Lecture on *Positioning*

Prof. Maria Papadopouli

University of Crete
ICS-FORTH
http://www.ics.forth.gr/mobile

Roadmap

- Location Sensing Overview
  - Location sensing techniques
  - Location sensing properties
  - Survey of location systems

Importance of Location Sensing

- Mapping systems
- Locating people & objects
- Emergency situations/mobile devices
- Wireless routing
- Supporting ambient intelligence spaces
  - location-based applications/services
  - assistive technology applications

Location System Properties

- **Location description**: physical vs. symbolic
- **Coordination systems**: Absolute vs. relative location
- **Methodology for estimating distances, orientation, position**
- **Computations**: Localized vs. remote
- **Requirements: Accuracy, Precision, Privacy, Identification**
- **Scale**
- **Cost**
- Limitations & dependencies
  - infrastructure vs. ad hoc
  - hardware availability
  - multiple modalities (e.g., RF, ultrasonic, vision, touch sensors)

Accuracy vs. Precision

- A result is considered ***accurate*** if it is **consistent with the *true* or *accepted* value** for that result

- Precision refers to the ***repeatability*** of measurement
  - Does ***not*** require us to know the correct or true value
  - Indicates how sharply a result has been defined

Location Sensing Techniques

- Distance- vs. signature-based approaches
  - **Distance-based**

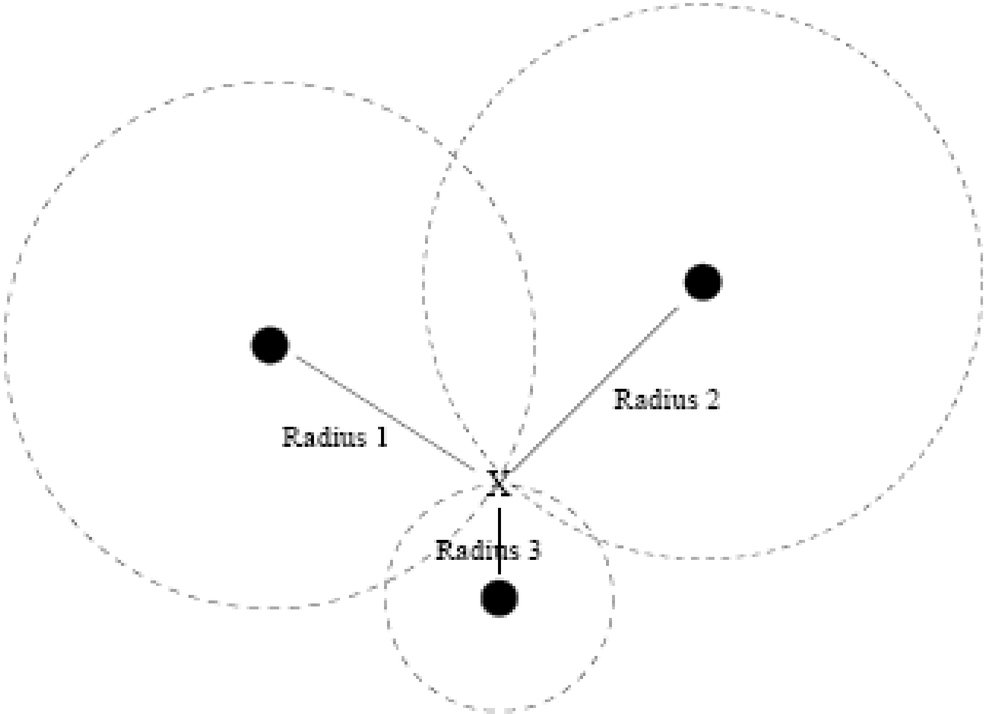
1. use radio propagation models to estimate distance from landmarks
2. apply lateration or angulation techniques

  - **Signature-based**

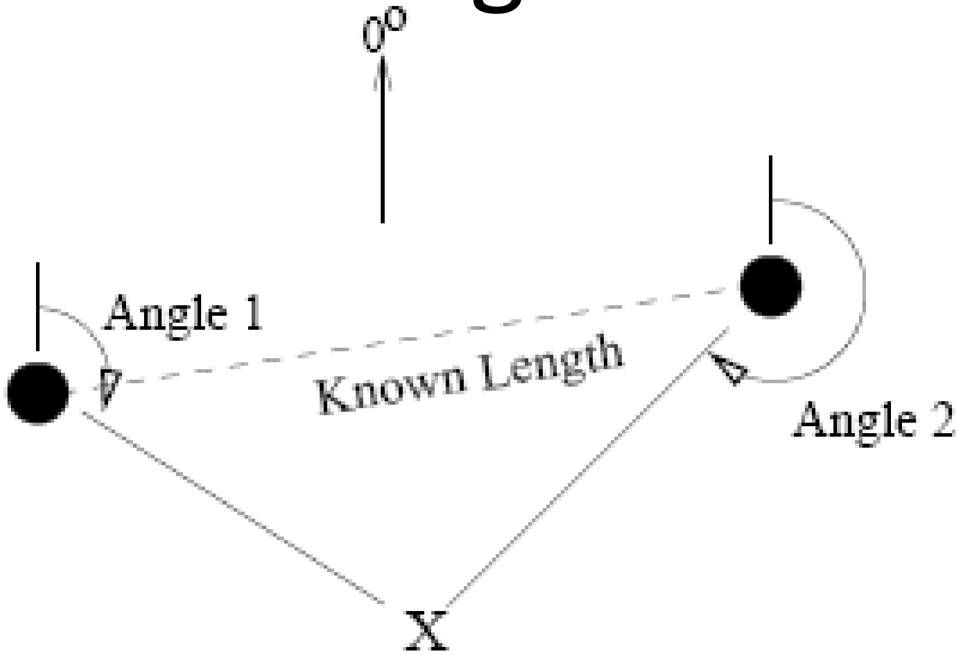
1. build maps of physical space enriched with measurements
2. apply pattern matching algorithms

- Proximity

Lateration



Angulation



- The angle between two nodes can be determined by estimating the AOA parameter of a signal traveling between two nodes

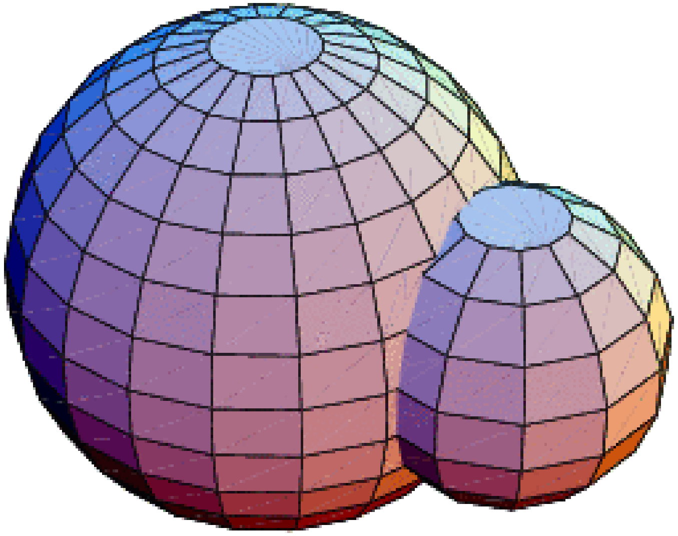
☞ Phased antenna array can be employed

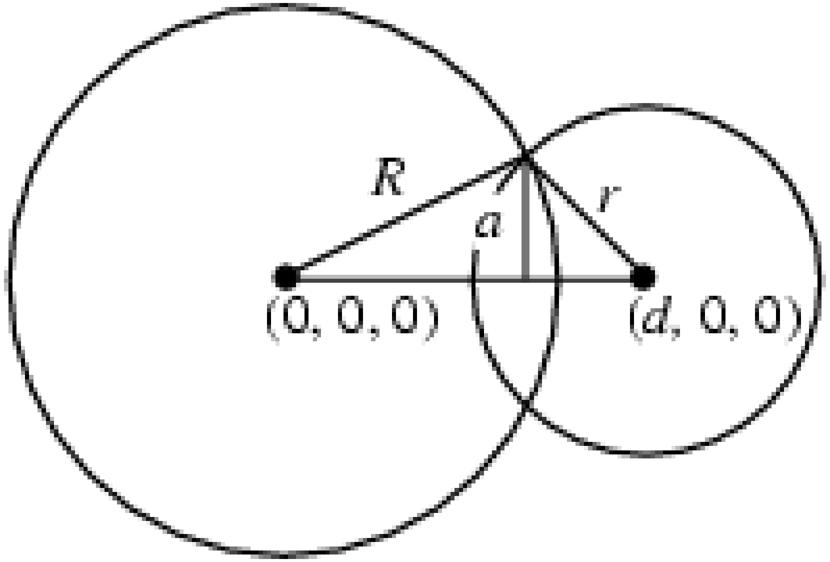
Phased Antenna Array

- Multiple antennas with ***known separation***
- Each measures ***time of arrival of signal***
- Given the **difference in time of arrival** **&** **geometry of the receiving array**, the angle from which the emission was originated can be computed
- If there are enough elements in the array with large separation, the angulation can be performed

Triangulation - Lateration

- Uses **geometric properties** of triangles to compute object locations
- Lateration: Measures distance from reference points
  - 2-D requires 3 non-colinear points
  - 3-D requires 4 non-coplanar points





Triangulation - Lateration

Types of Measurements

- **Direct touch, pressure**
- **Time-of-flight**

  (e.g., sound waves travel 344m/s in 21°C)
- **Signal attenuation**
  - calculate based on send and receive strength
  - attenuation varies based on environment

Time-of-Arrival Issues

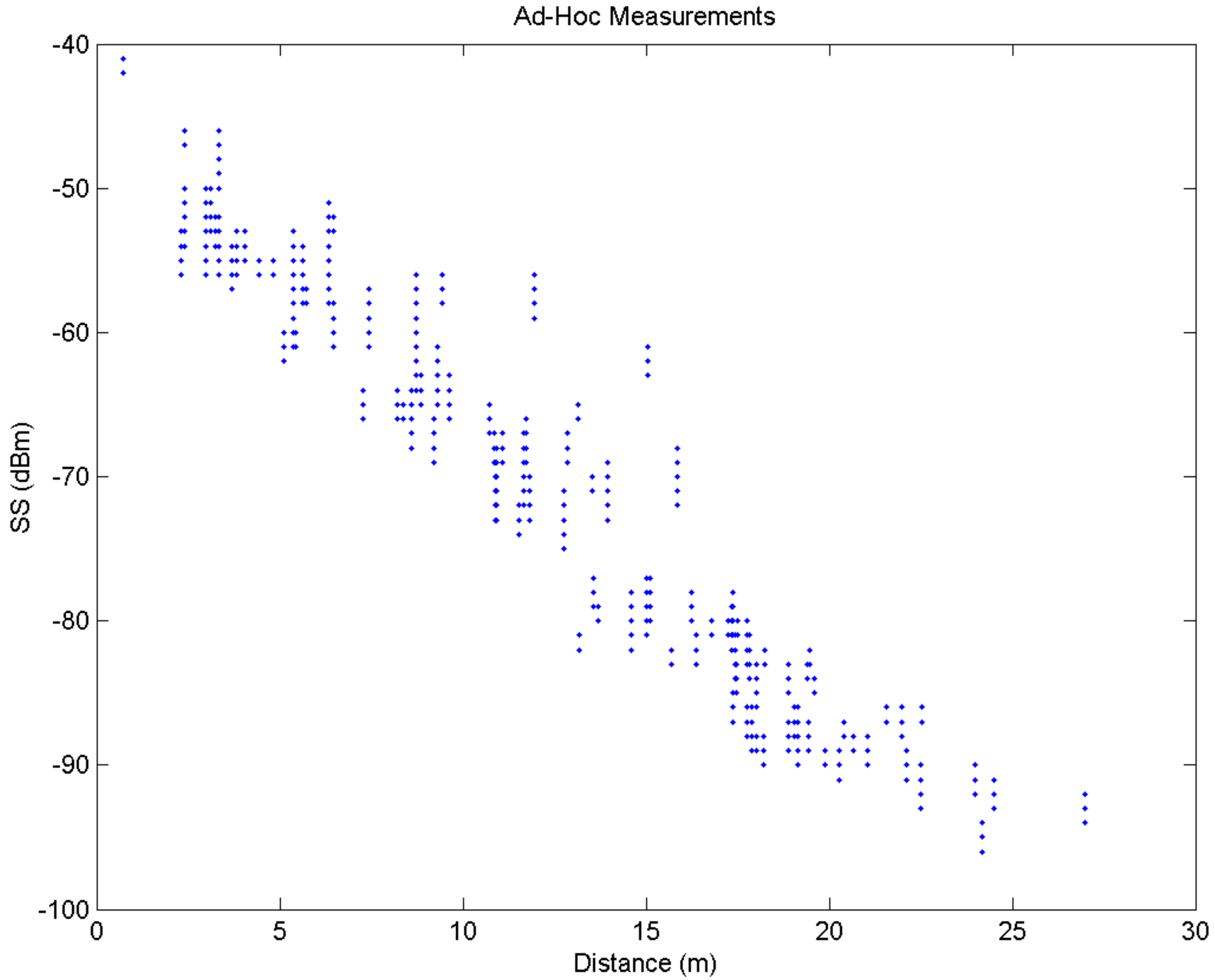
- Requires **known velocity**
- May require **high time resolution** (e.g., for light or radio)

A light pulse (with 299,792,458m/s) will travel the 5m in 16.7ns

Time of flight of light or radio requires clocks with much higher resolution (by 6 orders of magnitude) than those used for timing ultrasound

- ***Clock synchronization***
  - Possible solution ?

