TU Delft

Wi Fi™ Fingerprinting

Group 8: Ming-Chieh Hu, Ákos Sárkány, Daan Schlosser, Neelabh Singh

WiFi Fingerprinting

Survey signal strengths at locations

Create a fingerprint / ID of the location

Classify new signals to determine users location

Implement and use classifiers:

- Linear Least Squares
- Maximum Likelihood
- Advanced Classifiers with `scikit-learn`

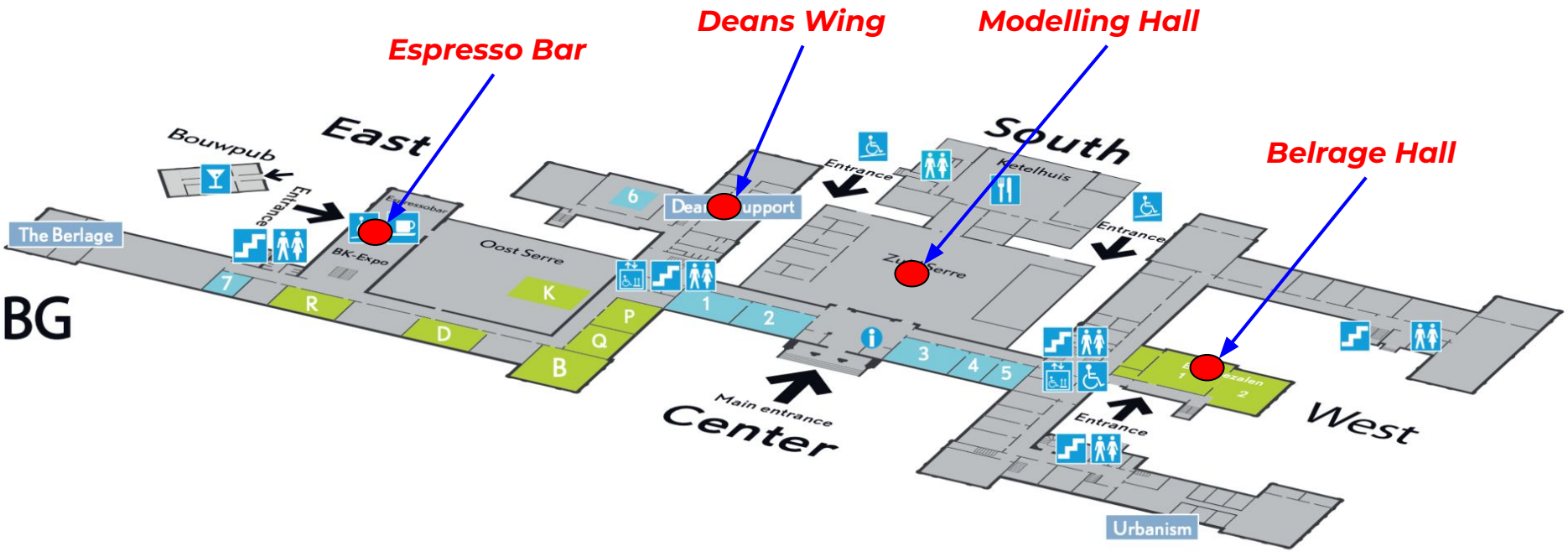
TU Delft

```
"id","MAC Address","SSID","APN","Connected AP","Signal S
1,"D0:4D:C6:F2:D9:B0","eduroam","","",-79,40,30,"Ch 116
1,"D0:4D:C6:F2:F0:F0","eduroam","","",-80,38,6,"Ch 44 [5
1,"D0:4D:C6:F2:C9:40","eduroam","","",-68,72,0,"Ch 6 [2.
1,"D0:4D:C6:F2:F0:E0","eduroam","","",-78,43,2,"Ch 6 [2.
1,"D0:4D:C6:F2:8D:D0","eduroam","","",-72,62,5,"Ch 132 |
1,"1C:28:AF:62:9E:30","eduroam","","",-73,60,0,"Ch 56 [5
1,"1C:28:AF:62:55:D0","eduroam","","Connected AP",-51,88
1,"D0:4D:C6:F2:8D:F0","eduroam","","",-84,29,6,"Ch 56 [5
1,"1C:28:AF:65:F5:40","eduroam","","",-78,43,0,"Ch 1 [2.
1,"1C:28:AF:66:03:B0","eduroam","","",-80,38,11,"Ch 161
1,"D0:4D:C6:F3:0B:40","eduroam","","",-65,80,0,"Ch 1 [2.
1,"1C:28:AF:62:55:C0","eduroam","","",-59,83,2,"Ch 1 [2.
1,"D0:4D:C6:F2:C9:50","eduroam","","",-73,60,0,"Ch 161 |
1,"1C:28:AF:62:9E:20","eduroam","","",-69,70,0,"Ch 11 [2
1,"D0:4D:C6:F3:0B:50","eduroam","","",-59,83,3,"Ch 108 |
1,"1C:28:AF:61:EE:20","eduroam","","",-67,75,0,"Ch 11 [2
1,"D0:4D:C6:F2:8D:C0","eduroam","","",-74,57,0,"Ch 1 [2.
1,"D0:4D:C6:F2:8F:20","eduroam","","",-77,44,1,"Ch 6 [2.
1,"D0:4D:C6:F2:B9:E0","eduroam","","",-78,43,1,"Ch 1 [2.
1,"D0:4D:C6:F2:64:D0","eduroam","","",-82,33,27,"Ch 64 |
1,"1C:28:AF:66:10:30","eduroam","","",-80,38,5,"Ch 149 |
1,"1C:28:AF:65:F5:E0","eduroam","","",-78,43,1,"Ch 11 [2
1,"D0:4D:C6:F3:07:F0","eduroam","","",-77,60,5,"Ch 36 [5
1,"1C:28:AF:65:F5:50","eduroam","","",-81,35,23,"Ch 108
1,"1C:28:AF:65:F9:F0","eduroam","","",-80,38,14,"Ch 48 |
1,"1C:28:AF:65:FF:F0","eduroam","","",-80,38,13,"Ch 52 |
1,"D0:4D:C6:F2:E3:60","eduroam","","",-80,38,1,"Ch 1 [2.
1,"1C:28:AF:65:EA:50","eduroam","","",-83,31,15,"Ch 104
```

Motivation:

Unreliability of indoor GNSS Positioning

A1. Gnss Accuracy In Indoor - BK city



- GNSS data collected at above locations.
- GNSS Data was converted to CRS of EPSG:3857 with units in metres to perform statistics in metres.

