ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

- Τόμος 9 -

ΕΠΙΛΕΓΜΕΝΕΣ

ΠΤΥΧΙΑΚΕΣ & ΔΙΠΛΩΜΑΤΙΚΕΣ ΕΡΓΑΣΙΕΣ

ΑΘΗΝΑ 2012

ΕΠΙΛΕΓΜΕΝΕΣ

ΠΤΥΧΙΑΚΕΣ & ΔΙΠΛΩΜΑΤΙΚΕΣ ΕΡΓΑΣΙΕΣ

- Τόμος 9 -

Εκδίδεται μία φορά το χρόνο από το:

Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών,
Πανεπιστημιούπολη, 15784 Αθήνα

Τηλ: 210 - 727 5190,   Φαξ: 210 - 727 5333
email: library@di.uoa.gr,   url: http://www.di.uoa.gr/lib

Επιμέλεια έκδοσης:

Επιτροπή Ερευνητικών και Αναπτυξιακών Δραστηριοτήτων

Θ. Θεοχάρης (υπεύθυνος έκδοσης), Καθηγητής, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Η. Μανωλάκος, Αναπληρωτής Καθηγητής, Τμήμα Πληροφορικής και Τηλεπικοινωνιών

Γραφιστική Επιμέλεια:

Λ. Χαλάτση, Γραφείο Προβολής, Τμήμα Πληροφορικής και Τηλεπικοινωνιών

ISSN 1792-8826

Περιεχόμενα

ΔΙΠΛΩΜΑΤΙΚΕΣ ΕΡΓΑΣΙΕΣ

Πρόλογος

Ο τόμος αυτός περιλαμβάνει περιλήψεις των καλύτερων διπλωματικών και πτυχιακών εργασιών που εκπονήθηκαν στο Τμήμα Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών κατά το διάστημα **1/1/2011 - 31/12/2011**. Πρόκειται για τον 9ο τόμο στη σειρά αυτή. Στόχος του θεσμού είναι η ενθάρρυνση της δημιουργικής προσπάθειας και η προβολή των πρωτότυπων εργασιών των φοιτητών του Τμήματος.

Η έκδοση αυτή είναι ψηφιακή και έχει δικό της ISSN. Αναρτάται στην επίσημη ιστοσελίδα του Τμήματος και έτσι, εκτός από τη μείωση της δαπάνης κατά την τρέχουσα περίοδο οικονομικής κρίσης, έχει και μεγαλύτερη προσβασιμότητα. Για το στόχο αυτό, σημαντική ήταν η συμβολή της Λήδας Χαλάτση που επιμελήθηκε και φέτος την ψηφιακή έκδοση και πέτυχε μια ελκυστική ποιότητα παρουσίασης, ενώ βελτίωσε και την ομοιογένεια των κειμένων.

Η συμμετοχή των φοιτητών ήταν και φέτος πολύ ικανοποιητική. Η στάθμη των εργασιών που εγκρίθηκαν (αλλά και αυτών που υποβλήθηκαν) ήταν υψηλή και τα γνωστικά αντικείμενα που πραγματεύονται οι εργασίες καλύπτουν σε μεγάλο βαθμό ολόκληρη την περιοχή της Πληροφορικής και των Τηλεπικοινωνιών.

Θα θέλαμε να ευχαριστήσουμε τους φοιτητές για το χρόνο που αφιέρωσαν για να παρουσιάσουν τη δουλειά τους στα πλαίσια αυτού του θεσμού και να τους συγχαρούμε για την ποιότητα των εργασιών τους. Ελπίζουμε η διαδικασία αυτή να προσέφερε και στους ίδιους μια εμπειρία που θα τους βοηθήσει στη συνέχεια των σπουδών τους ή της επαγγελματικής τους σταδιοδρομίας.

Η Επιτροπή Ερευνητικών και Αναπτυξιακών Δραστηριοτήτων

Θ. Θεοχάρης (υπεύθυνος έκδοσης), Η. Μανωλάκος

Αθήνα, Ιούνιος 2012

ΕΠΙΛΕΓΜΕΝΕΣ
ΠΤΥΧΙΑΚΕΣ ΕΡΓΑΣΙΕΣ

Μελέτη και Υλοποίηση Τεχνικών Κατάστρωσης Σχεδίου σε Περιβάλλον Unity3D

Βασίλειος - Μάριος Αναστασίου
astyanax@di.uoa.gr

Παναγιώτης Διαμαντόπουλος
panos_10d@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Τα τελευταία χρόνια, ακαδημαϊκές τεχνικές Τεχνητής Νοημοσύνης και πιο συγκεκριμένα τεχνικές κλασσικού σχεδιασμού για τον καθορισμό συμπεριφορών προσανατολισμένων στο στόχο έχουν αποδειχθεί ιδιαίτερα χρήσιμες στον προγραμματισμό ηλεκτρονικών παιχνιδιών. Οδηγούμενοι από το γεγονός ότι δεν υπάρχει σαφώς ορισμένο πρότυπο για την ανάπτυξη τέτοιων συμπεριφορών σε ηλεκτρονικά παιχνίδια, υλοποιήσαμε την iThink, μία βιβλιοθήκη που επιτρέπει τη χρήση ακαδημαϊκών τεχνικών για την εφαρμογή τέτοιων συμπεριφορών σε πραγματικά περιβάλλοντα ανάπτυξης βιντεοπαιχνιδιών. Στην παρούσα εργασία, εστιάζουμε στο φορμαλισμό STRIPS καθώς και στο περιβάλλον Unity3D, μία δημοφιλή μηχανή παιχνιδιών. Εκτός από τη χρηστικότητα της iThink για τους προγραμματιστές, θεωρούμε ότι μπορεί να αξιοποιηθεί και στην εκπαίδευση, παρέχοντας ένα μοντέρνο και διασκεδαστικό μέσο εκμάθησης και πειραματισμού με τον κλασσικό σχεδιασμό.

Λέξεις-Κλειδιά: Κατάστρωση σχεδίου, Μηχανή παιχνιδιών, Unity3D, C#, STRIPS

Επιβλέποντες:
Κουμπαράκης Μανώλης, Καθηγητής
Βάσσος Σταύρος, Μεταδιδακτορικός Ερευνητής

1. Εισαγωγή

Παραδοσιακά, οι τεχνικές Τεχνητής Νοημοσύνης που χρησιμοποιούνταν για τη δημιουργία βιντεοπαιχνιδιών δεν βασίζονταν σε τεχνικές που είχαν αναπτυχθεί από την πανεπιστημιακή κοινότητα. Στην ουσία, οι προγραμματιστές δημιουργούσαν ψευδαίσθηση νοημοσύνης, αξιοποιώντας προγραμματιστικά κόλπα χωρίς να βασίζονται στη μελέτη εκλεπτυσμένων τεχνικών από τον τομέα της ακαδημαϊκής Τεχνητής Νοημοσύνης [6, 13]. Αυτή η τάση αποδείχθηκε παρά ταύτα ιδιαίτερα αποδοτική, καθώς οδηγούσε σε ταχύτατη ανάπτυξη συμπεριφορών για τους πράκτορες του βιντεοπαιχνιδιού, συχνά υπό τη μορφή μιας σειράς από προγραμματισμένες ενέργειες που βασίζονταν σε συνθήκες που αφορούσαν τον κόσμο του παιχνιδιού και τους παίκτες.

Αυτό δεν αποτελούσε πρόβλημα για αρκετά χρόνια καθώς η βιομηχανία των παιχνιδιών μπόρεσε να εξελιχθεί εξαιρετικά καλά αξιοποιώντας τις ραγδαίες εξελίξεις σε άλλους τομείς όπως τα γραφικά, τομέας που έγινε ενθουσιωδώς αποδεκτός από την πλειοψηφία των παικτών. Η ανάγκη για τη δημιουργία πιο ανταγωνιστικής Τεχνητής Νοημοσύνης εξασφαλιζόταν παρέχοντας στους χαρακτήρες περισσότερη δύναμη από αυτή που είχαν οι άνθρωποι παίκτες (π.χ να έχουν άνομο πλεονέκτημα). Τα τελευταία χρόνια η βιομηχανία των παιχνιδιών έχει φτάσει σε ένα σημείο που είναι απαραίτητη πλέον η χρήση πιο εκλεπτυσμένων τεχνικών για τη συμπεριφορά των πρακτόρων. Αυτό έχει γίνει αντιληπτό τόσο από τους παίκτες που αναζητούν πιο ανταγωνιστικούς πράκτορες τεχνητής νοημοσύνης (NPC’s), ώστε να δίνουν την αντίληψη ενός ανθρώπου με στόχους που αντιδρά στο περιβάλλον του [4, 7], όσο και από τους προγραμματιστές οι οποίοι αναζητούν έναν επεκτάσιμο, επαναχρησιμοποιήσιμο και διατηρήσιμο πρότυπο για τη δημιουργία πρακτόρων Τεχνητής Νοημοσύνης, αφού πλέον τα ηλεκτρονικά παιχνίδια έχουν συνεχώς αυξανόμενη πολυπλοκότητα [9].

Μία αξιοσημείωτη τεχνική Τεχνητής Νοημοσύνης που προέρχεται από την ακαδημαϊκή κοινότητα και αποδείχθηκε ιδιαίτερα χρήσιμη για τον σκοπό αυτό, είναι ο κλασικός σχεδιασμός που συχνά αναφέρεται στη βιομηχανία παιχνιδιών ως κατάστρωση σχεδίου με ενέργειες προσανατολισμένες στο στόχο (goal oriented action planning). Η ιδέα στην οποία στηρίζεται αυτή η τεχνική είναι η παροχή ενός συνόλου ενεργειών, με την περιγραφή των επιδράσεων που αυτές έχουν στο περιβάλλον, καθώς και ένα στόχο στον οποίο καλείται να οδηγηθεί ο πράκτορας χρησιμοποιώντας αυτές τις ενέργειες. Δόθηκε ιδιαίτερη προσοχή σε αυτήν την τεχνική κυρίως λόγο του διάσημου παιχνιδιού F.E.A.R. [8] που χρησιμοποιούσε μία απλουστευμένη έκδοση του φορμαλισμού STRIPS [3] για τη μοντελοποίηση της συμπεριφοράς πρακτόρων τεχνητής νοημοσύνης.

Στην παρούσα εργασία εστιάζουμε στο βασικό πρόβλημα υλοποίησης μίας πλατφόρμας κλασικού σχεδιαστή η οποία μπορεί να χρησιμοποιηθεί σε πληθώρα προγραμματιστικών περιβαλλόντων ηλεκτρονικών παιχνιδιών. Αναλύουμε τη διαδικασία ανάπτυξης του iThink, μίας βιβλιοθήκης κλασικού σχεδιασμού βασισμένη στο STRIPS και η οποία στοχεύει στη γενίκευση της διαδικασίας σχεδιασμού, στην ευκολία χρήσης και στην ενσωμάτωση σε μία πραγματική μηχανή παιχνιδιών. Στην ενότητα 2, θα γίνει μία σύντομη περιγραφή του STRIPS ενώ στην ενότητα 3 θα γίνει η περιγραφή της βιβλιοθήκης iThink. Στην ενότητα 4 γίνεται αναφορά σε σχετικές δημοσιεύσεις που έχουν ολοκληρωθεί πρόσφατα, στην ενότητα 5 αναφέρονται μελλοντικές επεκτάσεις της βιβλιοθήκης iThink και καταλήγουμε στην ενότητα 6 με τα γενικά συμπεράσματα.

2. STRIPS planning

Σε αυτή την περιοχή της Τεχνητής Νοημοσύνης κάποιος αντιμετωπίζει τη ακόλουθη διαδικασία. Δεδομένης μιας πλήρως ορισμένης αρχικής κατάστασης, μιας κατάστασης στόχου και ενός συνόλου ενεργειών που περιγράφουν τις αλλαγές που μπορεί να επέλθουν στον κόσμο, είναι δυνατή η εύρεση μιας αλληλουχίας ενεργειών έτσι ώστε όταν τις εφαρμόσουμε διαδοχικά, η αρχική κατάσταση να μετατραπεί σε μία κατάσταση που να ικανοποιεί τον επιθυμητό στόχο.

Στην παρούσα εργασία εστιάζουμε στον φορμαλισμό STRIPS για την αναπαράσταση διαδικασιών σχεδιασμού. Στον STRIPS, η αναπαράσταση της αρχικής κατάστασης, των ενεργειών και των καταστάσεων στόχου βασίζεται σε λεκτικά από την κατηγορηματική λογική (predicate logic).

Η αρχική κατάσταση (initial state) περιγράφεται από ένα σύνολο από θετικά λεκτικά, τα οποία παρέχουν μία πλήρη προδιαγραφή της κατάστασης βασισμένη στην υπόθεση κλειστού κόσμου (closed world assumption). Οι ενέργειες (actions) επηρεάζουν την τρέχουσα κατάσταση προσθέτοντας ή αφαιρώντας λεκτικά.

Τα σχήματα ενεργειών (action schemas) είναι γενικοί κανόνες που καθορίζουν τις ενέργειες που μπορούν δυνητικά να εφαρμοστούν σε μία κατάσταση. Τα σχήματα ενεργειών περιγράφονται επιπλέον από δύο ακόμα στοιχεία, τις προϋποθέσεις και τις επιδράσεις τους. Οι προϋποθέσεις (preconditions) είναι ένα σύνολο από λεκτικά των οποίων η ύπαρξη είναι απολύτως απαραίτητη σε μία κατάσταση προκειμένου να μπορεί να εφαρμοστεί η αντίστοιχη ενέργεια. Οι επιδράσεις (effects) είναι ένα σύνολο από λεκτικά τα οποία καθορίζουν τις αλλαγές που θα επιφέρει η εφαρμογή της ενέργειας στην κατάσταση.

Καταλήγοντας η κατάσταση στόχου (goal condition) είναι ένα σύνολο από θετικά λεκτικά τα οποία πρέπει να εμφανίζονται σε μία κατάσταση για να ικανοποιεί αυτή τον επιθυμητό στόχο. Λύση σε ένα πρόβλημα σχεδιασμού αποτελεί μία ακολουθία ενεργειών που αν εφαρμοστεί σειριακά από μία αρχική κατάσταση, κάνοντας τους απαραίτητους ελέγχους για προϋποθέσεις και εφαρμόζοντας τις κατάλληλες επιδράσεις, τη μετατρέπει σε κατάσταση στόχου.

3. Η βιβλιοθήκη iThink στο περιβάλλον Unity3D

3.1 Περιγραφή του εργαλείου Unity3D

Το Unity3D είναι μια πλατφόρμα εργαλείων καθώς και μηχανή παιχνιδιών που προσφέρει υψηλής ποιότητας χαρακτηριστικά και εργαλεία διαχείρισης περιεχομένου, τα οποία την καθιστούν ιδανική για χρήση ως πλατφόρμα ταχείας ανάπτυξης βιντεοπαιχνιδιών καθώς και προσομοιώσεων. Ο προγραμματισμός σε αυτό το περιβάλλον βασίζεται σε scripts κώδικα γραμμένα σε γλώσσες C#, Javascript ή Boo, βασιζόμενος στην πλατφόρμα ανοιχτού κώδικα Mono.

Παρέχεται δωρεάν πακέτο χρήσης που την καθιστά προσιτή σε πληθώρα προγραμματιστών. Τα χαρακτηριστικά της, σε συνδυασμό με τη διαρκώς αυξανόμενη δημοτικότητα και υποστήριξή της, μας οδήγησαν στην επιλογή της ως ιδανικής πλατφόρμας ανάπτυξης για την εργασίας μας.

3.2 Περιγραφή της βιβλιοθήκης iThink

Στην πτυχιακή μας εργασία ορίσαμε δύο στόχους. Ο πρώτος ήταν η δημιουργία μιας πρακτικής, γενικής χρήσης και εύκολα επεκτάσιμης βιβλιοθήκης η οποία θα προσέφερε ευκολία ανάπτυξης τεχνικών κατάστρωσης σχεδίου στους προγραμματιστές. Ο δεύτερος ήταν το πακέτο να είναι κατανοητό στους χρήστες, πρακτικό για χρήση σε ερευνητικό πλαίσιο, καθώς και εκπαιδευτικό ως εργαλείο εκμάθησης τεχνικών κατάστρωσης σχεδίου. Αυτοί οι στόχοι αποδείχθηκαν αντιφατικοί, παρά ταύτα αναπτύξαμε έναν φορμαλισμό βασισμένο στο STRIPS με επαρκή εκφραστικότητα που όμως προσαρμόζεται εύκολα στις ανάγκες του εκάστοτε προγραμματιστή βιντεοπαιχνιδιών.

Τα στοιχεία της iThink που αναπτύξαμε επιτρέπουν την πλήρη περιγραφή ενός προβλήματος κλασικού σχεδιασμού. Τα λογικά λεκτικά μεταφράζονται σε γεγονότα (iThinkFact) και τα λογικά κατηγορήματα μεταφράζονται σε ενέργειες (iThinkAction). Το σύνολο των γεγονότων και των ενεργειών που είναι διαθέσιμα σε ένα πράκτορα κατάστρωσης σχεδίου αποτελούν το πεδίο σχεδιασμού (planning domain). Μια συλλογή γεγονότων αποτελεί την περιγραφή μιας κατάστασης στο χώρο αναζήτησης (iThinkState), συνεπώς μπορεί να χρησιμοποιηθεί για την περιγραφή της τωρινής/αρχικής καταστάσης (initial state) του πράκτορα καθώς και μια κατάσταση στην οποία θέλουμε να οδηγηθούμε μέσω ενός σχεδίου/πλάνου, την κατάσταση στόχου (goal state). Το σύνολο των ενεργειών και των γεγονότων, μαζί με την αρχική κατάσταση και καταστάσεις στόχου, αποτελούν ένα πρόβλημα σχεδιασμού (planning problem). Σκοπός του πράκτορα είναι να αξιοποιήσει τις διαθέσιμες ενέργειες ώστε να εκτελέσει αναζήτηση στο χώρο καταστάσεων και να βρει αλληλουχία ενεργειών – δηλαδή ένα σχέδιο/πλάνο ενεργειών (iThinkPlan) - που θα τον οδηγήσουν από την αρχική κατάσταση στην κατάσταση στόχου.

Η περιγραφή των ενεργειών γίνεται μέσω ειδικά διαμορφωμένων συμβολοσειρών, τις οποίες περιγράφουμε ως action schemas, οι οποίες περιγράφουν τη διαδικασία αυτόματης (ή και καθοδηγούμενης) παραγωγής διαθέσιμων ενεργειών (iThinkActionSchema) βάσει της γνωστής στον πράκτορα γνώσης. Αυτό είναι απαραίτητο ώστε να καθορίζεται η σύνδεση των οντοτήτων/αντικειμένων της προσομοίωσης με τα στοιχεία των ενεργειών. Η συλλογή αυτών των αντικειμένων γίνεται μέσω ενός συστήματος αισθητήρων (iThinkSensorySystem). Οι ενέργειες που παράγονται συλλέγονται σε λίστα διαθέσιμων ενεργειών (iThinkActionManager).

Ως σχεδιαστή (iThinkPlanner) περιγράφουμε τη λειτουργική μονάδα που αξιοποιεί τις διαθέσιμες πληροφορίες ώστε να εκτελέσει τεχνικές αναζήτησης και να παράγει πλάνο ενεργειών προς χρήση από τον πράκτορα. Όλα αυτά τα συστήματα συλλέγονται σε ένα ενιαίο πακέτο που απλοποιεί την υλοποίηση του πράκτορα και τη διαδικασία περιγραφής του προβλήματος σχεδιασμού αλλά και της μετέπειτα κατάστρωσης σχεδίου (iThinkBrain).

3.3 Παράδειγμα χρήσης του iThink

```
public class SimpleFPSAgent : MonoBehaviour
{
  iThinkBrain brain ;
  public string [ ] schemaList  =  {
  "ActionMove-3-Tag~loc-Tag~loc-Tag~dir" ,
  "ActionTurn-2-Tag~dir-Tag~dir" ,
  "ActionShoot-4-Tag~loc-Tag~loc-Tag~dir-Tag~gun" ,
  "ActionStab-2-Tag~loc-Tag~knife" ,
  "ActionPickUp-2-Tag~knife-Tag~loc" ,
  "ActionPickUp-2-Tag~gun-Tag~loc"} ;

public void Awake ( ) // executed when NPC is constructed
{
  brain = new iThinkBrain ( ) ;

  List<String> tags = new List<String>() ;
  tags.Add("dir ") ; tags.Add("loc ") ; tags.Add(" player ") ;
  tags.Add("npc") ; tags.Add( "gun" ) ; tags.Add(" knife ") ;
  brain.sensorySystem.OmniUpdate(this.gameObject , tags);

  brain.ActionManager = new iThinkActionManager ( ) ;
  brain.ActionManager.initActionList(
       this.gameObject ,
       schemaList ,
       brain.getKnownObjects ( ) ,
       brain.getKnownFacts ( )
  );
}

public void Update () // executed at every frame of the game
{
  // . . . code that specifies the initial state and goal

  Brain.planner.forwardSearch ( // invokes planner
       brain.startState,
       brain.goalState,
       brain.ActionManager,
       1); // specifies the search method

  brain.planner.getPlan().debugPrintPlan();

  // . . . code that uses the plan for NPC behavior
  }
}
```

Σχήμα 1: Παράδειγμα χρήσης του IThink σε ένα απλοποιημένο κόσμο

Η επίσημη τεκμηρίωση είναι διαθέσιμη στον ιστοχώρο της βιβλιοθήκης στη διεύθυνση: http://code.google.com/p/ithink-unity3d/

4. Σχετικές δημοσιεύσεις

Υπάρχει τεράστιο εύρος εφαρμογών ποικιλίας μεθόδων κατάστρωσης σχεδίου στη βιβλιογραφία που αφορά μεγάλο εύρος μεγέθους προβλημάτων καθώς και χρονικών ή υπολογιστικών περιορισμών. Κύριο χαρακτηριστικό αυτών είναι η συμβολική περιγραφή των οντοτήτων, πολλές φορές υπό τη μορφή απλών συμβολοσειρών, γεγονός που έρχεται σε αντίθεση με τις ανάγκες ενός πρακτικού περιβάλλοντος. Η προσέγγισή μας αποτελεί επιτυχή και πρωτοπόρα γεφύρωση αυτών των δύο τρόπων ανάπτυξης, συμβατό με το φορμαλισμό STRIPS και προσφέροντας ένα προγραμματιστικό περιβάλλον αναπαράστασης βασισμένο σε κατανοητές έννοιες, όπως η χρήση μεταβλητών (οντοτήτων) και ετικετών (ειδών οντοτήτων).

Το βιντεοπαιχνίδι "F.E.A.R." αποτελεί το πρώτο και χαρακτηριστικό παράδειγμα χρήσης κατάστρωσης σχεδίου στα πλαίσια βιντεοπαιχνιδιού, βασισμένο στη δουλειά του Jeff Orkin [8]. Ο Bjarnolf πρόσθεσε «ανάλυση απειλής» σε σύστημα GOAP με βάση τη μέθοδο "Observe, Orient, Decide, Act" [1]. Ο Long έκανε σύγκριση επιδόσεων μεταξύ υλοποιήσεων συστήματος GOAP και μηχανής πεπερασμένων καταστάσεων [2]. Ο Πεϊκίδης μελέτησε την εφαρμογή συστήματος GOAP σε περιβάλλον βιντεοπαιχνιδιού στρατηγικής [10]. Τέλος, ο Pittman εισήγαγε ιεραρχία εντολών σε σύστημα GOAP με σκοπό την απλοποίηση της διαδικασίας συντήρησης κώδικα [11].

Οι παραπάνω υλοποιήσεις εμπνέονται από τις αρχιτεκτονικές επιλογές που πρότεινε ο Orkin ούτως ώστε να επιτυγχάνεται ικανοποιητική επίδοση σε περιβάλλον πραγματικού χρόνου. Αυτές οδηγούν σε χρήση απλοποιημένου φορμαλισμού βασισμένο στον STRIPS, περιορίζοντας σημαντικά την εκφραστικότητα της μεθόδου. Επιπλέον, καθεμία από τις παραπάνω προσεγγίσεις εστιάζει σε οργάνωση και λεπτομέρειες υλοποίησης συγκεκριμένων βιντεοπαιχνιδιών, περιορίζοντας σημαντικά τη μεταφερσιμότητα του κώδικα. Αντιθέτως, η βιβλιοθήκη iThink είναι γενικής χρήσης και εκφραστικότητας, καθώς και εύκολα μεταφέρσιμη σε εναλλακτικά περιβάλλοντα ανάπτυξης. Αξιοποιώντας το λογικό υπόβαθρο του STRIPS, η βιβλιοθήκη προσφέρει μια ισχυρή πλατφόρμα αναπαράστασης σχεδίων σε περιβάλλον Unity3D.

5. Μελλοντικές επεκτάσεις της iThink

Τελικός σκοπός μας είναι η ανάπτυξη της βιβλιοθήκης iThink σε ένα πλήρες και αποδοτικό πακέτο δημιουργίας Τεχνητής Νοημοσύνης κατάστρωση σχεδίου, συνοδευόμενη από εργαλεία που απλοποιούν τη διαδικασία. Αρχικά, θα γίνει επέκταση με επιπλέον μεθόδους αναζήτησης και ευρεστικές συναρτήσεις για πληθώρα εφαρμογών σε περιβάλλοντα παιχνιδιών. Σημαντικό όφελος απόδοσης και ανταπόκρισης θα επέλθει έπειτα από εισαγωγή πολυνηματικού προγραμματισμού. Ταυτόχρονα, θα εισαχθούν και στοιχεία αρχιτεκτονικής που θα υποστηρίζουν την ταυτόχρονη λειτουργία και συνεργασία πολλαπλών πρακτόρων. Τέλος, έχει ήδη δρομολογηθεί η δημιουργία εργαλείων που θα διαχειρίζονται και θα απλοποιούν τη διαδικασία περιγραφής/δημιουργίας των γεγονότων και των ενεργειών του πεδίου σχεδιασμού μέσω του γραφικού περιβάλλοντος της Unity3D.

