

Temporalité dans la recommandation de spectacles

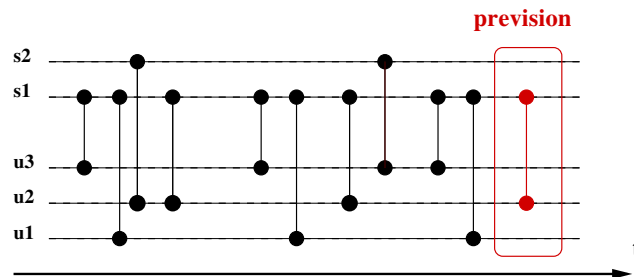
Lionel Tabourier (LIP6 – CNRS / UPMC – Paris)

stages@complexnetworks.fr, <http://complexnetworks.fr>

L'objet du stage est de construire une méthode de recommandation, qui soit basée sur un formalisme original tirant partie de l'information temporelle contenue dans les données, dans la perspective d'être appliquée à un contexte industriel.

Delight est une plateforme SaaS à destination des producteurs de spectacle vivant qui leur permet d'améliorer la connaissance de leurs publics, d'accroître la fréquentation et les recettes de billetterie de leurs spectacles. Ce type de recommandation présente plusieurs difficultés qui lui sont propres. D'abord, les spectacles restant peu de temps à l'affiche, il n'est pas possible de collecter les informations sur les achats avant de proposer la recommandation. Autrement dit, il s'agit d'une situation de *démarrage à froid*, dans lesquelles beaucoup d'approches classiques, tel que le filtrage collaboratif, sont inapplicables directement. Dans ce contexte, on cherchera à privilégier des méthodes hybrides [Bur07]. Par ailleurs, les pratiques de consommation du spectacle vivant sont extrêmement hétérogènes : une petite minorité d'individus génère l'essentiel des recettes. Un autre enjeu important du problème consiste donc à fidéliser un public moins averti.

Une piste pour résoudre ces difficultés consiste à rassembler les spectacles en groupes significatifs en fonction de leurs caractéristiques (genre, artistes, ...). On cherche alors à prévoir les interactions entre spectateurs et classes de spectacles. Un tel choix permet de représenter les données à l'aide d'un flot de liens biparti, c'est-à-dire une séquence de triplets (u, s, t) représentant une interaction (un achat) entre un utilisateur u , une classe de spectacle s et l'instant t où l'action est réalisée, comme représenté sur le schéma ci-dessous.



La recommandation peut alors être comprise comme un problème de prévision de liens dans ce flot : on cherche à inférer quelles interactions vont se produire et durant quelle période. Pour ce faire, on développe des méthodes qui tirent partie à la fois de la structure et des

régularités temporelles apparaissant dans le flot. Celles-ci sont au carrefour des techniques de prévisions de liens dans les graphes et de la prévision de séries temporelles [ATL17].

Dans un second temps, les recommandations seront distribuées dans les groupes, en fonction des spectacles disponibles dans chaque région géographique, et des spectateurs potentiellement amateurs de ces spectacles. Au fil du temps, en agrégeant de nouvelles informations dans le système, il sera possible d'affiner la définition des classes, et de sélectionner les caractéristiques des flots les plus pertinentes pour améliorer les performances du moteur de recommandation.

Le stage s'adresse de préférence à un ou une étudiante de Master 2 ou de fin de cycle d'école d'ingénieur. Il/elle travaillera de manière partagée entre le LIP6 (UPMC) et Delight, en répartissant son temps entre l'un et l'autre lieu en fonction des besoins pour le projet. Les principales compétences recherchées sont une connaissance générale de la fouille de données, une certaine capacité d'adaptation à des méthodes nouvelles et de l'intérêt pour les thématiques interdisciplinaires.

Références

- [ATL17] Thibaud Arnoux, Lionel Tabourier, and Matthieu Latapy. Combining structural and dynamic information to predict activity in link streams. In *International Symposium on Foundations and Applications of Big Data Analytics*, 2017.
- [Bur07] Robin Burke. The adaptive web. chapter hybrid web recommender systems. 2007.