SURF-2022076 Kickoff Meeting: Scalable Representation of RSSIs for Multi-Building and Multi-Floor Indoor Localization Based on Deep Neural Networks

Kyeong Soo (Joseph) Kim
Department of Communications and Networking
School of Advanced Technology
Xi'an Jiaotong-Liverpool University (XJTLU)

1

Outline

- XJTLU Camus Information and Visitor Service System
- Wi-Fi Fingerprinting
- Review of Related Projects
- Plans for This Year

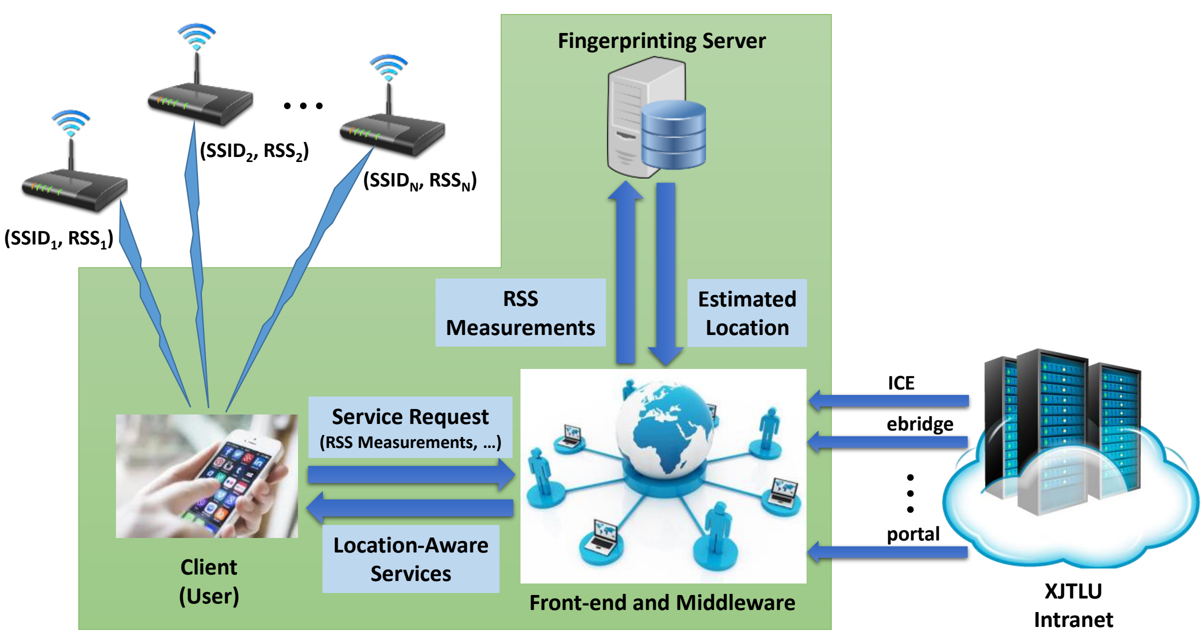
2

2

XJTLU Camus Information and Visitor Service System

3

3



XJTLU Camus Information and Visitor Service System

4

4

Examples: Indoor Navigation and Location-Aware Service

5

5

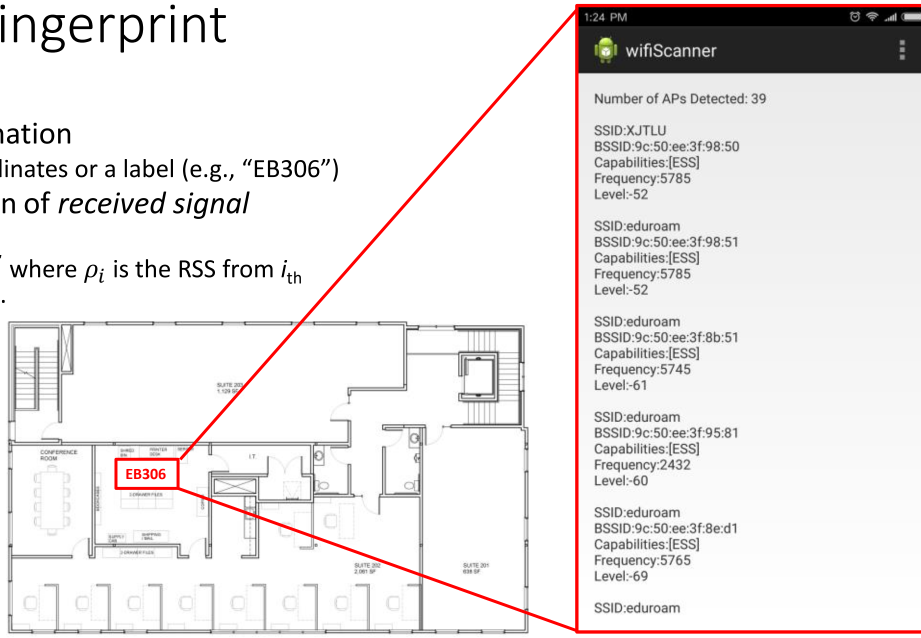
Wi-Fi Fingerprinting

6

6

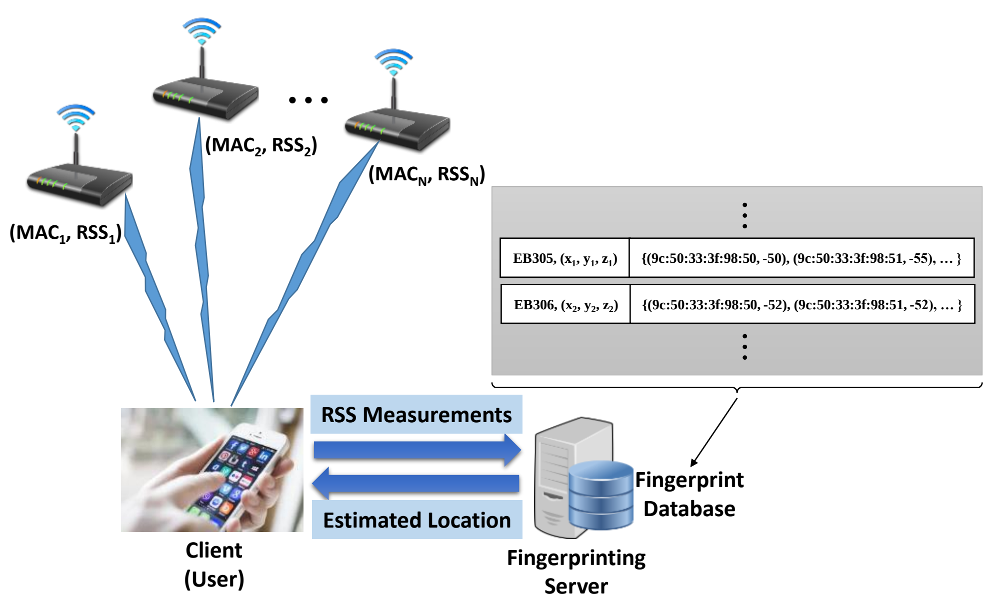
Location Fingerprint

- A tuple of $(\mathcal{L}, \mathcal{F})$
  - $\mathcal{L}$: Location information
    - Geographic coordinates or a label (e.g., "EB306")
  - $\mathcal{F}$: Vector/function of *received signal strengths (RSSs)*
    - e.g., $(\rho_1, \cdots, \rho_N)^T$ where ρ_i is the RSS from i_{th} access point (AP_i).



7

7



8

Location Estimation

- Deterministic
  - **Nearest Neighbour Methods**
  - Neural Network Methods
