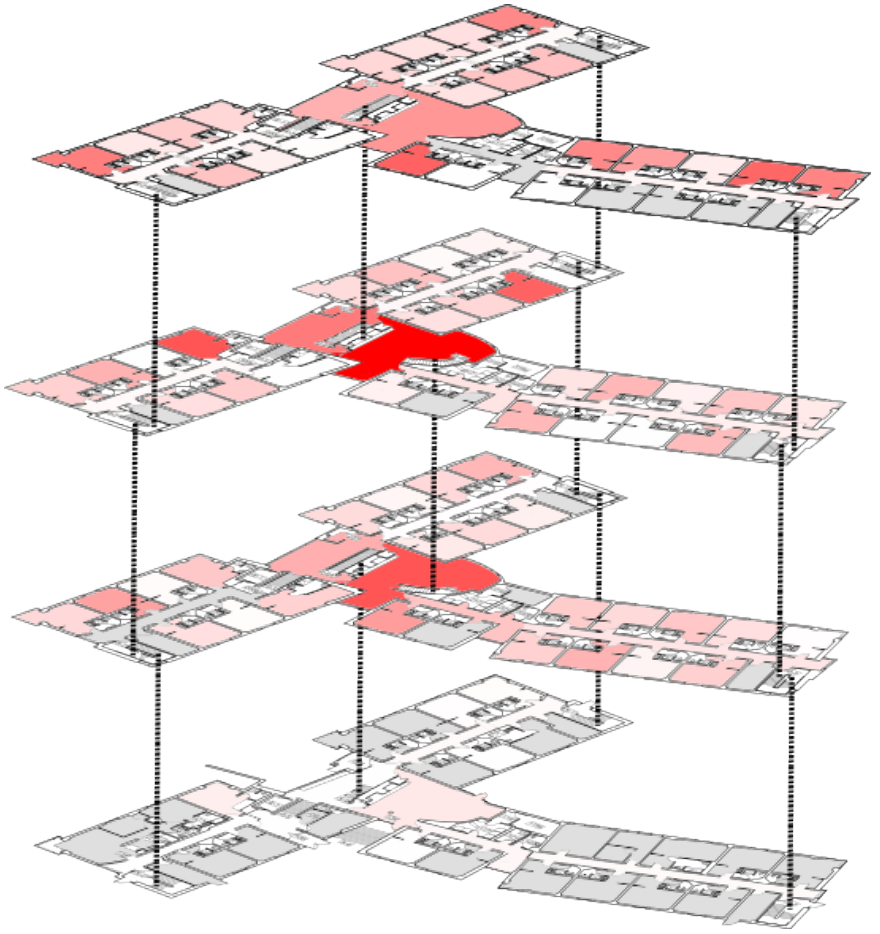
A Long Duration Study of User-Trained 802.11 Localization

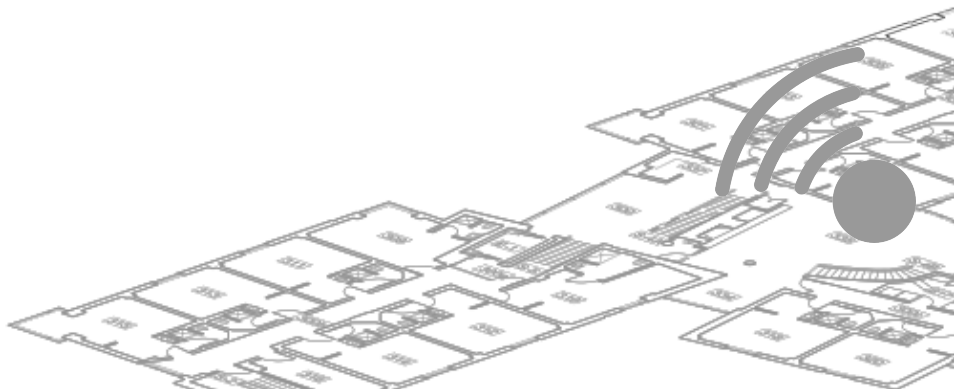
Andrew Barry
Olin College of Engineering



MELT 2009
Sept. 30, 2009

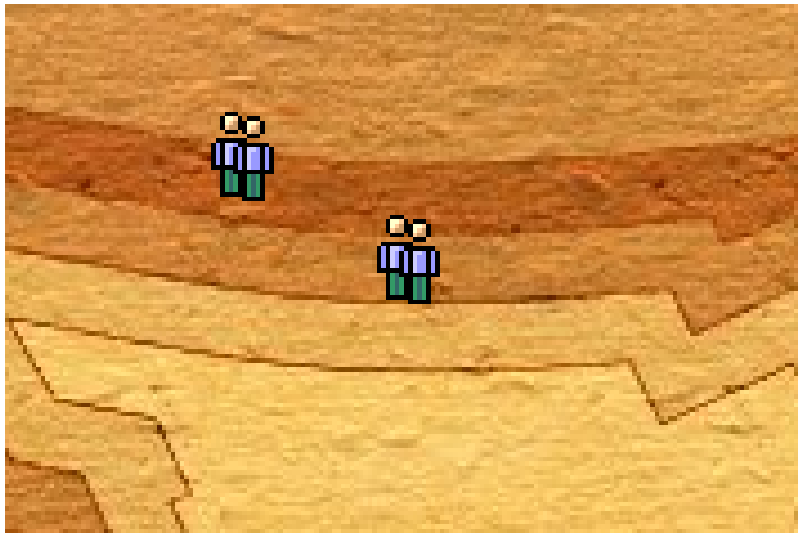
Overview

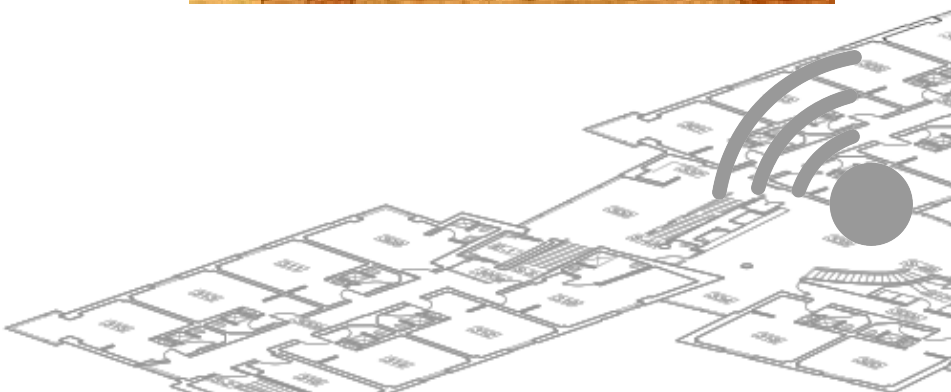
- Wireless localization with fixed 802.11 access points
- Euclidean distance—nearest neighbor in signal space (NNSS) algorithm.
- Running on personal laptops
- Crowdsourced data collection



Frontend Map Interface

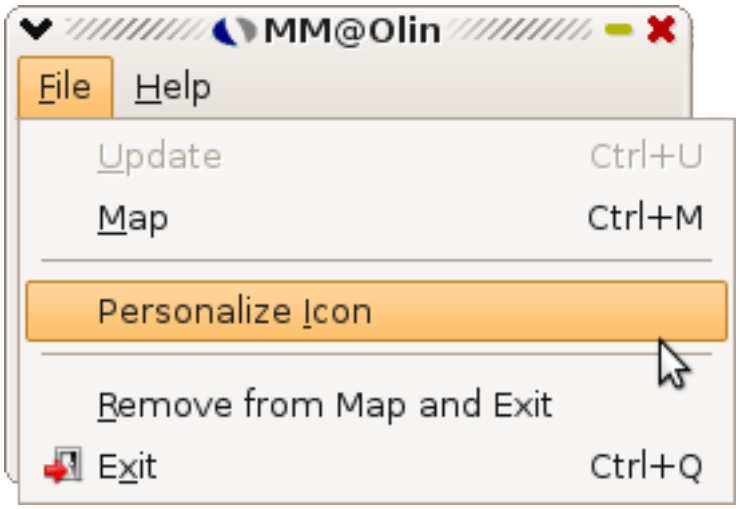
- Friend-finding search
- All active users on one screen
  - Repeated floor visualization shows vertical displacement

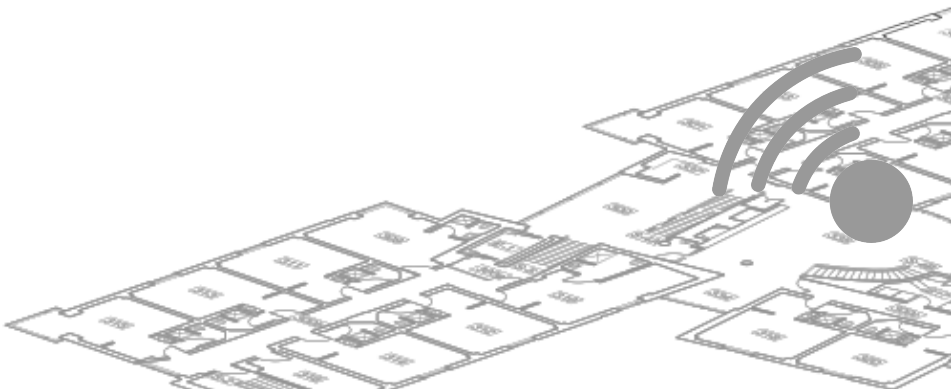




Frontend Map Interface

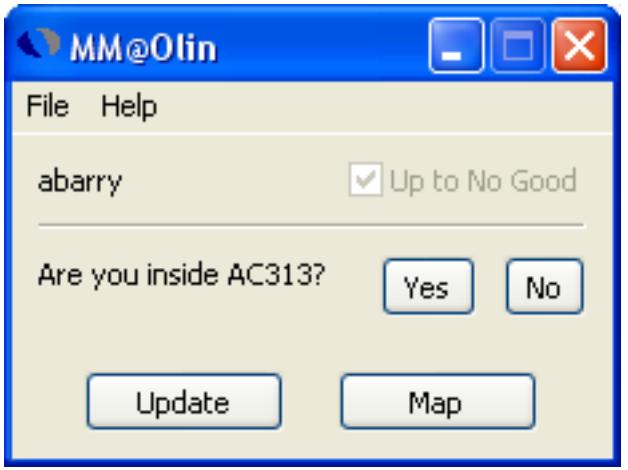
- Custom icons allows rapid user identification

