

Mesures d'hétérogénéité sémantique des systèmes P2P non-structurés

Thomas Cerqueus*, Sylvie Cazalens* et Philippe Lamarre*,**

*LINA - CNRS - UMR 6241, Université de Nantes, **INRIA
2, rue de la Houssinière - 44000 Nantes, France
{thomas.cerqueus, sylvie.cazalens, philippe.lamarre}@univ-nantes.fr

Résumé. L'autonomie des participants dans les systèmes P2P pour le partage de données peut conduire à une situation d'hétérogénéité sémantique dans le cas où les participants utilisent leurs propres ontologies pour représenter leurs données. Dans cet article nous commençons par définir des mesures de disparité entre participants en considérant leurs contextes sémantiques. En considérant la topologie du système et les disparités entre participants, nous proposons des mesures d'hétérogénéité sémantique d'un système P2P non-structuré.

1 Introduction

Dans les systèmes pair-à-pair conçus pour le partage de données, les participants sont autonomes. Dans le cas où ceux-ci utilisent des ontologies pour représenter leurs données, nous ne pouvons pas supposer qu'ils utilisent tous la même, car leurs objectifs et leurs besoins peuvent être différents. Dans de tels systèmes, il peut devenir difficile de faire en sorte que les pairs puissent partager des données et se comprendre. Dans cet article, nous focalisons sur la cause même du problème, à savoir le fait que des participants n'utilisent pas tous la même ontologie ou, même si tel était le cas, ne l'appréhendent pas exactement de la même façon. C'est ce que nous appelons l'*hétérogénéité sémantique* du système. Intuitivement, plus le système est hétérogène, plus il est difficile d'assurer l'interopérabilité en son sein. L'objectif de cet article est de définir des mesures d'hétérogénéité sémantique d'un système P2P.

La section 2 introduit un certain nombre de définitions. Puis la section 3 considère les disparités entre les contextes sémantiques de deux participants et propose deux mesures permettant de capturer différents aspects des disparités. La section 4 focalise sur la définition de l'hétérogénéité d'un système en exploitant les mesures définies dans la section 3. Là encore, plusieurs mesures sont proposées afin de prendre en compte différentes facettes de l'hétérogénéité. Nous concluons en section 5.

2 Définitions de base

Chaque participant d'une communauté ou d'un système peut vouloir utiliser sa propre ontologie. Et même en se référant à une ontologie commune, le contexte spécifique d'un participant (interprétation, usage, ...) peut amener des "distorsions" par rapport à un autre. Ainsi