DATA DRIVEN APPROACHES TO LOCALIZATION

Saravnan S V
Project Member
EC01-AR Indoor Tour Guide

ANALYTICAL APPROACHES TO LOCALIZATION

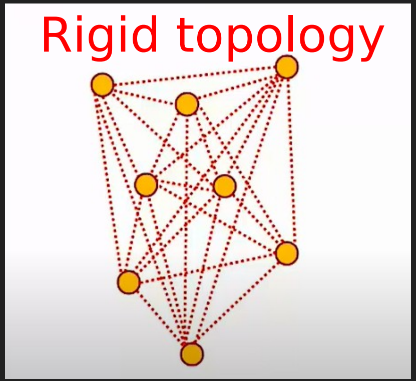
- **Localization (X,Y)** generally requires Ranges and information related to the positions of the Anchors. Ranging is done through **Time of Flight/Arrival** (ToF/ToA). Positioning of the anchors must be pre-known.
- However, at tough times like disasters, it is almost impractical to collect information regarding the anchors. Even otherwise, it requires a lot of effort and intensive infrastructure. So, we often restrict ourselves to the concept of **Relative Localization** where the infrastructure requirements are typically reduced.

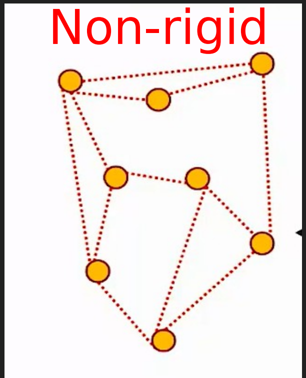
RELATIVE LOCALIZATION
EUCLIDEAN DISTANCE MATRIX (EDM)

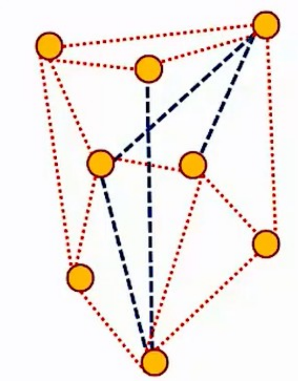
- The **Euclidean Distance Matrix (EDM)** is a triangular matrix where the element corresponding to n^{th} row and m^{th} column and is equal to the distance between the node labelled n and the node labelled m for n<m (since the distance is same between two points no matter from which point we measure it to the other, the lower half of the matrix is not needed). It is 0 for m=n for obvious reasons.
- The order of the EDM is roughly N^2 where N is the total no. of nodes. Here, a node may be fixed (like the beacons) or moving (like the device that is to be head-mounted).
- This EDM upon **Multi-Dimensional Scaling (MDS)** forms a **Topology** which is a 2-dimensional embedding. In practice however, we build the EDM from the topology. If the corresponding EDM does not change upon mirroring/flipping/rotating it, the topology is said to be rigid (The phenomenon is called **Rigidity**).
- In the cases where all the distances are specified/accessible, the topology is always rigid. If not, flipping the topology may not give out the same EDM. However, interestingly, a few **critical ranges** (distances) are sufficient to ensure rigidity. Using these critical ranges, the remaining ranges can be predicted approximately and a rigid topology can be built as a result (right most example).