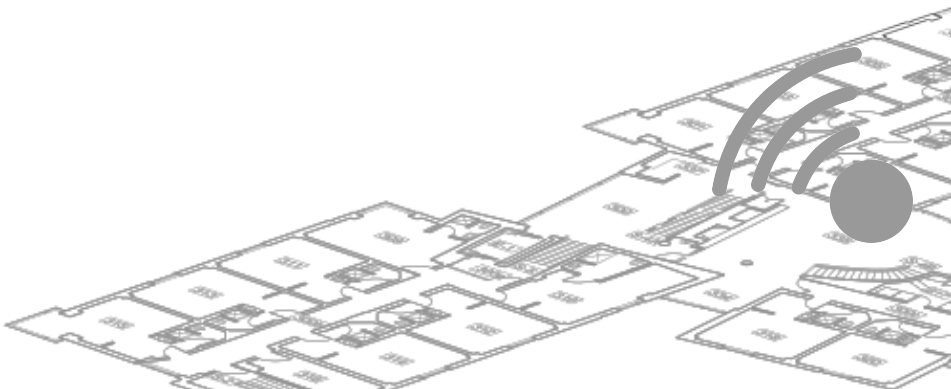




Client – Interface

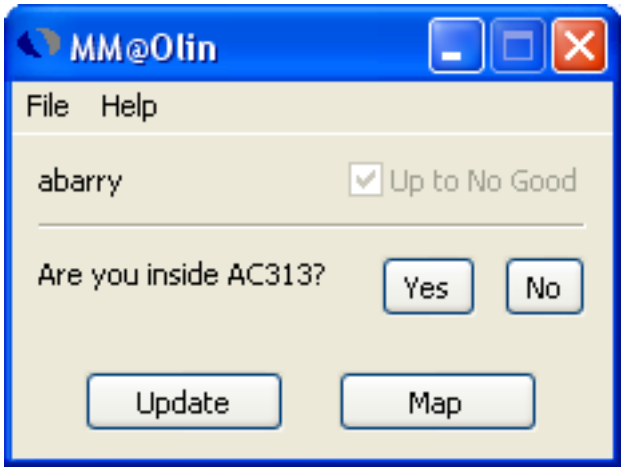
- Requirements
  - Non-intrusive
  - Intuitive
  - Lightweight
- Training support
  - Confirmation
  - New point creation

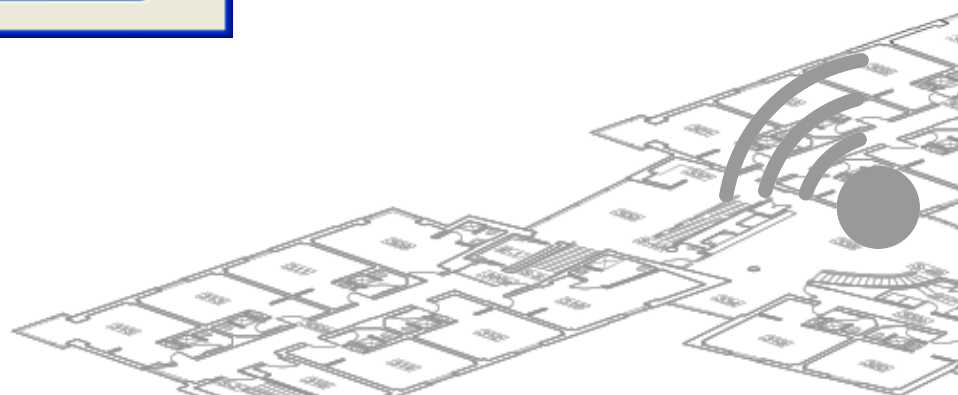




Client – Training

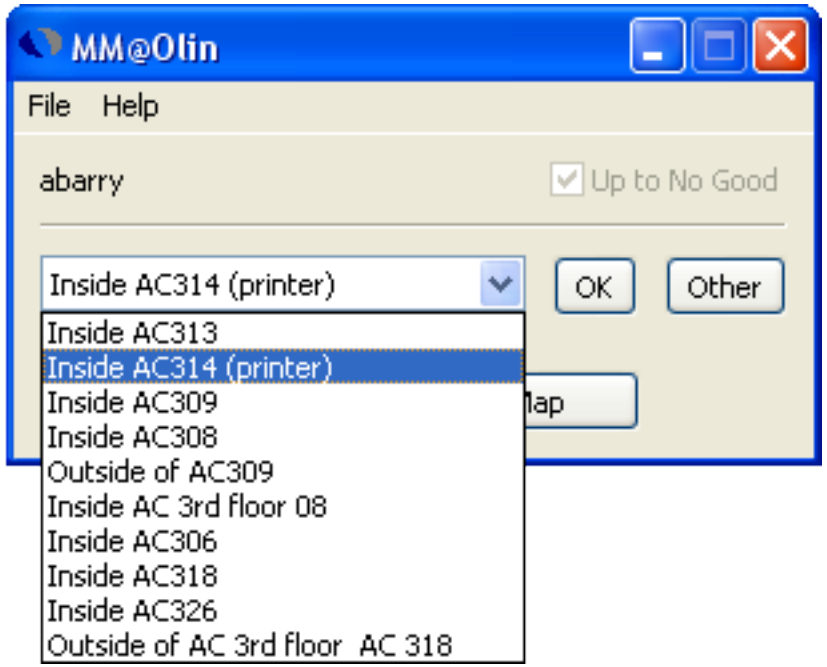
- Simple and easy
  - User is doing us a **favor**
- Start with a location estimate:

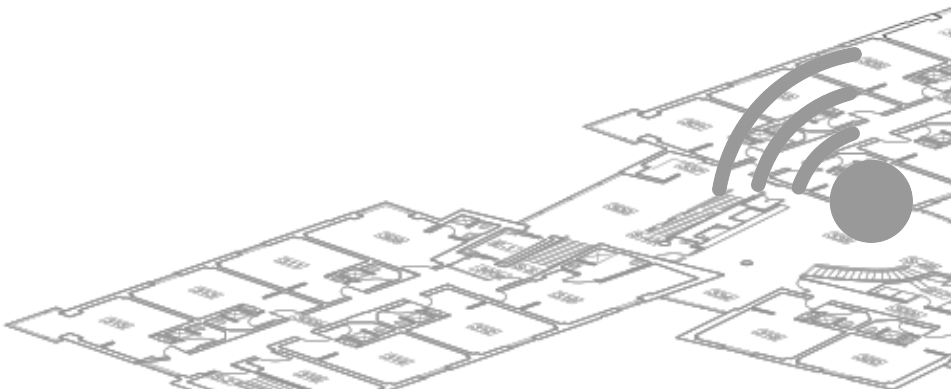




Client – Training

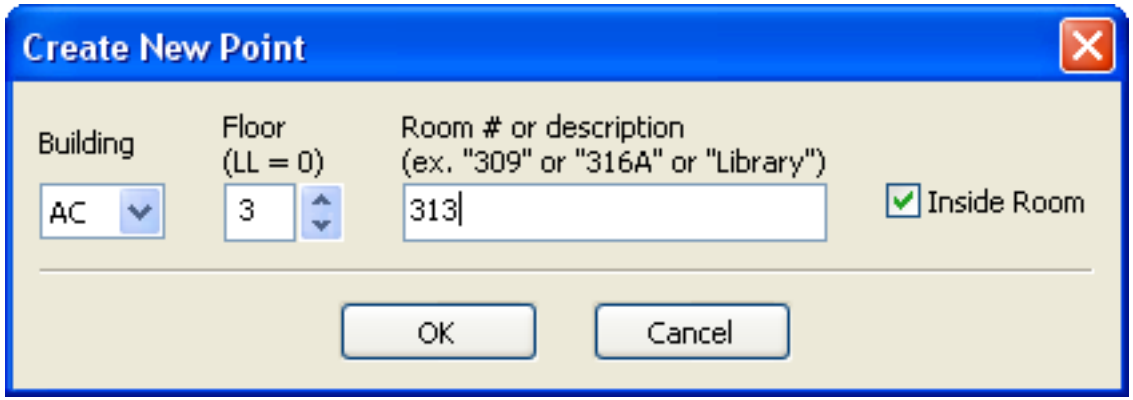
- If **Yes**, send data
- If **No**, offer nearby locations