Some Real-life Measurements



Signal Power Decay with Distance

- A signal traveling from one node to another experiences fast (multipath) fading, shadowing & path loss
- ***Ideally***, ***averaging RSS*** over sufficiently **long time interval**

excludes the effects of multipath fading & shadowing ⇨

general ***path-loss model***:

$$\bar{P}(d) = P_0 - 10n \log_{10} (d/d_o)$$

n: path loss exponent

$\bar{P}(d)$: the average received power in dB at distance d

P_0 is the received power in dB at a short distance d_0

GPS

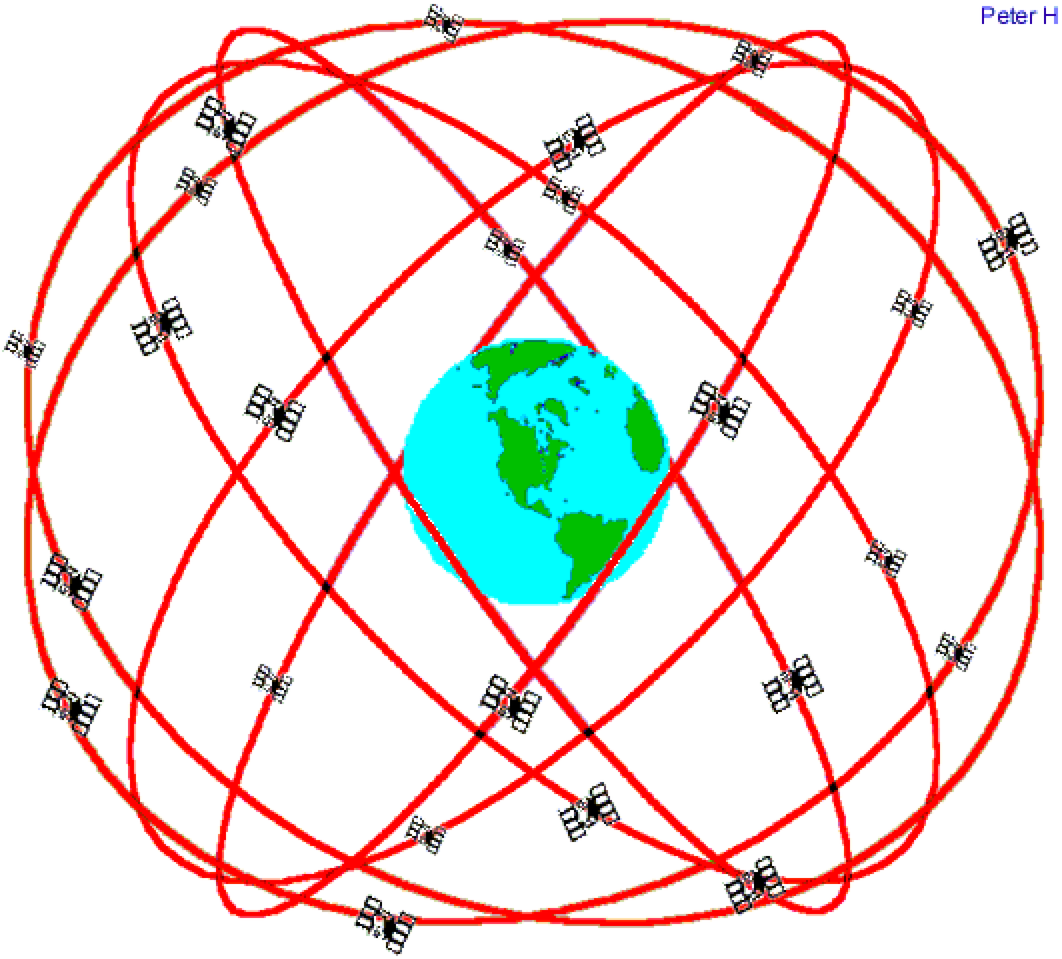
- 27 satellites
- The orbit altitude is such that the satellites ***repeat the same track*** and configuration over any point **approximately each 24 hours**
- Powered by **solar energy** (also have backup batteries on board)
- GPS is a ***line-of-sight*** technology

  the receiver needs a clear view of the satellites it is using

  to calculate its position
- Each satellite has ***4 rubidium atomic clocks***
  - locally averaged to maintain accuracy
  - updated daily by a **Master Control** facility

- Satellites are ***precisely synchronized with each other***
- Receiver is **not synchronized** with the satellite transmitter
- Satellites transmit their ***local time*** in the signal

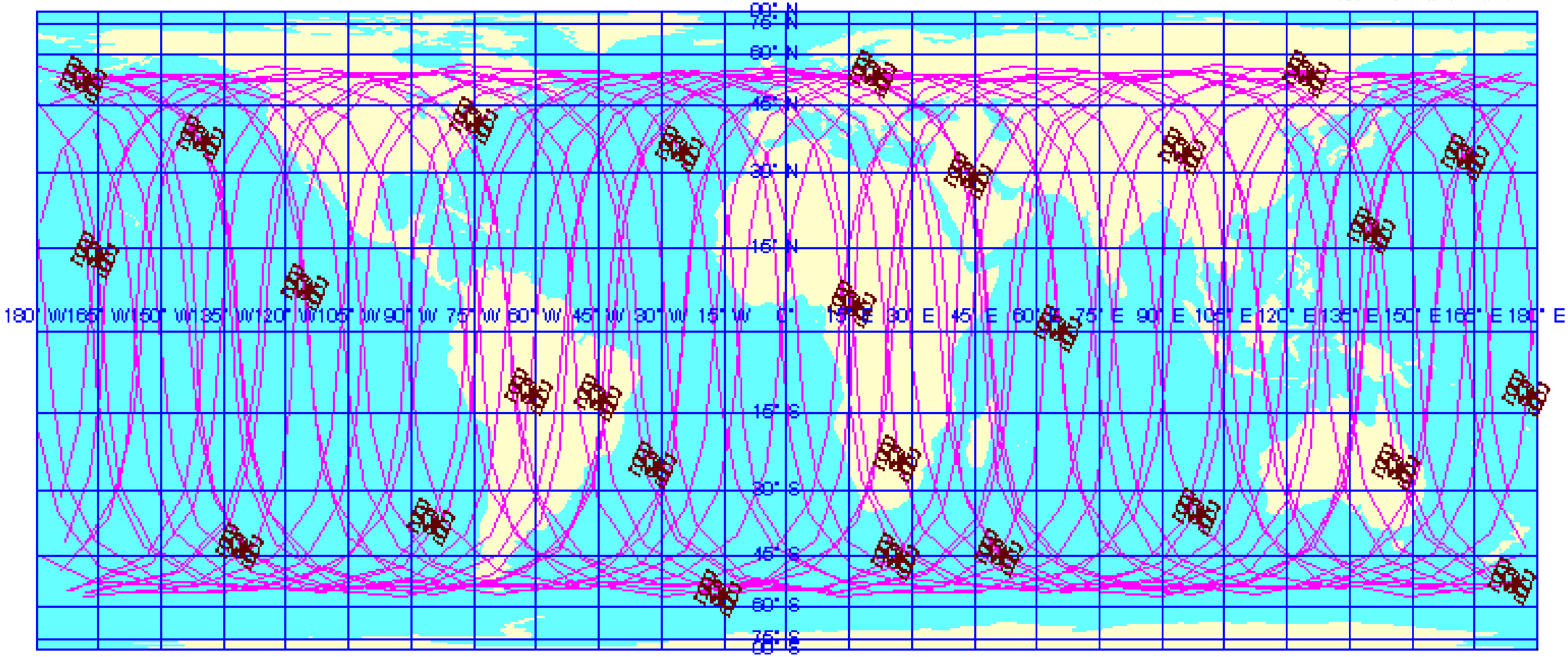
Satellites Orbits



Peter H. Dana 9/22/98

GPS Nominal Constellation
24 Satellites in 6 Orbital Planes
4 Satellites in each Plane
20,200 km Altitudes, 55 Degree Inclination

Satellites Positions and Orbits

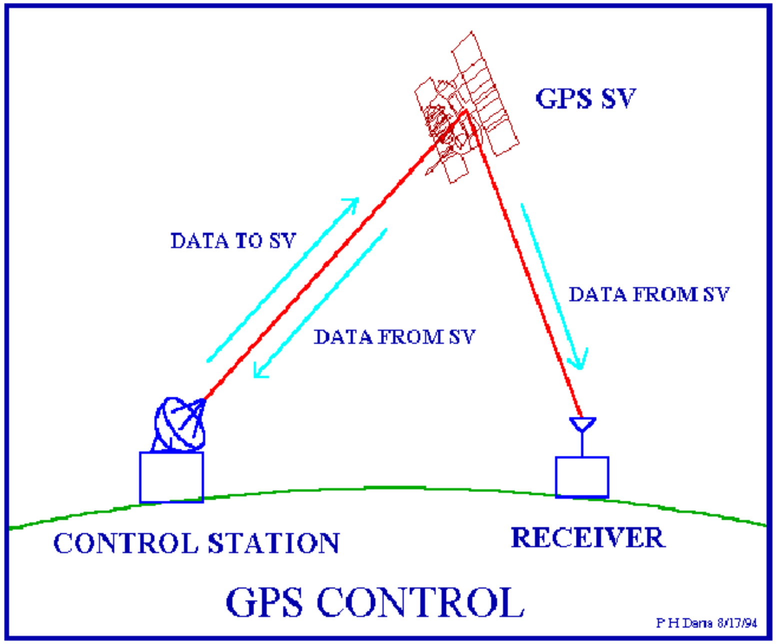


Peter H. Dana 10/6/98

Global Positioning System Satellites and Orbits
for 27 Operational Satellites on September 29, 1998
Satellite Positions at 00:00:00 9/29/98 with 24 hours (2 orbits) of Ground Tracks to 00:00:00 9/30/98

GPS (cont'd)

- Master Control facility monitors the satellites
- Computes
  - precise orbital data (i.e., ephemeris)
  - clock corrections for each satellite



GPS Receiver

- Composed of an antenna and preamplifier, radio signal microprocessor, control and display device, data recording unit, & power supply
- Decodes the timing signals from the 'visible' satellites (four or more)
- Calculates their distances, its own latitude, longitude, elevation, & time
- A continuous process: the position is updated on a sec-by-sec basis, output to the receiver display device and, if the receiver provides data capture capabilities, stored by the receiver-logging unit

GPS Satellite Signals

As light moves through a given *medium*, low-frequency signals get “refracted” or slowed more than high-frequency signals

Satellites transmit two microwave carrier signals:

- On *L1 frequency* (1575.42 MHz)

  it carries the navigation message (satellite orbits, clock corrections & other system parameters) & a unique identifier code
- On L2 frequency (1227.60 MHz)

  it uses to measure the ionospheric delay

☞ By **comparing the delays of the two different carrier frequencies** of the GPS signal L1 & L2, we can deduce what the medium is

GPS (cont'd)

- Receivers compute their difference in time-of-arrival
- Receivers estimate their position (longitude, latitude, elevation) using satellites
- 1-5m (95-99%)

GPS Error Sources

- Noise
- Satellites clock errors uncorrected by the controller (~1m)
- Ephemeris data errors (~1m)
- Troposphere delays due to weather changes

  e.g., temperature, pressure, humidity (~1m)

  Troposphere: lower part of the atmosphere, ground level to from 8-13km
- Ionosphere delays (~10m)

  Ionosphere: layer of the atmosphere that consists of ionized air (50-500km)
- Multipath (~0.5m)
  - caused by reflected signals from surfaces near the receiver that can either interfere with or be mistaken for the signal that follows the straight line path from the satellite
  - difficult to be detected and sometime hard to be avoided

GPS Error Sources (cont’d)

- Control segment mistakes due to computer or human error (1m-100s km)
- Receiver errors from software or hardware failures
- User mistakes

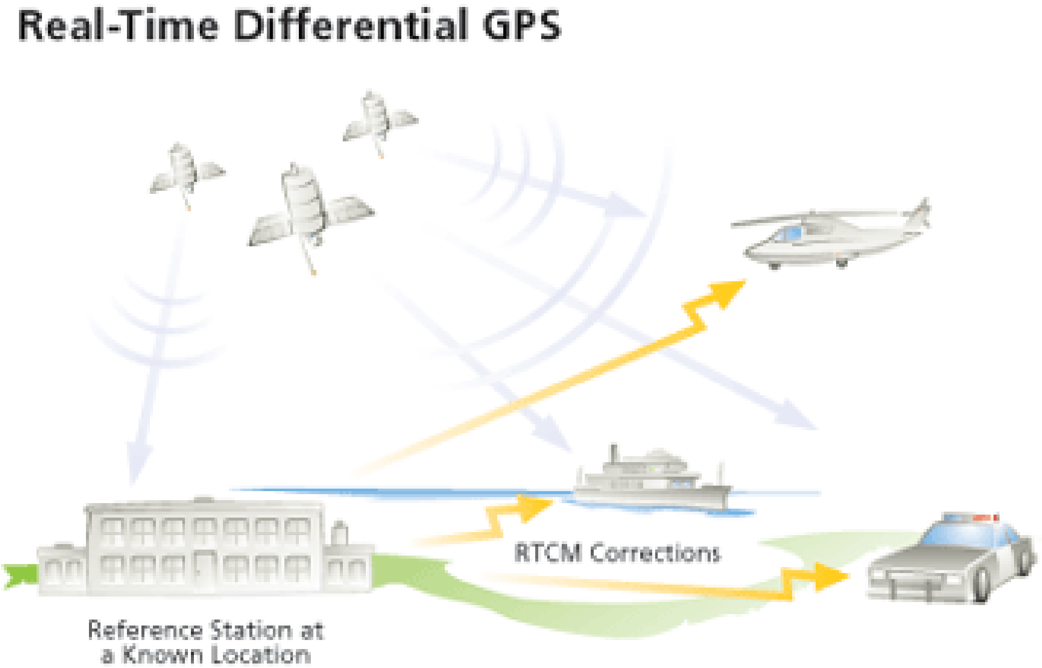
  e.g., incorrect geodetic datum selection (1-100m)

Differential GPS (DGPS)

- Assumes: any two receivers that are ***relatively close*** together will experience ***similar atmospheric errors***

- Requires ***reference station***: a GPS receiver been set up on a precisely known location

Reference stations calculate their position based on satellite signals and compares this location to the known location



Differential GPS (cont'd)

- The difference is applied to GPS data recorded by the roving receiver in real time in the field using radio signals or through postprocessing after data capture using special processing software

