Efficient Wi-Fi Localization in In-Door Environments

Internship Supervisors: Arnaud Legout, Chadi Barakat

Academic Supervisor: Giovanni Neglia

Student: Angelo Rodio

Motivations

- Rapid development of Location-based Services
- GPS unable to guarantee satisfactory positioning performance in indoor environments (multipath, fading, …)
- Various indoor positioning technologies have been proposed
  - Require dedicated hardware → *cost-prohibitive*
  - Require APs position knowledge → *user-unpractical*
  - Do not support unassociated clients → *privacy-unfriendly*

Objective

- Is it possible to estimate the position of a mobile device in indoor environments?

Requirements

- Do not require dedicated hardware → reuse existing infrastructure
- Do not require APs position knowledge
- Support unassociated clients

Challenge

- Satisfy all these requirements together

State of the art

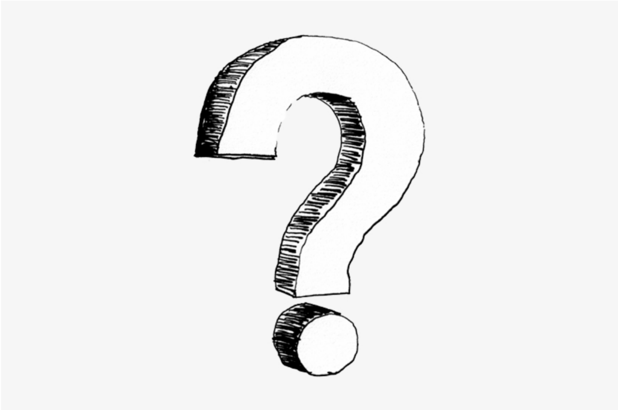
Positioning technologies

- GPS
- RFID
- UWB
- IR
- Wi-Fi
- Bluetooth

⚠ Does not work in indoor environments (multipath, fading, …)

⚠ Require dedicated hardware

👍 Ubiquitous networks



State of the art

Indoor positioning technologies

Wi-Fi

Fine-grained localization based on computation

- RTT, TSA-based methods
  1. Time of Arrival (ToA)
  2. Angle of Arrival (AoA)
  3. Time Difference of Arrival (TDoA)

  \+ Trilateration

⚠ Require dedicated hardware
⚠ Do not support unassociated clients

Coarse-grained localization based on estimation

- RSS-based methods
  1. Propagation model + trilateration
  2. Fingerprinting

- Reuse WLAN existing infrastructure

State of the art

Indoor positioning technologies

Wi-Fi

Coarse-grained localization based on estimation

- RSS-based methods

<u>Propagation model + trilateration</u>

⚠ Require APs position knowledge
⚠ Indoor environments unpredictable

<u>Fingerprinting</u>

👍 Reuse WLAN existing infrastructure
👍 Support unassociated clients
👍 Do not require APs position knowledge

All requirements satisfied

- Coarse-grained
- Time-consuming

Need enhancing

Outline

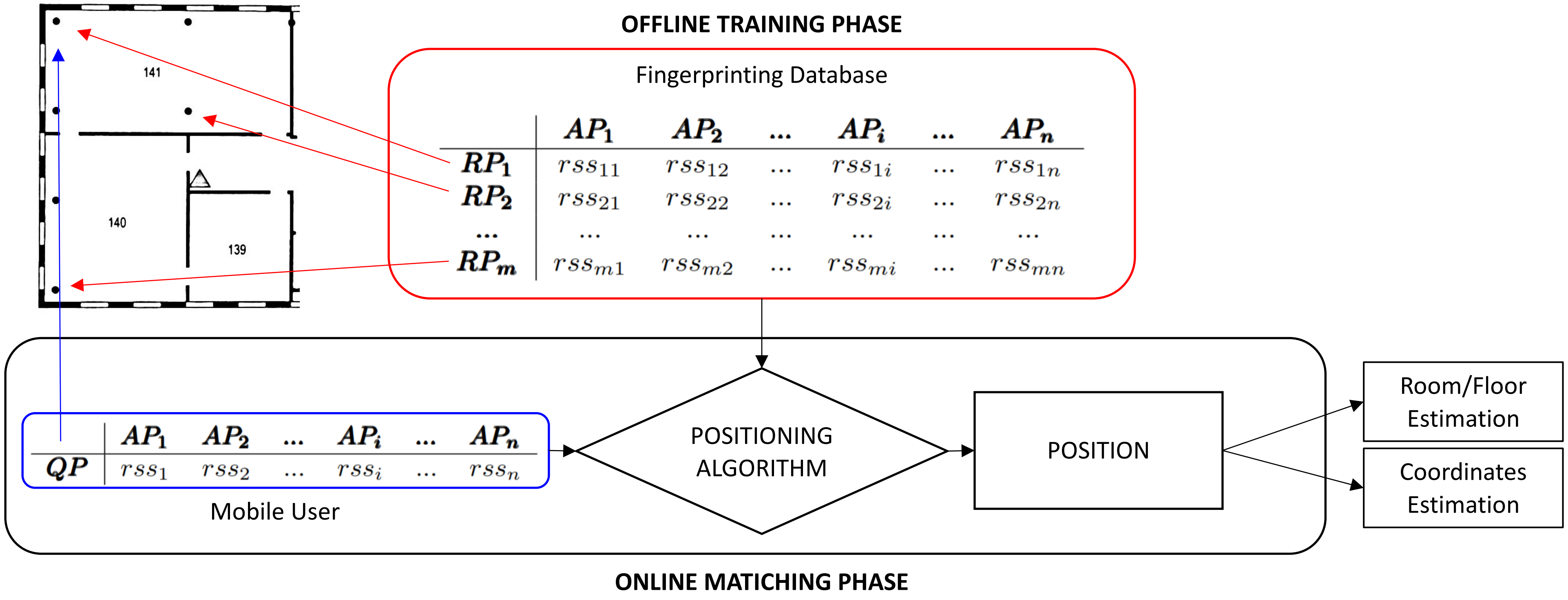
- Implement and compare Wi-Fi fingerprinting state-of-the-art techniques
- Problem: cost of data collection in current indoor positioning techniques
- Proposed solution: Wi-Fi Fingerprints generation through DCGAN (WF-DCGAN)
- Experiments and results
- Conclusions and current/future work

