HAL
open science

ASAX : Segmentation adaptative basée sur la quantité d'information pour SAX

Lamia Djebour, Reza Akbarinia, Florent Masseglia

To cite this version:

Lamia Djebour, Reza Akbarinia, Florent Masseglia. ASAX : Segmentation adaptative basée sur la quantité d'information pour SAX. BDA 2021 - 37e Conférence sur la Gestion de Données - Principes, Technologies et Applications, Oct 2021, Paris, France. lirmm-03468535

HAL Id: lirmm-03468535

https://hal-lirmm.ccsd.cnrs.fr/lirmm-03468535

Submitted on 7 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ASAX : Segmentation adaptative basée sur la quantité d'information pour SAX

Lamia Djebour
INRIA & LIRMM, Univ Montpellier
France
lamia.djebour@inria.fr

Reza Akbarinia
INRIA & LIRMM, Univ Montpellier
France
reza.akbarinia@inria.fr

Florent Masseglia
INRIA & LIRMM, Univ Montpellier
France
florent.masseglia@inria.fr

RÉSUMÉ

Les approches existantes pour le calcul de similitude entre séries temporelles sont au cœur de nombreuses tâches d'analyse de données. Étant donné les volumes de données considérés, ou simplement le besoin de les traiter rapidement, ces approches s'appuient souvent sur des représentations alternatives, plus courtes, qui résument les séries d'origine avec une perte d'information acceptable. Les comparaisons de séries temporelles qui se basent sur ces représentations sont alors approximatives, ce qui fait de la précision un enjeu majeur. Nous présentons et évaluons expérimentalement ASAX, une nouvelle approche pour la segmentation de séries temporelles avant qu'elles soient transformées en représentations symboliques. ASAX réduit de manière significative la perte d'information due aux fractionnements dans les différentes étapes du calcul de la représentation. Nous fournissons des garanties théoriques sur la borne inférieure des mesures de similitude entre séries temporelles, et nos expériences illustrent l'intérêt de notre méthode sur l'état de l'art, en particulier avec un gain de précision significatif.

© 2021, Copyright is with the authors. Published in the Proceedings of the BDA 2021 Conference (October 25-28, 2021, En ligne, France). Distribution of this paper is permitted under the terms of the Creative Commons license CC-by-nc-nd 4.0.
© 2021, Droits restant aux auteurs. Publié dans les actes de la conférence BDA 2021 (25-28 octobre 2021, En ligne, France). Redistribution de cet article autorisée selon les termes de la licence Creative Commons CC-by-nc-nd 4.0.