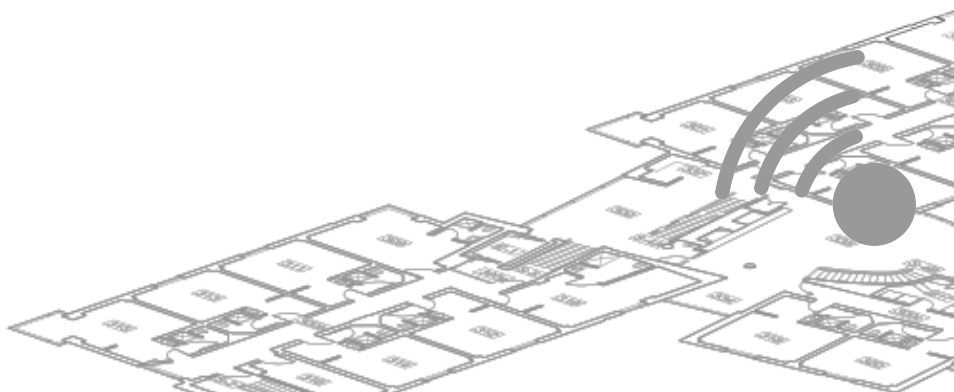




Client – Training

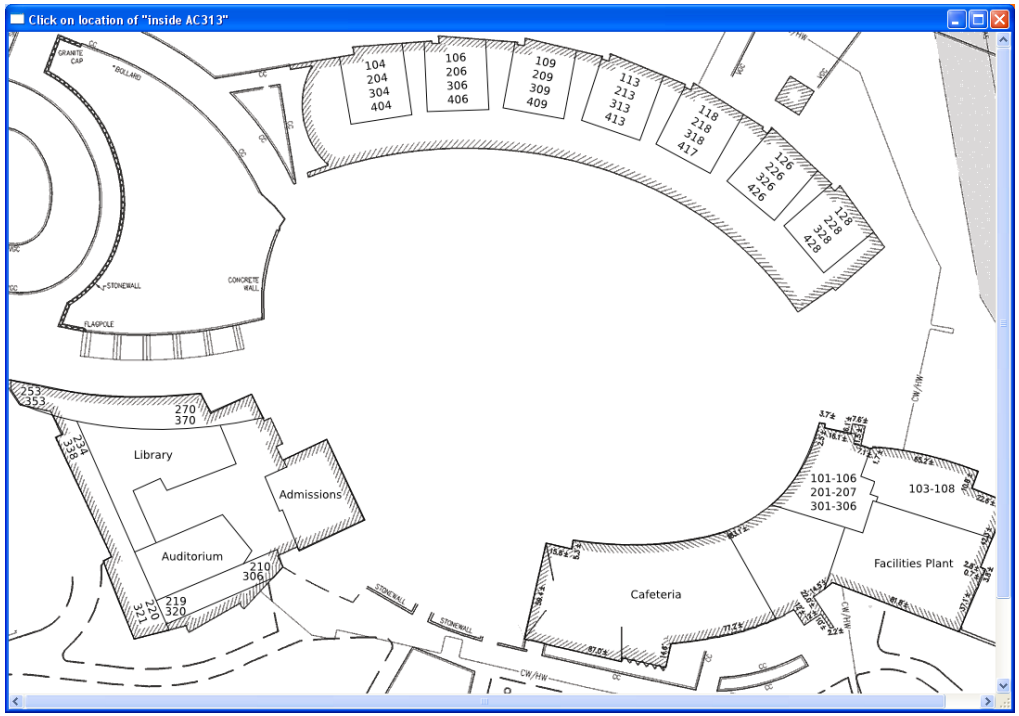
- If user finds location, send data
- If not, the current location is not in the system and should be added

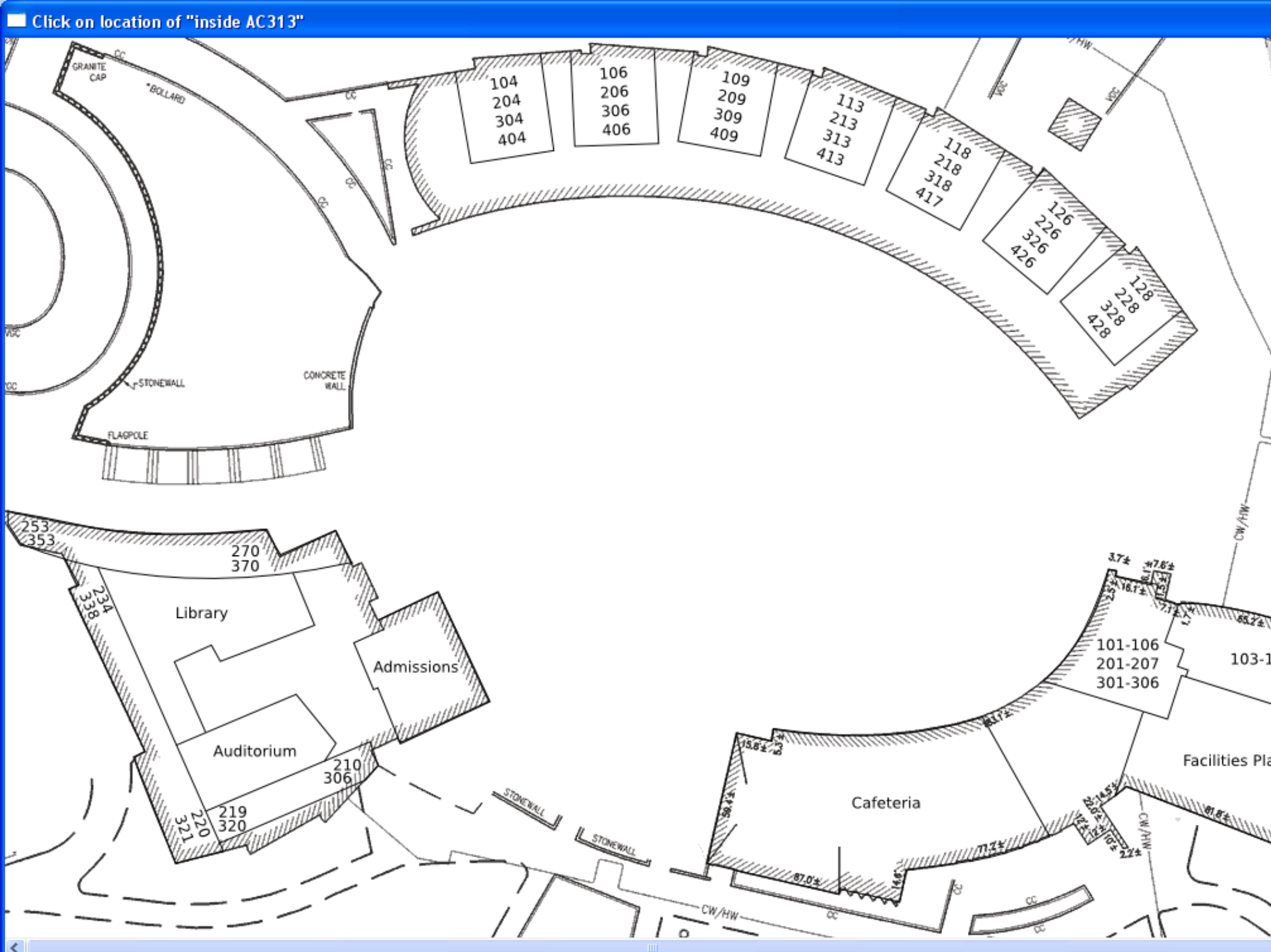




Client – Training

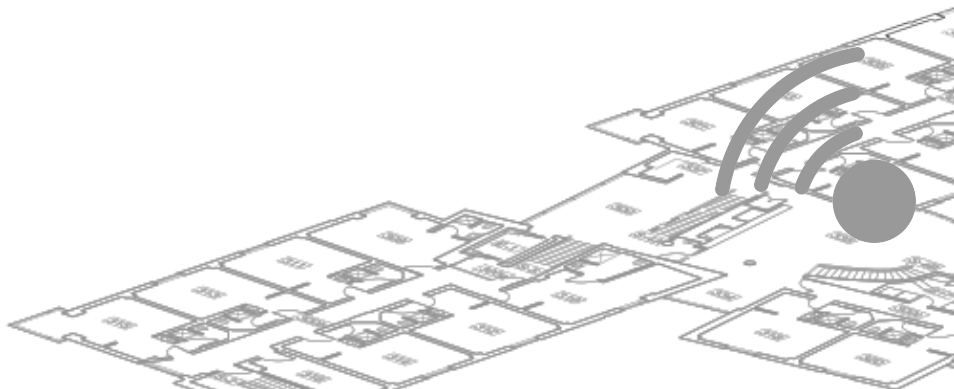
- New point creation
- Must be well labeled
  - Users do not know where they are on a floorplan




Click on location of "inside AC313"
GRANITE CAP
BOLLARD
STONEWALL
CONCRETE WALL
FLAGPOLE
104
204
304
404
106
206
306
406
109
209
309
409
113
213
313
413
118
218
318
417
126
226
326
426
128
228
328
428
253
353
270
370
234
338
Library
Admissions
Auditorium
210
306
219
320
220
321
STONEWALL
STONEWALL
CW/HW
Cafeteria
101-106
201-207
301-306
103-1
Facilities Pla
CW/HW

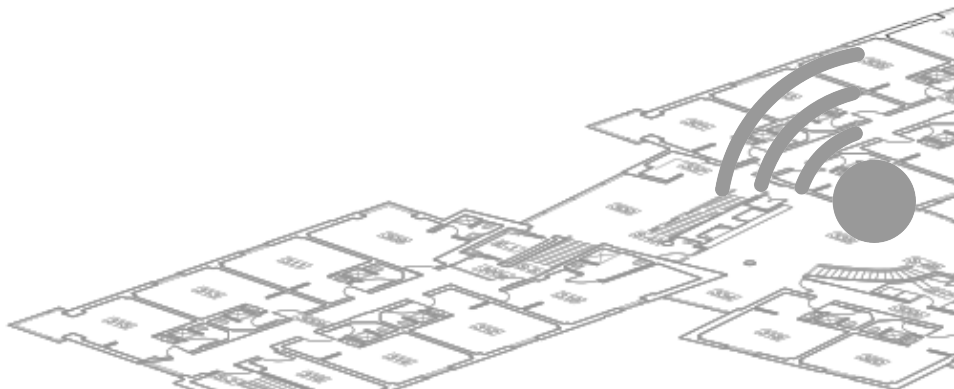
Localization Algorithm

- Euclidean distance in 76-dimensional space
- 76 = number of access points
- Similar to RADAR's Nearest Neighbor in Signal Space (NNSS) algorithm.