Wi-Fi fingerprinting technique

Fingerprinting

Offline phase:
data collection

Online phase:
positioning algorithm

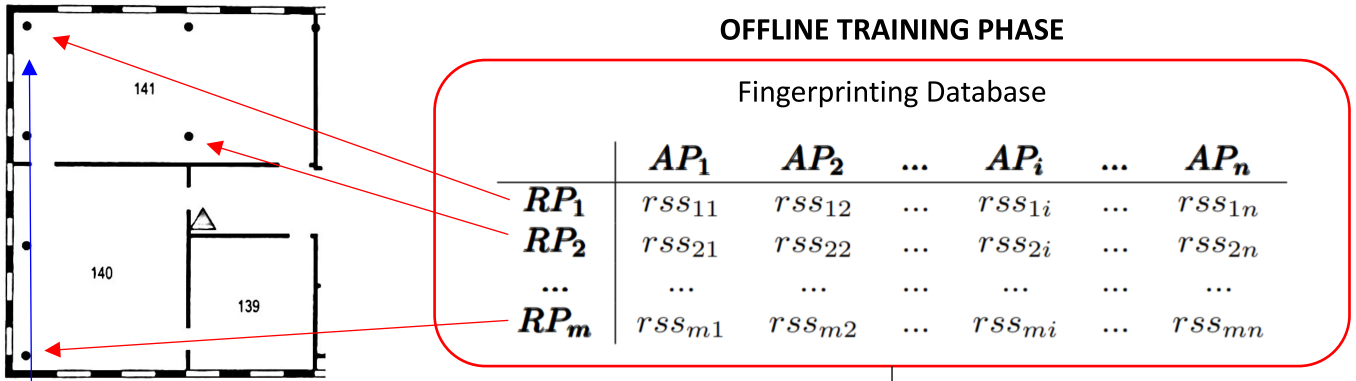


Wi-Fi fingerprinting technique

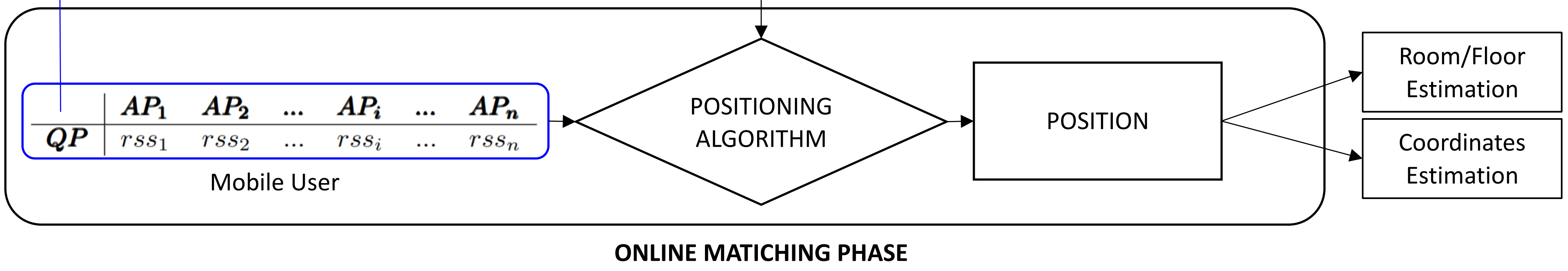
Fingerprinting

Offline phase:
data collection

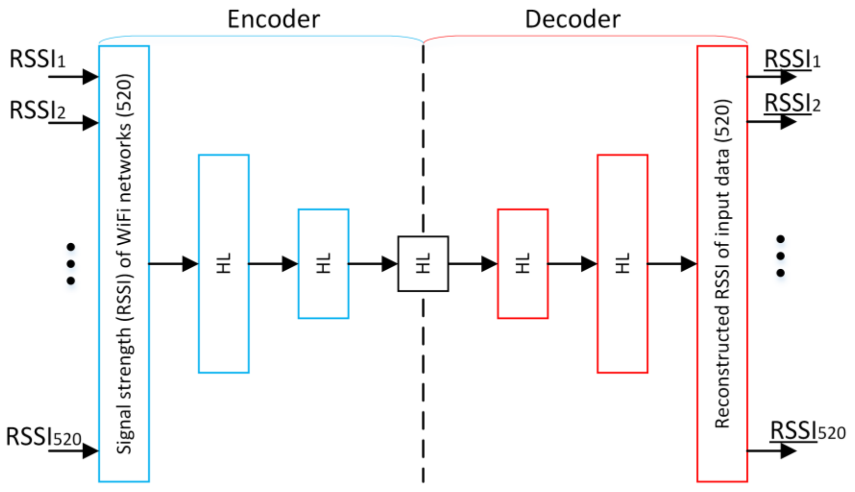
Online phase:
positioning algorithm



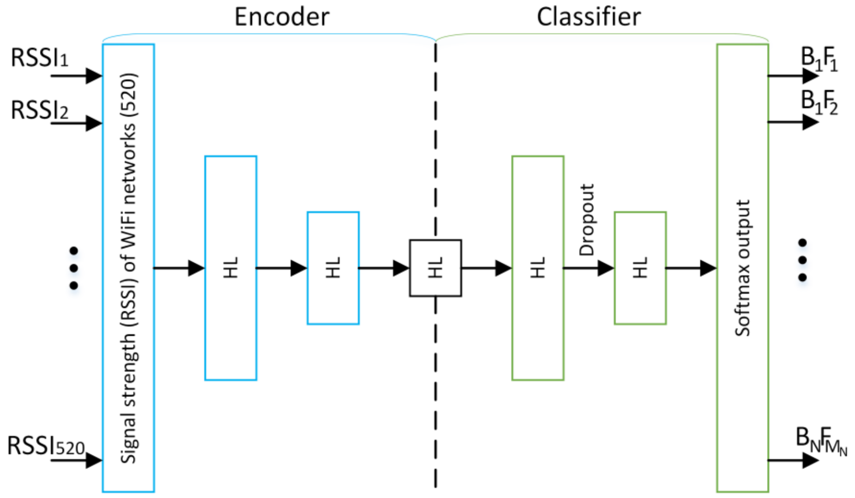
- Deep Neural Network with SAE
- K-Nearest Neighbors
- Random Forest classifier
- Support Vector Machine
- Naïve Bayes classifier



Deep Neural Network with SAE [1]



Unsupervised learning



