

# Extraction de motifs pour la détection d’anomalies dans des graphes : application à la fraude dans les marchés publics

Lucas Potin\*, Rosa Figueiredo\*, Vincent Labatut\*, Christine Largeron\*\*

\* Laboratoire Informatique d’Avignon, F-84911, Avignon, France  
{prénom.nom}@univ-avignon.fr,

\*\* Laboratoire Hubert Curien UMR 5516, F-42023, Saint-Etienne, France  
christine.largeron@univ-st-etienne.fr,

**Résumé.** Dans le cadre des marchés publics, il existe plusieurs indicateurs, appelés red flags, permettant d’estimer le risque de fraude. Ces red flags sont calculés en fonction des attributs spécifiques de chaque contrat et sont ainsi dépendants du bon remplissage des notices d’attributions. Nous proposons une méthode basée sur l’extraction de motifs pour la détection d’anomalies dans des graphes. Cette approche générique vise à identifier les sous-graphes associés à la présence de red flags, afin de construire un ensemble de nouveaux indicateurs. Ces motifs peuvent ensuite être utilisés dans les cas où les red flags sont manquants.

## 1 Introduction

Les marchés publics désignent l’achat de biens et de services par une autorité publique (le client), auprès d’une personne morale de droit public ou privé (le fournisseur). Lorsqu’un marché dépasse les seuils européens, l’appel d’offres ainsi que l’avis d’attribution de ce marché doivent être publiés au *Journal Officiel de l’Union Européenne* (JOUE). La version en ligne de ce journal, appelée *Tenders Electronic Daily*<sup>1</sup> (TED), publie plus de 650 000 avis de marché par an<sup>2</sup>. Par conséquent, le secteur des marchés publics fournit une énorme quantité de données accessibles au public.

Historiquement, les anomalies dans les marchés publics, qui font référence à des comportements suspicieux, sont liées à des caractéristiques spécifiques des contrats, appelées *red flags*, et utilisées comme indicateurs de fraude potentielle (Fazekas et Tóth, 2014; Ferwerda et al., 2017), par exemple, la modification du prix du contrat en cours de procédure, ou la réception d’une seule offre pour un appel d’offre donné (National Fraud Authority, 2016). Mais les informations nécessaires au calcul de ces red flags ne sont pas toujours disponibles. Dans les données françaises du TED, certains attributs sont largement absents (Potin et al., 2022), tel le nombre de réponses à un appel d’offres (vide pour 30% des lots), permettant uniquement le calcul de red flags *partiels*.

---

1. <https://ted.europa.eu/>

2. <https://ted.europa.eu/TED/main/HomePage.do>