

# Mannheim University Geo-location Lab



- Can we locate a person, medical equipment, etc. in a building via its wireless signal?
- 6 wireless base stations within a building.
- Person walks through the building and measures the signal from a hand-held wireless device to each of the base stations.
- Person's orientation is important, so make them turn through 360 degrees and get measurements.
- Given a new device, determine where it is?



# Data

- Each record has a
  - time stamp ( $t=1139692477303$ )
  - own MAC address ( $id=00:02:2D:21:0F:33$ )
  - position ( $pos=0.0,0.5,0.0$ )
  - orientation ( $degree=130.5$ )
  - for each base station it can detect, measures
    - MAC address of the base station
    - signal strength
    - channel frequency
    - whether it is an access point or ad hoc network device.



# Data

```
# timestamp=2006-02-11 22:14:37
# usec=250
# minReadings=110
t=1139692477303;id=00:02:2D:21:0F:33;pos=0.0,0.05,0.0;degree=130.5;00:14:bf:b1:97:8a=-43,2437000000,3;00:0f:a3:39:e1:c0=-52,2462000000,3;00:14:bf:3b:c7:c6=-62,2432000000,3;00:14:bf:b1:97:81=-58,2422000000,3;00:14:bf:b1:97:8d=-62,2442000000,3;00:14:bf:b1:97:90=-57,2427000000,3;00:0f:a3:39:e0:4b=-79,2462000000,3;00:0f:a3:39:e2:10=-88,2437000000,3;00:0f:a3:39:dd:cd=-64,2412000000,3;02:64:fb:68:52:e6=-87,2447000000,1;02:00:42:55:31:00=-85,2457000000,1
t=1139692477555;id=00:02:2D:21:0F:33;pos=0.0,0.05,0.0;degree=130.5;00:14:bf:b1:97:8a=-43,2437000000,3;00:14:bf:b1:97:8a=-43,2437000000,3;00:0f:a3:39:e1:c0=-52,2462000000,3;00:14:bf:b1:97:90=-57,2427000000,3;00:14:bf:b1:97:8d=-64,2442000000,3;00:0f:a3:39:e0:4b=-77,2462000000,3;00:0f:a3:39:dd:cd=-62,2412000000,3;02:00:42:55:31:00=-85,2457000000,1;02:64:fb:68:52:e6=-88,2447000000,1
t=1139692477807;id=00:02:2D:21:0F:33;pos=0.0,0.05,0.0;degree=130.5;00:0f:a3:39:e1:c0=-51,2462000000,3;00:14:bf:b1:97:90=-49,2427000000,3;00:14:bf:3b:c7:c6=-62,2432000000,3;00:14:bf:b1:97:8a=-44,2437000000,3;00:14:bf:b1:97:81=-68,2422000000,3;00:0f:a3:39:e0:4b=-75,2462000000,3;00:0f:a3:39:dd:cd=-66,2412000000,3;00:0f:a3:39:e2:10=-90,2437000000,3;02:00:42:55:31:00=-87,2457000000,1
t=1139692478059;id=00:02:2D:21:0F:33;pos=0.0,0.05,0.0;degree=130.5;00:0f:a3:39:e1:c0=-52,2462000000,3;00:14:bf:b1:97:90=-58,2427000000,3;00:14:bf:b1:97:8a=-39,2437000000,3;00:14:bf:b1:97:8d=-63,2442000000,3;00:0f:a3:39:dd:cd=-66,2412000000,3;00:14:bf:3b:c7:c6=-68,2432000000,3;00:0f:a3:39:e0:4b=-76,2462000000,3;00:0f:a3:39:e2:10=-88,2437000000,3;02:00:42:55:31:00=-85,2457000000,1;02:64:fb:68:52:e6=-88,2447000000,1
t=1139692478311;id=00:02:2D:21:0F:33;pos=0.0,0.05,0.0;degree=130.5;00:14:bf:b1:97:8a=-42,2437000000,3;00:0f:a3:39:e1:c0=-51,2462000000,3;00:14:bf:b1:97:81=-60,2422000000,3;00:14:bf:3b:c7:c6=-59,2432000000,3;00:14:bf:b1:97:8d=-64,2442000000,3;00:0f:a3:39:dd:cd=-61,2412000000,3;00:0f:a3:39:e0:4b=-76,2462000000,3;00:0f:a3:39:e2:10=-90,2437000000,3
t=1139692478563;id=00:02:2D:21:0F:33;pos=0.0,0.05,0.0;degree=130.5;00:14:bf:b1:97:90=-57,2427000000,3;00:0f:a3:39:e1:c0=-51,2462000000,3;00:14:bf:b1:97:8a=-43,2437000000,3;00:14:bf:3b:c7:c6=-62,2432000000,3;00:14:bf:b1:97:8d=-63,2442000000,3;00:0f:a3:39:dd:cd=-62,2412000000,3;00:0f:a3:39:e0:4b=-76,2462000000,3;00:14:bf:b1:97:81=-57,2422000000,3;00:0f:a3:39:e2:10=-89,2437000000,3;02:00:42:55:31:00=-85,2457000000,1;02:64:fb:68:52:e6=-87,2447000000,1
t=1139692478819;id=00:02:2D:21:0F:33;pos=0.0,0.05,0.0;degree=130.5;00:14:bf:b1:97:8a=-44,2437000000,3;00:0f:a3:39:e1:c0=-50,2462000000,3;00:14:bf:3b:c7:c6=-62,2432000000,3;00:14:bf:b1:97:90=-56,2427000000,3;00:0f:a3:39:dd:cd=-61,2412000000,3;00:14:bf:b1:97:8d=-63,2442000000,3;00:0f:a3:39:e0:4b=-76,2462000000,3;00:0f:a3:39:e2:10=-91,2437000000,3;02:00:42:55:31:00=-87,2457000000,1;02:64:fb:68:52:e6=-89,2447000000,1
```



# Student Tasks

- Read the data into an R data frame
  - Break each record into values of variables.
  - How to represent the data for easy computation?
- Validate and “clean” the data
- Explore the data graphically!
- Use different methods to create a predictor for a new signal strength vector to determine (x, y) of signal source.



- Model signal strength as  $1/\text{distance}^2$  or  $1/\log(x)$  ... and then triangulate
- K-nearest neighbors
  - cross-validation to determine
    - k
    - weights
  - avoid recomputing distance matrix for each training sample.



# Topics covered

- Text manipulation
- Data exploration
- Database access for the offline data
- statistical methodology - k-NN
- numerical computations