

Vers un algorithme multi-agents de clustering dynamique

Gaële Simon*, Dominique Fournier**, Bruno Mermet*

*GREYC CNRS UMR 6072, 6 Boulevard du Maréchal Juin 14050 CAEN cedex
{galele.simon, bruno.mermet}@univ-lehavre.fr,

**LITIS EA 4051, 25 rue Philippe Lebon, BP 540 76058 Le Havre cedex
dominique.fournier@univ-lehavre.fr

Résumé. Dans cet article, nous présentons un algorithme multi-agents de clustering dynamique. Ce type de clustering doit permettre de gérer des données évolutives et donc être capable d'adapter en permanence les clusters construits.

1 Introduction

Dans cet article, nous proposons une technique de clustering dynamique de données évolutives. Cette problématique est née de l'objectif initial de nos travaux visant à permettre, au cours de l'exécution d'un système multi-agents, de détecter des groupes d'agents liés à des phénomènes d'auto-organisation. On se trouve donc face à un problème de clustering dynamique qui présente les deux particularités suivantes : le cardinal de l'ensemble de données à clusteriser n'est pas constant et des données déjà clusterisées peuvent être modifiées du fait de l'évolution des agents correspondants.

Cela peut entraîner des modifications ou des réorganisations de l'ensemble existant de clusters. Ainsi, une méthode de clustering dynamique est nécessaire afin d'adapter continuellement l'ensemble des clusters afin qu'ils reflètent le mieux possible l'état courant des données.

2 Travaux connexes

Il existe de nombreux travaux portant sur les techniques de clustering où l'ensemble des données à clusteriser n'est pas totalement connu dès le départ comme en clustering classique. On trouve en particulier dans cette catégorie les techniques de clustering de flux de données et de flux de données évolutifs. Malheureusement, ces algorithmes ne prennent pas en compte le fait que des données déjà clusterisées puissent elles aussi évoluer.

Les travaux les plus proches de notre problématique concernent un algorithme de clustering de données mobiles présenté dans [Li et al. (2004)] : un micro-clustering est effectué en enrichissant les données d'un vecteur vitesse. Cependant, dans un deuxième temps, l'algorithme k-means doit être utilisé pour regrouper les micro-clusters, ce qui oblige à donner un nombre de clusters attendu et à ce que ce nombre soit constant.

3 Notre approche

Les algorithmes fournis de clustering semblent plus adaptés à la prise en compte de l'évolution des données. Ainsi, l'algorithme AntClass [Monmarché (2000)] associe successivement en quatre phases un algorithme de fourragement et l'algorithme k-means. Cette approche n'étant pas compatible avec l'aspect dynamique de notre problématique, nous avons décidé