Context-aware system for recognizing daily-activities using smartphone sensors

BELOUADAH Sakher - Université Paris Saclay

Abstract. *In recent decades, human activity recognition has been the subject of an important amount of research which enabled many applications in different areas, such as time management, healthcare and anomaly detection. Most of those works were based on using multiple special sensors and few address complex activities. In order to solve those issues, we propose a context-aware system based on the combination of ontological reasoning, GPS mining using k-nearest neighbors, and statistical recognition model using cascade neural networks. We first present some complex activity recognition models and discus their limitations. A general architecture of our approach is then presented along with a detailed description of each section of the system. Finally, we present the results obtained and discus the system's limitation and ideas that need to be addressed in future work.*

1. Proposed Approach

In order to address the problem of activity recognition, we propose a context-aware system using an ontological reasoning combined with cascade neural network and GPS mining technic. The statistical recognition model (neural network) is also combined with a symbolic location ontology which allows the system to infer the current activity of the user among the candidate activities. We prove that this is not only useful to refine the model by taking under consideration only the context related activities, but also to deal with the cold start when the user don't have sufficient data to train our system with. The main intuition of using the ontological reasoning is that the neural network infer the current user's activity by learning on raw data gathered from the set of smartphone sensors (Accelerometer, GPS, light...). But since it might be difficult to recognize the activity among a large set of possible activities, we use the ontological reasoning to refine the model by learning and recognizing only the set of activities of each context, for example: if the user never performs Gardening in the Kitchen, there is no need to consider it as a possible output in this context.

1.1. Architecture

Using the ontological model, we categories activities into two sets locomotion and complex daily activities. The former represents the set of activities that are recognizable using motion sensors features such as accelerometers, gyroscope, and GPS. The locomotion activities are detected by periodically waking up the device and reading short bursts of sensor data. It only makes use of the mentioned low power sensors in order to keep the power usage to a minimum. The later represents the set of complex daily human activities such as reading, eating, working, cooking...etc. In order to detect complex activities, a complex recognition model need to be used in order to be able to describe the context that helps recognizing current activity. Therefore, multiple sensors are combined together to generate data, and a non-parametric model is trained over those data.

1.2. Recognition

1-In case of locomotion, we start by detecting the user's motion state using the accelerometer,If the current state is "in Vehicle", we use a trajectory recognition process in order to detect the current trajectory and the type of vehicle the user is in. We address this task as a map matching problem, where each trajectory is represented by a sequence of GPS

points ordered based on the temporal information. The proposed model is based on K Nearest Neighbors algorithm that considers (1) the spatial-temporal information and (2) the marginal velocity of each trajectory. Many issues are not addressed in existing works, such as high volume of trajectory points, high variance of points gathering frequency, missing points and complexity optimization. In order to deal with the mentioned issues, we propose a feature mapping function Φ_1 that transforms the data representation from the GPS sequence space into R^{n*m} space. The new space is based on Military Grid Reference System coordinate system [16], where n and m are the height and width of the grid respectively. The new representation is used to process missing data using a linear interpolation, and improve accuracy using dilation. Then, another feature mapping function $\Phi_2 : R^{n*m} \mapsto N^d$ will map the sparse grid representation into a simple vectorial representation, which will be used to optimize the recognition process.
2-The second type of activities are **complex daily human activities**. Unlike locomotion activities, they are more complex to recognize due to: the amount of similarities between activities, none relevant (or useless) features used to recognize, amount of candidate activities and the limitation of using sensors of the smartphone only. In order to address this problem, we propose a combination of none-parametric model (Cascade Neural Networks) with a symbolic location ontology. The main intuition behind this combination is that statistical models cannot be used alone to distinguish similar activities among a huge set of activities (about 100) based only on raw data. Hence, the use of the symbolic location ontology helps refine the model by selecting a small set of candidate activities based on the correspondent context, and therefore infer the right activity.

1.3. Evaluation

The set of raw data values are represented as signals emitted by sensors embedded in the smartphone and/or other connected objects such as computers. These are time-related, hence the problem of recognition can be defined as follows: Having a set $T = \{T_0, .., T_{m-1}\}$ where the set of T_i represent Timeslots (time windows) of different sizes. And a set $A = \{A_0, .., A_{n-1}\}$ of possible activities. The goal is to find a function $f : T \mapsto A$ that approximates the real activity performed during T_i.
The recognition model is evaluated using a subject-dependent matter, in which a model is trained and tested for each user independently using his own data. The model's precision is then considered as the mean precision. We first started by performing a dry run using Decision tree model (DT) and Support vector machines (SVM), the results are then compared to the proposed model in term of precision. In order to improve the models precision, we run experiments using cross validation. The method used in cross validation is called K-fold with K equals to 6 in our case, and that is by dividing data into 6 subsets, we train our models 6 times. Each time, one of the 6 subsets is used as the test data and the 5 other subsets are combined to form the training data. The average error across all 6 runs is returned.

2. Conclusion

Our results show that by using the context-aware system, the problem of recognizing complex activities becomes simpler and that is by refining the model using ontology reasoning, and that's by considering smaller set of candidate activities based on the context.Then Combining different techniques for different contexts in order to recognize activities.