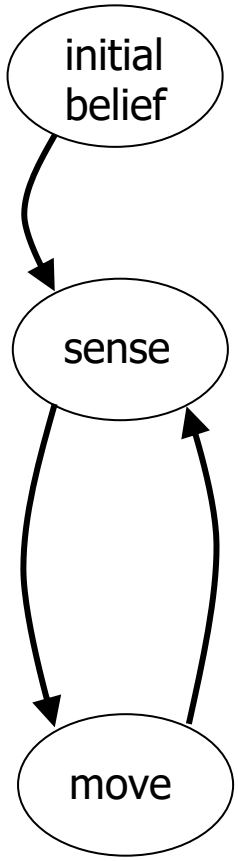
Lecture 03

SmartPhone Sensing

First assignment

- Confusion matrices
- Picture of App
- 1 page, please!
  - Don't explain things that were already explained in class.
  - If you use say, SVM, explained
- Code must be on phone!

Recap



current pdf (**posterior**) | perception model (sense) | pdf from last time step (prior)

$$p(X_k \mid Z_{1:k}) = \frac{p(Z_k \mid X_k)\; p(X_k \mid Z_{1:k-1})}{p(Z_k \mid Z_{1:k-1})}$$

normalization

current pdf (**posterior**) | motion model (move) | pdf from last time step (prior)

$$p(X_k \mid Z_{1:k-1}) = \int p(X_k \mid X_{k-1})\; p(X_{k-1} \mid Z_{1:k-1})\, dX_{k-1}$$

Sense

perception model

p(z="red" | Xi is "red") = 0.6 p(z="blue" | Xi is "blue") = 0.8

p(z="blue" | Xi is "red") = 0.4 p(z="red" | Xi is "blue") = 0.2

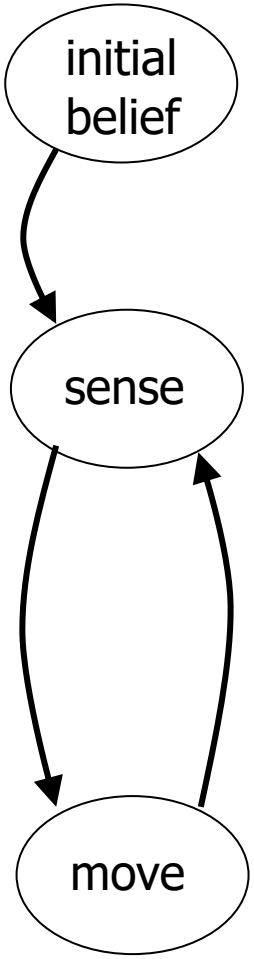
0.2	**0.2**	**0.2**	**0.2**	**0.2**	p(x) [prior]
0.2	0.6	0.6	0.2	0.2	p(z|x)
0.04	0.12	0.12	0.04	0.04	p(z|x)*p(x)
0.04/0.36	0.12/0.36	0.12/0.36	0.04/0.36	0.04/0.36	normalize
1/9	**1/3**	**1/3**	**1/9**	**1/9**	**p(x)** [posterior]

current pdf (**posterior**) = perception model (sense) × pdf from last time step (prior) / normalization

$$p(X_k \mid Z_{1:k}) = \frac{p(Z_k \mid X_k)\, p(X_k \mid Z_{1:k-1})}{p(Z_k \mid Z_{1:k-1})}$$

You need a good training phase to build your "perception model" This is key!

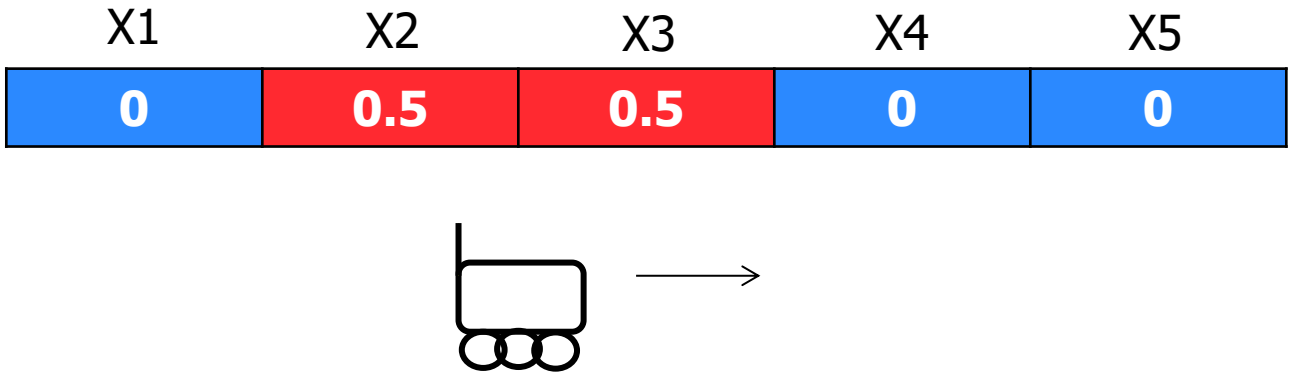
Now let's look at the movement part



current pdf (**posterior**) $p(X_i^t \mid X_j^{t-1}) =$ Σ motion model (move) $p(X_i \mid X_j)$ pdf from last time step (prior) $p(X_j^{t-1})$

Movement Model

movement model	p(Xi | Xi) = 0.1 p(Xi+1 | Xi) = 0.8 p(Xi+2 | Xi) = 0.1

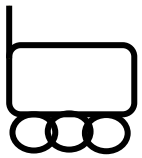


Exercise: the robot is moving right, what is the new distribution?

Move

movement model	p(Xi | Xi) = 0.1 p(Xi+1 | Xi) = 0.8 p(Xi+2 | Xi) = 0.1

X1	X2	X3	X4	X5
0	**0.5**	**0.5**	**0**	**0**
0	0.05	0.45	0.45	0.05



When using RF, you don't want to see people teletransporting in your localization app. Movements models (any model) help a ton!