6. Συμπεράσματα

Η κλασική κατάστρωση σχεδίου αποτελεί χρήσιμη τακτική για επεκτάσιμη δημιουργία αναδυόμενης συμπεριφοράς σε πράκτορες τεχνητής νοημοσύνης βιντεοπαιχνιδιών. Η προσέγγισή μας προσφέρει έναν πρωτοποριακό τρόπο εφαρμογής ακαδημαϊκών τεχνικών στο πρακτικό περιβάλλον των σύγχρονων βιντεοπαιχνιδιών. Η βιβλιοθήκη iThink αποτελεί έναν κατανοητό, εύκολα μεταφέρσιμο, εύκολα επεκτάσιμο και κυρίως γενικευμένο τρόπο περιγραφής και επίλυσης προβλημάτων κατάστρωσης σχεδίου. Εκτός από χρήσιμο εργαλείο, αποτελεί και χρήσιμο εκπαιδευτικό μέσο, προσφέροντας ένα διαδραστικό και διασκεδαστικό περιβάλλον πειραματισμού εκμάθησης και πειραματισμού με την Τεχνητή Νοημοσύνη, εκπληρώνοντας έτσι και τους δύο στόχους της πτυχιακής εργασίας.

Αναφορές

1. Bjarnolf, P.: Threat analysis using goal-oriented action planning. B.Sc. Thesis, University of Skovde, School of Humanities and Informatics (2008)

2. Long, E.: Enhanced NPC Behaviour using Goal Oriented Action Planning. Master’s

thesis, University of Abertay Dundee, School of Computing and Advanced Technologies, Division of Software Engineering (2007)

3. Fikes, R.E., Nilsson, N.J.: STRIPS: A new approach to the application of theorem proving to problem solving. Artificial Intelligence 2, 189–208 (1971)

4. Funge, J.D.: Artificial Intelligence For Computer Games: An Introduction. A. K. Peters, Ltd., MA, USA (2004)

5. Mcdermott, D., Ghallab, M., Howe, A., Knoblock, C., Ram, A., Veloso, M., Weld, D., Wilkins, D.: PDDL - the planning domain definition language. Tech. rep., CVC TR-98-003/DCS TR-1165, Yale Center for Comp. Vision and Control (1998)

6. Millington, I., Funge, J.: Artificial Intelligence for Games, Second Edition. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edn. (2009)

7. Nareyek, A.: Artificial intelligence in computer games - State of the art and future directions. ACM Queue 10(1), 58–65 (2004)

8. Orkin, J.: Three states and a plan: The AI of F.E.A.R. In: Proceedings of the Game Developer's Conference (GDC) (2006)

9. Orkin, J.: Agent architecture considerations for Real-Time planning in games. In: Artificial Intelligence & Interactive Digital Entertainment (AIIDE) (2005)

10. Peikidis, P.: Demonstrating the use of planning in a video game. B.Sc. Thesis, University of Sheffield, CITY Liberal Studies, Dept. of Computer Science (2010)

11. Pittman, D.L.: Practical Development of Goal-Oriented Action Planning AI. Master's thesis, University of Nebraska-Lincoln (2007)

12. Russell, S.J., Norvig, P.: Artificial Intelligence: A Modern Approach (2nd Edition). Prentice Hall (2002)

Ολοκληρωμένο Σύστημα Μετάφρασης Ερωτημάτων SQL σε Φυσική Γλώσσα

Παναγιώτης Βαγενάς
p.vagenas@di.uoa.gr

Αλέξανδρος Ζερβάκης
a.zervakis@di.uoa.gr

Ανδρέας Κόκκαλης
a.kokkalis@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Σε αυτή την εργασία παρουσιάζουμε ένα ολοκληρωμένο σύστημα για τη μετάφραση ερωτημάτων SQL σε φυσική γλώσσα. Αναπαριστούμε το ερώτημα ως κατευθυνόμενο γράφο και επισημαίνουμε τις ακμές του με ετικέτες, χρησιμοποιώντας έναν επεκτάσιμο μηχανισμό προτύπων. Το σύστημά μας χρησιμοποιεί διαφορετικές τεχνικές διάσχισης του γράφου για να τον προσπελάσει αποδοτικά και να συνθέσει το κείμενο φυσικής γλώσσας. Όλη η λειτουργικότητα, που αφορά τη μετάφραση αλλά και τον προσδιορισμό των προτύπων, παρέχεται μέσω μίας χρηστικής, βασισμένης στο δίκτυο, διεπαφής.

Λέξεις-Κλειδιά: Διεπαφές βάσεων δεδομένων σε φυσική γλώσσα, Επεξεργασία φυσικής γλώσσας, Ερώτημα, Βάση δεδομένων

Μέρος της εργασίας που παρουσιάζεται έχει δημοσιευτεί στο ACM SIGMOD 2012 ως:
Logos: A System for Translating Queries into Narratives.
(A. Kokkalis, P. Vagenas, A. Zervakis, A. Simitsis, G. Koutrika, Y. Ioannidis)

Επιβλέπων: Ιωάννης Ιωαννίδης, Καθηγητής
Με τη συμβολή της Γεωργίας Κούτρικα και του Άλκη Σιμιτσή

1. Εισαγωγή

Η χρήση φυσικής γλώσσας στο χώρο των βάσεων δεδομένων δεν είναι νέα ιδέα. Διάφορες προσεγγίσεις, εμπνευσμένες από τη φυσική γλώσσα ή βασισμένες σε αυτή, έχουν γίνει αντικείμενο μελέτης τις τελευταίες δεκαετίες. Ως χαρακτηριστικά παραδείγματα μπορούμε να αναφέρουμε την αναζήτηση με λέξεις-κλειδιά [1] ή τα ερωτήματα précis [2][3], καθώς και τη μετάφραση φυσικής γλώσσας σε εκφράσεις δομημένης μορφής [4][5]. Ενώ, όμως, η σύνθεση ερωτημάτων από κείμενο έχει μελετηθεί επανειλημμένα στο παρελθόν και συστήματα που την επιτρέπουν έχουν αναπτυχθεί, η αντίθετη κατεύθυνση, δηλαδή αυτή που θεωρεί την εξαγωγή φυσικής γλώσσας από δομημένα ερωτήματα, έχει μόλις πρόσφατα εξερευνηθεί ως περιοχή [6].

Η μετάφραση ερωτημάτων σε φυσική γλώσσα συνιστά μία πηγή πλούσια τόσο σε ερευνητικό ενδιαφέρον όσο και σε εφαρμογές. Καταρχάς, πριν ένα ερώτημα σταλεί για εκτέλεση, θα ήταν ωφέλιμο για το χρήστη να μπορούσε να το δει εκφρασμένο σε μία πιο προσφιλή μορφή, αυτή της φυσικής γλώσσας, ως μία επιβεβαίωση για το ότι έχει το επιθυμητό νόημα. Όσο δε πιο πολύπλοκη είναι η διατύπωσή του, τόσο πιο χρήσιμη παρουσιάζεται μία περιγραφή σε φυσική γλώσσα. Μία διαπιστωμένα σωστή μετάφραση θα μπορούσε να αποτελέσει μία πολύ ισχυρή επαλήθευση για την εγκυρότητα του ερωτήματος. Επιπρόσθετα, δεδομένης μίας τέτοιας μετάφρασης, θα ήταν πιο εύκολο για το χρήστη να καταλάβει τα κομμάτια του ερωτήματος που ευθύνονται για μία τυχόν κενή απάντηση ή, αντίστοιχα, για μία ιδιαίτερα μεγάλη. Πέρα, όμως, από τους κατεξοχήν χρήστες βάσεων δεδομένων, σημαντικά μπορούν να ωφεληθούν και όσοι δεν έχουν μεγάλη εμπειρία, καθώς η εξήγηση ερωτημάτων σε φυσική γλώσσα μπορεί να βελτιώσει αισθητά την καμπύλη μάθησης. Προκύπτουν έτσι ενδιαφέρουσες ιδέες για εφαρμογές στο χώρο της εκπαιδευτικής διαδικασίας. Παράλληλα, οι τεχνικές που παρουσιάζονται σε αυτή την εργασία μπορούν να επεκταθούν ώστε, επιπλέον των ερωτημάτων SQL (εντολή SELECT), να καλύπτουν και άλλες εντολές SQL όπως εισαγωγές (INSERT), διαγραφές (DELETE) και ενημερώσεις (UPDATE). Σε αυτή την περίπτωση, το Σύστημα Διαχείρισης Βάσης Δεδομένων θα παρείχε μία πιο ολοκληρωμένη διεπαφή φυσικής γλώσσας στο χρήστη. Έτσι, όταν, για παράδειγμα, κάποιος υποβάλλει μία φόρμα που ανάγεται στην εκτέλεση κάποιων εντολών SQL, θα ήταν χρήσιμο να υπάρχει η δυνατότητα παροχής μιας περιγραφής σε φυσική γλώσσα για τις εν λόγω εντολές, ειδικά όταν μία τέτοια φόρμα είναι μεγάλη και ο χρήσης δεν είναι εύκολο να γνωρίζει τον τρόπο με τον οποίο συνδέονται σημασιολογικά τα διάφορα εμπλεκόμενα πεδία. Τέλος, πέραν της SQL, η μεθοδολογία που παρουσιάζεται μπορεί να τροποποιηθεί ώστε να διαχειρίζεται ερωτήματα RDF σε SPARQL ή RQL, προγράμματα Datalog κ.α. Οι διάφορες πιθανές εφαρμογές, λοιπόν, είναι πολυάριθμες και σε κάθε περίπτωση ενδιαφέρουσες.

Η εξαγωγή μίας σωστής αλλά και φυσικής μετάφρασης για ένα ερώτημα δεν είναι τετριμμένη διαδικασία. Ας θεωρήσουμε, για παράδειγμα, το παρακάτω ερώτημα (Ερώτημα 1):

```
select a.name, m.title from MOVIES m, CAST c, ACTOR a,
            DIRECTED r, DIRECTOR d, GENRE g
where m.id = c.mid and c.aid = a.id and m.id = r.mid
and r.did = d.id
     and m.id = g.mid and d.name = `Coppola' and g.genre = `action'
```

Κάποιος θα μπορούσε να ισχυριστεί ότι μπορούμε να παράγουμε τη μετάφραση απλώς ακολουθώντας τις σχέσεις πρωτεύοντος-ξένου κλειδιού, ωστόσο λόγω του αυξημένου αριθμού σχέσεων ζεύξης, η απλοϊκή αυτή προσέγγιση θα δημιουργούσε μία βερμπαλιστική και αφύσικη μετάφραση. Ωστόσο, με χρήση κατάλληλων προτύπων, το σύστημα μας μπορεί να προσφέρει το εξής αποτέλεσμα:

```
"Find the action movies directed by Coppola and the actors that
play in these movies"
```

Τα βασικά εργαλεία που χρησιμοποιούμε για να αντιμετωπίσουμε αυτές τις δυσκολίες είναι, αφενός, ένας μηχανισμός προτύπων που επιτρέπει την κατάλληλη αναπαράσταση πολύπλοκων συντακτικών μοτίβων και, αφετέρου, μία τεχνική διάσχισης που προσπελαύνει το γράφο ερωτήματος για την εξαγωγή της μετάφρασης.

2. Θεωρητικό υπόβαθρο

Σε αυτή την ενότητα παρουσιάζουμε μία σύνοψη της προσέγγισής μας. Το θεωρητικό υπόβαθρο της εργασίας καλύπτεται σε μεγάλο βαθμό στο [7].

2.1 Γράφος σχήματος βάσης δεδομένων και γράφος ερωτήματος

Τόσο το σχήμα της βάσης δεδομένων όσο και τα ερωτήματα που υποβάλλονται, αναπαρίστανται ως κατευθυνόμενοι γράφοι. Ο γράφος σχήματος βάσης δεδομένων είναι ένας κατευθυνόμενος γράφος που αποτυπώνει τους κύριους ρόλους των σχέσεων και των πεδίων τους στη βάση. Περιλαμβάνει ένα κόμβο για κάθε σχέση ή πεδίο, ενώ οι ακμές του είναι ακμές μέλους (από τα πεδία προς στις αντίστοιχες σχέσεις, εκφράζοντας προβολές), επιλογής (από τις σχέσεις προς τα πεδία, εκφράζοντας επιλογές) ή κατηγορήματος (από πεδία σε πεδία, εκφράζοντας ζεύξεις). Ένας απλός γράφος ερωτήματος εκφράζει τη σημασιολογία ενός ερωτήματος επιλογής-προβολής-ζεύξης και ουσιαστικά συνιστά μία επέκταση του γράφου σχήματος βάσης δεδομένων. Για την ακρίβεια, έχει τους ίδιους τύπους ακμών με αυτόν, ενώ όσον αφορά στους κόμβους, περιλαμβάνει ένα κόμβο για κάθε εμφάνιση σχέσης, πεδίου ή τιμής στο ερώτημα. Για πιο πολύπλοκα είδη ερωτημάτων, ωστόσο, ο γράφος ερωτήματος μπορεί να περιέχει επιπρόσθετους τύπους κόμβων και ακμών, έτσι ώστε να αναπαριστά συναθροιστικές συναρτήσεις, ειδικούς τελεστές (για παράδειγμα DISTINCT), καθώς και τα τμήματα GROUP BY, HAVING και ORDER BY.

2.2 Ετικέτες και πρότυπα

Αποδίδουμε σημασιολογικά χαρακτηριστικά στα διάφορα μέρη του ερωτήματος, επισημαίνοντας τους κόμβους και τις ακμές του γράφου ερωτήματος με ετικέτες και χρησιμοποιώντας έναν επεκτάσιμο μηχανισμό προτύπων. Αρχικά, στον κάθε κόμβο του γράφου ερωτήματος αντιστοιχίζεται μια ετικέτα που εκφράζει τη σημασία του κόμβου σε φυσική γλώσσα. Έτσι, για παράδειγμα για τη σχέση MOVIES, μπορούμε να ορίσουμε την ετικέτα 'movies'. Παρόμοια, κάθε ακμή ή μονοπάτι που συνδέει δύο κόμβους μπορεί να επισημανθεί με μία ετικέτα που εκφράζει τη σημασία της σχέσης μεταξύ του κόμβου προέλευσης και του κόμβου προορισμού. Για παράδειγμα, η ακμή μέλους που συνδέει το πεδίο title της σχέσης MOVIES στην ίδια τη σχέση θα μπορούσε να έχει την ετικέτα 'of'. Οι ετικέτες αυτές αποθηκεύονται στο γράφο σχήματος βάσης, τόσο για τους κόμβους όσο και για τις ακμές, και στη συνέχεια κληρονομούνται από τον εκάστοτε γράφο ερωτήματος.

Οι αλγόριθμοι μετάφρασης που χρησιμοποιούμε διασχίζουν το γράφο ερωτήματος και κατασκευάζουν φράσεις συνθέτοντας κατάλληλα τις ετικέτες που βρίσκουν στη διαδρομή. Μία ετικέτα $\mathrm{l}\left((\mathrm{v},\mathrm{u})\right)$ αντιστοιχίζεται σε μία ακμή (v,u) ή, γενικότερα, σε ένα μονοπάτι από το v στο u. Έτσι ορίζεται ένα πρότυπο, που χρησιμοποιείται για τη μετάφραση της σχέσης μεταξύ του v και του u, κατά την παραγωγή του κειμένου φυσικής γλώσσας.

Οι ετικέτες προτύπων κατασκευάζονται μη αυτοματοποιημένα, από ανθρώπους, και μπορούν να παράγουν υψηλής ποιότητας κείμενο. Ως ένα σύντομο παράδειγμα, η ετικέτα για το μονοπάτι επιλογής από τη σχέση MOVIES στην τιμή του πεδίου year, στη γλώσσα προτύπων [7] , θα μπορούσε να είναι:

```
l(MOVIES) + " released in " + MOVIES.year.<val>
```

2.3 Μετάφραση ερωτήματος με διάσχιση γράφου

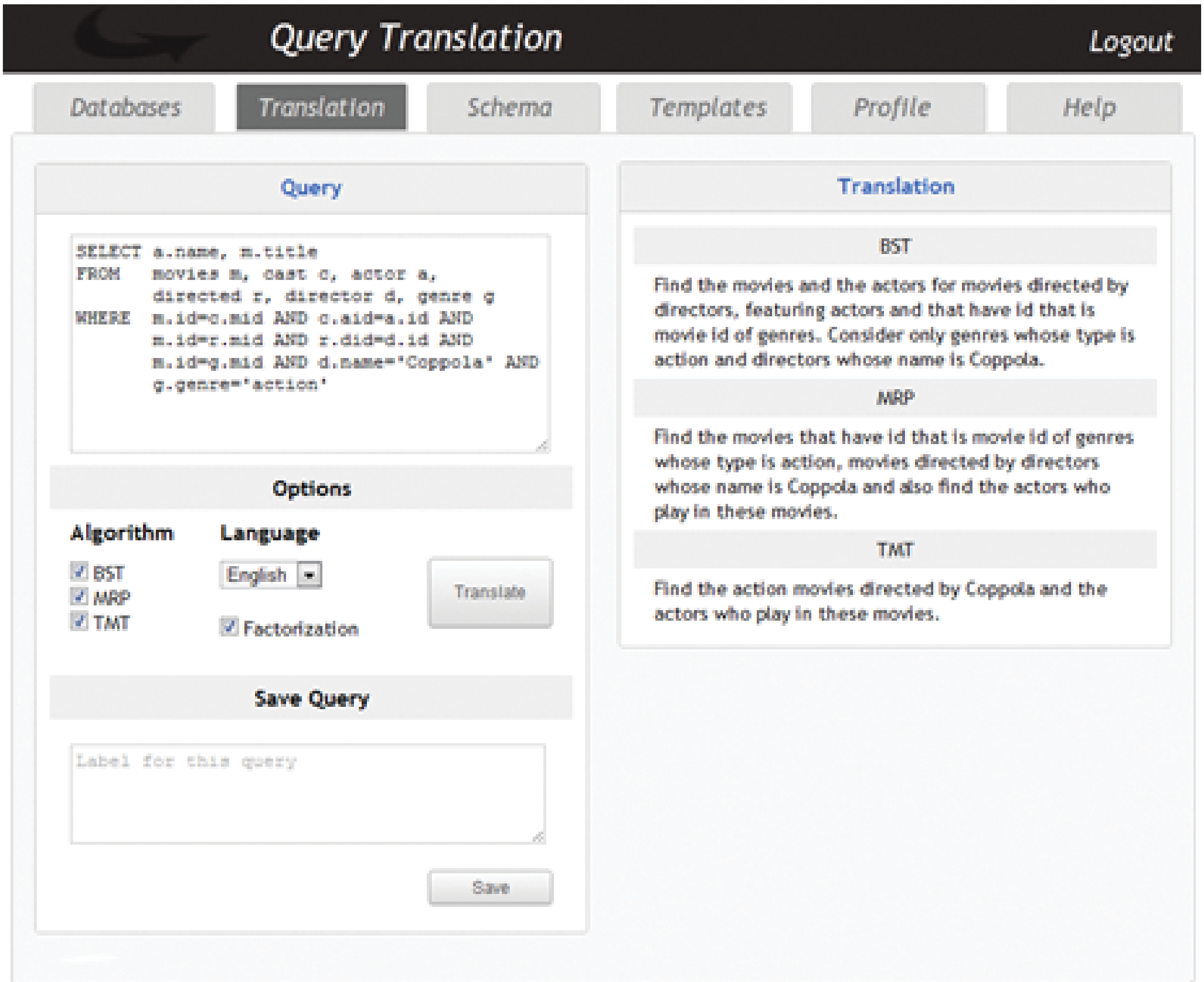
Η διάσχιση του γράφου ερωτήματος μπορεί να πραγματοποιηθεί με μία πληθώρα αλγορίθμων. Στα πλαίσια αυτής της εργασίας, μελετάμε τρεις διαφορετικές τεχνικές διάσχισης για την εξαγωγή της μετάφρασης. Θεμελιώδη σημασία έχει σε κάθε περίπτωση ο προσδιορισμός του κόμβου – σημείο αφετηρίας της διάσχισης, τον οποίον ονομάζουμε υποκείμενο του ερωτήματος. Αναλόγως με την επιλογή του υποκειμένου, η κατεύθυνση της διάσχισης και, τελικά, η παραγόμενη μετάφραση μπορεί να είναι πολύ διαφορετική.

Η πρώτη τεχνική που θεωρούμε, ο αλγόριθμος BST, κατασκευάζει διαφορετικές προτάσεις για το κάθε τμήμα του ερωτήματος. Έτσι, αρχικά μεταφράζει τις ακμές μέλους, στη συνέχεια συνδέει όλες τις σχέσεις στο υποκείμενο μέσω των ζεύξεων του ερωτήματος και, τέλος, καταναλώνει τα μονοπάτια που συνδέουν σχέσεις με κόμβους τιμής που έχουν προσδιοριστεί για πεδία των εν λόγω σχέσεων. Η διάσχιση σε αυτή την περίπτωση πραγματοποιείται πρώτα κατά βάθος, ξεκινώντας από το υποκείμενο του ερωτήματος και προς όλες τις εμπλεκόμενες σχέσεις, μέσω των ζεύξεων.

Σύμφωνα με τη δεύτερη τεχνική, τον αλγόριθμο MRP, η μετάφραση δημιουργείται με έναν ολιστικό τρόπο, όπου πληροφορίες από όλα τα τμήματα του ερωτήματος υπεισέρχονται στη μετάφραση καθώς διασχίζουμε το γράφο. Η βασική ιδέα είναι ότι, ενώ στον BST όλα τα στοιχεία του ερωτήματος αναφέρονται πάντα στο ίδιο υποκείμενο, εδώ χρησιμοποιούμε επιπρόσθετα σημεία αναφοράς με σκοπό να αποφύγουμε μακροσκελείς, πιθανώς μη φυσικές προτάσεις. Κατά αυτή την έννοια, διαιρούμε τη μετάφραση σημασιολογικά σε πολλαπλά σημεία.

Τέλος, η τρίτη μέθοδος, ο αλγόριθμος TMT, επιτρέπει τη χρήση πλουσιότερων προκαθορισμένων προτύπων, με σκοπό την εξαγωγή μεταφράσεων υψηλής ποιότητας. Πιο συγκεκριμένα, σε αυτή την περίπτωση μπορούμε να αναθέσουμε ετικέτες σε ολόκληρους γράφους, και επομένως, να αναπαραστήσουμε σημασιολογικά υπογράφους του γράφου ερωτήματος.

Παραδείγματα εκτέλεσης του Ερωτήματος 1 και με τους τρεις αλγορίθμους φαίνονται στο Σχήμα 1.



Σχήμα 1: Φόρμα μετάφρασης

3. Διεπαφή του συστήματος

Κάθε χρήστης χρησιμοποιεί τη διεπαφή για να συνδεθεί σε μια απομακρυσμένη βάση και να μεταφράσει ερωτήματα σε αυτήν. Τα δικαιώματα που έχει στη βάση κατηγοριοποιούν τον χρήστη είτε ως κανονικό χρήστη που ενδιαφέρεται για τη μετάφραση ερωτημάτων και τις φόρμες του προφίλ του, είτε ως διαχειριστή που μπορεί να παραμετροποιεί τις μεταφράσεις μέσω μηχανισμών προτύπων.

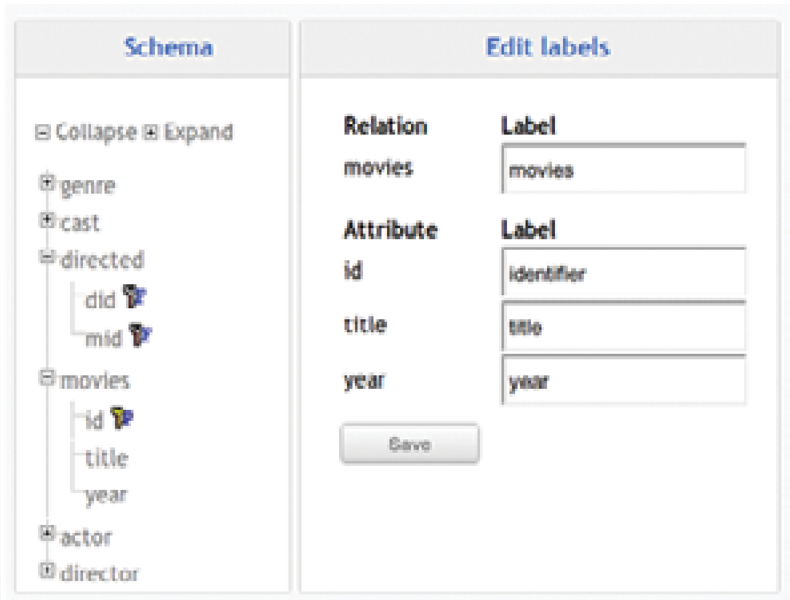
3.1 Διεπαφή χρήστη

Τυπικά, ένας κανονικός χρήστης ενδιαφέρεται για τη μετάφραση ερωτημάτων και τις φόρμες του προφίλ του.

Μετάφραση Ερωτημάτων:

Η μετάφραση ερωτημάτων καθίσταται δυνατή μέσω μίας απλής φόρμας, στην οποία ο χρήστης εισάγει το ερώτημα, για να του προβληθεί στη συνέχεια το παραχθέν κείμενο σε φυσική γλώσσα (Σχήμα 1). Η μετάφραση μπορεί να προσαρμοστεί θέτοντας παραμέτρους όπως η επιθυμητή γλώσσα μετάφρασης, ο αλγόριθμος μετάφρασης, καθώς και η παραγοντοποίηση. Η τελευταία επιλογή ελέγχει το αν κάποιες ειδικές λέξεις, όπως άρθρα ή αναφορικές αντωνυμίες, επαναλαμβάνονται ή όχι, προσδιορίζοντας με αυτό τον τρόπο το πόσο συμπαγής είναι η παραγόμενη μετάφραση. Παράλληλα, υπάρχει η δυνατότητα επιλογής παραπάνω του ενός αλγορίθμου μετάφρασης. Σε αυτή την περίπτωση, προβάλλονται στο χρήστη όλες οι σχετικές μεταφράσεις. Στη συνέχεια, ο χρήστης μπορεί είτε να εισάγει ένα νέο ερώτημα, είτε να τροποποιήσει τις παραμέτρους του υπάρχοντος. Τέλος, ένα ερώτημα μπορεί να αποθηκευτεί για μετέπειτα χρήση, συνοδευόμενο από μία κατάλληλη ετικέτα για εύκολη αναφορά.

