

Διπλωματικές Εργασίες Έτους 2014-2015

Υπεύθυνος καθ. Α.Τσακαλίδης επ. καθ. Μακρής Χρήστος

Ενημέρωση για τις διπλωματικές εργασίες του Εργαστηρίου Γραφικών Πολυμέσων και Γεωγραφικών Συστημάτων θα γίνει την Πέμπτη 25/9/2014, ώρα 14.00 – 17.00 στην αίθουσα Π200.

1. Τίτλος : Αξιοποίηση των Πολιτών ως Ανθρώπινων Αισθητήρων για τη Βελτίωση Κοινωνικών Υπηρεσιών

Επίβλεψη: Τζήμας Γ., Βιέννας Ε.

Στόχος: Στην παρούσα εργασία θα διερευνηθεί η δυνατότητα βελτίωσης του επιπέδου ζωής των πολιτών αξιοποιώντας δεδομένα που είτε παράγονται από τους πολίτες (crowdsensing), είτε συλλέγονται από τα διάφορα μέσα κοινωνικής δικτύωσης και αφορούν σε μια ευρεία γκάμα δραστηριοτήτων κοινωνικού και πολιτικού περιεχομένου. Μερικά παραδείγματα θα μπορούσε να είναι: Αναγνώριση συμβάντων και κοινοποίηση τους, ενεργή συμμετοχή στις δημοκρατικές διαδικασίες διαβούλευσης, κυκλοφοριακή δραστηριότητα και μέσα μαζικής μεταφοράς, αξιολόγηση προσφερόμενων υπηρεσιών δημοσίου και ιδιωτικού τομέα κ.α..

Αντικείμενο: Στην εποχή μας, περισσότερο από ποτέ ο πολίτης, μπορεί να αποτελέσει σημαντικό μοχλό βελτίωσης της κοινωνίας στην οποία ζει και δραστηριοποιείται, αξιοποιώντας τις δυνατότητες που του δίνει η σύγχρονη τεχνολογία πληροφορικής. Η ευρεία διάδοση των υπολογιστών και των έξυπνων συσκευών μπορεί να δώσει φωνή στους πολίτες αναπτυσσόμενων και αναδυόμενων χωρών, όπως η χώρα μας, και να οδηγήσει μέσω μεθοδικής μελέτης και ανάλυσης των διαθέσιμων τεχνολογιών (ανοικτού κώδικα, δεδομένων μεγάλης κλίμακας) στην πρόταση καινοτόμων λύσεων, αναδεικνύοντας τον ίδιο τον πολίτη ως τον κύριο φορέα αλλαγών στις κοινωνίες στις οποίες ζει και δραστηριοποιείται.

Στα πλαίσια της διπλωματικής εργασίας θα διερευνηθούν έτοιμες λύσεις για την υποστήριξη των παραπάνω δραστηριοτήτων, όπως το CitySDK και θα οριστεί ένα σενάριο υλοποίησης μίας εφαρμογής, η οποία θα μπορούσε να αξιοποιεί τεχνολογίες έξυπνων συσκευών και δεδομένα παραγόμενα από τους πολίτες σε συνδυασμό με μέσα Κοινωνικής Δικτύωσης για την επίτευξη της βελτιστοποίησης υπηρεσιών κοινωνικού χαρακτήρα σε αστικά περιβάλλοντα

και όχι μόνο. Λόγω του όγκου των δεδομένων που είναι πιθανό να συγκεντρωθούν θα εφαρμοστούν τεχνικές εξόρυξης γνώσης και θα αξιοποιηθούν τεχνολογίες ανάλυσης μεγάλου όγκου δεδομένων.

Παραπομπές:

- City Service Development Kit: <http://www.citysdk.eu/>
- <http://cities.media.mit.edu/>
- <http://smartcities.ieee.org/>

Όροι αναζήτησης: CitySDK, Human Sensors, Mobile Apps, Data Mining, Big Data

2. Τίτλος : Πειραματική αξιολόγηση τεχνολογιών μεγάλου όγκου δεδομένων

Περιγραφή

Στα πλαίσια της διπλωματικής αυτής θα αξιολογηθούν πειραματικά πλατφόρμες επεξεργασίας δεδομένων μεγάλου όγκου όπως τα ακόλουθα:

- Apache Tez - <http://tez.apache.org/>
- Apache Spark - <https://spark.apache.org/>
- Apache™ Hadoop - <http://hadoop.apache.org/>
- Apache Drill - <http://incubator.apache.org/drill/>
- Apache Storm - <https://storm.incubator.apache.org/>
- Apache Samza - <http://samza.incubator.apache.org/>
- Apache Giraph - <http://giraph.apache.org/>
- Impala - <http://www.cloudera.com/content/cloudera/en/products-and-services/cdh/impala.html>
-

Οι παραπάνω πλατφόρμες θα αξιολογηθούν σε μία από τις παρακάτω θεματικές ενότητες:

1. Αλγόριθμοι εξόρυξης γνώσης: ομαδοποίηση αντικειμένων, εξαγωγή κανόνων συσχέτισης
2. Αλγόριθμοι επεξεργασίας κειμένων: εύρεση προτύπων, summarization κ.α.
3. Αλγόριθμοι επεξεργασίας γράφων
4. κ.α.

Προαπαιτούμενα : Πολύ καλή γνώση προγραμματισμού, Ανάκτηση Πληροφορίας, Εξόρυξη Γνώσης, Βάσεις δεδομένων

Επιβλέπων: Αθανάσιος Τσακαλίδης

Συνεπιβλέπων: Μανόλης Βιέννας, Ευάγγελος Θεοδωρίδης

Επικοινωνία: biennas@ceid.upatras.gr, theodori@ceid.upatras.gr

Σχετικές Αναφορές:

1. hortonworks - <http://hortonworks.com/>
2. mapr - <https://www.mapr.com/>
3. cloudera - <http://www.cloudera.com/content/cloudera/en/home.html>

3. Τίτλος : Πειραματική αξιολόγηση τεχνολογιών συστημάτων διαχείρισης βάσεων δεδομένων νέας γενιάς

Περιγραφή

Στα πλαίσια της διπλωματικής αυτής θα αξιολογηθούν πειραματικά συστήματα διαχείρισης βάσεων δεδομένων όπως τα ακόλουθα:

- Column Store
 - <https://cassandra.apache.org/>
- Document Store
 - <http://www.mongodb.org/>
 - <http://couchdb.apache.org/>
- Key Value
 - <http://redis.io/>
 - <http://www.oracle.com/technetwork/database/database-technologies/berkeleydb/overview/index.html>
- Graph
 - <http://www.neo4j.org/>
 - <http://research.microsoft.com/en-us/projects/trinity/>
 - <http://giraph.apache.org/>
- NewSQL
 - <http://www.nuodb.com/>
 - <http://voltdb.com/>

Οι παραπάνω πλατφόρμες θα αξιολογηθούν σε μία από τις παρακάτω θεματικές ενότητες:

1. εκτέλεση απλών ερωτημάτων (select/where, aggregates)
2. εκτέλεση ερωτημάτων σύνδεσης (joins)
3. εκτέλεση σύνθετων ερωτημάτων αναζήτησης λέξεων κλειδιών και geolocation
4. κ.α.

Προαπαιτούμενα : Πολύ καλή γνώση προγραμματισμού, Ανάκτηση Πληροφορίας, Εξόρυξη Γνώσης, Βάσεις δεδομένων

Επιβλέπων: Αθανάσιος Τσακαλίδης

Συνεπιβλέπων: Μανόλης Βιέννας, Ευάγγελος Θεοδωρίδης

Επικοινωνία: biennas@ceid.upatras.gr, theodori@ceid.upatras.gr

Σχετικές Αναφορές:

1. <http://en.wikipedia.org/wiki/NoSQL>
2. <http://nosql-database.org/>
3. <http://en.wikipedia.org/wiki/MySQL>

4.Τίτλος : Ανάπτυξη εφαρμογής σε φορητές συσκευές εξόρυξης ομοιότητας χρηστών

Περιγραφή

Στις μέρες μας όλο και περισσότεροι είναι οι χρήστες που έχουν λογαριασμό σε κάποιο κοινωνικό δίκτυο και αυτοί που χρησιμοποιούν φορητές συσκευές. Στα πλαίσια της διπλωματικής αυτής θα αναπτυχθεί ένα πληροφοριακό σύστημα που ως στόχο θα έχει την εύρεση ομοιότητας μεταξύ προφίλ χρηστών και το οποίο θα:

- κάνει εξόρυξη γνώσης από το διαδίκτυο μαζεύοντας πληροφορίες για το προφίλ των χρηστών
- διασυνδέεται με τα προφίλ των χρηστών στα διάφορα κοινωνικά δίκτυα μαζεύοντας πληροφορίες για αυτά
- καταγράφει την τοποθεσία του χρήστη από την φορητή συσκευή
- βρίσκει την ομοιότητα μεταξύ προφίλ χρηστών σε διάφορες κατηγορίες και συνολικά

Πιο αναλυτικά θα:

- αναπτυχτεί μία εφαρμογή για φορητές συσκευές σε κάποια από τις πλατφόρμες (android, windows phone, iphone)
- αναπτυχτούν τα απαιτούμενα web services που θα λαμβάνουν τα δεδομένα από την φορητή συσκευή και θα τα αποθηκεύουν σε βάση δεδομένων κεντρικά
- θα εξεταστούν όλες οι τεχνολογίες που βοηθάνε στο να μπορεί η συγκεκριμένη υλοποίηση να υποστηρίξει μεγάλο αριθμό χρηστών (nosql databases, graph databases, mahout)

Το πρώτο βήμα στα πλαίσια της διπλωματικής αυτής είναι να εξαχθεί η ομοιότητα μεταξύ χρηστών. Αφού γίνει αυτό, το σύνολο των δεδομένων αυτών μπορεί να χρησιμοποιηθεί για διάφορους σκοπούς όπως να προτείνει σε έναν χρήστη άλλους χρήστες με παρόμοιο προφίλ στην περιοχή που βρίσκεται κ.α.

Προαπαιτούμενα : Πολύ καλή γνώση προγραμματισμού, Ανάκτηση Πληροφορίας, Εξόρυξη Γνώσης, Τεχνολογίες Διαδικτύου, Προγραμματισμός Κινητών Τηλεφώνων

Επιβλέπων: Αθανάσιος Τσακαλίδης

Συνεπιβλέπων: Μανόλης Βιέννας

Επικοινωνία: biennas@ceid.upatras.gr

Σχετικές Αναφορές:

1. <http://www.pipl.com/>
2. Mining the Social Web - <http://shop.oreilly.com/product/0636920010203.do>
3. https://research.microsoft.com/en-us/um/redmond/groups/connect/cscw_10/docs/p41.pdf
4. <http://dl.acm.org/citation.cfm?id=1463477>
5. <http://www.cs.ucdavis.edu/research/tech-reports/2009/CSE-2009-23.pdf>
6. http://islab.kaist.ac.kr/chungcw/InterConfPapers/user_sim.pdf
7. Keyword: mining user similarity

5. Τίτλος: Ανάπτυξη συστήματος με χρήση γεωγραφικών πληροφοριών (GIS) για φορητές συσκευές Android

Περιγραφή

Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη μιας εφαρμογής με χρήση γεωγραφικών πληροφοριών που να έχει εφαρμογή στην καθημερινότητα του χρήστη καθώς και στην προσέγγιση πελατών από χώρους εστίασης.

Συνολικά θα πραγματοποιηθούν οι παρακάτω δραστηριότητες:

1. Εκμάθηση και εμβάθυνση τεχνολογίας και σχετικών εργαλείων τεχνολογικής πλατφόρμας για φορητές συσκευές και κινητά τηλέφωνα (Eclipse, Android, Android SDK)
2. Μελέτη τεχνολογίας Google Maps.
3. Εκμάθηση υπηρεσιών διαδικτύου (web services).
4. Ανάπτυξη εφαρμογής για φορητές συσκευές που να συνδυάζει όλα τα παραπάνω.

Προαπαιτούμενα : Τεχνολογίες Εφαρμογών Διαδικτύου, Πληροφοριακά Συστήματα, Καλή γνώση Προγραμματισμού (java, eclipse)

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Μάριος Κενδέα, Βασιλική Γκαντούνα, Γιάννης Τζήμας

Επικοινωνία: kendea@ceid.upatras.gr, gkantoun@ceid.upatras.gr, tzimas@cti.gr

Σχετικές Αναφορές:

1. Android Development: <http://developer.android.com/index.html>
2. Google Maps for Android:
<https://developers.google.com/maps/documentation/android/hello-mapview>
3. RESTful Web Services:
<http://www.ibm.com/developerworks/webservices/library/ws-restful/>
4. Eclipse: <http://www.eclipse.org/downloads/>

6. Τίτλος : Υλοποίηση Interpolation Search Tree (IST)σε NoSQL Συστήματα Βάσεων Δεδομένων

Στόχος: Στόχος της διπλωματικής είναι η υλοποίηση του Interpolation Search Tree σε Map-Reduce Προγραμματιστικό Περιβάλλον (Parallel and Distributed Programming Framework).

