Chapter 3

Mining RFID Behavior Data using Unsupervised Learning

Guénaël Cabanes
LIPN-CNRS UMR 7030

Younès Bennani
LIPN-CNRS UMR 7030

Dominique Fresneau
LEEC, FRANCE

ABSTRACT

Radio Frequency IDentification (RFID) is an advanced tracking technology that can be used to study the spatial organization of individual's spatio-temporal activity. The aim of this work is firstly to build a new RFID-based autonomous system which can follow individuals' spatio-temporal activity, a tool not currently available. Secondly, the authors aim to develop new tools for automatic data mining. In this paper, they study how to transform these data to investigate the division of labor, the intra-colonial cooperation and conflict in an ant colony. They also develop a new unsupervised learning data mining method (DS2L-SOM: Density-based Simultaneous Two-Level - Self Organizing Map) to find homogeneous clusters (i.e., sets of individual which share a similar behavior). According to the experimental results, this method is very fast and efficient. It also allows a very useful visualization of the results.

INTRODUCTION

Radio Frequency IDentification (RFID) is an advanced tracking technology. The RFID tags, which consist of a microchip and an antenna, must be used with a reader that can detect simultaneously a lot of tags in a single scan. A computer has to be used to store the data about the position of each tag for each scan in a database. This allows different analyses.

RFID systems can be used to study animal societies. Animal societies are dynamic complex systems characterized by numerous interactions between individual members. Such dynamic structures stem from the synergy of these interactions, the individual capacities in information processing and the diversity of individual responses (Fresneau et al., 1989). RFID, thanks to miniaturization, of-
fers the advantage of automation and overcomes the constraints imposed by video analyzes. Indeed, video recording allows long-duration tracking, however the time for analyzes highly increases with the number of individuals monitored. It also imposes strong constraints (as the need of a minimum illumination and high contrast between the animals and the environment) and it does not work when the ant is moving in a reverse position which doesn’t allow individual identification. The aim of this work is to develop a new RFID-based autonomous system to follow the spatio-temporal activity of groups, which is currently very difficult to study in its entirety and to develop new tools for automatic data processing. These objectives have necessarily led to an interdisciplinary project combining behavioral and complex systems sciences with computer and engineering sciences.

Since dynamic experimental data are extremely difficult to collect, behavioral sciences are dominated by static approaches (optimality, game theory). In order to understand the functioning of insect societies, the integration of both individual and collective types of analyzes is necessary. However, there are two essential challenges: the presence of autonomous units and the great influence of fluctuations. These challenges imply that adequate observation tools are available. Therefore, traceability technologies are greatly needed, e.g., RFID, which allows the automated monitoring of the localizations and of the movements of many individuals simultaneously. Those systems are now well developed and miniaturized enough to be used in insect societies (Streit et al., 2003) and they are now almost operational.

However, experiments using RFID generate large datasets which need suitable analysis methods to allow a comprehensive understanding of the link between events and reveal behavioral patterns. In this paper, we investigate how to transform these data to study the division of labor, the intra-colonial cooperation and conflict in an ant colony. Ants, often caricatured and little known, have nevertheless a huge ecological impact and are considered as major energy catalysts. Their complex underground nests contribute to soil ventilation and ecosystem equilibrium because of their predatory and detritivore diets. Ants are very diverse and the Formicidae (11000 described species) exhibit a great variety of social structures (Passera & Aron, 2005). Division of labor is one of their fundamental characteristics (Hölldobler & Wilson, 1990). According to the needs of the colony, each individual in the colony can assume a basic behaviors, such as: nursing the queen and the brood (i.e: eggs, larva and cocoon), transporting food or building material, hunting, and so on. Ant colonies face rapid changes of environmental conditions and constraints through an important individual flexibility. The dynamic component of this phenomenon is the hallmark of our research on social organization and on colony performance during migrations (change of nest). ARFID device has been developed for these study organisms. Based on marketed products, it requires little development. It consists of a network of RFID readers in a constrained space with compulsory passageways in an artificial nest. These readers are connected to a detector which sends the information to a computer.

However, in this study, we don’t have any prior knowledge about the structure of the social organisation of the colony, for example, the usual behavior of each ant and how many different behaviors can be expressed by the colony. In that case, most usual statistical approaches fail, as they need at least two different samples to compare. To analyze the internal structure of a unique data set, unsupervised classification methods (clustering methods), are very powerful. They allow an automatic detection of relevant sub-groups (or clusters) in a data set and are particularly suitable for data mining from experimental studies, for which we have generally little a priori information. Thus, in this paper, we analyze the collected data by using a new method of data mining, based on a clustering algorithm (DS2L-SOM: Density-based Simultaneous Two-Level - Self Organizing Map,