    - Deep neural networks (DNNs) enabled by deep learning
- Probabilistic
  - Bayesian Inference
  - Support Vector Machine (SVM)
  - Gaussian Process Latent Variable Model (GP-LVM)

9

9

Nearest Neighbour Methods*

- A simple approach based on the notion of distance in the signal space:
  - Given a fingerprint of $(\mathcal{L}, (\rho_1, \cdots, \rho_N)^T)$ and an RSS measurement of $(s_1, \cdots, s_N)^T$, the *Euclidean distance* measure between them is defined as

$$\sqrt{\sum_{i=1}^{N} (s_i - \rho_i)^2}$$

  - Then, we find a fingerprint providing a minimum distance, $\mathcal{L}$ of which is the estimated location.

* *P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," Proc. of INFOCOM 2000, vol. 2, pp. 775-784, Mar. 2000.*

10

10

Major Challenges in Large-Scale Implementation

- ***Scalability***
- ***Localization accuracy***
- Non-stationarity of location fingerprints
  - Incremental/online learning algorithms with pruning/forgetting mechanisms*
- Passive vs. active location estimation
- Integration with other services
- Security/privacy issues

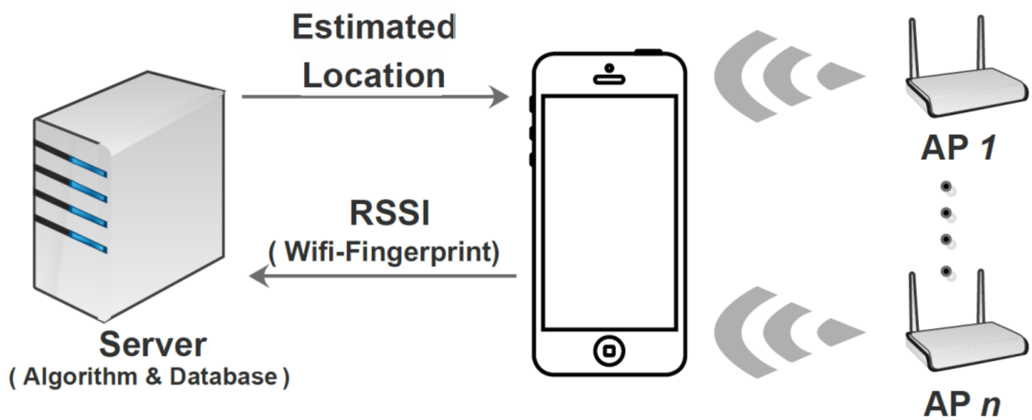
** R. Elwell and R. Polikar, "Incremental learning in nonstationary environments with controlled forgetting," Proc. IJCNN'09.*