Αντικείμενο: Αντικείμενο της διπλωματικής είναι η κατανομημένη υλοποίηση του Δέντρου Εύρεσης με Παρεμβολή (IST) σε 1D και 2D δεδομένα που είναι αποθηκευμένα σε NoSQL Βάσεις Δεδομένων (HBASE) οι οποίες εγκαθίστανται στους κόμβους ενός HADOOP Cluster. Το περιβάλλον υλοποίησης είναι το Map-Reduce. Ενδεικτικά περιλαμβάνονται:

- Μελέτη του IST σε RAM, Distributed(P2P) και I/O model
- Εγκατάσταση hadoop cluster το οποίο υποστηρίζεται από το HDFS (Hadoop Distributed File System). Το τελευταίο είναι η εξέλιξη του γνωστού GFS (Google File System)
- Εγκατάσταση NoSQL ΒΔ (HBASE) σε κάθε κόμβο του cluster

Προαπαιτούμενες Γνώσεις: Κατανομημένες Δομές Δεδομένων, Βάσεις Δεδομένων, Αλγόριθμοι και Πολυπλοκότητα, Τεχνολογίες Διαδικτύου, Αντικειμενοστραφής Προγραμματισμός

Παραπομπές:

1. Welcome to Apache Hadoop <http://hadoop.apache.org/>
2. Hadoop: The Definite Guide, Tom White, O' REILLY

3. www.sciencedirect.com/science/article/pii/S0306457310001044, Map-Reduce Indexing Strategies: Studying Scalability and Efficiency
4. Kaporis, Ch. Makris, S. Sioutas, A. Tsakalidis, K. Tsichlas, Ch. Zaroliagis, "Dynamic Interpolation Search revisited", International Colloquium in Automata, Languages and Programming (ICALP 2006), Part I, LNCS 4051, pp. 382-394.
5. G.Brodal, A. Kaporis, A. Papadopoulos, S. Sioutas, K. Tsakalidis, K. Tsichlas "Dynamic 3-sided Planar Range Queries with Expected Doubly Logarithmic Time", Theor. Comput. Sci. (TCS) 526:58-74 (2014), Elsevier.
7. S. Sioutas, P. Triantafyllou, G. Papaloukopoulos, E. Sakopoulos, K. Tsichlas, Y. Manolopoulos, "ART: Sub-Logarithmic Decentralized Range Query Processing with Probabilistic Guarantees", Distributed and Parallel Databases (DAPD), 31(1):71-109 (2013), Springer 2013.
8. A. Kaporis, Ch. Makris, S. Sioutas, A. Tsakalidis, K. Tsichlas, Y. Ch. Zaroliagis, "ISB-Tree: A New Indexing Scheme with Efficient Expected Behaviour", Journal of Discrete Algorithms (8), 2010, pp. 373-387, Elsevier.

Όροι αναζήτησης: Indexing Strategies, Input Distributions, Map-Reduce, Hadoop, HBase.

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέποντες: Σ. Σιούτας, Χ. Μακρής, Γ. Τζήμας

Επικοινωνία: sioutas@ceid.upatras.gr, sioutas@ionio.gr, makri@ceid.upatras.gr, tzimas@cti.gr

7. Τίτλος: Μελέτη Επιδημιολογικού Μοντέλου SIS σε αυθαίρετο γράφημα

Σκοπός της παρούσας διπλωματικής είναι η θεωρητική και πρακτική μελέτη του επιδημιολογικού μοντέλου SIS σε ένα αυθαίρετο γράφημα. Σε αυτή την εκδοχή του προβλήματος, ένας ιός εξαπλώνεται σε ένα γράφημα, όπου οι κόμβοι του γραφήματος παρουσιάζουν διαφορετικό βαθμό ευαισθησίας απέναντι στον ιό. Στόχος, είναι η μαθηματική περιγραφή του μοντέλου και ο υπολογισμός κατωφλίων που προσδιορίζουν κρίσιμα σημεία για την εξάπλωση ή μη του ιού. Επίσης, θα υλοποιηθεί προσομοίωση του μοντέλου σε συνθετικά γραφήματα αλλά και πειράματα σε πραγματικά δεδομένα. Η διπλωματική βρίσκει εφαρμογή σε τομείς όπως:

- Βιολογία, όπου ο ιός αναφέρεται σε μία ασθένεια που εξαπλώνεται σε μία ομάδα του πληθυσμού.
- Viral Marketing, όπου ο ιός έχει τη μορφή ενός προϊόντος, μίας φήμης (rumor) ή είδησης που εξαπλώνεται σε ένα δίκτυο χρηστών.

Προαπαιτούμενα: Αλγόριθμοι, Γραμμική Άλγεβρα, Γλώσσες Προγραμματισμού

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Α. Ράπτη, Γ. Τζήμας

Επικοινωνία: arapti@ceid.upatras.gr, tzimas@cti.gr

Σχετικές Αναφορές:

1. B. Aditya Prakash, Alex Beutel, Roni Rosenfeld, and Christos Faloutsos. Winner Takes All: Competing Viruses or Ideas on fair-play Networks. WWW 2012, April 16–20, 2012.
2. B. Aditya Prakash, Deepayan Chakrabarti, Michalis Faloutsos, Nicholas Valler, Christos Faloutsos. Threshold Conditions for Arbitrary Cascade Models on Arbitrary Networks. Knowledge and Information Systems Journal, 549-575, Springer. 2012.

8. Τίτλος : Ανάπτυξη εφαρμογής οδηγού διασκέδασης για φορητές συσκευές Android για την πόλη της Πάτρας με χρήση Google Maps

Περιγραφή

Πρόκειται για τη σχεδίαση και ανάπτυξη εφαρμογής όπου ένας χρήστης θα συνδέεται στην εφαρμογή δίνοντας κάποιο συνθηματικό και κάποιον κωδικό (εναλλακτικά να συνδέεται με το λογαριασμό του στο Facebook). Στη συνέχεια, θα μπορεί μέσω Google Maps να περιηγηθεί στην περιοχή της Πάτρας για να εντοπίσει σημεία διασκέδασης που έχουν προσθέσει οι χρήστες του συστήματος. Για κάθε σημείο στον χάρτη θα έχει τη δυνατότητα να δει τη βαθμολογία που έχουν δώσει οι χρήστες καθώς και διάφορες κριτικές. Επιπλέον, θα μπορεί και ο ίδιος κάποιο μέρος διασκέδασης στο σύστημα αν το επιθυμεί. Το πληροφοριακό σύστημα θα αποτελείται από την εφαρμογή στην φορητή συσκευή, έναν web εξυπηρετητή που θα αποθηκεύονται κεντρικά όλες οι πληροφορίες και ένα σύνολο υπηρεσιών διαδικτύου για την επικοινωνία.

Συνολικά θα πραγματοποιηθούν οι παρακάτω δραστηριότητες:

- Εκμάθηση και εμβάθυνση τεχνολογίας και σχετικών εργαλείων τεχνολογικής πλατφόρμας για φορητές συσκευές και κινητά τηλέφωνα (Eclipse, Android, Android SDK)
- Μελέτη τεχνολογίας Google Maps.
- Εκμάθηση υπηρεσιών διαδικτύου (web services).
- Ανάπτυξη εφαρμογής για φορητές συσκευές που να συνδυάζει όλα τα παραπάνω.

Προαπαιτούμενα : Τεχνολογίες Εφαρμογών Διαδικτύου, Προηγμένα Πληροφοριακά Συστήματα, καλή γνώση Java

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Νικόλαος Νοδαράκης

Επικοινωνία: nodarakis@ceid.upatras.gr

Σχετικές Αναφορές:

1. Διπλωματική εργασία κ. Νοδαράκη: «Ανάπτυξη συστήματος γεωγραφικών πληροφοριών (GIS) για φορητές συσκευές σε περιβάλλον Android»
2. Android Development: <http://developer.android.com/index.html>
3. Google Maps for Android:
<https://developers.google.com/maps/documentation/android/hello-mapview>
4. RESTful Web Services:
<http://www.ibm.com/developerworks/webservices/library/ws-restful/>
Eclipse: <http://www.eclipse.org/downloads/>

9. Τίτλος: Ανάπτυξη εφαρμογής ανταλλαγής μηνυμάτων μέσω Bluetooth και με χρήση της τεχνολογίας Push Notification για φορητές συσκευές Android.

Περιγραφή

Πρόκειται για τη σχεδίαση και ανάπτυξη εφαρμογής όπου 2 χρήστες φορητών συσκευών με λειτουργικό σύστημα Android θα μπορούν να ανταλλάσσουν άμεσα μηνύματα και φωτογραφίες μέσω σύζευξης των συσκευών με την τεχνολογία Bluetooth ή με χρήση της τεχνολογίας push notification. Οι χρήστες θα συνδέονται στην εφαρμογή αφού προηγουμένως έχουν δημιουργήσει ένα προφίλ. Στη συνέχεια, θα έχουν τη δυνατότητα να αναζητήσουν συσκευές μέσω Bluetooth και να επικοινωνήσουν με αυτές μέσω μηνυμάτων και να ανταλλάσουν και εικόνες. Επίσης, δίνεται η δυνατότητα αποστολής μηνυμάτων μεταξύ των χρηστών της εφαρμογής και με χρήση της τεχνολογίας Push Notification. Η εφαρμογή θα κρατάει ιστορικό των συνομιλιών τις οποίες θα μπορεί να αναζητήσει ο χρήστης.

Συνολικά θα πραγματοποιηθούν οι παρακάτω δραστηριότητες:

- Εκμάθηση και εμβάθυνση τεχνολογίας και σχετικών εργαλείων τεχνολογικής πλατφόρμας για φορητές συσκευές και κινητά τηλέφωνα (Eclipse, Android, Android SDK)
- Μελέτη τεχνολογίας Bluetooth.
- Μελέτη της τεχνολογίας Push Notification
- Εκμάθηση υπηρεσιών διαδικτύου (web services).
- Ανάπτυξη διαδραστικής εφαρμογής για φορητές συσκευές που να συνδυάζει όλα τα παραπάνω.

Προαπαιτούμενα : Τεχνολογίες Εφαρμογών Διαδικτύου, Προηγμένα Πληροφοριακά Συστήματα, καλή γνώση Java

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Νικόλαος Νοδαράκης

Επικοινωνία: nodarakis@ceid.upatras.gr

Σχετικές Αναφορές:

1. Android Development: <http://developer.android.com/index.html>
2. Bluetooth for Android:
<http://developer.android.com/guide/topics/connectivity/bluetooth.html>

3. RESTful Web Services:
<http://www.ibm.com/developerworks/webservices/library/ws-restful/>
4. Push Technology: http://en.wikipedia.org/wiki/Push_technology
5. Push Notification for Android:
<http://www.vogella.com/tutorials/AndroidCloudToDeviceMessaging/article.html>
<https://parse.com/tutorials/android-push-notifications>
6. Eclipse: <http://www.eclipse.org/downloads/>

10. Τίτλος: Εξαγωγή Αφαιρετικού Μοντέλου από Σύστημα Διαχείρισης Περιεχομένου με χρήση Αντίστροφης Μηχανίκευσης

Περιγραφή

Η ολοένα και αυξανόμενη πολυπλοκότητα των σύγχρονων εφαρμογών Παγκόσμιου Ιστού (Web applications), έχει οδηγήσει στην υιοθέτηση Model-Driven τεχνικών κατά την διαδικασία της ανάπτυξης και συντήρησης λογισμικού. Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη ενός συστήματος που θα εξάγει αυτόματα το αφαιρετικό μοντέλο (conceptual model) Web applications τα οποία έχουν αναπτυχθεί με βάση ένα σύστημα διαχείρισης περιεχομένου (CMS). Το μοντέλο που θα ανακτηθεί θα αναπαρασταθεί σύμφωνα με την γλώσσα μοντελοποίησης (WebML-IFML). Αντικείμενο, επίσης, της παρούσας διπλωματικής εργασίας θα είναι η ανάπτυξη εφαρμογής για την γραφική αναπαράσταση του ανακτηθέντος μοντέλου σύμφωνα με τα πρότυπα της WebML-IFML.

Προαπαιτούμενα: Τεχνολογίες Εφαρμογών Διαδικτύου, Εμπειρία στην ανάπτυξη εφαρμογών διαδικτύου.

