



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Πρότυπο Σύστημα Ομότιμων Κόμβων
Βασισμένο σε Σχήματα RDF

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΕΙΡΗΝΗΣ Κ. ΣΠΥΡΟΠΟΥΛΟΥ

Επιβλέπων: Τιμολέων Σελλής
Καθηγητής Ε.Μ.Π.

ΕΡΓΑΣΤΗΡΙΟ ΣΥΣΤΗΜΑΤΩΝ ΒΑΣΕΩΝ ΓΝΩΣΕΩΝ ΚΑΙ ΔΕΔΟΜΕΝΩΝ
Αθήνα, Οκτώβριος 2005



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών
Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων

Πρότυπο Σύστημα Ομότιμων Κόμβων Βασισμένο σε Σχήματα RDF

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΕΙΡΗΝΗΣ Κ. ΣΠΥΡΟΠΟΥΛΟΥ

Επιβλέπων: Τιμολέων Σελλής
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 19η Οκτωβρίου 2005.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....

Τιμολέων Σελλής
Καθηγητής Ε.Μ.Π.

.....

Ιωάννης Βασιλείου
Καθηγητής Ε.Μ.Π.

.....

Νικόλαος Μήτρου
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2005

(Υπογραφή)

.....

ΕΙΡΗΝΗ ΣΠΥΡΟΠΟΥΛΟΥ

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

© 2005 – All rights reserved



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών
Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων

Copyright ©–All rights reserved Ειρήνη Σπυροπούλου, 2008.

Με επιφύλαξη παντός δικαιώματος.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Ευχαριστίες

Θα ήθελα καταρχήν να ευχαριστήσω τον καθηγητή κ. για την επίβλεψη αυτής της διπλωματικής εργασίας και για την ευκαιρία που μου έδωσε να την εκπονήσω στο εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων. Επίσης ευχαριστώ ιδιαίτερα τον Δρ. για την καθοδήγησή του και την εξαιρετική συνεργασία που είχαμε. Τέλος θα ήθελα να ευχαριστήσω τους γονείς μου για την καθοδήγηση και την ηθική συμπαράσταση που μου προσέφεραν όλα αυτά τα χρόνια.

Περίληψη

Ένα σύστημα ομότιμων κόμβων αποτελείται από ένα σύνολο αυτόνομων υπολογιστικών κόμβων στο Διαδίκτυο, οι οποίοι συνεργάζονται με σκοπό την ανταλλαγή δεδομένων. Στα συστήματα ομότιμων κόμβων που χρησιμοποιούνται ευρέως σήμερα, η αναζήτηση πληροφορίας γίνεται με χρήση λέξεων κλειδιών. Η ανάγκη για πιο εκφραστικές λειτουργίες, σε συνδυασμό με την ανάπτυξη του Σημασιολογικού Ιστού, οδήγησε στα συστήματα ομότιμων κόμβων βασισμένα σε σχήματα. Στα συστήματα αυτά κάθε κόμβος χρησιμοποιεί ένα σχήμα με βάση το οποίο οργανώνει τα τοπικά διαθέσιμα δεδομένα. Για να είναι δυνατή η αναζήτηση δεδομένων στα συστήματα αυτά υπάρχουν δύο τρόποι. Ο πρώτος είναι όλοι οι κόμβοι να χρησιμοποιούν το ίδιο σχήμα κάτι το οποίο δεν είναι ευέλικτο. Ο δεύτερος τρόπος δίνει την αυτονομία σε κάθε κόμβο να επιλέγει όποιο σχήμα θέλει και απαιτεί την ύπαρξη κανόνων αντιστοίχισης μεταξύ των σχημάτων για να μπορούν να αποτιμώνται οι ερωτήσεις. Αυτός ο τρόπος προσφέρει ευελιξία όμως δεν υποστηρίζει την αυτόματη δημιουργία και τη δυναμική ανανέωση των κανόνων, που είναι απαραίτητες για ένα σύστημα ομότιμων κόμβων.

Στόχος της διπλωματικής εργασίας είναι η ανάπτυξη ενός συστήματος ομότιμων κόμβων βασισμένο σε σχήματα το οποίο (α) θα επιτρέπει μια σχετική ευελιξία στην χρήση των σχημάτων και (β) θα δίνει την δυνατότητα μετασχηματισμού ερωτήσεων χωρίς την ανάγκη διατύπωσης κανόνων αντιστοίχισης μεταξύ σχημάτων, χρησιμοποιώντας κόμβους με σχήματα RDF που αποτελούν υποσύνολα-όψεις ενός βασικού σχήματος (καθολικό σχήμα).

