

# Découverte de motifs intelligibles et caractéristiques d'anomalies dans les traces unitaires

Olivier Cavadenti<sup>\*,\*\*</sup>, Victor Codocedo<sup>\*</sup>, Mehdi Kaytoue<sup>\*</sup>, Jean-François Boulicaut<sup>\*</sup>

<sup>\*</sup>Université de Lyon. CNRS, INSA-Lyon, LIRIS. UMR5205, F-69621, France.

<sup>\*\*</sup>Actemium Saint-Etienne

contact : prenom.nom@insa-lyon.fr

**Résumé.** De nombreuses industries manufacturières s'intéressent aujourd'hui à l'exploitation des grandes collections de traces unitaires. Les applications sont multiples et vont du simple "reporting" à la détection de fraudes en passant par la gestion de retours ou encore la mise en évidence d'incohérences dans les circuits de distribution. Une étape importante consiste à détecter des anomalies dans des collections de traces. Si les travaux concernant la détection d'anomalies sont assez nombreux, peu permettent de caractériser les anomalies détectées par une description intelligible. Étant donné un ensemble de traces unitaires, nous développons une méthode d'extraction de motifs pour détecter et contextualiser des comportements non conformes à un modèle expert (fourni ou construit à partir des données). Le degré d'anomalie est alors quantifié grâce à la proportion du nombre de mouvements des objets qui ne sont pas prévus dans le modèle expert. Cette recherche est financée partiellement par un programme industriel qui ne permet ni de dévoiler le contexte concret ni de parler des données réelles. Ainsi, nous validons empiriquement la valeur ajoutée de la méthode proposée par l'étude de traces de mobilité dans un jeu vidéo : nous pouvons alors discuter d'un motif qui explicite les raisons de l'inexpérience de certains joueurs.

## 1 Introduction

Avec la dissémination de nombreux systèmes de capteurs, de très grands volumes de données sont accessibles sous la forme de collections de traces. Ces traces correspondent à la séquence des événements *captés* dans un système qui définit ce qui est *captable*. Les traces modélisent alors la vie des objets dans ce système. Le type de traces qui motive cette recherche est celui des traces unitaires de produits manufacturés qui peuvent être tracés depuis leur fabrication jusqu'à leur vente en transitant via un réseau logistique éventuellement complexe. Les traces unitaires codent les comportements des objets, dont certains sont prévisibles, qui permettent d'avoir accès à de nombreuses informations spatio-temporelles ou sémantiques sur les processus appliqués à ces objets. Parallèlement, ces systèmes de traçage possèdent eux-mêmes de nombreuses caractéristiques comme les distances et positions entre les capteurs ainsi que leur type, leur propriétaire, les données qu'ils fournissent, leur état, etc. On peut donc se demander si les comportements des objets vérifient bien les processus attendus par ceux qui ont