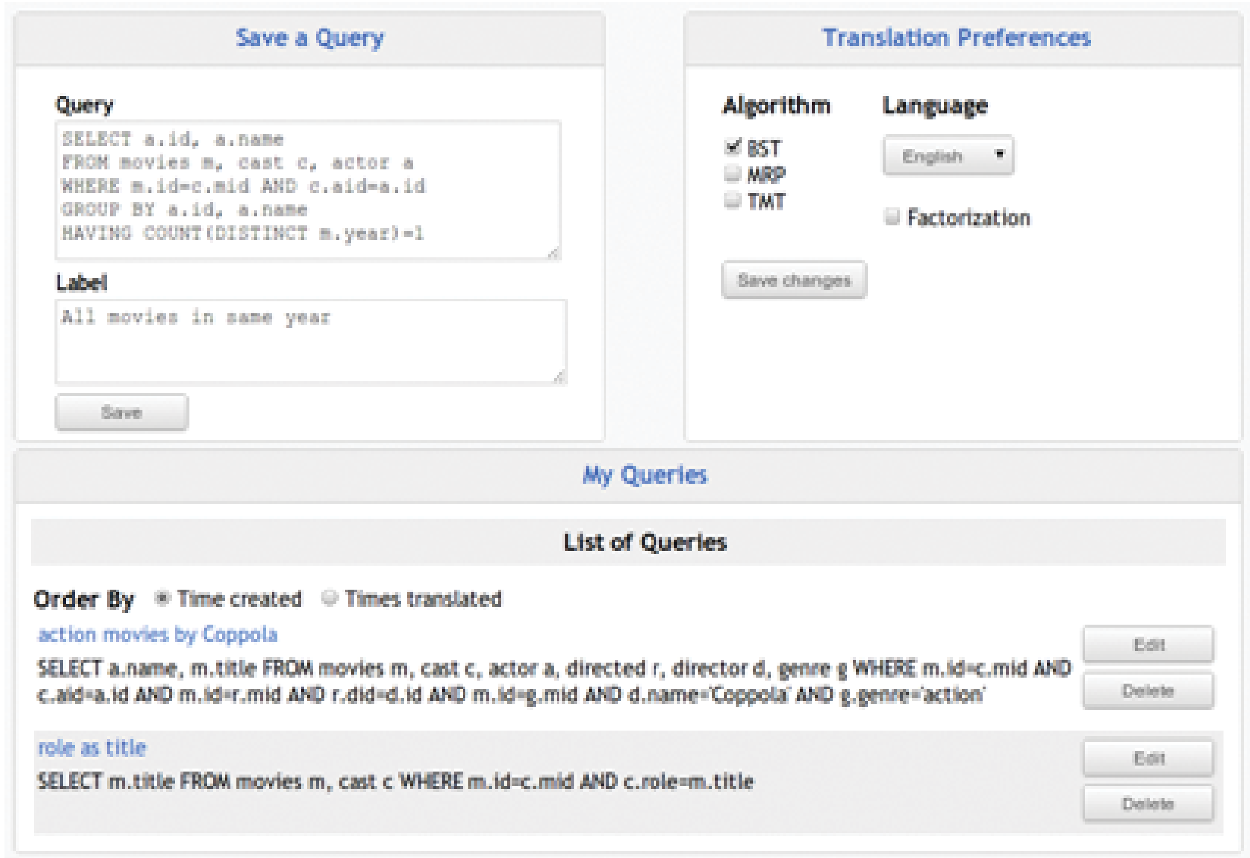
Για να εισάγει κάποιος ένα ερώτημα προς μετάφραση, θα πρέπει να γνωρίζει το σχήμα της βάσης δεδομένων. Το σύστημά μας παρέχει ένα δενδροειδή εξερευνητή σχήματος (Σχήμα 2) που επιτρέπει τη γρήγορη επισκόπηση του σχήματος. Καταρχάς, στο χρήστη παρουσιάζονται οι κόμβοι του άνω επιπέδου του δένδρου, δηλαδή οι σχέσεις της βάσης δεδομένων. Ο χρήστης μπορεί μέσω του εξερευνητή να αναπτύξει ένα κόμβο σχέσης ώστε να προβληθούν οι αντίστοιχοι κόμβοι πεδίων. Η επιλογή ενός συγκεκριμένου κόμβου προβάλλει την ετικέτα του.



Σχήμα 2: Δενδροειδές σχήμα βάσης και επεξεργασία ετικετών

Προφίλ Χρήστη

Η διεπαφή του συστήματος μας είναι διαδραστική και επιτρέπει παραμετροποιημένες μεταφράσεις. Προκειμένου να προσφέρει μία ολοκληρωμένη εμπειρία στο χρήστη, το σύστημά μας αποθηκεύει δεδομένα του προφίλ του. Τέτοια δεδομένα μπορεί να είναι οι προτιμήσεις μετάφρασης, όπως η γλώσσα, ο αλγόριθμος μετάφρασης και η επιλογή παραγοντοποίησης, καθώς και η λίστα των αποθηκευμένων ερωτημάτων. Μέσω της σελίδας προφίλ (Σχήμα 3), ο χρήστης μπορεί να τροποποιήσει τις προτιμήσεις μετάφρασής του ή να χρησιμοποιήσει τη λίστα ερωτημάτων για να μεταφράσει ή να επεξεργαστεί την ετικέτα κάποιου ερωτήματος.

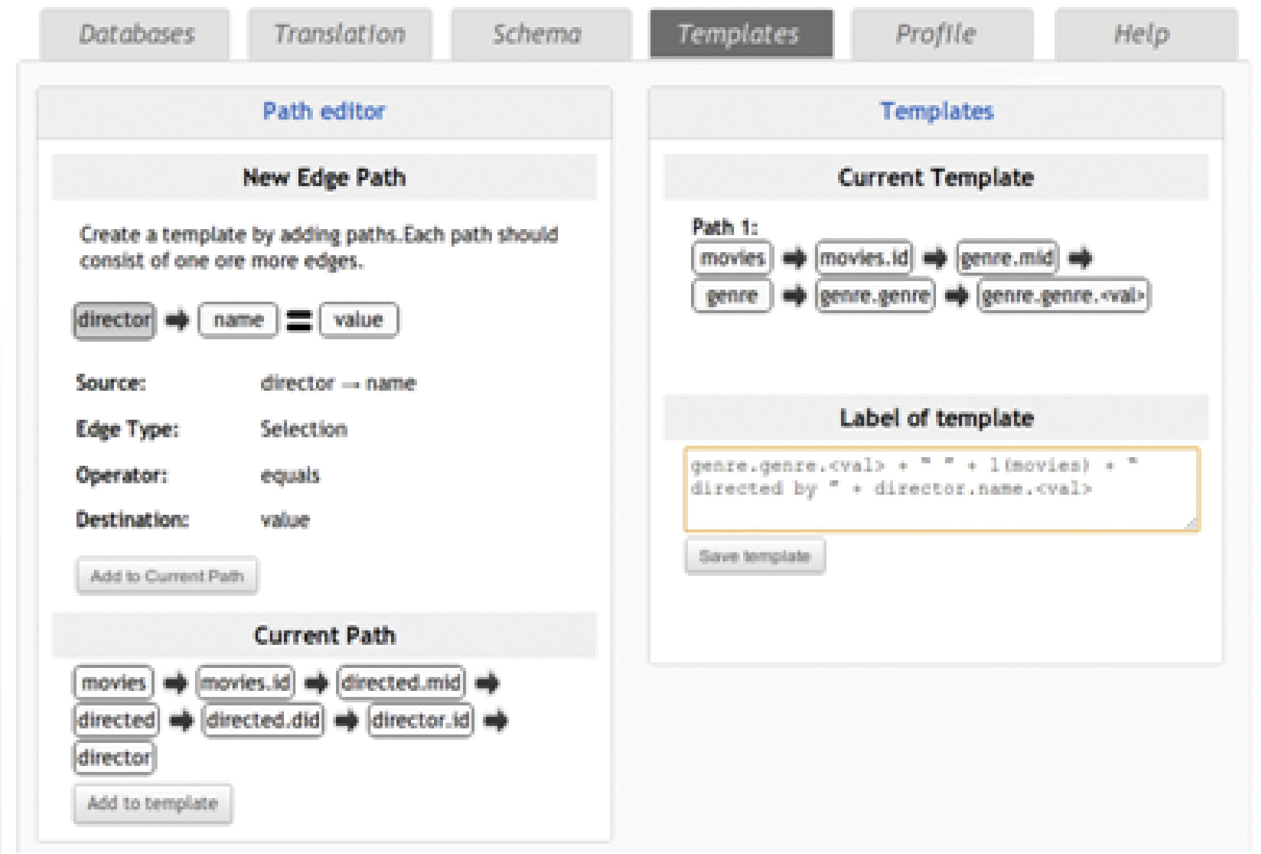


Σχήμα 3: Επιλογές χρήστη

3.2 Διεπαφή διαχειριστή

Προκειμένου να παράγει μεταφράσεις για ερωτήματα επί μίας βάσης δεδομένων, το σύστημά μας συνδέεται με το απομακρυσμένο Σύστημα Διαχείρισης Βάσης Δεδομένων και εξάγει τα απαραίτητα μεταδεδομένα από το αντίστοιχο σχήμα. Έχοντας συνδεθεί με μία συγκεκριμένη βάση δεδομένων, ο ορισμός προτύπων πραγματοποιείται διαμέσου της ενσωματωμένης κονσόλας διαχειριστή που παρέχει το σύστημα. Η υλοποίησή μας προσφέρει προεπιλεγμένες ετικέτες (λόγου χάρη, την ετικέτα "of" για ακμές μέλους), ωστόσο, όσο περισσότερες εξειδικευμένες ετικέτες παρέχει στο σύστημα ο σχεδιαστής, τόσο βελτιώνεται η ποιότητα των αποτελεσμάτων της μετάφρασης.

Η σύνθεση των γράφων προτύπων πραγματοποιείται βήμα-προς-βήμα και υποβοηθείται από το σύστημα. Το Σχήμα 4 δείχνει τη διαδικασία κατασκευής του προτύπου που αντιστοιχεί στο κομμάτι του Ερωτήματος 1 που μεταφράζεται ως "action movies directed by Coppola". Ο σχετικός γράφος προτύπου αποτελείται από δύο μονοπάτια. Το πρώτο ξεκινά από τον κόμβο σχέσης MOVIES, μεταβαίνει στο DIRECTOR μέσω του DIRECTED χρησιμοποιώντας κόμβους πεδίου με σχέση πρωτεύοντος-ξένου κλειδιού, στη συνέχεια μεταβαίνει στον κόμβο πεδίου name της σχέσης DIRECTOR και καταλήγει στον αντίστοιχο κόμβο τιμής. Το δεύτερο μονοπάτι έχει αφετηρία τον κόμβο σχέσης MOVIES, στη συνέχεια περνά στο GENRE χρησιμοποιώντας κόμβους πεδίου με σχέση πρωτεύοντος-ξένου κλειδιού, έπειτα στο πεδίο genre της σχέσης GENRE και, τέλος, στον αντίστοιχο κόμβο τιμής.

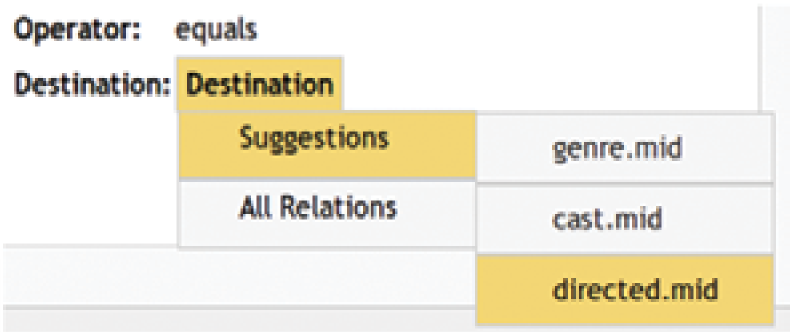


Σχήμα 4: Κατασκευή προτύπου

Για να δημιουργήσουμε το πρότυπο, εργαζόμαστε ως εξής. Καταρχάς, χρησιμοποιούμε τον επεξεργαστή μονοπατιού για να προσδιορίσουμε ένα μονοπάτι υποδεικνύοντας διαδοχικές ακμές.

Ας μελετήσουμε το πρώτο μονοπάτι. Χρησιμοποιώντας το δενδροειδή επιλογέα προέλευσης, καθορίζουμε τον κόμβο σχέσης προέλευσης (MOVIES) και το σχετικό κόμβο πεδίου του (id). Στη συνέχεια, ο τύπος της τρέχουσας ενέργειας ορίζεται ως ζεύξη. Έχοντας επιλέξει τον τελεστή ισότητας ως τελεστή κατηγορήματος, χρησιμοποιούμε το δενδροειδή επιλογέα προορισμού για να προσδιορίσουμε πώς ολοκληρώνεται η ζεύξη.

Το σύστημά μας εκμεταλλεύεται σχέσεις πρωτεύοντος-ξένου κλειδιού για να προσφέρει προτάσεις για τον κόμβο προορισμού (Σχήμα 5), διευκολύνοντας και επιταχύνοντας με αυτόν τον τρόπο ολόκληρη τη διαδικασία. Χρησιμοποιώντας αυτό το χαρακτηριστικό, μπορούμε με ευκολία να επιλέξουμε τη σχέση DIRECTED και το αντίστοιχο πεδίο mid και με αυτό τον τρόπο να ολοκληρώσουμε τη ζεύξη. Στη συνέχεια, η ζεύξη αυτή προστίθεται στο τρέχον μονοπάτι. Ακολούθως, επαναλαμβάνουμε τη διαδικασία δύο φορές για να κατασκευάσουμε ολόκληρο το μονοπάτι: αρχικά προσθέτουμε τη ζεύξη από το DIRECTED στο DIRECTORS και μετά, την επιλογή από το DIRECTORS στην τιμή του πεδίου name τους. Αφού ολοκληρωθεί το μονοπάτι, μπορεί να προστεθεί στο τρέχον πρότυπο.



Σχήμα 5: Προτάσεις προορισμού

Στη συνέχεια, το σύστημά μας συσσωρεύει πολλαπλά μονοπάτια ζεύξης. Για να προσθέσουμε το δεύτερο μονοπάτι του προτύπου, αρχικά προσδιορίζουμε τον κόμβο που μοιράζεται με το πρώτο μονοπάτι και στη συνέχεια επαναλαμβάνουμε τα ίδια βήματα.

Όταν ο γράφος προτύπου ολοκληρωθεί, μία κατάλληλη έκφραση της γλώσσας προτύπων πρέπει να ανατεθεί σε αυτόν πριν αποθηκευτεί. Η έκφραση σε αυτή την περίπτωση θα είναι:

```
GENRE.genre.<val> + " " + l(MOVIES) + "directed by" + DIRECTOR.
name.<val>
```

Τέλος, ο σχεδιαστής μπορεί να αναθέσει αυτή την ετικέτα στον παραχθέντα γράφο και να αποθηκεύσει το πρότυπο.

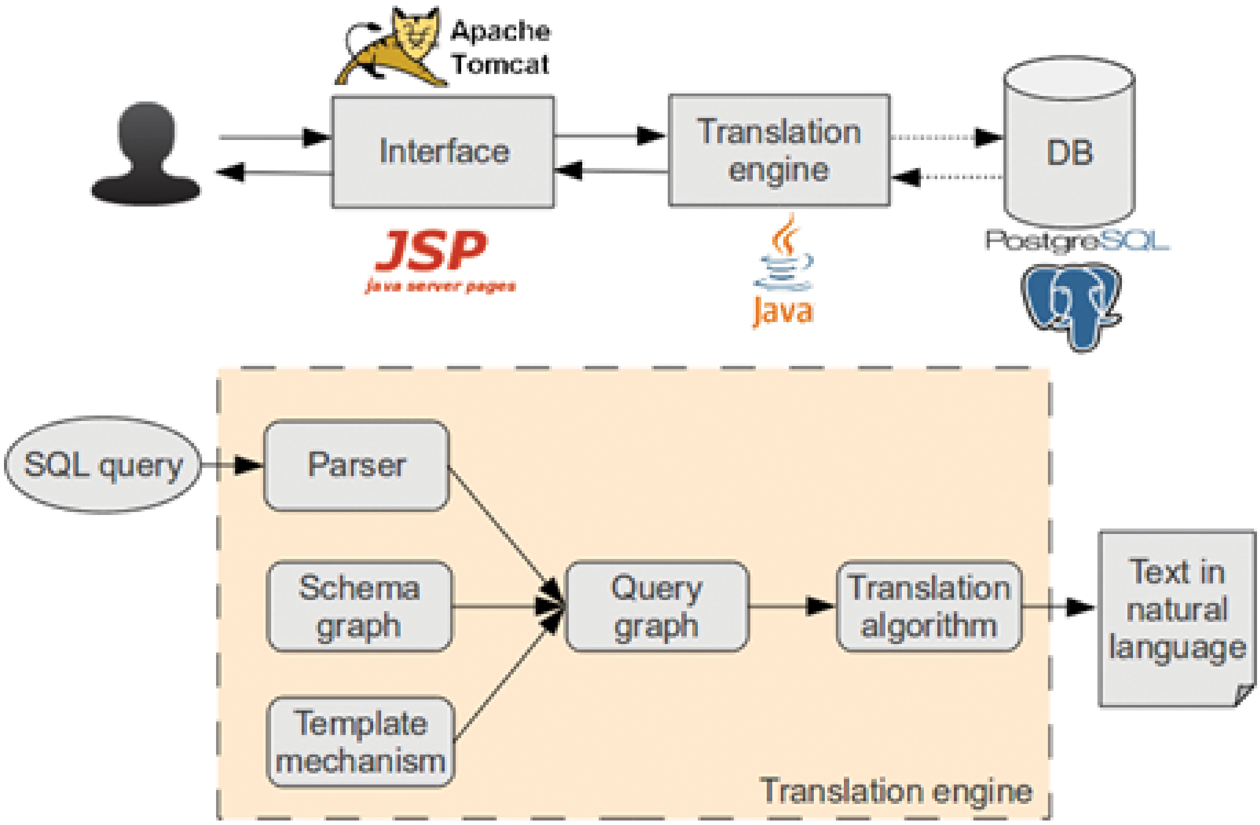
3.3 Πολυγλωσσικές μεταφράσεις

Το σύστημά μας προσφέρει μεταφράσεις και για άλλες γλώσσες πέραν των αγγλικών. Σε πολλά περιβάλλοντα τοπικοποίησης, η υποστήριξη μίας γλώσσας συνίσταται απλά σε ανάθεση συμβολοσειρών. Στην περίπτωση μας, ωστόσο, απαιτείται πλήρης εκφραστικότητα και ευελιξία. Καθώς κάθε γλώσσα έχει τις δικές τις γραμματικές και συντακτικές ιδιαιτερότητες, η απλή ανάθεση συμβολοσειρών δεν είναι επαρκής. Το σύστημά μας αντιμετωπίζει αυτή την πρόκληση θεωρώντας ένα επίπεδο αφαίρεσης που ενσωματώνει τόσο τα διάφορα γραμματικά, όσο και συντακτικά χαρακτηριστικά. Πιο συγκεκριμένα, ορίζουμε την κάθε γλώσσα ως μία προγραμματιστική μονάδα, μέσα στην οποία καθορίζονται τα ιδιαίτερα χαρακτηριστικά της, καθώς και η συμπεριφορά της ως προς τη μετάφραση. Επιπρόσθετα, η κληρονομικότητα προγραμματιστικών μονάδων ευνοεί την ευελιξία μέσω της επαναχρησιμοποίησης κώδικα. Με αυτή την προσέγγιση, μία νέα γλώσσα μπορεί να υποστηριχτεί με τη δημιουργία μιας τέτοιας προγραμματιστικής μονάδας και την εισαγωγή της στο σύστημα.

Το σύστημά μας, πέρα από τα αγγλικά, υποστηρίζει μεταφράσεις σε ελληνικά και ισπανικά. Για παράδειγμα, το Ερώτημα 1 μπορεί να μεταφραστεί στα ελληνικά ως: «Βρες τα 'movies' τύπου action που έχει σκηνοθετήσει ο Coppola και τα 'actors' που παίζουν σε αυτά τα 'movies'» και στα ισπανικά ως: «Buscar los 'movies' de género action que ha rodado Coppola y los 'actors' que actúan en estos 'movies'».

4. Αρχιτεκτονική

Το σύστημά μας αποτελείται από δύο κύρια συστατικά: το μηχανισμό μετάφρασης και τη γραφική διεπαφή χρήστη (Σχήμα 6). Ο μηχανισμός μετάφρασης είναι ο πυρήνας του συστήματος και περιλαμβάνει πέντε κύριες μονάδες: τους γράφους σχήματος και ερωτήματος, το μηχανισμό προτύπων, το μετασχηματιστή ερωτήματος σε γράφο και τη μονάδα αλγορίθμων μετάφρασης. Η γραφική διεπαφή συνιστά το διαμεσολαβητή που προσφέρει τη λειτουργικότητα του μηχανισμού μετάφρασης ερωτημάτων στο χρήστη. Όσον αφορά στις τεχνολογίες, ο μηχανισμός μετάφρασης υλοποιήθηκε σε Java, ενώ η γραφική διεπαφή χρήστη σε Java EE, JSP και servlets. Τέλος, ως Σύστημα Διαχείρισης Βάσης Δεδομένων χρησιμοποιήσαμε την PostgreSQL.



Σχήμα 6: Αρχιτεκτονική συστήματος

Αναφορές

1. A. Sanjay Agrawal, Surajit Chaudhuri, Gautam Das: DBXplorer: A System for Keyword-Based Search over Relational Databases. ICDE 2002: 5-16

2. Georgia Koutrika, Alkis Simitsis, Yannis E. Ioannidis: Précis: The Essence of a Query Answer. ICDE 2006: 69-78

3. Alkis Simitsis, Georgia Koutrika, Yannis Ioannidis: Précis: from unstructured keywords as queries to structured databases as answers. VLDB J. 17(1): 117-149 (2008)

4. Michael Minock: A STEP Towards Realizing Codd's Vision of Rendezvous with the Casual User. VLDB 2007:1358-1361

5. Michael Minock: C-Phrase: A system for building robust natural language interfac-es to databases. Data Knowl. Eng. 69(3): 290-302 (2010)

6. Yannis E. Ioannidis: From Databases to Natural Language: The Unusual Direction. NLDB 2008: 12-16

7. Georgia Koutrika, Alkis Simitsis, Yannis E. Ioannidis: Explaining structured queries in natural language. ICDE 2010: 333-344

Κατανεμημένοι Αλγόριθμοι Συσταδοποίησης Χρησιμοποιώντας το MapReduce και το Πλαίσιο Misco

Γεώργιος Γάσπαρης

g.gasparis@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Αντικείμενο της παρούσας πτυχιακής εργασίας είναι η εισαγωγή και ανάλυση της ιδέας εφαρμογής κατανεμημένων αλγορίθμων συσταδοποίησης σε περιβάλλοντα κινητών τηλεφώνων. Με βάση το πλαίσιο Misco, που υλοποιεί την τεχνική του MapReduce [1] σε περιβάλλον κινητών συσκευών, παρουσιάζουμε ένα αξιόπιστο σύστημα επικοινωνίας κατανεμημένων συσκευών με χαμηλό κόστος. Παρουσιάζεται επίσης και μια εφαρμογή συσταδοποίησης που εκτελείται πάνω από το προαναφερθέν σύστημα και η οποία υλοποιήθηκε για να εκτελεί μια παραλλαγή του γενικού αλγορίθμου K-means.

Λέξεις-Κλειδιά: Κατανεμημένοι αλγόριθμοι συσταδοποίησης, Εξόρυξη γνώσης, MapReduce, πλαίσιο Misco, K-means.

Επιβλέπων: Δημήτριος Γουνόπουλος, Αναπληρωτής Καθηγητής

1. Εισαγωγή

Από τη στιγμή που η κατανεμημένη επεξεργασία υιοθετήθηκε σε πολλούς τομείς της επιστήμης των υπολογιστών και τα κινητά τηλέφωνα εξελίχθηκαν σε πανίσχυρες συσκευές, ένα νέο περιβάλλον δημιουργήθηκε. Επιπροσθέτως, οι τεχνικές εξόρυξης γνώσης αναδεικνύονται σε ολοένα και πιο σημαντικό εργαλείο στη σημερινή εποχή. Ακόμη, η αξιοποίηση και η μετέπειτα εξέλιξη του πλαισίου Misco [2] συντέλεσαν στη συνένωση και επεξεργασία των τριών καινοτόμων προαναφερθέντων τομέων με στόχο τη δημιουργία ενός αξιόπιστου και εύρωστου συστήματος που βρίσκεται ένα επίπεδο υψηλότερα από το παρεχόμενο πλαίσιο και επιτρέπει στους χρήστες την εκτέλεση πολύπλοκων εφαρμογών συσταδοποίησης βασισμένων σε πληροφορίες που παράγονται από τις ίδιες τις συσκευές.

Πλήθος εργασιών έχουν δημοσιευτεί έχοντας ως αντικείμενο αυτό των κατανεμημένων περιβαλλόντων. Τελευταία, πολλές εξ' αυτών όπως οι [2], [3] εστιάζουν στην ανάπτυξη της τεχνικής MapReduce που προτάθηκε από την Google. Άλλες [4], στοχεύουν στο σχεδιασμό της λειτουργίας των περιβαλλόντων αυτών.

Στο σημείο αυτό θέλουμε να αναφέρουμε ότι η ιδέα μας αυτή συντέλεσε στην υλοποίηση μιας εφαρμογής που μπορεί να χρησιμοποιηθεί με σκοπό τον εντοπισμό περιοχών με μεγάλο τουριστικό ενδιαφέρον. Περιοχές, όπως η Ακρόπολη και το Big Ben, είναι δημοφιλείς τοποθεσίες και κατά συνέπεια πολλοί επισκέπτες τις έχουν αποθανατίσει. Αν κάποιος χρήστης εκτελέσει την εφαρμογή μας, αυτή θα ψάξει για κοντινούς συμμετέχοντες και έπειτα θα εκτελέσει τον αλγόριθμο συσταδοποίησης, τοποθετώντας στην ίδια ομάδα, φωτογραφίες που έχουν τραβηχτεί σε κοντινές τοποθεσίες. Με το πέρας της εκτέλεσης ο χρήστης θα γνωρίζει ποια είναι τα πιο δημοφιλή μέρη στον κόσμο και μπορεί έτσι να κανονίσει το επόμενο ταξίδι του.