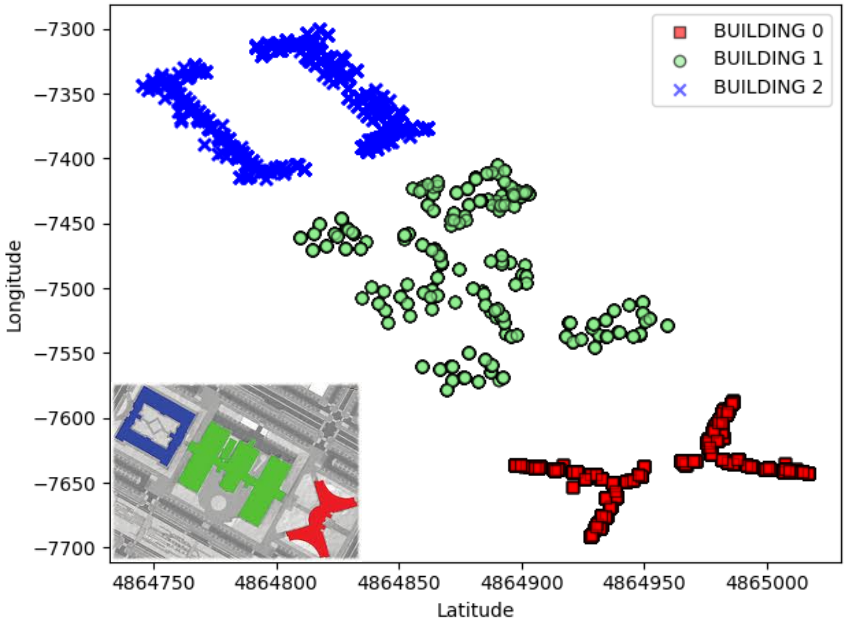
Supervised learning

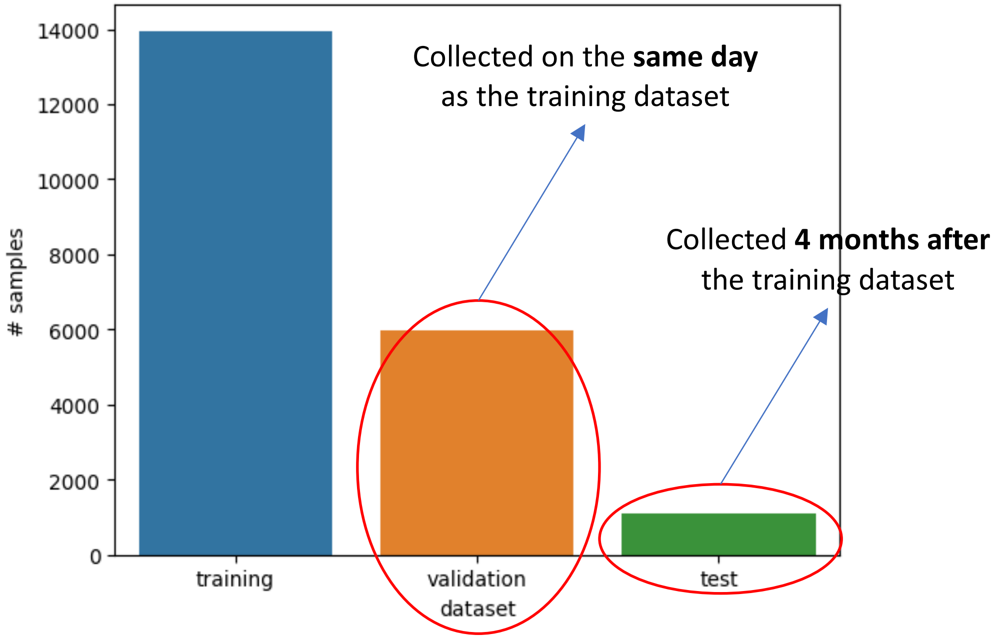
- Low-effort place recognition algorithm that uses deep learning to provide location recognition from Wi-Fi data
    - Input: vector representation of RSS fingerprints
    - Output: probabilities of current sample belonging to a Building and Floor (BF)
- Advantage
    - Significantly reduce the effort for manual tuning

[1]: Nowicki, Michał, and Jan Wietrzykowski. "Low-effort place recognition with WiFi fingerprints using deep learning." *International Conference Automation.* Springer, Cham, 2017.