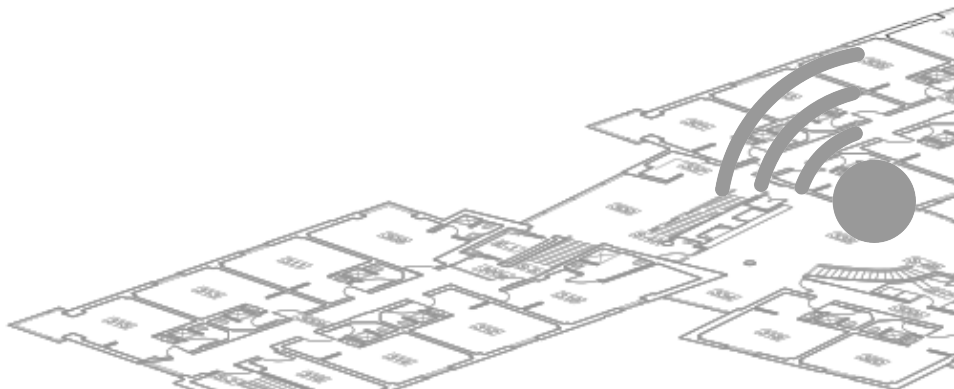
Training

- New fingerprint sent to the server
- Always append to known-location database
  - Duplicate points and confirmations added without regard to current database
  - Allows multiple training points for each location



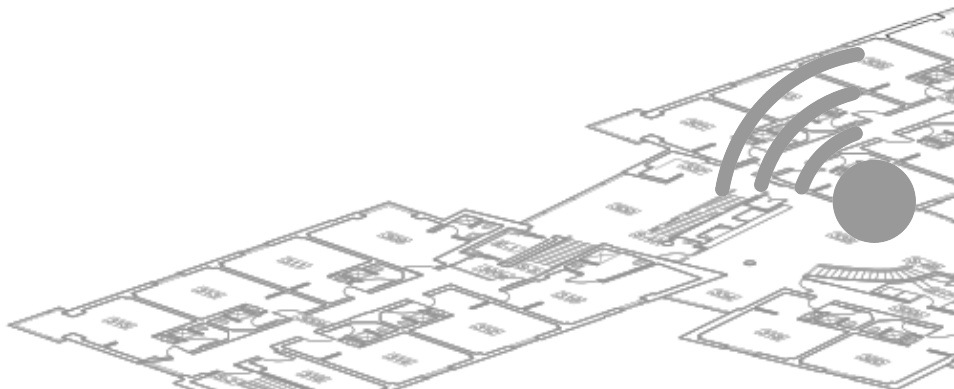
Initial Training

- 2 hours of training
  - About 200 points
- Convince users system works well enough to train it themselves
- 10-20 meter accuracy
- Initial set is now only 2.1% of location database



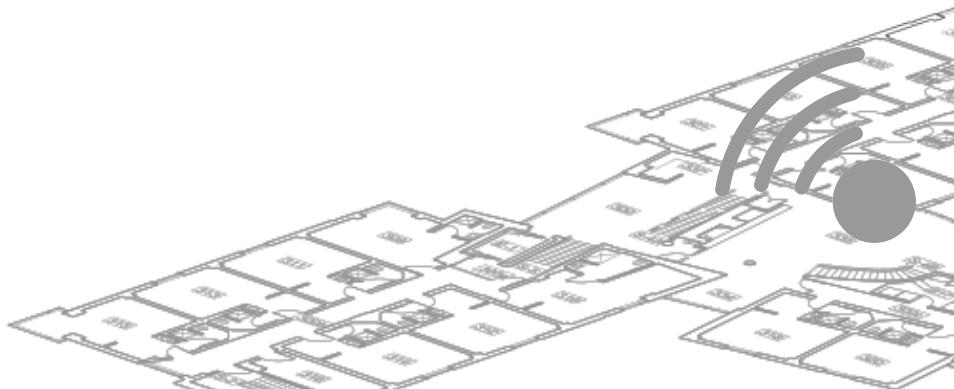
Deployment

- Launched in April 2008 at Olin College after short beta test
- Olin College
  - 300 students
  - 5 buildings enclosing 300,000 square feet
  - 76 wireless access points



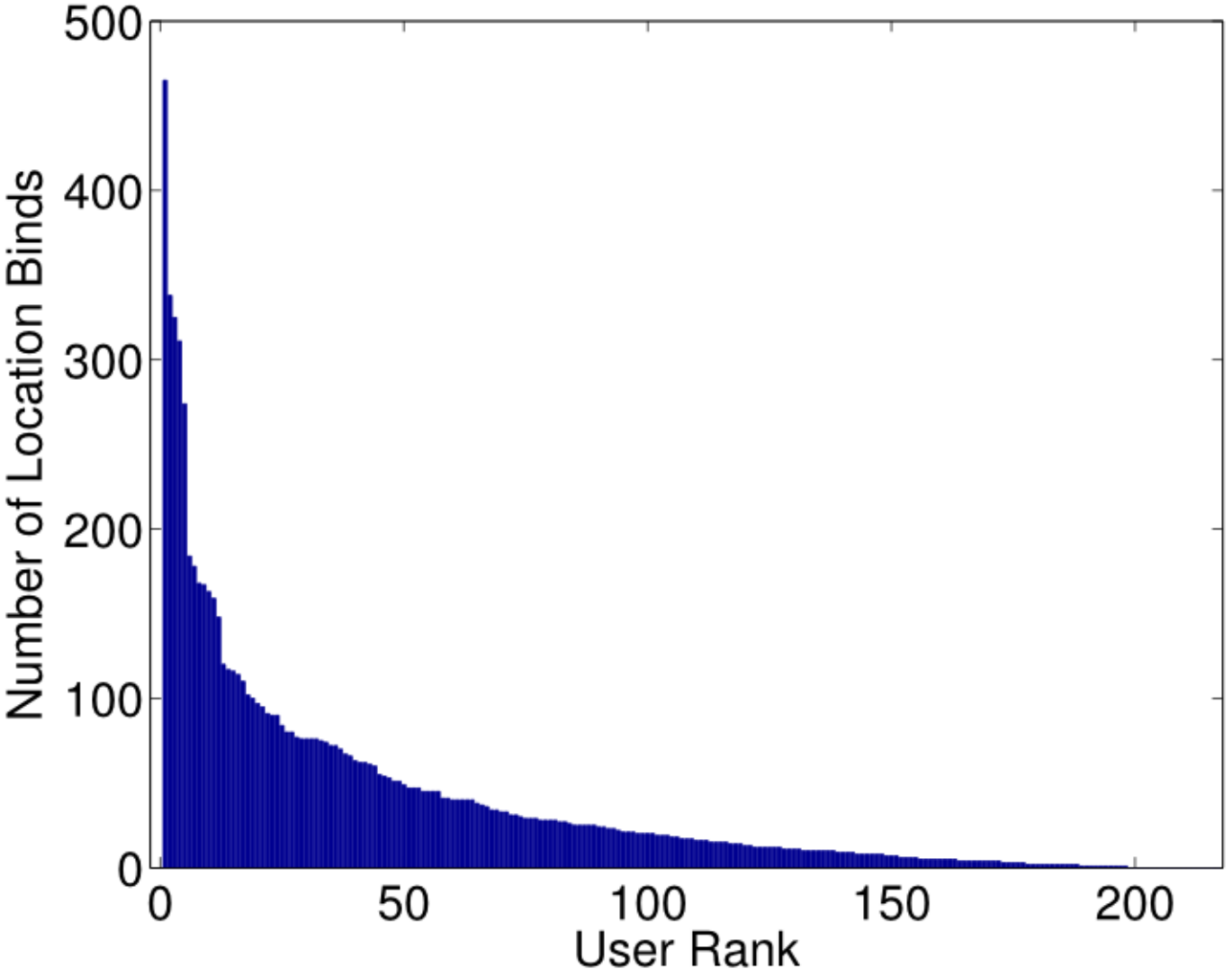
Deployment

- 200 total unique users
- Currently have approximately 100 active users
- 95% of users train the system
- Received 9,300 training updates
- Computed 1,000,000+ locations
- 14,000 friend-finding pages served