11

SURF 2017: Indoor Localisation Based on Wi-Fi Fingerprinting with Fuzzy Sets

12

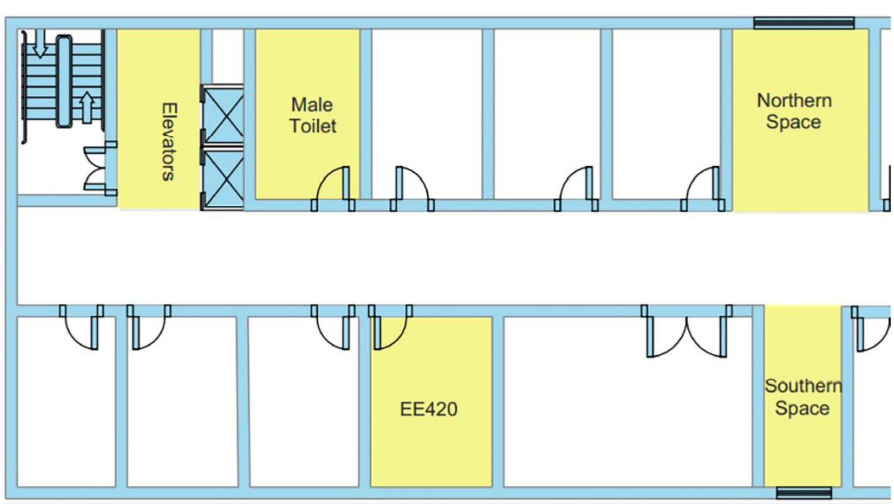
A Prototype of DNN-Based Indoor Localization System for Floor-Level Location Estimation



13

13

A Partial Layout of the Fourth Floor of EE Building



14

14

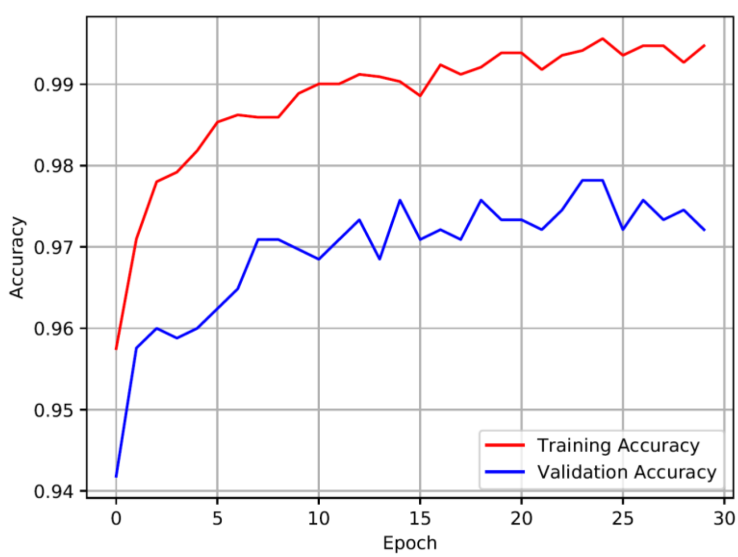
DNN Parameter Values for Floor-Level Location Estimation

DNN Parameter	Value
Ratio of Training Data to Overall Data	0.75
Batch Size	10
SAE Hidden Layers	128-64-8-64-128
SAE Activation	Hyperbolic Tangent (TanH)
SAE Optimizer	ADAM
SAE Loss	Mean Squared Error (MSE)
Classifier Hidden Layers	64-32-7
Classifier Activation	ReLU
Classifier Optimizer	AdaGrad
Classifier Loss	Cross Entropy
Classifier Dropout Rate	0.50
Classifier Epochs	30

