



Localisation de Données à travers une Architecture de Médiation Pair-à-Pair

Nicolas Lumineau - Anne Doucet - Stéphane Gançarski

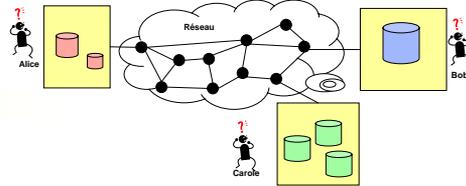
Mots-clés : Pair à pair , Médiation , interopérabilité, Regroupement sémantique

Sous – problème n°1 :
Offrir aux utilisateurs une vision homogène et pertinente du réseau

Médiation Pair à Pair

Problématique générale :

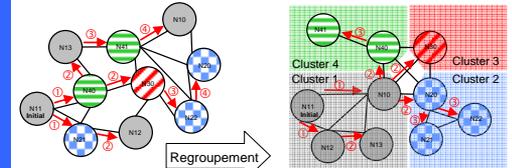
Partage de données hétérogènes distribuées à large échelle



Sous – problème n°2 :
Localiser efficacement un ensemble de ressources pertinentes.

Organisation sémantique d'un réseau pair à pair

Localisation et organisation



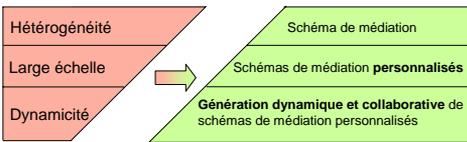
Proposition :

Construction d'un réseau pair à pair de médiateurs interopérables.

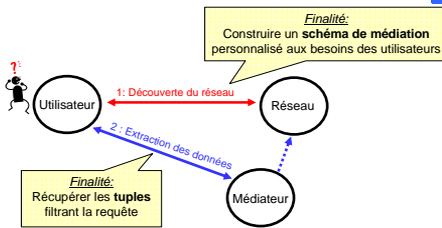
Médiateurs : solution pour traiter l'hétérogénéité des données
Pair à pair : solution pour traiter la large échelle

Verrous scientifiques et solutions

Le partage de données passe par la gestion de :



Principe



Modèle de schémas

Enrichissement sémantique des schémas via le modèle de schéma : **PAXschema** (PeerActiveXschema) : extension de schema XML permettant le stockage de méta-infomations :

- ✓ Identifiantes
- ✓ Localisatrices
- ✓ Qualitatives
- ✓ Descriptives

Deux types de schémas :

- ↳ schémas publiés (générés par fournisseurs)
- ↳ schémas thématiques (générés par système)

Solution linguistique : Algorithme d'intégration des schémas basé sur les descriptions des schémas publiés

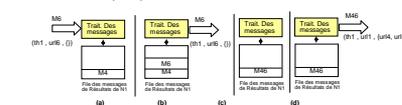
Intégration dynamique et collaborative

Maintien de la cohérence entre les schémas thématiques construits et la réalité du réseau.

Deux stratégies de collaboration :

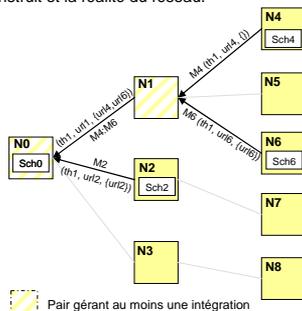
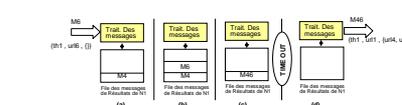
Stratégie optimiste :

Observation des messages retournés + Collaboration + Réconciliation par pair initiateur.



Stratégie pessimiste :

Rétention des messages retournés + Collaboration



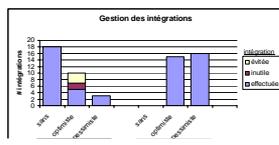
Transmission des messages de résultats

Résultats

↳ Evaluation de la faisabilité par prototypage et simulation

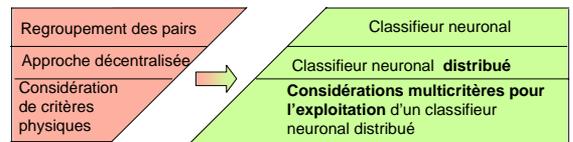
↳ Comparatifs des stratégies optimistes et pessimistes.

- Intérêt important de la collaboration
- Stratégie pessimiste permet une meilleure répartition de la collaboration: charge moyenne par pair collaborant : Opt 2.1 ; Pess 1.6
- Evaluation du temps imparti difficile à établir



Verrous scientifiques et solutions

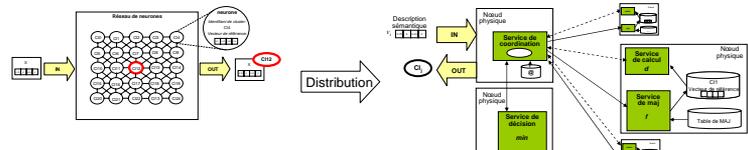
Notre objectif est de rapprocher logiquement les pairs sémantiquement proches (par rapport à leur contenu). Cela passe par :



Classifieur neuronal distribué

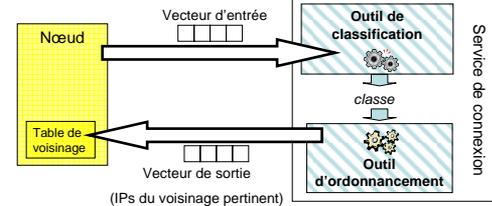
Principe des cartes auto-organisatrices de Kohonen.

Décomposition des différentes tâches liées à la classification en services pouvant être ainsi distribués. Les neurones sont alors distribués sur différents super-pairs.



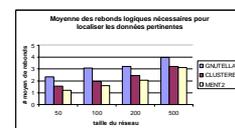
Interrogation du service de connexion

(représentation sémantique du contenu d'un nœud)



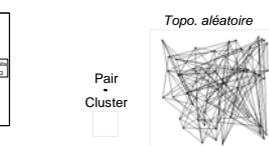
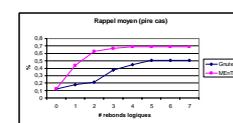
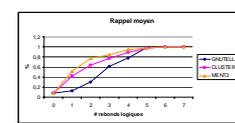
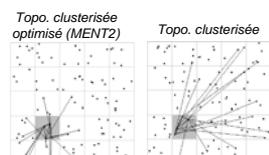
Résultats : comparatif des stratégies d'organisation

Evaluation des performances par simulation



Comparatifs des stratégies de clusterisation.

- Impact important de la clusterisation
- Intérêt de distinguer les liens intra-clusters des liens inter-clusters.



Publications sur ces travaux

- Lumineau, N., Doucet, A., Gançarski, S., Thematic Schemas Building for Mediation-based P2P Architecture, International Workshop On Database Interoperability (InterDB 2005), Namur, Belgium, April 2005.
- Lumineau, N., Doucet, A., Defude, B., Cluster Entries for Semantic Organization of Peer-to-Peer Network, Semantics for Grid Databases (ICSNW'04), Paris, June 2004.
- Lumineau, N., Doucet, A., Defude, B., VENISE: Content-based Clustering for Data Sharing in Peer-to-Peer Architecture, In Proc. Base de Données Avancées (BDA'04), Montpellier, October 2004