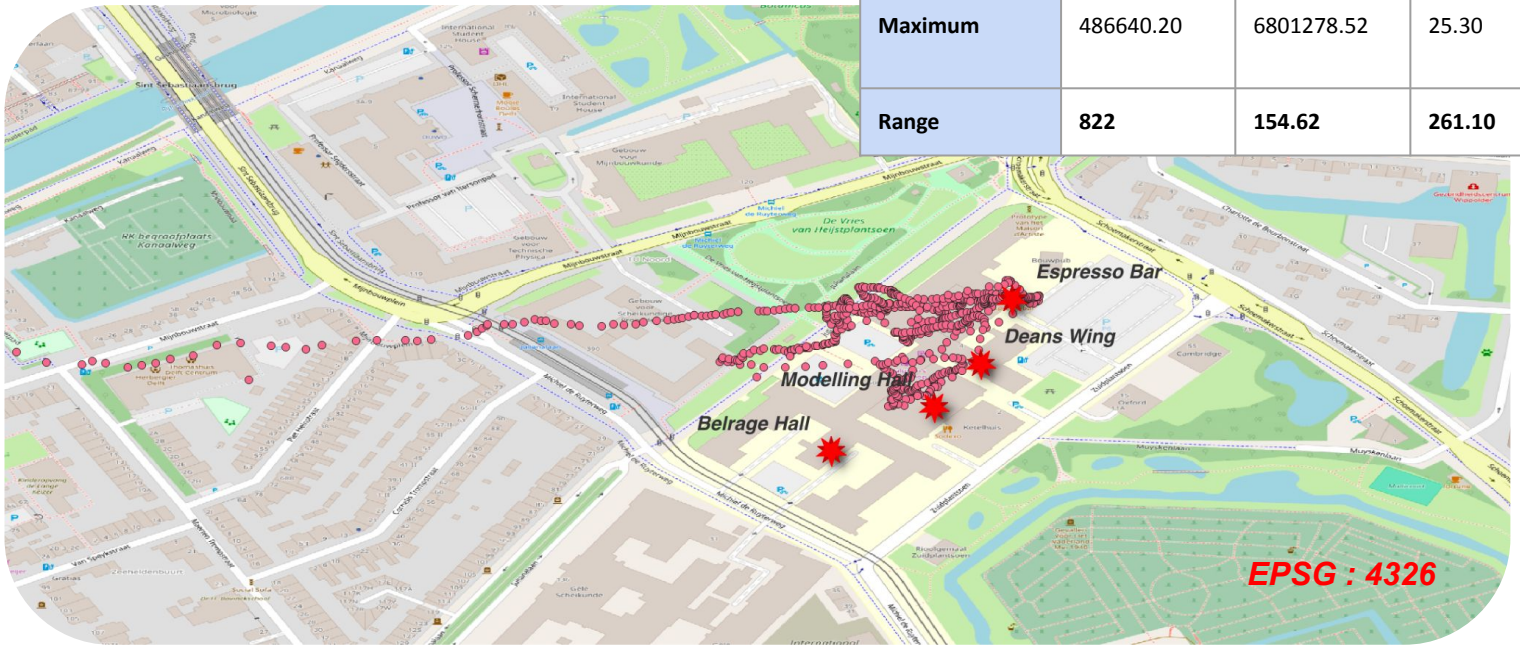
TU Delft

Espresso Bar

- Difference between x,y,z of mean and estimated values are 88.34m, 27.46m and 81.75m respectively.
- Highest DOP value is more than 10.

	X projected (m)	Y projected (m)	Z (m)	PDOP	HDOP	VDOP
Mean	486531.37	6801227.75	-79.95	1.70	0.92	1.37
Esti. value /difference	**486619.71 / 88.34**	**6801255.21 / 27.46**	**1.8 / 81.75**			
Standard Deviation	116.18	37.41	47.72	0.85	0.66	0.57
Minimum	485817.92	6801123.90	-235.80	1.20	0.60	0.80
Maximum	486640.20	6801278.52	25.30	16.20	13.40	9.20
Range	**822**	**154.62**	**261.10**	**15.00**	**12.80**	**8.20**

Scatter Plot of GNSS measurements at Espresso Bar

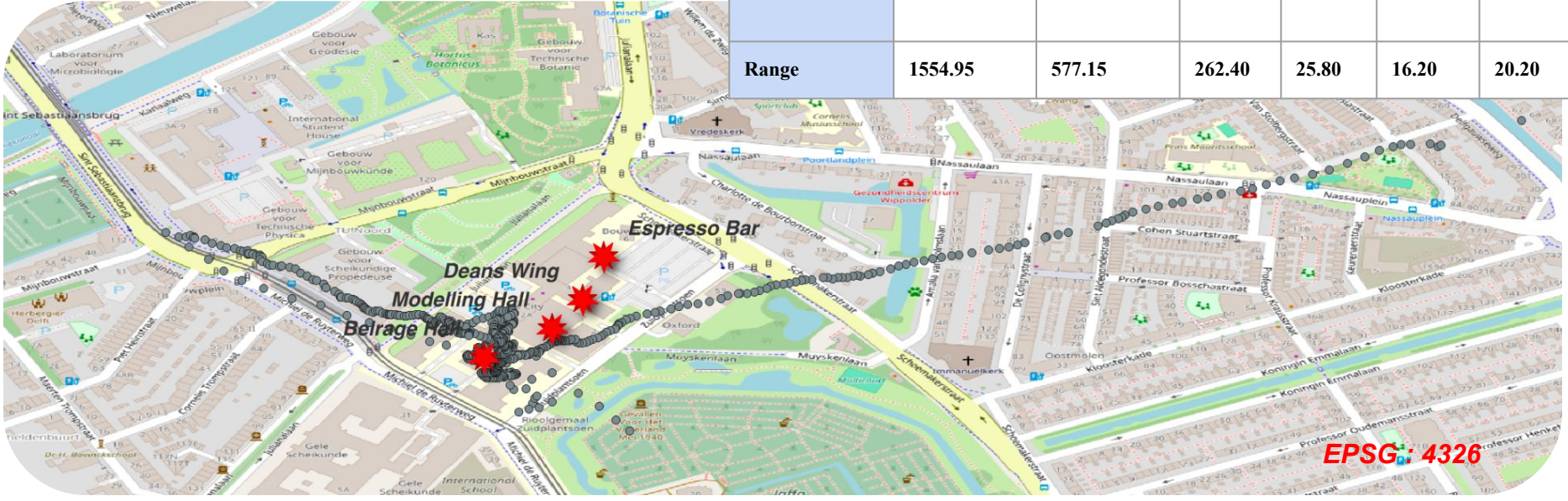


Berlage Hall 1

- Difference between x,y,z of mean and estimated values are -48.64m, -55.60m and 73.51m respectively.
- Highest DOP value is more than 20.

	X projected (m)	Y projected (m)	Z (m)	PDOP	HDOP	VDOP
Mean	486531.59	6801124.31	-73.01	2.22	1.22	1.79
Esti value difference	**486482.95 / -48.64**	**6801068.71 / -55.60**	**0.5 / 73.51**			
Standard Deviation	213.57	78.55	45.78	1.81	1.21	1.37
Minimum	486117.93	6800931.32	-198.80	1.20	0.60	0.80
Maximum	487672.88	6801508.49	63.60	27.00	16.80	21.00
Range	**1554.95**	**577.15**	**262.40**	**25.80**	**16.20**	**20.20**

Scatter Plot of GNSS measurements at Belrage Hall



Dean's Wing

- Difference between x,y,z of mean and estimated values are 26.09m, 0.97m and 38.83m respectively.
- Highest DOP value is more than 3.

	X projected (m)	Y projected (m)	Z (m)	PDOP	HDOP	VDOP
Mean	486621.68	6801176.78	-36.33	2.0	1.15	1.54
Esti. value /difference	**486595.59 / 26.09**	**6801175.81/ 0.97**	**2.5 / 38.83**			
Standard Deviation	38.0	32.97	34.36	0.75	0.36	0.67
Minimum	486528.88	6801122.10	-132.60	1.20	0.80	0.80
Maximum	486704.58	6801286.96	22.90	3.60	2.0	3.0
Range	**175.70**	**164.86**	**155.50**	**2.40**	**1.20**	**2.20**