Real-time DGPS

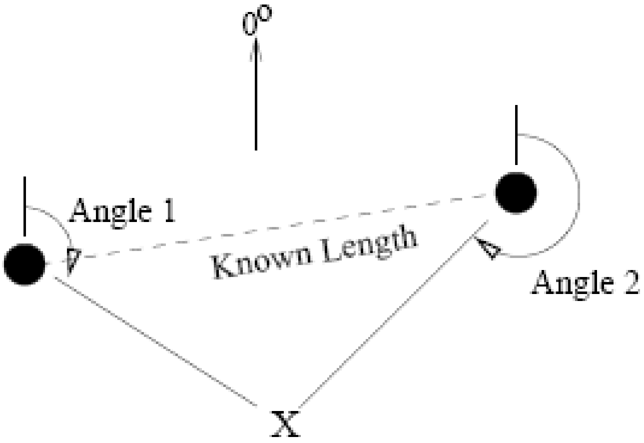
- Reference station calculates & **broadcasts corrections** for each satellite as it receives the data
- The correction is received by the roving receiver via a radio signal if the source is land based or via a satellite signal if it is satellite based and applied to the position it is calculating

Triangulation - Angulation

- 2D requires:

2 angles and 1 known distance

- Phased antenna arrays



Statistical-based Fingerprint

- ***Grid-based*** representation of physical space
- ***RSSI*** values collected from various APs @ ***cells*** of the space
- Statistical fingerprints based on:
  - Confidence intervals
  - Percentiles
  - Empirical distribution
  - Theoretical distribution (e.g., ***Multivariate Gaussian***)
- ***Training*** fingerprint
  - Formed for various cells during the training phase
- ***Runtime*** fingerprint
  - Formed @ the unknown position
- ***Estimated position***: cell whose ***training fingerprint*** has the ***minimum "distance"*** from the ***runtime one***

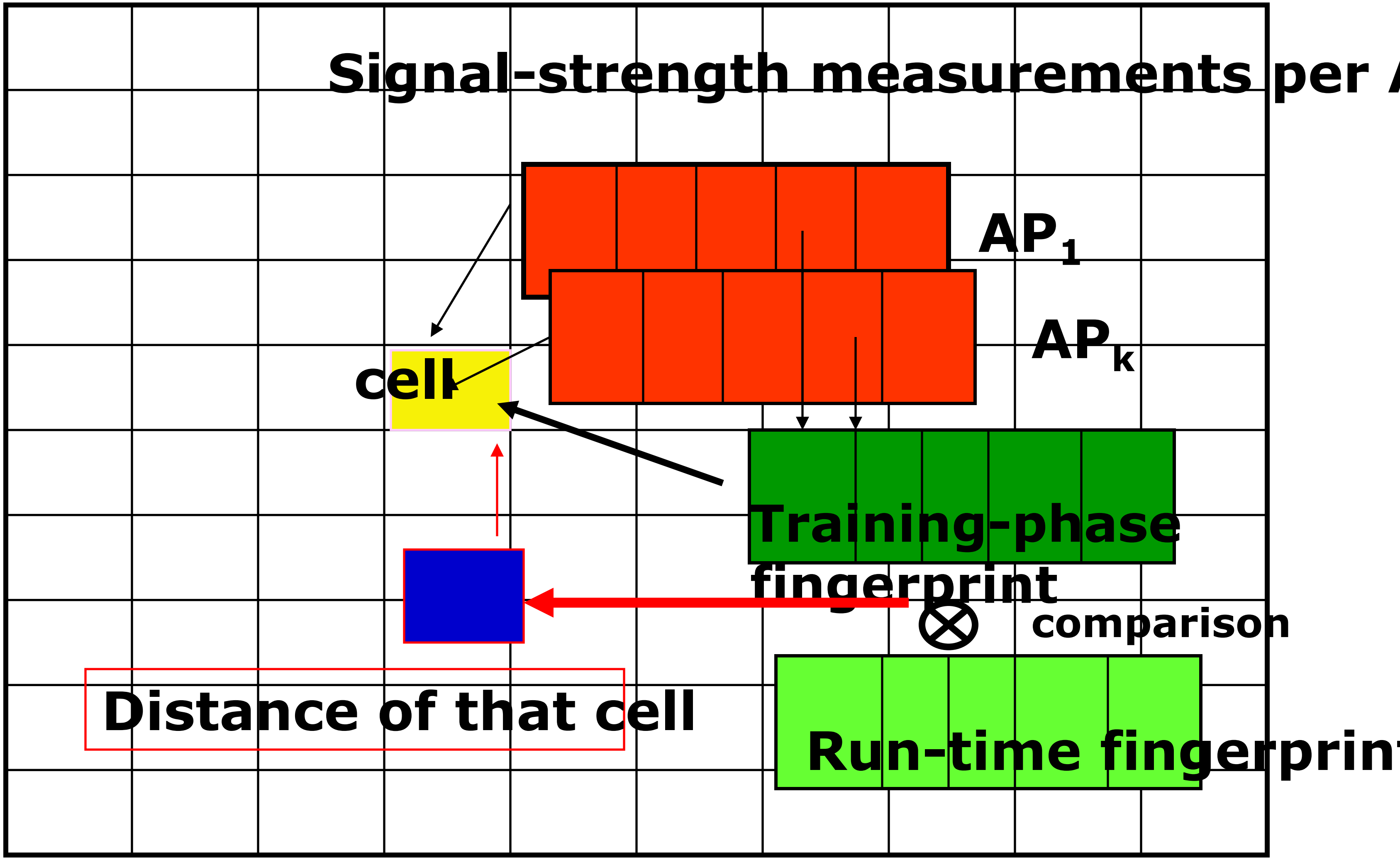
Fingerprint

- A fingerprint can be built using various statistical properties
  - **Mean, standard deviation**
  - **Percentiles**
  - **Empirical distribution** (entire set of signal strength values)
  - **Theoretical models** (e.g., multivariate Gaussian)

- Fingerprint comparison depends on the statistical properties of the fingerprint

  Examples:
  - Euclidean distances, Kullback-Leibler Divergence test

Training & Run-Time Signature Comparison


Signal-strength measurements per

AP$_1$

AP$_k$

cell

Training-phase fingerprint

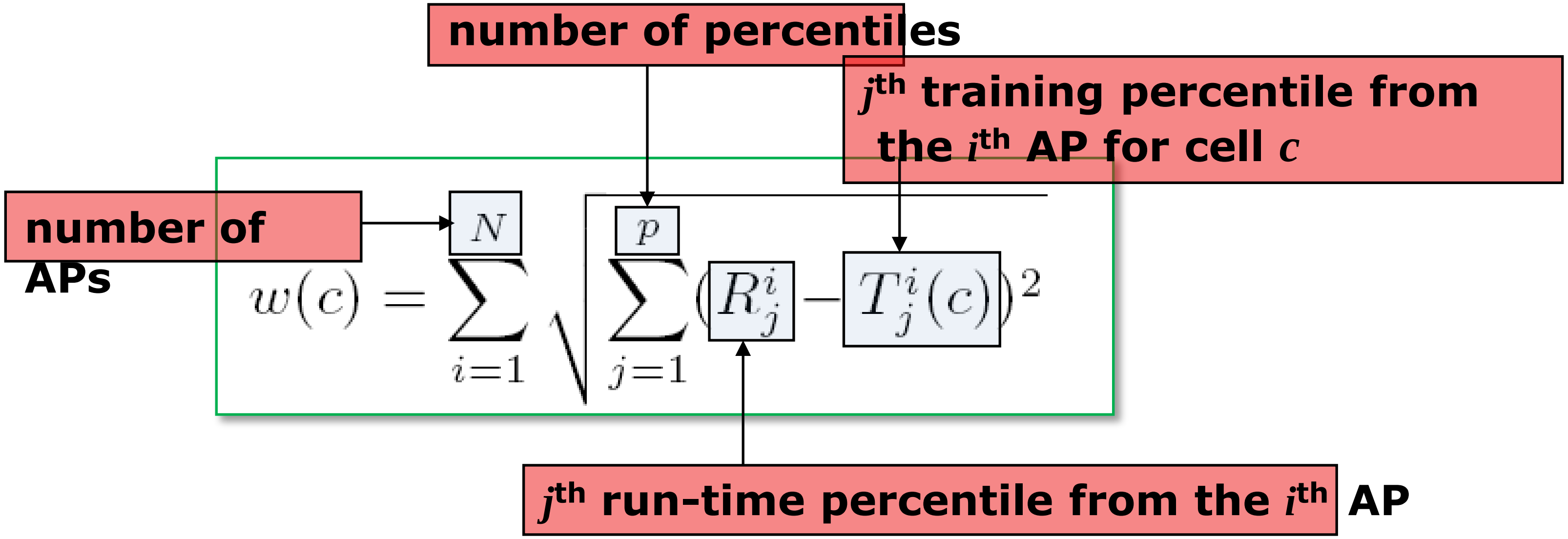
⊗ comparison

Run-time fingerprin

Distance of that cell

Fingerprint Method: Percentiles

Distance of cell *c*, *w(c)*, is computed as follows:



number of percentiles

j^{th} training percentile from the i^{th} AP for cell c

number of APs

$$w(c) = \sum_{i=1}^{N} \sqrt{\sum_{j=1}^{p} \left(R_j^i - T_j^i(c)\right)^2}$$

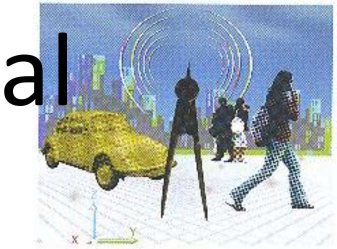
j^{th} run-time percentile from the i^{th} AP

- Estimated position: cell with minimum distance

☞ *Top 5 weighted percentiles:*

weighted centroid of the 5 cells with the smallest distance

Fingerprint Method: Empirical Distribution



- Only APs that appear in *both training and runtime* are used
- Signature uses ***all the RSSI measurements*** collected per AP
- Distance estimation: average ***Kullback-Leibler Divergence (KLD***s) for all APs (between training & runtime fingerprints)
- Select the cell with the smallest distance

Multivariate Gaussian Method: Main Idea

- Statistical-based fingerprint method
- ***Multivariate Gaussian Models*** for the signal strength measurements collected from different APs

✳ Exploit the *2nd order* ***spatial correlations*** between APs

☞ Perform in ***iterations &*** in multiple ***spatial scales*** (regions)

- Use ***Kullback-Leibler Divergence*** (KLD) for distance estimation

Multivariate Gaussian Distribution

Signature of **cell *i*** in ***training phase***: $c_i \mapsto \mathcal{S}_i = \{\vec{\mu}_i, \mathbf{\Sigma}_i\}$

- $\vec{\mu}_i$ *mean values* of the received RSSI measurements per AP
- $\mathbf{\Sigma}_i$: *covariance matrix* (measure of spatial correlation)

Signature of a cell in runtime phase: $c_R \mapsto \mathcal{S}_R = \{\vec{\mu}_R, \mathbf{\Sigma}_R\}$

- $I_R^{i,T}$ APs from which measurements were collected at both training & runtime phases

Multivariate Gaussian Distribution (cnt'd)

- Mean sub-vectors $\vec{\mu}^s_R, \vec{\mu}^s_{i,T}$d covariance sub-matrices $\mathbf{\Sigma}^s_R, \mathbf{\Sigma}^s_{i,T}$ are extracted according to $I^{i,T}_R$
- Multivariate Gaussian density function:

$$p(\vec{x}|\vec{\mu},\mathbf{\Sigma}) = \frac{1}{(2\pi)^{K/2}|\mathbf{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T\mathbf{\Sigma}^{-1}(\vec{x}-\vec{\mu})\right)$$

- KLD between runtime and *i*th training cell:

$$D(p_R||p_{i,T}) = \frac{1}{2}\Big((\vec{\mu}^s_{i,T}-\vec{\mu}^s_R)^T(\mathbf{\Sigma}^s_{i,T})^{-1}(\vec{\mu}^s_{i,T}-\vec{\mu}^s_R) + tr\left(\mathbf{\Sigma}^s_R(\mathbf{\Sigma}^s_{i,T})^{-1}-\mathbf{I}\right) - \ln|\mathbf{\Sigma}^s_R(\mathbf{\Sigma}^s_{i,T})^{-1}|\Big)$$

Estimated position: **Training cell** with *minimum KLD*

Multivariate Gaussian Method

- ☞ Apply Multivariate Gaussian Model in ***multiple spatial scales***
  - Physical space is divided into ***overlapping regions***
  - Signature of a region based on RSSI measurements collected from all APs at various positions in that region
  - Multivariate Gaussian model ***applied in each region***
  - ✳ **Select the region** with the minimum distance

In iterations: selected region also divided into sub-regions

Repeat the above process in that region

until the region becomes a cell

Kullback-Leibler divergence

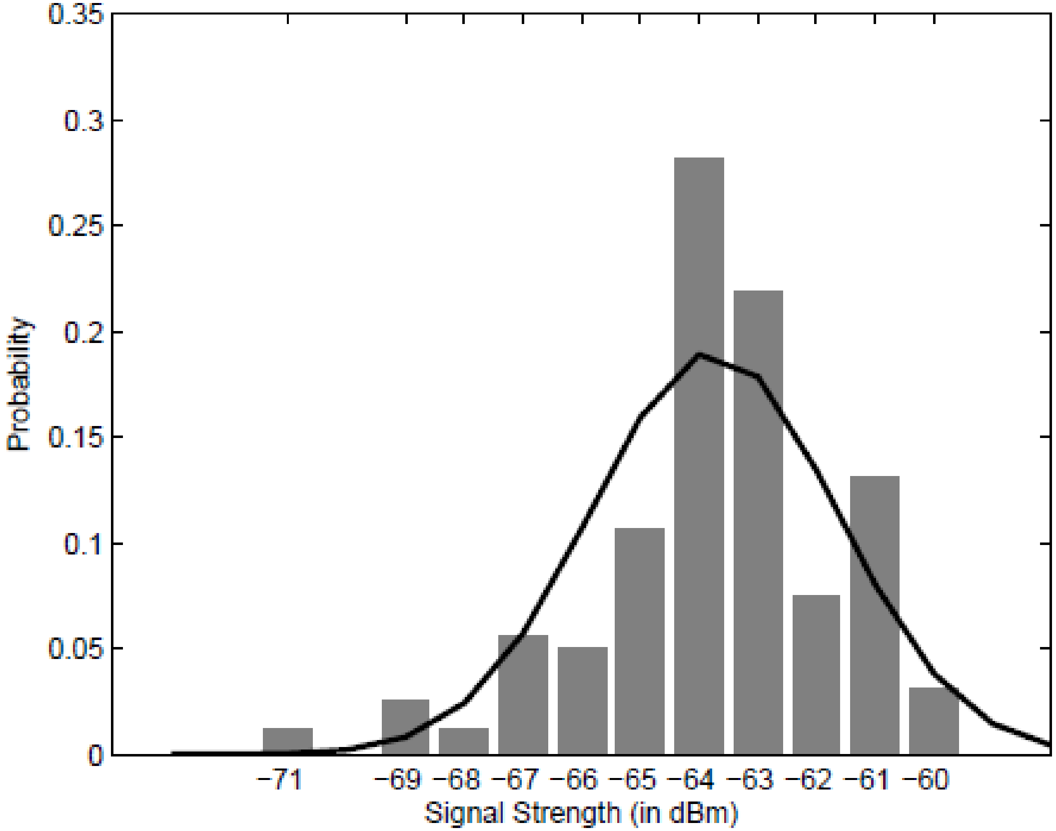
Information gain, relative entropy: a non-symmetric measure
of the **difference between two probability distributions P and Q**

