CHAPTER 4

# Probabilistic Ontology Model for Human Activity Recognition

*Pouya Foudeh and Naomie Salim*

## 4.1 INTRODUCTION

Human Activity Recognition (HAR) is one of the most applicable and yet challenging areas in context awareness. HAR is a multidisciplinary area of research, and it tackles several different issues. This research proposes an ontological model for sensor-based human activity recognition that facilitates dealing with uncertainty. In some applications, HAR systems must be able to process a significant amount of data in a reasonable time. As the proposed system is highly scalable, it can be used for batch processing. A model for storing ontologies in relational databases is proposed in the form of tables that contain ontology's semantic material with accompanying probability values.

Sensor-based HAR is preferred for many applications since it is more pervasive, has fewer privacy issues, and requires less computational process and storage space compared to camera-based HAR. In sensor-based recognition, Inertial Measurement Units (IMUs) and the Indoor Positioning System (IPS) track subjects' and environmental objects' movements and location. These sensors are basically analogue devices that generate continuous signals. IMU sensors' data are processed with signal processing techniques; data-driven, classification methods convert them to posture (e.g., sitting or walking), and low-level, fine-grained activities (e.g., taking a coffee cup with the right hand). IPS signals are also used to calculate subjects' location in a room or building. This information is used to find out high-level activities, such as resting time, performed by the subject.

Many algorithms have been presented for sensor based HAR, from simple statistical models to deep learning (Patel & Shah, 2019).

The main contributions of the current research are ability to deal with uncertainty, scalability and storing probabilistic ontological knowledge in ordinary relational databases (Foudeh & Salim, n.d). One of the challenges in HAR systems is uncertainty. The input information is essentially uncertain; sensors used in activity recognition are usually powered by unreliable batteries, data transmission is in a noisy wireless medium, and sensors might be displaced from their original position. Moreover, classification methods used for predicting low-level activities are not perfect. In short, there is no guarantee that whatever obtained from sensors' data is correct. However, the degree of belief, the probability of having correct information, is somehow calculable. When information such as low-level activities and location is available, one popular approach is to store them in a knowledge base system like an ontology system. Afterward, the knowledge base system does reason and infers high-level activities. In addition to probabilistic information obtained from sensors' data, known as ABox, the activity recognition knowledge base is also uncertain because the definition set, TBox, is also uncertain. For example, if Alice is certainly standing in place x in the kitchen taking a cup with her right hand and moving the chair with her left hand, she is probably in "tea-time" (80%), or she is in "cleaning time" (20%). This research models uncertainty with probabilistic representation and makes use of a probabilistic ontology to develop the proposed human activity recognition knowledge base. To the best of our knowledge, this is the first ontology-based activity recognition work with probabilistic observations from sensors.

Another aspect of HAR systems is their computing mode that can be real-time or batch processing. Real-time systems are for applications such as elderly monitoring and gaming, while batch processing is suitable for applications like employee

monitoring, on parole criminal monitoring, and medical or praxeological studies on people's behaviour. In real-time HAR systems, the processing time should be only less than or equal to the performing time and the window size should be small. However, batch processing recognition systems must be able to deal with a significant amount of data from several subjects, each performed in a long-time span. Therefore, scalability is one of the key challenges in these systems. This research proposes a probabilistic method for sensor-based human activity recognition to overcome both obstacles: uncertainty and scalability.

Storing data and reasoning about the knowledge base are difficulties of HAR batch processing. There are some knowledge management systems for storing and reasoning about conceptual knowledge bases. They are appropriate to use for a limited amount of data in research labs. On the other hand, Relational Database Management Systems (RDBMSs) are incredibly efficient, even though they are not designed to deal with complex knowledge structures. Decades of experience, billions of investments, and millions of active users have enabled RDBMSs to store and manage huge databases reliably and securely. For achieving scalability, this study designs a procedure to store activity recognition knowledge and do reasoning about them in a relational database. There are some research works on storing ontologies in databases. However, we didn't find any research on storing probabilistic ontologies on ordinary relational databases.

## 4.2 RELATED WORKS

A comprehensive review of the challenges and the state of the art of ontology-based human activity recognition, probabilistic data and knowledge, and ontology storage in relational databases can be found in this section.