Scatter Plot of GNSS measurements at Deans Wing

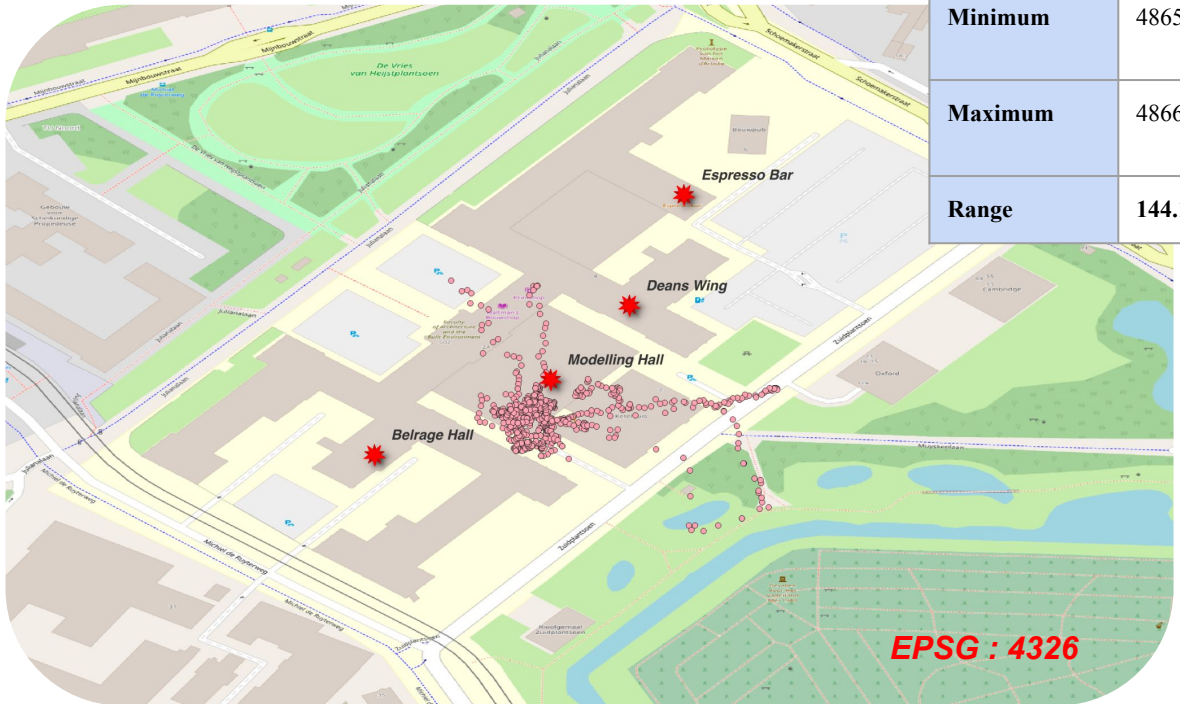
- ***With the above examples we can conclude that GNSS measurement for indoor positioning are highly inaccurate and could <u>not</u> be relied upon for this use.***
- ***Also, in many cases we saw that the receiver was not able to connect with GNSS and receive data. Especially, when inside halls with no direct connection with the satellites. Leading to none or very high DOP values .***

Modelling Hall

- Difference between x,y,z of mean and estimated values are -2.87m, 26.21m and 69.89m respectively.
- Highest DOP value is more than 10.

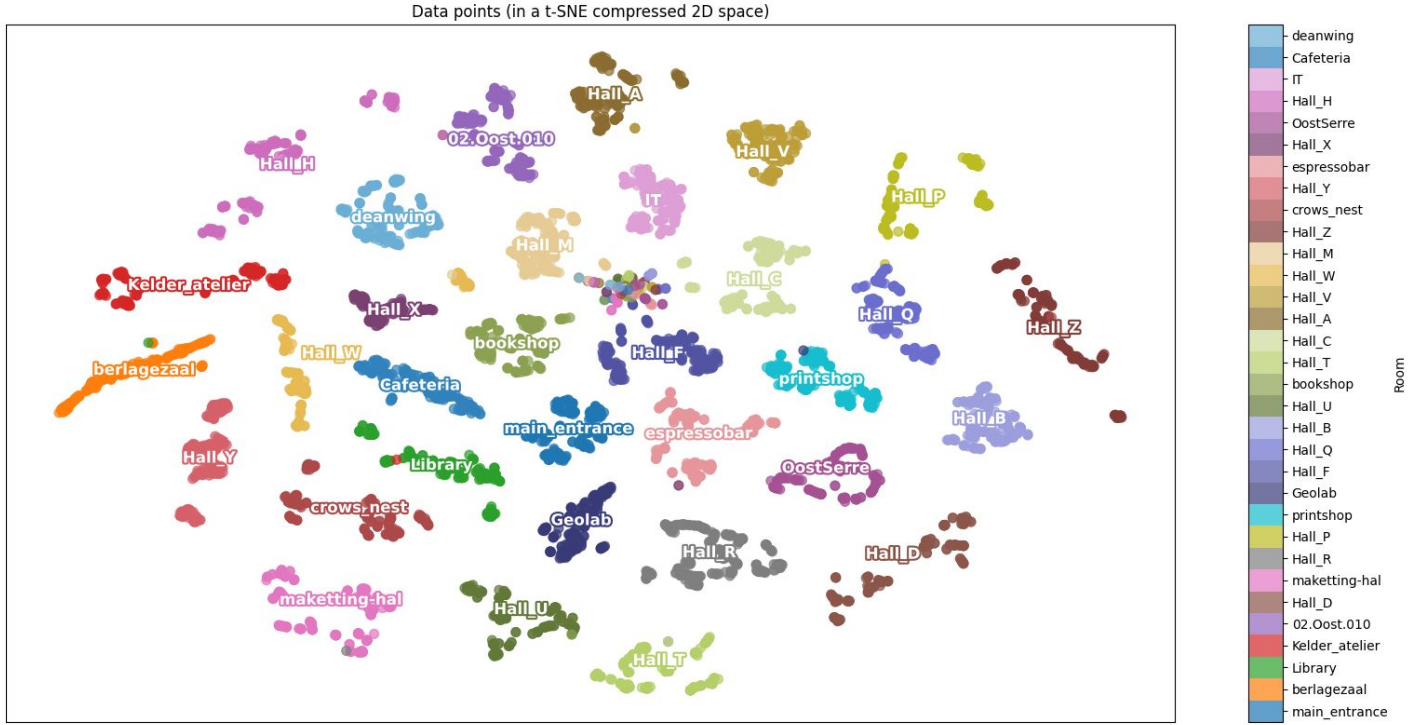
	X projected (m)	Y projected (m)	Z (m)	PDOP	HDOP	VDOP
Mean	486563.64	6801096.45	-69.49	1.30	0.72	1.03
Estimated and difference	**486560.77 / -2.87**	**6801122.66 / 26.21**	**0.4 / 69.89**			
Standard Deviation	28.03	20.73	35.48	0.64	0.45	0.46
Minimum	486516.83	6801013.90	-214.30	0.80	0.40	0.60
Maximum	486660.98	6801193.83	108.00	11.80	8.40	8.20
Range	**144.15**	**179.93**	**322.30**	**11.00**	**8.00**	**7.60**