- P represents the "true" distribution of data, observations, or a precise calculated theoretical distribution
- Q represents the theory, model, description or approximation of P

$$D_{\mathrm{KL}}(P\|Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}.$$

$$D_{\mathrm{KL}}(P\|Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} \, \mathrm{d}x,$$

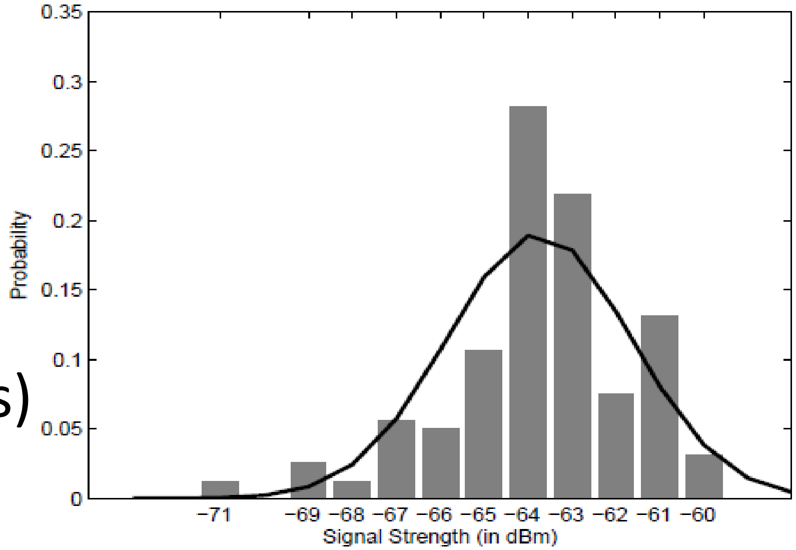
Example of a Fingerprint



Multivariate Gaussian Model

- Each cell corresponds to a Multivariate Gaussian distribution

- Measure the similarity of the Multivariate Gaussian distributions (MvGs) with the KLD closed form:



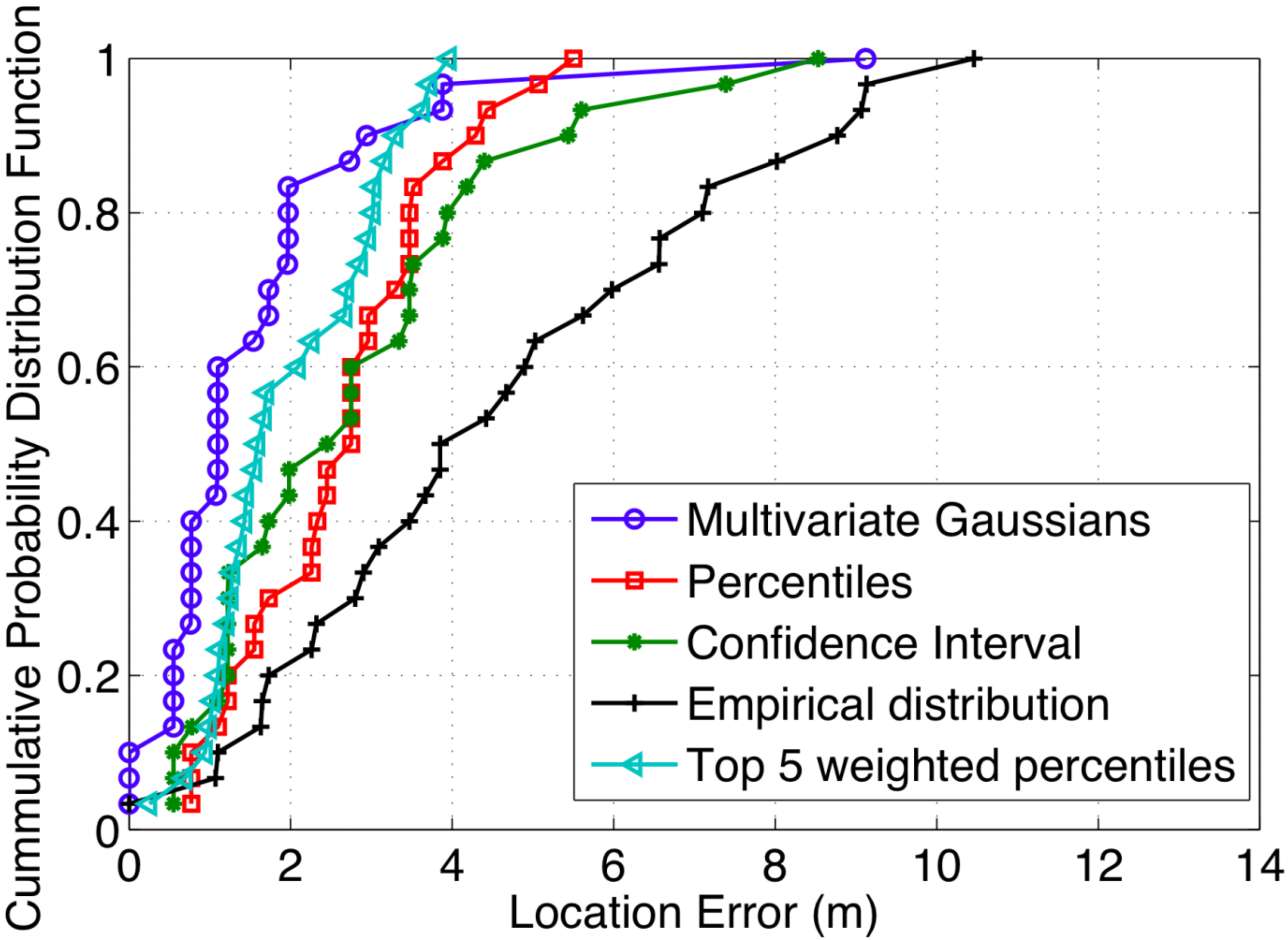
$$
\begin{aligned}
D(MvG_1 \| MvG_2) &= \frac{1}{2}\Big((\mu_2-\mu_1)^T \Sigma_2^{-1} (\mu_2-\mu_1) \\
&+ trace(\Sigma_1 \Sigma_2^{-1} - I) - ln|\Sigma_1 \Sigma_2^{-1}|\Big)
\end{aligned}
$$

Performance Analysis of Fingerprinting

Impact of various parameters

- Number of APs & other reference points (landmarks)
- Size of training set (e.g., number of measurements at various environment conditions (user populations, number of cells, cell size)
- Types of wireless technologies/modalities employed to form fingerprints
- Metrics for computing divergence/“distances”
- Knowledge of the environment
  - Floorplan
  - user mobility

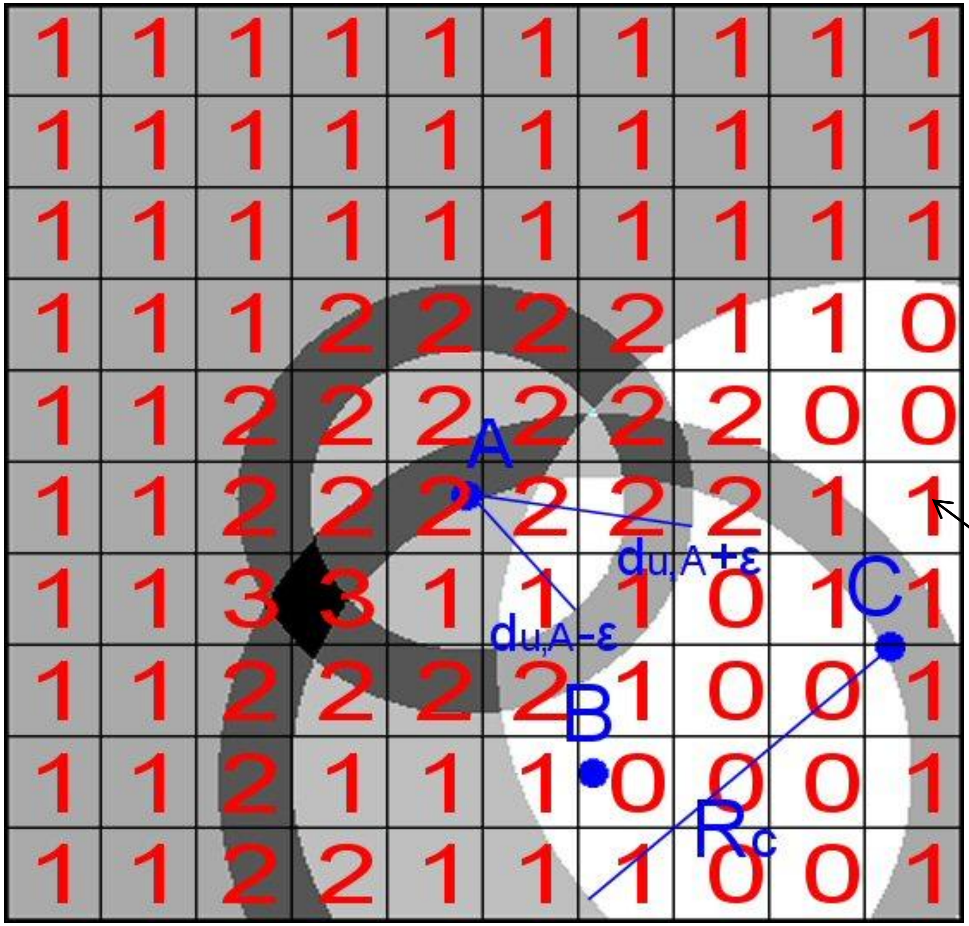
Empirical Results



Collaborative Location Sensing **(CLS)**

- Each host
  - estimates its distance from neighboring peers
  - refines its estimations iteratively as it receives new positioning information from peers
- Voting algorithm to accumulate and assesses the received positioning information
- Grid-representation of the terrain

Example of grid with accumulated votes



1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	2	2	2	2	1	1	0
1	1	2	2	2	2	2	2	0	0
1	1	2	2	2	2	2	2	1	1
1	1	3	3	1	1	1	0	1	1
1	1	2	2	2	2	1	0	0	1
1	1	2	1	1	1	0	0	0	1
1	1	2	2	1	1	1	0	0	1

Grid for host **u**
Corresponds to the **terrain**

Host **u** tries to position itself

A **cell** is a possible position

Peers **A, B, C**

The value of a cell in the grid is the sum of the accumulated votes
The higher the value, the more hosts it is likely position of the host

Multi-modal Positioning System: Cricket (1/4)

- ***Cricket "beacons"*** mounted on the ceiling and consists of:
  - a micro-controller running at 10MHz, with 68 byres of RAM and 1024 words of program memory, lower power ***RF-transmitter,*** and ***single-chip RF receiver***, both in 418MHz unlicensed band
  - ***Ultrasonic transmitter*** operating at 40Hz
- A similar interface at the **client** (e.g., laptop, printer)

Cricket (2/4)

- A cricket beacon sends **concurrently** an ***RF message*** (with info about the space) & an ***ultrasonic pulse***

When the **listener at a client receives the RF signal**, it performs the following:

1. uses the first few bits as training information
2. **turns on its ultrasonic receiver**
3. listens for the ultrasonic pulse
   which will usually arrive a short time later

4. **correlates** the RF signal & ultrasonic pulse

5. determines the distance to the beacon
   from the ***time difference*** between the ***receipt of the first bit RF*** information & the ***ultrasonic pulse***

Cricket (3/4)

- Lack of coordination can cause:
  - RF transmissions from different cricket beacons to collide
  - A listener may correlate incorrectly the RF data of one beacon with the ultrasonic signal of another, yielding false results
- Ultrasonic reception suffers from severe multi-path effect
- **Order of magnitude longer in time** than RF multi-path because of the ***relatively long propagation time*** of sound waves in air

Cricket (4/4)

- Handles the problem of collisions using randomization:

  beacon transmission times are chosen randomly with a uniform distribution within an interval

⇒ the broadcasts of different beacons are statistically independent, which avoids repeated synchronization & persistent collisions

- Statistical analysis of correlated RF, US samples

Proximity

- Physical contact

  e.g., with pressure, touch sensors or capacitive detectors
- Within range of an access point
- Automatic ID systems
  - computer login
  - credit card sale
  - RFID
  - UPC product codes

Sensor Fusion

- Seeks to improve accuracy and precision by aggregating many location-sensing systems (modalities/sources)

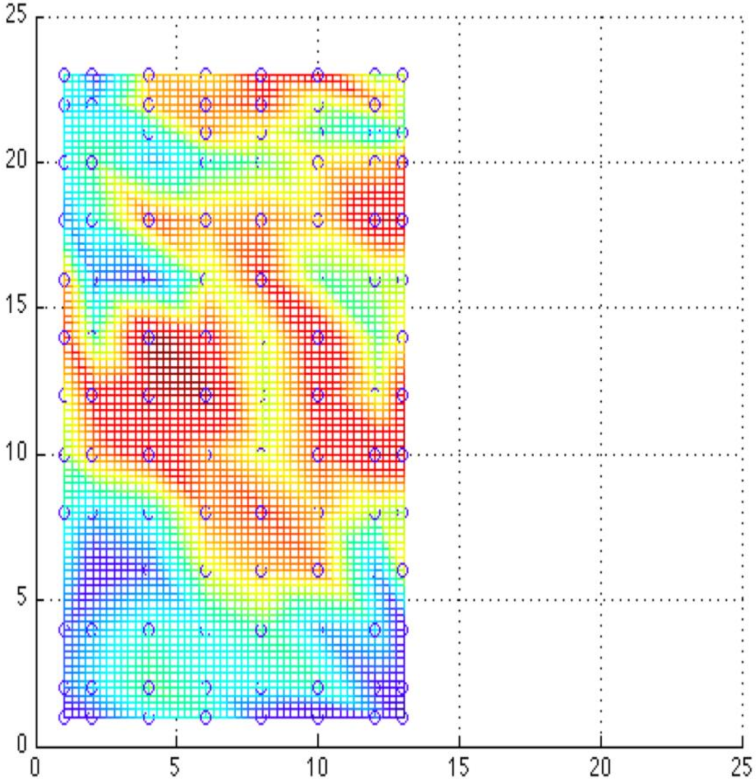
  to form hierarchical & overlapping levels of resolution
- Robustness when a certain location-sensing system (source) becomes unavailable