2. Μοντέλο του συστήματος

Η λειτουργία του συστήματος απαιτεί την ύπαρξη ενός συνόλου κόμβων-κινητών τηλεφώνων. Η σταθερή και υψηλού επιπέδου επικοινωνία των κόμβων διασφαλίζεται από το πλαίσιο Misco. Η συμπεριφορά των κόμβων ρυθμίζεται από έναν ειδικό κόμβο που ονομάζουμε συντονιστή. Αυτός είναι υπεύθυνος για τη διαχείριση των υπολογιστικών πόρων, την ανάθεση εργασιών στους λοιπούς κόμβους, και τη διατήρηση πληροφοριών σχετικά με την πρόοδο της εκτέλεσης. Ο αριθμός των εργαζόμενων κόμβων γίνεται γνωστός κατά τη διάρκεια εκτέλεσης.

Θεωρούμε κάθε δουλειά που υποβάλλεται στο σύστημά μας ως ένα σύνολο από φάσεις. Κάθε φάση εκτελείται παράλληλα και είναι κοινή για όλους τους συμμετέχοντες. Όταν όλοι οι κόμβοι έχουν ολοκληρώσει τους υπολογισμούς τους η δουλειά μεταβαίνει στην επόμενη φάση της. Αυτή είναι η υπόθεση που διαφοροποιεί το σύστημά μας από το Misco και τα άλλα περιβάλλοντα που υλοποιούν το MapReduce. Στην περίπτωσή μας, επιθυμούμε όλοι οι εργαζόμενοι να εκτελούν σειριακά τις δουλειές που τους ανατίθενται, με τρόπο τέτοιο ώστε κανένας να μην προπορεύεται των υπολοίπων. Οι εργαζόμενοι είναι υπεύθυνοι για την περάτωση της δουλειάς που τους ανατίθεται, και πέραν αυτής δεν διατηρούν πληροφορίες για ολοκληρωμένες δουλειές. Στο μοντέλο Misco, τα δεδομένα εισόδου κάθε εργασίας μεταδίδονται από το συντονιστή στους κόμβους. Εδώ, τα δεδομένα παράγονται από τους κόμβους. Σε αντίθεση με τις διάφορες άλλες υλοποιήσεις του MapReduce, στις οποίες ο συντονιστής έχει μονάχα ρόλο οργανωτή, εμείς εισαγάγουμε έναν αναβαθμισμένο συντονιστή, επιφορτισμένο με υπολογιστικές αρμοδιότητες.

3. Ανάλυση κόμβων

3.1 Ανάλυση κόμβου εργάτη

Ένας τέτοιος κόμβος αποτελείται από τρία βασικά μέρη. Το πρώτο είναι υπεύθυνο για την αλληλεπίδραση του κόμβου με το συντονιστή, την αποστολή αιτήσεων και τη λήψη και μεταμόρφωση δεδομένων. Όταν ένας εργάτης είναι άνεργος αποστέλλει αίτηση για εργασία στο συντονιστή μέσω ενός Ενιαίου Εντοπιστή Πόρων (URL). Όλα τα απαραίτητα δεδομένα, όπως αρχεία εισόδου και ενδιάμεσα αρχεία των διαφόρων φάσεων, αποθηκεύονται στο δεύτερο μέρος του εργάτη, την αποθήκη εργασιών. Τέλος, χρησιμοποιούμε ένα μέρος που είναι υπεύθυνο για την μετάδοση στατιστικών πληροφοριών (χρονική διάρκεια υπολογισμού, χρόνος μεταφόρτωσης, κ.α.) μεταξύ κόμβων και συντονιστή.

3.2 Ανάλυση κόμβου συντονιστή

Ο κόμβος αυτός αποτελείται από τέσσερα κύρια μέρη. Αρχικά, ένας χρονοπρογραμματιστής (scheduler) χρησιμοποιείται για την οριοθέτηση των διαφόρων εργασιών. Η αποθήκη εφαρμογής, που διατηρεί πληροφορία σχετικά με την είσοδο και την έξοδο της δουλειάς, αποτελεί το δεύτερο τμήμα του συντονιστή. Χρησιμοποιούμε ακόμη ένα διακομιστή Πρωτοκόλλου Μεταφοράς Υπερκειμένου

(HTTP server), ο οποίος εξυπηρετεί την επικοινωνία συντονιστή-κόμβων. Αυτός είναι υπεύθυνος για τη λήψη αιτήσεων και την παρουσίαση της κατάστασης της εφαρμογής στο χρήστη. Τέλος, υπάρχει ένας ακόμη διακομιστής (UDP) που ακούει για εισερχόμενα μηνύματα στατιστικής πληροφορίας από την πλευρά των εργαζομένων κόμβων.

3.3 Διεπαφή χρήστη

Οι διάφορες δουλειές που υποβάλλονται στο σύστημα είναι πλήρως διαχειρίσιμες από το χρήστη μέσω της διεπαφής φυλλομετρητή. Εκεί, θα βρει τις διάφορες παραμέτρους που σχετίζονται με αυτές, π.χ., μπορεί να καθορίσει τη χρονική διάρκεια κατά την οποία μπορούν οι εργαζόμενοι να δηλώσουν συμμετοχή σε μια δουλειά. Εκεί, επιπλέον θα του παρουσιαστεί και η στατιστική πληροφορία σχετικά με τη δουλειά.

4. Υλοποιήσεις

Όσον αφορά το τεχνικό κομμάτι, επιλέξαμε η επικοινωνία εργαζομένων-συντονιστή να υλοποιείται με βάση την polling-based προσέγγιση. Ακόμη, χρησιμοποιούμε το Πρωτόκολλο Μεταφοράς Υπερκειμένου (HTTP) προκειμένου να μεταφέρουμε αιτήματα εργασίας, πληροφορίες εργασιών και δεδομένα μεταξύ των κόμβων. Σχεδιάσαμε το σύστημα μας ώστε να λειτουργεί σε κινητά τηλέφωνα Nokia N95 8GB [5]. Αυτά ανήκουν στην κατηγορία S60 και υποστηρίζουν μια απλοποιημένη, δοκιμαστική έκδοση της γλώσσας προγραμματισμού Python.

4.1 Υλοποίηση κόμβου εργάτη

Ο κόμβος είναι αρμόδιος για την επεξεργασία των εισερχόμενων εργασιών και τη μετάδοση των αποτελεσμάτων στο συντονιστή. Εκτός από τη βασική λειτουργικότητα που παρέχεται, υποστηρίζουμε επίσης δυνατότητες όπως: εγκατάσταση του πλαισίου, ανανέωση του πηγαίου κώδικα και καθαρισμού των δεδομένων στη συσκευή. Προκειμένου μια συσκευή να συμμετάσχει σε μια δουλειά θα πρέπει να το δηλώσει εντός της περιόδου συμμετοχής. Μόλις ο εργαζόμενος κόμβος ενημερωθεί ότι συμμετέχει στη δουλειά, στέλνει αιτήσεις για εργασία στο συντονιστή. Μεταξύ διαδοχικών αιτημάτων έχουμε ορίσει μια περίοδο οκνηρίας προκειμένου να μην κάνουμε αλόγιστη χρήση των υπολογιστικών μας πόρων.

Ύστερα, χρησιμοποιεί τον Ενιαίο Εντοπιστή Πόρων, που βρίσκεται στην πληροφορία του συντονιστή σχετικά με την εργασία, έτσι ώστε να ανακτήσει, απ' αυτόν, όλα τα σχετικά δεδομένα. Εν συνεχεία, αρχικοποιεί τις εσωτερικές του δομές και ξεκινά την εκτέλεση της εργασίας. Τα αποτελέσματά του αποθηκεύονται τοπικά, ενώ τα κοινοποιεί στον ειδικό κόμβο μέσω ενός μηνύματος (τύπου POST). Παράλληλα με αυτό, αποστέλλει και στατιστική πληροφορία σχετικά με τους χρόνους που χρειάστηκαν για τα διάφορα κομμάτια της εργασίας.

4.2 Υλοποίηση συντονιστή

Ο χρήστης αλληλεπιδρά με τον συστονιστή μέσω μιας διεπαφής Γλώσσας Σήμανσης Υπερκειμένου (HTML). Ο διακομιστής εκτελείται σε τοπικό δίκτυο (LAN) και είναι προσβάσιμος σε διεύθυνση και θύρα καθοριζόμενη από το χρήστη. Μόλις η δουλειά υποβληθεί και ύστερα από τη δημιουργία των απαραίτητων αρχείων και καταλόγων γι' αυτή, ο συντονιστής αναμένει για χρονικό διάστημα ίσο με το χρόνο συμμετοχής. Κατά τη διάρκεια αυτού καταγράφει πληροφορία για τους εργάτες που έκαναν αίτηση να συμμετάσχουν στη δουλειά, ενώ παράλληλα τους στέλνει ένα μήνυμα περιγράφοντάς τη. Στους εκπρόθεσμους στέλνει μια αρνητική απάντηση συμμετοχής. Έτσι, γνωρίζει, τον αριθμό των απαντήσεων που πρέπει να λάβει προτού η δουλειά μεταβεί σε επόμενη φάση.

Ο συντονιστής μας συμπεριφέρεται διαφορετικά απ' ότι ο αντίστοιχος του Misco. Εδώ, έχει διττό ρόλο. Απαντά στις αιτήσεις των κόμβων για εργασίες παρέχοντας τους πληροφορία γι' αυτές, ενώ παράλληλα είναι υπεύθυνος για τη συγκέντρωση και επεξεργασία των αποτελεσμάτων τους, προετοιμάζοντας έτσι την είσοδο της επόμενης φάσης. Μόλις μια δουλειά ολοκληρωθεί ο συντονιστής είναι διαθέσιμος για λήψη νέας. Τέλος, παρέχει στο χρήστη στατιστική πληροφορία σχετικά με την δουλειά που ολοκλήρωσε.

5. Εφαρμογή

Η εφαρμογή που υλοποιούμε, είναι μια παραλλαγή του γενικού αλγορίθμου K-means. Αποτελείται από δύο βήματα, τα βήματα ανάθεσης και επαναϋπολογισμού. Στοχεύουμε στη συσταδοποίηση φωτογραφιών με άξονα το γεωγραφικό τους μήκος και πλάτος.

5.1 Περιγραφή αλγορίθμου

5.1.1 Βήμα ανάθεσης

Στο βήμα αυτό ο συντονιστής αναθέτει στους κόμβους τα αρχικά (τυχαία επιλεγμένα) κέντρα βάρους για τις διάφορες συστάδες. Οι εργάτες υπολογίζουν τις αποστάσεις των σημείων τους από τα κέντρα και αναθέτουν τα σημεία τους στη συστάδα της οποίας το κέντρο είναι πιο κοντά σε αυτά. Πρέπει εδώ να τονίσουμε ότι τα σημεία μας αποτελούν σημεία στην υδρόγειο και γι’ αυτό χρησιμοποιήσαμε μια ειδική μετρική για τον υπολογισμό αποστάσεων, που ονομάζεται απόσταση haversine [6]. Ο τύπος αυτός μας δίνει την απόσταση δυο σημείων στην υδρόγειο με βάση τα γεωγραφικά πλάτη και μήκη τους. Η χρήση της απόστασης αυτής εξασφαλίζει ποσοστό σφάλματος μικρότερο των τριών μέτρων το χιλιόμετρο.

5.1.2 Βήμα επαναϋπολογισμού

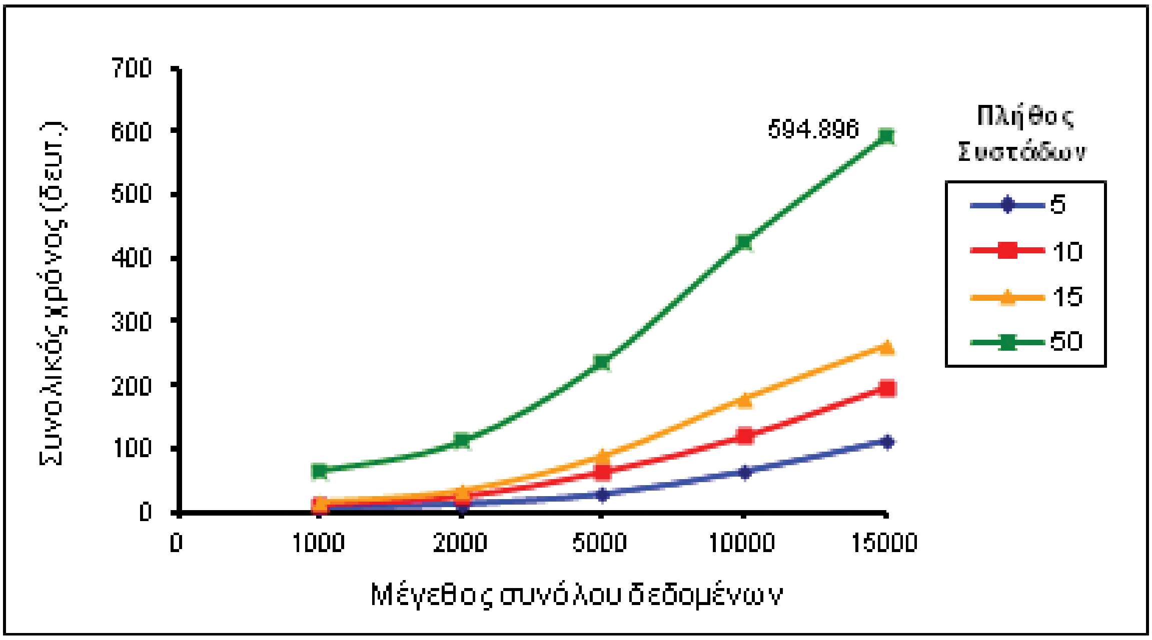
Αυτό συνίσταται από δύο μικρότερα βήματα. Το πρώτο εκτελείται σε κάθε κόμβο ξεχωριστά, ενώ το δεύτερο στο συντονιστή. Στο πρώτο υποβήμα, ο κάθε κόμβος υπολογίζει τα νέα κέντρα βάρους για κάθε συστάδα σύμφωνα με τα δικά του σημεία και τα αποστέλλει στο συντονιστή. Ο υπολογισμός αυτός γίνεται με βάση την τεχνική γεωγραφικού μέσου που υπολογίζει τις συντεταγμένες του σημείου που κείτεται μεταξύ ενός συνόλου σημείων. Στο δεύτερο υποβήμα, ο συντονιστής λαμβάνει όλη την πληροφορία και υπολογίζει τα νέα καθολικά κέντρα βάρους, που είναι κοινά για όλους τους κόμβους. Στην περίπτωση αυτή εισάγουμε την έννοια του βάρους. Συγκεκριμένα, δεν θεωρούμε ισάξιες τις απαντήσεις όλων των κόμβων, αλλά αναθέτουμε σε κάθε μία ένα βάρος. Το βάρος (ένα για κάθε κέντρο συστάδας για κάθε κόμβο) αντιπροσωπεύει το πλήθος των σημείων που αντιστοιχούν στη συγκεκριμένη συστάδα για κάθε κόμβο. Η πληροφορία του βάρους εμπεριέχεται, μαζί με το κέντρο βάρους για κάθε συστάδα, στο μήνυμα του εργάτη.

6. Πειραματικά αποτελέσματα

Διεξήγαμε ένα σύνολο από πειράματα προκειμένου να αξιολογήσουμε την απόδοση του συστήματός μας. Για τα πειράματα χρησιμοποιήσαμε 5 κινητά τηλέφωνα Nokia N95 8GB. Το σύνολο δεδομένων μας περιελάμβανε 15.580 σημεία που αντιστοιχούν σε φωτογραφίες ανά τον κόσμο που μεταμορφώσαμε από τον ιστότοπο Flick-r.

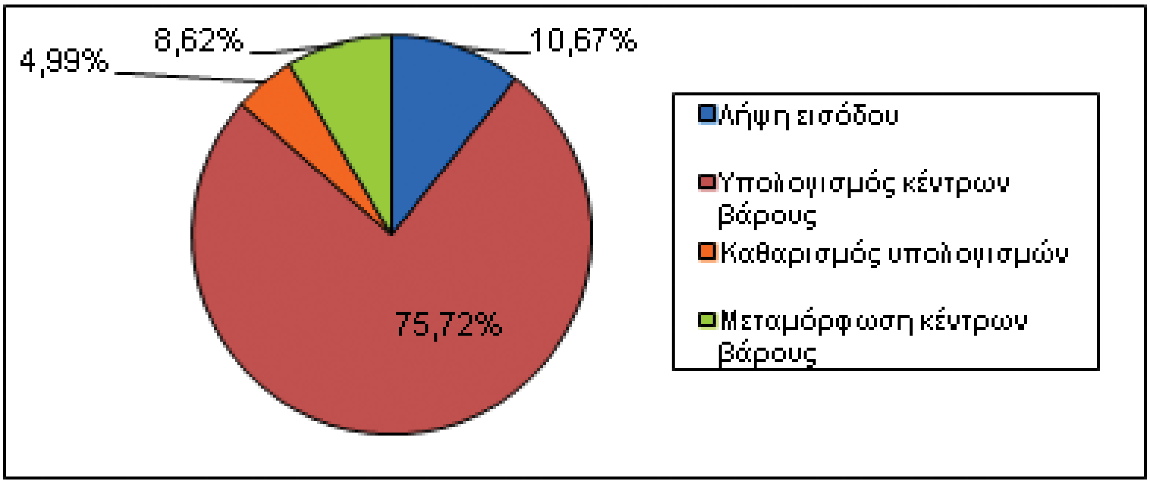
6.1 Πρώτο πείραμα

Στο πρώτο μας πείραμα, κατατμήσαμε το σύνολο των δεδομένων σε πέντε ισομεγέθη τμήματα και αναθέσαμε καθ' ένα εξ' αυτών σε κάθε τηλέφωνο. Ξεκινήσαμε έχοντας συνολικό μέγεθος 1000 σημεία και τελικά χρησιμοποιήσαμε όλα τα δεδομένα μας. Οι συνολικοί χρόνοι εκτέλεσης παρουσιάζονται στο παρακάτω διάγραμμα. Είναι εμφανές ότι οι χρόνοι εκτέλεσης αυξάνονται καθώς μεγαλώνει το μέγεθος των δεδομένων. Επιπλέον η απόδοση του συστήματος επηρεάζεται και από τον αριθμό των συστάδων. Γενικά, όσο περισσότερες συστάδες έχουμε τόσο μεγαλύτερος χρόνος απαιτείται, γιατί απαιτούνται περισσότερες φάσεις εκτέλεσης ώστε να επιτευχθεί σύγκλιση.



Σχήμα 1: Χρόνος σύγκλισης – μέγεθος δεδομένων.

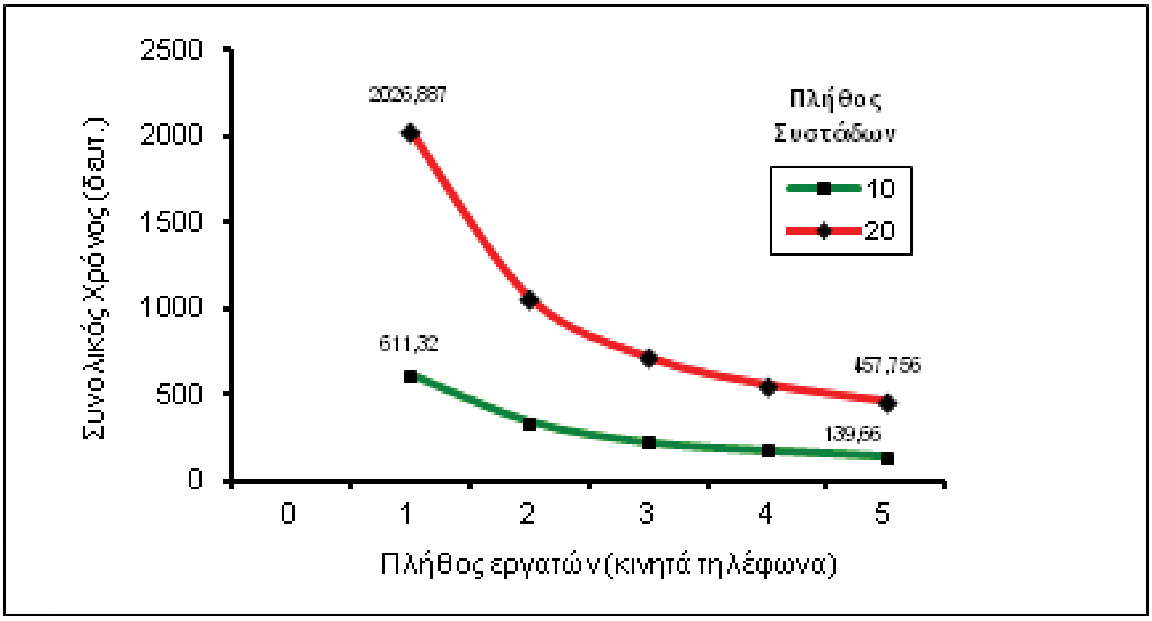
Παρουσιάζουμε επίσης ένα διάγραμμα (Σχήμα 2) που αποτελεί ανάλυση της χρήσης των τηλέφωνων. Ο περισσότερος χρόνος χρησιμοποιείται για τον υπολογισμό των κέντρων βάρους των συστάδων.



Σχήμα 2: Ανάλυση χρήσης κινητών τηλεφώνων.

6.2 Δεύτερο πείραμα

Σε αυτό το πείραμα αποδεικνύουμε ότι το μοντέλο μας συμμερίζεται την αρχή των κατανεμημένων συστημάτων. Εδώ, χρησιμοποιούμε ολόκληρο το σύνολο δεδομένων μας και το αναθέτουμε σταδιακά σε όλους τους εργάτες. Συγκεκριμένα ξεκινάμε με ένα κινητό στο οποίο αναθέτουμε όλα τα σημεία και στη συνέχεια αναθέτουμε τα σημεία εξίσου σε όλα τα κινητά.

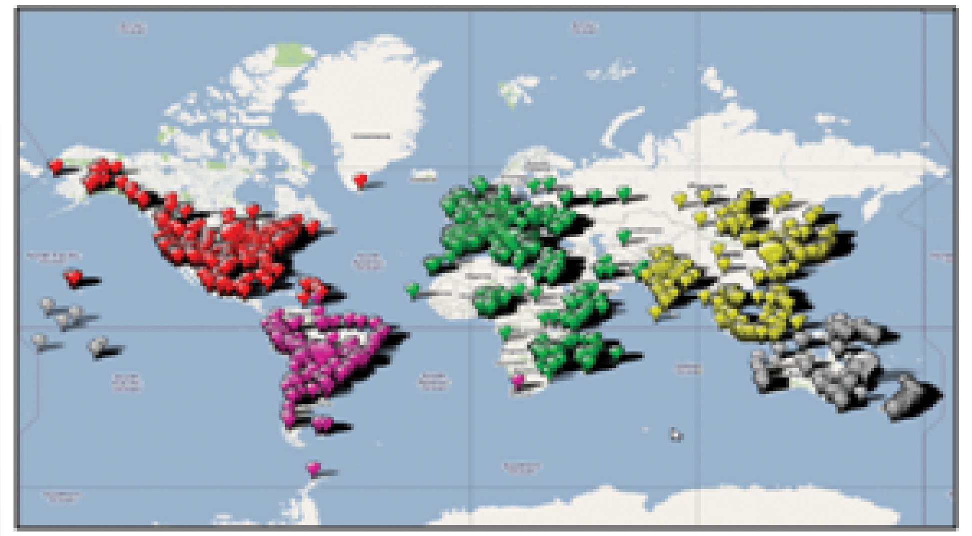


Σχήμα 3: Χρόνος σύγκλισης – πλήθος εργατών.

7. Απεικόνιση των αποτελεσμάτων συσταδοποίησης

Σε αυτό το κομμάτι της εργασίας παρουσιάζουμε γραφικά τα αποτελέσματα των πειραμάτων μας. Όπως έχει αναφερθεί, τα δεδομένα μας αποτελούν φωτογραφίες απ' όλο τον κόσμο. Εδώ, χρησιμοποιούμε την προγραμματιστική διεπαφή της υπηρεσίας Google maps [7] για να παρουσιάσουμε τη δουλειά μας. Στο πρώτο σχήμα εικονίζεται η αρχική κατάσταση των σημείων μας, προτού η συσταδοποίηση λάβει χώρα, ενώ στο δεύτερο παρουσιάζεται το αποτέλεσμα εκτέλεσης της εφαρμογής στο σύστημα μας για αριθμό συστάδων ίσο με 5.

Σχήμα 4: Αρχική κατάσταση των σημείων.



Σχήμα 5: Αποτελέσμα συσταδοποίησης για 5 συστάδες.

Αναφορές

1. Jeffrey Dean, Sanjay Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters", OSDI 2004.

2. A. Dou, V. Kalogeraki, D. Gunopulos, T. Mielikainen and V. H. Tuulos, "Misco: A MapReduce Framework for Mobile Systems", Petra 2010.

3. Disco, http://discoproject.org/.

4. Adam Dou, V. Kalogeraki, D. Gunopulos, T. Mielikinen and V. Tuulos, "Scheduling for Real-Time Mobile MapReduce Systems", New York, 2011.

