{k}'_{\tilde{\mathbf{x}}} \alpha,$$

    to project **s** onto the canonical vectors and obtain

    $$P(\tilde{\mathbf{s}}) = [P_1(\tilde{\mathbf{s}}), P_2(\tilde{\mathbf{s}}), \ldots, P_T(\tilde{\mathbf{s}})]'.$$

  - Find the K *Nearest N*eighbors of *P*(**s**) in the projections $P(\mathbf{s}_i)$ of training set with the weighted Euclidean distance :

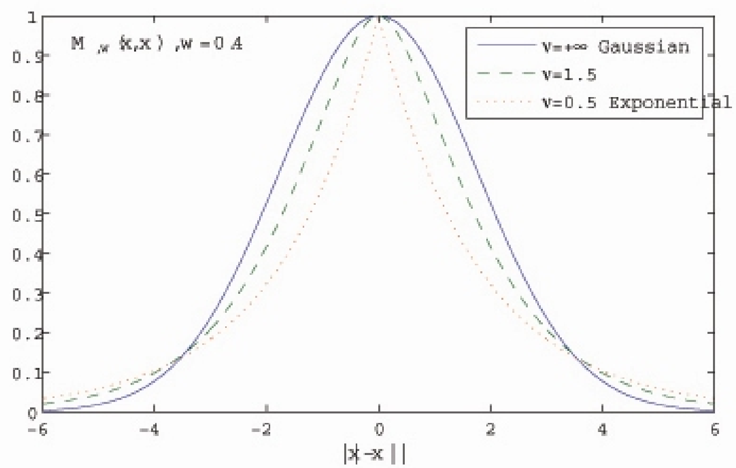
    $$d_i = \sum_{j=1}^{T} \lambda_j (P_j(\tilde{\mathbf{s}}) - P_j(\mathbf{s}_i))^2$$

  - Interpolate these neighbors' locations to predict the location of **s**
- Essentially, we are performing Weighted KNN in the feature space with which weights are obtained from the feedback of location information.

Choices of Kernels

- Kernel for Signal Space
  - Gaussian Kernel to smooth the noisy characteristics
  - Widely used : [Roos et al. 2002, Battiti et al. 2002]
- Kernel for Location Space
  - Matern Kernel to sense the change in location
  - Used in : GPPS [Schwaighofer et al., 2003]

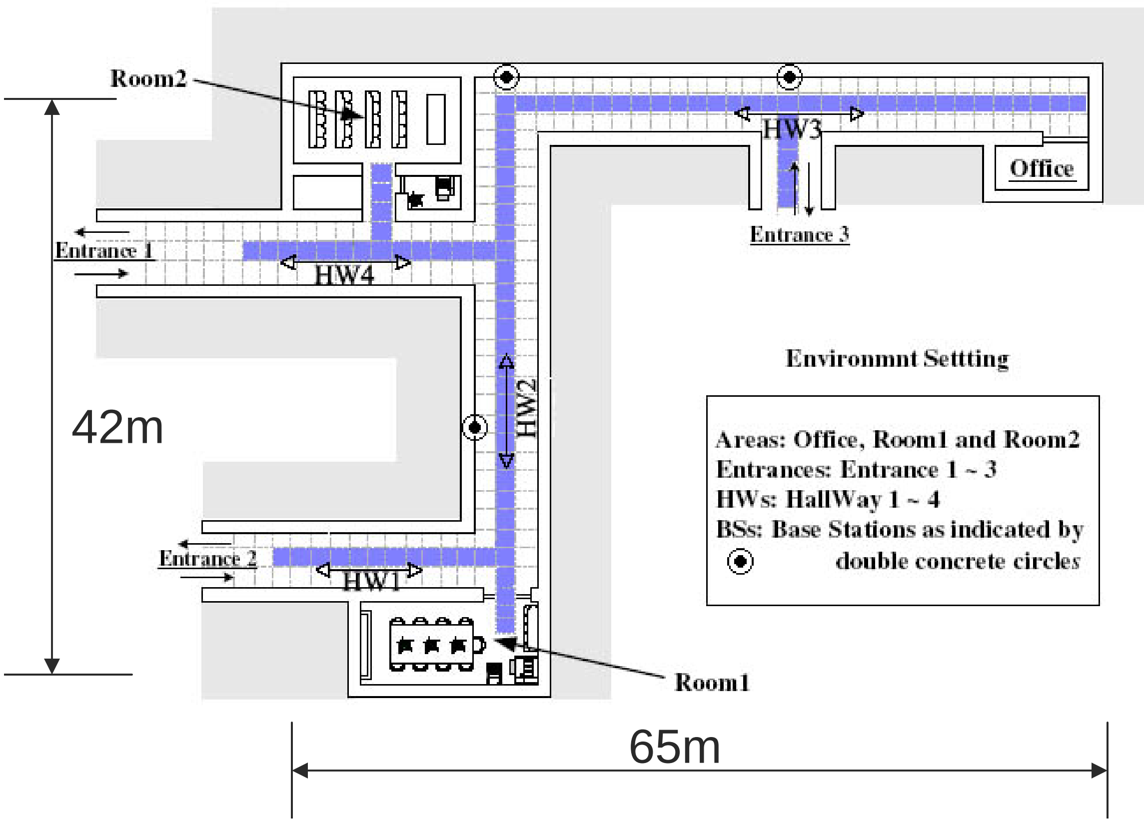
$$K_I(\|\mathbf{x}-\mathbf{z}\|)/K_I(0) = \frac{1}{2^{\nu-1}\Gamma(\nu)}\left(\frac{2\sqrt{\nu}\|\mathbf{x}-\mathbf{z}\|}{\theta}\right)^{\nu} H_\nu\left(\frac{2\sqrt{\nu}\|\mathbf{x}-\mathbf{z}\|}{\theta}\right)$$