Congratulations

you can now do discrete localization!!!

let's see how this can be used with phones

Paper 1: The Horus WLAN Location Determination System

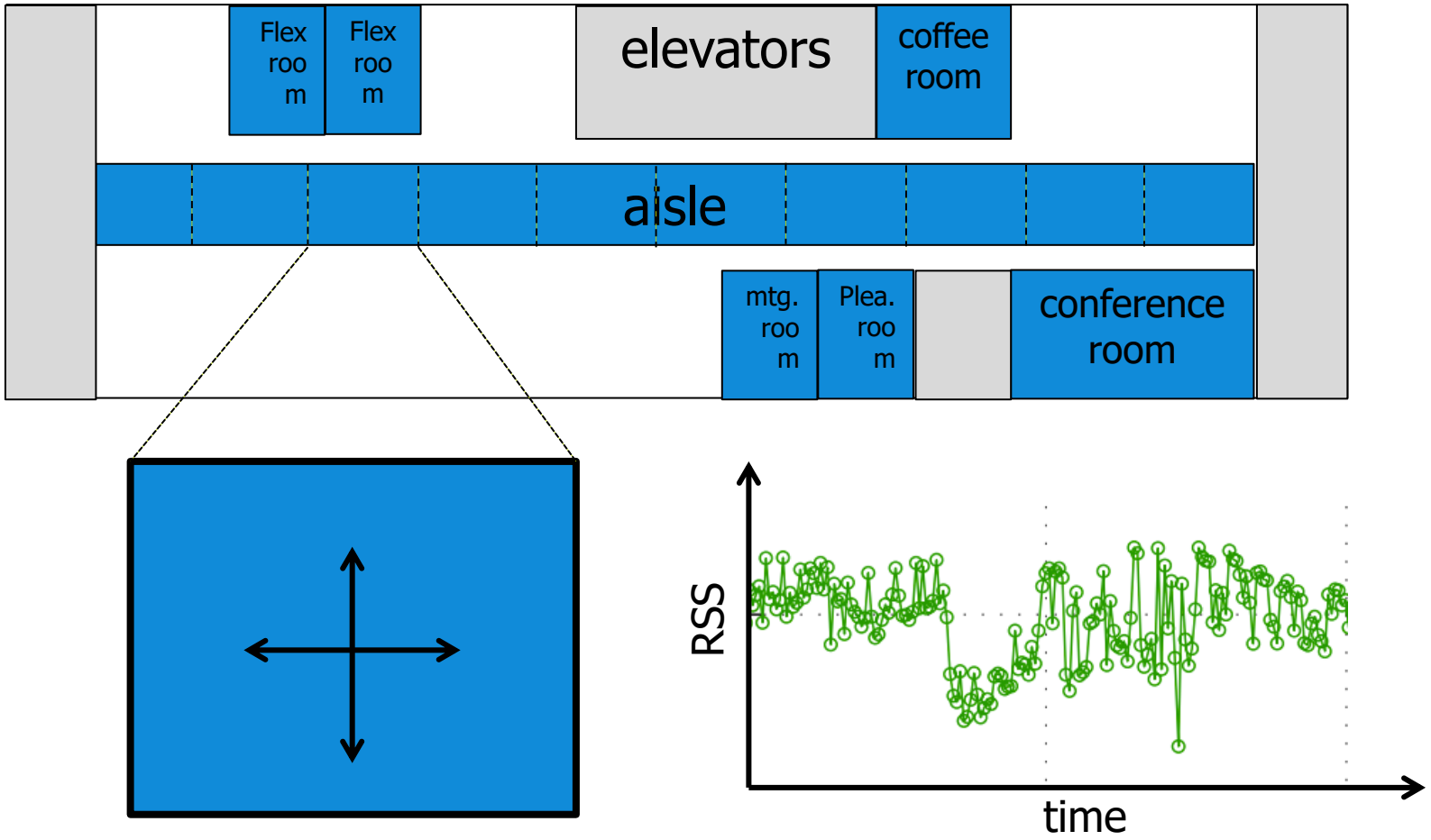
http://www.cs.umd.edu/~moustafa/papers/horus_usenix.pdf