5. Nokia N95 8 GB, http://www.forum.nokia.com/devices/N95_8GB

6. Wikipedia, "Haversine formula", http://en.wikipedia.org/wiki/Haversine_formula

7. Google maps api, http://code.google.com/intl/el-GR/apis/maps/index.html

Ενεργή Εκπαίδευση Φίλτρου Ανεπιθύμητης Αλληλογραφίας με Επαυξητική Συσταδοποίηση

Κλεάνθη Ε. Γεωργαλά

std06017@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Η παρούσα εργασία εισάγει μία μέθοδο αντιμετώπισης του προβλήματος των ανεπιθύμητων ηλεκτρονικών μηνυμάτων, βασισμένη στην ενεργό εκμάθηση, με χρήση μίας επαυξητικής μεθόδου συσταδοποίησης και έχοντας ως κίνητρο το γεγονός ότι ο χρήστης δεν μπορεί να παρέχει την σωστή κατηγορία για όλα τα μηνύματα που λαμβάνει. Τα μηνύματα ταξινομούνται από ένα φίλτρο που έχει εκπαιδευτεί με τα μηνύματα των προηγούμενων δεσμών-ημερών και στη συνέχεια επανεκπαιδεύουν το φίλτρο με δύο εναλλακτικούς τρόπους: α) μόνο τα μηνύματα για το οποία έχει ζητηθεί η σωστή κατηγορία τους (περιορισμένη εκπαίδευση). β) όλα τα μηνύματα, όπου για αυτά που δεν γνωρίζουμε τη σωστή κατηγορία εμπιστευόμαστε την κρίση του φίλτρου ταξινόμησης (ημι-επιβλεπόμενη εκπαίδευση). Συμπερασματικά, η μέθοδος ημι-επιβλεπόμενης εκπαίδευσης, χρησιμοποιώντας το 10% των κατηγοριών από το σύνολο των μηνυμάτων κατάφερε να αποδώσει εξίσου καλά με μία μέθοδο πλήρους επιβλεπόμενης εκπαίδευσης.

Λέξεις-Κλειδιά: Ενεργός εκμάθηση, Μηχανική μάθηση, Επαυξητική συσταδοποίηση, Ημι-επιβλεπόμενη μάθηση, Ανεπιθύμητη αλληλογραφία

Επιβλέποντες:

Παναγιώτης Σταματόπουλος, Επίκουρος Καθηγητής
Γεώργιος Παλιούρας, Ερευνητής

1. Εισαγωγή

Ως ανεπιθύμητη αλληλογραφία (spam) ορίζεται το σύνολο των μηνυμάτων που είναι αυτόκλητα, ανεπιθύμητα και αποστέλλονται μαζικά από ένα αποστολέα που δεν έχει άμεση σχέση με τους παραλήπτες [1]. Βασικό κίνητρο των ανεπιθύμητων μηνυμάτων ηλεκτρονικής αλληλογραφίας είναι η παροχή πληροφοριών και η διαφήμιση μη υπαρκτών προϊόντων και υπηρεσιών, που αποσκοπούν στις χρηματικές απάτες και στην εγκατάσταση βλαβερού λογισμικού για τον υπολογιστή του παραλήπτη. Άμεσο αποτέλεσμα της ανεπιθύμητης αλληλογραφίας είναι η χρήση ενός μέσου διαδικτυακού επικοινωνίας για κακόβουλη διαφήμιση, υπερφόρτωση του αποθηκευτικού χώρου του ηλεκτρονικού ταχυδρομείου του χρήστη, καθώς και λάθος χρήση του εύρους ζώνης του δικτύου. Μια πρώτη προσέγγιση για την αντιμετώπιση της ανεπιθύμητης αλληλογραφίας είναι η χρήση χειροκίνητων μεθόδων για τη μείωση και την άμεση αναγνώριση μηνυμάτων τέτοιου είδους με σκοπό την προστασία του χρήστη από επιτήδειους αποστολείς. Κάποιες μέθοδοι είναι η αναγνώριση συγκεκριμένων λέξεων, καταλήξεων αρχείων και κωδικοποιήσεων κειμένων από το φίλτρο προστασίας. Οι τακτικές αυτές δεν ανταποκρίθηκαν με τον καλύτερο δυνατό τρόπο στις ανάγκες των χρηστών ηλεκτρονικής αλληλογραφίας, μιας και προκαλούσαν απώλεια χρήσιμων μηνυμάτων ή εξαπάτηση του παραλήπτη, παρ' όλες τις προσπάθειες να προστατεύσουν την ακεραιότητα της ηλεκτρονικής του διεύθυνσης. Αυτό είχε ως αποτέλεσμα, να στραφεί η προσοχή σε αυτόματους ταξινομητές μηνυμάτων.

1.1 Τεχνικές φίλτρων ανεπιθύμητης αλληλογραφίας βασισμένες στη μηχανική μάθηση

Έχουν χρησιμοποιηθεί αρκετοί αλγόριθμοι ενεργούς εκμάθησης με σκοπό την αντιμετώπιση της ανεπιθύμητης αλληλογραφίας λόγω της προσπάθειας αντικατάστασης των χειροκίνητων μεθόδων [2]. Η ενεργός εκμάθηση είναι ένα είδος μηχανικής μάθησης που βασίζεται στην ημι-επιβλεπόμενη μάθηση. Η ενεργός εκμάθηση στοχεύει στην επίτευξη μεγαλύτερης ακρίβειας με λιγότερα δεδομένα προς εκπαίδευση, για τα οποία γνωρίζουμε την πραγματική τους κατηγορία. Σε προβλήματα μηχανικής μάθησης, όπου η ανάθεση ετικετών σε όλα τα δεδομένα είναι μία δύσκολη, χρονοβόρα και ακριβή διαδικασία, η ενεργός εκμάθηση αποτελεί μια καλή λύση.

Η μέθοδος των Κ-κοντινότερων γειτόνων είναι γνωστή μέθοδος ταξινόμησης κειμένων, η οποία αποφασίζει την ετικέτα ενός εγγράφου με βάση την ομοιότητα του με τα αποθηκευμένα παραδείγματα. Μια προσέγγιση, από την πλευρά της ενεργούς εκμάθησης, είναι η χρήση διαφορετικών βαρών σε κάθε λέξη του κειμένου ή στις αποστάσεις του από τα υπόλοιπα δεδομένα χρησιμοποιώντας το πληροφοριακό κέρδος (information gain). Δίνοντας διαφορετική βαρύτητα σε κάθε λέξη ή απόσταση, η ενεργός μάθηση αξιοποιεί ανάλογα την πληροφορία που μπορεί να παρέχει αυτή η λέξη στον ταξινομητή [3].

Οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines – SVM) αποτελούν μια αρκετά διαδεδομένη μέθοδο ταξινόμησης. Οι ταξινομητές SVM αναζητούν ένα υπερεπίπεδο που διαχωρίζει τα θετικά από τα αρνητικά παραδείγματα, μεγιστοποιώντας το περιθώριο (margin) μεταξύ των δεδομένων που ανήκουν στις δύο κατηγόριες [4]. Στόχος είναι να βρεθεί ένα καλό υπερεπίπεδο, το οποίο κατηγοριοποιεί με σωστό τρόπο τα δεδομένα εκπαίδευσης. Μια προσέγγιση SVMs που χρησιμοποιεί ενεργό μάθηση είναι η επιλεκτική δειγματοληψία, όπου η ενεργός μάθηση βοηθάει τον ταξινομητή να επιλέξει ένα υποσύνολο δεδομένων κοντά στο υπερεπίπεδο, με σκοπό να αυξηθεί η ακρίβεια του αλγόριθμου και να μειωθούν μελλοντικά λάθη. Επίσης, σε μια άλλη προσέγγιση, σκοπός είναι το επόμενο δεδομένο που θα επιλεγεί προς ταξινόμηση να μειώσει το μέγεθος του χώρου που περιλαμβάνει όλα τα πιθανά υπερεπίπεδα, με σκοπό να βρεθεί το υπερεπίπεδο που θα μπορούσε να σχηματιστεί αν είχαμε την πραγματική κλάση όλων των δεδομένων. Η ενεργός μάθηση επιλέγει κάθε φορά το στοιχείο που μπορεί να μειώσει τον χώρο όλων των πιθανών υπερεπιπέδων [5].

Μια τελευταία προσέγγιση αφορά την επιτροπή Ταξινομητών (Query by Committee –QBC) σε συνδυασμό με τη Βελτιστοποίηση Πιθανοφάνειας (Expectation Maximization– EM) [6]. Η μέθοδος QBC επιλέγει κείμενα προς ταξινόμηση μελετώντας τη διαφωνία μεταξύ της επιτροπής των ταξινομητών, ενώ ο αλγόριθμος EM χρησιμοποιείται με σκοπό να εκμεταλλευτούμε την πληροφορία που παρέχεται από κείμενα που δεν έχουν ταξινομηθεί. Σε αυτή την περίπτωση ο αλγόριθμος της μηχανικής μάθησης παρέχει την δυνατότητα επιλογής ετικετών από το σύνολο των μη ταξινομημένων κειμένων, κυρίως εκείνων που θα ενισχύσουν τη γνώμη του ταξινομητή και παρουσιάζουν μια διαφορετική συμπεριφορά από τα ήδη ταξινομημένα μηνύματα.

Οι παραπάνω προσεγγίσεις που βασίζονται στις μεθόδους K-κοντινότερων γειτόνων, SVMs και στην επιτροπή Ταξινομητών, είναι υπολογιστικά ανέφικτο να εφαρμοστούν, διότι σε ένα σύνολο υποψήφιων δεδομένων εκπαίδευσης η αξιολόγηση του καθενός είναι δαπανηρή και χρονοβόρα διαδικασία. Η ενεργός μάθηση βελτιώνει τις προσεγγίσεις αυτές, αλλά λόγω του παράγοντα της αβεβαιότητας και του τρόπου διαχείρισης των δεδομένων από τις τεχνικές αυτές, τείνει να επιλέγει δείγματα ακραίων τιμών επειδή διαφέρουν έντονα με τα προηγούμενα δεδομένα που είχε επιλέξει προς εκπαίδευση. Οι ακραίες τιμές είναι μεμονωμένες περιπτώσεις και η γνώση τους δεν ενδυναμώνει τον ταξινομητή εφόσον δεν αποτελούν μεγάλο κομμάτι των δεδομένων.

1.2 Τεχνικές ταξινόμησης κειμένων βασισμένες στην επαυξητική συσταδοποίηση

Τοποθετώντας κείμενα παρόμοιου περιεχομένου στην ίδια συστάδα βοηθάμε το χρήστη να περιηγηθεί με ευκολία σε διαφορετικές θεματολογίες και επιτρέπουμε σε μηχανές αναζήτησης να εξετάσουν αποτελεσματικά μεγάλες συλλογές δεδομένων. Οι περισσότεροι αλγόριθμοι συσταδοποίησης που έχουν παρουσιαστεί, όπως ο ιεραρχικός αλγόριθμος COBWEB [7], επεξεργάζονται τα κείμενα και ταυτόχρονα τα τοποθετούν σε συστάδες με επαναληπτικές μεθόδους βελτιστοποίησης. Αντιθέτως, στην επαυξητική συσταδοποίηση, καθώς λαμβάνονται νέα κείμενα, μεταβάλλονται οι συστάδες ανάλογα χωρίς να ανακατατάσσονται τα ήδη υπάρχοντα κείμενα, ταξινομώντας τα από την αρχή και αλλάζοντας έτσι τη μορφή της συσταδοποίησης. Ένα ακόμα σημαντικό θέμα της μεθόδου συσταδοποίησης, είναι ο τρόπος που χειρίζεται την αλλαγή θέματος σε τακτά χρονικά διαστήματα. Η θεματολογία και οι τεχνικές των ανεπιθύμητων μηνυμάτων αλλάζουν σε τακτά χρονικά διαστήματα και για αυτό το λόγο, τα φίλτρα που βασίζονται στην επαυξητική συσταδοποίηση δίνουν μεγαλύτερο βάρος στη χρονική περίοδο που έχει ληφθεί το μήνυμα.

1.3 Συνεισφορά εργασίας

Στην εργασία αυτή συνδυάζεται η ενεργός μάθηση και η μέθοδος επαυξητικής συσταδοποίησης έχοντας, ως στόχο, να δημιουργήσει μια νέα τεχνική φιλτραρίσματος της ανεπιθύμητης αλληλογραφίας του ηλεκτρονικού ταχυδρομείου. Το φίλτρο που αναπτύχθηκε δέχεται ως είσοδο δεδομένα για τα οποία δεν γνωρίζουμε εξ' αρχής την φύση τους, αν είναι δηλαδή επιθυμητά, ή όχι, και τα κατατάσσει σε συστάδες. Η προτεινόμενη μέθοδος συνδυάζει στοιχεία τόσο της επιβλεπόμενης, όσο και της μη-επιβλεπόμενης μάθησης, αφού υπάρχει ένα σύνολο δεδομένων για τα οποία αποκαλύπτεται η φύση τους και εκπαιδεύει με αυτό τον τρόπο τον ταξινομητή Naïve Bayes [8]. Το φίλτρο εκπαιδεύεται με δύο τρόπους : α) με τα μηνύματα για τα οποία είναι γνωστή η πραγματική τους φύση (περιορισμένη εκπαίδευση) και β) με όλα τα μηνύματα, όπου για αυτά που δεν γνωρίζουμε την πραγματική τους κατηγορία, στηριζόμαστε στην απόφαση του φίλτρου που αναπτύχθηκε τις προηγούμενες ημέρες (ημι-επιβλεπόμενη εκπαίδευση). Η απόφαση για την αποκάλυψη της πραγματικής φύσης ενός νέου μηνύματος γίνεται με την βοήθεια της ενεργούς εκμάθησης, όπου μέσα από ένα σύνολο περιπτώσεων, επιλέγουμε αν θα ζητηθεί από το χρήστη η φύση του μηνύματος καθώς το ταξινομούμε σε μία κατηγορία. Σκοπός της εργασίας είναι, με όσο το δυνατό λιγότερα επισημειωμένα δεδομένα εκπαίδευσης, να επιτύχουμε την προσδοκώμενη ακρίβεια, συγκρίνοντας πειραματικά τα αποτελέσματα μας με ένα ταξινομητή που εκπαιδεύεται σε όλα τα δεδομένα (πλήρως επιβλεπόμενη εκπαίδευση).

2. Αλγόριθμοι εκπαίδευσης φίλτρου ανεπιθύμητης αλληλογραφίας με επαυξητική συσταδοποίηση

Βασικό κίνητρο για την εκπόνηση της εργασίας αυτής, αποτελεί η εκπαίδευση ενός απλού ταξινομητή (Naive Bayes) για την εισερχόμενη αλληλογραφία, με όσο το δυνατό λιγότερα ταξινομημένα δεδομένα. Θεωρούμε ότι τα μηνύματα λαμβάνονται από το χρήστη σε δέσμες, π.χ. μίας ημέρας και ότι ο ταξινομητής, μπορεί να αποδώσει το ίδιο καλά χωρίς να χρειάζεται να εκπαιδευτεί πάνω σε όλα τα μηνύματα κάθε δέσμης. Σκοπός μας είναι να μειώσουμε τις περιπτώσεις λανθασμένης απόφασης, χωρίς να επιβαρυνθεί πολύ ο χρήστης.

2.1 Επαυξητική συσταδοποίηση

Θεωρούμε πως την πρώτη μέρα, μας γίνεται γνωστή η φύση όλων των μηνυμάτων που θα λάβουμε και με αυτό τον τρόπο κατασκευάζουμε τις δύο μεγάλες κλάσεις, την κλάση των ανεπιθύμητων μηνυμάτων (spam) και την κλάση των επιθυμητών μηνυμάτων (ham). Στην αρχικοποίηση του συστήματος, κάθε συστάδα περιλαμβάνει αποκλειστικά μηνύματα που έχουν κριθεί ως επιθυμητά ή ως ανεπιθύμητα.

Για κάθε άλλη μέρα εκτός της πρώτης, πραγματοποιείται η διαδικασία αναπαράστασης του μηνύματος σε διάνυσμα και στη συνέχεια γίνεται χρήση της μετρικής, με σκοπό να αποφασιστεί η συσχέτιση του μηνύματος με τις δύο κλάσης. Ο αλγόριθμος εξετάζει κάθε νέο εισερχόμενο μήνυμα με σκοπό να βρεθεί το πόσο μοιάζει με κάθε συστάδα των δύο κλάσεων. Δηλαδή, μελετάμε την εγγύτητα του μηνύματος στη συγκεκριμένη συστάδα και στις υπόλοιπες συστάδες της ίδιας κλάσης. Κάθε νέο μήνυμα, που εισάγεται στο σύστημα, αναπαρίσταται σε ένα διάνυσμα της μορφής $<X_1 , x_1> , <X_2 , x_2> , \dots , <X_\nu , x_\nu>$, όπου κάθε στοιχείο της μορφής x_i αποτελεί το πλήθος των φορών που εμφανίζεται η λέξη X_i στο μήνυμα. Η μετρική που χρησιμοποιήσαμε, με σκοπό να υπολογίσουμε την εγγύτητα του νέου μηνύματος με τις ήδη υπάρχουσες συστάδες, παρέχει καλύτερη ποιότητα και δημιουργεί πιο συμπαγείς συστάδες, δεδομένου ότι η αναπαράσταση των λέξεων γίνεται με τη βοήθεια αριθμητικών τιμών. Οι συστάδες που παράγονται με τη χρήση της μετρικής αυτής δείχνουν μικρή ευαισθησία στη σειρά με την οποία εμφανίζονται τα δεδομένα [9]. Συχνά εμφανιζόμενες λέξεις δεν χρησιμοποιούνται στο διάνυσμα, γιατί η πληροφορία που μεταφέρουν δυσκολεύει τη λειτουργία του φίλτρου. Για κάθε νέο μήνυμα, αφού τελειώσει ο έλεγχος για κάθε του λέξη, αθροίζονται οι παραπάνω ακέραιοι αντιστοίχως και βγαίνει ένας βαθμός που χαρακτηρίζει τη συστάδα σε σχέση με το νέο μήνυμα. Για κάθε κλάση, σκοπός είναι να βρούμε το μέγιστο πηλίκο, δηλαδή τη συστάδα spam ή ham που είναι κοντινότερα στο νέο μήνυμα. Στο τέλος του υπολογισμού, έχουμε δύο βαθμούς, $rtSpam$ και $rtHam$ στην spam και στην ham συστάδα που είναι πιο κοντά στο μήνυμα [8]. Οι τύποι συνοψίζονται παρακάτω :

$$rtHamC_1 = \frac{\sum_{x_i=x_1}^{x_\nu} \textit{πληθος μηνυματων της συσταδας } C_i \textit{ που δεν περιεχουν τη λεξη } \mathrm{X}_i}{\sum_{x_i=x_1}^{x_\nu} \textit{πληθος μηνυματων της κλασης ham, εκτος της συσταδας } C_i\textit{, που περιεχουν τη λεξη } \mathrm{X}_i}$$

και

$$rtSpamC_1 = \frac{\sum_{x_i=x_1}^{x_\nu} \textit{πληθος μηνυματων της συσταδας } C_i \textit{ που δεν περιεχουν τη λεξη } \mathrm{X}_i}{\sum_{x_i=x_1}^{x_\nu} \textit{πληθος μηνυματων της κλασης spam, εκτος της συσταδας } C_i\textit{, που περιεχουν τη λεξη } \mathrm{X}_i}$$

Για κάθε κλάση, αν το rt είναι πάνω από το κάποιο πάνω όριο, το νέο μήνυμα θα τοποθετηθεί στην υπάρχουσα συστάδα με το μεγαλύτερο σκορ, ενώ αν είναι κάτω από ένα κάτω όριο θα δημιουργήσουμε μία νέα συστάδα. Ενδιαφέρουσα είναι η περίπτωση που τα rt είναι μεταξύ των ορίων. Στην εργασία μας, θεωρούμε ότι ο αλγόριθμος συσταδοποίησης δεν είναι σε θέση να διαμορφώσει άποψη για το μήνυμα αυτό και το τοποθετεί στην δομή not-yet-buffer. Στο τέλος της ημέρας, ελέγχει ξανά τα μηνύματα που έχουν τοποθετηθεί εκεί, υπολογίζει ξανά τη μετρική rt και εφαρμόζει εκ νέου τις τρεις παραπάνω περιπτώσεις.

Στις περιπτώσεις που ένα νέο μήνυμα φτιάξει μία νέα συστάδα ή ενωθεί με μία υπάρχουσα, η νέα πληροφορία που φέρνει το μήνυμα πρέπει να αξιοποιηθεί αναλόγως. Η έννοια της επαυξητικής συσταδοποίησης, έγκειται στο γεγονός ότι δεν χρειάζεται να ανακατασκευάζουμε τις συστάδες από την αρχή όταν εμφανιστεί ένα νέο μήνυμα. Για αυτόν το λόγο, κάθε συστάδα περιέχει ένα διάνυσμα της μορφής $<N_1\ ,\ n_1>$, $<N_2\ ,\ n_2>$, … , $<N_\nu\ ,\ n_\nu>$, όπου τα N_i είναι οι διαφορετικές λέξεις όλων των κειμένων της συστάδας και n_i είναι το πλήθος των φορών που εμφανίζονται οι λέξεις στη συστάδα. Αν εισαχθεί νέο κείμενο σε μία συστάδα, τότε ενημερώνεται το διάνυσμα εισάγοντας νέες λέξεις στη συστάδα και ενημερώνοντας το πλήθος των φορών που εμφανίζεται κάθε λέξη στη συστάδα.

2.2 Ενεργός εκμάθηση

Την πρώτη ημέρα, το φίλτρο δεν μπορεί να αναπτύξει άποψη για κάποιο νέο μήνυμα γιατί δεν έχει προηγούμενη γνώση, με βάση την οποία θα μπορεί να το κρίνει. Συνεπώς, θεωρούμε πως την πρώτη μέρα μας γίνεται γνωστή η φύση όλων των μηνυμάτων που θα λάβουμε. Επειδή γνωρίζουμε την κατηγορία κάθε μηνύματος, υπολογίζουμε για την αντίστοιχη κλάση την ποσότητα rt και τη συγκρίνουμε με τα όρια που τέθηκαν από τον χρήστη, οδηγώντας το μήνυμα είτε σε νέα συστάδα, είτε σε με μία ήδη υπάρχουσα, είτε τοποθετώντας το στον not-yet-buffer της κλάσης. Στο τέλος της ημέρας, προσπαθούμε να αδειάσουμε τον not-yet-buffer, επανελέγχοντας με τον ίδιο τρόπο τα μηνύματα που είναι εκεί. Επίσης, την πρώτη μέρα όλα τα μηνύματα χρησιμοποιούνται για την εκπαίδευση του ταξινομητή των μηνυμάτων, ανεξάρτητα αν τοποθετήθηκαν σε κάποια συστάδα ή στον not-yet-buffer.

Για κάθε επόμενη μέρα, αφού αναπαρασταθεί το κάθε νέο μήνυμα σε διάνυσμα και υπολογιστούν τα $rtSpam$ και $rtHam$ των αντίστοιχων συστάδων spam και ham, με τις οποίες το μήνυμα παρουσιάζει μεγαλύτερη ομοιότητα, εξετάζονται οι ακόλουθες επτά (7) περιπτώσεις:

- <u>$rtSpam$ πάνω από το άνω όριο και $rtHam$ κάτω από το κάτω όριο.</u> Σε αυτή την περίπτωση, χωρίς να ζητήσουμε ετικέτα από το χρήστη, θεωρούμε πως το μήνυμα είναι spam και το τοποθετούμε με τη συστάδα της spam κλάσης με το υψηλότερο $rtSpam$.

- <u>$rtHam$ πάνω από το άνω όριο και $rtSpam$ κάτω από το κάτω όριο.</u> Σε αυτή την περίπτωση, χωρίς να ζητήσουμε ετικέτα από το χρήστη, θεωρούμε πως το μήνυμα είναι ham και το τοποθετούμε με τη συστάδα της ham κλάσης με το υψηλότερο $rtHam$.

- <u>$rtSpam$ και $rtHam$ κάτω από το κάτω όριο.</u> Σε αυτή τη περίπτωση, ζητάμε να μας αποκαλυφθεί η πραγματική φύση του μηνύματος, το τοποθετούμε στην αντίστοιχη κλάση και φτιάχνουμε νέα συστάδα.

- <u>$rtSpam$ και $rtHam$ πάνω από το άνω όριο.</u> Σε αυτή την περίπτωση, υπάρχουν δύο εκδοχές : Αν η απόλυτη διαφορά των $rtSpam$ και $rtHam$ είναι μικρότερη της προκαθορισμένης τιμής, τότε ζητάμε να μας αποκαλυφθεί η πραγματική φύση του μηνύματος, το τοποθετούμε στην αντίστοιχη κλάση και το ενώνουμε με την αντίστοιχη συστάδα. Αν δεν είναι, δεν ζητάμε ετικέτα, και αν το $rtSpam$ είναι μεγαλύτερο του $rtHam$, το θεωρούμε spam και το τοποθετούμε στην κλάση spam στην αντίστοιχη συστάδα, αλλιώς το θεωρούμε ham και το τοποθετούμε στην κλάση spam και στην αντίστοιχη συστάδα.

- $rtSpam$ μεταξύ κάτω και άνω ορίου και $rtHam$ μικρότερο του κάτω ορίου. Σε αυτή την περίπτωση, ζητάμε να μας αποκαλυφθεί η πραγματική φύση του μηνύματος και αν είναι spam μπαίνει στον not-yet-buffer της κλάσης spam, αλλιώς, τοποθετείται στην κλάση ham και φτιάχνει σε αυτή νέα συστάδα.

- $rtHam$ μεταξύ κάτω και άνω ορίου και $rtSpam$ μεγαλύτερο του πάνω ορίου. Σε αυτή την περίπτωση, δεν ζητάμε ετικέτα, θεωρούμε το μήνυμα spam, το τοποθετούμε στην κλάση spam και το ενώνουμε με την αντίστοιχη συστάδα.

- $rtSpam$ και $rtHam$ μεταξύ κάτω και άνω ορίου. Σε αυτή την περίπτωση, υπάρχουν δύο εκδοχές : Αν η απόλυτη διαφορά των $rtSpam$ και $rtHam$ είναι μικρότερη μίας προκαθορισμένης τιμής, τότε ζητάμε να μας αποκαλυφθεί η πραγματική φύση του μηνύματος, το τοποθετούμε στον αντίστοιχο not-yet-buffer. Αν δεν είναι, δεν ζητάμε ετικέτα, και αν το $rtSpam$ είναι μεγαλύτερο του $rtHam$, το θεωρούμε spam και το τοποθετούμε στον not-yet-buffer και στην κλάση των spam, αλλιώς, το θεωρούμε ham και το τοποθετούμε στον not-yet-buffer και στην κλάση των ham.

Παράλληλα με τις μεθόδους, που εκπαιδεύουμε τον ταξινομητή με μερική ή πλήρη επίβλεψη, χρησιμοποιούμε μία ακόμα μέθοδο ως μέτρο σύγκρισης των αλγόριθμων που προτείνουμε. Αυτή η τέταρτη μέθοδος, χρησιμοποιεί για την ταξινόμηση των μηνυμάτων μόνο την κρίση του αλγόριθμου επαυξητικής συσταδοποίησης.

