

Analyse de séquences d'événements avec TraMineR

Nicolas S. Müller*, Matthias Studer*, Alexis Gabadinho*, Gilbert Ritschard*

*Institut d'études démographiques et des parcours de vie, Université de Genève
{nicolas.muller, matthias.studer, alexis.gabadinho, gilbert.ritschard}@unige.ch
<http://mephisto.unige.ch/traminer/>

TraMineR et séquences d'événements

Les méthodes de fouille de motifs séquentiels fréquents ont donné lieu, depuis le travail fondateur de Agrawal et Srikant (1995), à un grand nombre de travaux de recherche. La librairie TraMineR pour l'environnement statistique R cherche à rendre ces méthodes accessibles aux chercheurs de divers domaines désirant analyser ou visualiser des ensembles de séquences d'événements. Il est à noter que TraMineR n'est pas adapté à l'étude d'une séquence unique d'événements. La librairie est disponible sur le CRAN (Comprehensive R Archive Network) et possède une aide en ligne ainsi qu'un manuel de l'utilisateur détaillé (Gabadinho et al., 2009).

TraMineR offre les fonctionnalités pour séquences d'événements suivantes :

- l'importation et la conversion de plusieurs formats de données et leur transformation en séquences d'événements,
- l'extraction des motifs séquentiels fréquents, en prenant en compte des contraintes temporelles telles que le temps maximum entre deux événements ou la taille de la fenêtre,
- le calcul de mesure de dissimilarité entre séquences d'événements afin de procéder, par exemple, à une classification non-supervisée des séquences (Studer et al., 2010)
- l'identification des motifs séquentiels fréquents les plus discriminants par rapport à une autre variable d'intérêt en se basant sur des tests du χ^2 de Pearson (Studer et al., 2010),
- la visualisation des motifs séquentiels discriminant à l'aide des résidus,

