ABSTRACT

Multimodal sensing has facilitated the development of more innovative context-aware services, which actually form the core idea behind smart infrastructure development. However, with thousands of sensors continuously sensing the context, there can be a significant context pressure on the users concerning the privacy of the information captured. Motivated by this concern, we in this thesis look into the problems that plague the development of such services. Definitely, with a plethora of sensors sensing the context, the first concern that appears is regarding the presence of side channels that may breach the users' privacy. These information leaks often appear through alternate modalities that the user is unaware of and yet potentially capture signatures regarding the user's context.

To elaborately study this, we first start with the analysis of the alternate modalities like typing and smartphone engagement patterns for detecting mobility context. For this, we first collect a large amount of crowdsourced keystroke and application usage data. With the insights gained from the detailed analysis of this crowdsourced data, we understand that keystroke patterns indeed carry signatures to characterize the mobility state and can be used to sniff the mobility context even if the user has explicitly turned off the access to the location and locomotive sensors. Based on this understanding, we design an energy-efficient machine learning (ML)-based adversarial framework that can sniff the mobility data from the users' smartphones using only the keystroke and application usage patterns.

However, irrespective of whether the objective is to develop a dedicated context-aware service or an adversarial framework, one mandatory requirement in all these cases is the need for a significant volume of labeled data. With multimodal sensing in place, a massive volume of data is generated, which is challenging to annotate with the conventional human-in-the-loop-based approach. Motivated by this concern, we next develop an automated annotation framework that can generate labeled inertial measurement unit (IMU) data for personalized human activity recognition (HAR). Under the hood, the system precisely performs cross-modal change detection across the locomotive sensors and the environmental acoustic signatures, as an auxiliary modality, to generate precise labels for the IMU data without involving any human-in-the-loop.

Next, we extend this idea of automatic annotation of personalized locomotive sensing for smart infrastructures with two occupants leveraging on the inherent behavior of human beings while performing complex activities of daily living (ADL).

Finally, we investigate the general semantics surrounding annotations concerning HAR. From a detailed analysis of the human-in-the-loop-based approaches and the machine-assisted annotation systems, including our own frameworks, we observe a prevailing concern regarding the overall informativeness of the obtained (or generated) labels. As an informative label can provide the model with critical information for precise HAR, detailed labels are more of a necessity than a luxury. To understand this, we develop a framework that can assess the informativeness of the labels and provide feedback to the annotators in case re-annotation is required.

To summarize, we perform a longitudinal study of human activity context in the purview of multimodal sensing. Through these detailed analyses, we show that multimodal sensing could indeed cause severe information leakage. However, at the same time, there are several opportunities to use the multimodal information as auxiliary sources that can help design systems that can facilitate the development of intelligent context-aware services through precise HAR.

Keywords: human activity recognition; mobile computing; keystroke patterns; human-in-the-loop; change-points; locomotive sensing; acoustic sensing