MMLoc+: A Transfer Learning based Multimodal Machine Learning Localization System for Dynamic Sensor Networks

Xijia Wei
UCL Interaction Centre
University College London

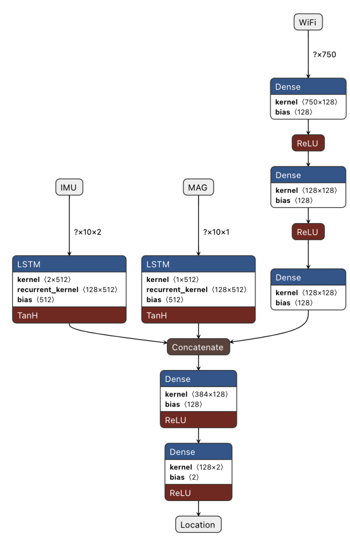
Valentin Radu
Dept. Computing Science
University of Sheffield

An increasing number of smartphone applications offer their users location-based services such as navigation or location tracking based on geographic locations. In outdoor environments, GPS exploits satellite signals to locate devices. However, in indoor environments, satellite signals have a poor penetration through building structure, which makes GPS unreliable. To produce alternative solutions for indoor location estimation, the smartphone built-in sensor signals, such as WiFi received signal strength, magnetic field, and inertial sensors (e.g., accelerometer, gyroscope), have been the focus of intensive research over the past years.

Two main approaches have been considered for indoor localisation, Pedestrian Dead Reckoning (PDR) and WiFi Fingerprinting [1]. Both of these alternatives require well-engineered solutions built on precise mathematical formulations to process the sensor signals. However, these heavily-engineered solutions often fail when deployed in different scenarios. This is due to environment change away from the ideal conditions of the lab, and varying sensor sensibility between these devices used when engineering the solution and those used in deployment. All these aspects affect the system robustness. As a result, human interventions for periodic calibration is essential for maintaining the accurate functioning of these systems, which makes wide adoption unattainable.

In the era of big data, we believe that only relying on data itself to deliver an end-to-end data-driven machine learning approach is a promising solution for robust indoor localization, instead of conventional specialised engineered approaches. Inspired by the success of multimodal machine learning in many modality-fusion tasks such as audio-vision speech recognition, in this study, we expand on the end-to-end multimodal machine learning architecture [2], proposing MMLoc+ with transfer learning for smartphone sensor fusion.

We use our own multimodal dataset collected from two indoor scenarios of A and B. Both datasets contain time-sequential IMU sensors and magnetic samplings as well as WiFi RSS (Received Signal Strength) fingerprints from corridors and ground truth location annotation. We categorise the multisensory dataset into two types: the infrastructure-free samplings (inertial sensors) and the infrastructure-based samplings (magnetic and WiFi signals). For processing infrastructure-free samplings, we pretrain a recurrent model (LSTM1) as a feature extractor and then integrate this model component into MMLoc+ architecture using transfer learning methods. For extracting infrastructure-based samplings, we adopt another LSTM model (LSTM2) and a DNN network to extract multi-sensor features. All extracted modality-specific features are then joined in a one-dimension vector, followed by additional multi-layer perceptrons to produce the joint location estimation. The MMLoc+ model is trained on the multi-sensor datasets (scenario A) to update its weights and bias, except for the transferred LSTM1 sub-component, which contains the learnt infrastructure-free data representations. To evaluate MMLoc+'s robustness and generalisation, we use data from scenario B for location prediction. We find that the MMLoc+ predicts the trajectory with a clear shape closed to the target path, with over 80% of the estimations having errors lower than 3 metres. Furthermore, by implementing the transfer learning approach, MMLoc+ is considered computationally efficient and can achieve better results being bootstrapped with just a small number of samples. This makes MMLoc+ ideal for deployment in various scenarios with minimal cost. We are expanding the model to accept more modalities as sensor input.



References

[1] Valentin Radu and Mahesh K. Marina. Himloc: Indoor smartphone localization via activity aware pedestrian dead reckoning with selective crowdsourced wifi fingerprinting. In *Proc. IEEE IPIN 2013*, 2013.

[2] Xijia Wei, Zhiqiang Wei, and Valentin Radu. Sensor-fusion for smartphone location tracking using hybrid multimodal deep neural networks. *Sensors*, 21(22):7488, 2021.