Project 1: Indoor Positioning System

Basic Background and Goals

Setting: There are six access points (i.e. wifi routers) on a certain floor of a building. Devices that are connected to the network can measure the signal strength of all access points within range.

Problem: The client wants to identify the physical location of devices that are connected to this network.

Goal: Create a model that takes a set of signal strengths of the relevant access points to a connected device and predicts the physical location of that device. Quantify the accuracy and precision of the model.

Deliverables:

- A report including appendices detailing your methodology
- A presentation summarizing your findings

Possible use cases:

- Track laptops that are owned by a university and lent to students
- Produce a real-time map of mobile medical equipment in a hospital (e.g. quickly locate a heart rate monitor)

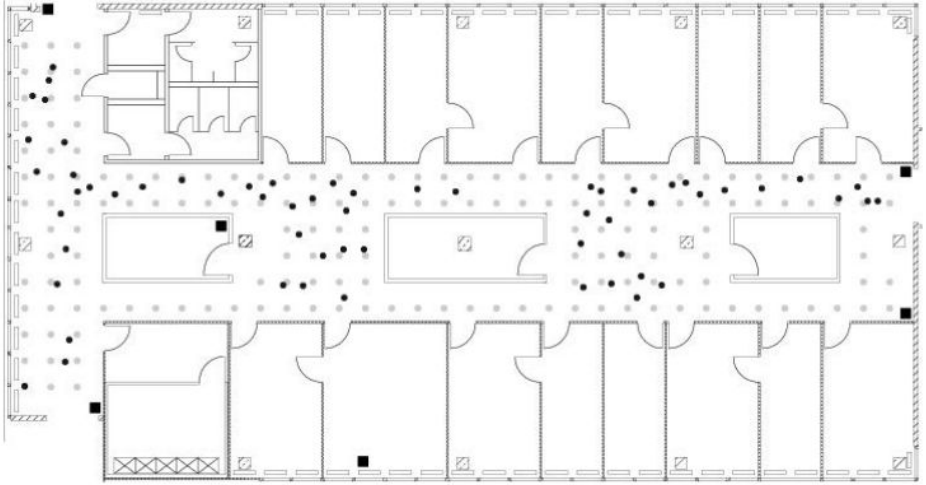
Data Details

Location:

- Floorplan is roughly 15mx36m

Data Collection Methodology

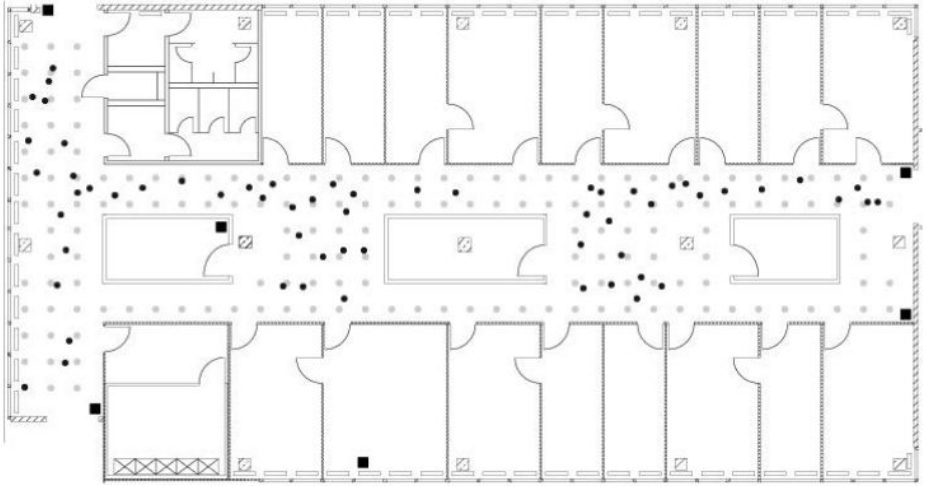
- Testing devices are connected to the network. They are taken to various locations and oriented in various directions.
- Testing device generates data by recording the signal strength of all access points
- In the figure, access points are squares, grey dots are *offline data* locations and black dots are *online data* locations



Data Details

Offline Data: Designed sampling on a grid

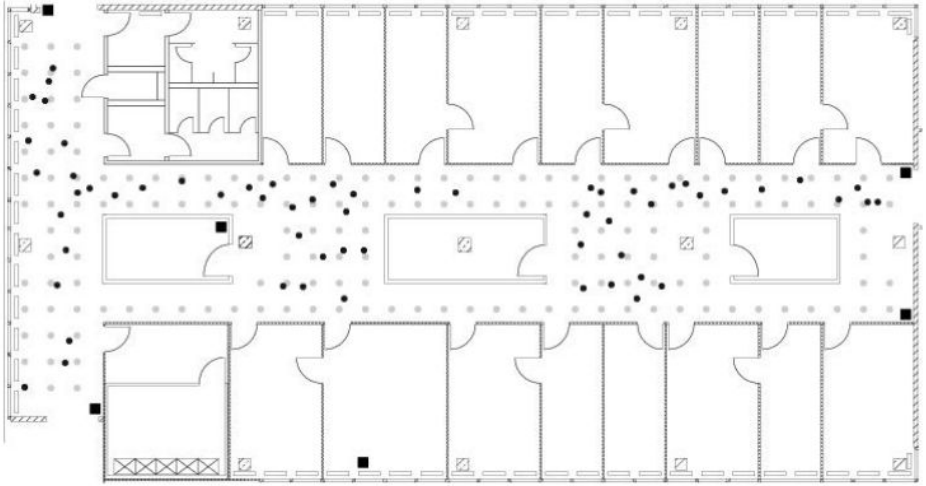
- Grid resolution is 1m
- 166 locations
- Eight orientations were measured at each location in 45-degree increments
- Every location/orientation combination was sampled 110 times
    - Implication: Each location was sampled 880 times total
    - All samples were measured with the same device
- Intended use: Training data