2.3 Αλγόριθμοι εκπαίδευσης

Οι διαφορετικοί τρόποι εκπαίδευσης του ταξινομητή Naïve Bayes είναι οι εξής:

- Εκπαίδευση μόνο με τα δεδομένα για τα οποία έχουμε ζητήσει ετικέτα (περιορισμένη εκπαίδευση - ΠΕ). Σε αυτή τη μέθοδο, ο ταξινομητής εκπαιδεύεται χρησιμοποιώντας μόνο τα μηνύματα για τα οποία είναι γνωστή η πραγματική τους ετικέτα. Με αυτόν τον τρόπο επιλέγουμε να αγνοήσουμε ένα μεγάλο σύνολο δεδομένων για τα οποία δεν είμαστε βέβαιοι και ενδέχεται να εισάγουν θόρυβο. Επίσης, στηριζόμαστε ιδιαίτερα στην ικανότητα του αλγόριθμου ενεργούς εκμάθησης για να επιλέγει «σημαντικά» μηνύματα, για τα οποία ζητά την πραγματική τους ετικέτα.

- Εκπαίδευση πάνω σε όλα τα μηνύματα (ημιεπιβλεπόμενη εκπαίδευση - ΗΕ). Η μέθοδος αυτή, εκπαιδεύει τον ταξινομητή πάνω σε όλα τα μηνύματα της ημέρας. Για όσα γνωρίζουμε την πραγματική τους ετικέτα την χρησιμοποιούμε. Για όσα δεν γνωρίζουμε, εμπιστευόμαστε την κρίση του ταξινομητή που είχε εκπαιδευτεί την προηγούμενη ημέρα. Η προσέγγιση αυτή εστιάζει στο μέγεθος των δεδομένων εκπαίδευσης, αποδεχόμενη χαμηλότερη ποιότητα λόγω θορύβου που εισάγεται από τον ταξινομητή.

- Εκπαίδευση πάνω σε όλα τα μηνύματα γνωρίζοντας τις κατηγορίες τους (πλήρως επιβλεπόμενη εκπαίδευση - ΠΕΕ). Η μέθοδος χρησιμοποιήθηκε ως μέτρο σύγκρισης των δύο παραπάνω μεθόδων που χρησιμοποιούν ενεργή εκμάθηση. Αφού χρησιμοποιεί όλα τα δεδομένα και χωρίς θόρυβο, αναμένεται να ξεπεράσει την απόδοση των δύο παραπάνω μεθόδων.

Χρησιμοποιήσαμε και μία τέταρτη μέθοδο που χρησιμοποιεί για την ταξινόμηση των μηνυμάτων μόνο την κρίση του αλγόριθμου επαυξητικής συσταδοποίησης (ΕΑΚ). Αναλυτικά, αφού υπολογιστούν οι ποσότητες *rtSpam* και *rtHam*, αν η *rtHam* είναι μεγαλύτερη ή ίση της *rtSpam* τότε θεωρούμε πως το μήνυμα είναι ham και αντίστροφα. Η μέθοδος αυτή αναμένεται να έχει χαμηλότερη απόδοση από τις άλλες, καθώς δεν εκπαιδεύει κάποιον ταξινομητή.

3. Πειράματα

Στα πειράματα μας χρησιμοποιήθηκε μία δημόσια ελεύθερη συλλογή κειμένων, που ονομάζεται Enron-Spam. Τα επιθυμητά μηνύματα της συλλογής μας προέρχονται από έξι εργαζόμενους στην εταιρία Enron και δόθηκαν στο κοινό μετά τη δημοσίευση του σκανδάλου της εταιρίας στην Αμερική. Τα μη επιθυμητά μηνύματα προέρχονται από 4 διαφορετικές πηγές: SpamAssasin, Honeypot project, τη συλλογή μηνυμάτων του Bruce Guenter και του κ. Παλιούρα [10]. Κάθε συλλογή μηνυμάτων εξετάστηκε και με τους τρεις τρόπους επιλογής δεδομένων προς εκπαίδευση.

Σκοπός της αξιολόγησης ήταν να ελέγξουμε αν ένα φίλτρο μπορεί να εκπαιδευτεί εξίσου καλά από δεδομένα, των οποίων γνωρίζουμε μόνο μερικώς την πραγματική ετικέτα. Για την αξιολόγηση, χρησιμοποιήσαμε μετρικές όπως το spam και ham recall. Υπενθυμίζουμε πως οι δύο αυτές μετρικές υπολογίζονται ως:

- Ως ham recall ορίζουμε το πηλίκο true positives προς το άθροισμα των true positives και false negatives, δηλαδή πόσα επιθυμητά μηνύματα ταξινομήθηκαν σωστά στο σύνολο όλων των επιθυμητών μηνυμάτων, και
- Ως spam recall το πηλίκο των true negatives προς το άθροισμα των true negatives και false positives, δηλαδή πόσα ανεπιθύμητα μηνύματα ταξινομήθηκαν σωστά στο σύνολο των ανεπιθύμητων μηνυμάτων.

Με βάση τις δύο παραπάνω μετρικές και με σκοπό να συγκρίνουμε τις μεθόδους εκπαίδευσης του ταξινομητή, χρησιμοποιήσαμε τα διαγράμματα ROC στα οποία ο άξονας των χ αντιστοιχεί στην τιμή 1 – ham recall και ο άξονας των ψ η τιμή spam recall. Στα αποτελέσματα του τεστ κάθε ημέρας, ο ταξινομητής επιστρέφει μία δεκαδική τιμή μεταξύ 0 και 1. Θεωρώντας ένα όριο Δ = 0.5, αποφασίζουμε πως τα μηνύματα με τιμή κάτω του 0.5 παίρνουν την ετικέτα ham, αλλιώς την ετικέτα spam . Δεδομένου αυτής της ιδιότητας του ορίου Δ, στις δικές μας πειραματικές αναλύσεις αποφασίσαμε για κάθε δεκαδική τιμή του Δ από το 0.1 μέχρι το 0.99, να βρίσκουμε ένα σημείο στην καμπύλη ROC. Κάθε καμπύλη χαρακτηρίζεται από το κάτω και το άνω όριο που εισάγουμε στην αρχή του προγράμματος καθώς και από την μέθοδο που χρησιμοποιούμε. Στην τρίτη μέθοδο, δεν παίζουν ρόλο τα άνω και κάτω όρια και στην τέταρτη μέθοδο η καμπύλη χαρακτηρίζεται από ένα μοναδικό σημείο μιας και υπολογίζουμε μόνο τα τελικά-συνολικά spam και ham recall. Επίσης, μελετήσαμε τη συμπεριφορά της συσταδοποίησης καθώς περνούν οι μέρες με διαγράμματα spam και ham recall per batch και τέλος δημιουργήσαμε διαγράμματα, στα οποία ο οριζόντιος άξονας είναι ο άξονας των ημερών εκπαίδευσης και ο κάθετος άξονας είναι ο άξονας που περιέχει τον αριθμό των ετικετών που έχουν ζητηθεί μέχρι και εκείνη την ημέρα.

Ένα αρχικό συμπέρασμα που προκύπτει από τα διαγράμματα ROC όλων των συλλογών, είναι πως η μέθοδος πλήρους επιβλεπόμενης μάθησης αποδίδει καλύτερα από τις δύο άλλες μεθόδους εκπαίδευσης που αναπτύχθηκαν. Όπως είναι αναμενόμενο, η πλήρης γνώση όλων των ετικετών των μηνυμάτων εκπαίδευσης επιτυγχάνει καλύτερα αποτελέσματα στην ταξινόμηση μηνυμάτων.

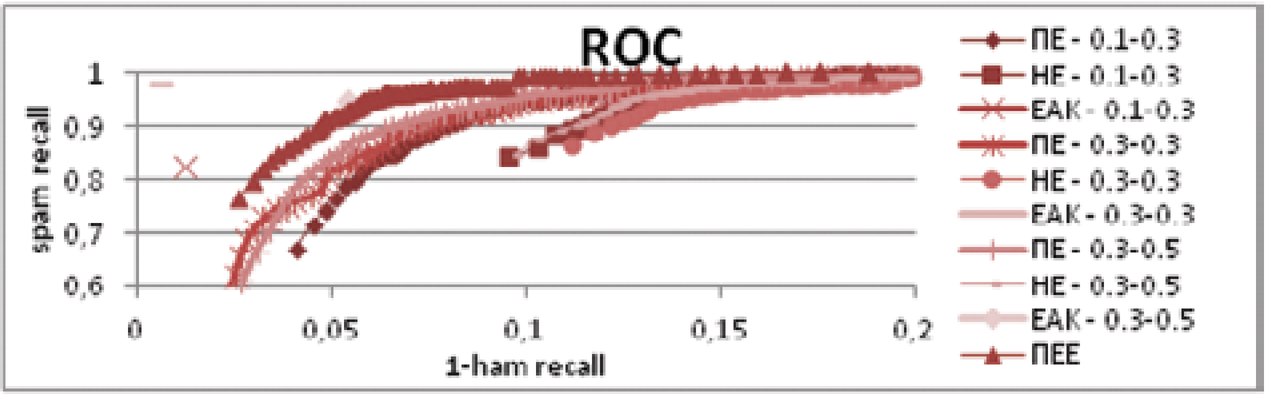
Ο παρακάτω πίνακας παρουσιάζει το ποσοστό ετικετών που ζητήθηκαν από τον αλγόριθμο για τα όρια 0.1 και 0.3, που ορίζονται στην αρχή από το χρήστη:

ham + spam	request for label	request for label in percentage
farmer-d +GP	1522	29.4%
kaminski-v +SH	293	5%
kitchen-l + BG	641	13%
williams-w3 + GP	393	7%
beck-s + SH	638	13%
lokay-m + BG	391	7%

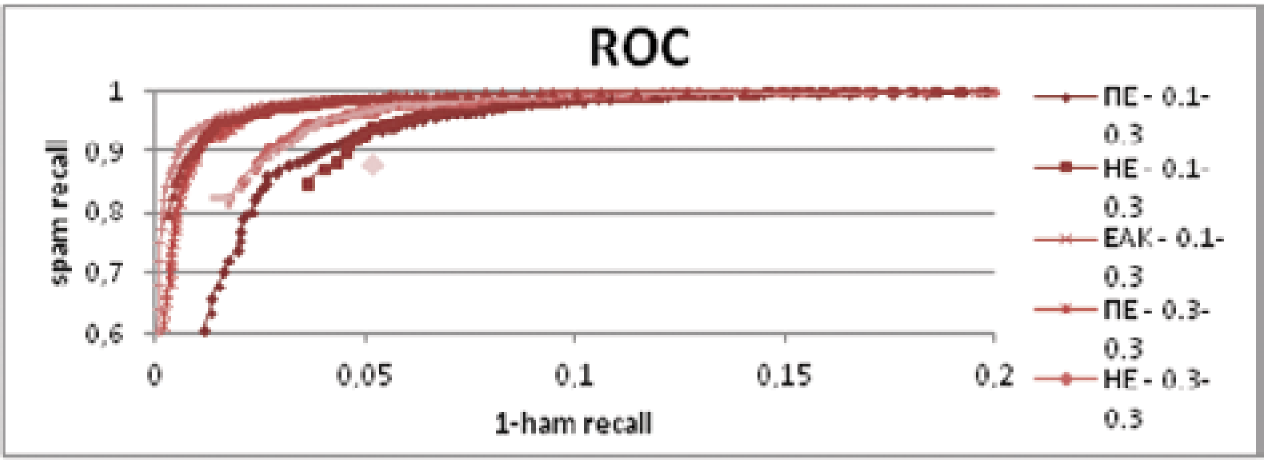
Πίνακας 1: Ετικέτες που ζητήθηκαν για κάθε συλλογή με όρια 0.1 και 0.3

Ενδιαφέρον παρουσιάζουν οι δύο πρώτες συλλογές, όπου η μέθοδος περιορισμένης εκπαίδευσης τείνει να έχει εξίσου καλά αποτελέσματα με τη μέθοδο πλήρους επιβλεπόμενης μάθησης. Στην πρώτη συλλογή (Σχήμα 1), παρουσιάζεται το μέγιστο ποσοστό ετικετών που έχει ζητηθεί από τον αλγόριθμο 29.4%, ενώ στην δεύτερη συλλογή (Σχήμα 2) παρουσιάζεται το ελάχιστο ποσοστό ετικετών 5%. Η ιδιαιτερότητα που παρουσίασαν οι δύο πρώτες συλλογές μας ώθησε να πραγματοποιήσουμε περεταίρω μετρήσεις, επιλέγοντας διαφορετικά άνω και κάτω όρια. Στην επόμενη συλλογή (Σχήμα 3), παρατηρούμε πως η μέθοδος ημιεπιβλεπόμενης εκπαίδευσης έχει εξίσου καλά αποτελέσματα με τη μέθοδο πλήρους επιβλεπόμενης εκπαίδευσης. Το ίδιο συμβαίνει και στις τρεις επόμενες συλλογές, που παρουσιάζουν το ίδιο διάγραμμα ROC με τη συλλογή 3 και δεν παρουσιάζονται εδώ για εξοικονόμηση χώρου. Τα ποσοστά των ετικετών που έχουν ζητηθεί από τον ταξινομητή κυμαίνονται μεταξύ 7% και 13%, πράγμα που στηρίζει την αρχική μας υπόθεση ότι ο ταξινομητής μπορεί να αποδώσει εξίσου καλά όταν χρησιμοποιεί ένα 10% των ετικετών από το σύνολο των μηνυμάτων.

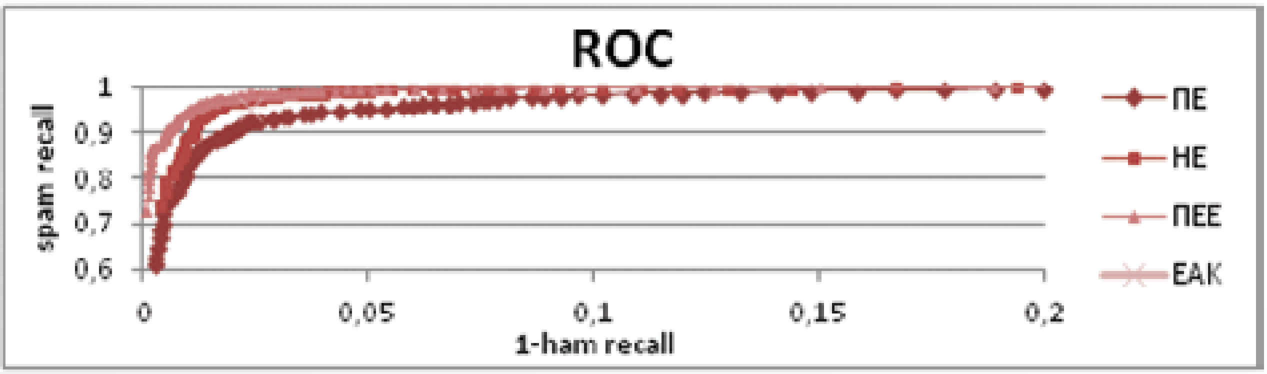
Στην πρώτη συλλογή, το ποσοστό των ετικετών που ζητούνται από το χρήστη ξεπερνάει κατά πολύ το 10% που ήταν ο αρχικός στόχος της εργασίας. Αντίθετα, στη συλλογή 2, το ποσοστό των μηνυμάτων που ζητήσαμε να μας αποκαλυφθεί η ετικέτα τους είναι μόλις 5%. Στις 4 επόμενες συλλογές, επιλέγεται ένα ποσοστό κοντά στο 10% των μηνυμάτων. Για την ακρίβεια, η μέθοδος ημιεπιβλεπόμενης εκπαίδευσης είναι εξίσου καλή με τη μέθοδο πλήρους επιβλεπόμενης εκπαίδευσης.



Σχήμα 1: ROC για την συλλογή farmer-d +GP με διαφορετικά όρια



Σχήμα 2: ROC για την συλλογή kaminsky-v +SH με διαφορετικά όρια



Σχήμα 3: ROC για συλλογή kitchen-l + BG

Σχήμα 4: Συστάδες μηνυμάτων ανά ημέρα

Η αναλογία μεταξύ επιθυμητών και μη επιθυμητών μηνυμάτων, φαίνεται επίσης να επηρεάζει την αναλογία μεταξύ των ham και spam συστάδων. Σε όλες τις συλλογές, το πλήθος των συστάδων που περιέχουν μηνύματα μη επιθυμητής αλληλογραφίας είναι μεγαλύτερο από των επιθυμητών μηνυμάτων. Αυτό οφείλεται στο γεγονός ότι τα μηνύματα μη επιθυμητής αλληλογραφίας, συνήθως, ποικίλουν ως προς το περιεχόμενο.

4. Συμπεράσματα

Από τα πειράματα που πραγματοποιήθηκαν, εξάγαμε τα εξής συμπεράσματα για τους αλγόριθμους που κατασκευάσαμε:

- Ο αλγόριθμος ημιεπιβλεπόμενης εκπαίδευσης καταφέρνει να αποδώσει εξίσου καλά με τον αλγόριθμο πλήρους επιβλεπόμενης εκπαίδευσης, ζητώντας μόνο ετικέτες για 7-10% του συνόλου των μηνυμάτων. Ο αλγόριθμος περιορισμένης εκπαίδευσης αγγίζει την απόδοση του αλγόριθμου πλήρους επιβλεπόμενης εκπαίδευσης, όταν το ποσοστό των ετικετών πλησίασε το 30%.

- Σημαντική παρατήρηση είναι πως όσο περισσότερα μη επιθυμητά μηνύματα λαμβάνουμε, τόσο περισσότερες συστάδες τέτοιων μηνυμάτων δημιουργούνται, γεγονός που οφείλεται στην ποικιλία της θεματολογίας των μηνυμάτων αυτών. Παρατηρήσαμε πως αλλάζοντας τα όρια στις 2 πρώτες συλλογές, η απόδοση της μεθόδου περιορισμένης εκπαίδευσης έφτασε την απόδοση της πλήρους επιβλεπόμενης εκπαίδευσης, αυξάνοντας τον αριθμό των συστάδων που δημιουργούνται.

Αναφορές

1. G. V. Cormack, «Foundations and Trends in Information Retrieval,» Email Spam Filtering: A Systematic Review, vol. 1, Απρίλιος 2007 2007.

2. K. Tretyakor, «Machine Learning Techniques in Spam Filtering,» in Data In Problem-oriented Seminar, 2004, pp. 60-79

3. G. Sakkis, «A Memory - Based Approach to Anti-Spam Filtering for Mailing Lists,» Information Retrieval, vol. 6, 2001.

4. Greg Schohn and D. Cohn, «Less is more: Active learning with support vector machines,» presented at the Proc. 17th International Conference on Machine Learning (ICML '00), San Francisco, CA, USA, 2000.

5. S. Tong, «Active Learning : Theory and Applications,» doctoral dissertation, Dept. Computer Science, Stanford University, California, 2001.

6. Andrew McCallum and K. Nigam, «Employing EM and Pool-Based Active Learning for Text Classification,» presented at the ICML '98 Proceedings of the Fifteenth International Conference on Machine Learning, San Francisco, CA, USA, 1998

7. D. H.Fisher, «Knowledge Acquisition Via Incremental Conceptual Clustering,» Machine Learning, vol. 2, 1987.

8. Ά. Κοσμόπουλος, «Διήθηση ανεπιθύμητης ηλεκτρονικής αλληλογραφίας με διαφορές μορφές του απλοϊκού ταξινομητή Bayes και διαμοιρασμό φίλτρων μεταξύ χρηστών,» Διπλωματική Εργασία, Τμήμα Επιστήμης των Υπολογιστών, Οικονομικό Πανεπιστήμιο Αθηνών, 2007

9. Dan Simovici , et al., «Metric Incremental Clustering of Nominal Data,» in Proc. 4th IEEE International Conference on Data Mining(ICDM '04), Washington, DC, USA, 2004

10. V. Metsis, «Spam Filtering with Naive Bayes -- Which Naive Bayes?,» in Proc. 3rd Conference on Email and Anti-Spam (CEAS), 2006.

Εξόρυξη Κοινωνικού Δικτύου από Διαδικτυακές Συζητήσεις με Χρήση Δομικών Χαρακτηριστικών

Θοδωρής Γεωργίου
t.georgiou@di.uoa.gr

Μάνος Καρβούνης
manosk@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Στην παρούσα εργασία προτείνουμε μια μέθοδο για τον προσδιορισμό σχέσεων συμφωνίας/διαφωνίας μεταξύ συνομιλητών, σε διαδικτυακά φόρα (web forum), και την ομαδοποίηση εκείνων που συμφωνούν μεταξύ τους. Για να μεγιστοποιήσουμε τη συνοχή των σχέσεων αυτών, προτείνουμε μια καινοτόμο μέθοδο για την ανάλυση της εξέλιξης του θέματος μιας συζήτησης, καθώς, σε πολλές περιπτώσεις, οι σχέσεις συμφωνίας/διαφωνίας αλλάζουν με την πρόοδο της συζήτησης. Η προσέγγισή μας βασίζεται αποκλειστικά σε δομικά στοιχεία, χωρίς εξόρυξη του κειμένου. Τα αποτελέσματα είναι πολύ ενθαρρυντικά και δείχνουν τις μεγάλες δυνατότητες της συνολικής προσέγγισης και των προτεινόμενων αλγορίθμων.

Λέξεις-Κλειδιά: Εξόρυξη δομής, Κατάτμηση συζήτησης, Εκτίμηση διαφωνιών, Κοινωνικά δίκτυα, Διαδικτυακά φόρα

Βραβείο "Best Undergraduate Poster" [11][12] στο συνέδριο SIGMOD 2010 και Βραβείο Καλύτερης Εργασίας [13] στο Πρώτο Διεθνές Φοιτητικό Συνέδριο Eureka! 2010.

Επιβλέπων: Ιωάννης Ιωαννίδης, Καθηγητής

1. Εισαγωγή

Οι διαδικτυακές συζητήσεις έχουν εξελιχθεί σε ένα αναπόσπαστο στοιχείο της καθημερινότητας του κοινωνικού ιστού. Στόχος της δουλειάς μας, είναι να εξάγουμε το πλούσιο δίκτυο σχέσεων συμφωνίας και διαφωνίας που σχηματίζονται κατά τη διάρκεια συζητήσεων πάνω σε «καυτά» θέματα. Με αυτό ως στόχο, προτείνουμε μια μέθοδο για την αναγνώριση των σχέσεων μεταξύ όλων των συνομιλητών και τη δημιουργία ομάδων των οποίων τα μέλη μοιράζονται όμοιες θέσεις και απόψεις. Πιστεύουμε ακράδαντα ότι η αυτόματη εξαγωγή ενός τέτοιου δικτύου μπορεί να εμπλουτίσει και να ωφελήσει την κατανόηση που έχουμε για το πως εξελίσσονται οι συζητήσεις με διάφορους τρόπους, όπως:

- Σαν βάση μιας βαθύτερης ανάλυσης του εξαγόμενου δικτύου: κεντρικότητα (centrality), «betweeness», γεφύρωση ακμών (bridge edges) κλπ. με στόχο την αναγνώριση βασικών χαρακτηριστικών μιας συζήτησης.
- Στα πλαίσια προ-επεξεργασίας για μια τεχνική Ανάλυσης Φυσικής Γλώσσας (Natural Language Processing - NLP) για την εξαγωγή της γνώμης των συμμετεχόντων. Είναι προφανές, ότι η διαδικασία εξόρυξης απόψεων γίνεται πολύ πιο εύκολη δεδομένου των σχέσεων συμφωνίας και διαφωνίας [8].
- Σαν ένας μηχανισμός επισκόπησης, για την αυτόματη δημιουργία περιλήψεων των συζητήσεων, όπου ο τελικός χρήστης διαβάζει αναρτήσεις συνομιλητών που διαφωνούν μεταξύ τους ώστε να έχει επίγνωση όλων των διαφορετικών απόψεων.

Αν και οι βασικές ιδέες της μεθόδου μας μπορούν να εφαρμοστούν σε πληθώρα μορφών συζήτησης, στην παρούσα δουλειά επικεντρωθήκαμε συγκεκριμένα σε συζητήσεις που λαμβάνουν χώρο σε διαδικτυακά φόρα (ή φόρουμς). Τέτοιου είδους συζητήσεις αναπτύσσονται στα πλαίσια αυτόνομων νημάτων (threads) τα οποία χαρακτηρίζονται από τον τίτλο τους. Τα δομικά στοιχεία ενός νήματος είναι οι αναρτήσεις οι οποίες χαρακτηρίζονται από το χρήστη που τις έκανε και την ημερομηνία που γίνανε. Ομοίως με τις ζωντανές συζητήσεις, το θέμα ενός νήματος μπορεί εύκολα να παρεκτραπεί από το αρχικό και νέες (διαφορετικές) σχέσεις μεταξύ των συμμετεχόντων μπορεί να δημιουργηθούν. Μια λίστα με τα μεγαλύτερα διαδικτυακά φόρα μπορεί να προσπελαστεί στο σύνδεσμο [10].

1.1 Προγενέστερες και σχετικές εργασίες

Η δουλειά των Agrawal et al. [2] πάνω στην εξόρυξη συζητήσεων σε Usenet

Newsgroups (διαδικτυακός ισότοπος με συζητήσεις πάνω σε θέματα της επικαιρότητας) είναι άμεσα σχετική με τη δική μας. Οι συγγραφείς προτείνουν μια μέθοδο για το διαχωρισμό των συμμετεχόντων σε δύο αντικρουόμενες ομάδες μέσω τεχνικών διαμερισμού γράφων. Αν και γίνεται χρήση μόνο δομικών στοιχείων η δουλειά τους διαφέρει από τη δική μας καθώς η μέθοδός μας μπορεί να εντοπίσει όλες τις πιθανές ομάδες, χωρίς περιορισμό στον αριθμό τους, καθώς και το σύνολο των σχέσεων μεταξύ τους. Βασιζόμενοι στην προαναφερθείσα δουλειά, οι Mukari και Raymond βελτίωσαν τη μέθοδο κάνοντας χρήση και του κειμένου των συζητήσεων [6]. Στη συνέχεια, οι Somasundaran και Wiebe καταφέρνουν παρόμοια αποτελέσμα κάνοντας, όμως, αποκλειστική χρήση του κειμένου, αγνοώντας κάθε μορφή δομικής πληροφορίας [7]. Παρόμοια δουλειά αποτελεί και των Xu et al. [9] που αναγνωρίζουν οποιοδήποτε πλήθος ομάδων αλλά θεωρούν τις σχέσεις συμφωνίας/διαφωνίας δεδομένες. Οι υπόλοιπες σχετικές εργασίες για τον εντοπισμό τέτοιων σχέσεων επικεντρώνονται καθαρά σε τεχνικές επεξεργασίας φυσικής γλώσσας.