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Γ. Τζήμας, Β. Γκαντούνα

Επικοινωνία: gkantoun@ceid.upatras.gr

Σχετικές Αναφορές:

1. WebML - <http://www.webml.org/>
2. IFML - <http://www.ifml.org/>
3. ASP.NET, <http://www.asp.net/>
4. Visual Studio 2013, msdn.microsoft.com/en-us/vstudio/default.aspx

11. Τίτλος : Συστήματα, εφαρμογές και αλγόριθμοι για web data management

Στόχος της εργασίας / των εργασιών είναι η χρήση συστημάτων ανοικτού κώδικα ή / και η ανάπτυξη τεχνικών και αλγορίθμων για ανάλυση και διαχείριση των δεδομένων στο Διαδίκτυο (web data management).

- Μελέτη διαθέσιμων συστημάτων, μεθοδολογιών, αλγορίθμων
- Εγκατάσταση, δοκιμή παραμετροποίηση ή / και ανάπτυξη νέων τεχνικών.

Προαπαιτούμενα: Τεχνολογίες Διαδικτύου, C++, Java, hadoop

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Δ. Τσώλης

Επικοινωνία: dtsolis@upatras.gr

12. Τίτλος : Εφαρμογές Επαυξημένης Πραγματικότητας στην Πολιτιστική Κληρονομιά και την Εκπαίδευση

Στόχος η ανάπτυξη εφαρμογών επαυξημένης πραγματικότητας στην Πολιτιστική Κληρονομιά με την αξιοποίηση μηχανών τρισδιάστατης αναπαράστασης (π.χ. unreal engine, unity κ.α.). Έμφαση θα δοθεί στην αξιοποίηση της παραγόμενης εφαρμογής για την εκπαίδευση και δια βίου μάθηση.

Προαπαιτούμενα: 3D rendering, 3D engines, Unreal, Unity κ.α.

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Δ. Τσώλης

Επικοινωνία: dtsolis@upatras.gr

13. Τίτλος : Δομές Δεδομένων και Δεικτοδότηση Κειμένων

Σκοπός της παρούσας διπλωματικής είναι η θεωρητική και πρακτική μελέτη αλγόριθμων και δομών δεδομένων για την αποδοτική αποθήκευση συλλογών εγγράφων και την γρήγορη αναζήτηση μέσα σε αυτά. Η διπλωματική μπορεί να εστιάσει σε διάφορα προβλήματα της περιοχής αυτής όπως:

- αποδοτική αποθήκευση και αναζήτηση σε ιεραρχικές μνήμες
- αναζήτηση γνωρίζοντας τις κατανομές των ερωτημάτων
- αποδοτική αποθήκευση και αναζήτηση σε επαναλαμβανόμενα κείμενα
- συμπίεση και αναζήτηση

Προαπαιτούμενα: Δομές Δεδομένων, Αλγόριθμοι, Γλώσσες Προγραμματισμού, Ανάκτηση Πληροφορίας

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Ευάγγελος Θεοδωρίδης

Επικοινωνία: theodori@ceid.upatras.gr

Σχετικές Αναφορές:

1. Paolo Ferragina, Rossano Venturini: Compressed Cache-Oblivious String B-tree. ESA 2013: 469-480
2. Paolo Ferragina, Jouni Sirén, Rossano Venturini: Distribution-Aware Compressed Full-Text Indexes. ESA 2011: 760-771
3. Andrés Abeliuk, Rodrigo Cánovas, Gonzalo Navarro: Practical Compressed Suffix Trees. Algorithms 6(2): 319-351 (2013)
4. Travis Gagie, Kalle Karhu, Gonzalo Navarro, Simon J. Puglisi, Jouni Sirén: Document Listing on Repetitive Collections. CPM 2013: 107-119
5. Djamel Belazzougui, Gonzalo Navarro, Daniel Valenzuela: Improved compressed indexes for full-text document retrieval. J. Discrete Algorithms 18: 3-13 (2013)
6. Antonio Fariña, Nieves R. Brisaboa, Gonzalo Navarro, Francisco Claude, Ángeles S. Places, Eduardo Rodríguez: Word-based self-indexes for natural language text. ACM Trans. Inf. Syst. 30(1): 1 (2012)
7. Francisco Claude and J. Ian Munro. 0013. Document Listing on Versioned Documents. In Proceedings of the 20th International Symposium on String Processing and Information Retrieval - Volume 8214 (SPIRE 2013),

14. Τίτλος : Εύρεση Προτύπων σε πολλαπλές ροές δεδομένων

Σκοπός της παρούσας διπλωματικής είναι η θεωρητική και πρακτική μελέτη αλγόριθμων και δομών δεδομένων για την αναζήτηση προτύπων σε ροές δεδομένων. Σε αυτή την εκδοχή του προβλήματος, ένα σύμβολο φτάνει κάθε χρονική στιγμή από τις ροές δεδομένων. Το πρόβλημα είναι να απαντηθεί αν εμφανίζεται ένα μοτίβο/πρότυπο σε κάποια/κάποιες από τις ροές δεδομένων χρησιμοποιώντας τον λιγότερο δυνατό χώρο. Επίσης θα μελετηθεί το πρόβλημα προσεγγιστικής εύρεσης προτύπου με σφάλματα. Οι αλγόριθμοι θα υλοποιηθούν σε πλατφόρμες επεξεργασίας ροών δεδομένων όπως το Apache Samza, Apache Spark, Apache Storm.

Προαπαιτούμενα: Δομές Δεδομένων, Αλγόριθμοι, Γλώσσες Προγραμματισμού, Ανάκτηση Πληροφορίας

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Ευάγγελος Θεοδωρίδης

Επικοινωνία: theodori@ceid.upatras.gr

Σχετικές Αναφορές:

1. Raphaël Clifford, Markus Jalsenius, Ely Porat, Benjamin Sach: Pattern Matching in Multiple Streams. CPM 2012: 97-109
2. Benny Porat , Ely Porat, Exact and Approximate Pattern Matching in the Streaming Model, Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science, p.315-323, October 25-27, 2009
3. <http://samza.incubator.apache.org/>
4. <https://spark.apache.org/>

5. <https://storm.incubator.apache.org/>

15. Υλοποίηση Συστήματος Εντοπισμού Λογοκλοπής (Plagiarism Detection) - Τεχνικές αποτίμησης ομοιότητας κειμένων

Σκοπός της παρούσας διπλωματικής είναι η θεωρητική και πρακτική μελέτη αλγόριθμων και δομών δεδομένων για την ανάλυση συλλογών κειμένου για τον εντοπισμό σημείων λογοκλοπής. Θα μελετηθούν σύγχρονες τεχνικές δεικτοδότησης εγγράφων, φυσικής επεξεργασίας γλώσσας, εξαγωγής οντοτήτων από κείμενα κλπ με στόχο τον εντοπισμό τμημάτων που έχουν υποστεί αντιγραφή (μερική ή πλήρης). Οι τεχνικές αυτές θα υλοποιηθούν και θα δοκιμασθούν στην πράξη σε περιβάλλοντα νέφους και σε πλατφόρμες επεξεργασίας μεγάλου όγκου δεδομένων (πχ Apache Hadoop, Apache Spark).

Προαπαιτούμενα: Δομές Δεδομένων, Αλγόριθμοι, Γλώσσες Προγραμματισμού, Ανάκτηση Πληροφορίας

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Ευάγγελος Θεοδωρίδης

Επικοινωνία: theodori@ceid.upatras.gr

Σχετικές Αναφορές:

1. Efstathios Stamatatos. 2011. Plagiarism detection based on structural information. In Proceedings of the 20th ACM international conference on Information and knowledge management (CIKM '11), pages 1221-1230.
2. Fernando SáNchez-Vega, Esaú Villatoro-Tello, Manuel Montes-Y-GóMez, Luis VillaseñOr-Pineda, and Paolo Rosso. 2013. Determining and characterizing the reused text for plagiarism detection. Expert Syst. Appl. 40, 5 (April 2013)
3. Salha Alzahrani, Vasile Palade, Naomie Salim, and Ajith Abraham. 2012. Using structural information and citation evidence to detect significant plagiarism cases in scientific publications. J. Am. Soc. Inf. Sci. Technol. 63, 2 (February 2012), 286-312.
4. <https://uima.apache.org/>
5. <http://lucene.apache.org/solr/>
6. <http://lucene.apache.org/>

16. Τίτλος : Εξόρυξη Ιατρικών/Βιολογικών Οντοτήτων από συλλογές Κειμένων

Περιγραφή

Στην παρούσα διπλωματική θα μελετηθούν και αξιολογηθούν πειραματικά αλγόριθμοι εξόρυξης ιατρικών και βιολογικών οντοτήτων και σχέσεων μεταξύ των οντοτήτων (πχ. πρωτεΐνες και πρωτεϊνικές αλληλεπιδράσεις ή γονιδίων, φαρμάκων και αλληλεπιδράσεων τους) από συλλογές βιολογικών κειμένων που βρίσκονται στο διαδίκτυο. Ποιο

συγκεκριμένα θα εστιάσει στην εφαρμογή τεχνικών εξόρυξης στο PUBMED, την μεγαλύτερη online βάση δεδομένων ιατρικών και βιολογικών επιστημονικών εργασιών.

Προαπαιτούμενα: Βιοπληροφορική, Ανάκτηση Πληροφορίας, Εξόρυξη Γνώσης, Αλγόριθμοι, Γλώσσες Προγραμματισμού

Επιβλέπων: Α.Τσακαλίδης

Συνεπιβλέπων: Θεοδωρίδης

Επικοινωνία: theodori@ceid.upatras.gr

Σχετικές Αναφορές:

1. Jorg Hakenberg, Robert Leaman, Nguyen Ha Vo, Siddhartha Jonnalagadda, Ryan Sullivan, Christopher Miller, Luis Tari, Chitta Baral, and Graciela Gonzalez. 2010. Efficient Extraction of Protein-Protein Interactions from Full-Text Articles. IEEE/ACM Trans. Comput. Biol. Bioinformatics 7, 3 (July 2010), 481-494.
2. Nate Sutton , Laura Wojtulewicz , Neel Mehta , Graciela Gonzalez, Automatic approaches for gene-drug interaction extraction from biomedical text: corpus and comparative evaluation, Proceedings of the 2012 Workshop on Biomedical Natural Language Processing, June 08-08, 2012, Montreal, Canada

17. Τίτλος : Αλγόριθμοι Ανάλυσης Βιο-ιατρικών Δεδομένων σε πλατφόρμες επεξεργασίας δεδομένων μεγάλου όγκου

Περιγραφή

Σκοπός της παρούσας διπλωματικής είναι η θεωρητική και πρακτική μελέτη αλγόριθμων και δομών δεδομένων για την αποδοτική ανάλυση μεγάλου πλήθους ακολουθιών DNA στην κύρια μνήμη και στα αποθηκευτικά μέσα υπολογιστικών συστημάτων. Θα εξεταστούν τεχνολογίες επεξεργασίας μεγάλου όγκου όπως το Hadoop, Apache Spark, SeqPig, SparkSeq και άλλες. Θα γίνει εστίαση είτε σε προβλήματα αντιστοίχισης ακολουθιών (Longest Common Prefix/Substring), σε αλγόριθμους ομαδοποίησης και κατηγοριοποίησης ακολουθιών (sequence clustering) ή σε αλγόριθμους επεξεργασίας μεγάλων δικτύων (protein-protein interaction networks).

Προαπαιτούμενα: Βιοπληροφορική, Ανάκτηση Πληροφορίας, Εξόρυξη Γνώσης, Αλγόριθμοι, Γλώσσες Προγραμματισμού

Επιβλέπων: Τσακαλίδης

Συνεπιβλέπων: Θεοδωρίδης

Επικοινωνία: theodori@ceid.upatras.gr

Σχετικές Αναφορές:

1. Markus J. Bauer, Anthony J. Cox, Giovanna Rosone, and Marinella Sciortino. 2012. Lightweight LCP construction for next-generation sequencing datasets. In Proceedings of the 12th international conference on Algorithms in Bioinformatics
2. Kamil Salikhov, Gustavo Sacomoto, Gregory Kucherov: Using Cascading Bloom Filters to Improve the Memory Usage for de Bruijn Graphs. 364-376 WABI 2013
3. http://www.informatik.hu-berlin.de/forschung/gebiete/wbi/research/publications/2012/dmc_in_ngs.pdf

18. Τίτλος : Χωρικές - Κειμενικές ενώσεις ομοιότητας και Spatio-Textual κατάλογοι

Οδηγούμενες από σύγχρονες εφαρμογές όπως τα κοινωνικά δίκτυα, τα κινητά τηλέφωνα κλπ οι βάσεις δεδομένων καλούνται να αποθηκεύσουν αντικείμενα που αποτελούνται τόσο από χωρικά δεδομένα (GPS στίγμα, διαδρομή κλπ) όσο και από στοιχεία κειμένου. Μελέτη της διπλωματικής αυτής είναι δομές δεδομένων και αλγόριθμοι για την αποδοτική αναζήτηση τέτοιων αντικειμένων συνδυάζοντας τόσο χωρικές δομές όσο και αλγόριθμους αναζήτησης σε κείμενα ή αλγόριθμους ενώσεων συνόλων.