15

15

Training and Validation Accuracy of Floor-Level Location Estimation



16

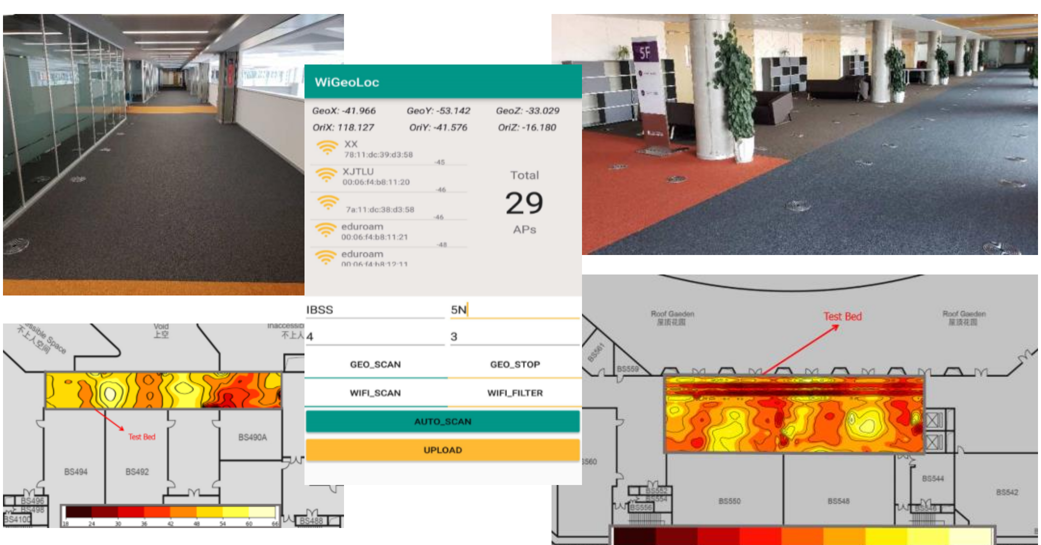
16

SURF 2018: Trajectory Estimation of Mobile Users/Devices Based on Wi-Fi Fingerprinting and Deep Neural Networks

17

17

Toward A Campus-Wide Indoor Localization System:
Multi-Floor Indoor Localization with RSS/Geomagnetic Field in 2018



4th Floor of IBSS Building

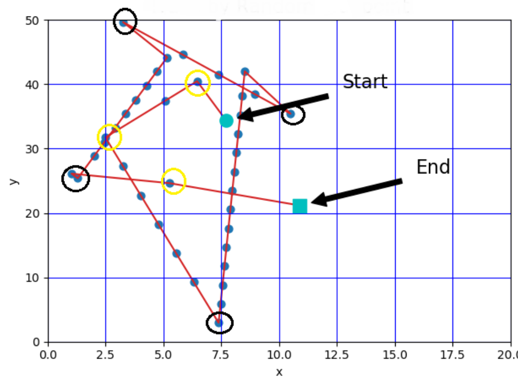
5th Floor of IBSS Building

18

18

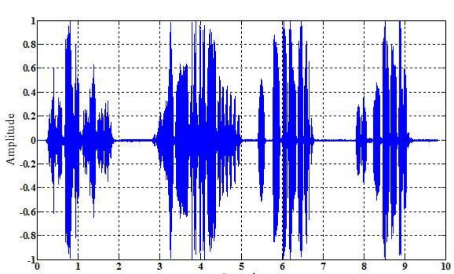
Trajectory Estimation Based on Human Walking Model and LSTM

Data Processing – Random Waypoint Model (RWM)

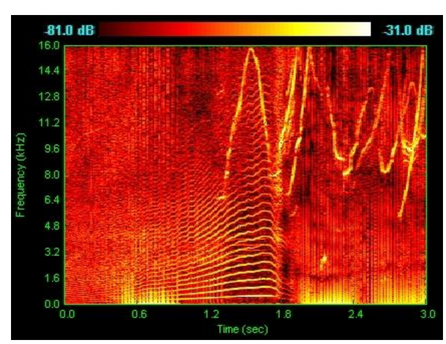


Use of CNN for Time Series Data (e.g., Audio)