Scatter Plot of GNSS measurements at Modelling Hall



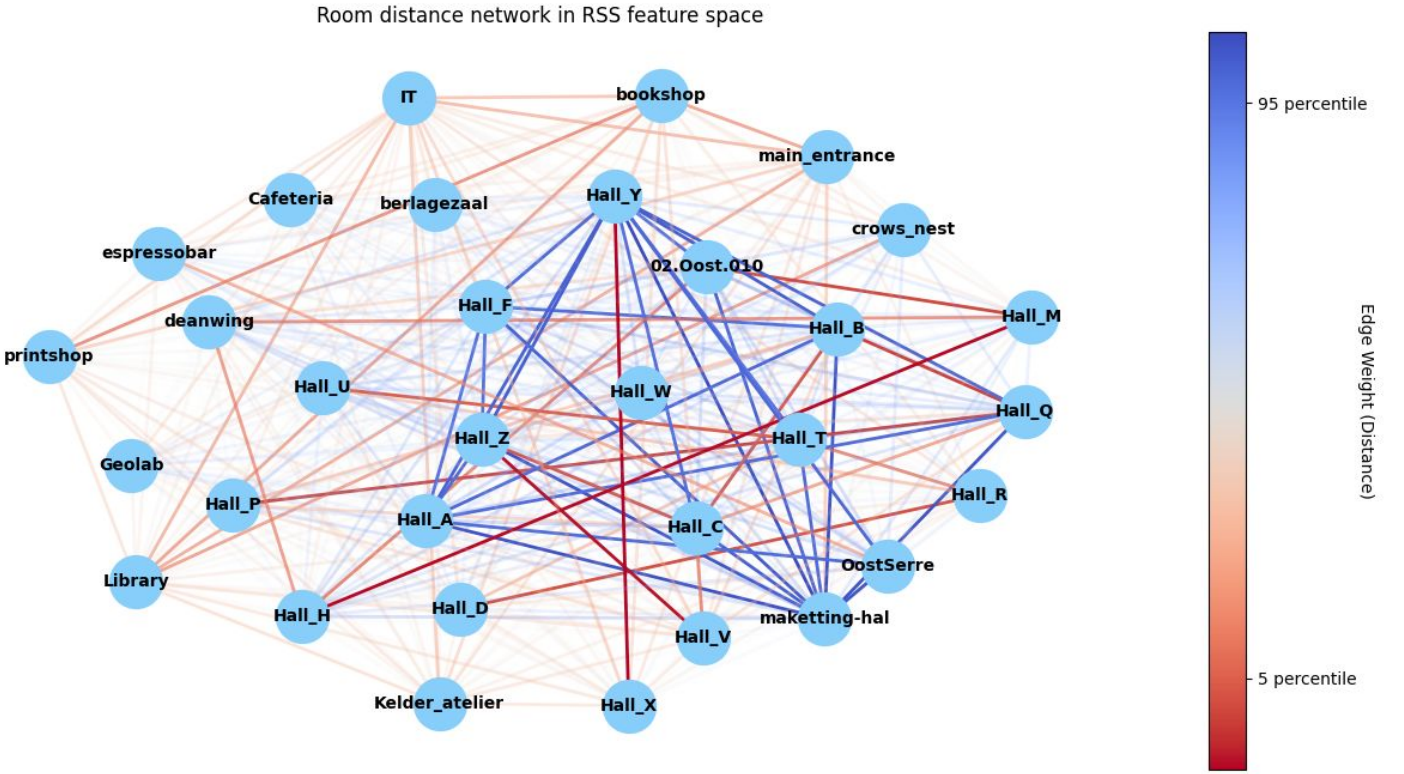
Indoor GNSS data doesn't have any of the following

A2. RSS-space Visualization



TU Delft

A2. RSS-space Distance



TU Delft

Before our analysis, we created a Wi-Fi fingerprint database by cleaning the HomeDale measurements:

- Retained only “eduroam” entries with signal strength > -85 dBm.
- Standardized timestamps to relative times.
- Computed average signal strength and standard deviation per MAC address per location for the Wi-Fi fingerprint.

TU Delft

Before (measurements as taken by HomeDale):

Timestamp	MAC Address	Vendor	SSID	Access Point Name	Connected AP	Signal Strength	Quality	Station Count	Frequency	Position	Info	Adapter
06-12-24 11:28	1C:28:AF:61:FB:E2	Hewlett Packard Enterprise	TUD-facility			-82	33	0	Ch 11 [2.462 GHz]		0 Stations, 8% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	1C:28:AF:61:FB:E0	Hewlett Packard Enterprise	eduroam			-80	38	0	Ch 11 [2.462 GHz]		0 Stations, 8% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	1C:28:AF:61:E3:A2	Hewlett Packard Enterprise	TUD-facility			-80	38	0	Ch 1 [2.412 GHz]		0 Stations, 7% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	1C:28:AF:5B:2F:62	Hewlett Packard Enterprise	TUD-facility			-80	38	0	Ch 6 [2.437 GHz]		0 Stations, 7% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	D0:4D:C6:F2:C8:A0	Hewlett Packard Enterprise	eduroam			-67	75	0	Ch 6 [2.437 GHz]		0 Stations, 10% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	D0:4D:C6:F2:3A:B2	Hewlett Packard Enterprise	TUD-facility			-77	46	0	Ch 153 [5.765 GHz]		0 Stations, 1% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	D0:4D:C6:F2:3A:B1	Hewlett Packard Enterprise	tudelft-dastud			-76	50	1	Ch 153 [5.765 GHz]		1 Station, 1% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	D0:4D:C6:F2:3A:B0	Hewlett Packard Enterprise	eduroam			-77	46	2	Ch 153 [5.765 GHz]		2 Stations, 1% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	D0:4D:C6:F2:C8:B2	Hewlett Packard Enterprise	TUD-facility			-70	67	0	Ch 52 [5.260 GHz]		0 Stations, 1% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	D0:4D:C6:F2:C8:B1	Hewlett Packard Enterprise	tudelft-dastud			-70	67	0	Ch 52 [5.260 GHz]		0 Stations, 1% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	D0:4D:C6:F2:C8:B0	Hewlett Packard Enterprise	eduroam			-70	67	1	Ch 52 [5.260 GHz]		1 Station, 1% Channel Utilization	Wi-Fi [Intel(
06-12-24 11:28	D0:4D:C6:F3:0D:52	Hewlett Packard Enterprise	TUD-facility			-82	33	0	Ch 116 [5.580 GHz]		0 Stations, 1% Channel Utilization	Wi-Fi [Intel(

After (measurements after cleaning up):

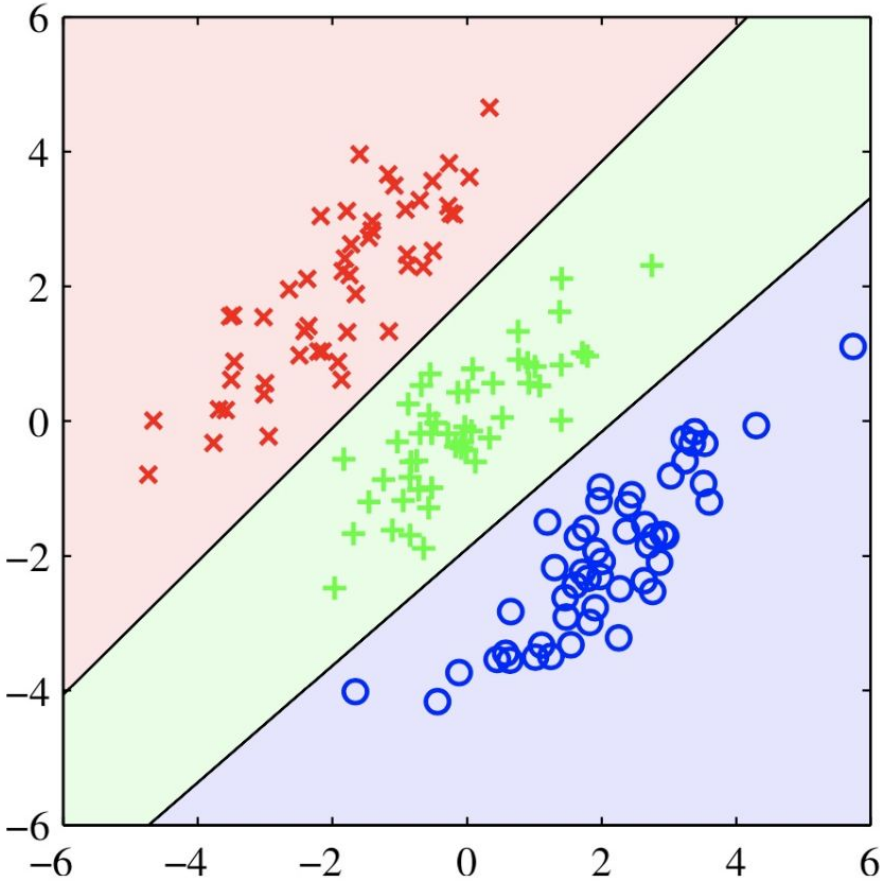
	Timestamp	MAC Address	SSID	Signal Strength
0		1C:28:AF:61:FB:E0	eduroam	-80
0		D0:4D:C6:F2:C8:A0	eduroam	-67
0		D0:4D:C6:F2:3A:B0	eduroam	-77
0		D0:4D:C6:F2:C8:B0	eduroam	-70
0		D0:4D:C6:F3:0D:50	eduroam	-82
0		D0:4D:C6:F2:D1:40	eduroam	-82
0		D0:4D:C6:F2:3A:A0	eduroam	-64
0		D0:4D:C6:F2:D9:50	eduroam	-77
0		1C:28:AF:5B:29:D0	eduroam	-78
0		1C:28:AF:5C:0C:70	eduroam	-79
0		D0:4D:C6:F2:D4:C0	eduroam	-83
0		1C:28:AF:5B:29:C0	eduroam	-77