Προηγούμενες δουλειές πάνω στην κατάτμηση συζητήσεων έχει γίνει από τους Kim et al. [3] και Adams et al. [1], οι οποίοι είναι και οι πρώτοι, απ' όσο ξέρουμε, που χρησιμοποιούν δομικά χαρακτηριστικά για να ομαδοποιήσουν αναρτήσεις σε θεματικές περιοχές. Η μέθοδός μας, παρόλα αυτά, διαφέρει στον τρόπο που γίνεται η τελική κατάτμηση (διακριτότητα υποσυζητήσεων) καθώς και στις ευριστικές που χρησιμοποιούμε για να αναγνωρίσουμε σε ποια σημεία αλλάζει το θέμα. Τέλος, στην εργασία των Lin et al. [5] γίνεται, επίσης για πρώτη φορά, ο συσχετισμός μεταξύ θέματος και δομής της συζήτησης (μοντέλο SMSS).

1.2 Συνεισφορά

Η καθαρά δομική προσέγγισή μας, έναντι αυτών που κάνουν μόνο χρήση επεξεργασίας φυσικής γλώσσας ή υβριδικών προσεγγίσεων, παρέχει τα εξής πλεονεκτήματα:

- Είναι τελείως ανεξάρτητη από την γλώσσα στην οποία γράφουν οι συμμετέχοντες. (Μπορέσαμε και εφαρμόσαμε τη μέθοδό μας και σε ελληνικά αλλά και σε αγγλικά διαδικτυακά φόρουμ χωρίς καμία αλλαγή.)
- Το κείμενο των αναρτήσεων δεν απαιτείται, λύνοντας έτσι σοβαρά θέματα ιδιωτικότητας (privacy).

Η βασική συνεισφορά της δουλειά μας αποτελεί την ένωση, σε ένα κοινό σύστημα, δύο επιμέρους περιοχών:

— Αναγνώριση εξέλιξης του θέματος συζήτησης (και κατάτμησης στα σημεία αλλαγής).

— Εκτίμηση διαφωνίας και συμφωνίας μεταξύ ομάδων συνομιλητών.

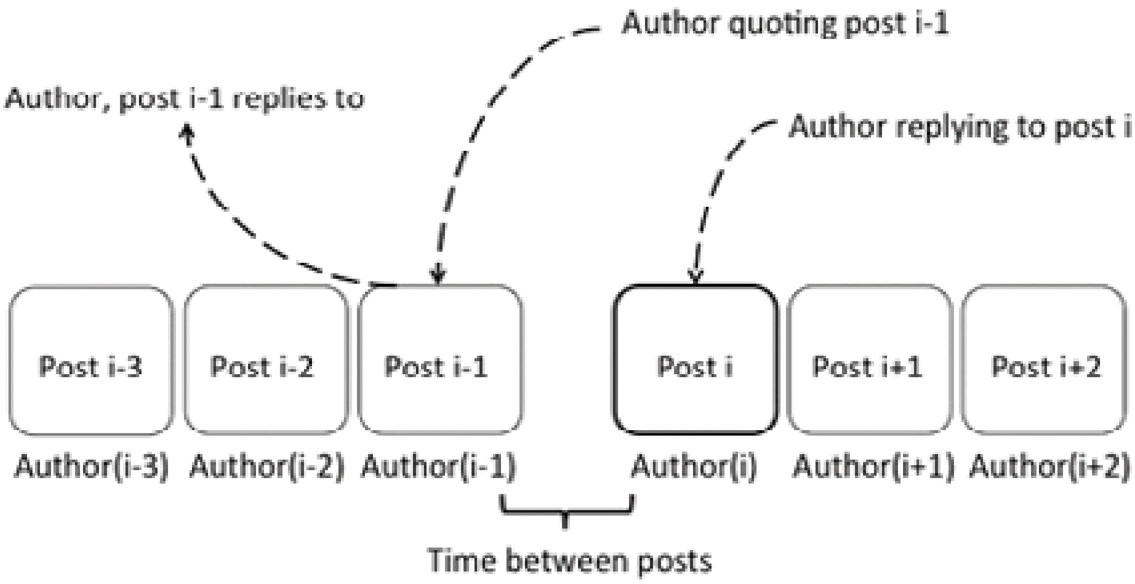
Επιπλέον, εκμεταλλευόμαστε τα πρότυπα κοινωνικής συμπεριφοράς (social behavior patterns) που παρατηρούνται στα πλαίσια μιας συζήτησης (ενότητες 2.1 και 2.2) για τη δημιουργία ευριστικών αλγορίθμων που εξάγουν την επιθυμητή πληροφορία μόνο από δομικά χαρακτηριστικά.

2. Μεθοδολογία

Στην καρδιά της μεθόδου μας βρίσκεται η εξαγωγή (και βελτίωση) της δομής των απαντήσεων στη μορφή ενός δέντρου (δέντρο απαντήσεων). Ως απάντηση θεωρούμε μια ανάρτηση ενός χρήστη η οποία αποτελεί την απάντησή του σε κάποια προγενέστερη ανάρτηση. Οι κόμβοι του δέντρου αποτελούν αναρτήσεις του νήματος και οι ακμές υποδεικνύουν σχέσεις απάντησης μεταξύ δύο αναρτήσεων. Όλα τα δομικά χαρακτηριστικά που χρησιμοποιούμε προέρχονται από το δέντρο αυτό.

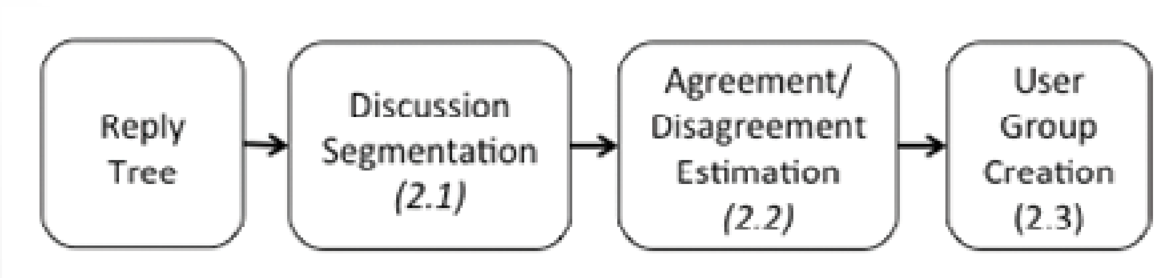
Η δημιουργία του δέντρου χωρίζεται σε δύο μέρη: α) τη συλλογή των φανερών σχέσεων απάντησης και β) την εξαγωγή των κρυφών σχέσεων απάντησης. Φανερή σχέση απάντησης δημιουργείται από μια ανάρτηση η οποία παραθέτει (quote) μέρος της αρχικής ανάρτησης στην οποία απαντάει, και τα διαδικτυακά φόρα έχουν ειδικό μηχανισμό για την υποστήριξή της. Ο ίδιος μηχανισμός μας επιτρέπει να συλλέξουμε αυτή την πληροφορία από τη σελίδα του φόρουμ, χωρίς να απαιτείται το κείμενο της ανάρτησης.

Για να είναι το δέντρο όσο πιο πλήρες γίνεται, αναγνωρίζουμε επίσης και ένα σύνολο από σχέσεις απαντήσεων που δεν είναι φανερές, δηλαδή, δεν έχουν γίνει με παράθεση. Εκπαιδεύσαμε ένα ταξινομητή διανύσματος υποστήριξης (Support Vector Machine - SVM) ώστε να ταξινομούμε ορφανές αναρτήσεις σε δύο κατηγορίες: αυτές που (α) δημιουργούν σχέση απάντησης (β) δεν δημιουργούν σχέση απάντησης. Μια ανάρτηση είναι ορφανή όταν δεν παραθέτει καμία άλλη, αλλά μπορεί με πιθανότητα 0.5 να απαντάει στην ακριβώς προηγούμενη. Συνολικά, σε ένα νήμα, το 10% των αναρτήσεων είναι ορφανές και οι μισές από αυτές απαντάνε στις ακριβώς προηγούμενες. Εφαρμόζοντας τον ταξινομητή σε όλες τις ορφανές αναρτήσεις αναγνωρίζουμε τις πιθανές σχέσεις απάντησης με επιτυχία 75,4%. Τα χαρακτηριστικά (features) εκπαίδευσης του ταξινομητή είναι επίσης δομικά και αφορούν τον εντοπισμό αλληλουχιών αναρτήσεων που καταλήγουν σε σχέσεις απάντησης. Στο Σχήμα 1 παρουσιάζουμε τα χαρακτηριστικά αυτά.



Σχήμα 1: Για μια ορφανή ανάρτηση i, ελέγχουμε την τοπική δομή συζήτησης (ποιος απαντάει σε ποιόν). Από τα εικονιζόμενα χαρακτηριστικά, δεν έγινε χρήση του χρόνου μεταξύ των αναρτήσεων, καθώς αποδείχθηκε παραπλανητικός για τη συγκεκριμένη ταξινόμηση.

Δεδομένου του δέντρου απαντήσεων η μέθοδος μας διαιρείται σε τρία επιμέρους βήματα, όπως παρουσιάζονται στο Σχήμα 2.



Σχήμα 2: Διάγραμμα ροής της μεθόδου μας: Μετά τη δημιουργία του δέντρου απαντήσεων, η συζήτηση κατακερματίζεται σε υποσυζητήσεις και σε κάθε υποσυζήτηση εξάγουμε τις σχέσεις συμφωνίας/διαφωνίας. Τέλος, οι χρήστες ομαδοποιούνται με βάση τις σχέσεις αυτές.

Περιγράφουμε αναλυτικά τα επιμέρους βήματα στις ακόλουθες υποενότητες 2.1, 2.2 και 2.3.

2.1 Κατάτμηση συζήτησης

Στην ενότητα αυτή παρουσιάζουμε τη μέθοδο και τις παρατηρήσεις/παραδοχές που κάναμε για την κατάτμηση μιας συζήτησης σε επιμέρους υποσυζητήσεις ώστε η κάθε μία να περιγράφεται από μόνο ένα και συγκεκριμένο θέμα.

1. **Παράλληλες Συζητήσεις:** Τα θέματα μιας συζήτησης σε μία συγκεκριμένη χρονική στιγμή δεν καλύπτουν περισσότερες από μία ή δύο θεματικές ενότητες [5]. Επιπλέον, για ένα συγκεκριμένο θέμα, οι αναρτήσεις που περιέχουν τη σχετική συζήτηση είναι συνήθως συγκεντρωμένες μαζί μέσα στο νήμα [1], κάτι που είναι φυσικό επακόλουθο της δυσκολίας να διατηρηθούν παράλληλες συζητήσεις για διαφορετικά θέματα. Τελικά, κάποια από τα «συναγωνιζόμενα» θέματα θα υπερισχύσει.

2. **Συμμετέχοντες:** Όταν σε μια συγκεκριμένη χρονική στιγμή υπάρχει μεγάλη συμμετοχή συνομιλητών, η πιθανότητα να αλλάξει το θέμα είναι μεγαλύτερη. Το αντίθετο ισχύει επίσης: Όταν ένα μικρό πλήθος χρηστών μόνο συμμετέχει η συζήτηση είναι πιο πιθανό να διατηρήσει το ίδιο θέμα [εμπειρική παρατήρηση].

3. **Σχέσεις Απάντησης:** Όταν το θέμα μιας συζήτησης αλλάζει, είναι λιγότερο πιθανό ένας χρήστης να απαντήσει σε μια ανάρτηση από τα προηγούμενα θέματα.

Το πρόβλημα της κατάτμησης συζητήσεων μπορεί να λυθεί προσεγγιστικά αναγνωρίζοντας όλα τα σημεία μέσα στη συζήτηση όπου η συμμετοχή είναι υψηλή και οι σχέσεις απάντησης σύντομες. Δεδομένου ενός συνόλου από σχέσεις απαντήσεων C, ορίζουμε ως Πυκνότητα Απαντήσεων (Post Connection Density - PCD) την ποσότητα:

$$pcd_x = \sum_{i=x}^{(x+\lambda)} connectionExists_x\,(\mathrm{i})\,|_{True} \qquad (1)$$

όπου x είναι ο αριθμός της ανάρτησης στην οποία υπολογίζουμε τη πυκνότητα, με $x < N$, και N ο συνολικός αριθμός αναρτήσεων του νήματος. Η συνάρτηση $connectionExists(.)$ ορίζεται ως εξής:

$$connectionExists_x\,(\mathrm{i}) = (\exists\, j)(x - \kappa \leq x) \cap C\,(\,i, j\,) \qquad (2)$$

Οι τιμές των κ και λ είναι εμπειρικές παράμετροι και καθορίζουν πόσο «πίσω» στη συζήτηση κοιτάμε για απαντήσεις. Πρακτικά, για μια ανάρτηση x μετράμε πόσες αναρτήσεις από τις επόμενες λ απαντάνε σε αναρτήσεις μεταξύ των κ προηγούμενων.

Επίσης, ορίζουμε ως Πυκνότητα Συνομιλητών (Discussant Density - DD):

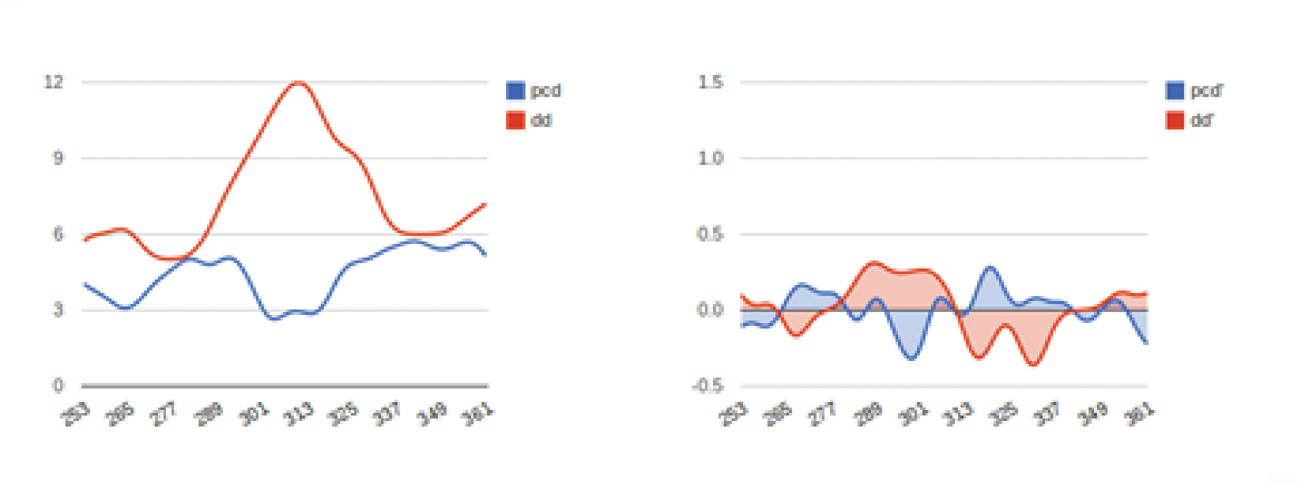
$$dd_x = |\{Author_i \mid i \in [x - \kappa,\ x + \lambda]\}| \qquad (3)$$

όπου σαν $Author_i$ συμβολίζουμε τον συνομιλητή που έκανε την ανάρτηση i. Και εδώ, μετράμε πόσοι χρήστες κάνανε έστω και μια ανάρτηση στο διάστημα των $[x - \kappa,\ x + \lambda]$αναρτήσεων.

Για τον εντοπισμό των τοπικών μέγιστων και ελάχιστων υπολογίζουμε τις παραγώγους των δύο πυκνοτήτων PCD και DD. Ενδιαφερόμαστε για περιοχές στη συζήτηση στις οποίες ένα τοπικό μέγιστο της Πυκνότητας Συνομιλητών (1) είναι πολύ κοντά (δ αναρτήσεις) σε ένα τοπικό ελάχιστο της Πυκνότητας Απαντήσεων (3). Δεδομένου ότι και οι δύο πυκνότητες είναι διακριτές μεταβλητές ορίζουμε την παράγωγό f' σε μία ανάρτηση x ως εξής:

$$f'_x = \frac{\Delta f}{\Delta x} = \frac{f_x - f_{x-1}}{x-(x-1)} = f_x - f_{x-1} \qquad (4)$$

Στο Σχήμα 3 παραθέτουμε ένα παράδειγμα κανονικοποιημένων πυκνοτήτων και των παραγώγων τους. Η μέθοδός μας αναγνωρίζει ότι στην περιοχή μεταξύ των αναρτήσεων 303 και 317 υπάρχει αλλαγή του θέματος. Η εξόρυξη των περιοχών αυτών έχει γραμμική πολυπλοκότητα και αγνοεί τα (σχετικά) σύντομα τοπικά μέγιστα/ελάχιστα.



Σχήμα 3: Παράδειγμα πυκνοτήτων και των παραγώγων τους σε μια πραγματική συζήτηση, μεταξύ των αναρτήσεων 253 και 361.

2.2 Αναγνώριση συμφωνίας/διαφωνίας

Για κάθε ζευγάρι συνομιλητών ορίζουμε ένα διάνυσμα το οποίο μετά ταξινομούμε σαν σχέση συμφωνίας ή διαφωνίας με χρήση ενός εκπαιδευμένου ταξινομητή. Αν και υπάρχει πάντα η πιθανότητα για ουδέτερες σχέσεις (ούτε συμφωνίας ούτε διαφωνίας) επιλέξαμε να μην εισάγουμε τέτοια κλάση στην ταξινόμηση. Οι κυριότεροι λόγοι για την απόφαση αυτή είναι οι εξής: (α) καθιστά την ταξινόμηση πιο εύκολη, και, το πιο σημαντικό, (β) μπορεί να θεωρηθεί ότι όλοι οι συμμετέχοντες στη συζήτηση έχουν κάποιο επίπεδο συμφωνίας ή διαφωνίας με όλους τους άλλους δεδομένου ότι εστιάζουμε σε ένα συγκεκριμένο θέμα (μετά την κατάτμηση) και όλοι οι συνομιλητές στο συγκεκριμένο τμήμα εκφράζουν την άποψή τους.

Η βασική υπόθεση/παρατήρηση στην οποία στηριζόμαστε για τη δημιουργία των χαρακτηριστικών των διανυσμάτων είναι η εξής:

- Κατά τη διάρκεια μιας συζήτησης, ένας συμμετέχων θα αφιερώσει τους περισσότερους από τους περιορισμένους πόρους του (χρόνος, προσπάθεια) αλληλεπιδρώντας με τους χρήστες που απειλούν τα επιχειρήματα του, δηλαδή, με αυτούς με τους οποίους διαφωνεί. (Παρόμοιες ιδέες έχουν ήδη χρησιμοποιηθεί με επιτυχία σε προηγούμενες εργασίες [2]).

Για τη δημιουργία των χαρακτηριστικών (συνιστώσες του διανύσματος) ορίζουμε τις εξής βοηθητικές συναρτήσεις:

— $replies(ui, uj)$: Ο αριθμός των αναρτήσεων του χρήστη ui που απαντούν σε αναρτήσεις του χρήστη uj.

— $posts(u)$: Ο αριθμός των αναρτήσεων του χρήστη u.

Κάθε διάνυσμα αποτελείται από τις εξής τρεις συνιστώσες:

Ομοιότητα Στόχων - Target Similarity

$$targetSimilarity\ (u_i, u_j) = cosineSimilarity\ (targets\ (u_i), targets\ (u_j)) \quad (5)$$

όπου $targets\ (.)$ είναι το διάνυσμα:

$$targets(u_i) = \left(\frac{replies(u_i, u_1)}{posts(u_i)}, \ldots, \frac{replies(u_i, u_n)}{posts(u_i)} \right) \quad (6)$$

Προφανώς, εφόσον ένας συνομιλητής δεν μπορεί να απαντάει στον εαυτό του:

$$\frac{replies(u_i, u_i)}{posts(u_i)} = 0$$

Συγκέντρωση Απαντήσεων - Reply Focus

$$replyFocus\left(u_i, u_j\right) = \max\left(\frac{replies\left(u_i, u_j\right)}{posts\left(u_i\right)}, \frac{replies\left(u_j, u_i\right)}{posts\left(u_j\right)}\right) \tag{7}$$

Κάλυψη Απαντήσεων - Reply Coverage

$$replyCoverage\left(u_i, u_j\right) = \frac{\min\left(replies\left(u_i, u_j\right), replies\left(u_j, u_i\right)\right)}{\max\left(replies\left(u_i, u_j\right), replies\left(u_j, u_i\right)\right)} \tag{8}$$

2.3 Δημιουργία ομάδων

Στο τελικό βήμα χρησιμοποιούμε την εξαγόμενη πληροφορία με τις σχέσεις συμφωνίας/διαφωνίας μεταξύ των συνομιλητών για να δημιουργήσουμε ομάδες χρηστών που συμφωνούν μεταξύ τους. Το τελικό προϊόν της μεθόδου είναι (για κάθε υποσυζήτηση) ένας μη-κατευθυνόμενος γράφος όπου οι κόμβοι αποτελούν ομάδες χρηστών και οι ακμές υποδεικνύουν ποιες ομάδες συμφωνούν/διαφωνούν μεταξύ τους.

Αρχικά δημιουργούμε τον γράφο όπου κάθε συνομιλητής αποτελεί μια ομάδα (κόμβο) από μόνος του και οι κόμβοι ενώνονται μεταξύ τους με τις σχέσεις που έχουμε ήδη εξάγει. Για τη συγχώνευση κόμβων, εφαρμόζουμε την εξής άπληστη (greedy) μέθοδο: Ενώνουμε ζεύγη κόμβων οι οποίοι ενώνονται με ακμές συμφωνίας και των οποίων οι ακμές προς όλους τους άλλους κόμβους περιγράφονται με το ίδιο τύπο σχέσης (όπου η μια ομάδα έχει συμφωνία έχει και η άλλη, και το ίδιο για τη διαφωνία). Όταν πλέον δεν υπάρχουν κόμβοι που να μπορούν να συγχωνευθούν, ο αλγόριθμος τερματίζει.

Καθώς ο παραπάνω αλγόριθμος είναι άπληστος δεν καταλήγει αναγκαστικά στη βέλτιστη λύση. Για να το βελτιώσουμε, ορίζουμε μια μετρικής κέρδους (gain) ως το μέσο όρο των μεγεθών κάθε ομάδας. Στη συνέχεια τρέχουμε την άπληστη μέθοδο αρκετές φορές (όπου κάθε φορά η επιλογή των κόμβων για συγχώνευση είναι τυχαία) και κρατάμε τη λύση με το μεγαλύτερο κέρδος. Η ιδέα πίσω από τη μετρική είναι ότι προτιμούμε γράφους με λίγες ομάδες και μεγάλο πλήθος χρηστών ανά ομάδα.

3. Αξιολόγηση μεθόδου

Σαν προκαταρκτική αξιολόγηση της μεθόδου μας και των επιμέρους προτεινόμενων αλγορίθμων, και για να μπορέσουμε να εξετάσουμε ένα μεγάλο σύνολο από διαφορετικές δομές, δημιουργήσαμε συνθετικές συζητήσεις. Καθώς η μέθοδός μας χρησιμοποιεί μόνο δομικά χαρακτηριστικά αυτό ήταν εφικτό χωρίς να χρειάζεται δημιουργία κειμένου, αλλά μόνο δέντρων συζήτησης. Για αυτό το σκοπό, επεκτείναμε δύο μοντέλα συζητήσεων από τη βιβλιογραφία ([4], [5]) ώστε να δίνουμε σαν είσοδο (α) το μέγεθος μια συζήτησης, (β) τη κατανομή πιθανότητας το θέμα να αλλάξει, και (γ) τη κατανομή πιθανότητας δύο συνομιλητές να διαφωνούν/συμφωνούν.

Χρησιμοποιώντας το μοντέλο αυτό μπορέσαμε να δημιουργήσουμε ένα σύνολο από 11.000 συζητήσεις (3.250.000 αναρτήσεις) για την αξιολόγηση της μεθόδου κατάτμησης. Για τη μέθοδο εκτίμησης συμφωνιών/διαφωνιών δημιουργήσαμε (α) ένα μικρό σύνολο από συζητήσεις για την εκπαίδευση του ταξινομητή (συνολικά 80) και (β) ένα σύνολο από 14.400 συζητήσεις (1.440.000 αναρτήσεις) για την αξιολόγηση.

Οι μετρικές ακρίβειας (precision) και ανάκλησης (recall) φαίνονται στον παρακάτω πίνακα.

Αλγόριθμος	**Αξιολόγηση**
Κατάτμηση Συζήτησης	Precision: *82%* Recall: *75%*
Εκτίμηση Συμφωνίας Διαφωνίας	Precision: *76%*

Τα αποτελέσματα αυτά συμφωνούν με αξιολόγηση μικρής κλίμακας που κάναμε με δεδομένα πραγματικών συζητήσεων.

4. Μελλοντικές κατευθύνσεις

Ένα απαραίτητο βήμα για την ολοκλήρωση της παρούσας δουλειάς είναι η δημιουργία μιας συλλογής δεδομένων μεγάλου μεγέθους από πραγματικές συζητήσεις με τις κατάλληλες επισημειώσεις (labeled dataset). Απ' όσο γνωρίζουμε τέτοια δεδομένα δεν υπάρχουν, αλλά είναι αναγκαία για τη σωστή και πλήρη αξιολόγηση της μεθόδου μας.

Επίσης, αν και εστιάσαμε σε διαδικτυακά φόρα στη συγκεκριμένη δουλειά, επιθυμούμε να γενικεύσουμε τους αλγορίθμους μας ώστε να εφαρμόζονται σε κάθε είδους διαδικτυακή συζήτηση, όπως διαλόγους στα Twitter/Facebook, σχόλια σε Blogs, διαδικτυακές κουβέντες (web chats) κλπ.