Λέξεις Κλειδιά

Σύστημα ομότιμων κόμβων, Σύστημα ομότιμων κόμβων βασισμένο σε σχήματα, Σημασιολογικός Ιστός, RDF/S, RQL, Jxta

Abstract

A peer-to-peer system is a set of autonomous computing nodes (the peers) which cooperate in order to exchange data. The peers in the peer-to-peer systems that are widely used today, rely on simple keyword selection in order to search for data. The need for richer facilities in exchanging data, as well as, the evolution of the Semantic Web, led to the evolution of the schema-based peer-to-peer systems. In those systems every node uses a schema to organize the local data. So there are two ways in order for data search to be feasible. The first but not so flexible way implies that every node uses the same schema. The second way gives every node the flexibility to choose a schema according with its needs, but on the same time requires the existence of mapping rules in order for queries to be replied. This way though, doesn't offer automatic creation and dynamic renewal of the mapping rules which would be essential for peer-to-peer systems.

This diploma thesis aims to the development of a schema-based peer-to-peer system that allows a certain flexibility for schema selection and on the same time enables query transformation without the use of mapping rules. The peers use RDF schemas that are subsets (views) of a big common schema called global schema.

Keywords

Peer-to-peer, Schema-based peer-to-peer, Semantic Web, RDF/S, RQL, Jxta

Περιεχόμενα

Ευχαριστίες	1
Περίληψη	3
Abstract	5
Περιεχόμενα	8
Κατάλογος Σχημάτων	9
Κατάλογος Πινάκων	11
1 Εισαγωγή	13
1.1 Αντικείμενο της διπλωματικής	14
1.2 Οργάνωση του τόμου	15
2 Θεωρητικό υπόβαθρο	17
2.1 Συστήματα ομότιμων κόμβων	17
2.1.1 Τι είναι τα συστήματα ομότιμων κόμβων	17
3 Περιγραφή θέματος	19
3.1 Σχετικές εργασίες	19
4 Ανάλυση και σχεδίαση	21
4.1 Ανάλυση - περιγραφή αρχιτεκτονικής	21
4.1.1 Διαχωρισμός υποσυστημάτων	21
4.1.2 Περιγραφή υποσυστημάτων	21
5 Υλοποίηση	23
5.1 Λεπτομέρειες υλοποίησης	23
5.1.1 Αλγόριθμοι	23
5.2 Περιγραφή κλάσεων	24
5.2.1 public class FirstUi	24

6 Έλεγχος	25
6.1 Μεθοδολογία Ελέγχου	25
6.2 Αναλυτική παρουσίαση ελέγχου	25
7 Επίλογος	27
7.1 Συμπεράσματα	27
7.2 Μελλοντικές Επεκτάσεις	28
8 Μεταφράσεις Ξένων όρων	31
Βιβλιογραφία	30

Κατάλογος Σχημάτων

4.1	Αριτεκτονική Απλού Κόμβου	22
-----	-------------------------------------	----

Κατάλογος Πινάκων

Κεφάλαιο 1

Εισαγωγή

Ο Παγκόσμιος Ιστός αποτελεί χώρο διακίνησης τεράστιου όγκου πληροφοριών. Ωστόσο, η συντριπτική πλειοψηφία των πληροφοριών του Ιστού, είναι προσανατολισμένη προς τον άνθρωπο-χρήστη και δεν είναι κατανοητή από τις εφαρμογές. Για να αξιοποιηθεί λοιπόν η διαθέσιμη πληροφορία και να γίνει πιο εύκολη η ανταλλαγή και η επεξεργασία της, ο Παγκόσμιος Ιστός εξελίσσεται στο Σημασιολογικό Ιστό.

Ο Σημασιολογικός Ιστός, είναι μια εξέλιξη του σημερινού Ιστού, μέσα στον οποίο δίνεται καλά ορισμένο νόημα στην πληροφορία που διακινείται, διευκολύνοντας τη συνεργασία μεταξύ υπολογιστή και ανθρώπου [;]. Πιο συγκεκριμένα δίνει τη δυνατότητα καλύτερης πρόσβασης σε μεγάλο όγκο πηγών πληροφορίας, καθώς και πιο αποτελεσματικής διακίνησης των πληροφοριών, χρησιμοποιώντας δεδομένα που τις περιγράφουν και ονομάζονται “μεταδεδομένα”. Η καλύτερη γνώση της σημασίας, της χρήσης και της ποιότητας των πηγών διευκολύνει σημαντικά τη δυνατότητα πρόσβασης σε πηγές του Ιστού και την αυτόματη επεξεργασία του περιεχομένου που υπάρχει διαθέσιμο στο Διαδίκτυο βάσει του νοήματος και όχι μόνο της μορφής της πληροφορίας.

Ένα από τα πιο βασικά θέματα για την ανάπτυξη του Σημασιολογικού Ιστού είναι το να μπορούν οι υπολογιστές να ανταλλάσσουν δεδομένα μεταξύ εφαρμογών. Σε ένα ανοιχτό περιβάλλον όπως είναι ο Σημασιολογικός Ιστός χρειάζεται ένα ευέλικτο και δυναμικό μοντέλο ανταλλαγής δεδομένων όπως είναι τα συστήματα ομότιμων κόμβων (Peer-to-Peer systems).

