Wi-Fi Device Localization in an Indoor Environment Using Graph Mapping

Abheek Saha
Abheek.saha@hsc.com
Hughes Systique Corporation

Mayank Maheshwari
Mayank.maheshwari@hsc.com
Hughes Systique Corporation

Abheek Saha

Interested in applications of mathematical modelling and optimization in the real world

Mayank Maheshwari

Interested in Wi-Fi and Bluetooth based Indoor positioning systems

Motivation

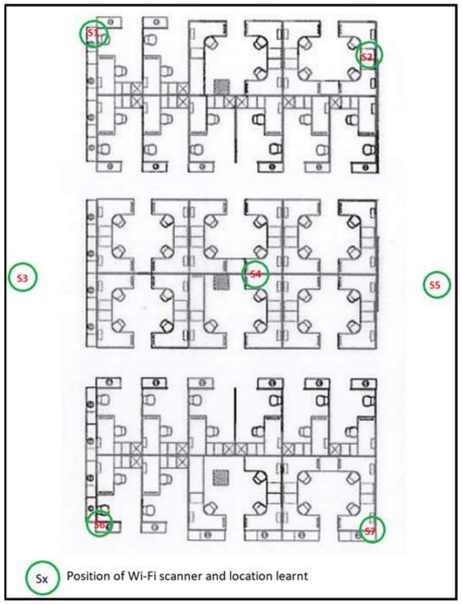
- What are we doing here?
  - Our interest is in the study of crowds in indoor locations using sampled position data
- Key points
  - Why crowds?
    - Security
    - Distancing in the era of CoVID 19
    - Modeling of crowd movement
    - Identification of 'hotspots' and 'coldspots'
  - Mass behavior
    - We are not interesting in collecting movement of individual users
    - We don't want to track or keep data regarding individuals

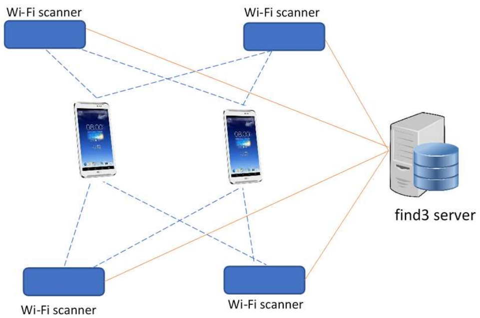
Prior Art

- Cellular Based
  - Accuracy is low, generally in the range of 50-20 meter
- Bluetooth
  - Smaller range as compared to Wi-Fi
- UWB
  - High accuracy
  - Not available in most of the mobile phones so not suitable for crowd tracking
- Wi-Fi based lateration techniques
  - Time of Arrival and Time Difference of Arrival
    - Requires time synchronization between Wi-Fi transmitter and receiver or among receivers
    - Very accurate measurement of time of Time of arrival or Time difference of arrival
  - RSSI Propagation loss model
    - RSSI propagation loss model is used to calculate distance between transmitter and receiver
    - Distance between transmitter and three or more receiver is used to find the location of the transmitter
    - These techniques do not work very well because of multipath in the indoor environment

Location patterning using RSSI fingerprinting

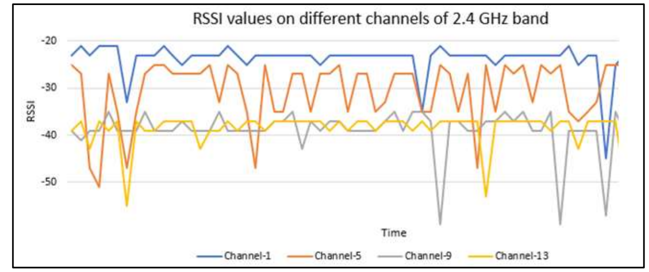
- Location patterning technique is based on the sampling and recording of radio signal patterns in specific environments
- Location patterning techniques fundamentally assumes that each potential device location ideally possesses a distinctly unique RF “signature”
- Location patterning solutions typically base such signatures on received signal strength (RSSI). This technique involves two phases
- Calibration
  - RSSI data is collected to determine the RF signatures of desired locations
  - RSS values associated with the device are recorded into a database known
  - Because of fading and other phenomena, the observed RSSI of a device at a location is not static but vary over time. As a result, multiple samples of RSSI for a device are collected during the calibration phase.
  - RSSI signature DB is used to training various ML classifier algorithms
- Tracking
  - Group of receiving sensors provide signal strength measurements of tracked the mobile device and forwards that information to a location tracking server
  - The location server uses a trained ML algorithms and the RSSI signature to estimate the location of the device