Issue: assign weight/importance to the different location-sensing systems

Technology Name	Properties						
	Technique	Phys	Symb	Abs	Rel	LLC	Recognition
GPS	Radio time-of-flight lateration	●		●		√	
Active Badges	Diffuse infrared cellular proximity		●	●			√
Active Bats	Ultrasound time-of-flight lateration	●		●			√
MotionStar	Scene analysis, lateration	●		●			√
VHF Omnidirectional Ranging (VOR)	Angulation	●		●		√	
Cricket	Proximity, lateration		●	○	○	√	
MSR RADAR	802.11 RF scene analysis & triangulation	●		●			√
PinPoint 3D-iD	RF lateration	●		●			√
Avalanche Transceivers	Radio signal strength proximity	●			●		
Easy Living	Vision, triangulation		●	●			√
Smart Floor	Physical contact proximity	●		●			√
Automatic ID Systems	Proximity		●	○	○		√
Wireless Andrew	802.11 cellular proximity		●	●			√
E911	Triangulation	●		●			√
SpotON	Ad hoc lateration	●			●		√

Backup Slides

Signal-Strength



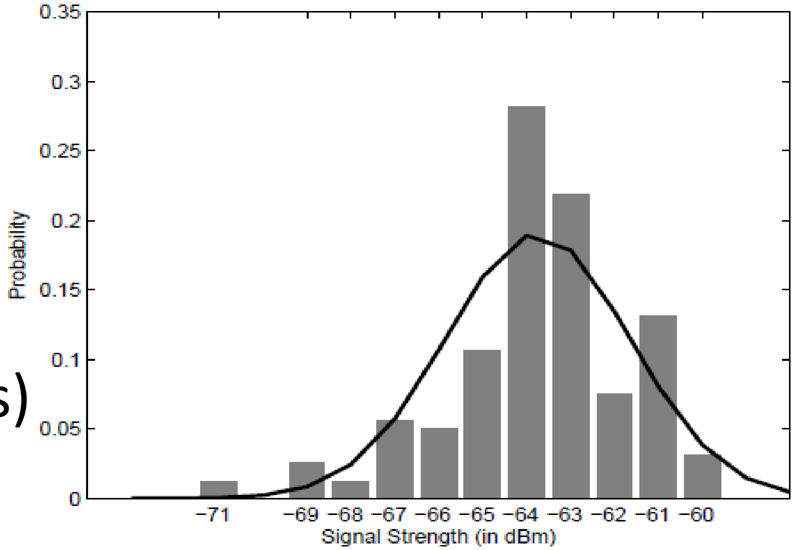
Left plot: Busy period , Right plot: Quiet period (TNL’s AP)

Multivariate Gaussian Model

- Fingerprint using signal-strength measurements from each AP and the interplay (covariance) of measurements from pairs of Aps
- Signature comparison is based on the Kullback-Leibler Divergence

Multivariate Gaussian Model

- Each cell corresponds to a Multivariate Gaussian distribution

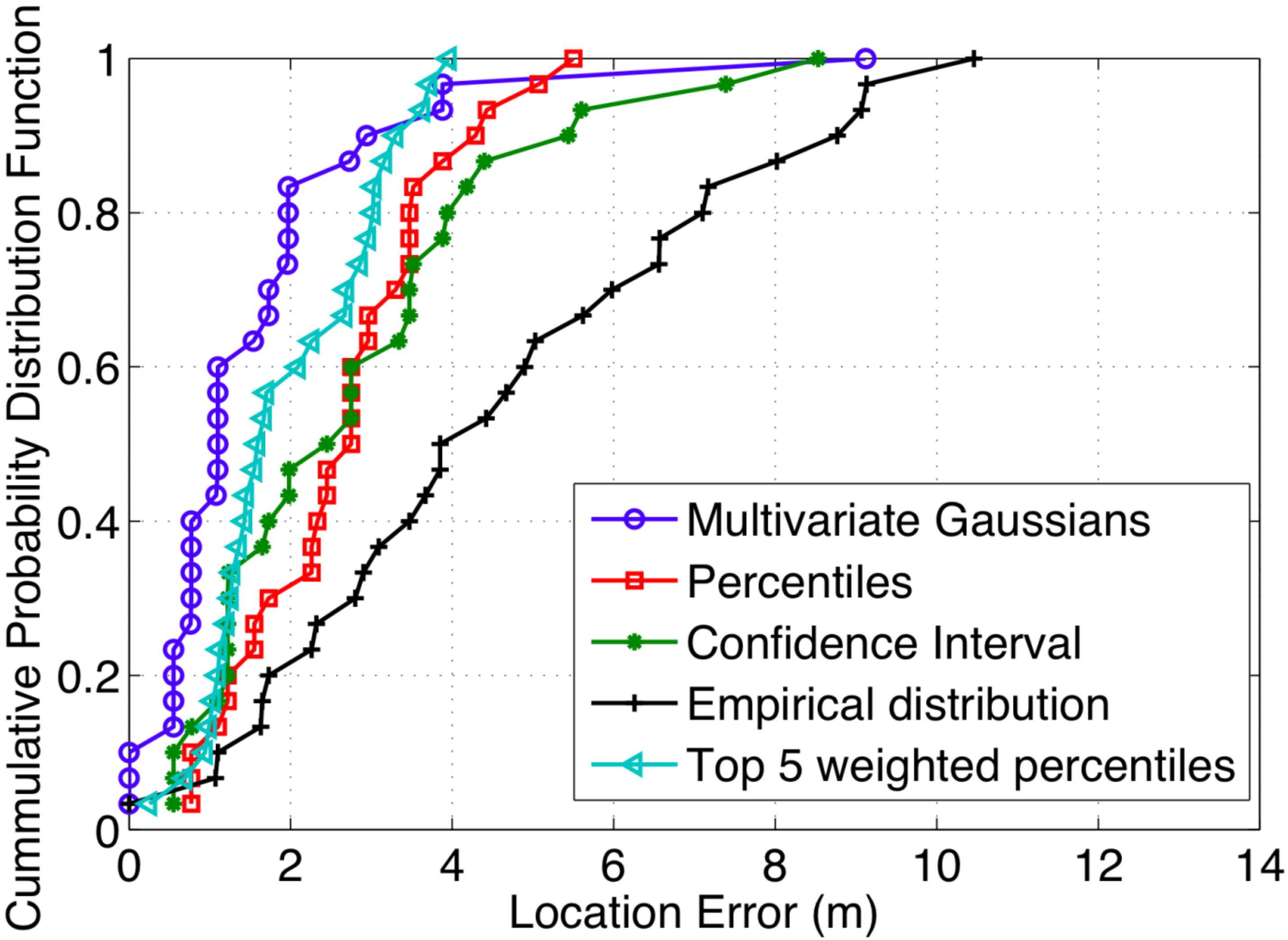


- Measure the similarity of the Multivariate Gaussian distributions (MvGs) with the KLD closed form:

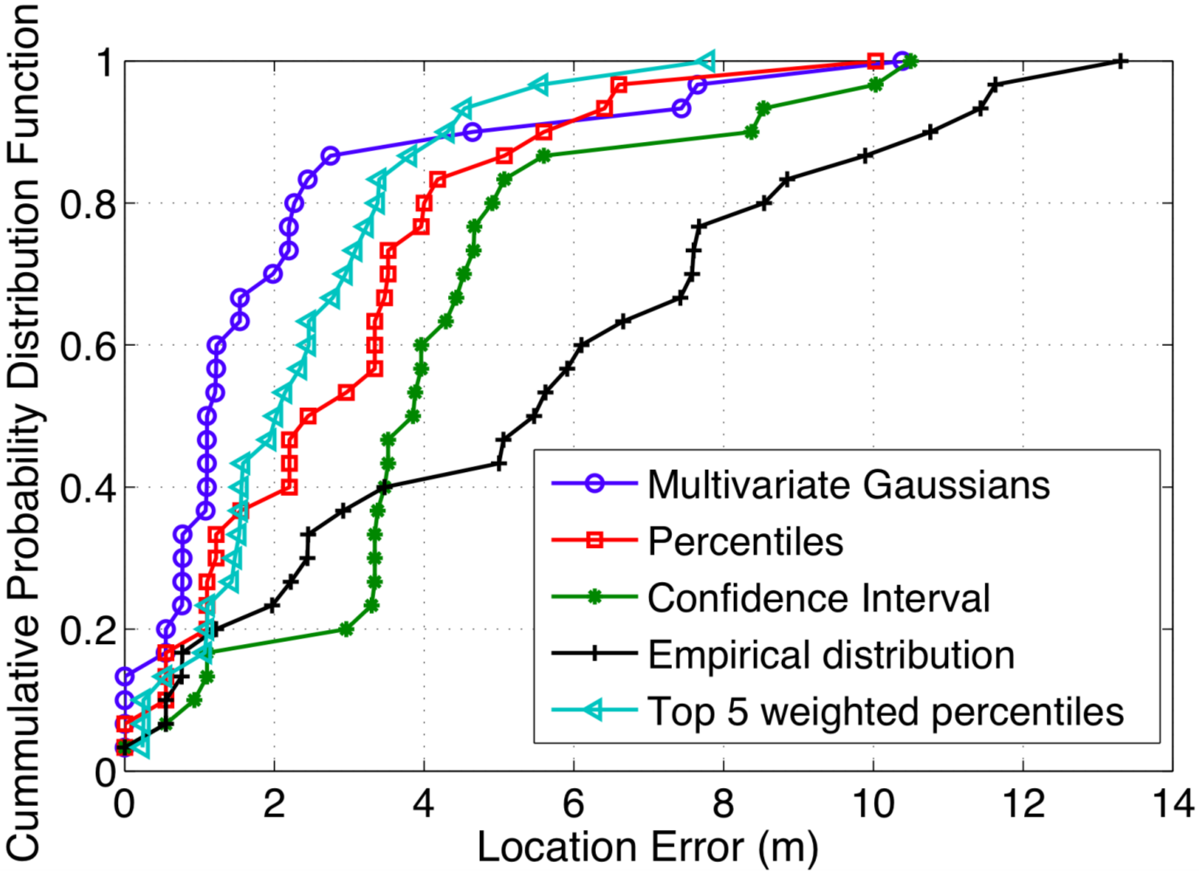
$$
\begin{aligned}
D(MvG_1 \| MvG_2) &= \frac{1}{2}\Big((\mu_2-\mu_1)^T \Sigma_2^{-1} (\mu_2-\mu_1) \\
&+ trace(\Sigma_1 \Sigma_2^{-1} - I) - ln|\Sigma_1 \Sigma_2^{-1}|\Big)
\end{aligned}
$$

Cretaquarium

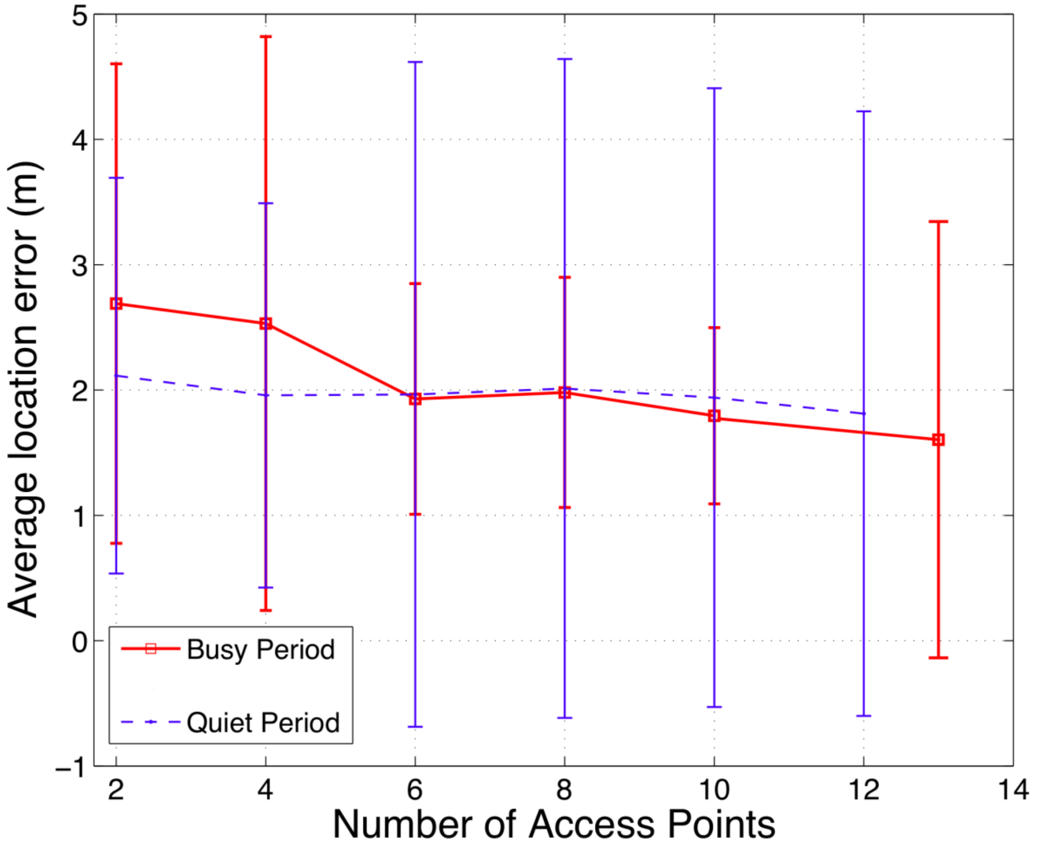
Empirical Results (Busy Period)