Watershed[a]	T1	T2	T3	O	G	V	B	S	F	H	N	NI	K2	K5	K3	P	K1	K4
T1	0.00	0.90	2.55	6.37	7.08	3.90	6.82	5.45	5.75	3.19	5.69	3.73	4.16	3.51	3.39	3.78	2.99	3.16
T2		0.00	2.72	6.36	7.17	4.15	6.94	5.51	5.99	3.62	6.10	3.75	4.54	3.84	3.82	4.06	3.01	3.28
T3			0.00	5.95	7.14	3.82	7.03	4.53	5.96	3.55	6.31	5.39	3.94	2.93	3.59	4.63	3.33	3.45
O				0.00	3.65	3.98	3.40	2.98	3.55	7.02	7.50	7.66	4.89	6.47	7.29	7.20	6.42	5.95
G					0.00	4.52	4.81	3.98	2.29	7.85	8.16	8.10	5.99	7.32	8.21	7.95	7.33	6.88
V						0.00	5.77	3.17	2.96	4.86	6.87	5.87	4.34	3.76	4.14	4.11	4.50	4.31
B							0.00	5.20	4.69	7.43	6.23	7.72	4.82	7.83	8.52	8.52	7.27	6.50
S								0.00	3.79	6.00	7.74	7.23	4.78	5.10	6.35	6.43	5.56	5.30
F									0.00	6.83	7.50	7.40	5.31	6.29	6.68	6.57	6.60	6.18
H										0.00	4.74	3.87	3.52	2.70	3.83	4.15	2.75	2.81
N											0.00	4.99	3.45	6.12	6.76	7.19	5.13	4.27
NI												0.00	4.88	4.69	4.76	4.26	2.76	2.88
K2													0.00	3.89	5.16	5.61	3.58	2.83
K5														0.00	2.77	3.22	2.50	2.84
K3															0.00	1.86	3.58	3.96
P																0.00	3.69	4.16
K1																	0.00	1.11

EDM







STATISTICAL/DATA-DRIVEN APPROACHES TO LOCALIZATION

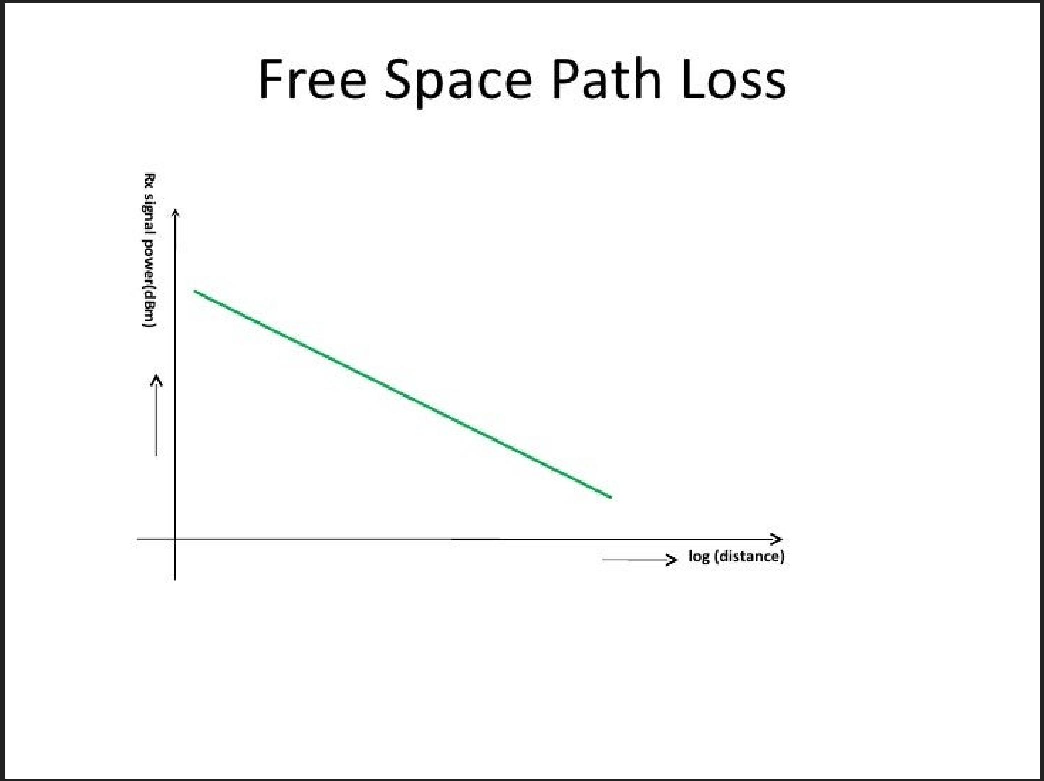
- In cases where the no. of dynamic nodes is higher, in the Relative Localization approach using EDM, the matrix needs to be updated frequently. This can be a very expensive process at times. Moreover, the accuracies are lower due to obstacles (obstacles slow down the waves, thereby increasing the ToA falsely).
- With Data-driven approaches (which may involve **Received Signal Strength Indicator (RSSI),** fingerprinting, statistical filters, RSSI, and ML models), the errors in ToA and hence, distance can be reduced significantly. Infrastructure might not even be required if we rely on the **Inertial Movement Unit (IMU)** sensors which usually come within a smart phone.