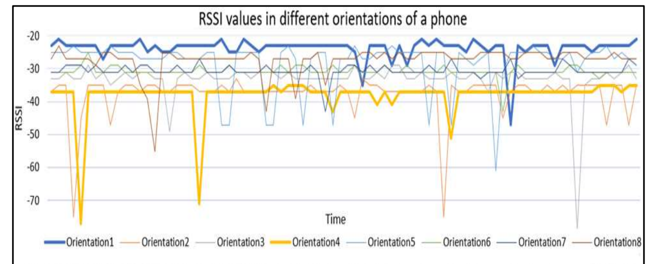




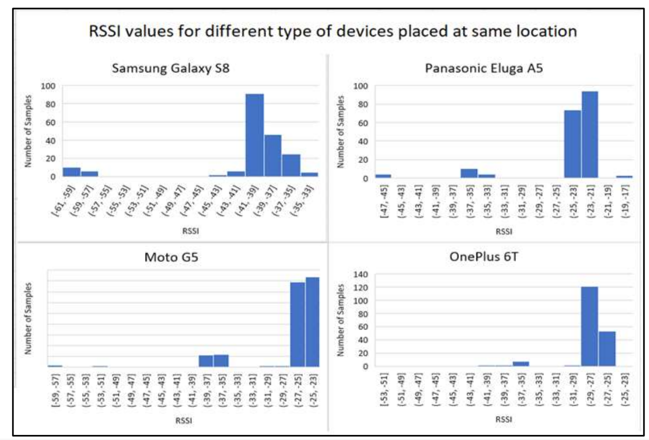
Challenges of Indoor Position Determination

- Dependency on Wi-Fi Channel
  - Considerable difference in RSSI values of probes on different channels, even when the location of the Wi-Fi transmitter and Wi-Fi receiver remains the same
  - So Wi-Fi RSSI based indoor localization system should account for RSSI differences on different Wi-Fi channels.
- Dependency on device orientation
  - RSSI values change significantly with the change in device orientation without any change in device location
  - The angle between Orientation1 and Orientation4 was 180^0, and RSSI values on these two orientations differ by about 18 dB
- Dependency on device orientation
  - RSSI values change significantly with change in Wi-Fi device type. Graph shows that average RSSI value from OnePlus 6T (-27 dB) and average RSSI value from Motorola G5(-39 dB) differ by about -12 dB
  - If the type of the device used during calibration is different from device used during tracking, the accuracy of the prediction deteriorate





Orientation1: Angle between transmitter at phone and receiver at scanner = θ OrientationN: Angle between transmitter at phone and receiver at scanne = θ + (N*45°)