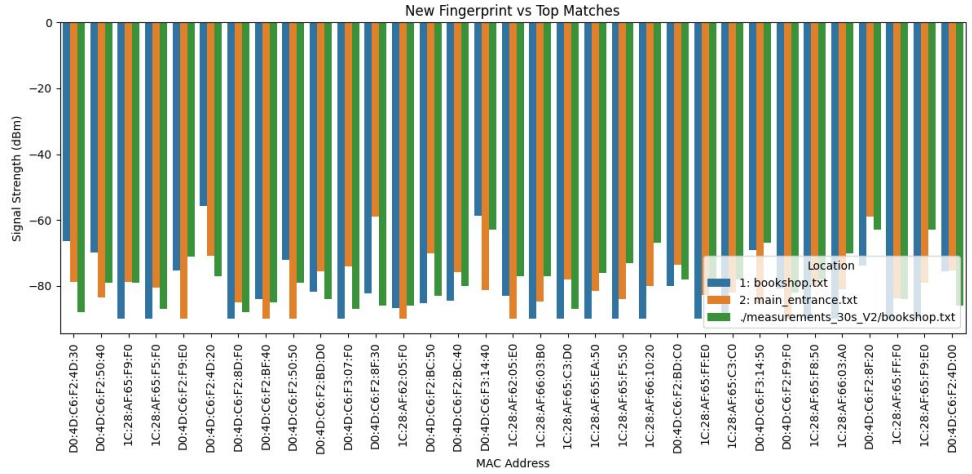
Wi-Fi Fingerprint database:

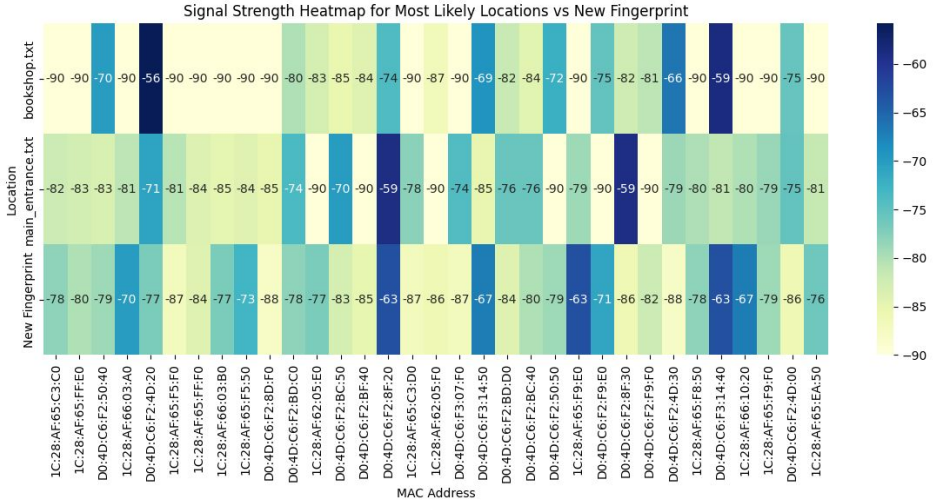
		02_Oost.010.txt	berlagezaal.txt	bookshop.txt	Cafeteria.txt	crows_nest.txt	deanwing.txt	espressobar.txt	Geolab.txt	Hall_A.txt	Hall_B.txt	Hall_C.txt	Hall_D.txt	Hall_F.txt	Hall_H.txt	Hall_M.txt	Hall_P.txt	Hall_Q.txt	Hall_R.txt	Hall_T.txt
38	1C:28:AF:65:F8:40	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
39	1C:28:AF:5B:83:F0	-86.511	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-83.659	NaN	NaN	NaN	NaN	-82.175	NaN	NaN	NaN
40	1C:28:AF:5B:EE:B0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
41	D0:4D:C6:F2:C3:D0	NaN	NaN	NaN	NaN	-78.106	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
42	D0:4D:C6:F2:D9:40	-77.37	NaN	NaN	NaN	NaN	-69.508	NaN	NaN	-80.118	NaN	NaN	NaN	NaN	-71.826	-66.333	NaN	NaN	NaN	NaN
43	D0:4D:C6:F2:F8:F0	NaN	-80.2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
44	D0:4D:C6:F1:D2:00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-67.548	NaN	NaN	NaN	NaN	NaN	NaN
45	D0:4D:C6:F3:0D:50	-60.154	NaN	NaN	NaN	NaN	-83.856	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-78.862	-73.159	NaN	NaN	NaN	NaN
46	D0:4D:C6:F2:D1:50	-58.549	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-73.952	-78.28	NaN	NaN	NaN	NaN
47	D0:4D:C6:F2:99:40	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-76.236	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-76.0
48	1C:28:AF:5B:FA:30	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-85.102	NaN	NaN	NaN	NaN	-81.205	-85.652	NaN	NaN
49	D0:4D:C6:F2:D2:D0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-85.343	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-84.0	NaN	NaN
50	1C:28:AF:5C:0C:70	NaN	NaN	NaN	NaN	NaN	-83.495	NaN	NaN	-78.86	NaN	NaN	NaN	NaN	-81.577	-88.125	NaN	NaN	NaN	NaN
51	D0:4D:C6:F2:F2:10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-68.349	-52.329	NaN	NaN	NaN	NaN	-85.661	-76.806	NaN	NaN
52	D0:4D:C6:F2:B6:60	NaN	NaN	NaN	NaN	NaN	-86.243	NaN	NaN	-77.882	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
53	D0:4D:C6:F2:9B:D0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN913	NaN	NaN	NaN	NaN	NaN	NaN
54	D0:4D:C6:F2:BC:50	NaN	-81.0	-85.337	-82.0	NaN	-84.0	-79.855	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
55	D0:4D:C6:F2:B9:F0	NaN	NaN	NaN	NaN	NaN	-78.575	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
56	D0:4D:C6:F2:C8:B0	-75.469	NaN	NaN	NaN	NaN	-73.367	NaN	NaN	-86.092	NaN	NaN	NaN	NaN	-63.714	-52.646	NaN	NaN	NaN	NaN
57	D0:4D:C6:F2:9A:60	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-78.771	-71.036	NaN	NaN	NaN	NaN	-65.151	-59.741	NaN	NaN
58	1C:28:AF:5B:58:30	-86.488	NaN	NaN	NaN	NaN	-44.734	NaN	NaN	-85.742	NaN	NaN	NaN	NaN	-68.407	-81.947	NaN	NaN	NaN	NaN
59	1C:28:AF:5C:01:10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-86.316
60	D0:4D:C6:F2:7B:F0	NaN	NaN	NaN	NaN	-83.547	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
61	1C:28:AF:5B:39:D0	NaN	NaN	NaN	-82.174	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
62	D0:4D:C6:F2:A9:90	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	-75.234	-88.425	-83.297	NaN	NaN	NaN	NaN	-82.335	-67.679	NaN