Experimental Setup

HKUST

- Test-bed : Department of Computer Science, Hong Kong University of Science and Technology

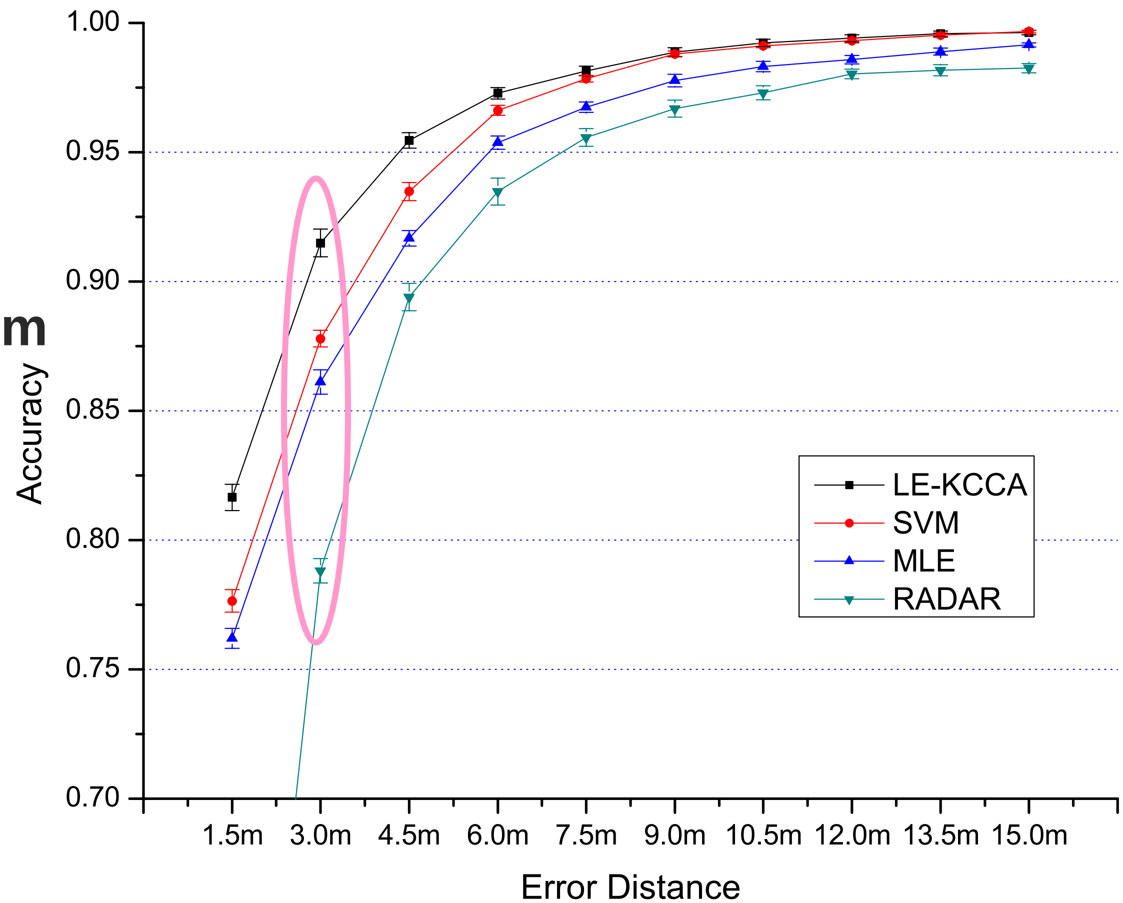


- **99 locations (1.5 × 1.5 meter)**
- **100 samples per location**
- **65% for training, 35% testing**