Paper2: Practical Robust Localization over Large-Scale 802.11 Wireless Networks

http://www.kavrakilab.rice.edu/sites/default/files/mobicom2004.pdf

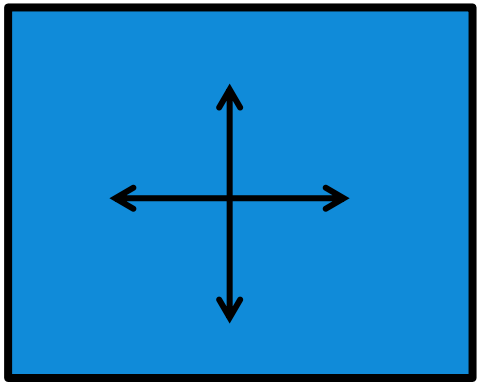
STEP BY STEP: BAYESIAN LOCALIZATION

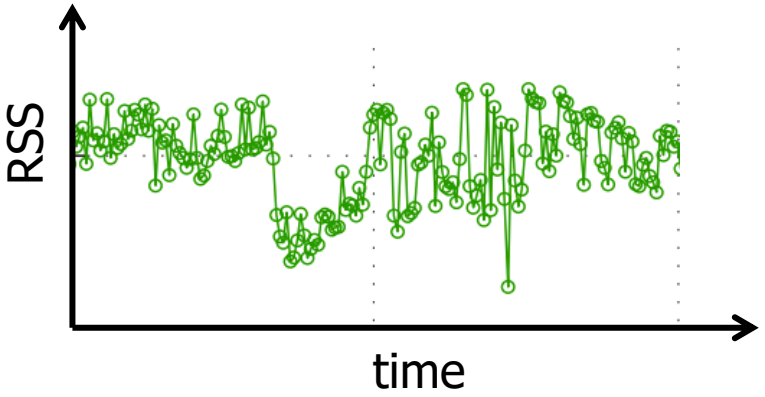
Step 1) Get RSS signals for each cell



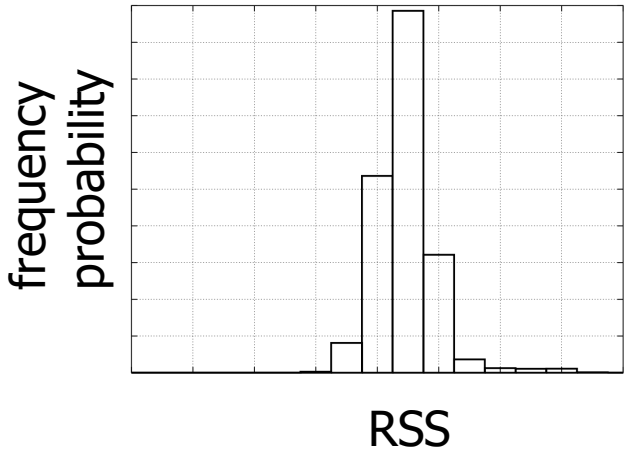
position yourself in the center of each cell to gather RSS data

Step 2) Process signal & get histogram

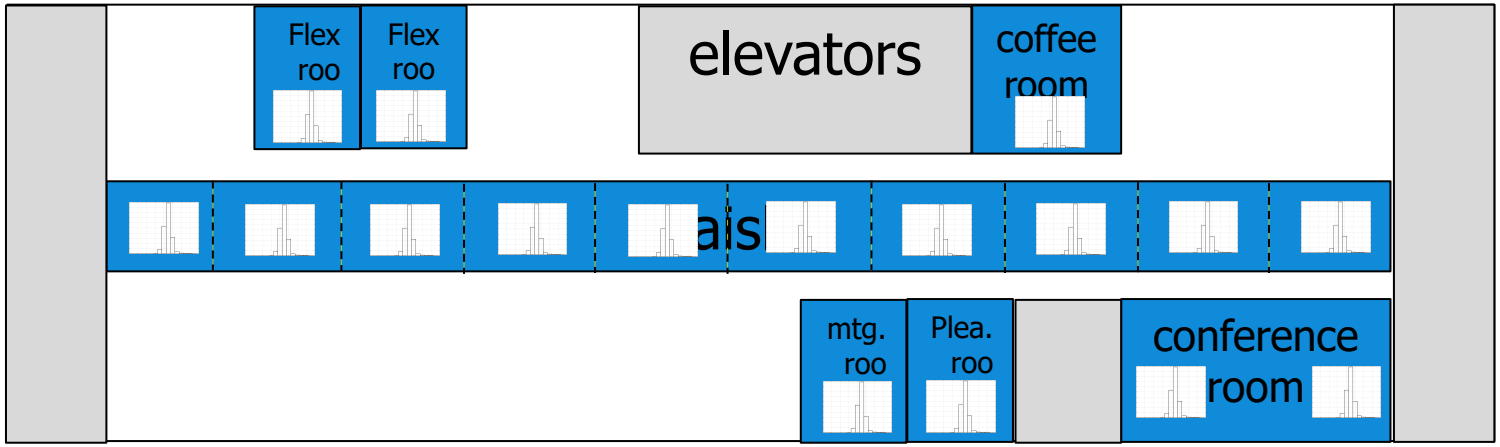




Sub-steps:
2.1) filter signal
2.2) get features (**RSS or avg RSS**)
2.3) get histogram
2.4) get pmf



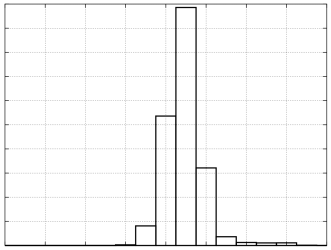
Radio Map



- Each cell has now a probability mass function (pmf)
- The higher the difference among pmf's the better
- Notice that the big room has two cells.

Step 3) Store data in phone

RSS values [0,255]



pmf of cells

	0	…					…	255
C1								
			0.1	0.2	0.5	0.2		
C17								

Σpi = 1

one table per access point

localization won't be trained on the spot

Considering the table below and given that a user measures an rss value of *r*, where is the user more likely to be?

RSS values [0,255]

pmf of cells

	…	r-2	r-1	r	r+1	r+2	r+3	…
C1		0.3	0.4	0.2	0.1			
C2		0.1	0.2	0.4	0.2	0.1		
C3			0.5		0.5			
C4			0.2	0.6	0.2			
C5			0.1	0.2	0.5	0.2		
C6				0.3	0.4	0.3		
C7			1.0					

$\Sigma pi = 1$

one table per access point

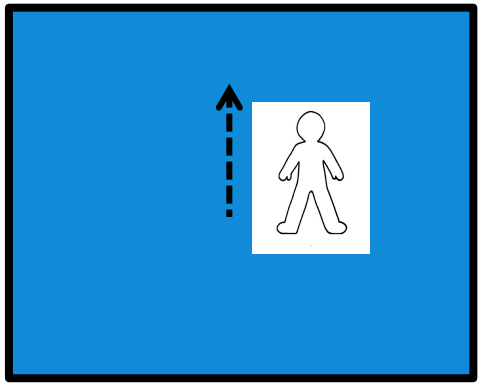
localization won't be trained on the spot