UJIIndoorLoc Dataset [2]

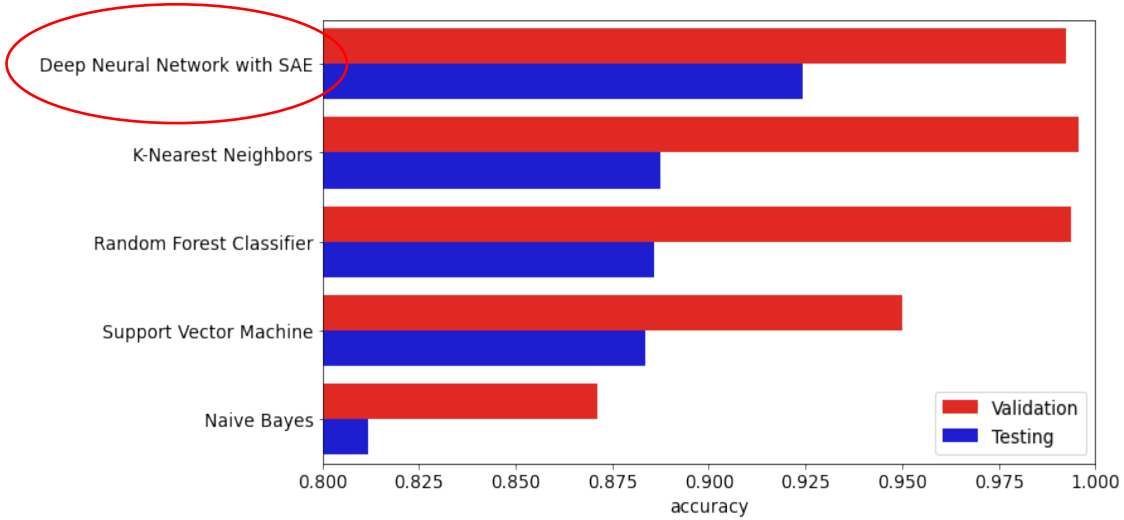
- 3 buildings with 4 floors, 110.000 m², from Universitat Jaume I, Valencia, Spain
- 19 different users and 25 Android devices





[2]: Torres-Sospedra, Joaquín, et al. "UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems." *2014 international conference on indoor positioning and indoor navigation (IPIN).* IEEE, 2014.

Comparison of different algorithms

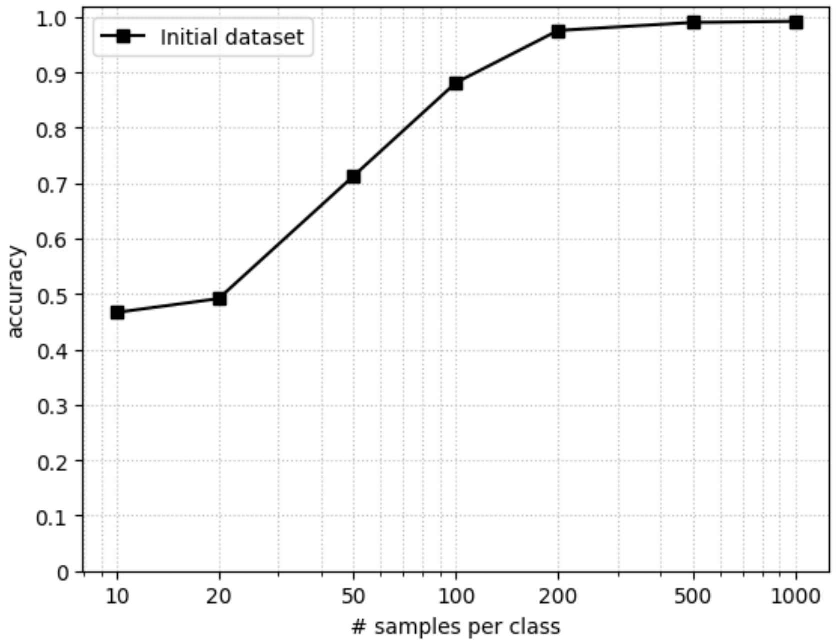


Comparison of correct recognition ratios for different classification algorithms evaluated in building and floor classification problem obtained for validation samples (collected on the **same day** as the training set) and testing samples (collected **four months after** the training set).

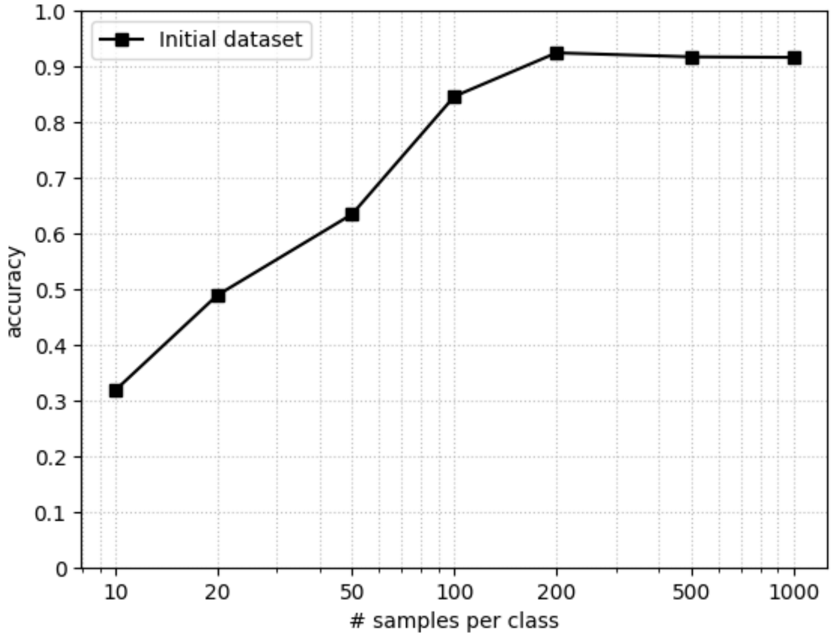
CONCLUSIONS:
- Neural Networks allow Wi-Fi localization algorithm to achieve very high accuracy
- Neural Networks perform better if we have lots of fresh samples