FREE-SPACE PATH-LOSS MODEL
(based on RSS & RSSI)

- The Received Signal Strength (RSS) when plotted against the distance (on semi-log plane), a curve that is close to a straight line is observed. This curve is actually obtained by using least square fit. This suggests that RSS, which is measured/accessed using an RSSI, can be mapped to range (distance) and hence might be an effective approach to our problem.
- *; n is the path loss exponent; d is the distance*

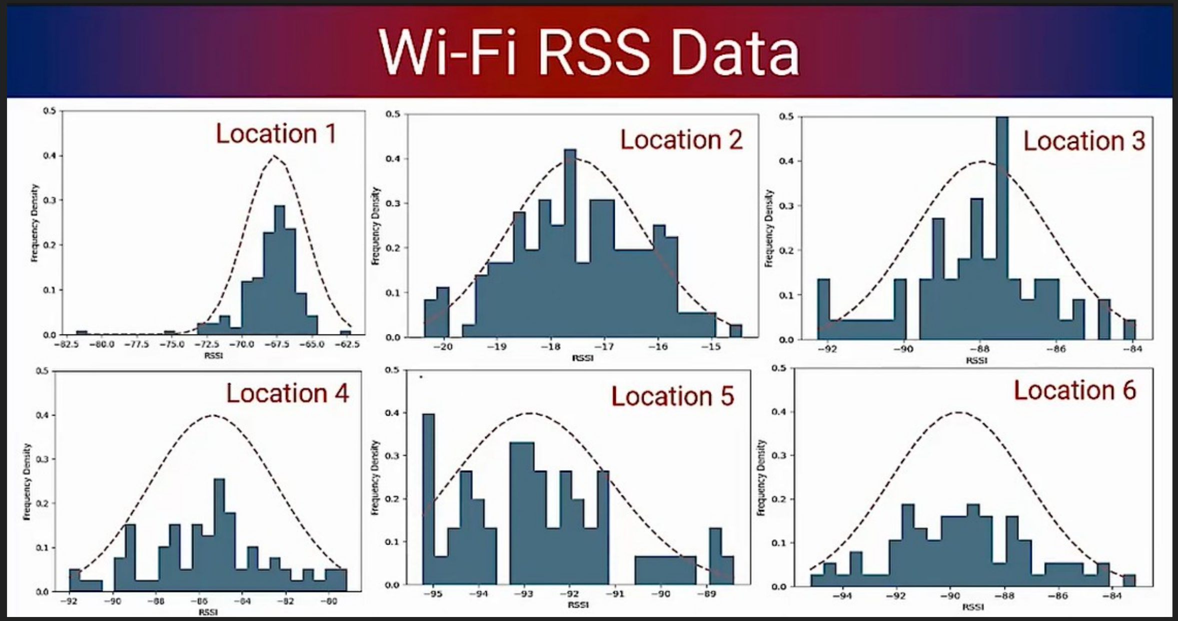
PL(d) & PL(d0) are path losses at distances d and d0.

- It is however important to note that the distance and PL calculations correspond to least square fit and hence not accurate. Looking at it physically, this happens due to noise - **shadowing/fading** due to the presence of obstacles.