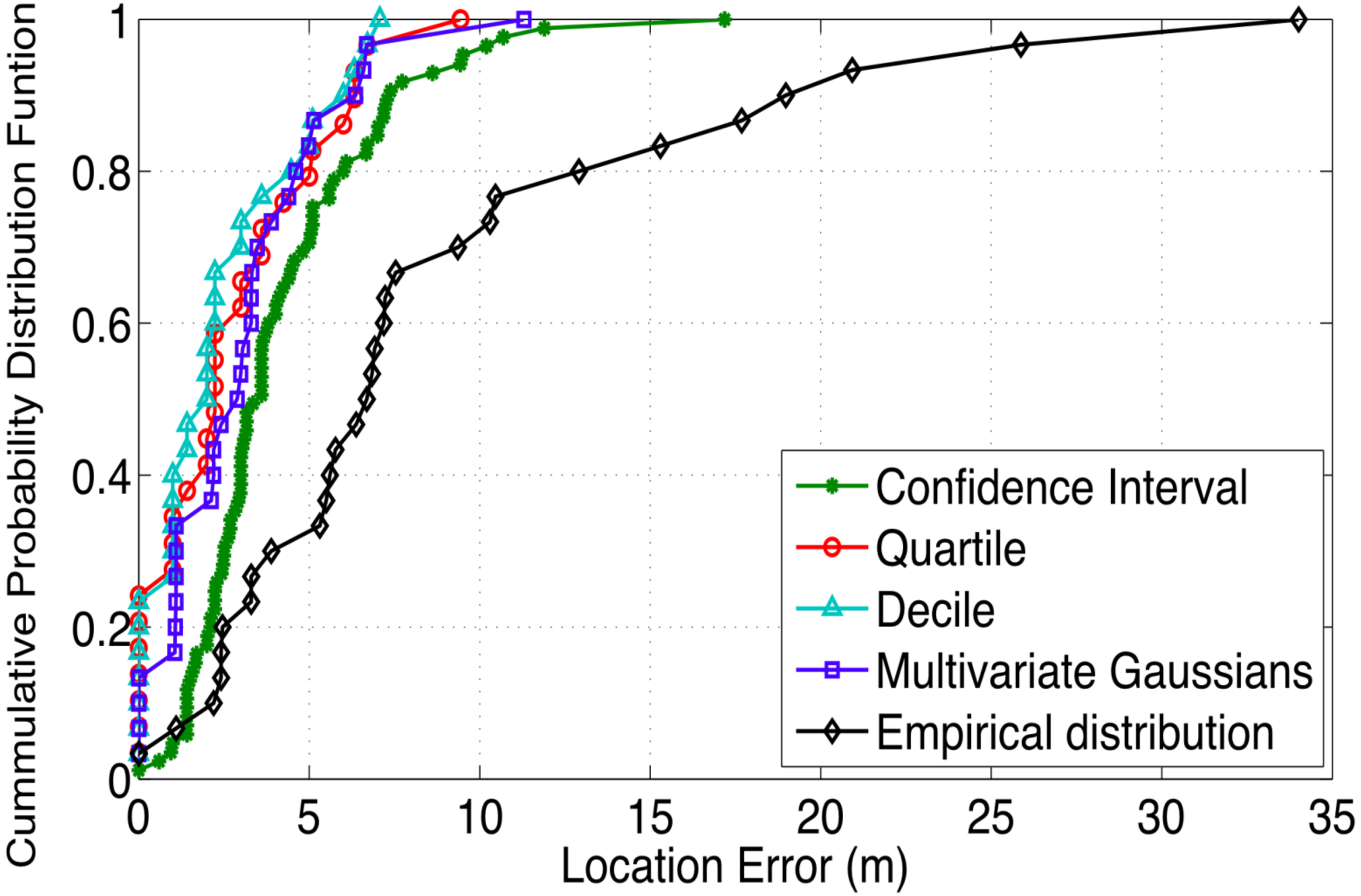
Empirical Results (Quiet Period)



Impact of the Number of APs



Empirical Results @ Cretaquarium



Collaborative Location Sensing **(CLS)**

- Each host
  - estimates its distance from neighboring peers
  - refines its estimations iteratively as it receives new positioning information from peers
- Voting algorithm to accumulate and assesses the received positioning information
- Grid-representation of the terrain

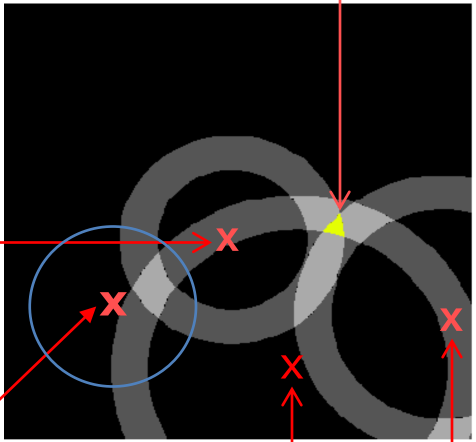
Example of voting process @ host u

Host u with unknown position

Peers A, B, C, and D have positioned themselves

Most likely position

Host **A** positioned at the center of the co-centric disks



Host A (positive vote)

Host **C** **(negative vote)**

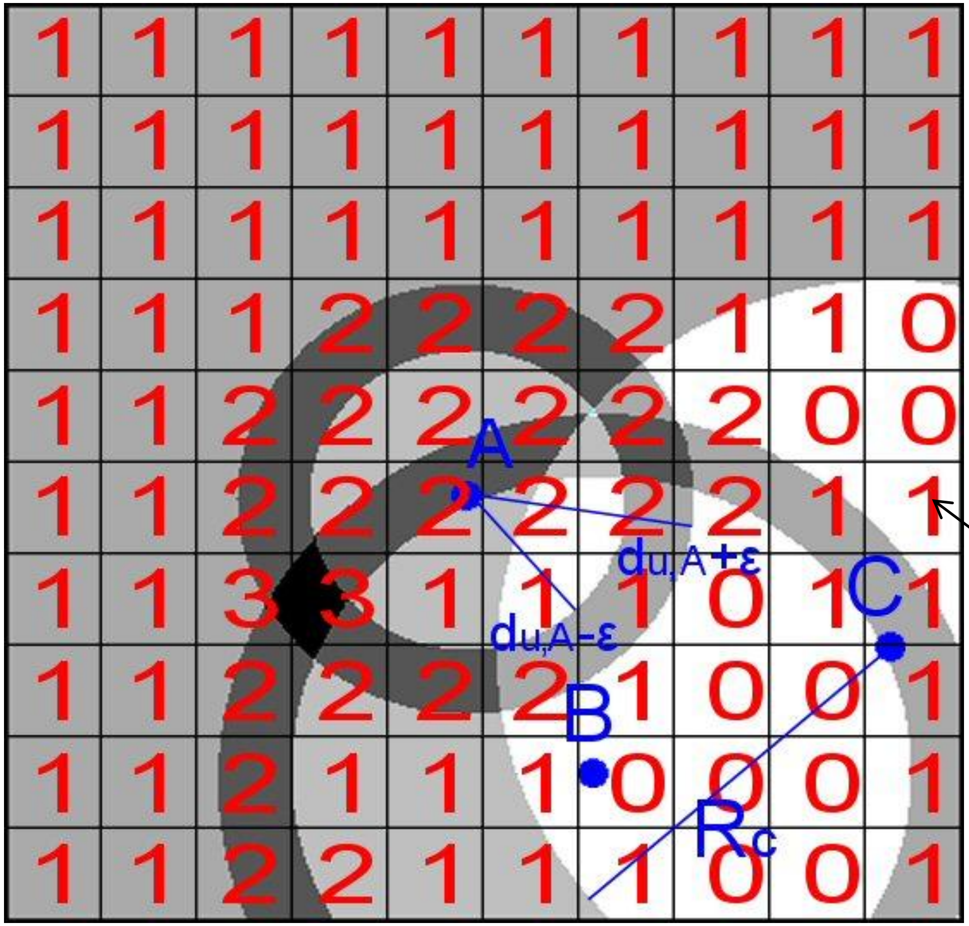
Host **D** votes (positive vote)

Host B votes (positive vote)

positive votes from peers A, B, D

negative vote from peer C

Example of grid with accumulated votes



1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	2	2	2	2	1	1	0
1	1	2	2	2	2	2	2	0	0
1	1	2	2	2	2	2	2	1	1
1	1	3	3	1	1	1	0	1	1
1	1	2	2	2	2	1	0	0	1
1	1	2	1	1	1	0	0	0	1
1	1	2	2	1	1	1	0	0	1

Grid for host **u**
Corresponds to the **terrain**

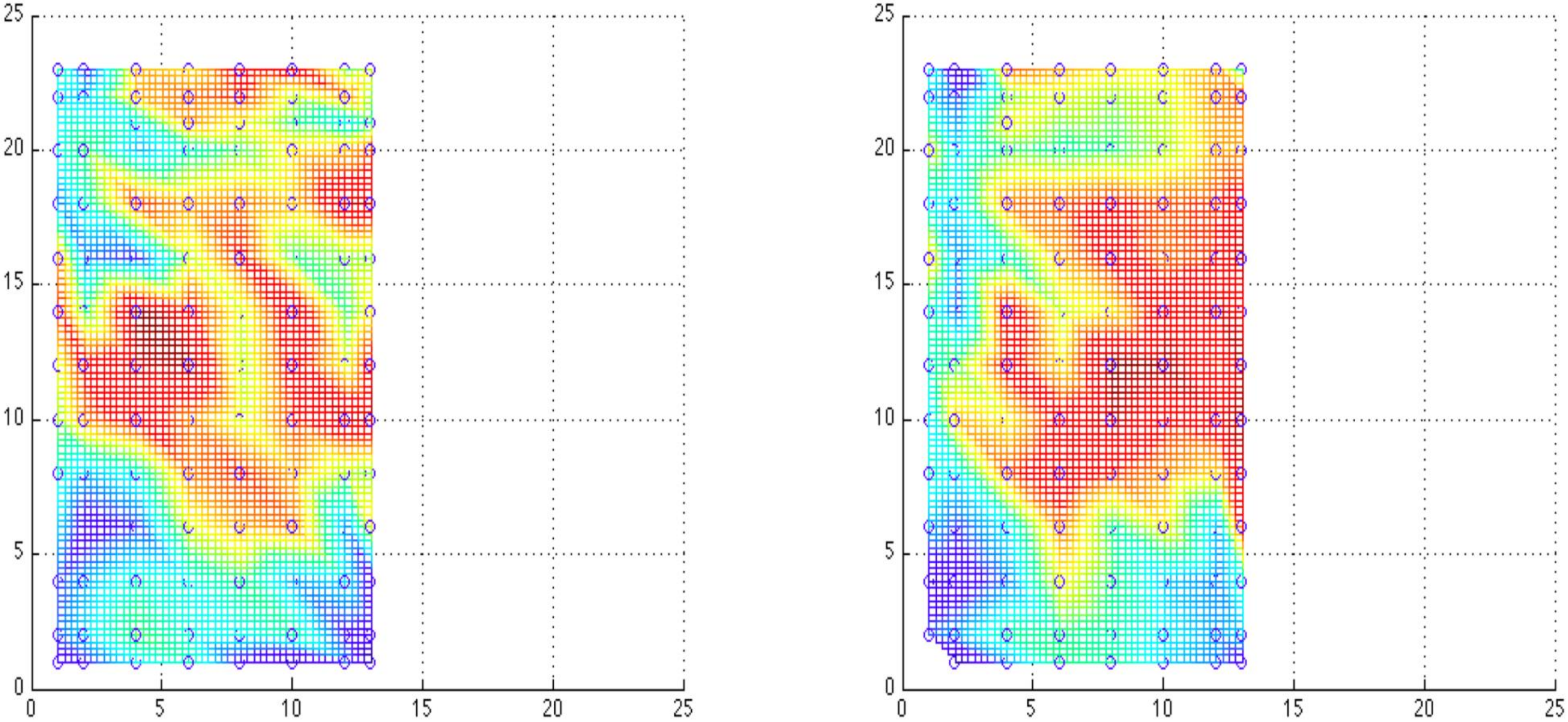
Host **u** tries to position itself

A **cell** is a possible position

Peers **A, B, C**

The value of a cell in the grid is the sum of the accumulated votes
The higher the value, the more hosts it is likely position of the host

Signal-Strength



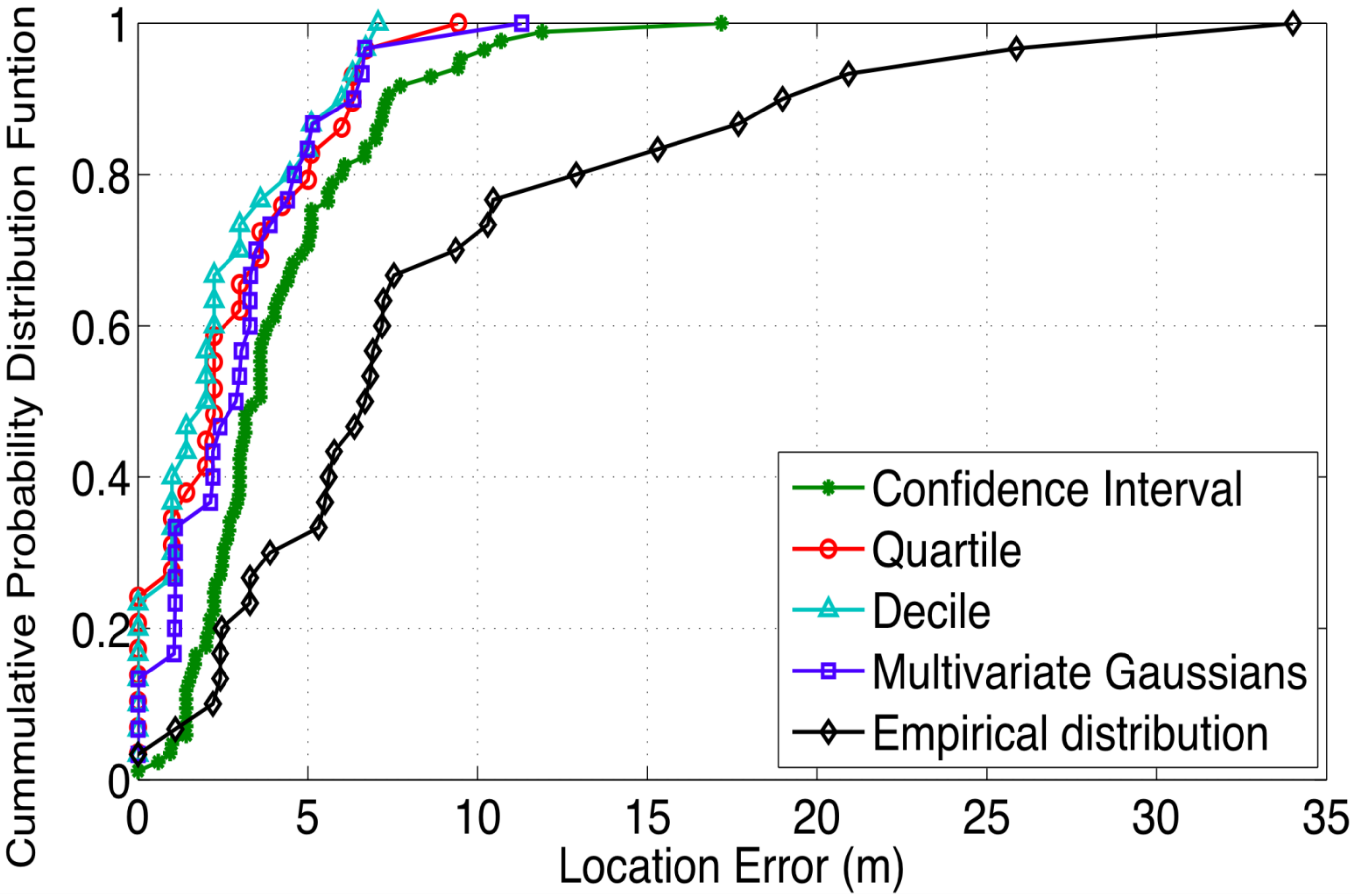
Left plot: Busy period , Right plot: Quiet period (TNL’s AP)