QUESTIONS:
- What is the price to pay for data collection?
- How many samples do we need to collect? More data = higher accuracy?

More data = higher accuracy



Localization accuracy for different sizes of the fingerprinting database on validation dataset collected on the **same day** as the training dataset.



Localization accuracy for different sizes of the fingerprinting database on test dataset collected **four months after** the training dataset.

We would like to collect **around 200 samples per location**

Cost of current fingerprinting techniques

Fingerprinting

Offline phase:
data collection

More data = Higher accuracy

Data collection is:
- Time-consuming
- User labor-intensive

Significant human effort = Low popularity

Online phase:
positioning algorithm

The following limitations apply to the frequency of scans using `WifiManager.startScan()`.

Android 8.0 and Android 8.1:
- Each background app can scan one time in a 30-minute period.

Android 9 and later:
- Each foreground app can scan **two times in** a **1-minute** period. → 100 min
- All background apps combined can scan **one time in** a **30-minute** period. → 100 h

GOAL: Reduce the data collection cost

How literature faces the problem

- Use modelling to choose in a ‘clever’ way which sample collect [3], [4], [5]
  - Semi-supervised learning / active learning techniques to reduce the labeling cost of fingerprint collection

PROBLEMS:
- Still assume the availability of a considerable number of samples (unlabeled rather that labeled)
- Reduce the user labor but do not actually reduce the time of data collection
- Reduce also the number of reference points → dramatically decrease the accuracy

- Use modelling to recover fingerprints based on existing partial fingerprints [6], [7]
  - Fingerprint interpolation / gradient-based fingerprint generation
  - Interesting approach: generate new samples that somehow integrate the radiomap, BUT...

PROBLEM: Rely on mobile trajectories whose collection is not feasible with Wi-Fi cards

[3]: Gu, Zhuan, et al. "Reducing fingerprint collection for indoor localization." *Computer Communications* 83 (2016): 56-63.
[4]: Pulkkinen, Teemu, et al. "Semi-supervised learning for wlan positioning." *International Conference on Artificial Neural Networks*. Springer, Berlin, Heidelberg, 2011.
[5]: Liu, Shaoshuai, et al. "A low-cost and accurate indoor localization algorithm using label propagation based semi-supervised learning." *Fifth International Conference on Mobile Networks*. IEEE, 2009.
[6]: Shu, Yuanchao, et al. "Gradient-based fingerprinting for indoor localization and tracking." *IEEE Transactions on Industrial Electronics* 63.4 (2015): 2424-2433.
[7]: Cho, Youngsu, et al. "GPR based Wi-Fi radio map construction from real/virtual indoor dynamic surveying data." *2013 13th International Conference on Control, Automation and Systems.* IEEE, 2013.

How do I propose to face the problem

WF-DCGAN: method to increase the amount of Wi-Fi Fingerprints collected at each location based on Deep Convolutional Generative Adversarial Network (DCGAN)

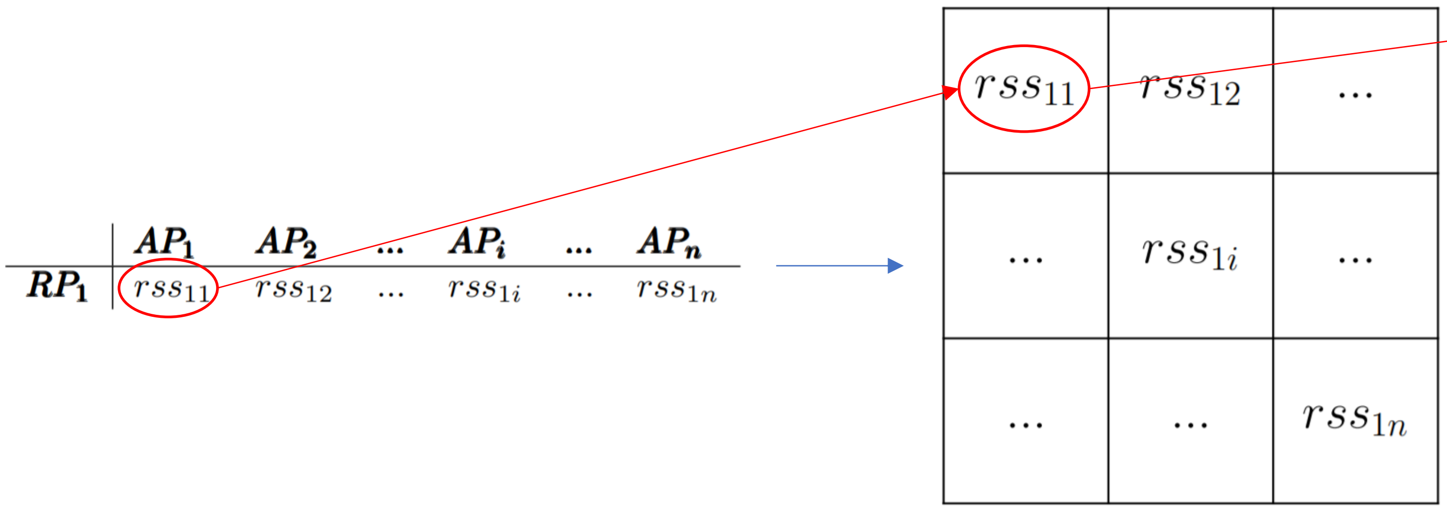
1. Transform RSS data into feature vectors
2. Transform feature vectors into feature images
3. Generate similar feature images with GANs
4. Merge initial fingerprints with generated feature images