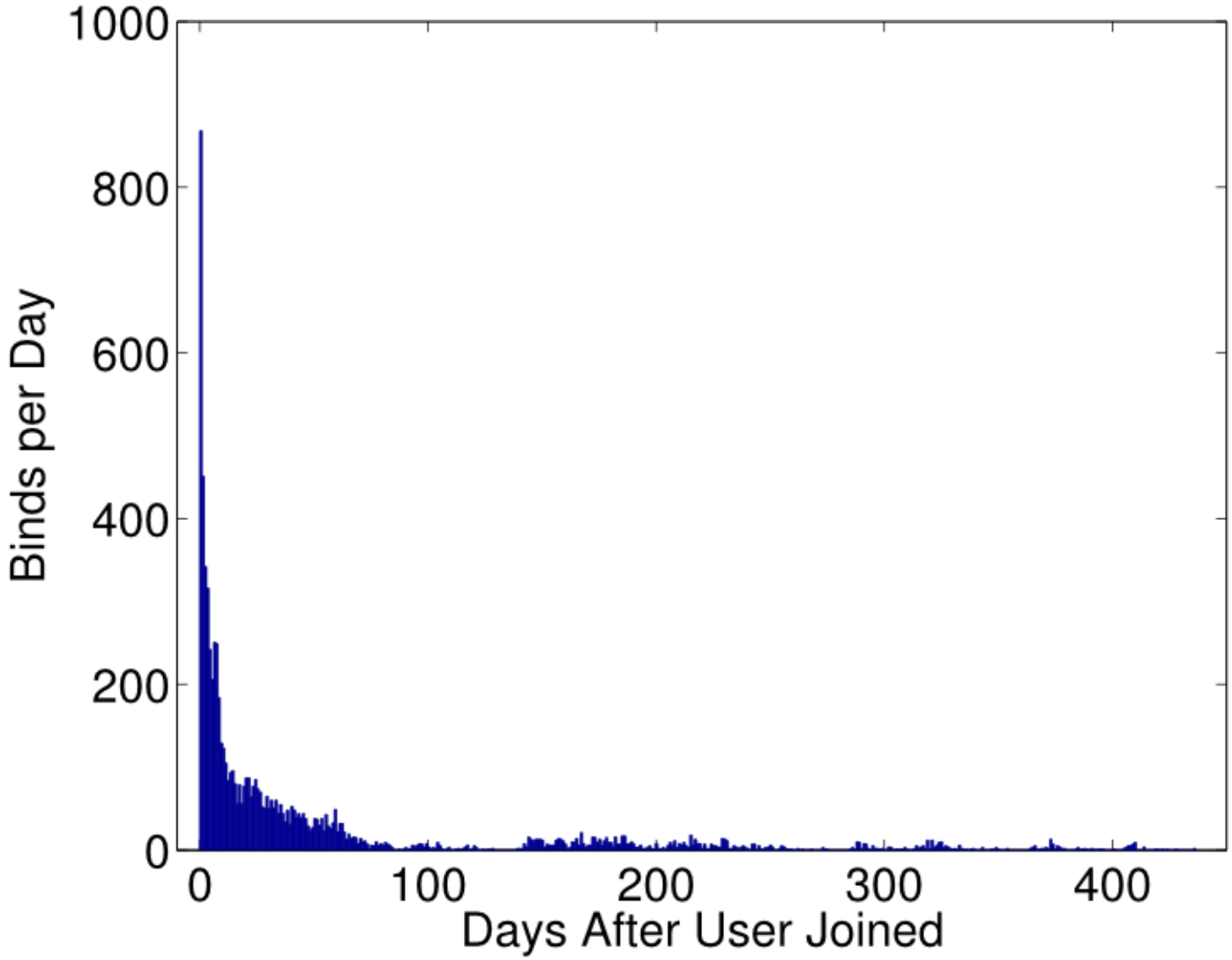
Who Trains?

- 20% of users bind two-thirds of the data



Who Trains?

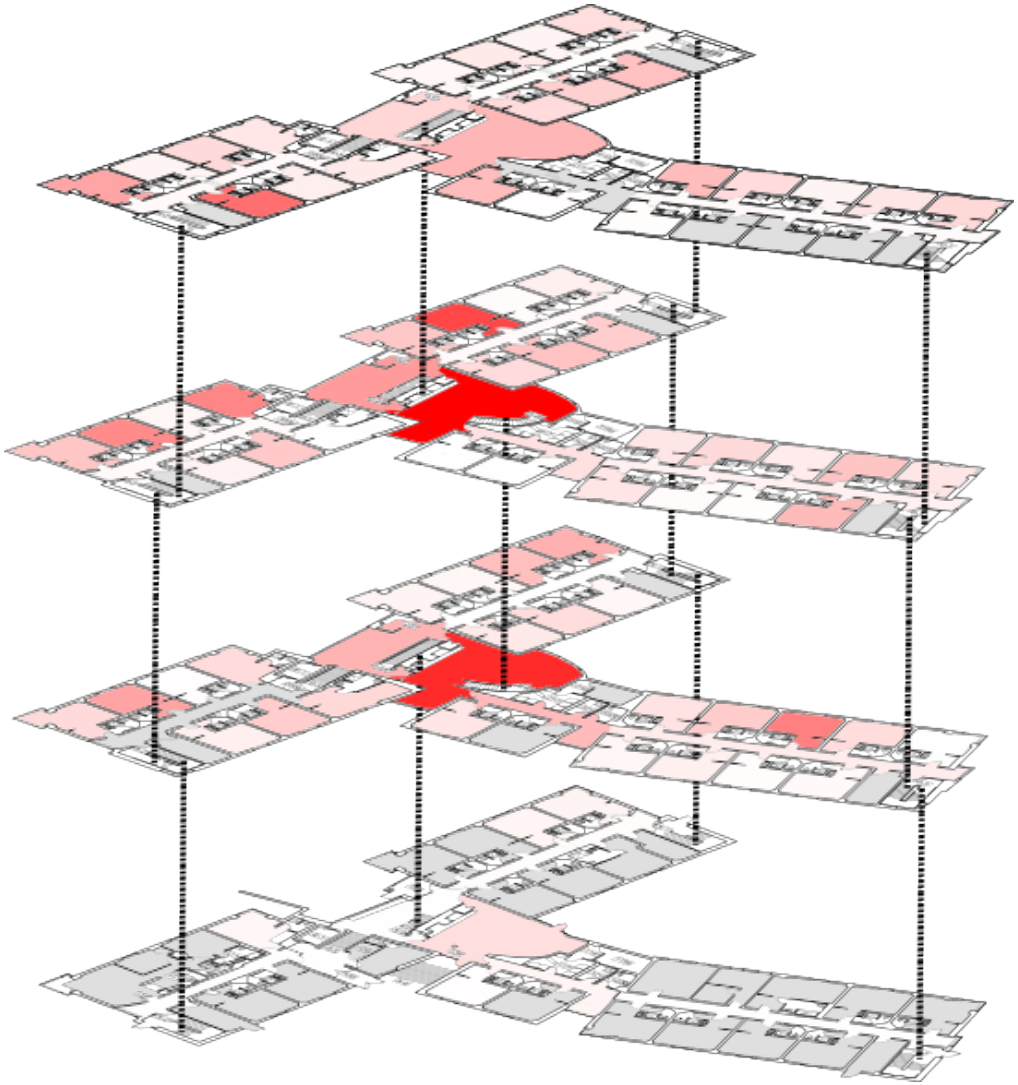
- Especially when those users are new



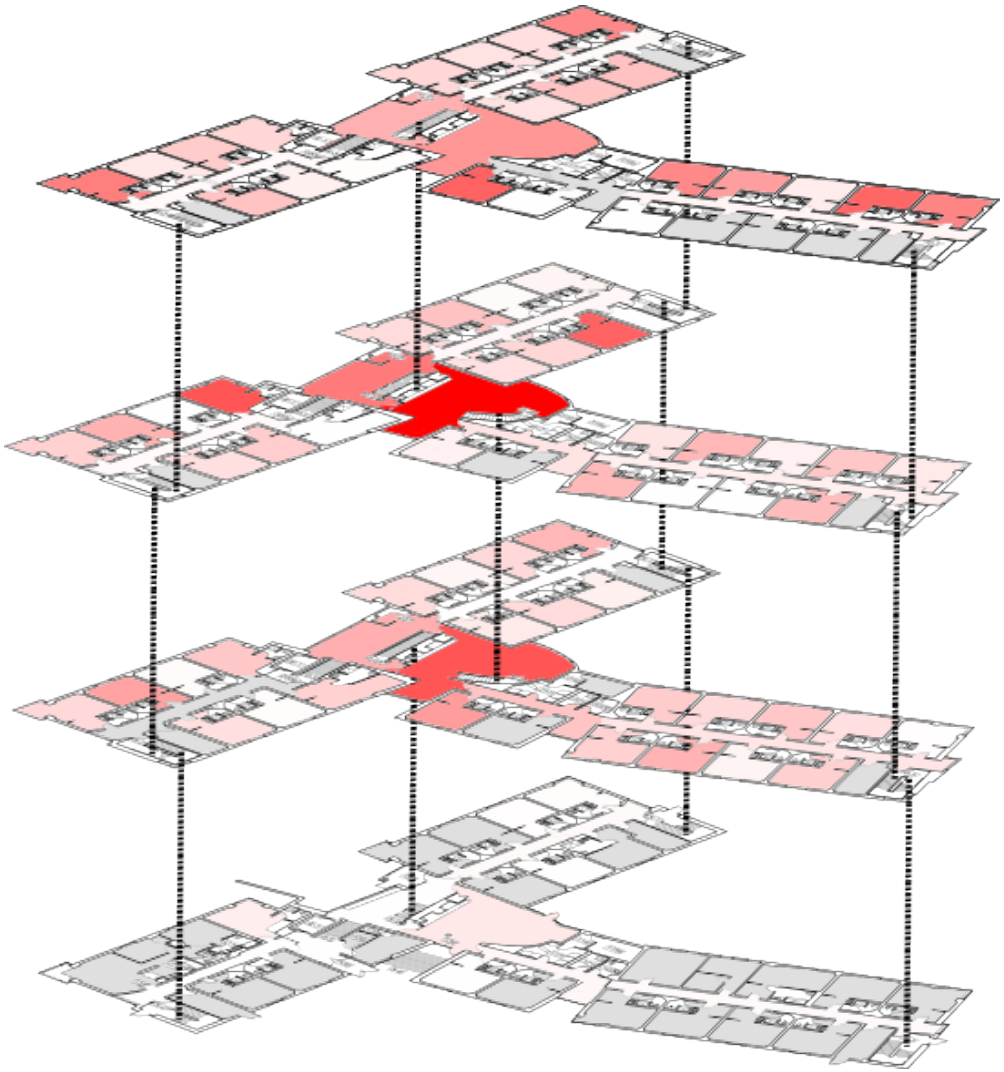
Where do they train?

