Abstract

In the past decade, indoor localization has played a significant role in advancing location-based services across a wide range of domains. Since indoor areas are deprived of satellite signals, the indoor localization process is dependent on indoor-available signal sources like WiFi or Bluetooth. Fingerprint is the mostly used technique nowadays for identifying location based on these signal data, applying machine learning methods.

Despite the prominent success of this field, locating a user or object within a large public indoor space based on these indoor-available sensors face various challenges. One of the major challenges is repetitive site surveys, i.e., manually collecting labeled ground truth data to train the localization model. Although some of the existing works have addressed this issue by generating synthetic data for training localization models, there is a lack of qualitative analysis regarding the mapping of the generated data to different locations. Another genuine concern is determining how to partition an entire indoor environment into smaller sub-regions—each serving as a unique location unit for user or object identification—thereby defining the appropriate spatial granularity. Existing works are mostly focused on uniform granularity distribution, either in coarse-grained or fine-grained manner. But, is it a feasible approach for real-world solution? Especially when the access points are deployed in non-uniform manner?

In this thesis, we have investigated these two research directions in detail. An ensemble of clustering is proposed to groupify grossly labeled data having similar fingerprint characteristic, into sub-regions of non-uniform granularity. Using that clustering information, around

97% accurate localization is achieved on average. To address the challenge of signal fluctuations in case of fine-grained uniform granularity, a signal-to-image encoding algorithm is proposed to improve the localization performance using spatial analysis. When the encoded images are classified using proposed lightweight CNN model, around 99% accuracy is achieved. A new algorithm, GO-kDN, is proposed to update the granularity level of any fingerprint dataset based on instance hardness analysis to increase localization accuracy. For experimentation, a new dataset is collected from two metro stations, and it is made publicly available for further research. Applying the proposed approach, localization performance improved by 15% for at-grade-level and 35% for underground-level metro stations. To generate synthetic data and map it with proper location based on qualitative analysis, a Vectored labeled Generative Adversarial Network (VL-GAN) model is proposed. Up to 10% improvement in localization accuracy is observed after augmenting synthetic data generated by VL-GAN.