Προαπαιτούμενα: Δομές Δεδομένων, Αλγόριθμοι, Γλώσσες Προγραμματισμού, Ανάκτηση Πληροφορίας

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Ευάγγελος Θεοδωρίδης

Επικοινωνία: theodori@ceid.upatras.gr

Σχετικές Αναφορές:

1. Panagiotis Bouros, Shen Ge, and Nikos Mamoulis. 2012. Spatio-textual similarity joins. Proc. VLDB Endow. 6, 1 (November 2012), 1-12.
2. Xin Cao, Gao Cong, Christian S. Jensen, and Beng Chin Ooi. 2011. Collective spatial keyword querying. In Proceedings of the 2011 ACM SIGMOD International Conference on Management of data (SIGMOD '11). ACM, New York, NY, USA, 373-384.
3. Lisi Chen, Gao Cong, Christian S. Jensen, and Dingming Wu. 2013. Spatial keyword query processing: an experimental evaluation. In Proceedings of the 39th international conference on Very Large Data Bases (PVLDB'13),

19. Τίτλος : Σχεδιασμός και ανάπτυξη εφαρμογής για συλλογή δεδομένων από αισθητήρες ενσωματωμένους σε smartphone σε εθελοντική βάση.

Περιγραφή

Σκοπός της διπλωματικής είναι η ανάπτυξη μιας εφαρμογής για Android smartphone μέσω της οποίας θα συλλέγονται δεδομένα από αισθητήρες που είναι ενσωματωμένοι στα κινητά αυτά. Τα κινητά θα προέρχονται από μια δυνητική βάση εθελοντών χρηστών, οι οποίοι θα είναι διατεθειμένοι να καταγράψουν δεδομένα και να τα διαθέσουν προς δημόσια χρήση. Για παράδειγμα, θα μπορούν να συλλέγονται πληροφορίες για τα επίπεδα θορύβου στο κέντρο της Πάτρας μέσω του μικροφώνου των κινητών, ή της ποιότητας του οδοστρώματος στους δρόμους της πόλης μέσω των επιταχυνσιόμετρών τους, κ.α. Θα αξιοποιηθεί σχετική προεργασία πάνω σε παρόμοια εφαρμογή.

Συνολικά θα πραγματοποιηθούν οι παρακάτω δραστηριότητες:

- Εκμάθηση και εμβάθυνση τεχνολογιών που σχετίζονται με την ανάπτυξη εφαρμογών σε Android και Web Services.
- Μελέτη τεχνολογικής πλατφόρμας για φορητές συσκευές και κινητά τηλέφωνα καθώς και υπάρχουσας αντίστοιχης εφαρμογής.
- Ανάπτυξη εφαρμογής για φορητές συσκευές και διενέργεια πειραματικής αξιολόγησης στους δρόμους της Πάτρας.

Προαπαιτούμενα : Τεχνολογίες Εφαρμογών Διαδικτύου, Εμπειρία στην ανάπτυξη εφαρμογών διαδικτύου, προγραμματισμός κινητών τηλεφώνων

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Ε. Θεοδωρίδης, Γ.Μυλωνάς

Επικοινωνία: {theodori, mylonasg}@ceid.upatras.gr

Σχετικές Αναφορές:

1. <http://ambientdynamix.org/>
2. Georgios Chatzimilioudis, Andreas Konstantinidis, Christos Laoudias, Demetrios Zeinalipour-Yazti: Crowdsourcing with Smartphones. IEEE Internet Computing 16(5): 36-44 (2012)

20. Τίτλος : Σχεδιασμός και ανάπτυξη διαδραστικής εγκατάστασης σε δημόσιο χώρο με τη χρήση αισθητήρων, smartphone και υπηρεσιών κοινωνικής δικτύωσης.

Περιγραφή

Στόχος είναι ο σχεδιασμός και η ανάπτυξη μιας διαδραστικής εγκατάστασης (interactive installation) στο χώρο του Τμήματος, η οποία θα αξιοποιεί ένα συνδυασμό από τεχνολογίες όπως αισθητήρες Arduino, smartphone (Android, iOS ή Windows Phone) καθώς και συστήματα κοινωνικής δικτύωσης (π.χ., Facebook). Η εγκατάσταση θα λειτουργήσει για κάποιο χρονικό διάστημα ή σε μόνιμη βάση, ως παρέμβαση στο χώρο του Τμήματος με εικαστικές, ηχητικές και κοινωνικές αποχρώσεις. Πιθανή κατεύθυνση υλοποίησης η δημιουργία μίας έξυπνης-διαδραστικής βιβλιοθήκης.

Συνολικά θα πραγματοποιηθούν οι παρακάτω δραστηριότητες:

- Εκμάθηση και εμπάθυνση τεχνολογιών που σχετίζονται με την ανάπτυξη εφαρμογών σε smartphones καθώς και συναφών δικτυακών τεχνολογιών.
- Εκμάθηση και εμπάθυνση τεχνολογιών που σχετίζονται με την ανάπτυξη εφαρμογών με τη χρήση Arduino.
- Ανάπτυξη διαδραστικής εγκατάστασης και διενέργεια πειραματικής αξιολόγησης.

Προαπαιτούμενα : Τεχνολογίες Εφαρμογών Διαδικτύου, Εμπειρία στην ανάπτυξη εφαρμογών διαδικτύου, ψηφιακά/βασικά ηλεκτρονικά, ενδιαφέρον για εικαστικά/μουσική

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Ε. Θεοδωρίδης, Γ.Μυλωνάς

Επικοινωνία: {theodori, mylonasg}@ceid.upatras.gr

Σχετικές Αναφορές:

1. Kafebook: <http://ru1.cti.gr/aigaion/?page=publication&kind=single&ID=892>
2. <http://www.mediamatic.nl/216/en/social-media-installations>

21. Τίτλος : Σχεδιασμός και ανάπτυξη εφαρμογής για την υποστήριξη αθλητικών διοργανώσεων και outdoor δραστηριοτήτων με χρήση smartphone και αισθητήρων

Περιγραφή

Στόχος είναι ο σχεδιασμός και η ανάπτυξη πλατφόρμας που θα συνδυάζει τη χρήση smartphone και συναφών τεχνολογιών για την ομαλή διεξαγωγή και λεπτομερέστερη παρακολούθηση αθλητικών δραστηριοτήτων, με εστίαση σε αγώνες τρεξίματος, αντοχής, σε βουνό, κτλ., στους οποίους είναι χρήσιμο να υπάρχει σαφής εικόνα για την πορεία των

συμμετεχόντων κατά τη διάρκεια ενός τέτοιου αγώνα. Το σύστημα θα έχει και μία δικτυακή διεπαφή χρήστη για πληρέστερη κάλυψη ενός τέτοιου γεγονότος.

Συνολικά θα πραγματοποιηθούν οι παρακάτω δραστηριότητες:

- Εκμάθηση και εμπάθунση τεχνολογιών που σχετίζονται με την ανάπτυξη εφαρμογών σε Android/iOS και Web Services, Web sites.
- Μελέτη τεχνολογικής πλατφόρμας για φορητές συσκευές και κινητά τηλέφωνα, NFC, RFID, QRtags.
- Ανάπτυξη εφαρμογής για φορητές συσκευές και διενέργεια πειραματικής αξιολόγησης στο πλαίσιο κάποιου σχετικού αγώνα.

Προαπαιτούμενα : Τεχνολογίες Εφαρμογών Διαδικτύου, Εμπειρία στην ανάπτυξη εφαρμογών διαδικτύου, προγραμματισμός κινητών τηλεφώνων

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Ε. Θεοδωρίδης, Γ.Μυλωνάς

Επικοινωνία: {theodori, mylonasg}@ceid.upatras.gr

Σχετικές Αναφορές:

1. <http://www.utmb.livetrail.net/parcours.php?course=ccc>
2. <http://www.utmb.livetrail.net/>
3. <http://802timing.com/>
4. <http://itzabouttime.com/>

22. Τίτλος : Σχεδιασμός και ανάπτυξη εφαρμογής για την εξοικονόμηση ενέργειας σε κτήρια

Περιγραφή

Στόχος είναι ο σχεδιασμός και η ανάπτυξη πλατφόρμας που θα συνδυάζει τη χρήση smartphone και συναφών τεχνολογιών για την εξοικονόμηση ενέργειας σε κτήρια (είτε δημόσια όπως πχ. Σχολεία, Πανεπιστήμια κλπ ή σε κατοικίες). Η πλατφόρμα καθώς και οι εφαρμογές που θα χρησιμοποιούν οι χρήστες θα εκτελούν τις ακόλουθες λειτουργίες:

- Θα βοηθούν στην αποτίμηση της ενεργειακής κατανάλωσης. Π.χ. με profiling της κατανάλωσης των ηλεκτρικών συσκευών, με χρησιμοποίηση κατάλληλων έξυπνων μετρητών ή με εθελοντική εισαγωγή των στοιχείων.
- Μέσω απλών και κατανοητών προτροπών θα συμβουλεύουν τους χρήστες να εκτελούν ενέργειες που θα οδηγούν σε εξοικονόμηση ενέργειας.

Συνολικά θα πραγματοποιηθούν οι παρακάτω δραστηριότητες:

- Εκμάθηση και εμπάθунση τεχνολογιών που σχετίζονται με την ανάπτυξη εφαρμογών σε Android/iOS και Web Services, Web sites και διασύνδεση έξυπνων μετρητών

- Μελέτη τεχνολογικής πλατφόρμας για φορητές συσκευές και κινητά τηλέφωνα, NFC, RFID, QRtags.
- Ανάπτυξη εφαρμογής για φορητές συσκευές και διενέργεια πειραματικής αξιολόγησης

Προαπαιτούμενα : Τεχνολογίες Εφαρμογών Διαδικτύου, Εμπειρία στην ανάπτυξη εφαρμογών διαδικτύου, προγραμματισμός κινητών τηλεφώνων

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Ε. Θεοδωρίδης, Γ.Μυλωνάς

Επικοινωνία: {theodori, mylonasg}@ceid.upatras.gr

Σχετικές Αναφορές:

1. M. Weiss, T. Staake, F. Mattern, and E. Fleisch. PowerPedia: changing energy usage with the help of a community-based smartphone application. Personal and Ubiquitous Computing, 16(6):655–664, 2012
2. <http://intelen.com/>
3. Opower social portal - Save energy with your friends, <http://social.opower.com/>.
4. T. Erickson, M. Li, Y. Kim, A. Deshpande, S. Sahu, T. Chao, P. Sukaviriya, and M. Naphade. The Dubuque electricity portal: Evaluation of a city-scale residential electricity consumption feedback system. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13, pages 1203–1212, New York, NY, USA, 2013. ACM
5. B. Faltings, J. J. Li, and R. Jurca. Incentive Mechanisms for Community Sensing. IEEE Transaction on Computers, 63(1):115–128, January 2014.,

23. Τίτλος : Μελέτη αποδοτικών αλγορίθμων για διαχείριση συμβολοσειρών με χρήση τεχνικών συμπίεσης ανεστραμμένων αρχείων

Περιγραφή

Αντικείμενο της διπλωματικής εργασίας είναι η μελέτη αποδοτικών δομών δεδομένων για διαχείριση ανεστραμμένων αρχείων σε ένα δυναμικό περιβάλλον. Μελετώνται βασικές τεχνικές και θα γίνει πειραματική αξιολόγησή τους. Με τις τεχνικές αυτές χρησιμοποιούνται επαναληπτικά υπο-συμβολοσειρές που κατασκευάζουν τις τελικές συμβολοσειρές.

Προαπαιτούμενα: Ανεστραμμένα Αρχεία, Γνώση Προγραμματισμού (C++, Java, κ.α.), Δομές Δεδομένων, Εξόρυξη Γνώσης, n-grams

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Α. Καναβός

Επικοινωνία: kanavos@ceid.upatras.gr

Σχετικές Αναφορές:

1. Kim, M. S., Whang, K. Y., Lee, J. G. and Lee M. J. n-Gram/2L: A Space and Time Efficient Two-Level n-Gram Inverted Index Structure
2. Iliopoulos, C. S., Makris, Ch., Panagis. Y., Perdikuri K., Theodoridis, E. and Tsakalidis, A. The Weighted Suffix Tree: An Efficient Data Structure for Handling Molecular Weighted Sequences and its Applications

24. Τίτλος : Εξόρυξη “σημαντικών” υπο-δικτύων από κοινωνικά δίκτυα

Περιγραφή

Στη διπλωματική αυτή θα ερευνηθούν μέθοδοι εξόρυξης γνώσης από κοινωνικά δίκτυα (linkedin, twitter, facebook). Το πρόβλημα που θα εστιάσουμε είναι η εξόρυξη ενός αριθμού χρηστών, που σχηματίζουν μεταξύ τους υπο-δίκτυο, και που να έχουν κάποια συγκεκριμένα γνωρίσματα ώστε το τελικό υπο-δίκτυο να θεωρείται efficient για μεταφορά κάποιου μηνύματος σε όλο το εύρος του συνολικού δικτύου.

