Abstract

In the past decade, WiFi-based Indoor Localization has emerged as a pivotal domain, offering transformative possibilities for location-based services in indoor environments. Most indoor public places are covered by WiFi today. So, reusing the distance sensitivity property of such signals is a sensible approach for localization where GPS signals are not available.

This research explores the requirement of dimensionality reduction within the context of WiFi-based Indoor Localization Systems (ILS). The objective of the work is to propose robust localization approach utilizing the existing Wi-Fi infrastructure. This involves refining the prediction accuracy by gearing towards distilling relevant information while mitigating the influence of noise and irrelevant features.

Dimensionality reduction encompasses both feature and instance space. In the context of feature space both Wrapper and Filter-based AP selections have been investigated. One of the core contribution of this thesis is that it combines meta-heuristic based feature selection with the notion of ensemble learning for addressing the dynamic contexts of the localization phase. The conditional ensemble has been constituted with the feature subsets obtained from the meta-heuristic algorithm-based feature selection approaches. This enhances the system's ability to adapt to diverse contexts, including device heterogeneity and environmental fluctuations. Both Genetic Algorithm and Binary Particle Swarm Optimization have been considered for the purpose. Further innovations unfold in the domain of dataset distillation, involved delving into Convolutional Autoencoding (CAE) principles and k-disagreeing neighbor scores. To tackle the effect of outliers that get induced during data collection, meta-heuristic based instance selection approaches have been proposed. This also ensures effective localization performance in constrained environments.

Experimentation has been carried out on collected dataset through multiple smartphone devices. Results shows that the feature-based pipeline achieved over 95% accuracy, reducing access points (APs) by 50-65%. Error deviation of 2.68m is achieved, which is acceptable for indoor user localization. The instance-based pipeline cuts dataset size by 40%, with a slight 2-3% accuracy dip but significant reduction in error deviation. The designed machine learning approaches are also found to be significantly applicable to other smartphone signal modalities and application domains such as, Human Activity Recognition, illuminating the transformative capacity of refined instance selection techniques. Edge level training analysis has also been employed to validate the performance of the model in constrained environment.