Time Domain

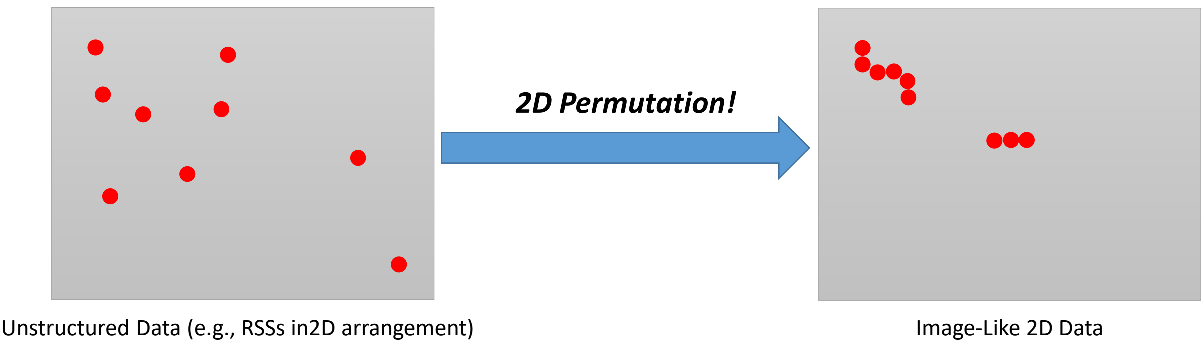


Spectral Domain

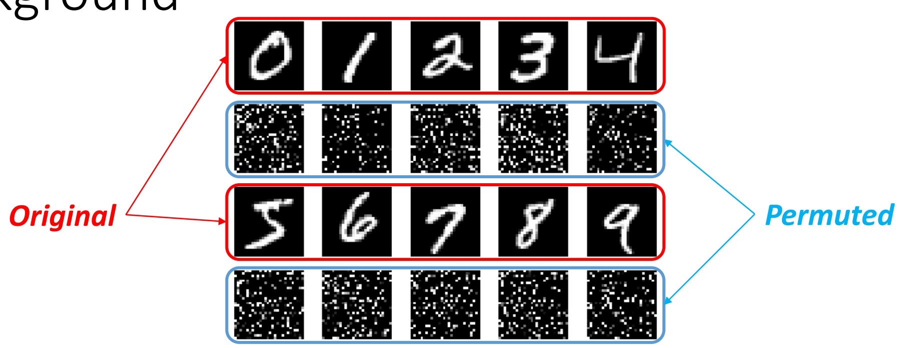


Treat the above as 2-dimensional image!

Mapping of Unstructured Data into Images



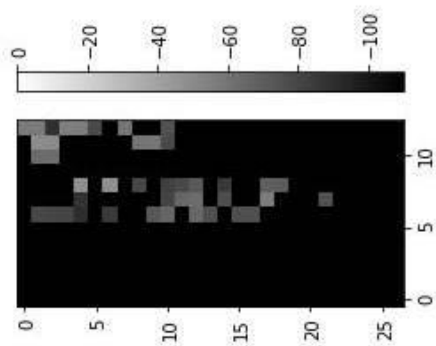
Mapping of Unstructured Data into Images: Background



- CNN
  - With original data: 0.99
  - With permuted data: 0.98
    - 1% drop in accuracy
- Multi-layer perceptron (MLP)
  - With original data: 0.98
  - With permuted data: 0.98
    - Virtually no difference

Mapping of Unstructured Data into Images: Challenges

- How to quantify the ***image-likeness***?
  - Number of connected regions (e.g., *skimage.measure.label*)
  - …
- How to overcome the extremely huge ***size of the search space***?
  - e.g., # of possible permutation for MNIST image = $28^2! \approx 10^{1930.50\ldots}$



Results from SURF-201830

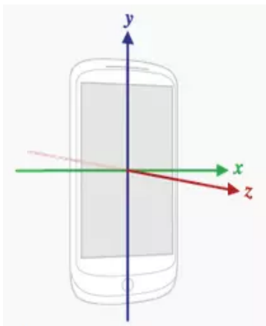
23

23

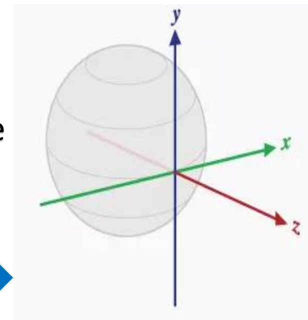
SURF 2019: Analysis of XJTLUIndoorLoc Multivariate Dataset for DNN-Based Indoor Localization

24

24

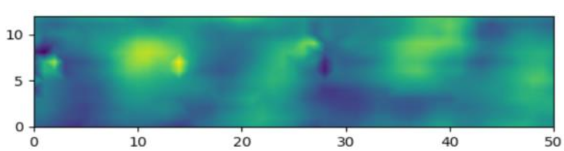


Rotation Matrix: a mathematical method can transform the phone coordinate to world coordinate



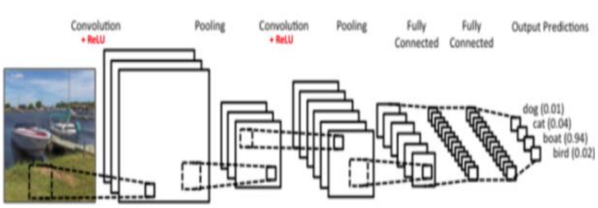
25

Generated Image from RSS Data Set



Interpolation method can generate additional data for the whole image.

Working Principle of CNN



CNN can use the generated image to predict the location.