Ένα σύστημα ομότιμων κόμβων αποτελείται από ένα σύνολο αυτόνομων υπολογιστικών κόμβων, οι οποίοι συνεργάζονται με σκοπό την ανταλλαγή δεδομένων. Τα συστήματα ομότιμων κόμβων που χρησιμοποιούνται ευρέως σήμερα κυρίως για την ανταλλαγή αρχείων μουσικής, έχουν πολύ μικρές δυνατότητες διαχείρισης δεδομένων. Η αναζήτηση πληροφορίας στα περισσότερα από αυτά γίνεται με χρήση λέξεων κλειδιών (keyword-based search).

Η ανάγκη για πιο εκφραστικές λειτουργίες, σε συνδυασμό με την ανάπτυξη του Σημασιολογικού Ιστού, οδήγησε στα συστήματα ομότιμων κόμβων που είναι βασισμένα σε σχήματα (schema-based peer-to-peer systems). Στα συστήματα αυτά κάθε κόμβος χρησιμοποιεί ένα σχήμα με βάση το οποίο οργανώνει τα τοπικά διαθέσιμα δεδομένα. Οι τεχνολογίες του Σημασιολογικού Ιστού δίνουν τη δυνατότητα οργάνωσης των δεδομένων μέσω σχημάτων που τα περιγράφουν.

Το πλαίσιο RDF είναι ένα τέτοιο εργαλείο αναπαράστασης μεταδεδομένων. Σε ένα RDF αρχείο ορίζονται δηλώσεις για αντικείμενα του Ιστού όπως σελίδες, συγγραφείς, προγράμματα κ.τ.λ. Μια επέκταση του πλαισίου RDF είναι το RDF Schema το οποίο παρέχει μηχανισμούς περιγραφής σχετικών αντικειμένων του Ιστού καθώς και των σχέσεων μεταξύ τους. Το RDF Schema βασίζεται σε κλάσεις και ιδιότητες έννοιες γνωστές από το χώρο των Αντικειμενοστρεφών συστημάτων. Η βασική διαφορά είναι ότι στο πλαίσιο RDF οι ιδιότητες ορίζονται ανεξάρτητα από τις κλάσεις.

Χρησιμοποιώντας λοιπόν τις τεχνολογίες του Σημασιολογικού Ιστού μπορούμε να δημιουργήσουμε συστήματα ομότιμων κόμβων με αυξημένη διαλειτουργικότητα τα οποία θα ανταλλάσσουν μεταξύ τους πληροφορία με νόημα και θα έχουν τη δυνατότητα διατύπωσης ερωτήσεων πιο εκφραστικών από αυτές που βασίζονται σε λέξεις κλειδιά.

1.1 Αντικείμενο της διπλωματικής

Το βασικό ζήτημα που προκύπτει για τα συστήματα ομότιμων κόμβων που είναι βασισμένα σε σχήματα, είναι πώς θα μπορούν οι κόμβοι να αναζητούν και να ανταλλάσσουν δεδομένα, διατηρώντας την αυτονομία τους. Δύο προσεγγίσεις έχουν προταθεί στην βιβλιογραφία:

1. Η πρώτη προσέγγιση απαιτεί να υπάρχει ένα κεντρικό σχήμα το οποίο θα χρησιμοποιούν όλοι οι κόμβοι [;]. Οι ερωτήσεις διατυπώνονται και αποτιμούνται με βάση το ίδιο σχήμα. Μια τέτοια λύση θα ήταν καλή για περιβάλλοντα με καθορισμένα όρια, όπως για παράδειγμα το τοπικό δίκτυο ενός οργανισμού. Όμως σε ένα ανοιχτό περιβάλλον όπως είναι ο Παγκόσμιος Ιστός χρειάζεται ένα πιο ευέλικτο μοντέλο που να επιτρέπει την χρήση πολλών σχημάτων.
2. Η δεύτερη προσέγγιση δίνει την αυτονομία σε κάθε κόμβο να επιλέγει όποιο σχήμα θέλει. Οι ερωτήσεις διατυπώνονται με βάση ένα σχήμα και αποτιμούνται με βάση άλλα σχήματα, μέσω μιας διαδικασίας μετασχηματισμού ερωτήσεων (query reformulation). Η διαδικασία αυτή απαιτεί την ύπαρξη κανόνων αντιστοίχισης (mapping rules) [;]. Όμως, σε ένα σύστημα ομότιμων κόμβων οι κόμβοι μπορούν να μπαίνουν και να βγαίνουν στο δίκτυο συνεχώς. Δεν είναι γνωστό επομένως εκ των προτέρων τα ζευγάρια των κόμβων μεταξύ των οποίων πρέπει να υπάρχουν κανόνες αντιστοίχισης. Επίσης, οι κανόνες αυτοί φτιάχνονται χειρωνακτικά και είναι δύσκολη η συντήρησή τους.