We have the radio map in the phone …

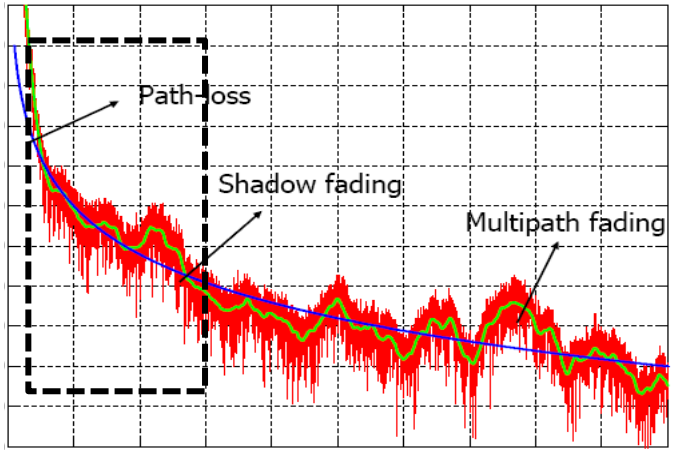
now, let's localize

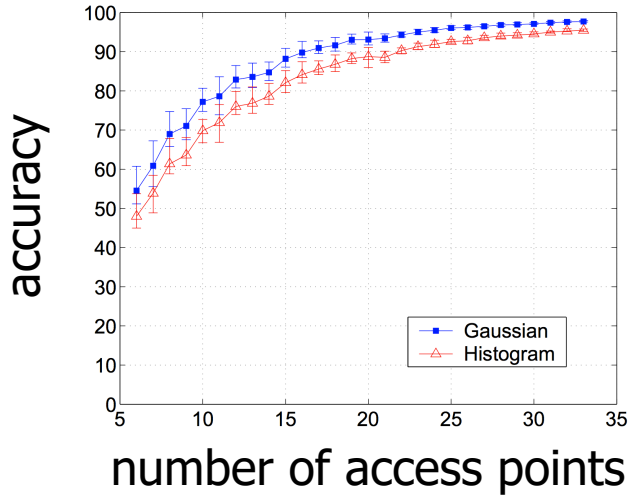
Step 4) Get testing data

- Start with initial belief
- Do WiFi scan
- Sort Access Points in decreasing order based on RSS
- Choose highest RSS, then second, etc



We will test only one direction

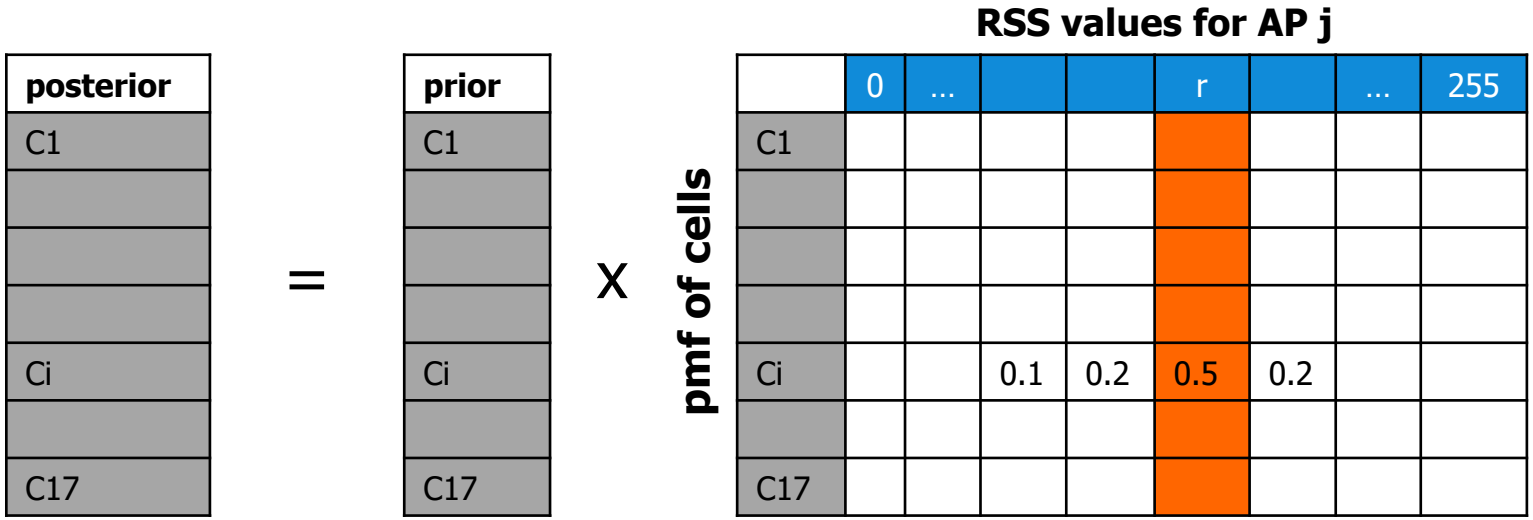




Step 5) Apply Bayes

probability I am in cell i given that I got an RSS measurement r from access point j:

$$p(cell_i / rss_j^r) = p(cell_i)\ p(rss_j^r / cell_i) \ / \ p(rss_j^r)$$



don't forget to normalize!

Step 6) When to stop iterations?

Step 6) When to stop iterations?

- No clear answer
- At every step you can update *prior* with
  - data from other access points
  - new scans
- Stop when you
  - pass a threshold: say 0.95
  - reach `steady state': oscillation around a max p

Step 7) Motion model

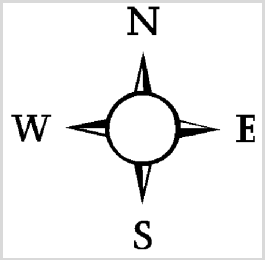
- Differentiate between idle and walking (we won't run :)
- Count number of steps s. Two options:
  - Preferred: TYPE_STEP_COUNTER (API 19 and higher)
  - Autocorrelation/Fourier transform (will be taught in class)
- For stride/time length assume
  - stride length (s) = height x 0.4 + noise
- Hence given location x, the new location distribution at each step is:

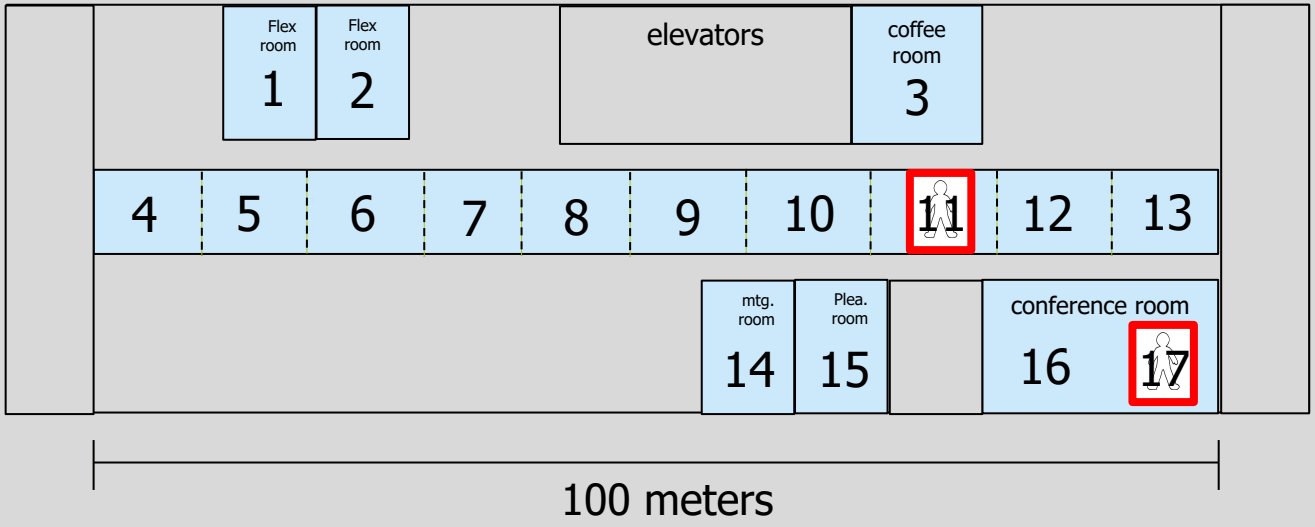
$$[(x-s)+noise,\ (x-s)-noise] \ \cup \ [(x+s)-noise,\ (x+s)+noise]$$

- Use compass to identify exact direction. Two options
  - Preferred: TYPE_ROTATION_VECTOR
  - Sensor fusion (will be taught in class)

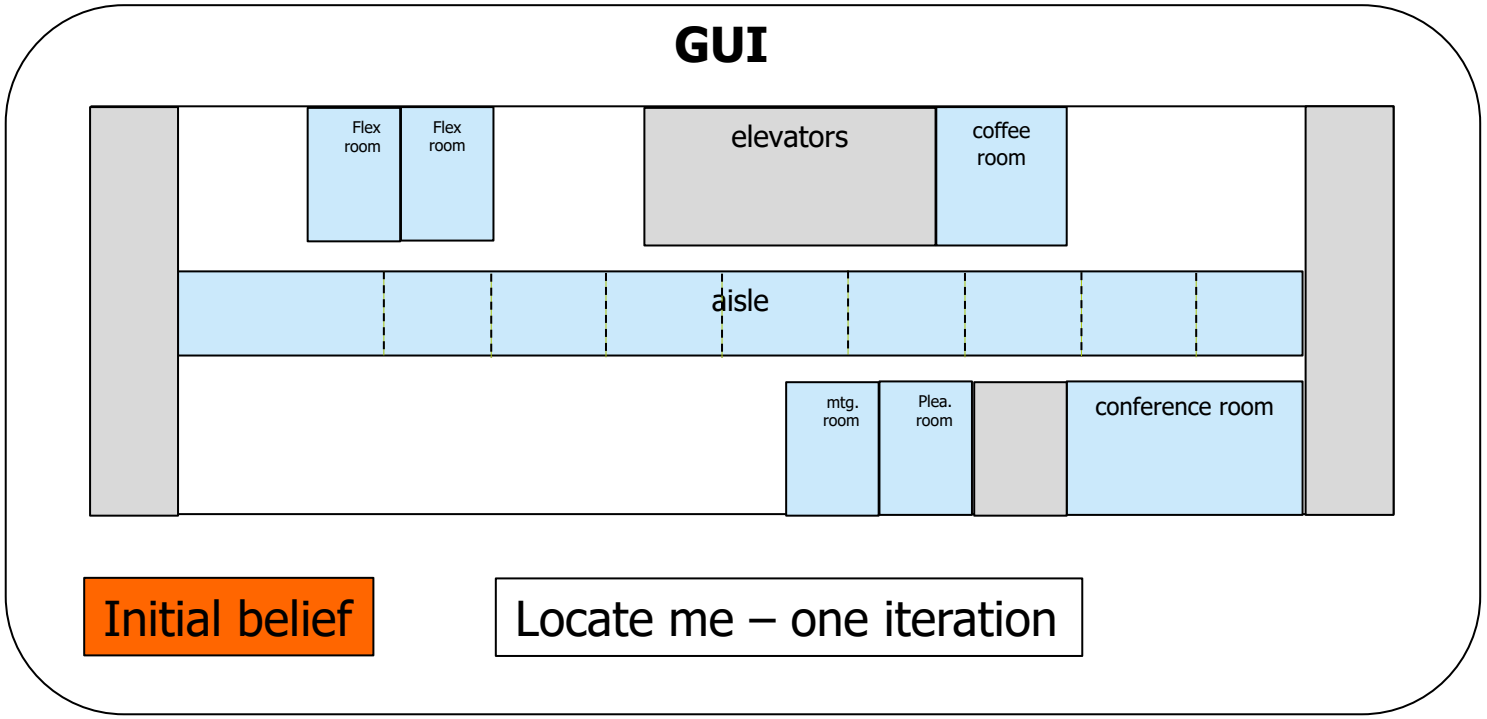
Advantages of motion model

A user has equal probability of being in cells 11 and 17 (p=0.5 each). The user walks 40 meters West, which cell or cells have the highest probability of being the correct location of the user after the 40 meter walk? Why?