26

Scalable DNN-Based Multi-Building and Multi-Floor Indoor Localization

27

27

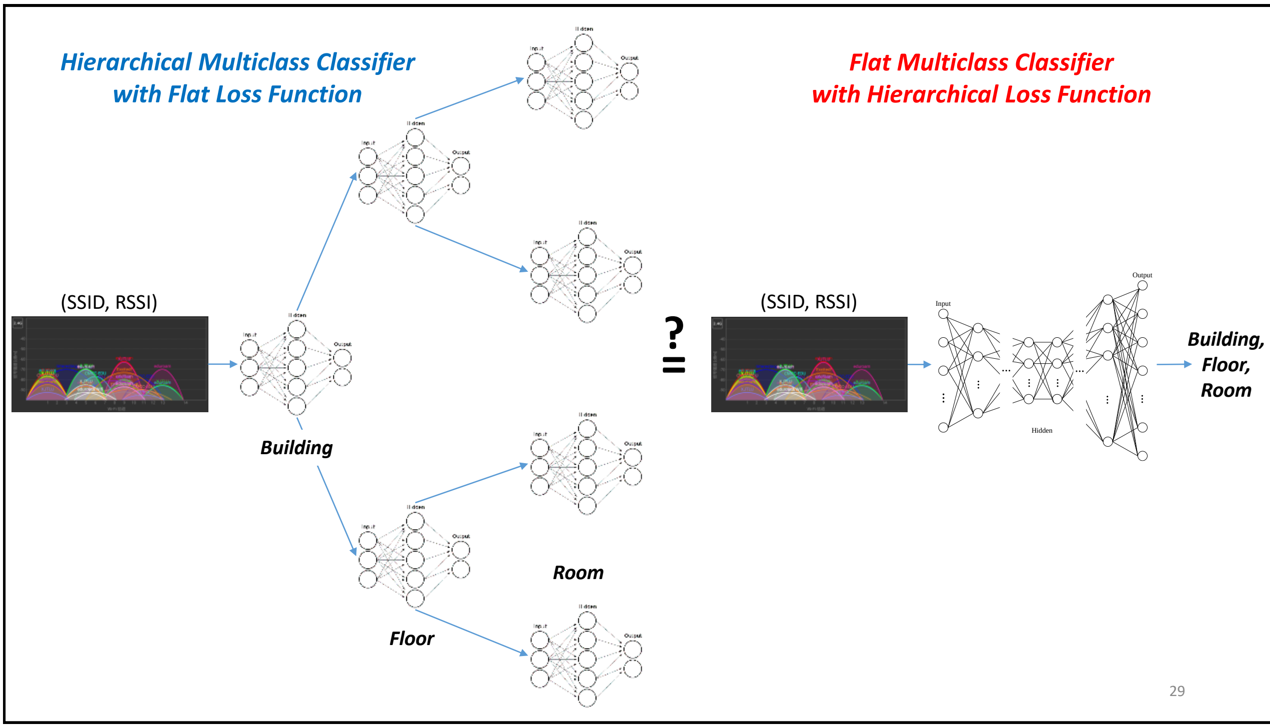
Changes in XJTLU Campuses

2006

2017~

28

28



29

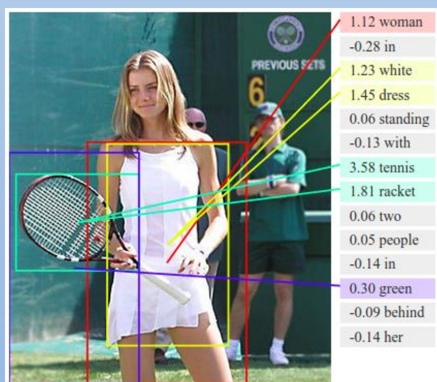
Two Ways of Representing Locations - 1

- Flattened labels
  - As one-dimensional vectors
    - e.g., "EB-3-06" (→ "Building_ID-Floor_ID-Location_ID")
  - For ***multi-class classification***.
- Multi-labels
  - As multi-dimensional vectors
    - e.g., ("EB", "3", "06")
  - For ***multi-label classification***.

30

30

Multi-Label vs. Multi-Class Classification



Multi-Label Classification
- Multiple labels can be assigned to each instance.

Multi-Class Classification
- Classifying an instance into (*only*) one of multiple classes.
- A special case of multi-label classification.
  - Also called *single-label classification*.

	Image #1	Image #2	Image #3
Dog	-0.39	-4.61	1.03
Cat	1.49	3.28	-2.37
Horse	4.21	1.46	-2.27

31

31

Two Ways of Representing Locations - 2

- Scalability of the two representations
  - In machine learning, categorical data containing labels are ***one-hot encoded*** (see the table on the right).