Αντικείμενο της διπλωματικής είναι η ανάπτυξη ενός συστήματος ομότιμων κόμβων βασισμένο σε σχήματα το οποίο (α) θα επιτρέπει μια σχετική ευελιξία στην χρήση των σχημάτων και (β) θα δίνει την δυνατότητα μετασχηματισμού ερωτήσεων χωρίς την ανάγκη διατύπωσης κανόνων αντιστοίχισης μεταξύ σχημάτων. Το σύστημα δηλαδή βρίσκεται ανάμεσα στα δύο μοντέλα που περιγράφηκαν παραπάνω, από πλευράς ευελιξίας και δίνει τη δυνατότητα αυτόματου μετασχηματισμού ερωτήσεων. Χρησιμοποιεί κόμβους με σχήματα RDFS που αποτελούν υποσύνολα-όψεις (views) ενός βασικού σχήματος (καθολικό σχήμα).

Ένα παράδειγμα εφαρμογής του συστήματος αυτού θα ήταν η ανταλλαγή βιβλιογραφικών δεδομένων μεταξύ των ερευνητών. Κάθε ερευνητής θα συμμετείχε σε αυτό το σύστημα ομότι-

μων κόμβων με ένα δικό του RDF σχήμα σύμφωνα με το οποίο θα οργάνωνε τις δημοσιεύσεις του και ταυτόχρονα θα μπορούσε να αναζητήσει ανάλογα δεδομένα από άλλους κόμβους. Σ' ένα τέτοιο σύστημα θα μπορούσαν να συμμετέχουν ως κόμβοι εκτός από μεμονωμένοι ερευνητές και εργαστήρια ή και συνέδρια.

1.2 Οργάνωση του τόμου

Η εργασία αυτή είναι οργανωμένη σε επτά κεφάλαια: Στο Κεφάλαιο 2 δίνεται το θεωρητικό υπόβαθρο των βασικών τεχνολογιών που σχετίζονται με τη διπλωματική αυτή. Αρχικά περιγράφονται τα δίκτυα ομότιμων κόμβων, στη συνέχεια το πλαίσιο RDF και τέλος δίνεται μια μελέτη των γλωσσών ερωτήσεων για RDF. Στο Κεφάλαιο 3 αρχικά περιγράφονται οι σχετικές με το θέμα εργασίες και στη συνέχεια δίνεται ο στόχος της συγκεκριμένης εργασίας. Στο Κεφάλαιο 4 παρουσιάζεται η ανάλυση και η σχεδίαση του συστήματος, δηλαδή η περιγραφή των υποσυστημάτων και των εφαρμογών του. Η περιγραφή της υλοποίησης του συστήματος, με ανάλυση των βασικών αλγορίθμων καθώς και λεπτομέρειες σχετικά με τις πλατφόρμες και τα προγραμματιστικά εργαλεία που χρησιμοποιήθηκαν δίνεται στο Κεφάλαιο 5. Στο Κεφάλαιο 6 παρουσιάζεται ο έλεγχος καλής λειτουργίας του συστήματος με βάση ένα συγκεκριμένο σενάριο χρήσης. Τέλος στο Κεφάλαιο 7 δίνεται η συνεισφορά αυτής της διπλωματικής εργασίας, καθώς και μελλοντικές επεκτάσεις.

Κεφάλαιο 2

Θεωρητικό υπόβαθρο

Στο κεφάλαιο αυτό παρουσιάζονται αναλυτικά οι τρεις βασικές τεχνολογίες που έχουν σχέση με την εργασία αυτή, δηλαδή τα συστήματα ομότιμων κόμβων, το πλαίσιο RDF και οι γλώσσες ερωτήσεων για RDF.

2.1 Συστήματα ομότιμων κόμβων

2.1.1 Τι είναι τα συστήματα ομότιμων κόμβων

Στα μεγάλα καταναμεμημένα συστήματα όπως είναι ο Παγκόσμιος Ιστός, γίνονται εμφανή τα προβλήματα του παραδοσιακού μοντέλου πελάτη/εξυπηρετητή: Οι πηγές πληροφορίας βρίσκονται μαζεμένες σε λίγους κόμβους (εξυπηρετητές) στους οποίους συνδέονται πάρα πολλοί πελάτες [;].

Οι αρχές που διέπουν τα συστήματα ομότιμων κόμβων είναι οι εξής:

- Η αρχή του μοιράσματος των πόρων.
- Η αρχή της αυτοοργάνωσης.

Σύμφωνα με το συντακτικό αυτό, το παράδειγμα που αναπαριστάται στο Σχήμα ;; γράφεται ως εξής:

```
1.<?xmlversion="1.0"?>
2.<rdf:RDFxmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3.    xmlns:dc="http://purl.org/dc/elements/1.1/"
4.    xmlns:extterms="http://www.example.org/terms/">
5.  <rdf:Description rdf:about="http://www.example.org/index.html">
6.    <extterms:creation-date>August 16, 1999</extterms:creation-date>
7.    <dc:language>en</dc:language>
8.    <dc:creator rdf:resource="http://www.example.org/staffid/85740"/>
9.  </rdf:Description>
10.</rdf:RDF>
```


Κεφάλαιο 3

Περιγραφή θέματος

Στο κεφάλαιο αυτό αρχικά γίνεται μια περιγραφή των συστημάτων ομότιμων κόμβων που είναι βασισμένα σε σχήματα (schema-based peer-to-peer systems). Στη συνέχεια περιγράφονται τρία βασικά συστήματα που ανήκουν σε αυτή την κατηγορία, καθώς και ένα σύστημα για τη διαχείριση RDF σχημάτων, και τέλος αναλύεται ο στόχος της παρούσας εργασίας.