Multivariate Gaussian Model

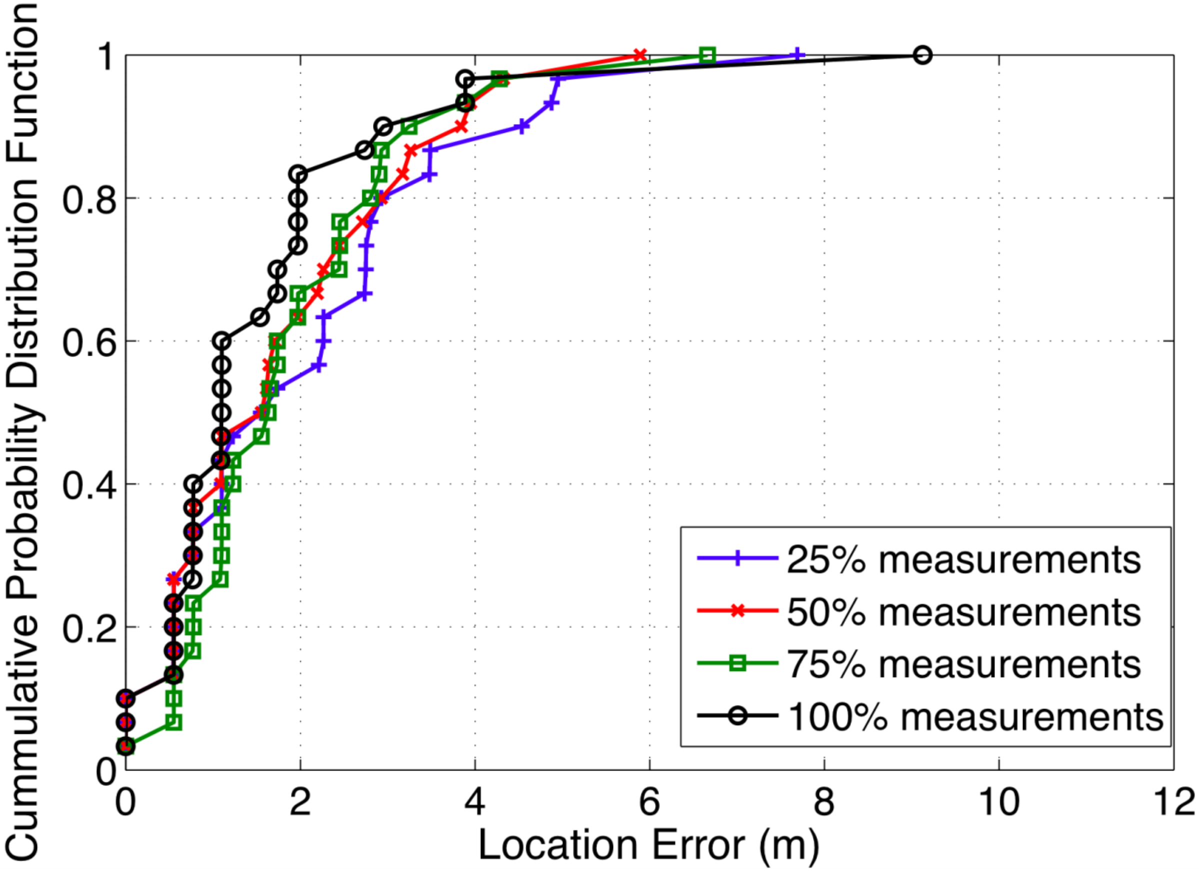
- Fingerprint using signal-strength measurements from each AP and the interplay (covariance) of measurements from pairs of Aps
- Signature comparison is based on the Kullback-Leibler Divergence

Cretaquarium

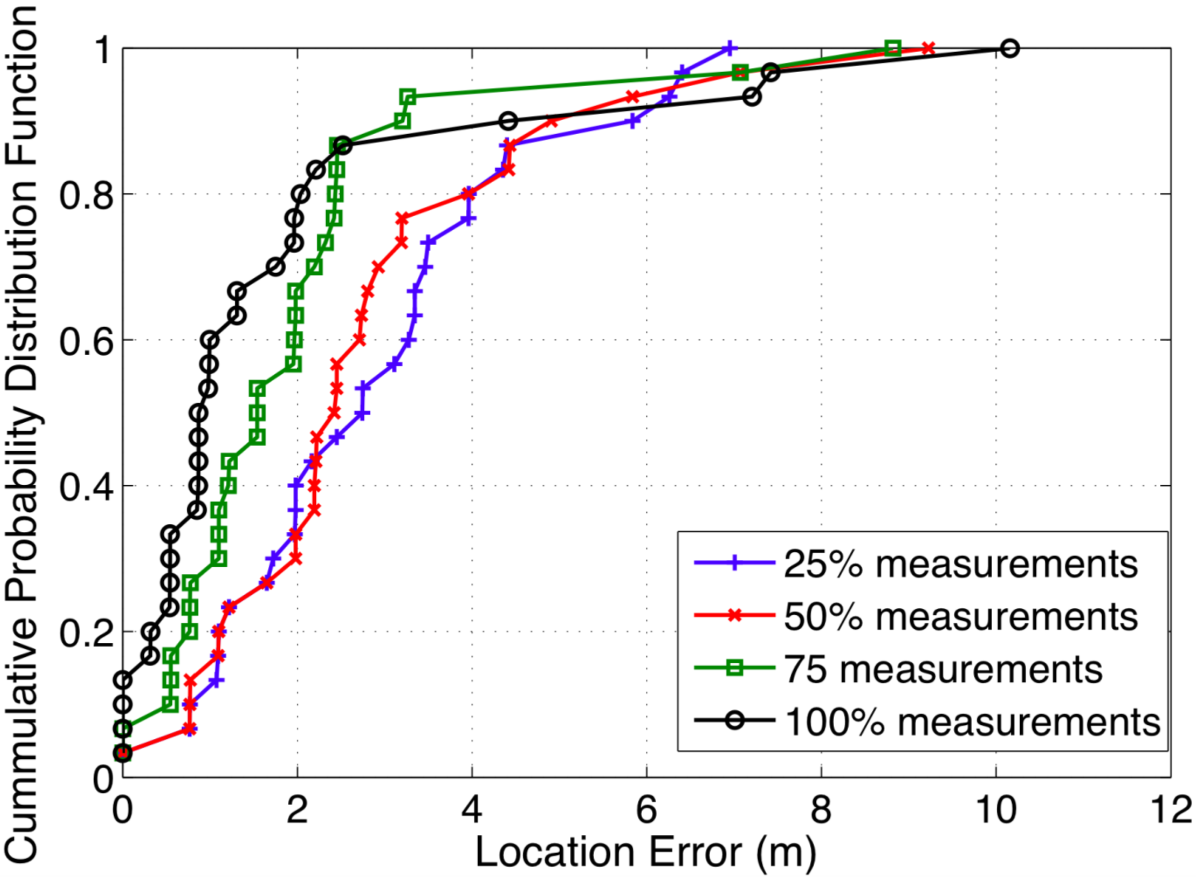
Empirical Results @ Cretaquarium



Empirical Results (Busy Period)

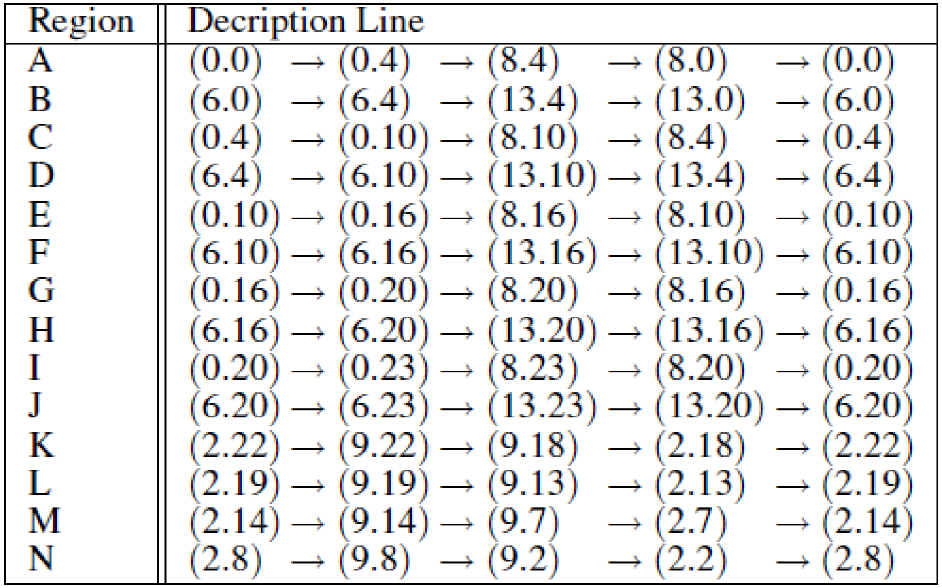


Experimental Results – Quiet Period (%)



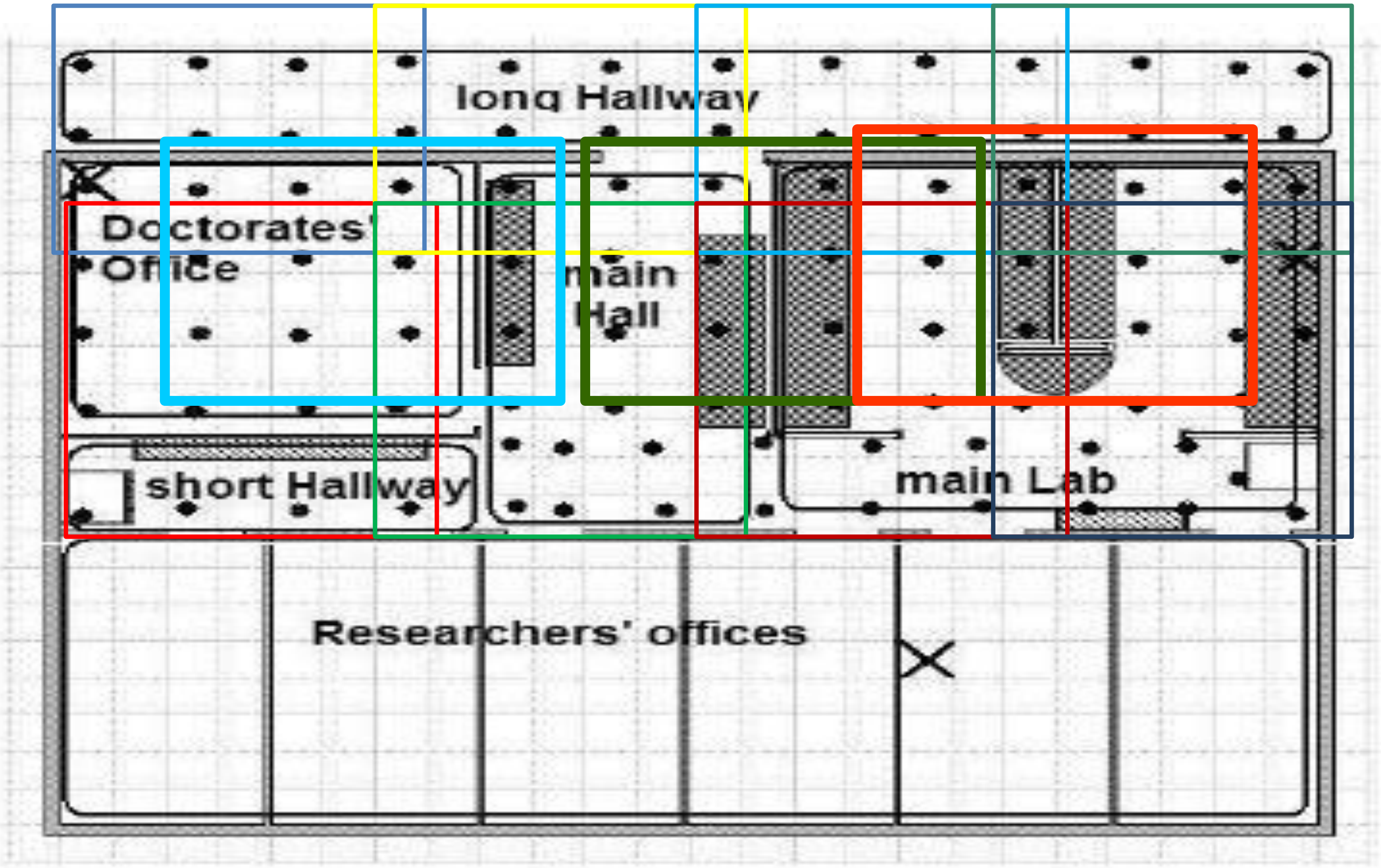
Splitting into areas of cells - TNL (1/2)