Προαπαιτούμενα: Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Α. Καναβός

Επικοινωνία: kanavos@ceid.upatras.gr

Σχετικές Αναφορές:

1. Blondel V. D., Guillaume J.-L., Lambiotte R. and Lefebvre E. Fast unfolding of community hierarchies in large networks.

25. Τίτλος : Εξόρυξη πρωτεϊνικών αλληλεπιδράσεων με χρήση οντολογιών στο Παγκόσμιο Ιστό

Περιγραφή

Στη διπλωματική αυτή θα μελετηθούν αλγόριθμοι εξόρυξης βιολογικών δεδομένων (πρωτεϊνικών αλληλεπιδράσεων) από συλλογές δεδομένων του Παγκόσμιου Ιστού. Ειδικότερα χρησιμοποιούνται κάποιες οντολογίες (Gene Ontology και Molecular Interaction Ontology) στο BioCreative.

Προαπαιτούμενα: Ανάκτηση Πληροφορίας, Βιοπληροφορική, Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Α. Καναβός

Επικοινωνία: kanavos@ceid.upatras.gr

Σχετικές Αναφορές:

1. <http://www.readcube.com/articles/10.1093/database/bas017?locale=en>
2. <http://www.biocreative.org/>

26. Τίτλος : Εξόρυξη χαρακτηριστικών από κοινωνικά δίκτυα

Περιγραφή

Στη διπλωματική αυτή θα ερευνηθούν μέθοδοι εξόρυξης χαρακτηριστικών γνωρισμάτων από κοινωνικά δίκτυα (linkedin, twitter, facebook). Το πρόβλημα που θα εστιάσουμε είναι η εξόρυξη χαρακτηριστικών, όπως το φύλο, η γλώσσα, η τοποθεσία, με χρήση τεχνικών machine learning.

Προαπαιτούμενα: Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Α. Καναβός

Επικοινωνία: kanavos@ceid.upatras.gr

Σχετικές Αναφορές:

1. Sloan L., Morgan J., Housley W., Williams M., Edwards A., Burnap P. and Rana O. Knowing The Twitters - Deriving Sociologically Relevant Demographics From Twitter
2. Friedhorsky R., Culotta A. and Del Valle L. Inferring the Origin Locations of Tweets with Quantitative Confidence

27. Τίτλος : Μελέτη τεχνικών και μεθόδων ανάλυσης συναισθήματος και εφαρμογή τους σε κοινωνικά δίκτυα

Περιγραφή

Στη βιβλιογραφία γνωρίζουμε ότι για ένα κείμενο, μπορούμε να αποφασίσουμε εάν έχει θετική ή αρνητική χροιά, ή αν είναι ουδέτερο. Στη διπλωματική αυτή λοιπόν ερευνώνται μέθοδοι ανάλυσης συναισθήματος και ακολούθως γίνεται εφαρμογή τους σε διαθέσιμα κοινωνικά δίκτυα, όπως το twitter.

Προαπαιτούμενα: Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Α. Καναβός

Επικοινωνία: kanavos@ceid.upatras.gr

Σχετικές Αναφορές:

1. Liu B. Sentiment Analysis and Opinion Mining
2. <http://www.fastcoexist.com/3029598/using-citizens-as-human-sensors-to-improve-public-services>

28. Τίτλος: Ανάπτυξη εφαρμογής γρήγορης ανακάλυψης υπηρεσιών παγκόσμιου ιστού μέσω υπολογισμού διατεμνουσών υπεργραφημάτων

Περιγραφή

Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη μιας εφαρμογής για την αυτόματη ανακάλυψη συλλογών από υπηρεσίες του παγκόσμιου ιστού (web services) που ικανοποιούν το σύνολο των ενδιαφερόντων ενός χρήστη, σύμφωνα με τους προκαθορισμένους περιορισμούς κόστους. Το πρόβλημα θα μοντελοποιηθεί ως ένα πρόβλημα εύρεσης ελαχιστοτικών διατεμνουσών ενός υπεργραφήματος (hypergraph transversals), θα υλοποιηθεί ένας αποτελεσματικός αλγόριθμος για την επίλυση του, ο οποίος και θα αποτιμηθεί σε ένα σύνολο πειραματικών δεδομένων.

Προαπαιτούμενα: Διακριτά Μαθηματικά, Θεωρία Γραφημάτων, Τεχνολογίες Εφαρμογών Διαδικτύου, Πολύ καλή γνώση Προγραμματισμού (java, eclipse, matlab)

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Ηλίας Σταυρόπουλος

Επικοινωνία: estavrop@ceid.upatras.gr

Σχετικές Αναφορές:

1. Z. Zhu and J. Bailey. Fast Discovery of Interesting Collections of Web Services. In Proc. of 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI'06).
2. G. Antoniou and F. van Harmelen. A semantic Web Premier, 2nd Edition, MIT Press, 2008.
3. C. Berge. Hypergraphs, North-Holland Mathematical Library, Vol. 45, Elsevier Science Publishers B.V., Amsterdam, 1989.

29. Τίτλος: Ανάπτυξη εφαρμογής δυναμικής ανακάλυψης ηλεκτρονικών υπηρεσιών μέσω εύρεσης βέλτιστων καλυμμάτων ενός υπεργραφήματος

Περιγραφή

Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη μιας εφαρμογής για την δυναμική ανακάλυψη συλλογών από ηλεκτρονικές υπηρεσίες (e-services) που ικανοποιούν το σύνολο των ενδιαφερόντων ενός χρήστη, σύμφωνα με τους προκαθορισμένους

περιορισμούς κόστους. Το πρόβλημα θα περιγραφεί στα πλαίσια Περιγραφικής Λογικής (Description Logics), και θα μοντελοποιηθεί ως ένα πρόβλημα εύρεσης βέλτιστων καλυμμάτων (best covering problem). Για την επίλυση του θα υλοποιηθεί ένας αποτελεσματικός αλγόριθμος που θα βασίζεται στην έννοια του υπεργραφήματος (hypergraph), ο οποίος και θα αποτιμηθεί σε ένα σύνολο πειραματικών δεδομένων.

Προαπαιτούμενα: Διακριτά Μαθηματικά, Θεωρία Γραφημάτων, Τεχνολογίες Εφαρμογών Διαδικτύου, Πολύ καλή γνώση Προγραμματισμού (java, eclipse, matlab)

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέπων: Ηλίας Σταυρόπουλος

Επικοινωνία: estavrop@ceid.upatras.gr

Σχετικές Αναφορές:

1. M.S. Hacid, A. Léger, C. Rey, and F. Toumani. Dynamic Discovery of E-services: A Description Logics Based Approach. Report, LIMOS, Clermont-Ferrand, France, 2002.
2. B. Benatallah, M.S. Hacid, A. Léger, C. Rey, and F. Toumani. On Automating Web Services Discovery, VLDB Journal 14: 84 – 96, 2005.
3. G. Antoniou and F. van Harmelen. A semantic Web Premier, 2nd Edition, MIT Press, 2008.
4. C. Berge. Hypergraphs, North-Holland Mathematical Library, Vol. 45, Elsevier Science Publishers B.V., Amsterdam, 1989.

30. Τίτλος : Υλοποίηση Fully Persistent B- Tree στο I/O μοντέλο (Δίσκου-Μνήμης)

Στόχος: Στόχος της διπλωματικής εργασίας είναι η υλοποίηση του Fully Persistent B-tree στο μοντέλο δίσκου-μνήμης (I/O model).

Αντικείμενο: Αντικείμενο της διπλωματικής εργασίας είναι η κεντροποιημένη (centralized) υλοποίηση του Fully Persistent B-tree για μονοδιάστατα δεδομένα που είναι αποθηκευμένα στο Δίσκο. Η υλοποίηση θα γίνει κάνοντας χρήση αρχείων (προγραμματισμός σε I/O) σε C, C++ ή Visual C++. Συγκεκριμένα θα δημιουργηθεί ένα αρχείο ακέραιων αριθμών με μέγεθος μεγαλύτερο από αυτό της κύριας μνήμης του υπολογιστή μας. Αρχικά θα υλοποιηθεί ένα B+ δέντρο αναζήτησης/ εισαγωγής/ διαγραφής/ τροποποίησης δεδομένων του παραπάνω αρχείου. Ο βαθμός διακλάδωσης του δέντρου (fun-out or branching factor) θα είναι ανάλογος του Page Size του υπολογιστή μας. π.χ. για Page size 4Kbytes το maximum fun-out είναι 100, δεδομένου ότι θέλουμε 40 bytes / data entry. ΠΡΟΣΟΧΗ!!! Κάνουμε την υπόθεση ότι το δέντρο δεν χωράει στην κύρια μνήμη (η υλοποίηση του B+ tree στη main memory είναι διαθέσιμη στο MMLAB). Επομένως, δυνατότητα για δυναμική δέσμευση μνήμης με εντολές malloc και δεικτοδότηση μέσω pointers σε node-structs, ΔΕΝ ΥΠΑΡΧΕΙ!!!! Τα key-value structs όλων των κόμβων πρέπει να «ΞΕΔΙΠΛΩΘΟΥΝ» σε αρχεία (FILES). Κάθε αρχείο «σπάει» σε disk-blocks και κάθε disk-block έχει δικό του ID. Με κατάλληλες εντολές fread, fwrite και fseek μπορούμε να διαβάσουμε, να γράψουμε και να μετακινηθούμε σε block με συγκεκριμένο ID. Το B+ δέντρο θα πρέπει να μετατραπεί σε FULLY PERSISTENT κάνοντας χρήση των τεχνικών της παραπομπής 1. Για τη σωστή χρήση των σχετικών I/O εντολών διαχείρισης αρχείων, δίνεται σχετικός ΟΔΗΓΟΣ ΧΡΗΣΗΣ (δείτε το ΠΑΡΑΡΤΗΜΑ που ακολουθεί).

Προαπαιτούμενες Γνώσεις: Δομές Δεδομένων, Βάσεις Δεδομένων, Αλγόριθμοι και Πολυπλοκότητα,

Αντικειμενοστραφής Προγραμματισμός.

Παραπομπές:

1. [Gerth Stølting Brodal, Konstantinos Tsakalidis, Spyros Sioutas, Kostas Tsichlas](#): Fully persistent B-trees. [SODA 2012:602-614](#)

Όροι αναζήτησης: Indexing, I/O model, Files, B-trees, B+ trees, Persistent Data Structures.

Επιβλέπων: Α. Τσακαλίδης

Συνεπιβλέποντες: Σ. Σιούτας, Κ. Τσίχλας, Κ. Τσακαλίδης

Επικοινωνία: sioutas@ceid.upatras.gr, sioutas@ionio.gr, tsichlas@csd.auth.gr, tsakalid@cs.au.dk

31. Τίτλος: Μελέτη γλωσσικών μοντέλων (language models) και τεχνικών μηχανικής μάθησης για εξάλειψη πλεονασμού στην ανάκτηση πληροφορίας και πειραματική αξιολόγηση τους με χρήση του συνόλου δεδομένων Clueweb12.

Υπεύθυνος επ. καθ.: Μακρής Χρήστος

Περιγραφή

Η χρήση Γλωσσικών Μοντέλων στην Ανάκτηση Πληροφορίας αρχίζει να επεκτείνεται τα τελευταία χρόνια καθώς η απαίτηση για καλύτερη κατανόηση της συμπεριφοράς των χρηστών συστημάτων ανάκτησης πληροφορίας και για ποιοτικότερη ανακτώμενη πληροφορία αρχίζει να γίνεται ολοένα και πιο έντονη. Παράλληλα αρχίζει να γίνεται κοινός τόπος η χρήση τεχνικών μηχανικής μάθησης για ποιοτικότερη (χωρίς πλεονασμούς) παρουσίαση πληροφορίας στον χρήστη που την αναζητά. Αντικείμενο της συγκεκριμένη διπλωματικής εργασίας είναι η μελέτη και πειραματική επιβεβαίωση των πιο πρόσφατων σχετικών τεχνικών στο ClueWeb12 σύνολο δεδομένων, και η σύγκριση τους με κλασσικά μοντέλα. Θα επιχειρηθεί να γίνει καταγραφή των βασικών χαρακτηριστικών των σχετικών αλγορίθμων, και ίσως προσπάθεια επέκτασής τους.