B1. Linear Least Square Classification

$$S = \sum_{i=1}^{n} r_i^2.$$

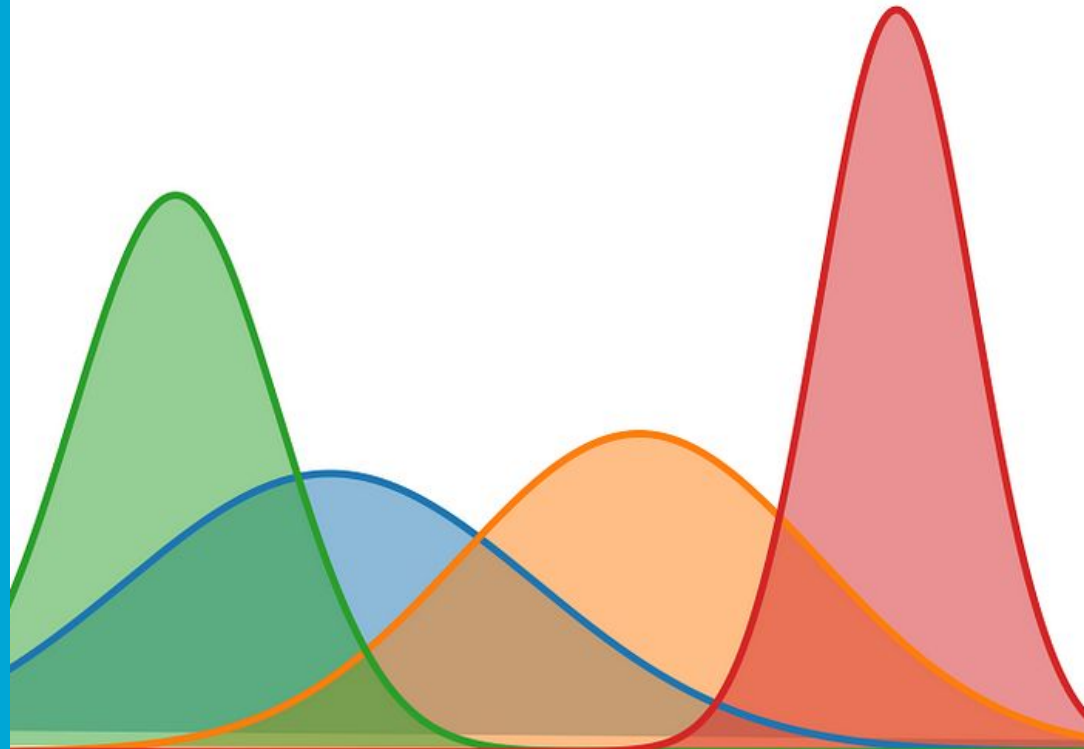



New Fingerprint vs Top Matches
Signal Strength (dBm)
0
−20
−40
−60
−80
Location
1: bookshop.txt
2: main_entrance.txt
./measurements_30s_V2/bookshop.txt
D0:4D:C6:F2:4D:30
D0:4D:C6:F2:50:40
1C:28:AF:65:F9:F0
1C:28:AF:65:F5:F0
D0:4D:C6:F2:F9:E0
D0:4D:C6:F2:4D:20
D0:4D:C6:F2:8D:F0
D0:4D:C6:F2:BF:40
D0:4D:C6:F2:50:50
D0:4D:C6:F2:BD:D0
D0:4D:C6:F3:07:F0
D0:4D:C6:F2:8F:30
1C:28:AF:62:05:F0
D0:4D:C6:F2:BC:50
D0:4D:C6:F2:BC:40
D0:4D:C6:F3:14:40
1C:28:AF:62:05:E0
1C:28:AF:66:03:B0
1C:28:AF:65:C3:D0
1C:28:AF:65:EA:50
1C:28:AF:65:F5:50
1C:28:AF:66:10:20
D0:4D:C6:F2:BD:C0
1C:28:AF:65:FF:E0
1C:28:AF:65:C3:C0
D0:4D:C6:F3:14:50
D0:4D:C6:F2:F9:F0
1C:28:AF:65:F8:50
1C:28:AF:66:03:A0
D0:4D:C6:F2:8F:20
1C:28:AF:65:FF:F0
1C:28:AF:65:F9:E0
D0:4D:C6:F2:4D:00
MAC Address


Signal Strength Heatmap for Most Likely Locations vs New Fingerprint
Location
bookshop.txt
main_entrance.txt
New Fingerprint
-90 -90 -70 -90 -56 -90 -90 -90 -90 -90 -80 -83 -85 -84 -74 -90 -87 -90 -69 -82 -84 -72 -90 -75 -82 -81 -66 -90 -59 -90 -90 -75 -90
-82 -83 -83 -81 -71 -81 -84 -85 -84 -85 -74 -90 -70 -90 -59 -78 -90 -74 -85 -76 -76 -90 -79 -90 -59 -90 -79 -80 -81 -80 -79 -75 -81
-78 -80 -79 -70 -77 -87 -84 -77 -73 -88 -78 -77 -83 -85 -63 -87 -86 -87 -67 -84 -80 -79 -63 -71 -86 -82 -88 -78 -63 -67 -79 -86 -76
−60
−65
−70
−75
−80
−85
−90
1C:28:AF:65:C3:C0
1C:28:AF:65:FF:E0
D0:4D:C6:F2:50:40
1C:28:AF:66:03:A0
D0:4D:C6:F2:4D:20
1C:28:AF:65:F5:F0
1C:28:AF:65:FF:F0
1C:28:AF:66:03:B0
1C:28:AF:65:F5:50
D0:4D:C6:F2:8D:F0
D0:4D:C6:F2:8D:C0
1C:28:AF:62:05:E0
D0:4D:C6:F2:BC:50
D0:4D:C6:F2:BF:40
D0:4D:C6:F2:8F:20
1C:28:AF:65:C3:D0
1C:28:AF:62:05:F0
D0:4D:C6:F3:07:F0
D0:4D:C6:F2:F9:F0
D0:4D:C6:F2:BD:D0
D0:4D:C6:F2:BC:40
D0:4D:C6:F2:50:50
1C:28:AF:65:F9:E0
D0:4D:C6:F2:F9:E0
D0:4D:C6:F2:8F:30
D0:4D:C6:F2:F9:F0
D0:4D:C6:F2:4D:30
1C:28:AF:65:F8:50
D0:4D:C6:F3:14:40
1C:28:AF:66:10:20
1C:28:AF:65:F9:F0
D0:4D:C6:F2:4D:00
1C:28:AF:65:EA:50
MAC Address

TU Delft

B1. Maximum Likelihood Classification (MLC)



TU Delft

MLC — Concept

Assume the *Signal Strength* from an antenna follows **Normal Distribution** curve

From 15 min measurements compute necessary values for each location (training):

- Mean **(μ)**: *The average signal strength during survey period per antenna*
- Standard Deviation **(σ)**: *The standard deviation of the values from the mean*
- Significance weight **(ω)**: *The number of logs the antenna is present relative to the total amount of logs during the surveyed period.*

Given a log from an unknown location:

- for each known location compute how likely it is that the unknown location is that location