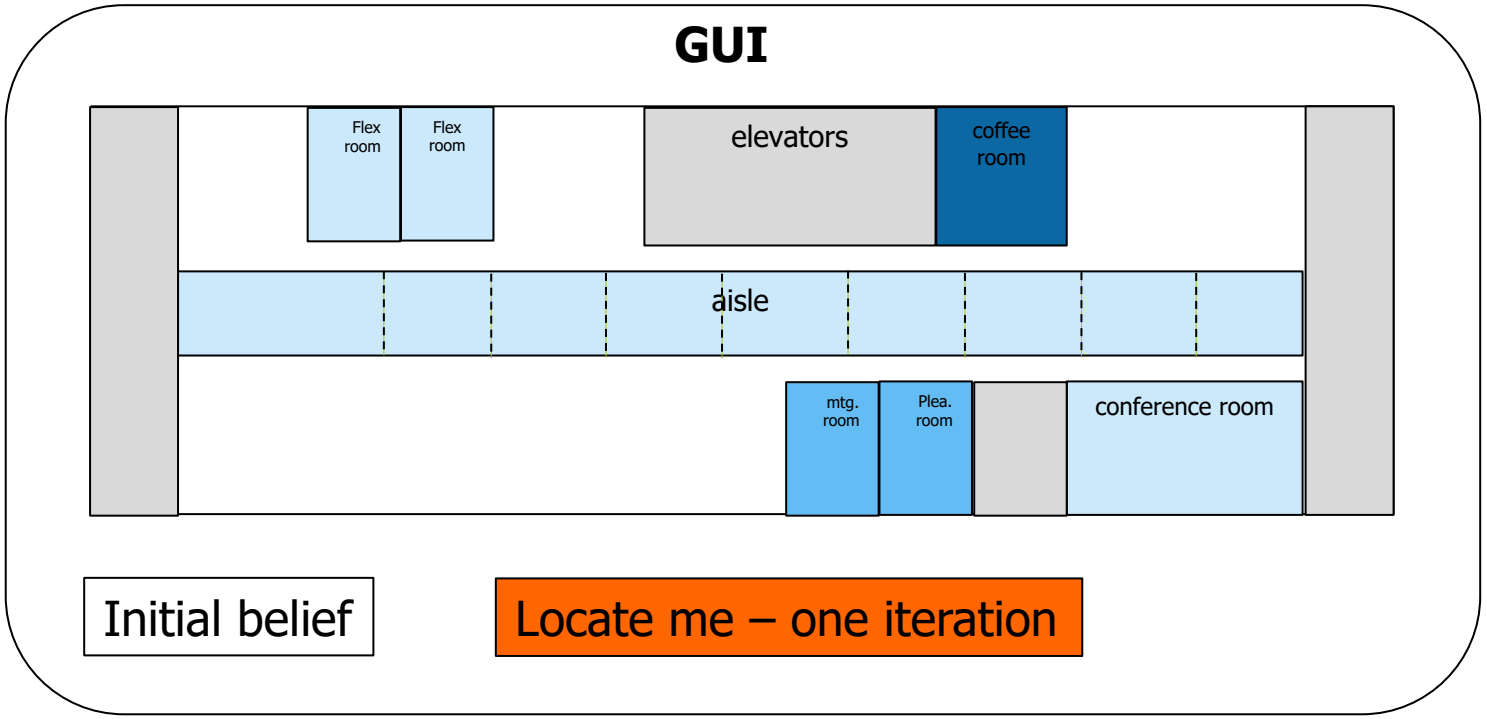




App: remember no training on the spot

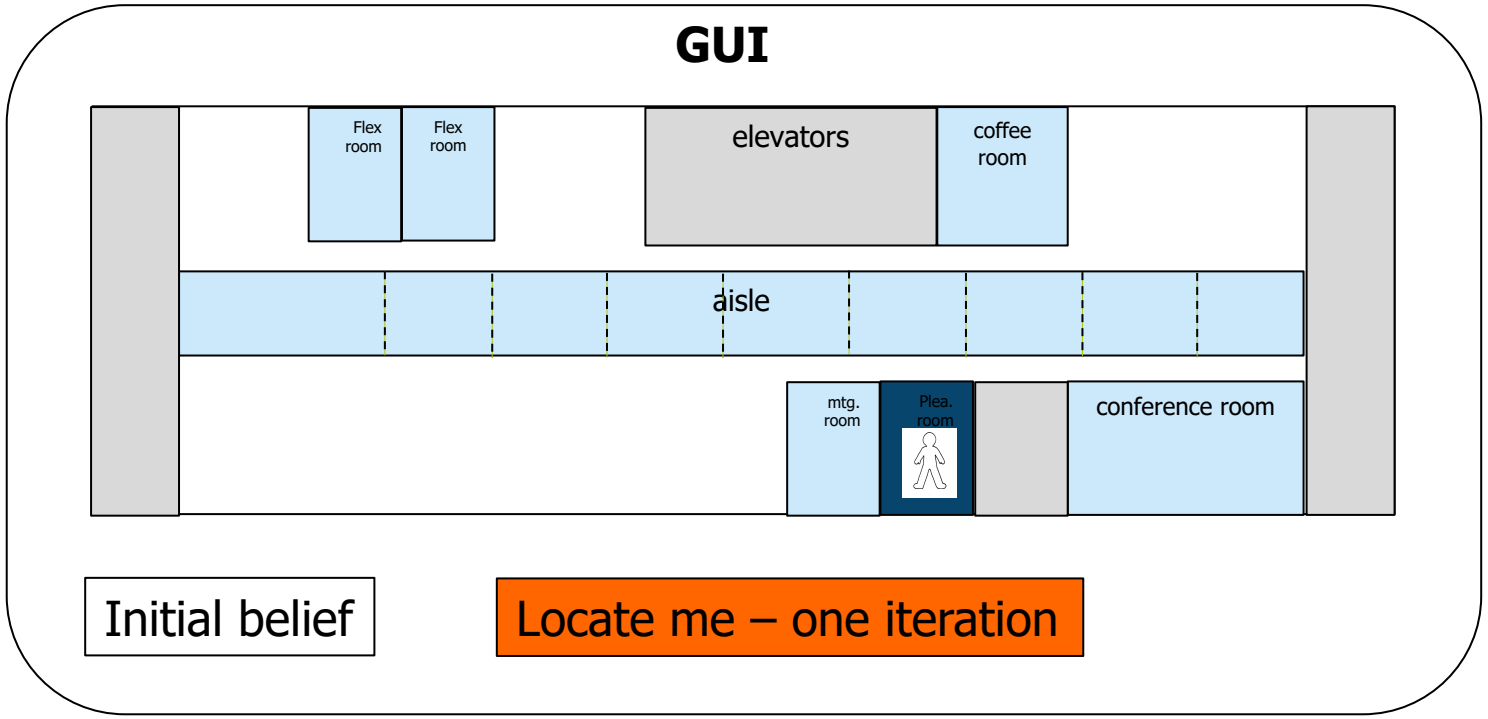


App: update belief



we will push "Locate me" until the app converges

App: update belief



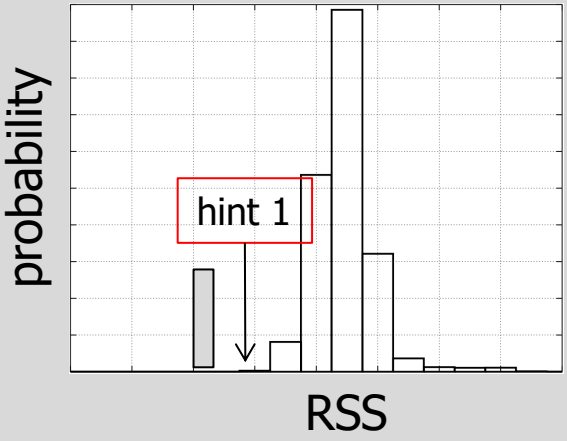
Off line processing

- Off line processing can be very useful to have a good understanding of a method.

- Gather data for each cell, say 3 minutes per cell.
- Divide samples at random in a training and testing sets.
- Build Bayesian filter in Java/Python/Matlab with training set and check accuracy with testing set.
- Gather data at different days (multipath fading)

NEAT TRICKS FROM PRIOR WORK

Modeling radio maps



RSS values for AP j

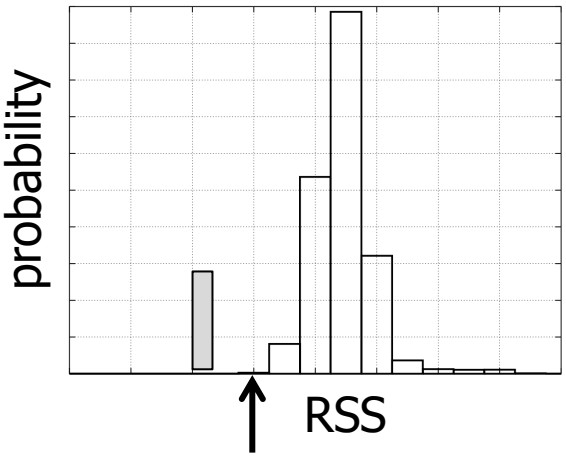