Προαπαιτούμενα:

Δομές Δεδομένων, Ανάκτηση Πληροφορίας, Αλγόριθμοι, Τεχνολογίας Υλοποίησης , Αλγορίθμων, C++, Java, C#

Αναφορές

- R. Baeza-Yates, B. Ribeiro-Neto, Modern Information Retrieval: the concepts and technology behind search, second edition, ACM Press, 2011.
- Zhai, C.X., Statistical Language Models for Information Retrieval, A critical Review, Foundations and Trends in Information Retrieval, 2(3), 2008, 137-213.

32. Τίτλος: Μελέτη αποδοτικών αλγορίθμων για δυναμική διαχείριση ανεστραμμένων αρχείων.

Υπεύθυνος επ. καθ.: Μακρής Χρήστος

Περιγραφή

Αντικείμενο της συγκεκριμένης διπλωματικής εργασίας είναι η μελέτη αποδοτικών δομών δεδομένων για την διαχείριση ανεστραμμένων αρχείων σε ένα δυναμικό περιβάλλον όπου μπορεί να γίνει ένθεση και διαγραφή αρχείων. Θα μελετηθούν θεωρητικά οι βασικότερες τεχνικές και θα γίνει πειραματική επιβεβαίωση της συμπεριφοράς τους. Το θέμα είναι αρκετά σημαντικό και με ερευνητική διάσταση καθώς οι τεχνικές που προτείνονται είναι λίγες σε πλήθος και βασίζονται κυρίως σε προσεκτικά δρομολογημένα επαναχτισίματα της δομής δεικτοδότησης.

Προαπαιτούμενα:

Δομές Δεδομένων, Ανάκτηση Πληροφορίας, Αλγόριθμοι, Τεχνολογίας Υλοποίησης, Αλγορίθμων, C++, Java, C#

Αναφορές

- Stefan Buttcher and Charles L. A. Clarke. A Hybrid Approach to Index Maintenance in Dynamic Text Retrieval Systems. In Proceedings of the 28th European Conference on Information Retrieval, pages 229{240, London, UK, April 2006.
- Stefan Buttcher and Charles L. A. Clarke. Indexing Time vs. Query Time Trade-offs in Dynamic Information Retrieval Systems. University of Waterloo Technical Report CS-2005-31, October 2005.

33. Τίτλος: Αποδοτικά B-δένδρα/Λεξικά για μεταβλητού μήκους κλειδιά και τεχνικές βελτιστοποίησης τους.

Περιγραφή:

Οι περισσότερες παραλλαγές των B- δένδρων λειτουργούν για ένα σύνολο κλειδιών σταθερού μήκους. Στη παρούσα διπλωματική μελετώνται (θεωρητικά και πειραματικά) παραλλαγές των B-δένδρων για μεταβλητού μήκους κλειδιά. Οι δομές αυτές εκπίπτουν σε λεξικά που αντιστοιχούν τα αρχικά κλειδιά σε ατομικά αντικείμενα επιλέγοντας με κατάλληλο τρόπο τα στοιχεία διαχωρισμού έτσι ώστε να είναι δυνατή η αναζήτηση στο B-δένδρο. Θα μελετηθούν τεχνικές που έχουν ήδη προταθεί και θα γίνει προσπάθεια βελτίωσης τους και στη συνέχεια υλοποίησης τους. Η βασική δομή δεδομένων πάνω στην οποία θα κινηθούμε είναι η δομή που περιγράφεται στο paper: M. Bender, H. Hu, B. Kuszmaul, Performance guarantees for B-trees with different-sized atomic keys, PODS 2010, 305-316 (κοιτάξτε ενδεικτική βιβλιογραφία).

Σχετικά Μαθήματα: Δομές Δεδομένων, Αλγόριθμοι, Οντοκεντρικός Προγραμματισμός

Επιβλέπων: Μακρής **Συνεπιβλέπων:** Θεοδωρίδης

Επικοινωνία: theodori @ceid.upatras.gr, makri@ceid.upatras.gr

Ενδεικτική. Βιβλιογραφία: <http://dl.acm.org/citation.cfm?doid=1807085.1807125>

34. Τίτλος: Μελέτη αποδοτικών αλγορίθμων για διαχείριση δεδομένων συμβολοσειρών στη δευτερεύουσα μνήμη, και πειραματική επιβεβαίωσή τους.

Υπεύθυνος επ. καθ.: Μακρής Χρήστος

Περιγραφή

Αντικείμενο της συγκεκριμένης διπλωματικής εργασίας είναι η ανάπτυξη καινοτόμων αλγορίθμων διαχείρισης συμβολοσειρών, στη κυρία και στη δευτερεύουσα μνήμη, για την επίλυση κυρίως προβλημάτων στο χώρο της Βιοπληροφορικής.

Ο έλεγχος των αλγορίθμων που θα αναπτυχθούν για την διαχείριση βιολογικών ακολουθιών θα γίνει με τη χρήση δεδομένων από βιολογικές βάσεις δεδομένων όπως η Genbank (<http://en.wikipedia.org/wiki/GeneBank>) που περιέχει δεδομένα βιολογικών ακολουθιών σε όρους DNA και η Swissprot (<http://en.wikipedia.org/wiki/Swiss-Prot>) που περιέχει δεδομένα σε όρους πρωτεϊνών. Επίσης, αν χρειαστεί θα χρησιμοποιούν και δεδομένα από διαθέσιμες βάσεις δεδομένων μικροσυστοιχιών όπως [ArrayExpress](http://www.ebi.ac.uk/microarray-as/ae/) (<http://www.ebi.ac.uk/microarray-as/ae/>), και [Stanford Microarray Database \(SMD\)](http://smd.stanford.edu/) (<http://smd.stanford.edu/>).

Προαπαιτούμενα:

Δομές Δεδομένων, Ανάκτηση Πληροφορίας, Εισαγωγή στη Βιοπληροφορική, Αλγόριθμοι, Τεχνολογία Υλοποίησης, Αλγορίθμων, C++, Java, C#

Αναφορές

- Iliopoulos, C.S., Makris, C., Panagis, Y., Perdikuri, K., Theodoridis, E., Tsakalidis, A.K.: The Weighted Suffix Tree: An Efficient Data Structure for Handling Molecular Weighted Sequences and its Applications. Fundam. Inform. 71(2-3): 259-277 (2006)
- Askitis, N., Zobel, J.: B-tries for disk-based string management. VLDB J. 18(1): 157-179 (2009).
- Algorithms on Strings, Trees and Sequences, D. Gusfield, Cambridge University Press, 1999

35. Τίτλος : Αλγόριθμοι και τεχνικές για ευθυγράμμιση οντολογιών με εφαρμογή σε προβλήματα βιοπληροφορικής.

Υπεύθυνος επ. καθ.: Μακρής Χρήστος

Περιγραφή

Αντικείμενο της διπλωματικής είναι η συγκριτική επισκόπηση των διαθέσιμων αλγορίθμων ταιριάσματος / ευθυγράμμισης οντολογιών και αξιολόγησή τους με βάση τα ιδιαίτερα χαρακτηριστικά της εφαρμογής στην οποία οι αλγόριθμοι αυτοί αναπτύσσονται (κυρίως Βιοπληροφορικής. Στη συνέχεια θα επιλεγούν και θα εφαρμοστούν στις επιμέρους οντολογίες οι κατάλληλοι αλγόριθμοι ώστε να επιτυγχάνεται κάθε φορά η μεγαλύτερη σημασιολογική ομογενοποίηση κατά την αναζήτηση ιατρικής πληροφορίας.

Προαπαιτούμενα:

Δομές Δεδομένων, Ανάκτηση Πληροφορίας, Εισαγωγή στη Βιοπληροφορική, Αλγόριθμοι, Τεχνολογίας Υλοποίησης , Αλγορίθμων, C++, Java, C#

Αναφορές

- K. Eilbeck, S. Lewis, C. Mungall, M. Yandell, L. Stein, R. Durbin, M. Ashburner. The sequence ontology: a tool for the unification of genome annotations. In: Genome Biology, Vol. 6, No. 5, pp. R44, 2005.

36. Τίτλος: eΑνεστραμμένο Ευρετήριο Δύο Επιπέδων για DNA ακολουθίες πρόσδεσης μεταγραφικών παραγόντων

Υπεύθυνος επ. καθ.: Μακρής Χρήστος

Περιγραφή

Οι μεταγραφικοί παράγοντες (Transcription Factors - TF) είναι πρωτεΐνες οι οποίες προσδένουν στο DNA και καθορίζουν τη γονιδιακή έκφραση [1]. Ο ρόλος αυτών των πρωτεϊνών είναι καθοριστικός καθώς μόνο η ακριβής πρόσδεσή τους οδηγεί στην παραγωγή της σωστής ποσότητας mRNA, το οποίο εν συνεχεία θα μεταφραστεί σε πρωτεΐνη [2]. Με τον όρο ακριβή πρόσδεση, εννοούμε την ακριβή ακολουθία DNA που τα TF θα προσδέσουν. Αυτή η DNA ακολουθία έχει μήκος μεταξύ 5 και 15 δεσοξυριβονουκλεοτιδίων. Η τελευταία αποτελεί και το μοτίβο του TF.

Η ιδέα είναι να χρησιμοποιήσουμε το ανεστραμμένο ευρετήριο δύο επιπέδων [3, 4] για να αποθηκεύσουμε αυτά τα μοτίβα σε συνδυασμό με τις συντεταγμένες του DNA στις οποίες το συγκεκριμένο μοτίβο εντοπίζεται. Οι συντεταγμένες αυτές αποτελούνται από το χρωμόσωμα (chr1-chr22, chrY, chrX, chrM), την έναρξη (στο κάθε χρωμόσωμα η αρίθμηση ξεκινάει από το 0. Ως εναρκτήριο σημείο του μοτίβου θεωρείται η αρχική συντεταγμένη στην οποία είναι η πρώτη βάση του μοτίβου) και η λήξη (στο κάθε χρωμόσωμα η αρίθμηση ξεκινάει από το 0. Ως τερματικό σημείο του μοτίβου θεωρείται η τελική συντεταγμένη στην οποία είναι η πρώτη βάση του μοτίβου).

Αυτά τα μοτίβα αποθηκεύονται σε δύο ειδών αρχεία (BED files και TXT files). Λόγο του πολύ μεγάλου αριθμού των σημείων πρόσδεσης και των πιθανών μοτίβων του κάθε TF τα αρχεία στα οποία αποθηκεύονται αυτές οι πληροφορίες καταλήγουν να είναι πολύ μεγάλα (π.χ. bed file 50GB) και η προσπέλασή τους μπορεί να γίνει μόνο σειριακά. Αν και έχει γίνει η αναπαράσταση των αρχείων αυτών σε δυαδική μορφή, αυτός ο τρόπος αναπαράστασης δεν προτιμάται ιδιαίτερα διότι σε αυτή την περίπτωση δεν μπορούν να αναγνωσθούν [5, 6, 7].

Η ιδέα είναι να συνδυάσουμε και τα δύο προαναφερθέντα αρχεία και τις πληροφορίες τους σε ένα ανεστραμμένο ευρετήριο. Έχουμε αποδείξει πως η χωρική αναπαράσταση των βεβαρυμένων DNA δεδομένων είναι αποδοτική [3]. Επιπλέον, αυτό που θα χρειαστεί να πετύχουμε είναι να καταφέρουμε εκτός από τις βεβαρυμένες ακολουθίες, να εντάξουμε στα ανεστραμμένα αρχεία και τις πληροφορίες σχετικά με τις γονιδιακές συντεταγμένες του κάθε μοτίβου. Τα ανεστραμμένα αρχεία είναι εύκολα και γρήγορα προσβάσιμα καθώς από όλες τις γλώσσες προγραμματισμού προσφέρονται βιβλιοθήκες προσπέλασης τέτοιων αρχείων. Έτσι, από τη μία πλευρά κάνουμε αποδοτική διαχείριση του χώρου αποθήκευσης που χρειαζόμαστε και από την άλλη πλευρά χρησιμοποιούμε μία αναπαράσταση φιλική (αναγνώσιμη) και εύκολα διαχειρίσιμη από το χρήστη. Ένας ακόμα στόχος μας θα πρέπει να είναι η χρονικά αποδοτική προσπέλαση των ευρετηρίων αυτών. Ακόμα είναι πολύ πιθανή η χρήση προηγμένων δομών δεδομένων (π.χ. tries, suffix tress κ.α.). Τέλος, να αναφέρουμε ότι αποσκοπούμε και σε μία online εφαρμογή η οποία θα αναπαριστά τα δεδομένα αυτά αποδοτικά και θα είναι εύκολα προσπελάσιμη (π.χ. [8, 9]).