MLC — Processed Data

```
"Location","MAC","Significance","Mean","Standard Deviation"
"02.Oost.010","1C:28:AF:5B:2F:60",1.0,-76.47428571428571,1.8041765604353943
"02.Oost.010","1C:28:AF:5B:58:20",1.0,-80.94285714285714,2.5314350209527645
"02.Oost.010","1C:28:AF:5B:29:D0",1.0,-73.66857142857143,2.521651548214433
"02.Oost.010","D0:4D:C6:F2:B9:A0",1.0,-80.34285714285714,2.453485650403257
"02.Oost.010","D0:4D:C6:F2:FC:C0",1.0,-45.29142857142857,2.5299770104626225
"02.Oost.010","D0:4D:C6:F3:0D:40",1.0,-57.76571428571429,3.2279089080033447
"02.Oost.010","1C:28:AF:61:E3:A0",1.0,-85.68571428571428,1.9620896815302171
"02.Oost.010","D0:4D:C6:F2:D1:40",1.0,-58.114285714285714,1.781996129087163
"02.Oost.010","D0:4D:C6:F2:C8:A0",1.0,-71.21142857142857,2.535662373714093
"02.Oost.010","D0:4D:C6:F2:FC:D0",1.0,-45.92,1.0925200226998129
"02.Oost.010","D0:4D:C6:F2:D1:50",1.0,-58.54857142857143,2.813576415329422
"02.Oost.010","D0:4D:C6:F3:0D:50",1.0,-60.15428571428571,2.514909013650151
"02.Oost.010","D0:4D:C6:F2:98:D0",1.0,-66.86857142857143,1.6732810247400471
"02.Oost.010","D0:4D:C6:F2:D9:50",1.0,-83.44,1.6126641666155153
"02.Oost.010","D0:4D:C6:F2:C8:B0",1.0,-75.46857142857142,1.8323710523443084
"02.Oost.010","D0:4D:C6:F2:B9:B0",1.0,-81.61714285714285,1.828892828897264
"02.Oost.010","D0:4D:C6:F2:D4:D0",1.0,-69.71428571428571,1.851199280334293
"02.Oost.010","1C:28:AF:5B:29:C0",0.9934640522875817,-70.48275862068965,2.1674820868
"02.Oost.010","D0:4D:C6:F2:98:C0",1.0,-62.61142857142857,2.9622164894909764
"02.Oost.010","D0:4D:C6:F2:D4:C0",1.0,-65.58857142857143,2.8146323209330237
"02.Oost.010","D0:4D:C6:F2:D9:40",0.9869281045751634,-77.3699421965318,2.74580468155
```

TU Delft

MLC — Computing Likelihood

$$pdf(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

For each antenna at a known location, and for each antenna at the unknown location compute the probability the unknown signal is coming from that antenna:

- If MAC addresses **do not** match: the probability is *zero.*
- If MAC addresses **do** match:
  - μ_k and σ_k is the mean and std. dev. of the antenna at the known location
  - compute μ_u and σ_u for the antenna at the unknown location, then combine σ_k and σ_u to obtain the overall std. dev.: σ_c
  - use μ_k as μ, μ_u as x, and σ_c as σ in *pdf* equation to obtain the probability that the signal is coming from the given location
  - use the antenna's significance weight ω_a to weight the probability when combining the probabilities derived for each antenna at the known location