3.1 Σχετικές εργασίες

Οι βάσεις δεδομένων εισήγαγαν ένα τρόπο αποθήκευσης και ανάκτησης των δεδομένων που βασιζόταν στο σχήμα [;]. Τα πρώτα συστήματα ομότιμων κόμβων που περιγράψαμε στην Υποενότητα 2.1.2 έδιναν μεγάλη σημασία στην αρχιτεκτονική του συστήματος και την δρομολόγηση των ερωτήσεων και λιγότερη στον τρόπο αναπαράστασης και τις δυνατότητες αναζήτησης. Η αναζήτηση σε αυτά τα συστήματα ομότιμων κόμβων γίνεται με βάση προκαθορισμένα χαρακτηριστικά - δείκτες, ή με προσπάθεια αντιστοίχισης μιας λέξης κλειδί.

Η ανάγκη λοιπόν για πιο εκφραστικές λειτουργίες οδήγησε στα συστήματα ομότιμων κόμβων τα οποία είναι βασισμένα σε σχήματα (schema based peer-to-peer systems). Πρόκειται για ομότιμες υποδομές διαχείρισης δεδομένων που όμως διατηρούν όλα τα χαρακτηριστικά των συστημάτων ομότιμων κόμβων.

Κεφάλαιο 4

Ανάλυση και σχεδίαση

Στο κεφάλαιο αυτό παρουσιάζεται η μελέτη που έγινε για την υλοποίηση του συστήματος. Αρχικά περιγράφεται η αρχιτεκτονική του συστήματος και γίνεται ο διαχωρισμός του στα επιμέρους υποσυστήματα, ενώ στη συνέχεια περιγράφονται οι εφαρμογές του συστήματος.

4.1 Ανάλυση - περιγραφή αρχιτεκτονικής

Στην ενότητα αυτή παρουσιάζεται η ανάλυση του συστήματος και ο χωρισμός του σε υποσυστήματα όσον αφορά την αρχιτεκτονική.

4.1.1 Διαχωρισμός υποσυστημάτων

Το σύστημα αποτελείται από τους απλούς κόμβους και ένα κόμβο διαχειριστή. Στο σημείο αυτό αναλύουμε το σύστημα ενός απλού κόμβου, το οποίο αποτελείται από τα εξής υποσυστήματα:

- Υποσύστημα δημιουργίας σχήματος.
- Υποσύστημα ενσωμάτωσης δεδομένων στο σχήμα.
- Υποσύστημα επικοινωνίας κόμβου.