Data Details

Online Data: Random location/orientations

- Simulates real-world data
- 60 location/orientation combinations were chosen randomly
  - Both locations and angles sampled continuously with a finite precision
- Each location/orientation combination was sampled 110 times
  - Total of 660 measurements
- Intended use: Testing data



Data Format

File format: .txt files

- Online and offline data given in separate files that share the same format

```
t="Timestamp";
id="MACofScanDevice";
pos="RealPosition";
degree="orientation";
MACofResponse1="SignalStrengthValue,Frequency,Mode"; ...
MACofResponseN="SignalStrengthValue,Frequency,Mode"
```

How to interpret data:

- Lines beginning with “#” are comments, and every set of 110 samples begin with three comment lines
- In the figure to the right, “\” indicates the continuation of a line of data

```
# timestamp=2006-02-11 08:31:58
# usec=250
# minReadings=110
t=1139643118358;id=00:02:2D:21:0F:33;pos=0.0,0.0,0.0;degree=0.0;\
00:14:bf:b1:97:8a=-38,2437000000,3;\
00:14:bf:b1:97:90=-56,2427000000,3;\
00:0f:a3:39:e1:c0=-53,2462000000,3;\
00:14:bf:b1:97:8d=-65,2442000000,3;\
00:14:bf:b1:97:81=-65,2422000000,3;\
00:14:bf:3b:c7:c6=-66,2432000000,3;\
00:0f:a3:39:dd:cd=-75,2412000000,3;\
00:0f:a3:39:e0:4b=-78,2462000000,3;\
00:0f:a3:39:e2:10=-87,2437000000,3;\
02:64:fb:68:52:e6=-88,2447000000,1;\
02:00:42:55:31:00=-84,2457000000,1
```

Data Format

MAC addresses:

- Unique identifier of electronic device
- Format is `mm:mm:mm:ss:ss:ss`
  - Each pair of digits is hexadecimal (0, 1, ..., 9, a, b, c, d, e, f)
  - First three pairs of digits indicate manufacturer of device. Last three indicate model and unique device

Relevant access point locations:

- Not included in either signal measurement data set. Included in a separate .txt file.

Macs	x	y
00:0f:a3:39:e1:c0	7.5	6.3
00:14:bf:b1:97:8a	2.5	-0.8
00:14:bf:3b:c7:c6	12.8	-2.8
00:14:bf:b1:97:90	1.0	14.0
00:14:bf:b1:97:8d	33.5	9.3
00:14:bf:b1:97:81	33.5	2.8

```
t="Timestamp";
id="MACofScanDevice";
pos="RealPosition";
degree="orientation";
MACofResponse1="SignalStrengthValue,Frequency,Mode"; ...
MACofResponseN="SignalStrengthValue,Frequency,Mode"
```

```
# timestamp=2006-02-11 08:31:58
# usec=250
# minReadings=110
t=1139643118358;id=00:02:2D:21:0F:33;pos=0.0,0.0,0.0;degree=0.0;\
00:14:bf:b1:97:8a=-38,2437000000,3;\
00:14:bf:b1:97:90=-56,2427000000,3;\
00:0f:a3:39:e1:c0=-53,2462000000,3;\
00:14:bf:b1:97:8d=-65,2442000000,3;\
00:14:bf:b1:97:81=-65,2422000000,3;\
00:14:bf:3b:c7:c6=-66,2432000000,3;\
00:0f:a3:39:dd:cd=-75,2412000000,3;\
00:0f:a3:39:e0:4b=-78,2462000000,3;\
00:0f:a3:39:e2:10=-87,2437000000,3;\
02:64:fb:68:52:e6=-88,2447000000,1;\
02:00:42:55:31:00=-84,2457000000,1
```

Data Format

Things to watch out for:

- Mode indicates whether the signal originates from access point or an ad hoc device (i.e. another device on the network such as someone's phone)
- There's no guarantee that all six access points we're interested in will appear in every sample
  - E.g. if the device is too far away from an access point, it may not detect any signal from it
- The same access point can be measured multiple times within the same reading

```
t="Timestamp";
id="MACofScanDevice";
pos="RealPosition";
degree="orientation";
MACofResponse1="SignalStrengthValue,Frequency,Mode"; ...
MACofResponseN="SignalStrengthValue,Frequency,Mode"
```

```
# timestamp=2006-02-11 08:31:58
# usec=250
# minReadings=110
t=1139643118358;id=00:02:2D:21:0F:33;pos=0.0,0.0,0.0;degree=0.0;\
00:14:bf:b1:97:8a=-38,2437000000,3;\
00:14:bf:b1:97:90=-56,2427000000,3;\
00:0f:a3:39:e1:c0=-53,2462000000,3;\
00:14:bf:b1:97:8d=-65,2442000000,3;\
00:14:bf:b1:97:81=-65,2422000000,3;\
00:14:bf:3b:c7:c6=-66,2432000000,3;\
00:0f:a3:39:dd:cd=-75,2412000000,3;\
00:0f:a3:39:e0:4b=-78,2462000000,3;\
00:0f:a3:39:e2:10=-87,2437000000,3;\
02:64:fb:68:52:e6=-88,2447000000,1;\
02:00:42:55:31:00=-84,2457000000,1
```

```
t=1139644637174;id=00:02:2D:21:0F:33;pos=2.0,0.0,0.0;degree=45.5;\
00:14:bf:b1:97:8a=-33,2437000000,3;\
00:14:bf:b1:97:8a=-38,2437000000,3;\
00:0f:a3:39:e1:c0=-54,2462000000,3;\
```