- **Repeat each experiment 10 times**

Experimental Result - 1

HKUST

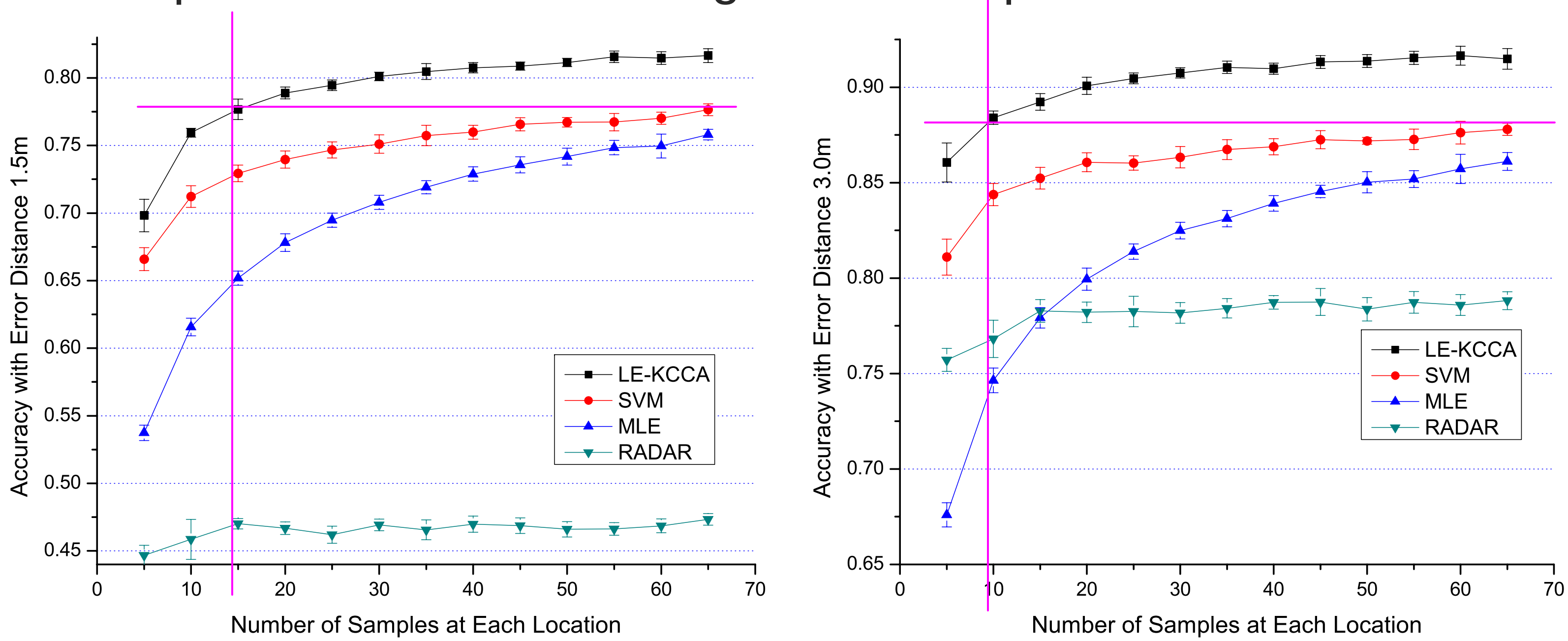
- Accuracy
  - **Data Set**
    - 65% training
    - 35% testing
  - **Error Distance is 3.0m**
    - **LE-KCCA 91.6%**
    - **SVM 87.8%**
    - **MLE 86.1%**
    - **RADAR 78.8%**



