

La simulation relaxée de graphes pour la recherche de motifs

IEEE/ACM ASONAM 2018, August 28-31, 2018, Barcelona, Spain

Abdelmalek Habi*, Brice Effantin*, Hamamache Kheddouci*

*Univ Lyon, Université Claude Bernard Lyon 1, CNRS, LIRIS, F-69622
Lyon, France

{abdelmalek.habi, brice.effantin-dit-toussaint, hamamache.kheddouci}@univ-lyon1.fr

Résumé. La recherche de motifs de graphe est l'une des opérations principales de la recherche des correspondances d'une requête dans un graphe donné. Dans ce contexte, trouver des solutions garantissant l'optimalité en termes de précision et de temps de calcul est un problème de recherche difficile et d'actualité. Différents modèles ainsi que leurs algorithmes appropriés ont été proposés pour la recherche de motifs dans les graphes de données. Cependant, l'inconvénient majeur est leur limitation à trouver des réponses significatives entraînant le problème des réponses vides. Dans cet article nous introduisons un nouveau modèle pour la recherche de motifs de graphe permettant un certain type d'assouplissement de requêtes afin d'éviter ce problème. Ensuite nous développons un algorithme efficace basé sur des techniques d'optimisation pour trouver les k -meilleurs réponses selon notre modèle. Nos expérimentations sur quatre ensembles de données réelles démontrent l'efficacité de notre approche.

1 Introduction

Les graphes sont des structures mathématiques constituant un outil de modélisation et de représentation universel utilisé dans une large gamme d'applications réelles. La recherche de motifs de graphes (*RMG*) est l'une des opérations fondamentales sur laquelle reposent la recherche et l'analyse des graphes de données. Soit $G(V, E, l, \Sigma)$ un graphe de données et $Q(V_q, E_q, f_v)$ un motif de graphe (requête) où : V (V_q), E (E_q), l (f_v) et Σ représentent respectivement l'ensemble de nœuds, l'ensembles des arêtes, la fonction des étiquettes et l'univers des étiquettes dans le graphe (la requête). Le problème *RMG* consiste à trouver toutes les correspondances de Q dans G , notées par $M(Q, G)$. Typiquement, ce problème est défini en termes de :

- l'isomorphisme de sous-graphes (Ullmann, 1976) : $M(Q, G)$ est constitué de tous les sous-graphes G' de G auxquels Q est isomorphe, *i.e.*, il existe une fonction bijective $h : V_q \mapsto V$ telle que $(u, u') \in Q$ si et seulement si $(h(u), h(u')) \in G'$; ou
- la simulation de graphes (Henzinger et al., 1995) : $M(Q, G)$ est une relation binaire $R \subseteq V_q \times V$ qui vérifie : (1) $\forall u \in V_q, \exists v \in V \mid (u, v) \in R$, et (2) $\forall (u, v) \in R$ et $\forall (u, u') \in Q, \exists (v, v') \in G \mid (u', v') \in R$.