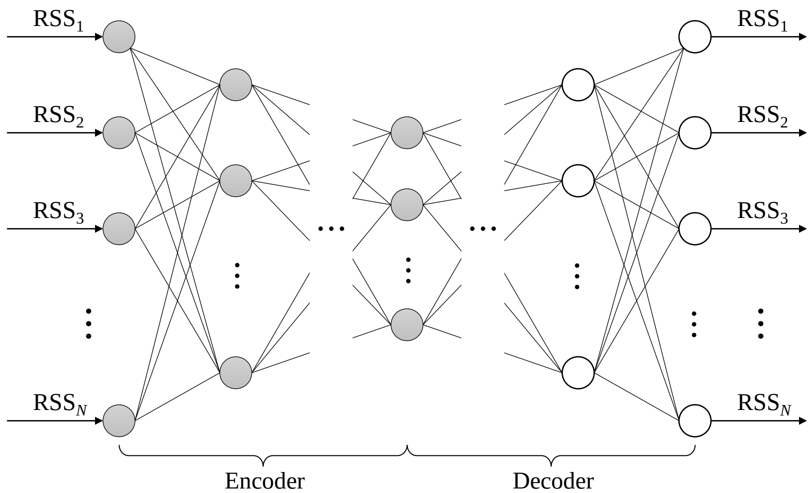
Dog	Cat	Horse
1	0	0
0	1	0
0	0	1

  - A flattened label is one-hot encoded as a whole, while each component of a multi-label can be one-hot encoded independently, e.g.,
    - "EB-3-06" → "0…010….0" vs ("EB", "3", "06") → (0…01, 010…0, 10…0)
  - For a campus with 10 buildings, 10 floors/building, and 10 locations/floor, ***the number of bits*** required for each representation with one-hot encoding is as follows:
    - Flattened labels: 10×10×10 = <u>1,000</u>
    - Multi-labels: 10+10+10 = <u>30</u>

32

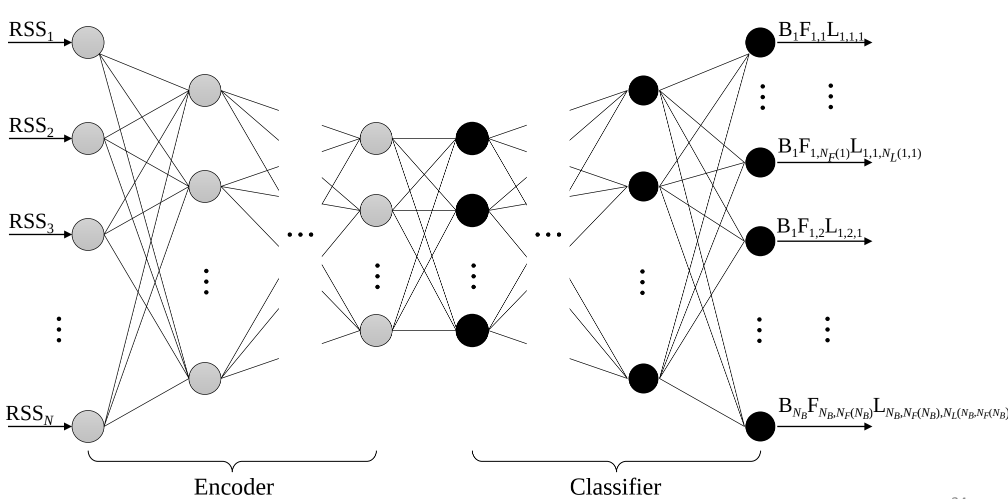
32

Stacked Autoencoder (SAE) for the reduction of feature space dimension



33

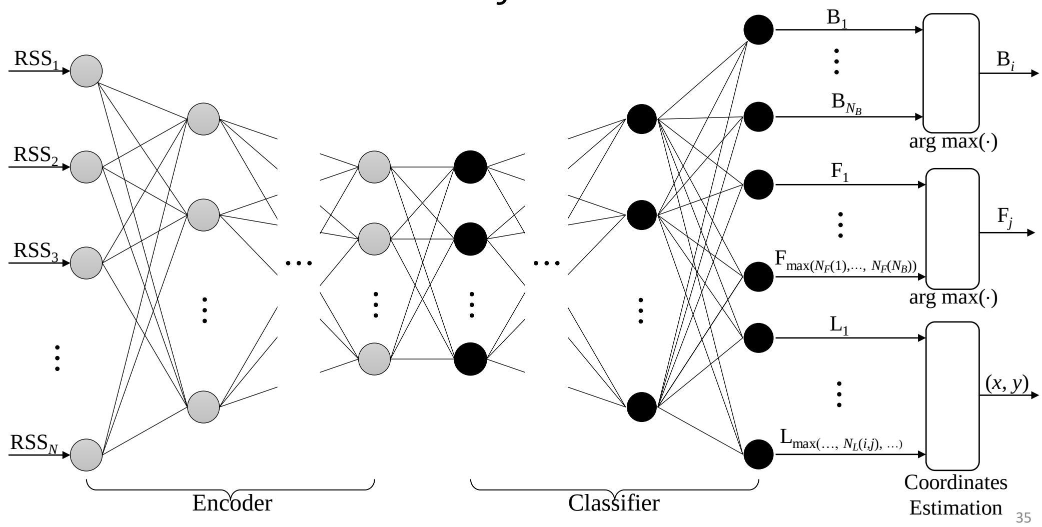
DNN Architecture for Combined Estimation of Building, Floor, and Location based on *Multi-Class Classifier* and Flattened Labels