- In the same places they localize
- 51% of all localization attempts are in areas where the localizing user has provided data

Where do they train?



Training Density (West Hall)



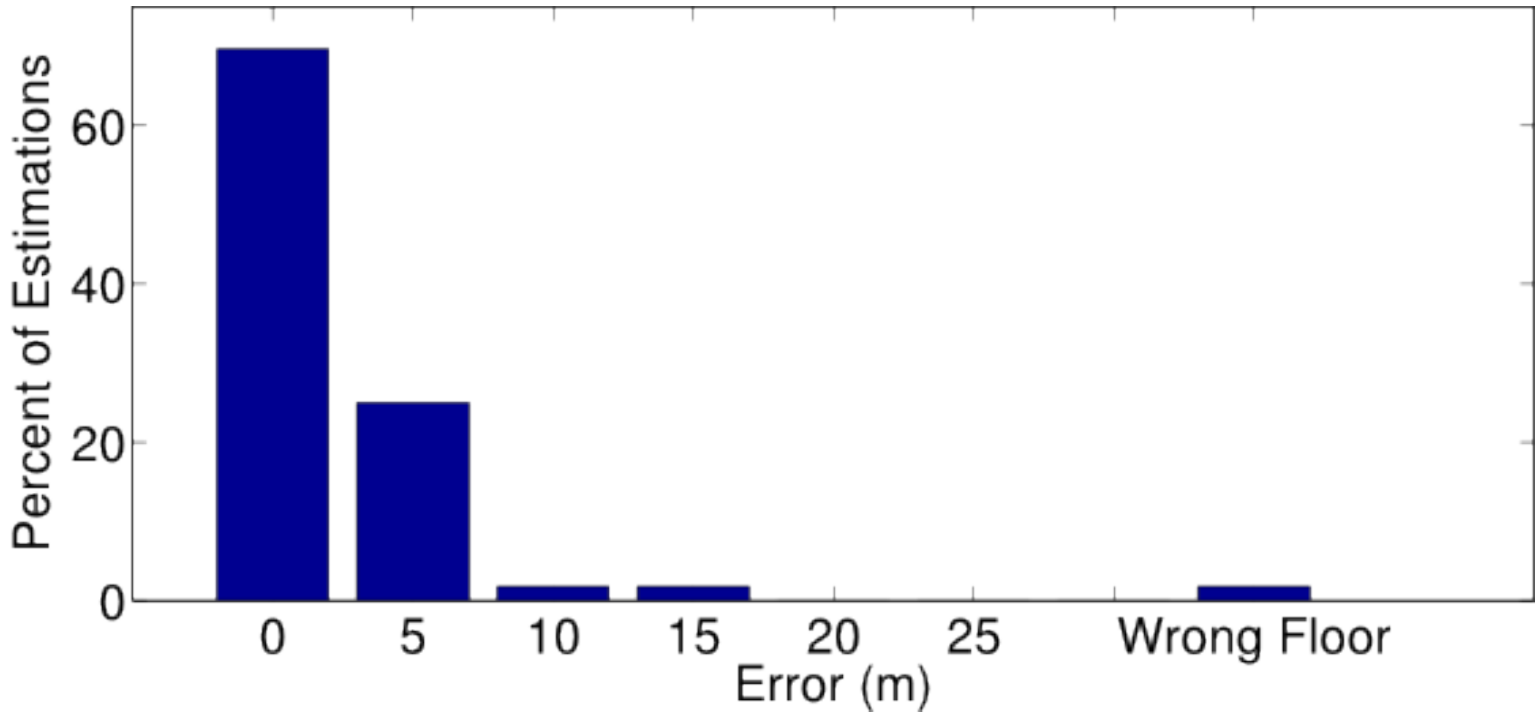
Localization Density (West Hall)

Accuracy

- True accuracy is not a random selection of rooms
  - Accuracy in a small trash-room is not important
  - How do we measure that?
    - Ask our users to test accuracy where they are at one moment

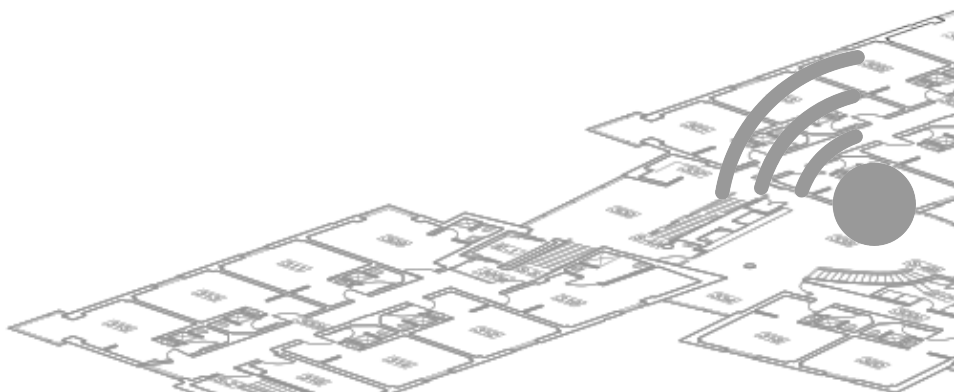
Accuracy

- Within 10 meters in 94% of cases



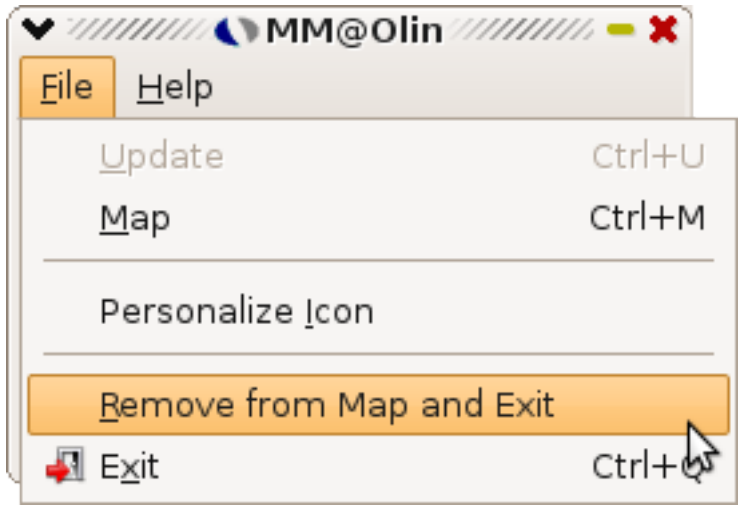
Errors

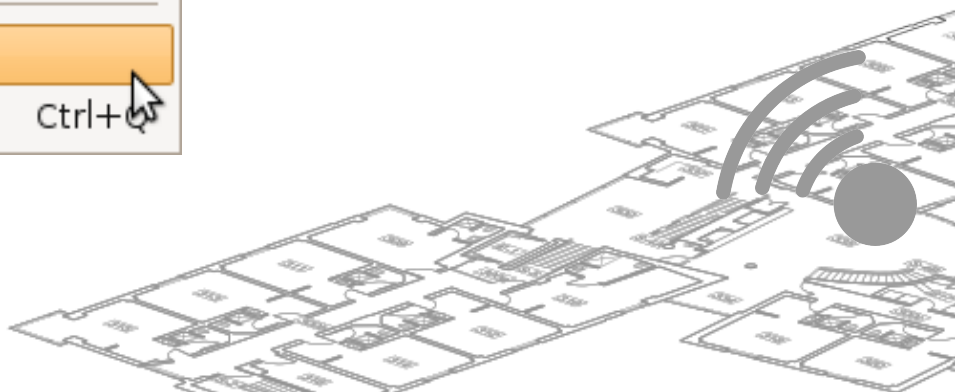
- Platform specific radios
  - Calibration needed
- Significant number of MAC addresses changed in firmware maintenance
- Access point locations moved
  - Old data does not expire
- User error when training



Privacy

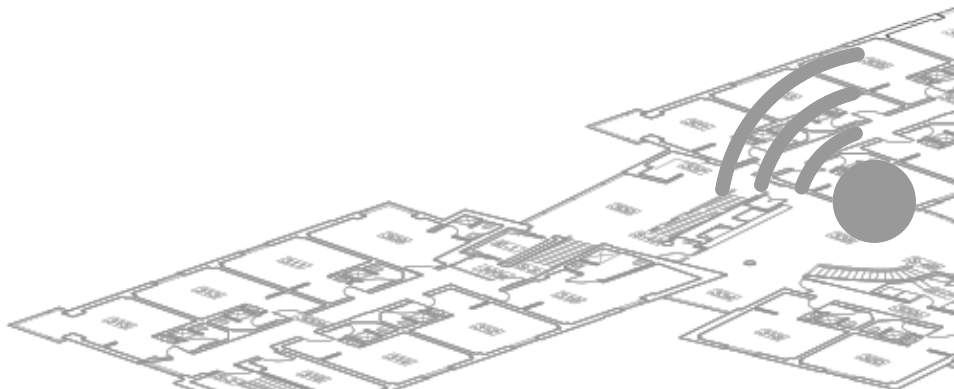
- Opt-in
- Internal to campus network
- Users can remove themselves at any time





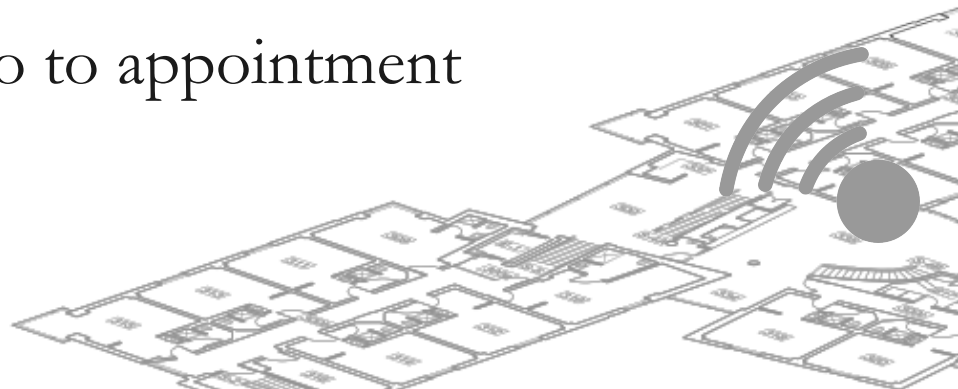
Looking Ahead

- Ad-hoc and prescribed calibration
- Multiple devices per user
  - Port to more devices (Andriod complete)
  - observations/predictions
    - Is the user with their phone or PC?
    - When will the user return to their PC?



