

Evaluation rapide du diamètre d'un graphe

Christian Belbeze^{*,**}, Max Chevalier^{*,***}
Chantal Soule-Dupuy^{*,**}

*Institut de Recherche en Informatique de Toulouse

**Université Toulouse 1 Capitole, 2 rue du Doyen Gabriel Marty, 31042 Toulouse Cedex 9

***Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse Cedex 9
christian@belbeze.com, chevalier@irit.fr, soule@irit.fr

Résumé. Lors de l'analyse de graphes, il est important de connaître leurs propriétés afin de pouvoir par exemple identifier leur structure et les comparer. Une des caractérisations importante de ces graphes repose sur le fait de déterminer s'il s'agit ou non d'un "petit monde". Pour ce faire, la valeur du diamètre du graphe est essentielle. Or la mesure du diamètre est pour un très grand graphe, une opération extrêmement longue. Nous proposons un algorithme en deux phases qui permet d'obtenir rapidement une estimation du diamètre d'un graphe avec une proportion d'erreur faible. En réduisant cet algorithme à une seule phase et en acceptant une marge d'erreur plus élevée, nous obtenons une estimation très rapide du diamètre. Nous testons cet algorithme sur deux grands graphes de terrain (plus d'un million de nœuds) et comparons ses performances avec celles d'un algorithme de référence BFS (Breadth-First Search). Les résultats obtenus sont décrits et commentés.

1 Introduction

Depuis quelques années, la recherche sur les grands graphes a soulevé de plus en plus d'intérêt. Quelle que soit la nature des réseaux représentés, il est indéniable que le nombre d'éléments (nœuds et liaisons) les constituant et pris en compte dans leurs études n'a cessé d'augmenter. L'identification de communautés dans les réseaux sociaux est un de ces domaines d'étude sur lequel nous nous sommes focalisés dans nos travaux et qui nous a amené à étudier les caractéristiques des graphes sous-jacents à ces réseaux (Belbeze et al., 2009).

Bien que de nature différente, les réseaux de "grande taille" en général (réseaux d'ordinateurs, réseaux sociaux, réseaux de mots, pages Web reliées par des hyperliens et autres réseaux d'échange), ont des caractéristiques qui comportent un grand nombre de similitudes. Ces similitudes ont permis de créer un type de graphe nommé "Grand graphe de terrain". Par définition, ces graphes venus du monde réel ne sont donc pas issus d'une formule mathématique. Ils existent sur le terrain et les nœuds se doivent d'avoir une existence physique. La phrase de Watts et Al. en 1998 (Watts et Strogatz, 1998) "la plupart des graphes de terrain ont des propriétés non-triviales en commun" peut être considérée comme la consécration de leur domaine d'étude. Parmi ces propriétés "non-triviales en commun", la valeur du diamètre d'un graphe