pmf of cells	0	...			r		...	255
C1								
Ci			0.1	0.2	0.5	0.2		
C17								

We suggested to use tables to store radio maps. But tables have limitations. Which ones?

hint 1: empty space cause by gray bar in pmf

hint 2: you may detect hundreds of Access Points at each scan.

Modeling radio maps



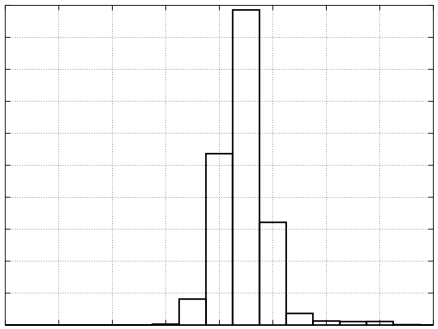
RSS values for AP j

pmf of cells

	0	...			r		...	255
C1								
Ci			0.1	0.2	0.5	0.2		
C17								

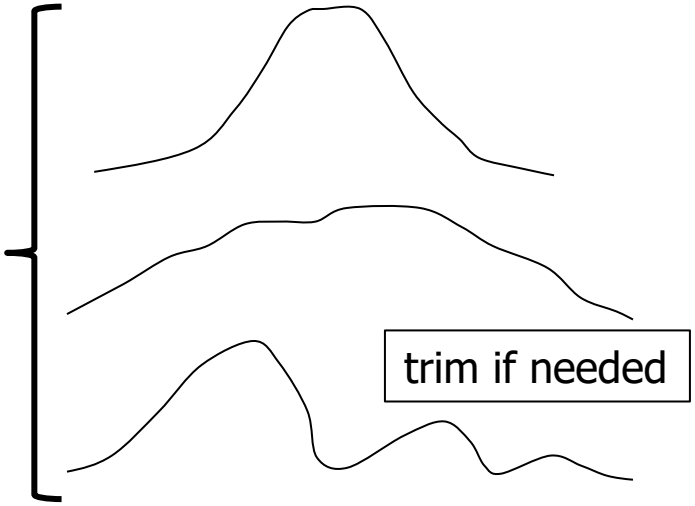
- We suggested to use tables to store radio map
- But tables have limitations
  - Sampling granularity
    - What if you get an RSS value that is not present?
  - Memory space
    - If you have N access points, you need 255xCxN elements