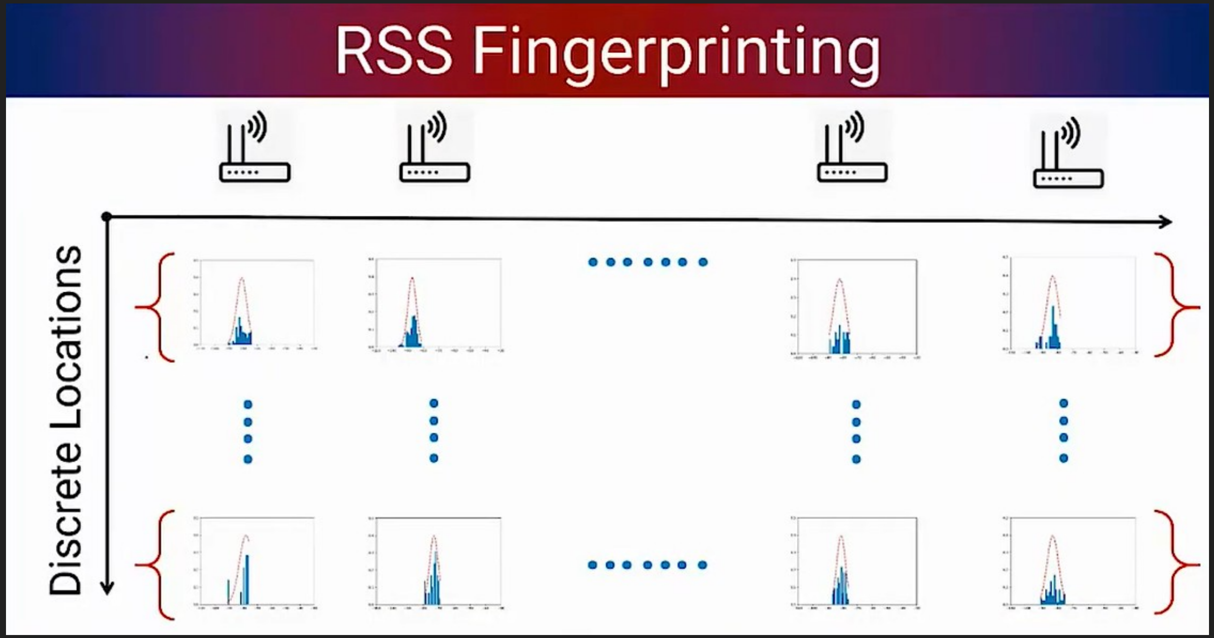
RSS FINGERPRINTING

- In the Free-space path loss model, a single value of RSS is chosen. This alternately means that we have assumed the signals coming from an **Access Point (AP)** to be composed majorly of a particular frequency and the RSS is calculated subsequently. However, in practice, corresponding to any AP, there is usually a distribution of frequencies within its bandwidth. In RSS fingerprinting, we consider all such frequency components that are present and have significant amplitude.
- This distribution of the frequency density along the y-axis and RSS (dB) along the x-axis looks like a Gaussian Distribution with a specific mean and standard deviation. This distribution is typically different for each pair of (Location, AP).



LOCALIZATION USING THE FINGERPRINTS

- The location is thus a function of the N-dimensional vector <RSS> : (RSS1, RSS2,....RSSN). With the distributions obtained at any discrete location for every AP, we form a matrix as shown. Here, for each of the discrete location-AP pair, there is a corresponding location and it is the fingerprint of that pair. N is the no. of discrete locations and it is set based on the accuracy and granularity requirements.
- Now, whenever a request/query is made, (a query is vector/row containing the RSS values at a random location from each of the APs requesting for the location information), it is compared with the matrix that has been constructed and stored in the database using a model. This model can either be parametric/ non-parametric. This model now outputs the required information.
- Supposing there are some locations where the RSS data is missing or in cases where there are continuous location requirements, techniques like interpolation can be used to estimate the RSS and its distribution.



Parametric and Non-parametric Models

Parametric Model

- This model is based on **Gaussian Naive Bayes Classification**.
- A **likelihood function** is exploited here and the implementation is based on classes (each discrete location is modelled as a class).

Non-parametric Model

- This model is based on **k-nearest neighbours**.
- Here, an input <RSS> vector is compared with each of the <RSS> vectors in the matrix and the nearest neighbour is found.
- This approach is one of the oldest and is costlier due to a large no. of comparisons.

Ranging using Commercial off the Shelf (COTS) Bluetooth Low Energy (BLE) beacons

- Estimote is one such company that manufactures BLE beacons. These can be interfaced with Android Application Programming Interface (API).
- The locations (co-ordinates) of the estimotes (beacons) are to be specified and this works based on the strength of Bluetooth packets.

THANK YOU