SURF-201913 Preparation Meeting: Analysis of XJTLUIndoorLoc Multivariate Dataset for DNN-Based Indoor Localisation

Kyeong Soo (Joseph) Kim
Department of Electrical and Electronic Engineering
Centre of Smart Grid and Information Convergence
Xi’an Jiaotong-Liverpool University (XJTLU)

Outline

• XJTLU Camus Information and Visitor Service System
• Wi-Fi Fingerprinting
• SURF 2017
• SURF 2018
• Scalable DNN-Based Multi-Building and Multi-Floor Indoor Localisation
• Plans
XJTLU Camus Information and Visitor Service System

RSS Measurements
Estimated Location

Service Request (RSS Measurements, ...)

Location-Aware Services

Front-end and Middleware

Fingerprinting Server

Client (User)

XJTLU Camus Information and Visitor Service System
Examples: Indoor Navigation and Location-Aware Service

Wi-Fi Fingerprinting
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., \( (\rho_1, \ldots, \rho_N)^T \) where \( \rho_i \) is the RSS from \( i \)th access point \((AP_i)\).
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)

**Nearest Neighbour Methods***

• A simple approach based on the notion of distance in the signal space:
  • Given a fingerprint of \( (L, (\rho_1, \ldots, \rho_N)^T) \) and an RSS measurement of \( (s_1, \ldots, s_N)^T \), 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, \( L \) of which is the estimated location.

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


SURF 2017: Indoor Localisation Based on Wi-Fi Fingerprinting with Fuzzy Sets
A Prototype of DNN-Based Indoor Localization System for Floor-Level Location Estimation

A Partial Layout of the Fourth Floor of EE Building
DNN Parameter Values for Floor-Level Location Estimation

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

Training and Validation Accuracy of Floor-Level Location Estimation
SURF 2018: Trajectory Estimation of Mobile Users/Devices Based on Wi-Fi Fingerprinting and Deep Neural Networks

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

Unstructured Data (e.g., RSSs in 2D arrangement) → 2D Permutation! → Image-Like 2D Data

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 = $28^2! \approx 10^{1930.50}$...

Scalable DNN-Based Multi-Building and Multi-Floor Indoor Localization
Changes in XJTLU Campuses

Hierarchical Multiclass Classifier with Flat Loss Function

Flat Multiclass Classifier with Hierarchical Loss Function

(SSID, RSSI) → Building → Floor → Room

(SSID, RSSI) → Building, Floor, Room
Plans

• WP1: Statistical analysis of XJTLUIndoorLoc dataset.
  • To quantify the dependency of measurement data on mobile devices.
  • To investigate the impact of mobile devices on indoor localization/trajectory estimation performance
  • To do additional measurements with new mobile devices.

• WP2: Handling device orientation information for geomagnetic field intensity.
  • To study the device coordinate frame and rotation data of smartphones based on their built-in accelerometer, gyroscope, and compass.
  • To investigate how to handle mismatch between the device orientation of geomagnetic filed data in the dataset and that of a new measurement during the online indoor localization/trajectory estimation phase.