ADVANTAGES:

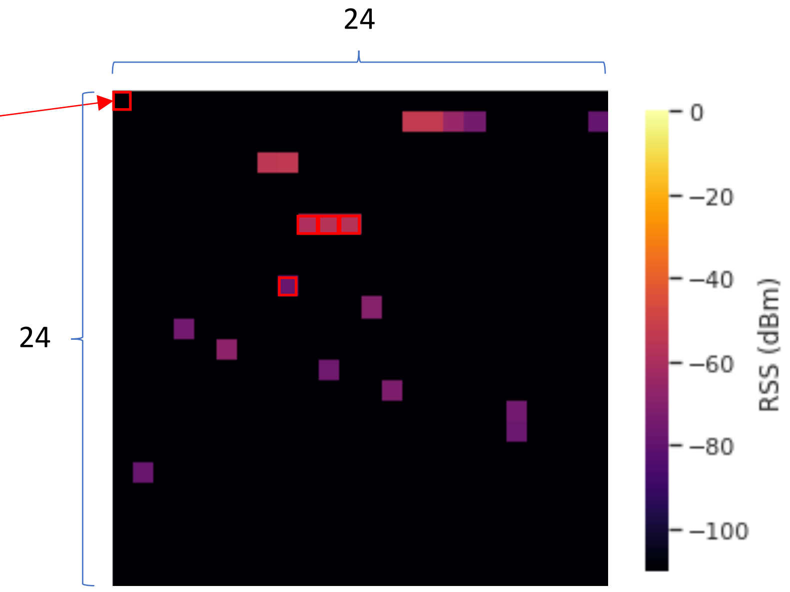
- Truly expand the number of fingerprints in the database
- Reduce human effort
- Increase diversity

WF-DCGAN

- Map feature vectors into feature images



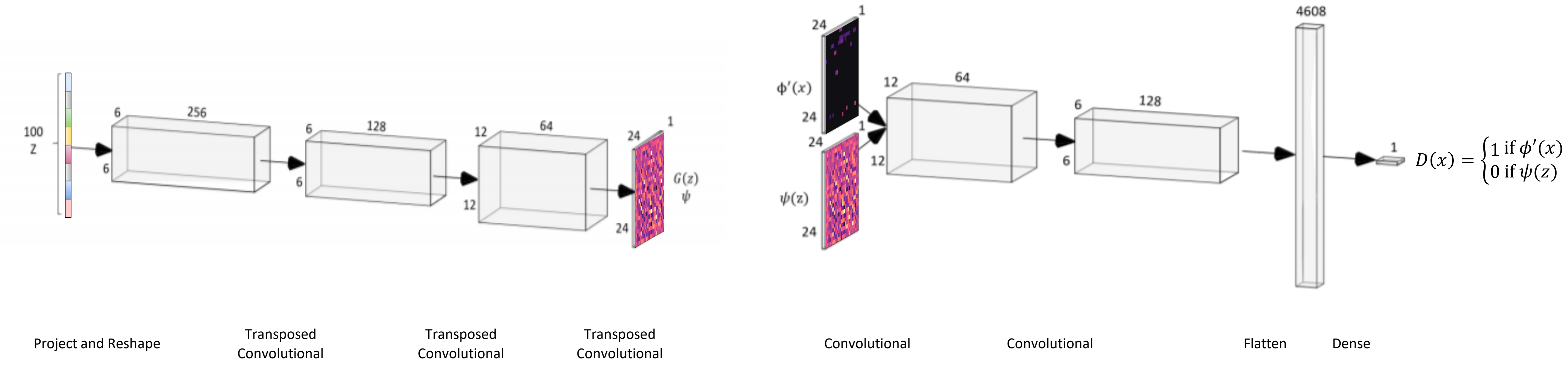
Mapping of a Wi-Fi scan (vector of length n, where n is the number of APs detected) into an image of n pixels, with shape $\sqrt{n} \cdot \sqrt{n}$. Each pixel represents a RSS measurement, whose value is in the range $[0, -110]$ dBm.



Example of image of shape 24x24 pixels representing a real Wi-Fi fingerprint measured at ground floor of first building.

WF-DCGAN

- Generate similar feature images with DCGAN



(a) Generative model

(b) Discriminative model

Network structure of the DCGAN model.
G is the generative model (a), implemented by a deconvolutional network;
D is the discriminative model (b), implemented by a convolutional network.