Προαπαιτούμενα:

Δομές Δεδομένων, Ανάκτηση Πληροφορίας, Εισαγωγή στη Βιοπληροφορική, Αλγόριθμοι, Τεχνολογίας Υλοποίησης, Αλγορίθμων, C++, Java, C#

Αναφορές

1. Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. *Nature* 386, 569-577.
2. Carey, M., and Smale, S.T. (2000). *Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques* (Cold Spring Harbor, NY: Cold Spring Harbor Lab).
3. Diamanti K., et. al (2014). Handling Weighted Sequences employing Inverted Files and Suffix Trees. In *Proceedings of the 10th International Conference on Web Information Systems and Technologies*, pp. 231-238.
4. Kim, M.-S., et. al (2005). n-gram/2l: A space and time efficient twolevel n-gram inverted index structure. In *International Conference on Very Large Databases (VLDB)*, pp. 325-336.
5. Kheradpour, P., and Kellis, M. (2014). Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Research* 42(5):2976-2987.
6. Daily, K., et. al (2011). MotifMap: integrative genome-wide maps of regulatory motif sites for model species. *BMC Bioinformatics* 12:495.
7. Xie X., et. al (2009). MotifMap: a human genome-wide map of candidate regulatory motif sites. *Bioinformatics* 25:167-174.
8. <http://compbio.mit.edu/encode-motifs/>

9. <http://motifmap.ics.uci.edu/>

37. Τίτλος : Πολιτικές και αλγόριθμοι διαχείρισης ουρών προτεριότητας σε λειτουργικά συστήματα μικροπυρήνα.

Υπεύθυνος επ. καθ.: Μακρής Χρήστος

Περιγραφή

Το Minix ανήκει στην κατηγορία των λειτουργικών συστημάτων μικροπυρήνα (microkernel operating systems) και ένα από τα βασικά μειονεκτήματά του είναι ότι είναι σχετικά αργό λόγω του μεγάλου αριθμού μηνυμάτων που ανταλλάσσονται ανάμεσα από τα διάφορα τμήματά του (servers και kernel). Παρολαυτά έχουν δημιουργηθεί παραλλαγές του Minix που προσπαθούν να είναι real time μειώνοντας τον χρόνο εξυπηρέτησης των διάφορων μηνυμάτων στις ουρές παραλαβής και αποστολής. Αντικείμενο της διπλωματικής είναι η μελέτη των βασικών πολιτικών διαχείρισης και αποστολής μηνυμάτων που έχουν προταθεί για να κάνουν το Minix γρήγορο και αν εφικτό on-line και η πρόταση νέων πολιτικών διαχείρισης ουρών προτεραιότητας.

Προαπαιτούμενα:

Δομές Δεδομένων, Λειτουργικά Συστήματα, Εργαστήριο Λειτουργικών Συστημάτων.

Αναφορές

- http://www.rtminix3.org/documents/rtminix3_report.pdf (βασική πηγή, κοιτάξτε και το συναφές <http://www.rtminix3.org>)
- <http://www.minix3.org/theses/moolenbroek-multimedia-support.pdf>
- <http://www.minix3.org/theses/herder-true-microkernel.pdf>

38. Τίτλος : Αξιολόγηση τεχνικών απόκρυψης προσωπικών δεδομένων κατά την εξόρυξη δεδομένων από ΒΔ συναλλαγών

Υπεύθυνος επ. καθ.: Μακρής Χρήστος, Βασίλειος Βερύκιος

Περιγραφή

Η απόκρυψη γνώσης αποτελεί μια αναδυόμενη περιοχή έρευνας η οποία απευθύνεται στην απόκρυψη ευαίσθητων προσωπικών δεδομένων με σκοπό την αποφυγή εξόρυξής τους. Οι διάφορες τεχνικές απόκρυψης δεδομένων που έχουν αναπτυχθεί τα τελευταία χρόνια αποσκοπούν στην αποτελεσματική απόκρυψη συνολο-στοιχείων από βάσεις δεδομένων συναλλαγών ώστε να μην επηρεάζονται τα υπόλοιπα δεδομένα. Σκοπός αυτής της διπλωματικής εργασίας αποτελεί η ανάπτυξη ενός ενιαίου μοντέλου αξιολόγησης τέτοιων τεχνικών. Το μοντέλο αυτό θα πρέπει να συνδυάζει διάφορες συνιστώσες επιτυχημένης ή αποτυχημένης απόκρυψης δεδομένων, ώστε να εξάγει ορθά συμπεράσματα.

Προαπαιτούμενα:

Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java), Εξόρυξη Γνώσης

Αναφορές:

- [1] Y. Wu, C. Chiang, and A. L. P. Chen, "Hiding sensitive association rules with limited side effects," *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, no. 1, pp. 29–42, 2007.
- [2] A. Gkoulalas-Divanis and V. S. Verykios, "Exact knowledge hiding through database extension," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 5, pp. 699–713, 2009.
- [3] Chun-Wei Lin, Tzung-Pei Hong, and Hung-Chuan Hsu, "Reducing Side Effects of Hiding Sensitive Itemsets in Privacy Preserving Data Mining," *The Scientific World Journal*, vol. 2014, Article ID 235837, 12 pages, 2014. doi:10.1155/2014/235837
- [4] C. Dwork, "Differential privacy," in *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)(2)*, 2006, pp. 1–12.

39. Τίτλος : **Big data real time analytics στο Twitter για την εύρεση χρηστών με επιρροή**

Υπεύθυνος επ. καθ.: Μακρής Χρήστος, Καφέζα Ελεάννα

Περιγραφή

Ο στόχος αυτής της εργασίας είναι η εύρεση των χρηστών που έχουν επιρροή στα κοινωνικά δίκτυα σε πραγματικό χρόνο. Στη διπλωματική αυτή θα χρησιμοποιηθούν εργαλεία και βιβλιοθήκες για real time machine learning. Κάθε χρήστης στο twitter έχει μια προσωπικότητα την οποία μπορούμε να εξάγουμε με βάση το μοντέλο των Big Five. Η διπλωματική αυτή θα βασιστεί σε έρευνα που ήδη έχει γίνει και δημοσιευτεί η οποία προβλέπει την προσωπικότητα του χρήστη και αποφασίζει για το Pattern διάδοσης της πληροφορίας που θα ακολουθήσει στο Twitter. Θα επεκτείνει αυτήν την εργασία στην εύρεση της προσωπικότητας του χρήστη σε πραγματικό χρόνο . Το σύστημα θα συλλέγει τα τελευταία tweets του χρήστη και άλλη σχετική πληροφορία και θα αποφασίζει αν επηρεάζει ή όχι την μετάδοση της πληροφορίας στο δίκτυο. Η εφαρμογή της εύρεσης χρηστών με επιρροή είναι μεγάλη για την διάδοση διαφήμισαν αλλά και άλλων μηνυμάτων στα κοινωνικά δίκτυα.

Προαπαιτούμενα:

Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java), Εξόρυξη Γνώσης

Αναφορές:

T-PICE: Twitter Personality based Influential Communities Extraction System (BigData2014-5028) Eleanna Kafeza , Andreas Kanavos, Christos Makris, and Pantelis Vikatos

40. Τίτλος : Topic-based- Real time community detection for mobile users

Υπεύθυνος επ. καθ.: Μακρής Χρήστος, Καφέζα Ελεάννα

Περιγραφή

Ο στόχος αυτής της εργασίας είναι η δημιουργία σε πραγματικό χρόνο κοινοτήτων χρηστών με βάση το θέμα συζήτησης. Η εργασία αυτή θα έχει σαν είσοδο ένα σύνολο από tweets καθώς και τη θέση των χρηστών με βάση τις συντεταγμένες που θα εξαχθούν από τη χρήση του κινητού τηλεφώνου. Με βάση το θέμα συζήτησης, την χρονική στιγμή του tweet, την ομοιότητα των tweets, την συχνότητα των tweets του χρήστη και άλλες παραμέτρους, το σύστημα θα αποφασίζει αν οι συγκεκριμένοι χρήστες ανήκουν σε μια τοπική κοινότητα που θα ήθελε να συνδεθεί. Στην διπλωματική θα χρησιμοποιηθούν εργαλεία για real time analytics στο twitter.

Προαπαιτούμενα:

Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java), Εξόρυξη Γνώσης

Αναφορές:

Garzo, A; Benczur, AA; Sidlo, C.I; Tahara, D.; Wyatt, E.F., "Real-time streaming mobility analytics," Big Data, 2013 IEEE International Conference on , vol., no., pp.697,702, 6-9 Oct. 2013

L. A. Tang, Y. Zheng, J. Yuan, J. Han, A. Leung, C.-C. Hung, and W.-C. Peng. On discovery of traveling companions from streaming trajectories. In ICDE, pages 186–197, 2012

41. Τίτλος : Η εφαρμογή community detection αλγορίθμων με χρήση του Mahout

Υπεύθυνος επ. καθ.: Μακρής Χρήστος, Καφέζα Ελεάννα

Περιγραφή

Ο στόχος της διπλωματικής αυτής είναι η διερεύνηση του Mahout, για community detection καθώς και την ανακάλυψη των patterns συμπεριφοράς των χρηστών σε κάθε κοινότητα για μεγάλο όγκο δεδομένων του Twitter. Ένα από τα βασικά προβλήματα των αλγορίθμων για εύρεση κοινοτήτων στα κοινωνικά δίκτυα είναι ο χρόνος εκτέλεσης και οι απαιτήσεις σε μνήμη. Για αυτό τον λόγο έχουν προταθεί λύσεις στην υλοποίηση τέτοιων αλγορίθμων με χρήση MapReduce. Αυτή η διπλωματική θα εξετάσει την χρήση του Mahout για την επίλυση του προβλήματος.

Προαπαιτούμενα:

Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java), Εξόρυξη Γνώσης

Αναφορές:

Shi, J.; Xue, W.; Wang, W.; Zhang, Y.; Yang, B.; Li, J., "Scalable community detection in massive social networks using MapReduce," IBM Journal of Research and Development , vol.57, no.3/4, pp.12:1,12:14, May-July 2013

42. Τίτλος : Επαλήθευση (verification) ορθής διάχυσης πληροφορίας σε κοινωνικά δίκτυα

Περιγραφή

Σύμφωνα με τη βιβλιογραφία, το πρότυπο διάχυσης της πληροφορίας είναι ένα σύνολο από δέντρα που έχουν μικρό βάθος. Επομένως, βασιζόμενοι εκεί, μπορούμε να πούμε ότι τα communities που έχουν δημιουργηθεί από τον κερματισμό ενός κοινωνικού δικτύου/γράφου, δεν θα έχουν μεγάλες αλυσίδες από influence, αλλά ότι θα υπάρχουν πολλά μικρά δέντρα. Οπότε για να ελέγξουμε την αποδοτικότητα της λύσης μας, θα ακολουθήσουμε την παρακάτω μέθοδο.

Αρχικά, θα μετρήσουμε σε κάθε ένα από τα δίκτυα που σχηματίσαμε, τον αριθμό των retweets. Στην παρακάτω εργασία μελετάται εμπειρικά η μετάδοση των μηνυμάτων. Έτσι, μπορούμε να εξετάσουμε τον αριθμό των “μικρών δέντρων” μετάδοσης των μηνυμάτων στα δίκτυα μας. Δηλαδή να μετρήσουμε πόσες μεταδόσεις μηνυμάτων 1 βαθμού έχουμε σε κάθε υπο-δίκτυο μας.

Προαπαιτούμενα: Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Α. Καναβός

Επικοινωνία: kanavos@ceid.upatras.gr

Σχετικές Αναφορές:

1. Goel S., Watts D. & Goldstein D. The Structure of Online Diffusion Networks

43. Τίτλος : Προσωπικότητα (personality) κόμβων και πως επιδράει στη διάχυση των πληροφοριών στα κοινωνικά δίκτυα

Περιγραφή

Η διπλωματική έχει σχέση με κοινωνικά δίκτυα και θέλουμε να χρησιμοποιήσουμε κάποια από τα μοντέλα για διάχυση των πληροφοριών (diffusion) που να λαμβάνει υπόψη το personality του κόμβου. Υπάρχουν δυο βασικά μοντέλα, το The Linear Threshold και το Independent Cascade. Αυτά τα μοντέλα είναι πιθανοτικά και βασίζονται σε τυχαίες επιλογές.

Στη διπλωματική αυτή θέλουμε να εξετάσουμε μήπως η πιθανότητα να μεταδώσει κάποιος το μήνυμα εξαρτάται και από το personality. Επιπλέον θέλουμε να μελετήσουμε το αν η μετάδοση σ' ένα κοινωνικό δίκτυο ακολουθεί κάποιο από αυτά τα δύο μοντέλα.