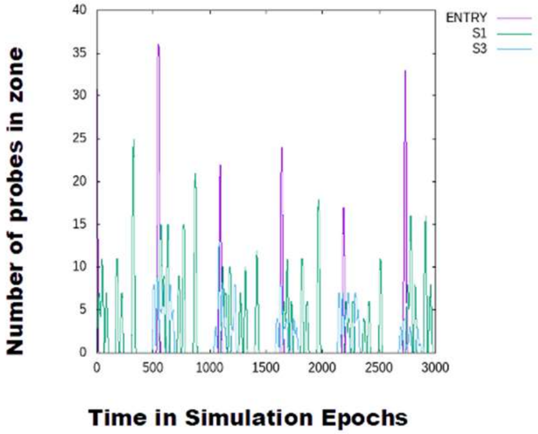
ML Algorithms

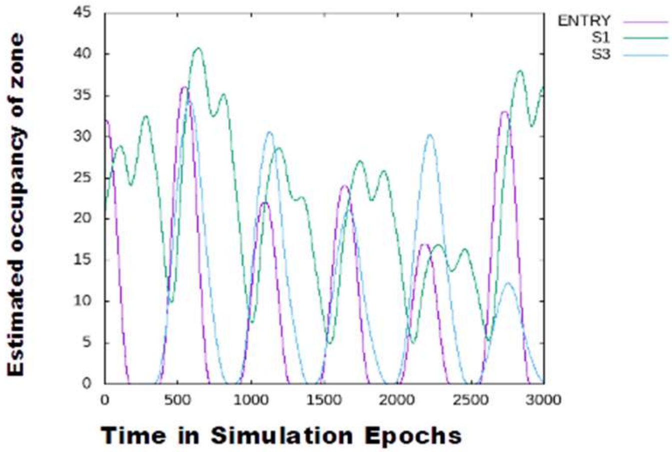
- We used the Machine Learning package find3 for location prediction
- The find3 package runs multiple machine learning algorithms in parallel then chooses the best among them using the Youden's J statistic diagnostic metric
- Algorithms used include include the K-nearest neighbor, linear SVM, Decision tree, Random Forest, and Extend Naive Bayes
- Using the labeled data provided, each algorithm is trained with a subset of the data and then tested using the remaining part of the data. The prediction is in the form of a probability factor *PL* for each location *L*. Based on the predictions by ML algorithms Youden's J statistic is calculated for each location and each ML algorithm
- We obtained more than 80% prediction accuracy on all of these devices, within a 3 meter radius of the calibration positions
- In future we would like to refine the algorithm to handle incomplete input i.e. the situation when a probe request is not received by all the Wi-Fi scanners
- Retraining of the ML for each change in interior topology is CPU intensive and slow; hence, we would like to find ways to augment existing algorithms for minor changes, rather than retrain the entire ML

Find3 -https://github.com/schollz/find3

From Raw Data to Occupancy

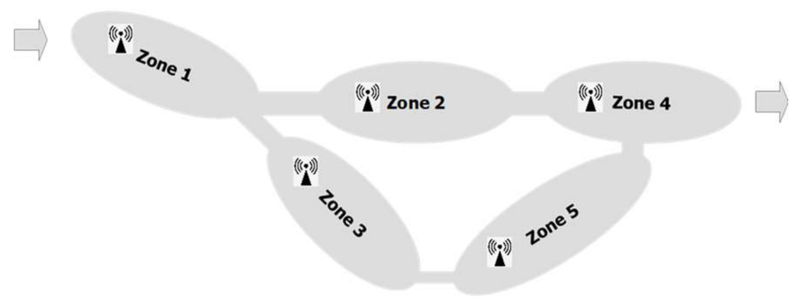
- Despreading
  - Converting sampled time-series data to continuous measures
- Modelling
  - Parameter identification
  - Identification of the state space
  - Modelling of mass behaviour
- Verification
  - Mapping of mathematical model against real world data
  - Boundary conditions