Looking Ahead

- Predicting user movement
  - Estimate location without current data
  - Trend identification—tell people their schedules
- New training methods
  - Calendar integration
    - Assign ground-truth data when user goes to appointments that have location tags
    - Must determine if user did go to appointment (and brought wireless device)



Acknowledgements

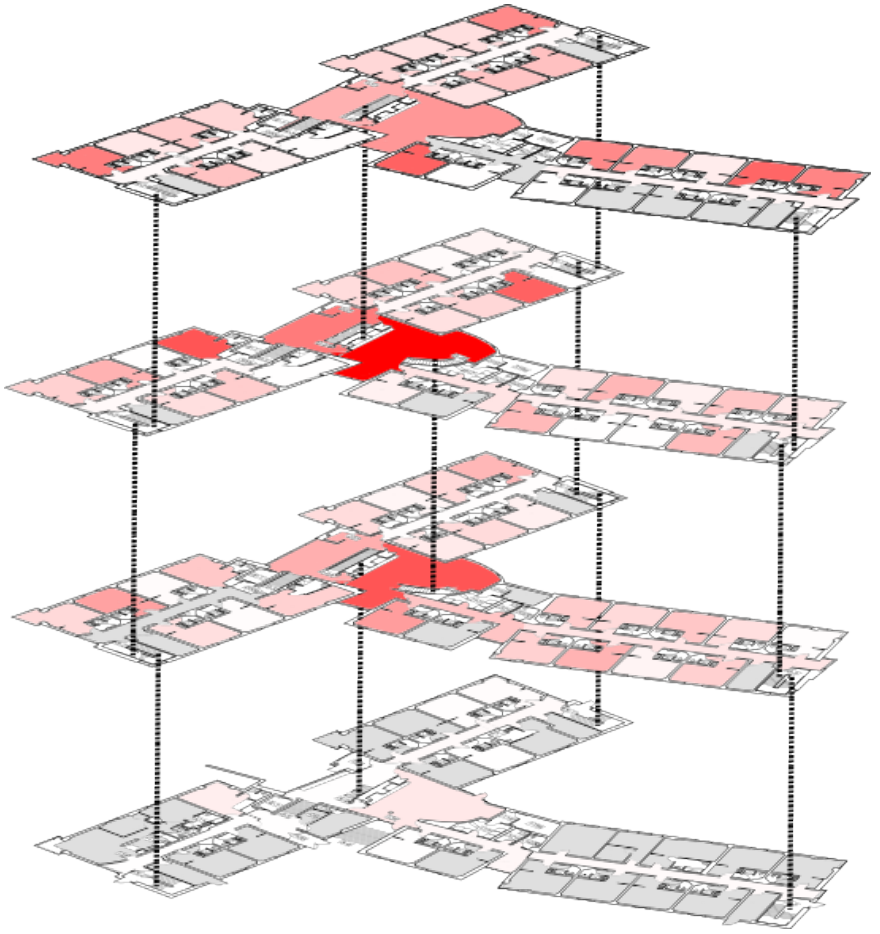
- Ben Fisher
- Mark L. Chang
- F. W. Olin College

the developers at

A Long Duration Study of User-Trained 802.11 Localization

Andrew Barry

MELT 2009
Sept. 30, 2009



Localization Algorithm

- Euclidean distance in 76-dimensional space
- 76 = number of access points

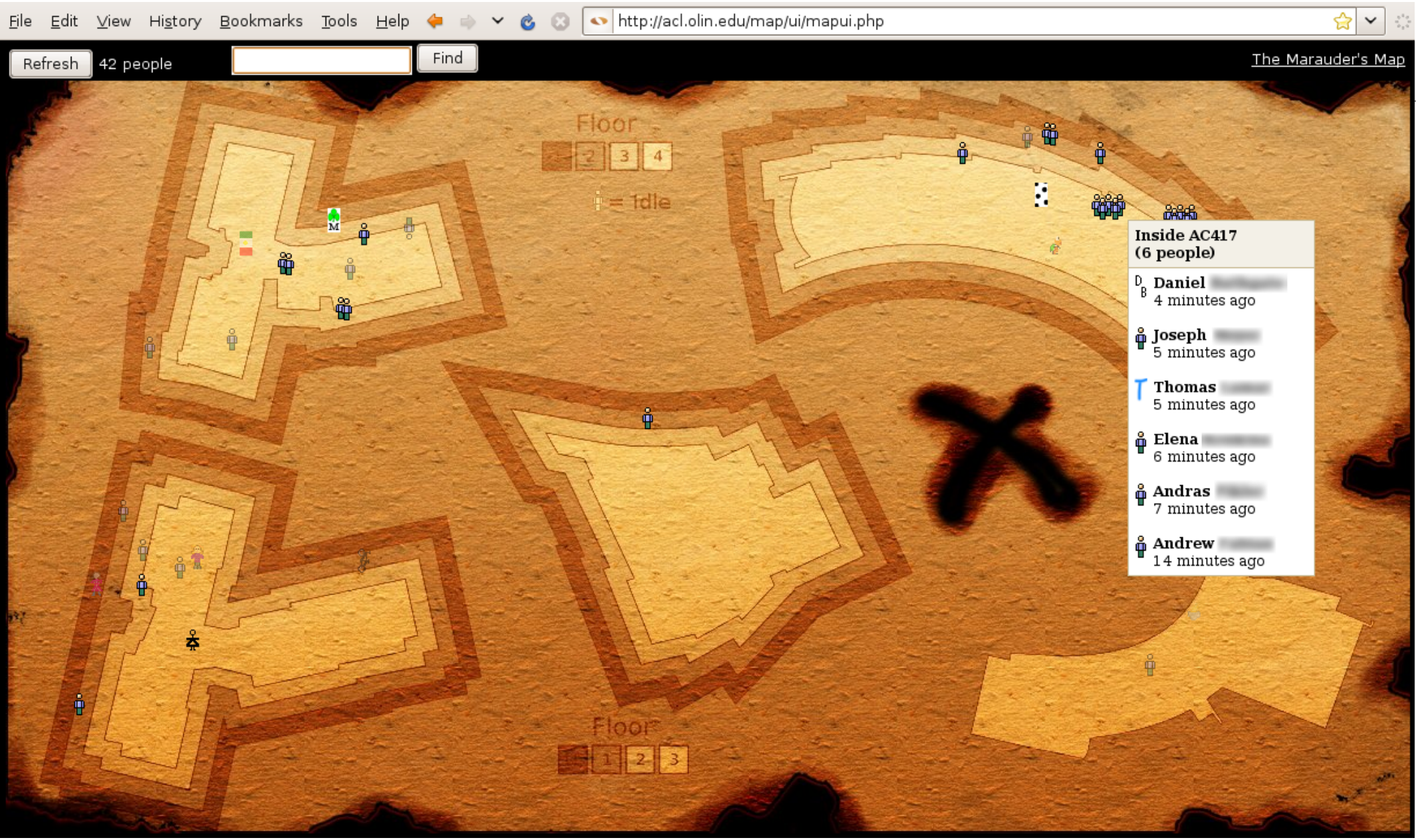
$$D(N) = \sqrt{(C[1] - F[1])^2 + (C[2] - F[2])^2 + ... + (C[76] - F[76])^2}$$

- Minimum D(N) is best location estimate
- C[x]: Array of candidate location's signal strengths
- F[x]: Array of user's location signal strengths
- Similar to RADAR's Nearest Neighbor in Signal Space (NNSS) algorithm.