Modeling radio maps with Gaussians



This looks like a Normal/Gaussian distribution

use Gaussians for all cases

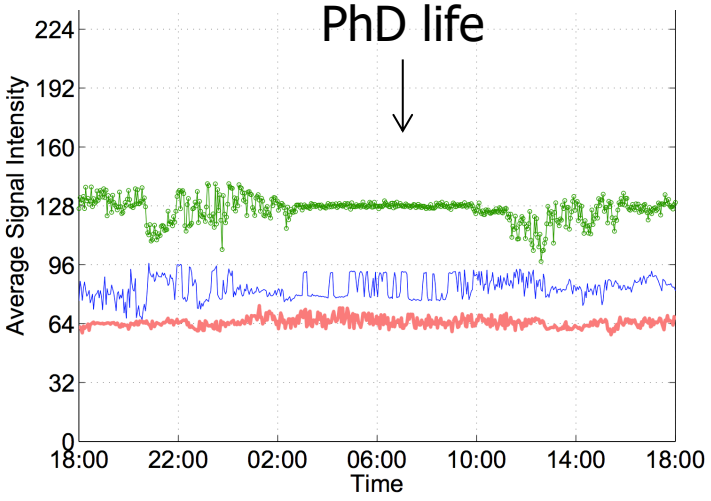


- Less memory space: 2xCxN (2 from 255)
- More granularity: $f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Other students have used Gaussian Kernels and interpolation

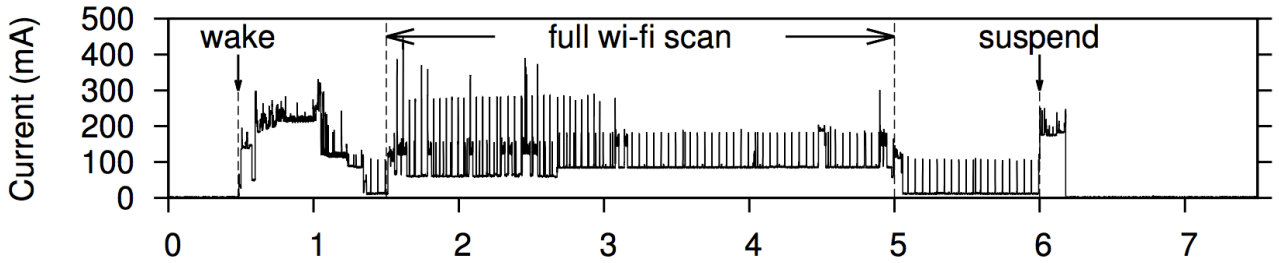
Modeling variance

- Check your testing phase at different days!
  - 2.4 GH-> 12.5 cm.
  - Multipath fading occur due to small changes
  - People moving, opening closing doors



Concrete sampling guidelines

- Horus (Bayesian grid)
  - 100 samples per cell spaced 300 ms, cell size:1.5m or 2.0 m
  - Error 2 meters
- Practical Robust Localization (Bayesian room)
  - 60 seconds sampling per office, cell size: 2.5 x 5.0 meters (500 offices)
  - 95% accuracy with 2 or 3 RSS measurements
- Both methods provide few meters error for 1-minute sampling per cell.
- WiFi scans take a few seconds
  - Dual band phones may bring new results. Extra: compare 2.4 and 5 GHZ bands?



Diminishing returns

- Work hard but smart in your localization app.
- More training and testing data does not necessarily mean much better results, but it may mean more work, time and use of resources (memory, processor)

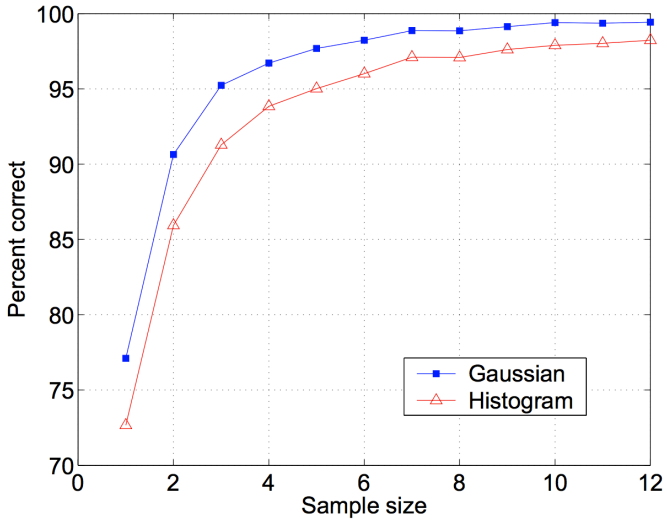


Figure 3: Bulk accuracy of localization methods after different numbers of observations.

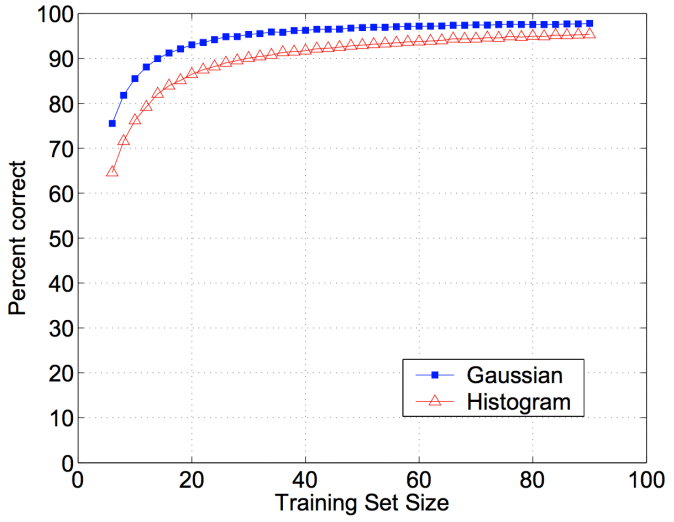


Figure 4: Training set size versus accuracy for the Gaussian and histogram methods.

Still major changes can happen on the day of evaluation affecting wireless channels. Do not despair! We will consider several attempts and external causes.

But if you have your app working ahead of time, make an appointment earlier to test it. It will help all of us.