Προαπαιτούμενα: Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Α. Καναβός

Επικοινωνία: kanavos@ceid.upatras.gr

Σχετικές Αναφορές:

1. Kempe D., Kleinberg J. & Tardos E. Maximizing the Spread of Influence through a Social Network

44. Τίτλος: Εξατομίκευση περιεχομένου με αξιοποίηση γνώμης - κριτικής και των κοινωνικών δικτύων (aspect based opinion mining)

Περιγραφή

Στόχος της διπλωματικής είναι η ενσωμάτωση σ' ένα μοντέλο των κριτικών των χρηστών αλλά και του κοινωνικού δικτύου τους ώστε να καταλήγει στις καλύτερες προτάσεις.

Προαπαιτούμενα: Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Α. Καναβός

Επικοινωνία: kanavos@ceid.upatras.gr

Σχετικές Αναφορές:

1. Community detection in graphs, Santo Fortunato
2. <http://dl.acm.org/citation.cfm?id=2010006>

45. Τίτλος: Θεωρητική και πειραματική μελέτη διάχυση πληροφορίας σε κοινωνικά δίκτυα

Περιγραφή

Σε αυτή την διπλωματική θα εξεταστούν μοντέλα διάχυσης πληροφορίας σε γραφήματα κοινωνικών δικτύων και θα γίνει συγκριτική μελέτη μεταξύ τους σε πραγματικά και συνθετικά δεδομένα.

Προαπαιτούμενα: Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης, Θεωρία Αλγορίθμων

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Π. Βικάτος

Επικοινωνία: vikatos@ceid.upatras.gr

Σχετικές Αναφορές:

1. <http://hal.archives-ouvertes.fr/docs/00/84/80/50/PDF/survey.pdf>
2. <http://lilianweng.github.io/papers/weng-thesis-single.pdf>

46. Τίτλος: Δημιουργία universal crawler για κοινωνικά δίκτυα και οπτικοποίηση δεδομένων και γραφημάτων

Περιγραφή

Δημιουργία universal crawler για κοινωνικά δίκτυα και οπτικοποίηση δεδομένων και γραφημάτων. Αντικείμενο αυτής της εργασίας είναι η δημιουργία crawler για την συλλογή δεδομένων από κοινωνικά δίκτυα (Facebook, Twitter, Linkedin), καθώς και η οπτικοποίηση των χαρακτηριστικών και των γραφημάτων.

Προαπαιτούμενα: Ανάκτηση Πληροφορίας, Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Π. Βικάτος

Επικοινωνία: vikatos@ceid.upatras.gr

Σχετικές Αναφορές:

1. <https://akela.mendelu.cz/~xfejfar/GIS/Mining%20the%20Social%20Web,%20Second%20Edition%20-%20Russell,%20Matthew%20A.pdf>
2. <http://www.amazon.com/Recipes-Mining-Twitter-Matthew-Russell/dp/1449303161>
3. http://mkurant.com/publications/papers/Gjoka_JSAC_2011_Practical.pdf

47. Τίτλος: Θεωρητική και πειραματική μελέτη αλγορίθμων εξαγωγής κοινοτήτων και ανίχνευσης influential users από κοινωνικά δίκτυα

Περιγραφή

Αντικείμενο αυτής της διπλωματικής είναι η μελέτη των διαφόρων τεχνικών ανίχνευσης κοινοτήτων σε γραφήματα κοινωνικών δικτύων, η συγκριτική μελέτη και η οπτικοποίηση κοινοτήτων από γραφήματα.

Προαπαιτούμενα: Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης, Θεωρία γραφημάτων

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Π. Βικάτος

Επικοινωνία: vikatos@ceid.upatras.gr

Σχετικές Αναφορές:

1. Community detection in graphs, Santo Fortunato
2. Fragkiskos D. Malliaros, Michalis Vazirgiannis: Clustering and Community Detection in Directed Networks: A Survey. CoRR abs/1308.0971 (2013)

48. Τίτλος: Πρόβλεψη σύνδεσης χρηστών κοινωνικών δικτύων

Περιγραφή

Αντικείμενο αυτής της διπλωματικής είναι η θεωρητική και πειραματική μελέτη αλγορίθμων εξόρυξης δεδομένων για την πρόβλεψη σύνδεσης ανάμεσα σε κόμβους κοινωνικών δικτύων.

Προαπαιτούμενα: Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης, Θεωρία γραφημάτων

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Π. Βικάτος

Επικοινωνία: vikatos@ceid.upatras.gr

Σχετικές Αναφορές:

1. <http://www.cs.rpi.edu/~zaki/PaperDir/SNDA11.pdf>
2. <https://www.cs.cornell.edu/home/kleinber/link-pred.pdf>

49. Τίτλος: Ανίχνευση μοτίβων χρονοσειρών από κοινωνικά δίκτυα

Περιγραφή

Αντικείμενο αυτής της διπλωματικής αποτελεί η μελέτη τεχνικών ανίχνευσης μοτίβων από χρονοσειρές μετρικών από κοινωνικά δίκτυα.

Προαπαιτούμενα: Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης, Θεωρία γραφημάτων

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Π. Βικάτος

Επικοινωνία: vikatos@ceid.upatras.gr

Σχετικές Αναφορές:

1. http://www.philippsinger.info/masterarbeit_psinger.pdf
2. http://cs.gmu.edu/~jessica/Lin_motif.pdf
3. <http://www.cc.gatech.edu/~isbell/papers/minnen-icdm2007.pdf>

50. Τίτλος : Θεωρητική και πειραματική μελέτη text mining τεχνικών σε περιβάλλοντα υπολογιστικού νέφους

Περιγραφή

Αντικείμενο αυτής της διπλωματικής είναι η μελέτη διαφόρων τεχνικών εξόρυξης δεδομένων με χρήση περιβαλλόντων υπολογιστικού νέφους.

Προαπαιτούμενα: Γνώση Προγραμματισμού (C++, Java, κ.α.), Εξόρυξη Γνώσης, mapR, Hadoop, HBASE

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Π. Βικάτος

Επικοινωνία: vikatos@ceid.upatras.gr

Σχετικές Αναφορές:

1. Hadoop: The Definite Guide, Tom White, O' REILLY
2. Welcome to Apache Hadoop <http://hadoop.apache.org/>
3. Survey of Text Mining: Clustering, Classification, and Retrieval (2004)
4. A Survey of Text Mining Techniques and Applications (2009)

51. Τίτλος : Δυναμική δρομολόγηση (dynamic scheduling) για βελτίωση του IAAS scheduling στο cloud

Περιγραφή

Αυτή η διπλωματική θα εξετάσει ήδη υπάρχουσες λύσεις σε Scheduling όπως τη χρήση First Fit και Round robin από τον Eucalyptus, τη χρήση του Haizea VM resource management με τεχνικές διαχείρισης ουρών και preemption scheduling από το Open Nebula και την χρήση εργαλείων δρομολόγησης από το Nimbus. Στη συνέχεια θα υλοποιηθούν νέες τεχνικές dynamic scheduling. Τα πειράματα θα γίνουν με την χρήση του CloudSim.

Προαπαιτούμενα: Γνώση Προγραμματισμού, Ανάκτησης πληροφορίας

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Ελεάννα Καφέζα

Επικοινωνία: kafeza@aueb.gr

Σχετικές Αναφορές

1. Hai Zhong; Kun Tao; Xuejie Zhang, "An Approach to Optimized Resource Scheduling Algorithm for Open-Source Cloud Systems," ChinaGrid Conference (ChinaGrid), 2010 Fifth Annual , vol., no., pp.124,129, 16-18 July 2010
2. Vijindra, Sudhir Shenai, Survey on Scheduling Issues in Cloud Computing, Procedia Engineering, Volume 38, 2012, Pages 2881-2888, ISSN 1877-7058

52. Τίτλος : Παρακολούθηση καλής εκτέλεσης συμβάσεων σε περιβάλλον cloud

Περιγραφή

Στην διπλωματική αυτή θα μελετηθούν οι συμβάσεις που κάνει ο πάροχος cloud υπηρεσιών με τους πελάτες του και οι περιορισμοί που θέτουν τα SLAs. Με βάση υπάρχοντα εργαλεία παρακολούθησης της λειτουργίας και της απόδοσης στο cloud (IAAS εργαλεία) θα μοντελοποιηθεί η εκτέλεση της σύμβασης και θα δημιουργηθούν alerts και μηχανισμοί οι οποίοι θα επιχειρούν να την νομική προστασία και των δύο μερών καθώς και εναλλακτικές λύσεις σε περίπτωση που κάποιους από τους όρους δεν ικανοποιείται.

Προαπαιτούμενα: Γνώση Προγραμματισμού, Ανάκτησης πληροφορίας

Επιβλέπων: Χ. Μακρής

Συνεπιβλέπων: Ελεάννα Καφέζα

Επικοινωνία: kafeza@aueb.gr

Σχετικές Αναφορές

1. <http://cloudstack.apache.org/>
2. <http://www.cs.man.ac.uk/~rizos/papers/europar11b.pdf>
3. Freitas, A.L. ; Parlavantzas, N. ; Pazat, J.-L. An Integrated Approach for Specifying and Enforcing SLAs for Cloud Services, Cloud Computing (CLOUD), 2012

ΣΥΝΕΡΓΑΣΙΕΣ ΜΕ ΙΙΒΕΑΑ (Ίδρυμα Ιατροβιολογικών Ερευνών της Ακαδημίας Αθηνών (ΙΙΒΕΑΑ)) – παρατίθεται 4 ενδεικτικό θέματα)

1. Τίτλος : Computer Aided Drug Design

Περιγραφή

Since drug discovery is a time consuming and quite expensive process, novel, innovative drug design approaches and methodologies are constantly developed. Recently, the scientific interest in the drug design field has been focused on computer-based bio-simulations and 3D molecular modeling. Herein, we aim to establish the 3D structure of a series of viral enzymes (in the HCV, Dengue, Yellow fever family), which will allow and facilitate initially the in silico design and evaluation of novel and potent antiviral agents, using structure based de novo drug design techniques. The ultimate goal is to establish a pipeline that will aid scientists in the screening and design of antiviral agents towards the eradication of major "killer" viral infections & diseases worldwide. The student will be acquainted with a series of diverse programming skills currently used in this field. Basic Perl or Python programming skills are required for this project.

2. Τίτλος : Machine Aided Drug Design

Περιγραφή

The development of uncountable software packages for drug design has somehow turned a highly complex process into a "black box" application, where the scientist prepares the input files, chooses some parameters for the analysis and leaves the rest to the computer. Inevitably, basing the drug discovery process to the computer's pseudo-knowledge marginalizes human intervention and researcher knowledge. Towards this end we are currently developing new technological hardware aids that may be used to facilitate rational drug design, by focusing on an interactive informed human intervention, where the computer is dethroned into a more assistive role. The user uses specialized 3D visual and handling devices that allow him to experience a full scale 3D molecular interaction with the simulation environment. New software is currently being developed so that real time tracking of the user's movement may be incorporated to the molecular simulation in a dynamic way that eliminates the need for computationally expensive calculations. Programming skills are a pre-requisite in any of the following: C, Perl, AWK or CUDA for the development of GPU based code.

3. Τίτλος : Evaluation of web based platform

Περιγραφή

As part of an ongoing EU project that will be delivering a suite of data mining, collaboration and decision support services, we aim to evaluate the usability and efficiency of those services. Our principal goal is to familiarize with the provided services by employing them to clinico-genomics, medical treatment, and opinion mining data. Additionally, sampling methodologies will be studied in order to decide upon the evaluation strategy followed and the adequate sample size. As part of this study, questionnaires will be constructed, web distributed and results will be analysed based on standard statistical methodologies as well as presented in graphical format. Methodologies for handling missing values will be also considered.

4. Τίτλος : Statistical Analysis of Genomic Data

Περιγραφή

Given the rapid advances in genomics and bioinformatics that have taken place in the past few years, there is a growing need for analysing vast amount of data and interpreting their results. Large-scale cancer genome studies have successfully applied some preliminary integration approaches. To this end, we aim to analyse multi-source data using association statistics to estimate pair wise as well as group dependencies in the data. Multivariate analysis and likelihood-based inference will be also employed to estimate data patterns. The ultimate goal is to establish a methodology to integrate different data which measure multiple genomic features and discover or validate findings that would not be discovered by analysing each data independently. Good programming skills are a pre-requisite (e.g. R statistical software, C/C++).

Προαπαιτούμενα	Δομές Δεδομένων, Αλγόριθμοι, Βιοπληροφορική, C++/Java
Άτομα	1
Υπεύθυνος Καθηγητής	επ. καθ. Χ Μακρής, καθ. Α. Τσακαλίδης, Δρ. Σοφία Κοσσίδα (ΙΙΒΕΑΑ)
Contact person	skossida@bioacademy.gr