ALGeoSPF: Un modèle de factorisation basé sur du clustering géographique pour la recommandation de POI

Jean-Benoît Griesner*, Talel Abdessalem*,**
Hubert Naacke*** Pierre Dosne*

*LTCI, Télécom ParisTech
Paris, France
griesner@telecom-paristech.fr,

**UMI CNRS IPAL, National University of Singapore
talel.abdessalem@telecom-paristech.fr

***Sorbonne Universités, UPMC Univ Paris 06, CNRS, LIP6 UMR 7606
Hubert.Naacke@lip6.fr

Résumé. La recommandation de points d'intérêts est devenue une caractéristique essentielle des réseaux sociaux géo-localisés qui a accompagné l'émergence des échanges massifs de données digitales. Cependant les faibles densités de points d'intérêts visités par les utilisateurs rendent le problème difficile à traiter, d'autant plus que les espaces de mobilité des utilisateurs sont très hétérogènes, allant de la ville au monde entier. Dans ce papier nous explorons l'impact d'une approche de clustering spatial sur la qualité de la recommandation. Notre approche est basée sur un modèle de factorisation de matrices de Poisson et un réseau social inféré des différents comportements de mobilité. Nous avons conduit une évaluation comparative des performances de notre approche sur un jeu de données réaliste. Les résultats expérimentaux montrent que notre approche permet une précision supérieure aux techniques de recommandation alternatives.

1 Introduction

La recommandation de *points d'intérêts* (ou POI pour *Point Of Interest*) consiste à proposer à un utilisateur une liste de POI qu'il pourrait être intéressé de visiter. Ce problème est devenu une composante majeure des réseaux sociaux géo-localisés (ou LBSN pour *Location-Based Social Networks*) permettant aux utilisateurs de découvrir de nouveaux POI mais également aux POI d'augmenter leur attractivité. Il faut toutefois tenir compte de différentes caractéristiques des LBSN, à savoir : (i) **sparsité** : la densité de la matrice utilisateur-POI est très faible, (ii) **fréquences de visite** : nous savons seulement combien de fois un utilisateur a été localisé dans un lieu, et (iii) **feedback implicite** : nous ne disposons pas d'une évaluation explicite du POI par l'utilisateur.