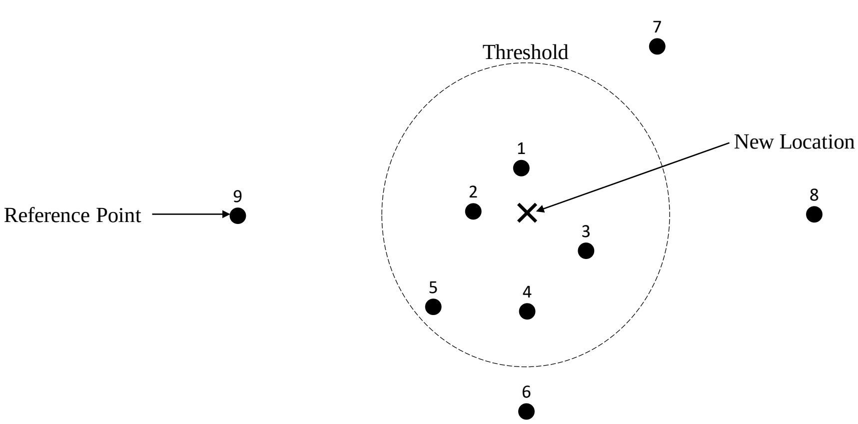
34

34

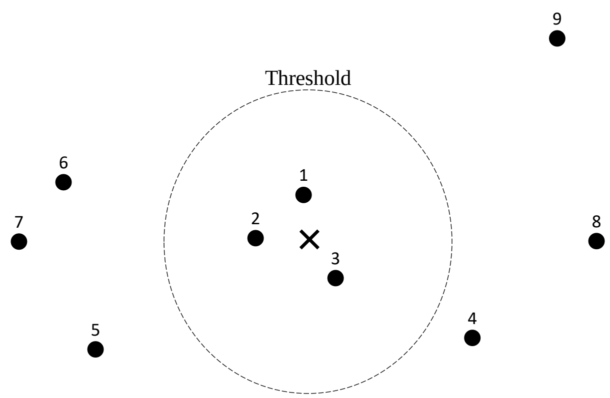
DNN Architecture for *Scalable* Building/Floor Classification and Floor-Level Coordinates Estimation based on *Multi-Label Classifier*



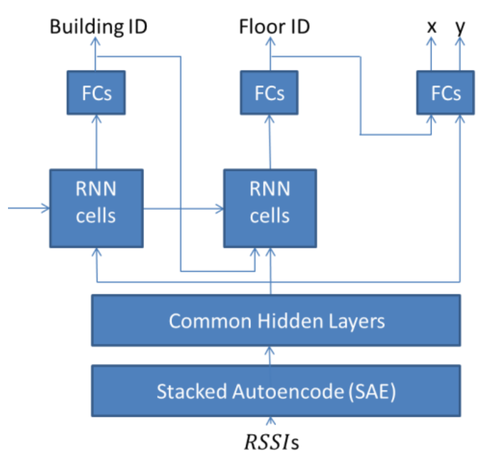
Location Coordinates Estimation: Many Reference Points Centred around A New Location



Location Coordinates Estimation: Only Few Reference Points Centred around A New Location



Hierarchical Multi-Building And Multi-Floor Indoor Localization Based On Recurrent Neural Networks



Plans for This Year

39

39

Backgrounds

- Statistics of the RSSIs in UJIIndoorLoc database
  - Only few APs are detected at a given reference point among the campus-wide 520 APs.
    - This is a typical characteristic of multi-building and multi-floor Wi-Fi RSSI datasets unlike those of large-scale but open-space structures like arenas, auditoriums, and halls where there are no hierarchical structures.
  - There are even reference points without any RSSI, which could cause issues during the prediction.
    - They should be removed during the preprocessing.
  - Inconsistencies in RSSI reported by different HWs (e.g., Samsung vs Xiaomi) and OSs (e.g., Android vs iOS) in different units and scales.
- Possible scalable representation of RSSIs
  - Based on stacked autoencoders (SAEs).
    - Depends on the statistics of a training set.
  - Based on ordering of RSSIs.
    - Note that sorting numbers is one of the hardest tasks for artificial neural networks.
    - Further truncation based on RSSI values (e.g., K-strongest selection).
    - Percentage of total energy (e.g., 90% of total energy; similar to FM).
    - Threshold (e.g., discard RSSIs less than -90)

40

40

Research Questions

- How can we represent in a scalable way large-dimensional RSSIs (e.g., 520-dimensional vectors in the UJIIndoorLoc database) as inputs to a DNN model for multi-building and multi-floor indoor localization?
- What are best DNN architectures for scalable representation of RSSIs (e.g., time series representation)?

Project Plans

- WP1: Scalable representation of RSSIs.
  - To investigate the statistical properties of the RSSIs in the UJIIndoorLoc database.
  - To investigate the representation of RSSI data with a smaller dimension.
  - To investigate the effects of truncating week RSSI measurements.
- WP2: DNN models for scalable representation of RSSIs.
  - To implement and evaluate the performance of indoor localization of DNN models based on various architectures for the proposed RSSI representation from WP1.