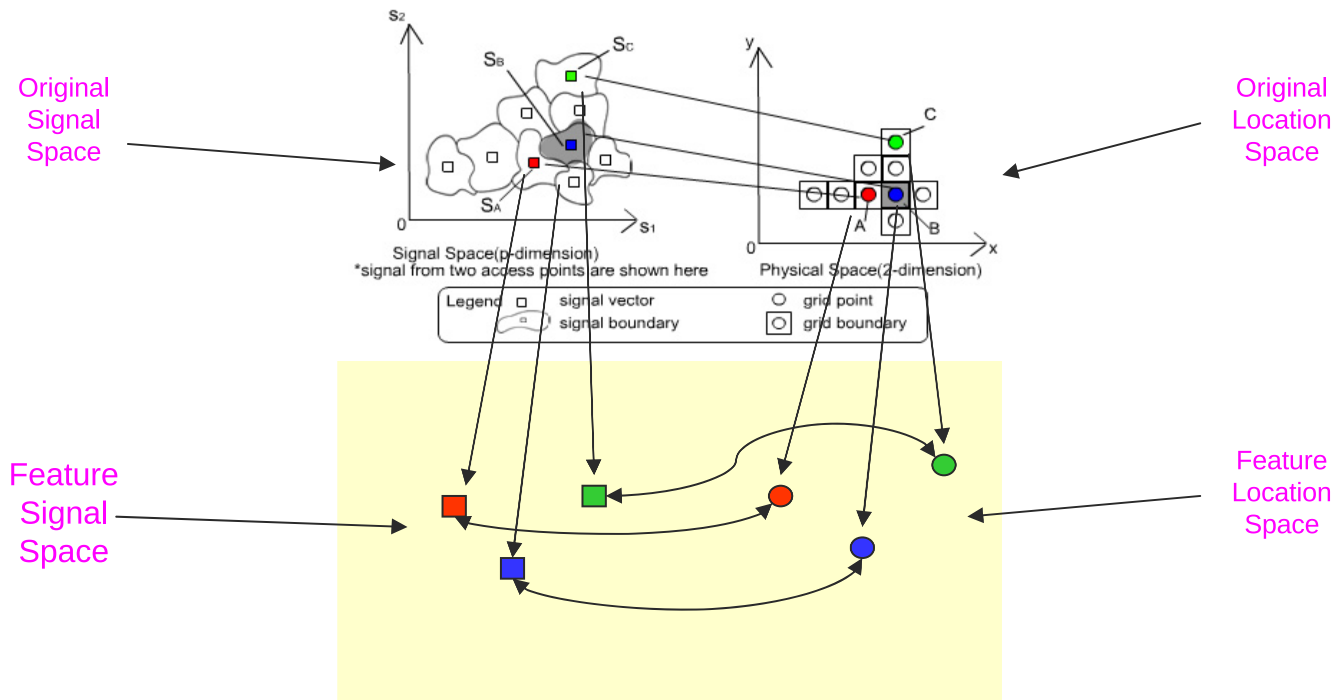
Experimental Result - 2

HKUST

- Reduce Calibration Effort
  - Incrementally Use a small subset of the the 65% training data
  - Outperform the others using 10-15 samples from each location