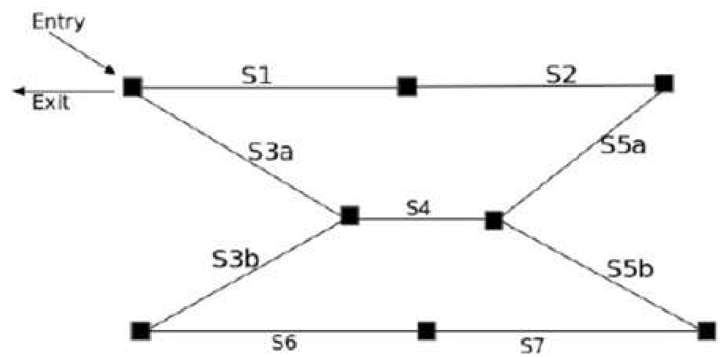




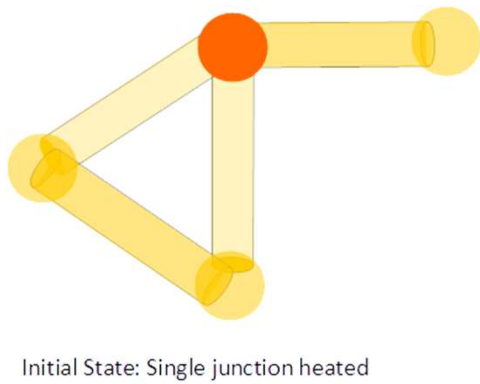
Crowd Modelling in an Indoor Area

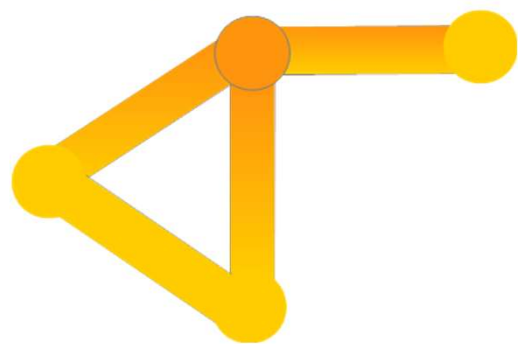
- Methodology
  - We consider an indoor arena as a set of zones connected by passages
  - The obvious mathematical model is a graph
    - Position is always measured in terms of a specific zone
    - To increase the zones, we increase the number of Access Points
    - We will predict crowd behaviour in terms of occupancy of edges
      - A zone maps to an edge
      - A hotspot is when an edge contains a large number of people etc.
  - Transition matrix ensures Kirchoff transition conditions at the nodes





Heat Diffusion Model

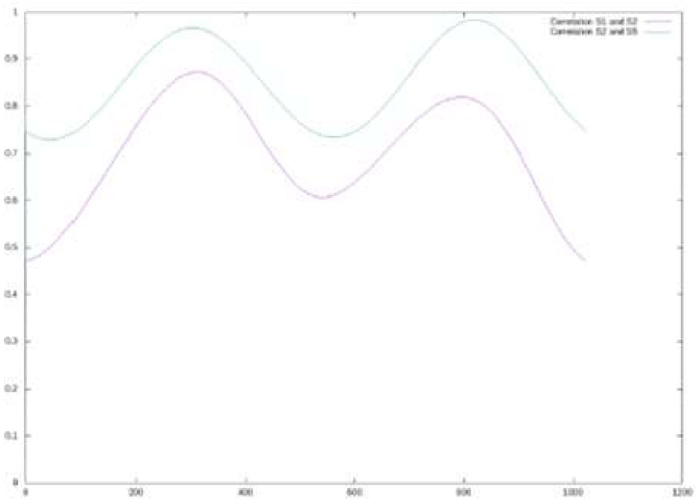




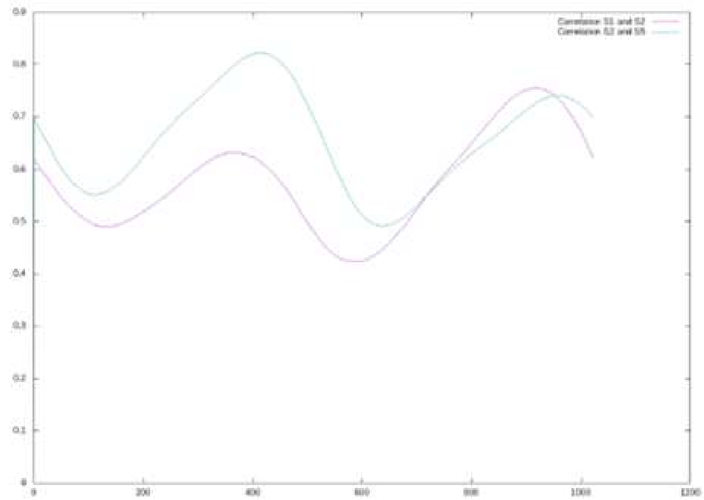
- Crowd entering through a door is equivalent to a single junction being heated.
- The dispersion model shows us how the intensity dissipates over time.
- Its natural to dissipate from higher to lower intensity

$$\nabla^2 x^i + \eta^i \frac{\partial x^i}{\partial t} + g(x^j) = 0, j \in Inc(i)$$

Mass behaviour



(a). Correlation between adjacent edges - phone always on



(b). Correlation between adjacent edges - random probing intervals

- We track correlation between occupancy of adjacent paths
  - For each path in the graph we can compare predicted occupancy vs actual occupancy over time.

Thank you