

Extraction de séquences fréquentes avec intervalles d'incertitude

Asma Ben Zakour^{*,**}, Sofian Maabout^{*}, Mohamed Mosbah^{*}, Marc Sistiaga^{**}

^{*}LaBRI, Université Bordeaux 1, CNRS, FRANCE

maabout@labri.fr, mosbah@labri.fr

^{**}2MoRO Solutions, Bidart, FRANCE

marc.sistiaga@2moro.fr, asma.ben-zakour@2moro.fr

Résumé

Lors de l'extraction des séquences, la granularité temporelle est plus ou moins importante selon les besoins des utilisateurs et les contraintes du domaine d'application. Nous proposons un algorithme d'extraction de séquences fréquentes par intervalles à partir de séquences à estampilles temporelles discrètes. Nous intégrons une relaxation des contraintes temporelles en introduisant la définition de "séquences temporelles par intervalles" (STI). Ces intervalles reflètent une incertitude sur les occurrences précises des événements. Nous formalisons ce nouveau concept en exhibant certaines de ses propriétés et nous menons quelques expériences afin de comparer (qualitativement) nos résultats avec une autre proposition assez proche de la nôtre.

Summary

Sequential patterns mining has been used in several domains. We note that time granularity is more or less important regarding the application domain. In this paper we propose a frequent *interval time sequences* (ITS) extraction technique from discrete temporal sequences using a sliding window approach in order to relax time constraints. The extracted sequences offer an interesting overview of the original data by allowing a temporal leeway on the extraction process. We formalize the ITS extraction under classical time and support constraints and conduct some experiments on synthetic data for validating our proposal.