Localizer Implementation

- Implemented Euclidean distance algorithm in SQL

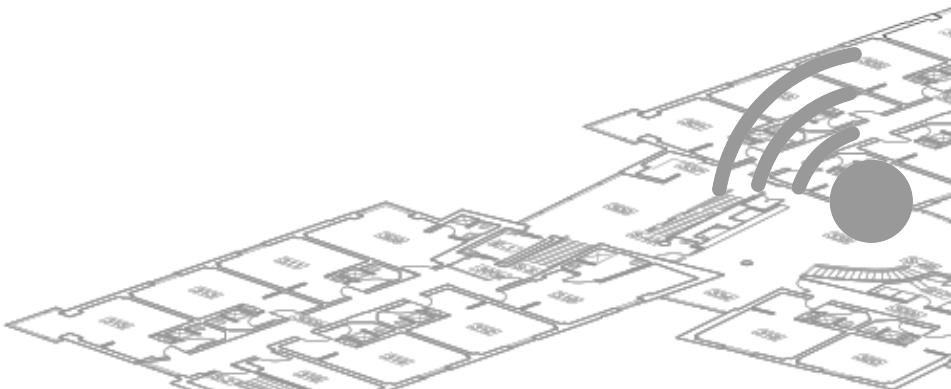
```
SELECT placename, min(pow(C1 -
  F1, 2) + pow(C2 - F2, 2)
  + ... + pow(C76 - F76, 2) AS
  score FROM point WHERE 1
  GROUP BY placename ORDER BY
  score ASC LIMIT 10
```


File Edit View History Bookmarks Tools Help
http://acl.olin.edu/map/ui/mapui.php
Refresh
42 people
Find
The Marauder's Map
Floor
1 2 3 4
= Idle
Inside AC417
(6 people)
Daniel
4 minutes ago
Joseph
5 minutes ago
Thomas
5 minutes ago
Elena
6 minutes ago
Andras
7 minutes ago
Andrew
14 minutes ago
Floor
1 1 2 3

Composition of a Fingerprint

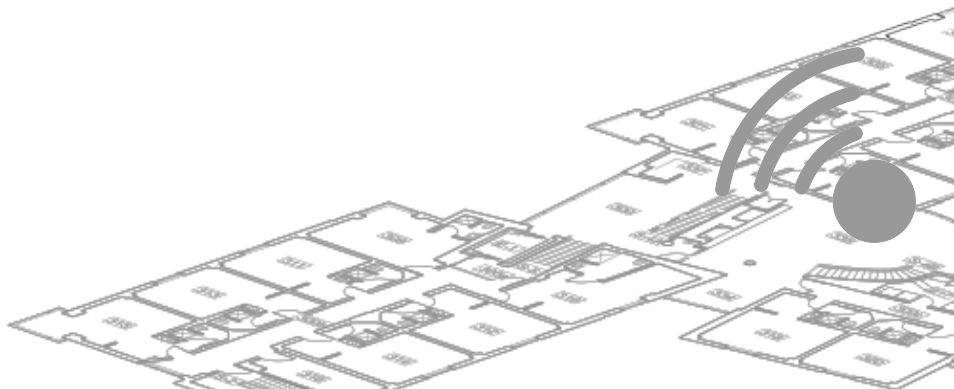
- List of
  - MAC addresses
  - Signal strengths

MAC Address	Signal Strength (dBm)	SSID
00:0B:0E:11:9B:80	-57	OLIN_EH
00:0B:0E:11:82:00	-74	OLIN_EH
00:0B:0E:11:8C:40	-63	OLIN_WH



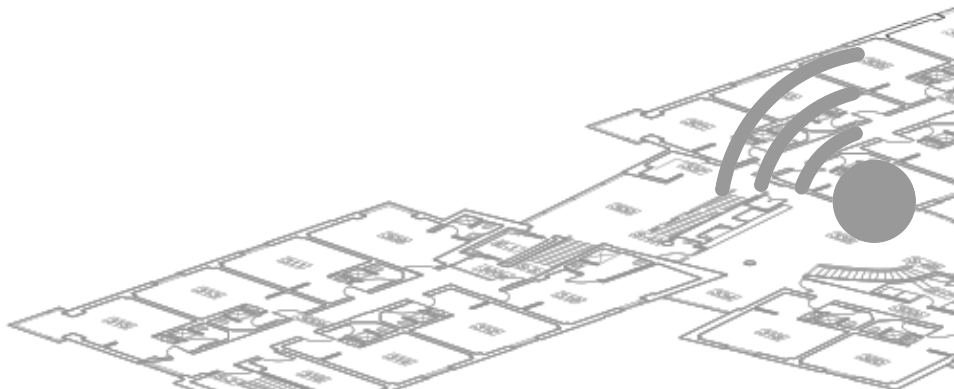
Client Implementation

- wxPython allows cross-platform codebase
- Emphasis on lightweight, non-intrusive, and easy.

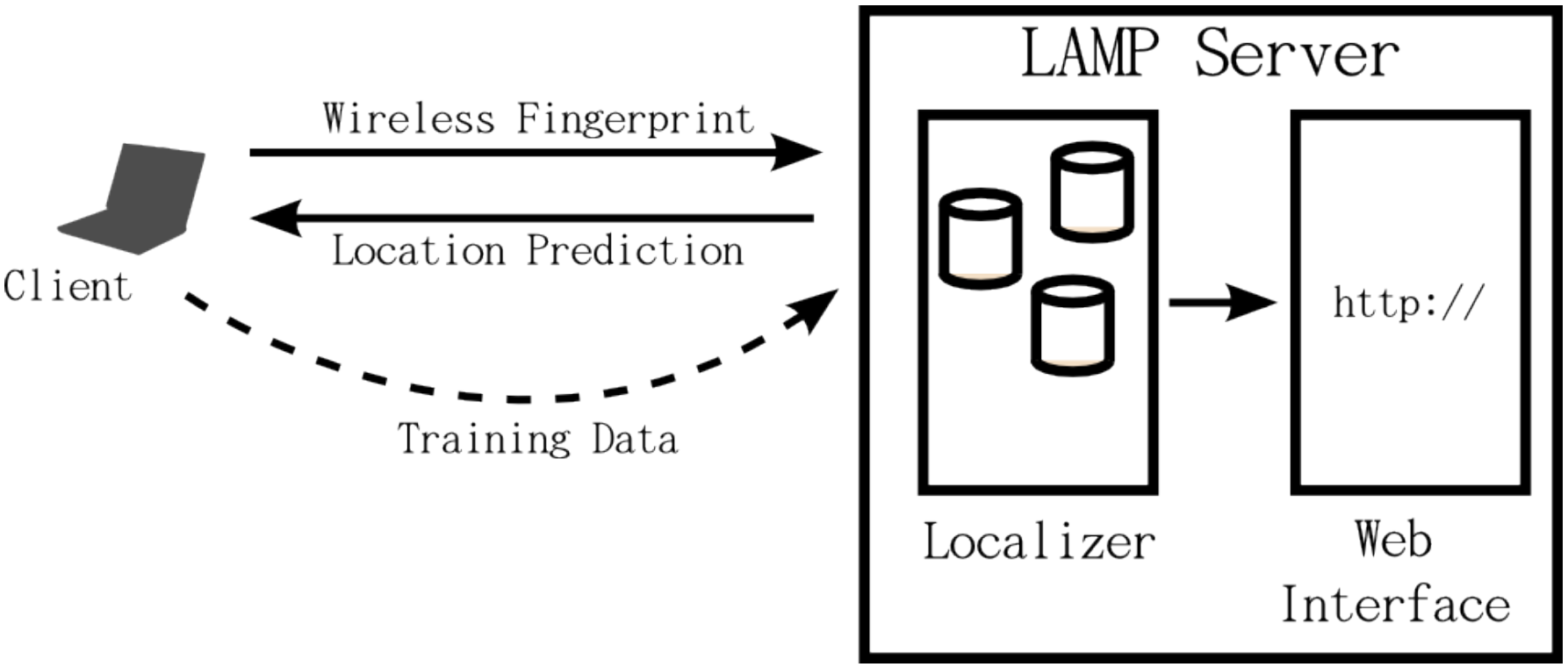


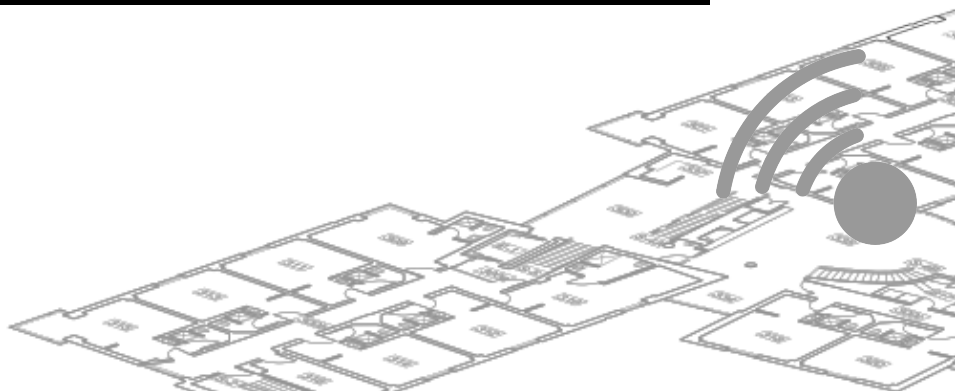
Communication Protocol

- System communication via HTTP GET
  - Same interface used to load webpages
- Information (fingerprints, etc) embedded in URL
- Server responds by producing an HTML page that the client interprets



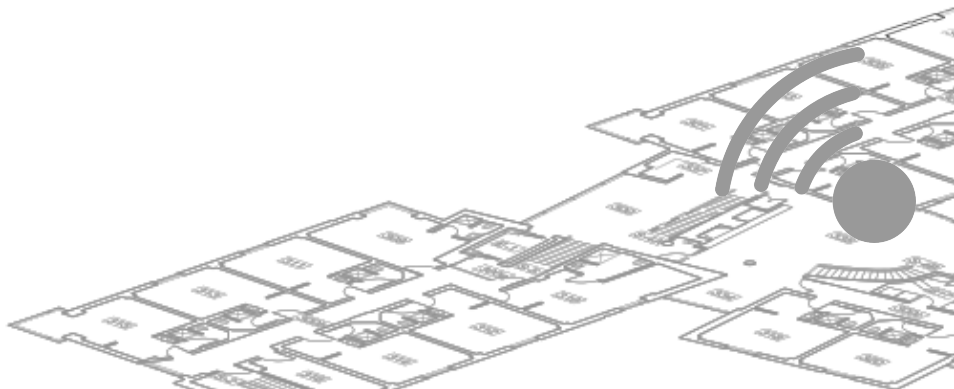
Communication Protocol (cont.)





Backend Services

- Implemented on a LAMP (Linux, Apache, MySQL, PHP) stack
- Database of all known locations and associated fingerprints
- Renders frontend map interface



Database makeup

Location	AP 1	AP 2	...	AP 76
AC312	34	55	...	23
AC128	56	63	...	52
AC109	25	23	...	46