WF-DCGAN

- Generate similar feature images with DCGAN [8]

$$\min_G \max_D V(D,G) = E_{x\sim P_{Data}(x)}\{\log D(x)\} + E_{z\sim P_Z(z)}\{\log(1 - D(G(z)))\}$$

G = generator

D = discriminator

$P_{Data}(x)$ = distribution of real data

$P_Z(z)$ = distribution of generator

x = sample from real data

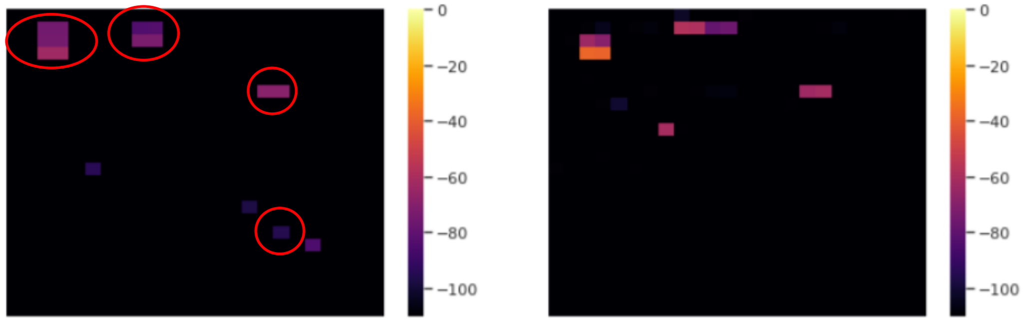
z = sample from generator

$D(x)$ = discriminator network

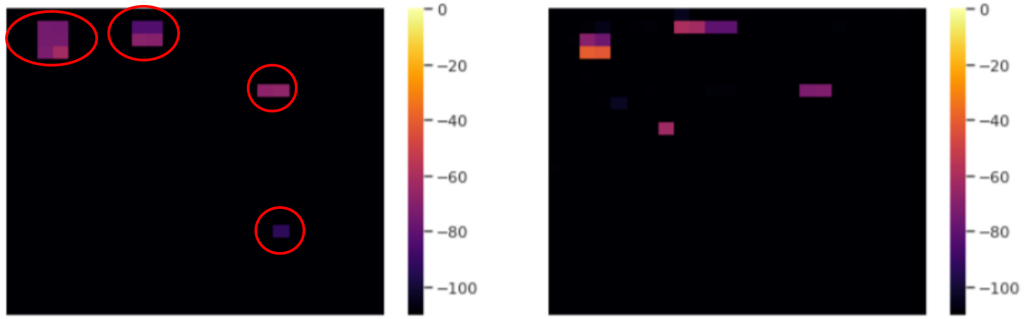
$G(z)$ = generator network

[8]: Goodfellow, Ian, et al. "Generative adversarial nets." *Advances in neural information processing systems.* 2014.

Real vs Generated fingerprints

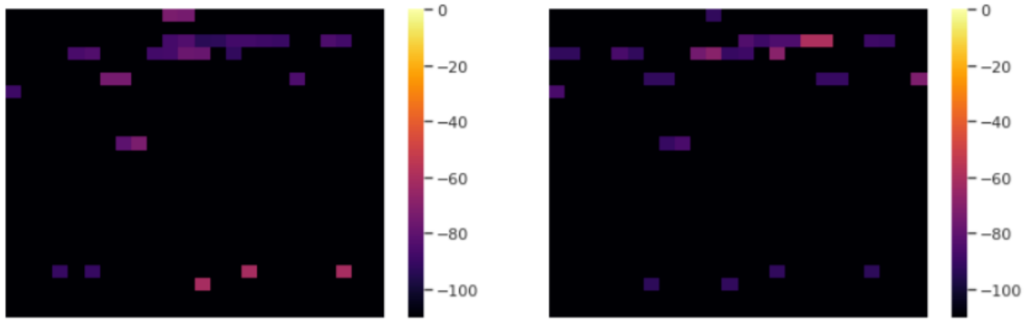


(a) Real images representing Wi-Fi fingerprints from class B0F0

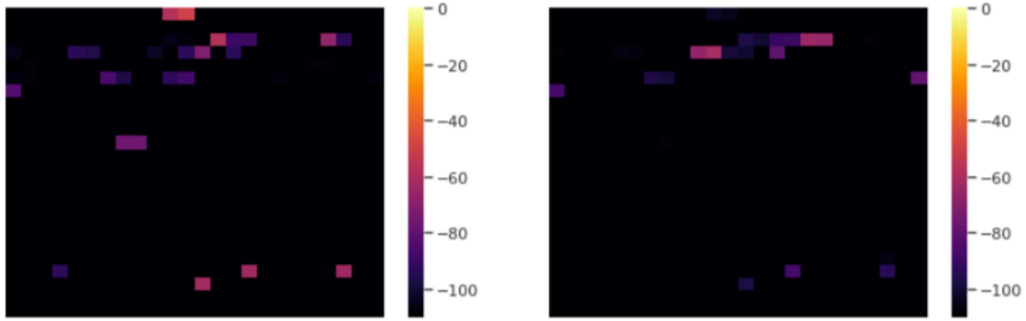


(b) Generated images representing Wi-Fi fingerprints from class B0F0

Comparison between real (a) and generated (b) images representing Wi-Fi fingerprints measured in the same environment, i.e., ground floor of first building.