MLC — Evaluation

30 second samples from 14 locations. 9 *versions* of probability combination

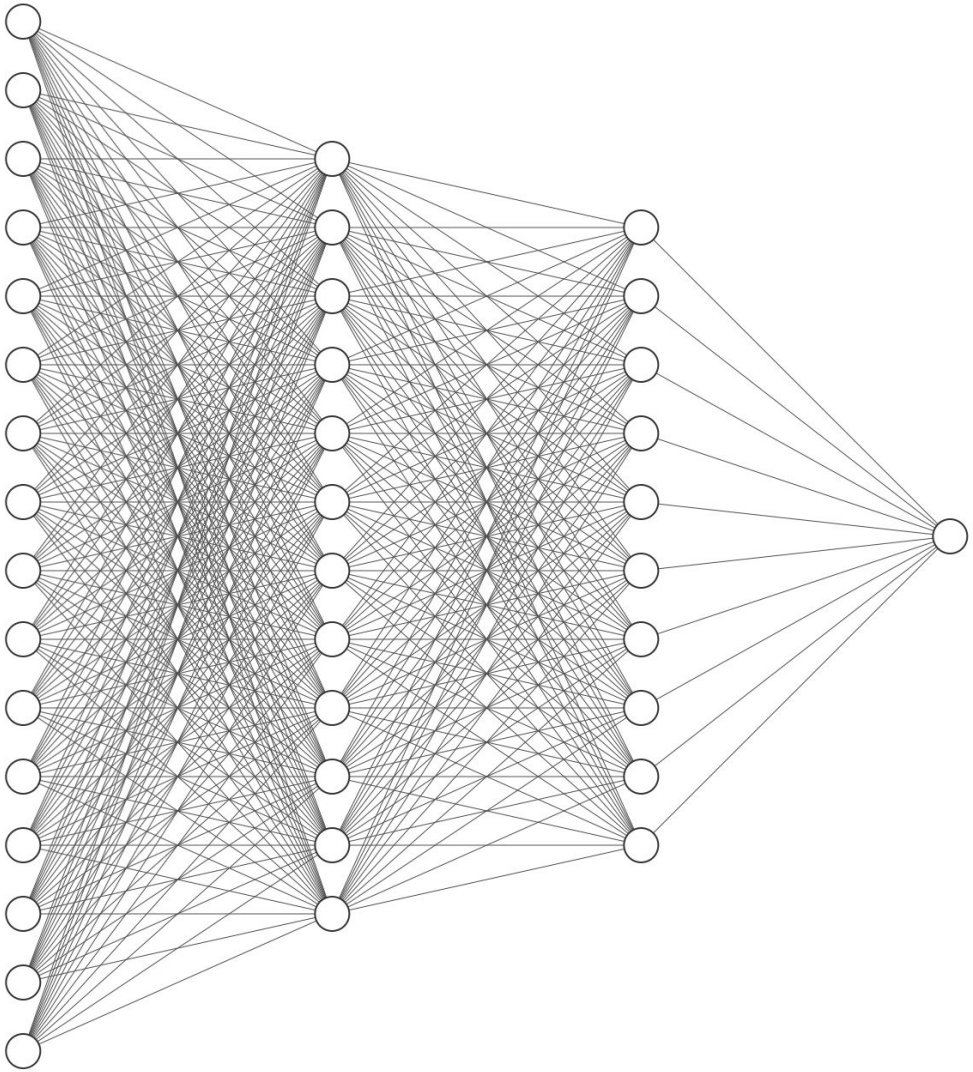
Overall accuracy: **66.66% — 73.33%**

Main errors: Hall A, Library, @HOK

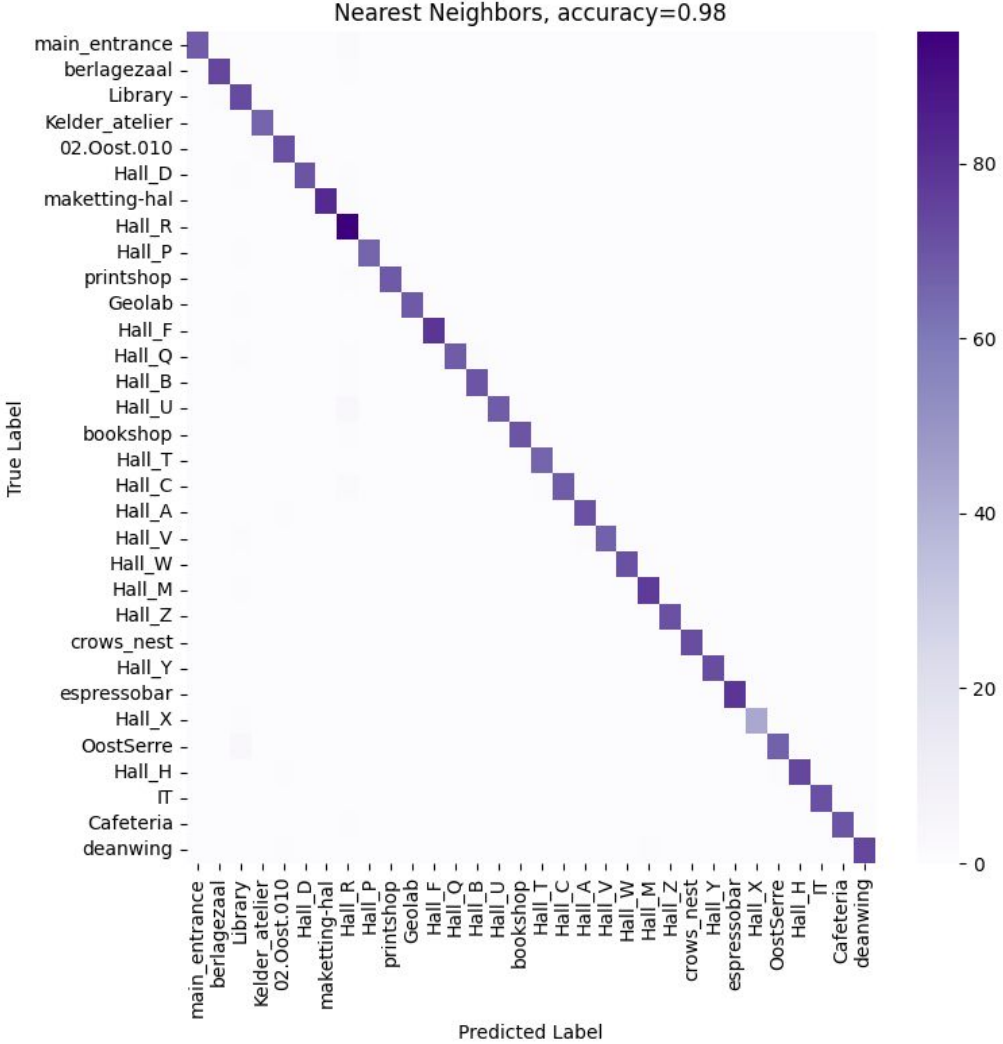
Conclusions:

- For each location there was at least one *version* which yields the correct result
- For maximum likelihood it is not the best approach to survey a room in a single position. By standing still, the signal strength only influenced by random noise, not by actual position
- Spanning larger area at a location while surveying provides higher standard deviation in the *pdf.* If one location's *pdf* overlaps with another (i.e. they are physically close) the one with less noise (smaller σ) wins

Advanced Classification with `scikit-learn`



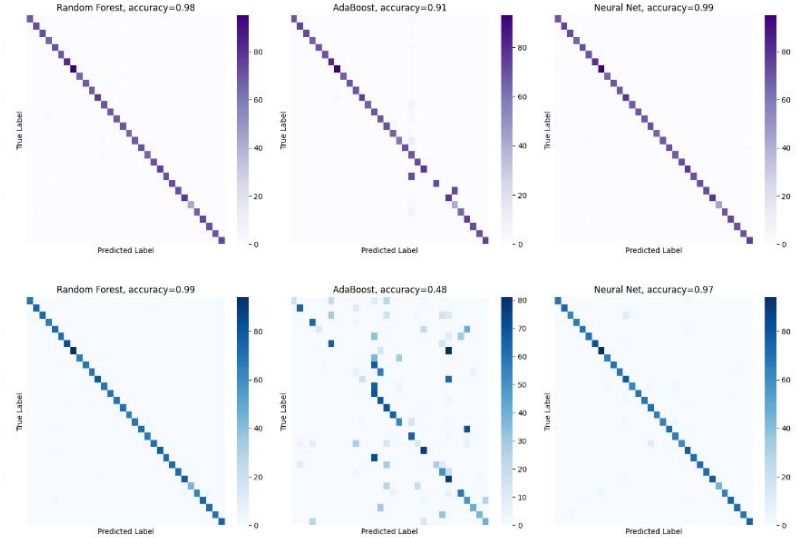
TU Delft



Random Forest

Adaptive Boosting

Neural Net



uncompressed 379D, Blue: t-SNE reduces to 3D

Accuracy Test

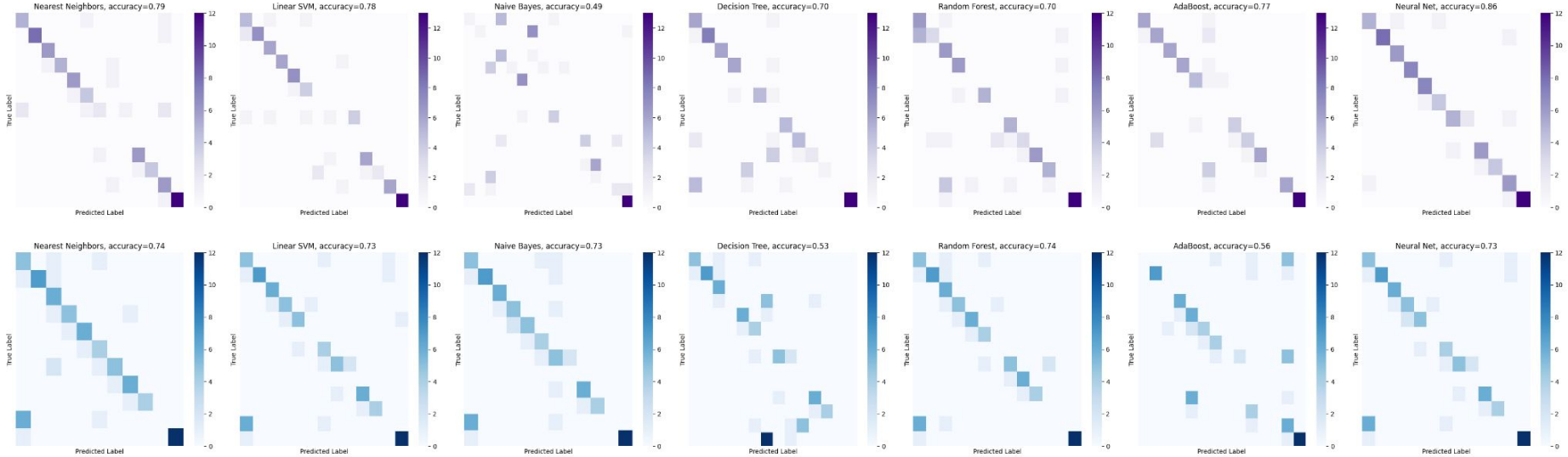
Feature Space Dimension	Nearest Neighbors	Linear SVM	Naive Bayes	Decision Tree	Random Forest	Adaptive Boosting	Neural Net
379	0.79	0.78	0.49	0.70	0.70	0.77	0.86
3 (t-SNE)	0.74	0.73	0.73	0.53	0.74	0.56	0.73
	Simple, works well		**Fine in 3D**	**Have to take care of more parameters**			**Overkill**

```
CLASSIFIERS = [
    KNeighborsClassifier(10, weight="distance"),
    SVC(kernel="linear", C=0.01, random_state=42),
    GaussianNB(),
    DecisionTreeClassifier(criterion="entropy", max_depth=20, random_state=42),
    RandomForestClassifier(criterion="entropy", max_depth=10, n_estimators=10, max_features=1, random_state=42),
    AdaBoostClassifier(algorithm="SAMME", random_state=42, n_estimators=200),
    MLPClassifier(alpha=1, max_iter=1000, learning_rate_init=0.001, learning_rate="adaptive", random_state=42),
]
```


These classification were done using test dataset (an entirely separate measurement)

Accuracy Test: 11/32 locations

Nearest Neighbors	Linear SVM	Naive Bayes	Decision Tree	Random Forest	Adaptive Boosting	Neural Net



TU Delft

Purple: uncompressed 379D, Blue: t-SNE reduces to 3D