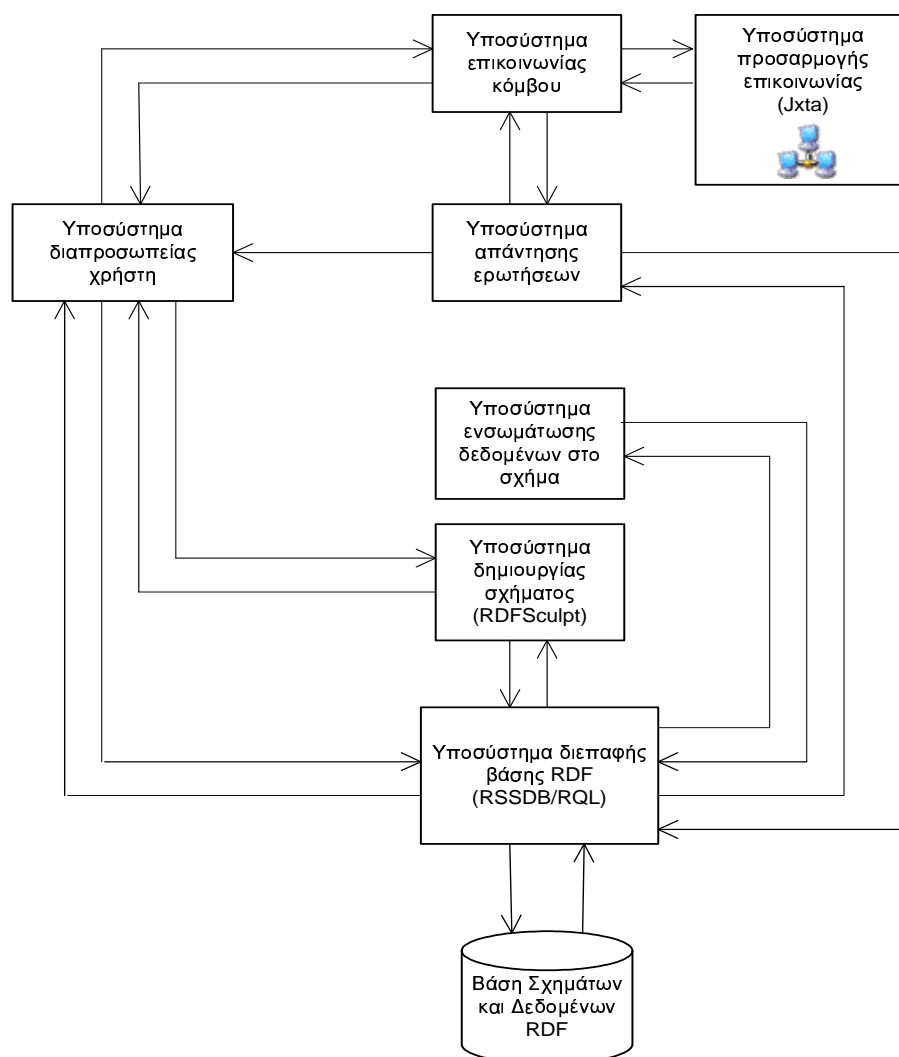
Το Σχήμα 4.1 απεικονίζει

4.1.2 Περιγραφή υποσυστημάτων

Παρακάτω δίνεται λεπτομερής περιγραφή για καθένα από τα συστήματα που αναφέραμε. Η περιγραφή αυτή γίνεται με βάση τα διαγράμματα ροής δεδομένων.

Υποσύστημα δημιουργίας σχήματος

Το υποσύστημα αυτό



Σχήμα 4.1: Αριτεκτονική Απλού Κόμβου

Κεφάλαιο 5

Υλοποίηση

Στο κεφάλαιο αυτό περιγράφεται η υλοποίηση του συστήματος, με βάση τη μελέτη που παρουσιάστηκε στο προηγούμενο κεφάλαιο. Αρχικά παρουσιάζεται η πλατφόρμα και τα προγραμματιστικά εργαλεία που χρησιμοποιήθηκαν. Στη συνέχεια δίνονται οι λεπτομέρειες υλοποίησης για τους βασικούς αλγόριθμους του συστήματος καθώς και η δομή του κώδικα.

5.1 Λεπτομέρειες υλοποίησης

Στην ενότητα αυτή παρουσιάζονται οι βασικοί αλγόριθμοι που αναπτύχθηκαν καθώς και λεπτομέρειες σχετικά με την υλοποίηση της επικοινωνίας των κόμβων.

5.1.1 Αλγόριθμοι

Αλγόριθμος εισαγωγής δεδομένων

Όταν ένας κόμβος εισέρχεται για πρώτη φορά στο σύστημα, αρχικά δημιουργεί το σχήμα που θέλει χρησιμοποιώντας το `RDFSculpt`. Στη συνέχεια.....

Κατασκευή του διανύσματος `groupedMapping`.

Περιέχει ομαδοποιημένα τα στοιχεία του `mapping` που ανήκουν στην ίδια κλάση.

Το διάνυσμα `groupedMapping` έχει τη μορφή:

```
[[[Κλάση1,Κυριολεκτικό1],Χαρακτηριστικό1],[Κλάση1,Κυριολεκτικό2],Χαρακτηριστικό2],...],[[Κλάση2,Κυριολεκτικό3],Χαρακτηριστικό3],[Κλάση2,Κυριολεκτικό4],Χαρακτηριστικό4],...]]
```

Για κάθε εγγραφή

Δημιούργησε αντίγραφο του `groupedMapping`, που ονομάζεται `imapping`

Όσο το `imapping` έχει στοιχεία

 Πάρε το πρώτο στοιχείο του διανύσματος έστω `classMapping`

 Βάλε την κλάση που ανήκει στο πρώτο στοιχείο, στο διάνυσμα `classesToWrite`

 Όσο το διάνυσμα `classesToWrite` έχει στοιχεία

 Πάρε το στοιχείο-κλάση που βρίσκεται στην αρχή του διανύσματος

Παράδειγμα

Έστω ότι ο κόμβος έχει επιλέξει να συμμετέχει στο σύστημα με το RDF σχήμα που φαίνεται στο Σχήμα. Έστω επίσης ότι από το SQL ερώτημα που έχει κάνει στη σχεσιακή βάση, έχει προκύψει η όψη που φαίνεται στον Πίνακα ;;. Για τις ανάγκες του παραδείγματος θεωρούμε ότι η όψη αυτή περιέχει μόνο μία εγγραφή.

.....

5.2 Περιγραφή κλάσεων

Στην ενότητα αυτή δίνεται μια σύντομη περιγραφή των κλάσεων, των πεδίων και των μεθόδων που τις απαρτίζουν.

5.2.1 `public class FirstUi`

Η κλάση αυτή κατασκευάζει την οθόνη εισαγωγής του χρήστη στο σύστημα.

Πεδία

- `private GridBagLayout blayout`
Το layout για όλα τα Panel.
- `private GridBagConstraints con`
Τα constraints για το layout.
- `private Icon arrowR`
Εικονίδιο για το κουμπί Next.

