End-to-End Machine Learning for Smartphone-based Indoor Localisation and Tracking using Recurrent Neural Networks

Xijia Wei and Valentin Radu

School of Informatics, University of Edinburgh

Indoor localization with smartphones is a favourite research topic for Mobile Systems researchers. This is due to the complexity of estimating accurate locations from noisy sensor data, and the potential benefit this can have on many location-based services. Various solutions to estimate indoor locations have been proposed to compensate for the obvious limitation of GPS inside buildings. Predominantly, these solutions rely on built-in sensors, such as the accelerometer, magnetometer, gyroscope and WiFi received signal strength, although integrating their signal streams to achieve a single accurate location estimation is not easy.

Currently, most indoor localization systems build on top of two main techniques, WiFi Fingerprinting and Pedestrian Dead Reckoning (PDR) [Radu and Marina, 2013]. Each of these two techniques functions on a firm set of equations defining the possible mobility frame, such as step counting, direction estimation and WiFi fingerprint matching. Their accuracy inherently depends on the well-defined mathematical formulation of these models, which identify the relationship between sensor data and location information. These are generally hard to determine to start with and are commonly fitted based on a small set of observations. We believe this is limiting the potential of indoor localization systems. Instead, we propose to avoid any intervention and hand-tuned formulation by relying on data alone and machine learning end-to-end to extract the relevant patterns that are not trivial for us to observe.

To replicate the construction of PDR, which starts from a known location and estimates consecutive locations based on sensor observations, we employ a range of sequential machine learning algorithms, commonly referred to as Recurrent Neural Networks (RNN), which estimates location from time windows of sensor data as input and an intermediary representation of previous estimations. Multimodal neural network constructions [Radu et al., 2018] permit the integration of a diverse set of sensors in different time window settings of the sensor signal.

Data is crucial to the proposed solutions, so we collected training sensor data in several runs over a long path on corridors, together with their ground truth location. We find that even with a reduced memory structure within the Long Short-Term Memory (LSTM) networks, estimations approximate the shape of the path well, with more than 80% of location estimations having error below 6 meters, while taking as input only inertial sensors (accelerometer, gyroscope and magnetometer) data. From a system perspective, our solution is energy-efficient without compromising on inference accuracy, which is achieved by reducing input data dimensionality (down sampling and PCA), making this ideal for mobile phones. We are expanding the system to capture more user mobility patterns including moving between floors.

References

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

[Radu et al., 2018] Radu, V., Tong, C., Bhattacharya, S., Lane, N., Mascolo, C., Marina, M., and Kawsar, F. (2018). Multimodal deep learning for activity and context recognition. *IMWUT*, 1(4).