- Split the grid in 14 regions, namely from A to N
  - The regions are overlapped
  - Collect the data from each cell that belongs in this region
  - Concat them in a new file named Region{A to N}
  - 16 APs average in every region



Region	Decription Line
A	(0.0) → (0.4) → (8.4) → (8.0) → (0.0)
B	(6.0) → (6.4) → (13.4) → (13.0) → (6.0)
C	(0.4) → (0.10) → (8.10) → (8.4) → (0.4)
D	(6.4) → (6.10) → (13.10) → (13.4) → (6.4)
E	(0.10) → (0.16) → (8.16) → (8.10) → (0.10)
F	(6.10) → (6.16) → (13.16) → (13.10) → (6.10)
G	(0.16) → (0.20) → (8.20) → (8.16) → (0.16)
H	(6.16) → (6.20) → (13.20) → (13.16) → (6.16)
I	(0.20) → (0.23) → (8.23) → (8.20) → (0.20)
J	(6.20) → (6.23) → (13.23) → (13.20) → (6.20)
K	(2.22) → (9.22) → (9.18) → (2.18) → (2.22)
L	(2.19) → (9.19) → (9.13) → (2.13) → (2.19)
M	(2.14) → (9.14) → (9.7) → (2.7) → (2.14)
N	(2.8) → (9.8) → (9.2) → (2.2) → (2.8)

Splitting into areas of cells - TNL (2/2)



Testbed description - Aquarium

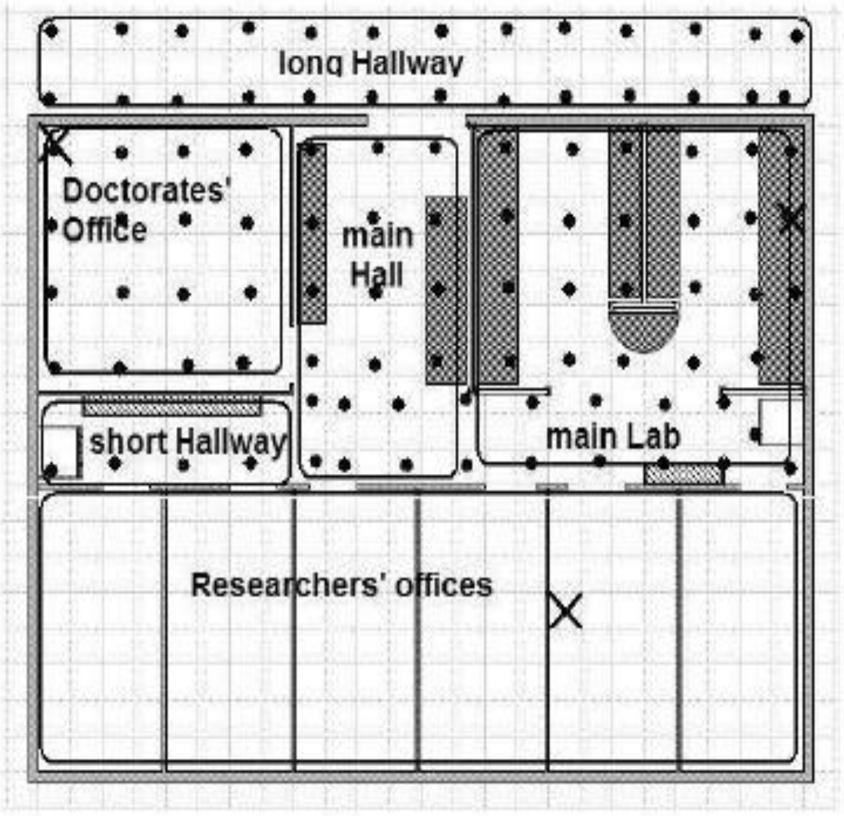
- 1760 m^2
- 30 tanks (extra 25 will be installed)
- 8 APs
- Cell’s size: 1m x 1m
- 5.7 APs on average were collected
- About 150 visitors

Test bed description - TNL

- 7 x 12 m
- Cell's size: 55 x 55 cm
- 13 APs
- 6 APs average detected at a cell

108 training cells

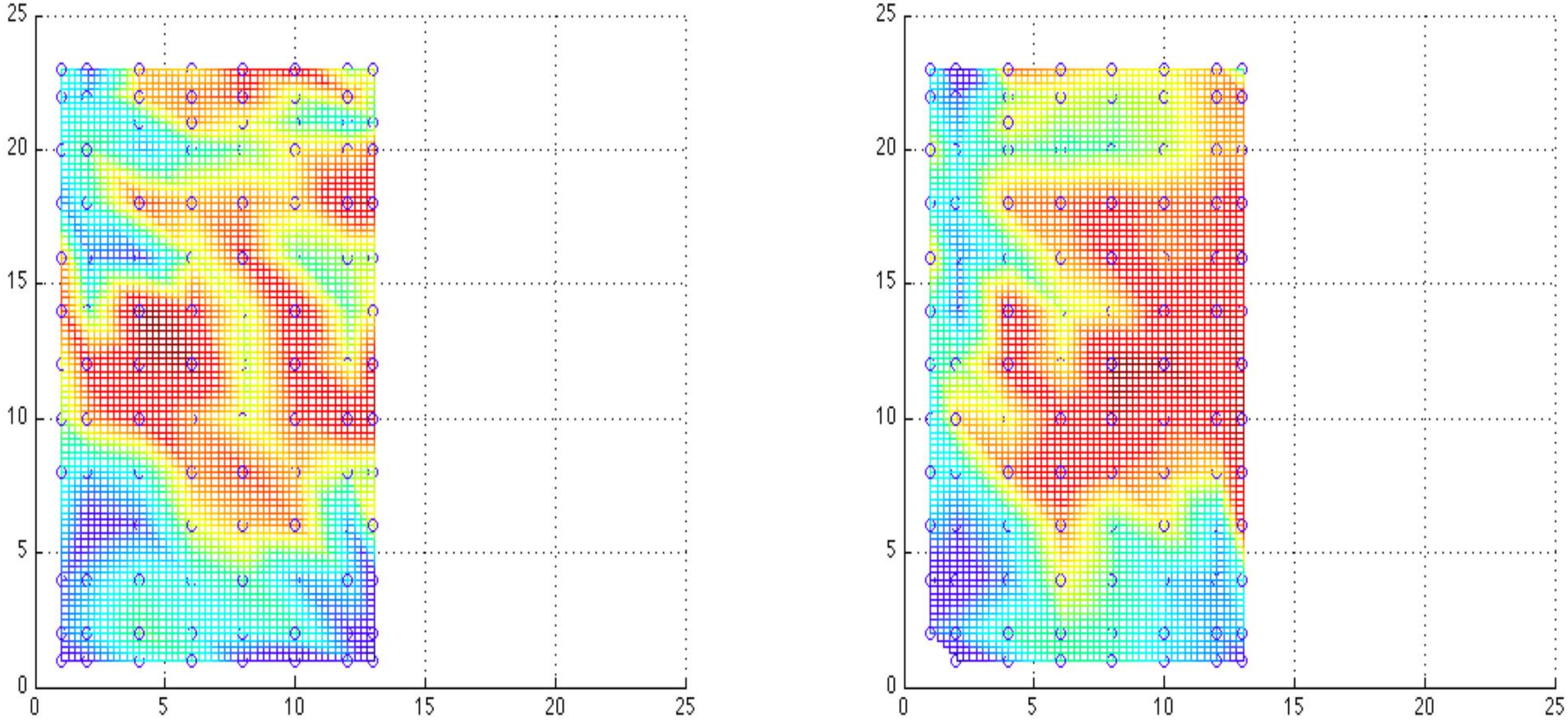
30 run-time cells



Experimental Results

- Two real map databases obtained from TNL
  - Busy period data
  - Quiet period data
- Real database obtained from Cretaquarium (Normal period data)
- Performance of positioning in terms of localization error.
- Measured by averaging the Euclidean distance between the estimated location of the mobile user and its true location

Signal-Strength



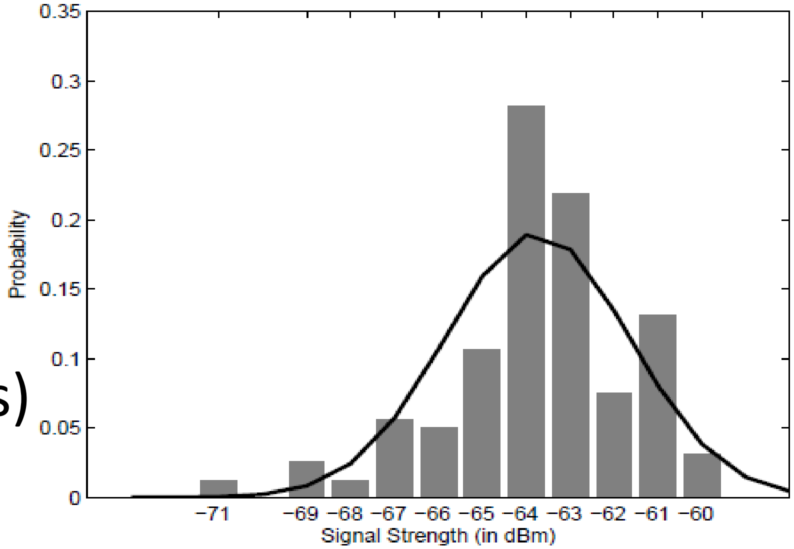
Left plot: Busy period , Right plot: Quiet period (TNL’s AP)

Multivariate Gaussian Model

- Fingerprint using signal-strength measurements from each AP and the interplay (covariance) of measurements from pairs of Aps
- Signature comparison is based on the Kullback-Leibler Divergence

Multivariate Gaussian Model

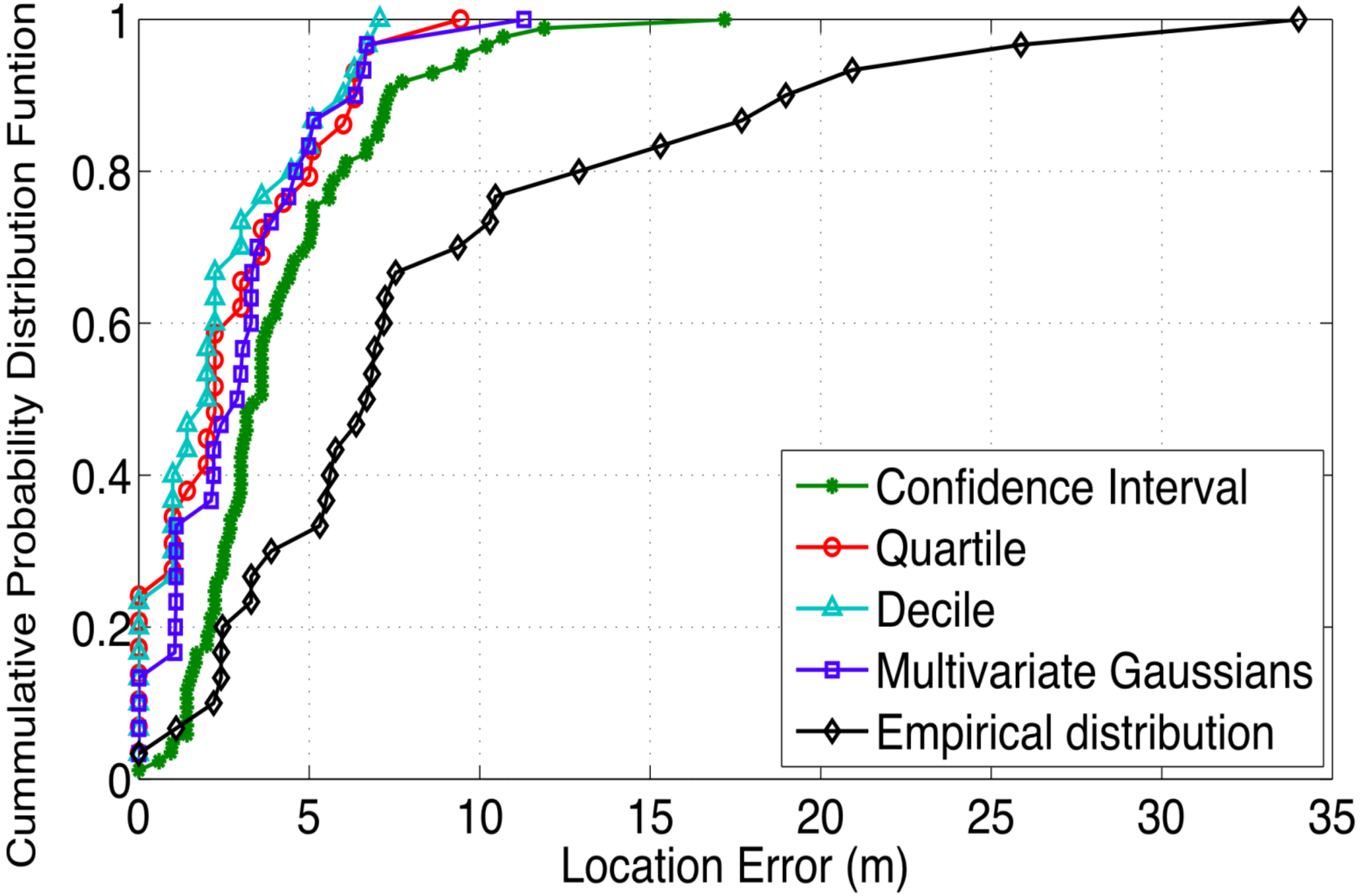
- Each cell corresponds to a Multivariate Gaussian distribution
- Measure the similarity of the Multivariate Gaussian distributions (MvGs) with the KLD closed form:



$$
\begin{aligned}
D(MvG_1\|MvG_2) &= \frac{1}{2}((\mu_2-\mu_1)^T\Sigma_2^{-1}(\mu_2-\mu_1) \\
&+ trace(\Sigma_1\Sigma_2^{-1}-I)-ln|\Sigma_1\Sigma_2^{-1}|)
\end{aligned}
$$

Cretaquarium

Empirical Results @ Cretaquarium



Empirical Results (Quiet Period)