Μέθοδοι

- `public FirstUi()`
Ο κατασκευαστής της κλάσης ο οποίος καλεί την `createEntryFrame()`.
- `private void createEntryFrame()`
Μέθοδος που κατασκευάζει το ενφραμε.

Κεφάλαιο 6

Έλεγχος

Στο κεφάλαιο αυτό γίνεται ο έλεγχος καλής λειτουργίας του συστήματος.

6.1 Μεθοδολογία Ελέγχου

Ο έλεγχος του συστήματος αυτού πραγματοποιήθηκε με τη χρήση ενός σεναρίου λειτουργίας. Σύμφωνα με το σενάριο αυτό θεωρούμε ότι στο σύστημα υπάρχουν τρεις κόμβοι (peer1,peer2,peer3). Θεωρούμε επίσης ότι οι κόμβοι peer2 και peer3 έχουν ήδη σχήμα και δεδομένα. Το σχήμα του peer2 φαίνεται στο Σχήμα ;;

Επίσης η τοπολογία του συστήματος έχει ως εξής: ο peer2 είναι γείτονας του peer1 και ο peer3 γείτονας του peer2.

Αρχικά λοιπόν θα δημιουργήσουμε σχήμα για τον κόμβο peer1 και στη συνέχεια θα εισάγουμε σε αυτό δεδομένα εξετάζοντας έτσι την καλή λειτουργία του υποσυστήματος δημιουργίας σχήματος και του υποσυστήματος εισαγωγής δεδομένων. Στη συνέχεια από τον κόμβο αυτό στέλνουμε ερωτήσεις στους υπόλοιπους για τον έλεγχο του υποσυστήματος απάντησης ερωτήσεων και επικοινωνίας κόμβων.

6.2 Αναλυτική παρουσίαση ελέγχου

Στην ενότητα αυτή παρουσιάζουμε αναλυτικά τον έλεγχο του συστήματος σύμφωνα με το σενάριο που περιγράφηκε στην προηγούμενη ενότητα.

Όταν ξεκινάει να λειτουργεί ο κόμβος peer1 αρχικά εμφανίζεται η οθόνη που φαίνεται στο Σχήμα ;;. Η οθόνη αυτή είναι η οθόνη εισαγωγής στο σύστημα στην οποία ο χρήστης εισάγει το όνομα της βάσης στην οποία είναι αποθηκευμένα τα δεδομένα του καθώς και το όνομα της βάσης στην οποία θα αποθηκεύει τα RDF δεδομένα.

Μόλις ο χρήστης πατήσει το κουμπί Next τότε εισάγεται στο σύστημα

Κεφάλαιο 7

Επίλογος

7.1 Συμπεράσματα

Τα συστήματα ομότιμων κόμβων, προκειμένου να υποστηρίζουν πιο εκφραστικές λειτουργίες αναπαράστασης και αναζήτησης δεδομένων, εξελίχθηκαν στα συστήματα ομότιμων κόμβων τα οποία βασίζονται στις τεχνολογίες του Σημασιολογικού Ιστού για την αναπαράσταση των δεδομένων μέσω σχημάτων που τα περιγράφουν (Schema-based peer-to-peer systems).

Στα συστήματα αυτά κάθε κόμβος χρησιμοποιεί ένα σχήμα για την αναπαράσταση των δεδομένων του. Όμως σε ένα σύστημα ομότιμων κόμβων, κάθε κόμβος έχει διαφορετικές απαιτήσεις αναπαράστασης δεδομένων. Επομένως πρέπει να υπάρχει ευελιξία στην επιλογή σχήματος. Τα συστήματα που έχουν προταθεί μέχρι τώρα και παρέχουν αυτή την ευελιξία, για να είναι δυνατή η αναζήτηση πληροφορίας, απαιτούν την ύπαρξη κανόνων αντιστοίχισης μεταξύ των σχημάτων με βάση τους οποίους να μετασχηματίζονται οι ερωτήσεις. Όμως δεν υποστηρίζεται ακόμα αυτόματη δημιουργία και δυναμική ανανέωση των κανόνων, που είναι απαραίτητα για τα συστήματα ομότιμων κόμβων.

Η συνεισφορά της παρούσας διπλωματικής εργασίας έχει δύο σκέλη. Το πρώτο αφορά τη δημιουργία ενός πλήρους συστήματος ομότιμων κόμβων βασισμένο σε σχήματα RDF το οποίο παρέχει: (α) την υποδομή για την επικοινωνία των κόμβων, (β) μηχανισμό δημιουργίας σχήματος, (γ) μηχανισμό ενσωμάτωσης σχεσιακών δεδομένων στο σχήμα με τη χρήση αντιστοιχίσεων που δημιουργεί ο χρήστης με τη βοήθεια ειδικής διαπροσωπείας, (δ) ευέλικτη διαπροσωπεία χρήστη για τη διατύπωση ερωτημάτων και (ε) μηχανισμό απάντησης και επεξεργασίας ερωτήσεων.