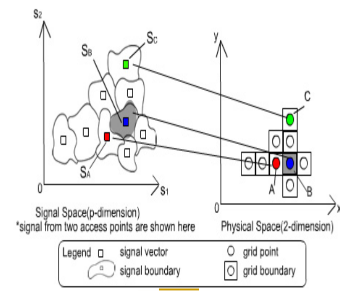


Recall our Motivation………



We could see on the next page……

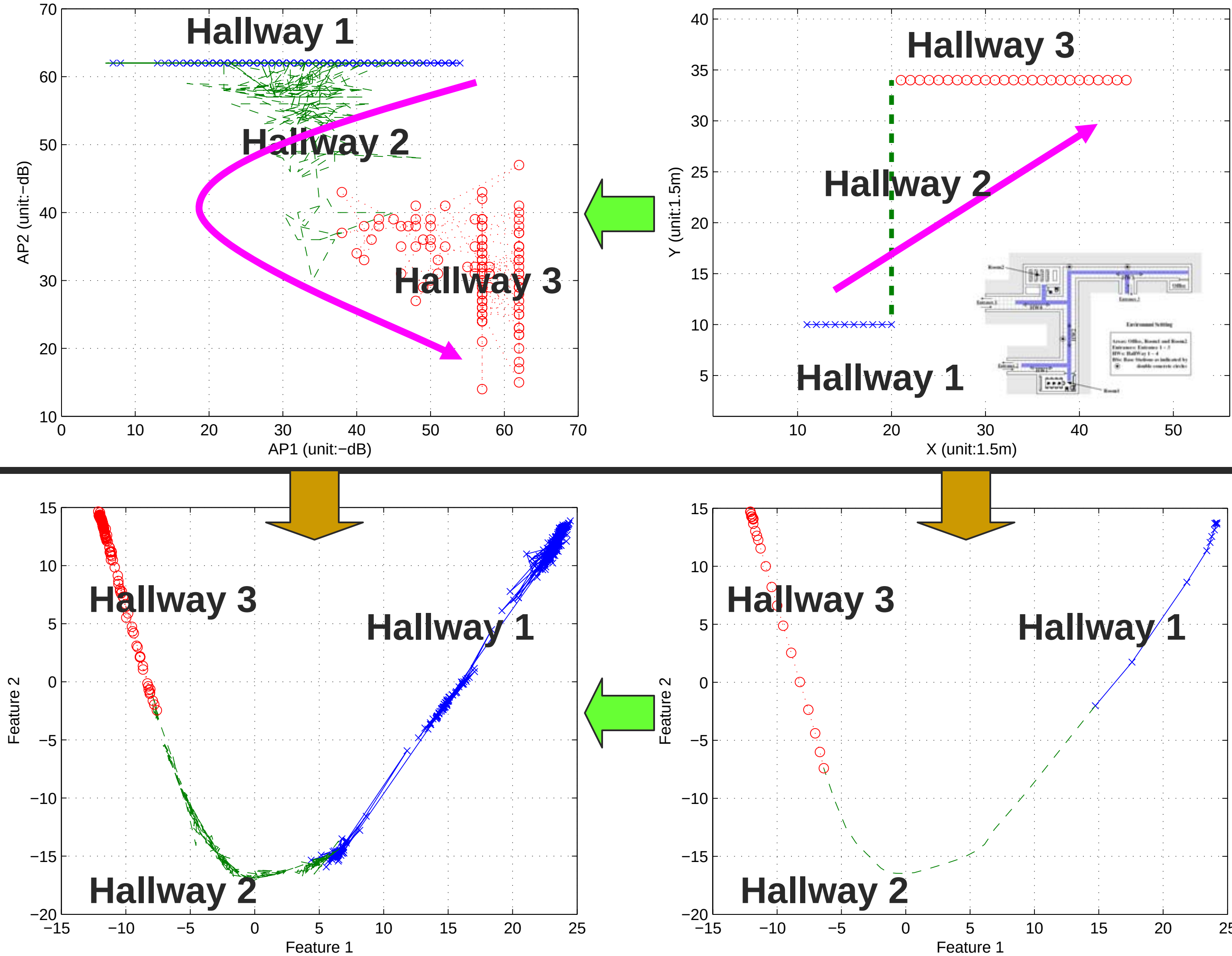
Visualization of Tracking in Both Original and Feature Spaces



HKUST

Original Signal Space

Original Location Space



Feature Signal Space

Feature Location Space

Strength and Weakness

HKUST

- Strength
  - Higher Accuracy
  - Reduced Calibration Effort (Low-cost)
- Weakness
  - Generally 50-100 times slower than RADAR

Future Work

HKUST

- **Consider Environment Dynamics to Reduce Uncertainty**
  - J. Yin et al. Adaptive temporal radio maps for indoor location estimation. PerCom'05
- **Consider User Dynamics to Reduce Uncertainty**
  - M. Berna et al. A Learning Algorithm for Localizing People Based On Wireless Signal Strength That Uses Labeled and Unlabeled Data. IJCAI'03
  - A. Ladd et al. Robotics-based location sensing using wireless ethernet, MobiCom'02
- **Speed up for Large-Scale Localization**
  - J. Letchner et al. Large Scale Localization from Wireless Signal Strength. AAAI'05
  - A. Haeberlen et al. Practical Robust Localization over Large-Scale 802.11 Wireless Networks. MobiCom'04

Acknowledge

HKUST

- Hong Kong RGC
  - HKUST6187/04E
- Thanks Jie Yin and Xiaoyong Chai
  - Data collection
  - Helpful discussion

HKUST

Thank You

Question ?