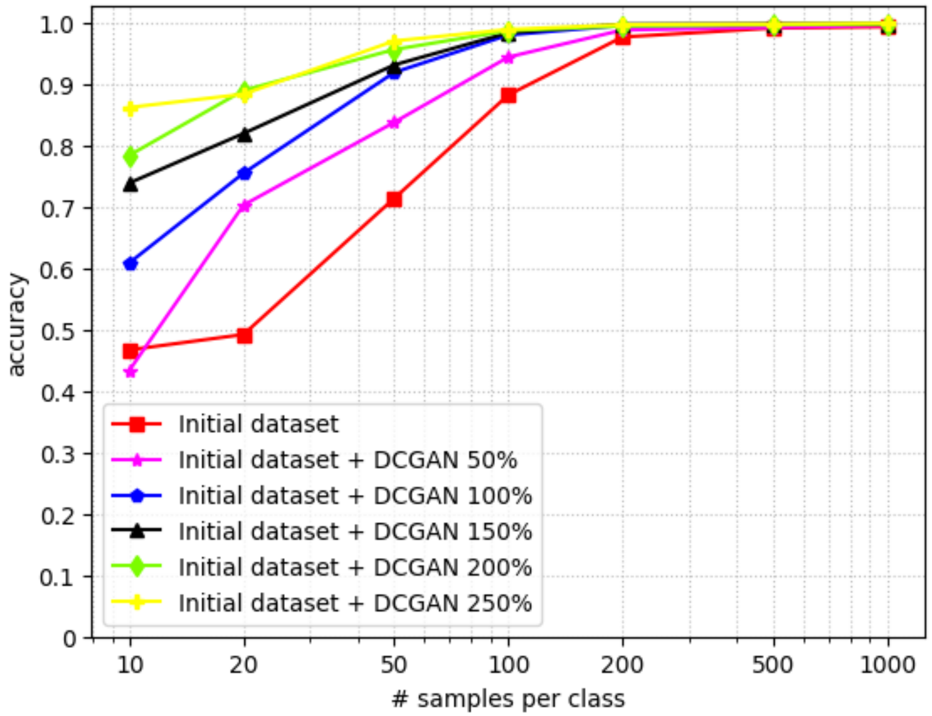
(a) Real images representing Wi-Fi fingerprints from class B2F4



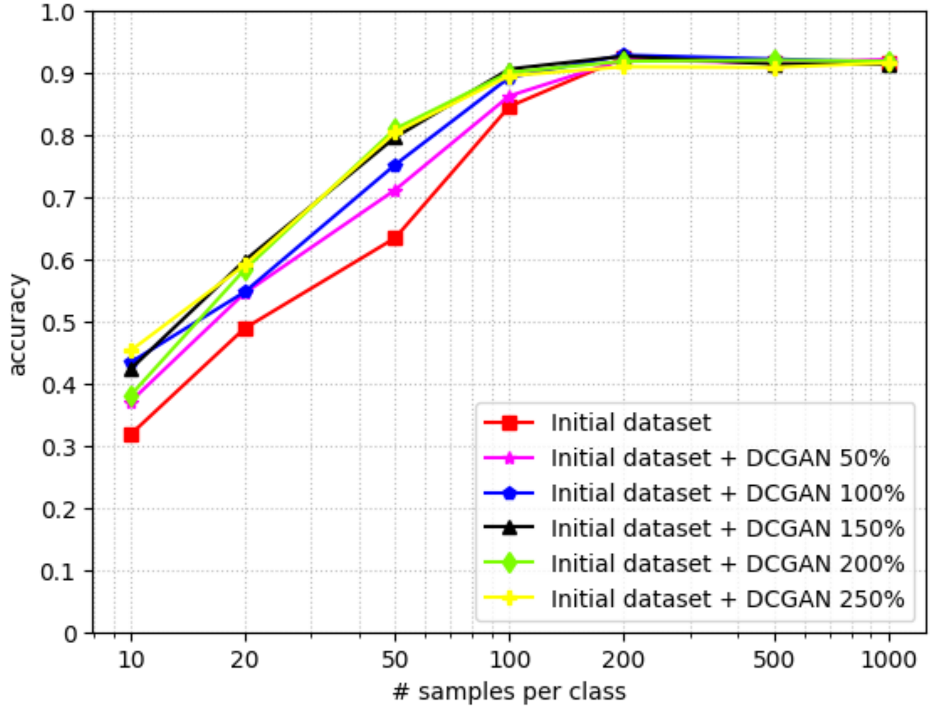
(b) Generated images representing Wi-Fi fingerprints from class B2F4

Comparison between real (a) and generated (b) images representing Wi-Fi fingerprints measured in the same environment, i.e., top floor of third building.

Experimental results



Localization accuracy for different sizes of the fingerprinting database and for different number of samples generated by DCGAN on validation dataset (collected on the **same day** as the training dataset).



Localization accuracy for different sizes of the fingerprinting database and for different number of samples generated by DCGAN on test dataset (collected **four months after** the training dataset).

Conclusions

- The low popularity of current Wi-Fi fingerprinting techniques is not due to poor accuracy in the positioning but to the high cost of fingerprint collection in terms of time and human-labor
- Most of state-of-the-art methods face the problem implementing techniques that reduce the number of reference points, thus decreasing the accuracy
- WF-DCGAN generates additional Wi-Fi fingerprints for each location, truly increasing the number of fingerprints in the database with no human effort
- Experiments show that WF-DCGAN increases the accuracy of the implemented localization algorithm on the UJIIndoorLoc dataset up to **40%** when the test samples are collected the **same day** as the training samples and up to **18%** if the test samples are collected **4 months after** the training samples

Current work: crowdsourcing

- Problem: all Wi-Fi fingerprinting techniques suffer from **time-aging**
  - Radiomaps becomes obsolete with time
  - It is crucial for the users to periodically resample the database
- **Crowdsourcing** allows scan sharing among multiple users
  - Facilitate radiomap creation through user collaboration
  - Drastically accelerate the fingerprint collection process
- My work already considers crowdsourcing on 25 Android devices, but do not take into account the Wi-Fi hardware variance problem

THE WI-FI HARDWARE VARIANCE PROBLEM

- Different devices measure different RSS readings at same location and from same Wi-Fi AP
- Varying RSS can degrade the positional accuracy of Wi-Fi location systems

PROPOSED SOLUTION: apply a transformation function based on signal-patterns of different devices