Το δεύτερο σκέλος αφορά το γεγονός ότι το συγκεκριμένο σύστημα προσφέρει μια σχετική ευελιξία ως προς την επιλογή του σχήματος από τον κάθε κόμβο, ενώ ταυτόχρονα δίνει τη δυνατότητα μετασχηματισμού ερωτήσεων χωρίς τη χρήση κανόνων αντιστοίχισης. Συγκεκριμένα, τα σχήματα των κόμβων αποτελούν υποσύνολα—όψεις (views) ενός βασικού σχήματος που ονομάζεται καθολικό σχήμα. Εχμεταλλευόμενοι λοιπόν το γεγονός ότι τα σχήματα αυτά είναι συμβατά μεταξύ τους, έχουμε τη δυνατότητα ελέγχου της ικανοποιησιμότητας μιας ερώτησης και μετατροπής της όπου χρειάζεται, χρησιμοποιώντας τόσο το σχήμα του κόμβου όσο και το καθολικό σχήμα.

Συμπερασματικά το σύστημα που αναπτύχθηκε στα πλαίσια αυτής της διπλωματικής είναι ένα πλήρες σύστημα ομότιμων κόμβων βασισμένο σε σχήματα, το οποίο καθιστά δυνατή την αναζήτηση της πληροφορίας με ένα διαφορετικό τρόπο απ' ό,τι τα προϋπάρχοντα συστήματα.

7.2 Μελλοντικές Επεκτάσεις

Το σύστημα που αναπτύχθηκε στα πλαίσια αυτής της διπλωματικής εργασίας θα μπορούσε να βελτιωθεί και να επεκταθεί περαιτέρω, τουλάχιστον ως προς τρεις κατευθύνσεις. Συγκεκριμένα, αναφέρονται τα ακόλουθα:

- Ενσωμάτωση διαδικασίας επιλογής σχήματος με βάση το οποίο ο κόμβος θα συμμετέχει στο σύστημα. Έτσι όπως έχει σχεδιαστεί το σύστημα, κάθε κόμβος έχει τη δυνατότητα να δημιουργήσει πολλά σχήματα και να αποθηκεύσει δεδομένα σε περισσότερα από ένα. Ως σχήμα του κόμβου (με βάση το οποίο απαντάει τις ερωτήσεις), θεωρείται το τελευταίο στο οποίο αποθήκευσε δεδομένα. Η δυνατότητα επιλογής θα του παρείχε περισσότερη ευελιξία.
- Δυνατότητα αντιστοίχισης δεδομένων τα οποία να μην είναι αποθηκευμένα σε βάση δεδομένων αλλά σε αρχεία. Η αποδέσμευση από τη βάση δεδομένων θα έκανε το σύστημα πιο εύκολο στην εγκατάσταση και τη χρήση.
- Αξιολόγηση του συστήματος ως προς τη συμπεριφορά του αν συμμετέχει σε αυτό μεγάλος αριθμός κόμβων (scalability testing) και αν χρησιμοποιηθεί ένα πολύ μεγάλο καθολικό σχήμα. Η αξιολόγηση αυτή αφορά την ταχύτητα με την οποία ένας κόμβος παίρνει απαντήσεις σε μια ερώτηση καθώς και την ποιότητα των απαντήσεων.

Βιβλιογραφία

- [1] Tim Berners-Lee, James Hendler και Ora Lassila. The semantic web. *Scientific American*, 284(5):34–43, 2001.
- [2] Alon Y. Halevy, Zachary G. Ives, Peter Mork και Igor Tatarinov. Piazza: data management infrastructure for semantic web applications. Στο *WWW*, σελίδες 556–567, 2003.
- [3] Giorgos Kokkinidis και Vassilis Christophides. Semantic query routing and processing in p2p database systems: The ics-forth sqpeer middleware. Στο *EDBT Workshops*, τόμος 3268 στο *Lecture Notes in Computer Science*, σελίδες 486–495. Springer, 2004.
- [4] Wolfgang Nejdl, Wolf Siberski και Michael Sintek. Design issues and challenges for rdf- and schema-based peer-to-peer systems. *SIGMOD Rec.*, 32(3):41–46, 2003.
- [5] Έλλη Ανδρουλάκη. Υλοποίηση Ενεργού Μηχανισμού σε σύστημα Ομότιμων Βάσεων. KDBS Lab, NTU Athens, 2005.

Κεφάλαιο 8

Μεταφράσεις Ξένων όρων

Μετάφραση

αδερφός
αμεταβλητότητα
ανάκτηση πληροφορίας
αντιμεταθετικότητα
απόγονος
απορρόφηση
βάση δεδομένων
γνώρισμα
διαπροσωπεία
διαφορά
δικτυακός κατάλογος
δικτυωτή δομή
δομικές επερωτήσεις
δομικές σχέσεις
δομικό σχήμα
εγκυρότητα
ένωση

Αγγλικός όρος

sibling
idempotency
information retrieval
commutativity
descendant
absorption
database
attribute
interface
difference
portal catalog
lattice
structural queries
structural relationships
schema
validity
union