5. Συμπεράσματα

Στην παρούσα εργασία προτείναμε μια καινοτόμο μέθοδο για την εξαγωγή σχέσεων διαφωνίας/συμφωνίας μεταξύ χρηστών, ακόμα και όταν αυτές αλλάζουν καθώς η συζήτηση εξελίσσεται. Οι προτεινόμενοι αλγόριθμοι κάνουν χρήση αποκλειστικά δομικής πληροφορία που εξάγεται από το Δέντρο Απαντήσεων μιας συζήτησης.

Η εκτενής αξιολόγηση της μεθόδου μας με συνθετικά δεδομένα προσφέρει πολύ ενθαρρυντικά αποτελέσματα και το επόμενο βήμα είναι μιας επίσης εκτενής αξιολόγηση σε πραγματικές συζητήσεις. Ελπίζουμε ότι η δουλειά μας θα μπορέσει να βοηθήσει στην εδραίωση της χρησιμότητας δομικών στοιχείων για εργασίες σε σχετικούς τομείς.

Αναφορές

1. Adams, P., Martell, C.: Topic Detection and Extraction in Chat. In Proceedings of the 2008 IEEE International Conference on Semantic Computing (ICSC ‘08). August 4-7, Santa Clara, California, USA.

2. Agrawal, R., Rajagopalan, S., Srikant, R., Xu., Y.: Mining newsgroups using networks arising from social behavior. In Proceedings of the 12th international conference on World Wide Web (WWW ‘03). ACM, New York, NY, USA, 869-875.

3. Kim, J., Candan, S., Donderler, M.: Topic Segmentation of Message Hierarchies for Indexing and Navigation Support. In Proceedings of the 14th international conference on World Wide Web (WWW ‘05). May 10-14, 2005, Chiba, Japan.

4. Kumar, R., Mahdian, M., McGlohon, M.: Dynamics of conversations. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD ‘10). ACM, New York, NY, USA, 553-562.

5. Lin, C., Yang, J., Cai, R., Wang, X., Wang, W., Zhang L.: Simultaneously Modeling Semantics and Structure of Threaded Discussions: A Sparse Coding Approach and Its Applications. In Proceedings of the 32nd international ACM SIGIR conference on research

and development in Information Retrieval (SIGIR '09). July 19-23, 2009, Boston, USA.

6. Murakami, A., Raymond, R.: Support or oppose?: Classifying positions in online debates from reply activities and opinion expressions. In Proceedings of the 23rd International Conference on Computational Linguistics: Posters (COLING '10). Association for Computational Linguistics, Stroudsburg, PA, USA, 869-875.

7. Somasundaran, S., Wiebe, J.: Recognizing stances in ideological on-line debates. In Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text (CAAGET '10). Association for Computational Linguistics, Stroudsburg, PA, USA, 116-124.

8. Tan, C., Lee, L., Tang, J., Jiang, L., Zhou, M., Li, P.: User-level sentiment analysis incorporating social networks. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge Discovery and data mining (KDD '11). ACM, New York, NY, USA, 1397-1405.

9. Xu, K., Li, J., Liao, S.: Sentiment community detection in social networks. In Proceedings of the 2011 iConference (iConference '11). ACM, New York, NY, USA, 804-805.

10. Top web forums, http://wikipedia.org/wiki/List_of_internet_forums

11. Georgiou, T., Karvounis, M., Ioannidis, Y.: Extracting Topics of Debate between Users on Web Discussion Boards. In Undergraduate Research Poster Competition of the ACM SIGMOD International Conference on Management of Data (SIGMOD '10), Indianapolis, Indiana, USA, June 2010.

12. Karvounis, M., Georgiou, T., Ioannidis, Y.: Estimating Agreement in Online Discussions. In Undergraduate Research Poster Competition of the ACM SIGMOD International Conference on Management of Data (SIGMOD '10), Indianapolis, Indiana, USA, June 2010.

13. M Karvounis, M., Georgiou, T., Ioannidis, Y.: Utilizing the Quoting System of Online Web Forums to Estimate User Agreement. Proceedings of the 1st International Conference for Undergraduate and Postgraduate Students in Computer Engineering, Informatics, related Technologies and Applications (EUREKA! '10), Patras, Greece, October 2010.

ARcropolis: Διαδικτυακό Εκπαιδευτικό Παιχνίδι Επαυξημένης Πραγματικότητας με Θέμα την Ακρόπολη Αθηνών

Εμμανουήλ Γιαννισάκης
em.giannisakis@gmail.com

Παναγής Παπαδάτος
panagis.papadatos@gmail.com

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Η παρούσα εργασία αποτελεί μια περιγραφή του σχεδιασμού και της υλοποίησης ενός διαδραστικού εκπαιδευτικού παιχνιδιού Επαυξημένης Πραγματικότητας, με γενικό θέμα την Ακρόπολη Αθηνών. Ο σκοπός του παιχνιδιού είναι τόσο ψυχαγωγικός, όσο και εκπαιδευτικός. Πρόκειται για ένα παιχνίδι γνώσεων που βασίζεται σε ιστορίες σχετικά με την Ακρόπολη Αθηνών, τις οποίες θα πρέπει να «συνθέτουν» οι παίκτες, σε συνεργασία μεταξύ τους. Ωστόσο, το πλαίσιο που σχεδιάστηκε μπορεί να χρησιμοποιηθεί και για άλλα –εκπαιδευτικά ή μη- παιχνίδια.

Λέξεις-Κλειδιά: Επαυξημένη Πραγματικότητα, Σχεδίαση Διεπαφών, Διαδικτυακή Εφαρμογή, Εκπαιδευτική Τεχνολογία, Ψηφιακά Παιχνίδια

Επιβλέποντες:
Ιωάννης Ιωαννίδης, Καθηγητής
Δρ. Μαρία Ρούσσου

1. Εισαγωγή

Το αντικείμενο της παρούσας εργασίας είναι η σχεδίαση και η ανάπτυξη ενός διαδικτυακού εκπαιδευτικού παιχνιδιού Επαυξημένης Πραγματικότητας (ΕΠ) [1]. Το παιχνίδι παίζεται από δύο παίκτες, σε δύο διαφορετικούς υπολογιστές, οι οποίοι είναι πιθανό να βρίσκονται σε διαφορετικούς χώρους και ίσως σε διαφορετικές πόλεις ή χώρες. Οι παίκτες, είναι δυνατόν, να είναι και διαφορετικής εθνικότητας, καθότι παρέχεται η δυνατότητα επιλογής γλώσσας. Οι διαθέσιμες γλώσσες του παιχνιδιού είναι προσωρινά τα Ελληνικά και τα Αγγλικά.

Το παιχνίδι απευθύνεται κυρίως σε παιδιά Δημοτικού σχολείου. Γι’ αυτό έχει σχεδιαστεί με τέτοιο τρόπο ώστε να προσελκύει το ενδιαφέρον των παιδιών με άμεσο και απλό τρόπο, παρέχοντας πολλά ελκυστικά γραφικά στοιχεία και εύκολα κατανοητή ανατροφοδότηση. Ο βασικός τρόπος αλληλεπίδρασης των παικτών και του υπολογιστή γίνεται μέσω συγκεκριμένων χάρτινων καρτών, τις οποίες οι παίκτες “δείχνουν” στην κάμερα του υπολογιστή τους (webcam). Βασικός γνώμονας για τη σχεδίαση της εφαρμογής ήταν οι αρχές ευχρηστίας.

Πρόθεσή μας ήταν η δημιουργία ενός παιχνιδιού που θα μπορεί να προφέρει έναν εναλλακτικό τρόπο μάθησης σε μαθητές Δημοτικού σχολείου, χρησιμοποιώντας νέες και μοντέρνες τεχνολογίες, όπως αυτή της ΕΠ.

2. Θεωρητικό υπόβαθρο

Η Επαυξημένη Πραγματικότητα (ΕΠ) (ή Augmented Reality – AR) αφορά στην άμεση ή έμμεση προβολή ενός φυσικού περιβάλλοντος του πραγματικού κόσμου, του οποίου τα στοιχεία έχουν επαυξηθεί με αισθητήρια είσοδο προερχόμενη από υπολογιστή, όπως ήχος ή γραφικά στοιχεία. Σχετίζεται με τη γενικότερη ιδέα που ονομάζεται διαμεσολαβημένη πραγματικότητα (mediated reality), στην οποία μια όψη της πραγματικότητας τροποποιείται (πιθανώς ίσως και να μειώνεται αντί να επαυξάνεται) από έναν υπολογιστή. Κατ’ αυτόν τον τρόπο, η τεχνολογία επηρεάζει την αντίληψη της πραγματικότητας.

Ο τρόπος με τον οποίο η ΕΠ χρησιμοποιήθηκε στη συγκεκριμένη εργασία είναι μέσω κάποιων ειδικών patterns (Εικόνα 1), δηλαδή κάποιων μοτίβων επάνω σε κάποιο υλικό, π.χ. εκτυπωμένων πάνω σε μια κόλλα χαρτί. Η είσοδος λαμβάνεται μέσω μίας κάμερας (webcam) και η έξοδος, δηλαδή η προβολή της εισόδου της εικόνας που υπάρχει στην οθόνη, επαυξάνεται αντικαθιστώντας το μοτίβο αυτό με κάποια δισδιάστατη εικόνα (Εικόνα 2).

Εικόνα 1: Pattern σε κάρτα

Εικόνα 2: Παράδειγμα Επαυξημένης Πραγματικότητας με κάμερα και κάρτα

Gameplay είναι το σύνολο των μεθόδων που χρησιμοποιούνται για να επιτευχθεί η διάδραση μεταξύ κάποιου αριθμού παικτών και ενός παιχνιδιού. Ορίζεται μέσω των κανόνων του παιχνιδιού, της διασύνδεσής του με τον παίκτη, των προκλήσεων που εμπεριέχει και της πλοκής του.

Βασικές λέξεις-κλειδιά είναι:

1. Η παρουσία ενός περιβάλλοντος
2. Η ικανότητα των αντικειμένων που ανήκουν στο περιβάλλον να αλλάζουν.
3. Κανόνες για την αλλαγή της κατάστασης των αντικειμένων, όπως η θέση τους, με βάση την κατάσταση άλλων αντικειμένων ή αποφάσεων των παικτών.
4. Οι ανταμοιβές και οι ποινές των παικτών, οι οποίες προκύπτουν από τις αλλαγές της κατάστασης του παιχνιδιού.
5. Οι παίκτες να έχουν την ικανότητα να αλληλεπιδρούν με το περιβάλλον και τα στοιχεία που βρίσκονται μέσα σε αυτό.

Ως **σοβαρό παιχνίδι** ορίζεται ένα παιχνίδι το οποίο έχει σχεδιαστεί με πρωταρχικό στόχο άλλον εκτός αυτού της καθαρής ψυχαγωγίας. Το επίθετο «σοβαρό» αναφέρεται, συνήθως, σε προϊόντα που χρησιμοποιούνται από βιομηχανίες όπως άμυνα, εκπαίδευση, επιστημονική εξερεύνηση, υγεία, διαχείριση κινδύνου, σχέδια πόλεων, μηχανική, θρησκεία και πολιτική. Τα σοβαρά παιχνίδια σχεδιάζονται με στόχο να λύσουν ένα πρόβλημα. Παρότι μπορεί να είναι διασκεδαστικά, ο κύριος σκοπός τους είναι να εκπαιδεύσουν, να ερευνήσουν ή να διαφημίσουν. Κάποιες φορές ένα παιχνίδι σκόπιμα «θυσιάζει» στοιχεία διασκέδασης και ψυχαγωγίας προκειμένου να κάνει κάποια σοβαρή δήλωση. Μια διάσημη υποκατηγορία, στην οποία εντάσσεται και η παρούσα εργασία, είναι τα εκπαιδευτικά διαδικτυακά παιχνίδια.

3. Περιγραφή του ARcropolis

3.1 Βασική ιδέα

Οι παίκτες καλούνται να συνθέσουν μαζί μια εικονογραφημένη ιστορία με θέμα την Ακρόπολη Αθηνών. Η σύνθεση πραγματοποιείται τοποθετώντας τα κομμάτια της ιστορίας στη σωστή σειρά, σαν πάζλ. Η ιστορία που χρησιμοποιήθηκε είναι ο μύθος της γέννησης της Αθηνάς και της Αθήνας. Μελλοντικά η «πλατφόρμα» θα μπορεί να δέχεται οποιαδήποτε ιστορία.

3.2 Φανταστικό πλαίσιο

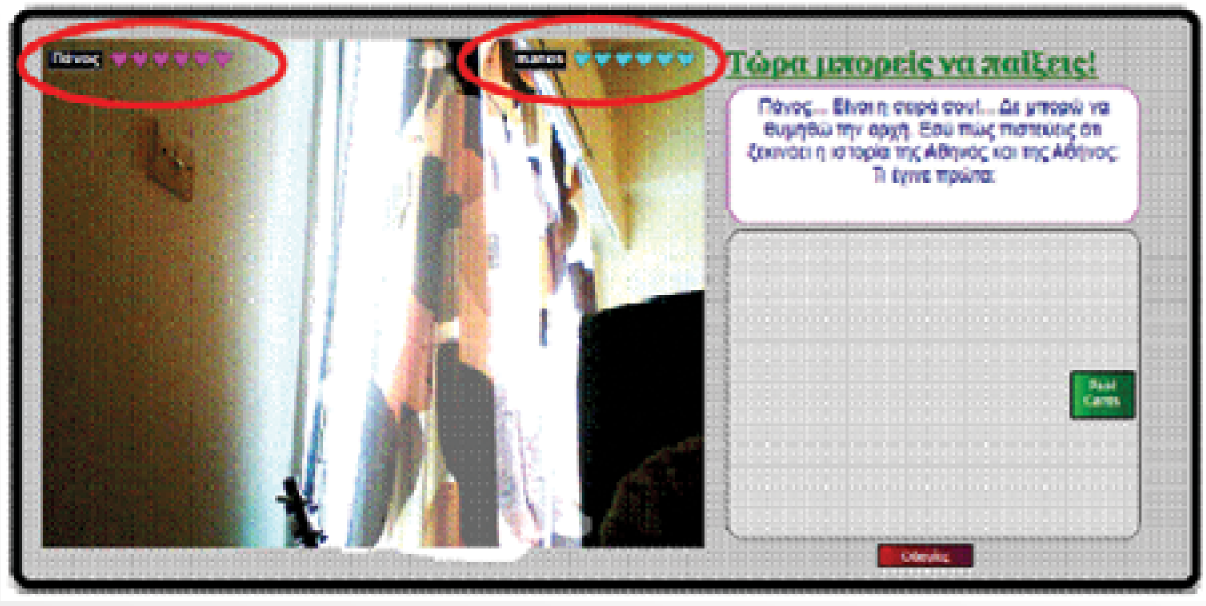
Ο βασικός φανταστικός χαρακτήρας είναι ένα «παραμυθάς». Ο παραμυθάς προσπαθεί να διηγηθεί μια ιστορία στους δύο παίκτες, αλλά δεν μπορεί να τη θυμηθεί σωστά. Στόχος των δύο παικτών (ως ακροατήριο) είναι να βοηθήσουν τον παραμυθά να θυμηθεί αυτή την ιστορία για να μπορέσει να τη διηγηθεί σωστά.

3.3 Φάσεις του παιχνιδιού

Το παιχνίδι χωρίζεται σε 2 φάσεις. Η πρώτη φάση είναι ένα προκαταρκτικό παιχνίδι το «Rock - Paper - Scissors - Lizard - Spock», το οποίο αποτελεί παραλλαγή του παιχνιδιού «πέτρα - ψαλίδι - χαρτί». Η φάση αυτή έχει τρεις βασικούς στόχους: Να αποφασιστεί ποιος θα παίξει πρώτος, να προκληθεί το ενδιαφέρον των παικτών με κάτι καθαρά ψυχαγωγικό, και τέλος οι παίκτες να αποκτήσουν μια πρώτη επαφή με την Επαυξημένη Πραγματικότητα.

Εικόνα 3: Οδηγίες Rock Paper Scissors Lizard Spock

Η δεύτερη φάση αποτελεί το βασικό παιχνίδι. Ο κάθε παίκτης έχει αρχικά έξι (6) «ζωές». Στη εκάστοτε σειρά του, ο κάθε παίκτης αρχικά μελετάει τις διαθέσιμες κάρτες και επιλέγει την κάρτα που πιστεύει ότι συνεχίζει σωστά την ιστορία. Αν κάνει λάθος χάνει μια ζωή, αλλά έχει τη δυνατότητα να ξαναπροσπαθήσει αμέσως. Αν μείνει χωρίς ζωές, χάνει το παιχνίδι. Εκτός από «χρονικά» λάθος κάρτες, υπάρχουν και κάποιες «κάρτες-παγίδες» οι οποίες δεν αντιστοιχούν σε κανένα κομμάτι της ιστορίας. Εάν επιλεγεί μία τέτοια κάρτα από κάποιον παίκτη, το αποτέλεσμα είναι πάλι να χάσει μία από τις ζωές του.



Εικόνα 4: Αρχή ARcropolis: Ζωές

Εικόνα 5: ARcropolis: Αναμονή του αντιπάλου

Το παιχνίδι τελειώνει όταν κάποιος παίκτης χάσει όλες τις ζωές του ή όταν η ιστορία ολοκληρωθεί επιτυχώς.

4. Κίνητρο

Βασικό κίνητρο για τη δημιουργία αυτού του παιχνιδιού ήταν το γεγονός ότι δεν καταφέραμε να βρούμε ένα παιχνίδι που να συνδυάζει όλα τα παραπάνω χαρακτηριστικά. Θεωρούμε πως η αλληλεπίδραση μέσω Επαυξημένης Πραγματικότητας, ενισχύει σημαντικά την εμπειρία χρήστη. Επιπλέον, η Επαυξημένη Πραγματικότητα, σε συνδυασμό με στοιχεία συναγωνισμού, ανταγωνισμού, φαντασίας και πρόκλησης, καθώς και το γεγονός ότι το παιχνίδι βρίσκεται στο διαδίκτυο, θα μπορούσε να προσελκύσει το ενδιαφέρον των μαθητών και να τους κάνει να θέλουν να συμμετέχουν ενεργά σε αυτό.

5. Υλοποίηση

Για την υλοποίηση του παιχνιδιού χρησιμοποιήσαμε το Adobe Flex SDK, στο οποίο χρησιμοποιούνται MXLML και Actionscript. Επίσης, χρησιμοποιήσαμε PHP, mySQL, HTML, CSS. Για το κομμάτι της Εικονικής Πραγματικότητας χρησιμοποιήσαμε το προγραμματιστικό πλαίσιο (framework) FLARmanager [5]. Το παιχνίδι για το έτος 2012 είναι διαθέσιμο στην ιστοσελίδα http://arcropolis.heypano.com.

6. Σχεδιασμός

Ο σχεδιασμός της διεπαφής του παιχνιδιού έγινε με γνώμονα των 10 αρχών ευχρηστίας κατά Nielsen [2], αλλά και των καλών πρακτικών στο σχεδιασμό εφαρμογών για περιβάλλοντα άτυπης εκπαίδευσης [3]. Ιδιαίτερη έμφαση δόθηκε στη μινιμαλιστική εικαστική σχεδίαση, καθότι η απόσπαση της προσοχής του χρήστη με γραφικά στοιχεία που δεν εξυπηρετούν κάποιο σκοπό, δεν ήταν κάτι επιθυμητό. Όπου κρίθηκε απαραίτητο, δόθηκε η ελευθερία στο χρήστη να επιστρέφει σε μία προηγούμενη κατάσταση. Παραδείγματος χάριν, ο χρήστης, ενώ βρίσκεται στην οθόνη που πρέπει να επιλέξει ένα δωμάτιο στο οποίο επιθυμεί να παίξει, έχει τη δυνατότητα να μεταφερθεί πίσω, στην αρχική οθόνη. Ωστόσο, από τη στιγμή που θα αρχίσει το παιχνίδι, δεν μπορεί να επιστρέψει σε κάποια προηγούμενη οθόνη. Μεγάλη προσοχή δόθηκε επίσης, τόσο στην ενημέρωση του παίκτη σχετικά με την κατάσταση του συστήματος, όσο και στην αποφυγή ανάγκης απομνημόνευσης πολλών πληροφοριών. Αυτό επιτεύχθηκε παρέχοντας στον παίκτη τη δυνατότητα να βρίσκει εύκολα όποια πληροφορία δεν θυμάται. Ειδικότερα, καθ’ όλη τη διάρκεια του παιχνιδιού, υπάρχει συνεχής ανατροφοδότηση για το τι συμβαίνει, ειδικά κουμπιά τα οποία παρέχουν στους παίκτες αναλυτικά τις οδηγίες, και δυνατότητα προβολής των καρτών που έχουν ήδη παιχτεί.

Ο σχεδιασμός του παιχνιδιού έγινε με τέτοιο τρόπο, ώστε να ικανοποιεί τα κριτήρια που τίθενται από τους Malone και Lepper [4] για τη δημιουργία του εγγενούς κινήτρου, εστιάζοντας κατά κύριο λόγο στα παρακάτω:

Πρόκληση: Το ARcropolis ζητά από τους παίκτες να λύσουν ένα πρόβλημα του οποίου τη λύση θα πρέπει, σε πολλά σημεία, να εικάσουν. Η επιτυχία δεν είναι βέβαιη, καθώς ο αριθμός των ζωών είναι περιορισμένος.

Περιέργεια: Το ARcropolis κάνει τους παίκτες να αναρωτιούνται για την πλοκή της ιστορίας και για το ποιά εικόνα θα εμφανιστεί στη συνέχεια.

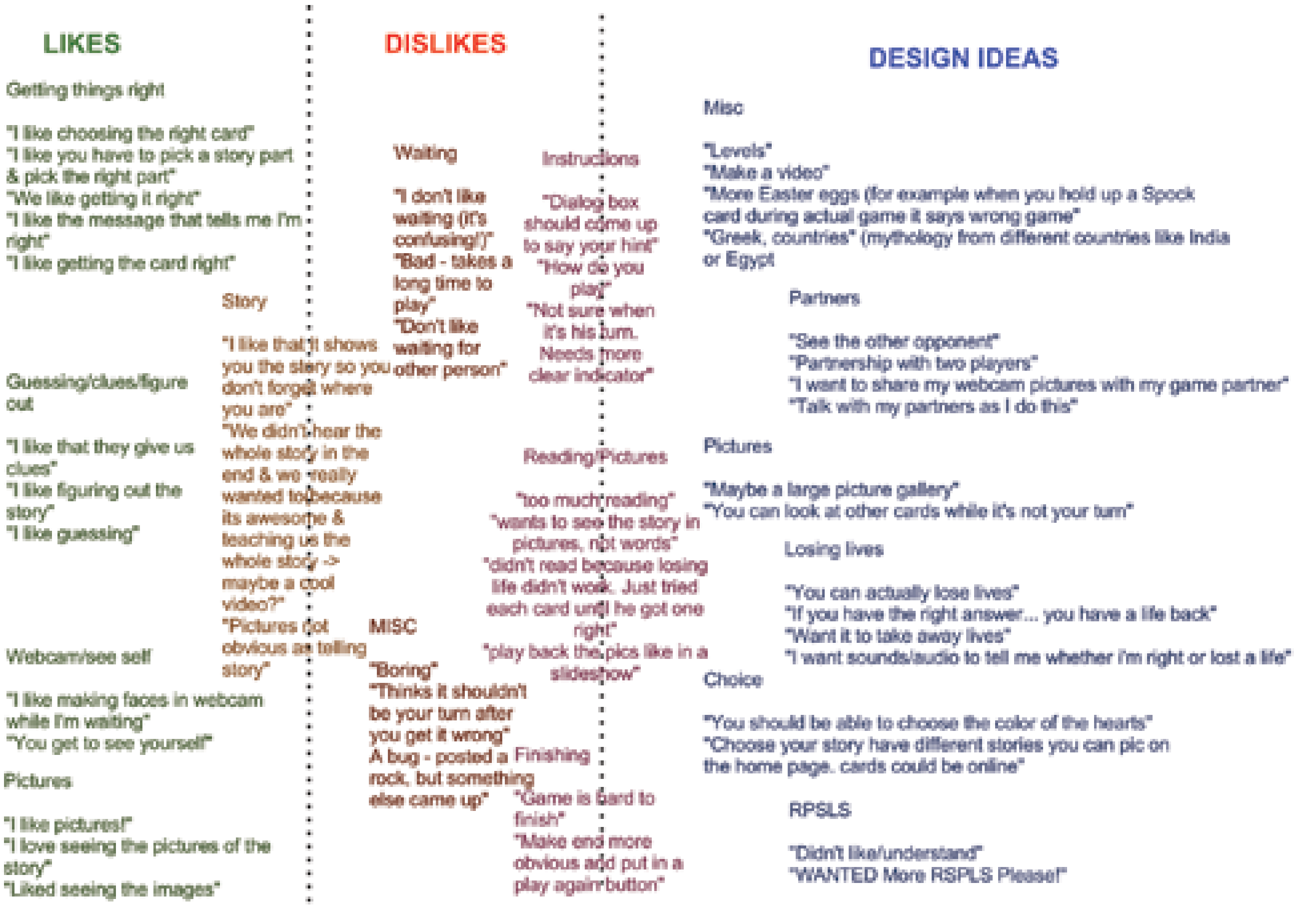
Ανταγωνισμός και Συνεργασία: Οι παίκτες συνεργάζονται για να «κτίσουν» την ιστορία, αλλά ανταγωνίζονται για το ποιος θα κάνει τα λιγότερα λάθη.

7. Εμπειρίες χρήσης - Αξιολόγηση

Το παιχνίδι παρουσιάστηκε στα πλαίσια τεσσάρων διαφορετικών οργανωμένων συνεδρίων/συνεδριάσεων: Ημερίδα Greek ACM SIGCHI (2011), Serious Games Showcase and Best Practices (2011), Σεμινάριο Πληροφοριακών Συστημάτων, Τμήμα Πληροφορικής και Τηλεπικοινωνιών (2011), 3rd IEEE Intelligence Conference in Games and Virtual Worlds for Serious Applications (VS-Games 2011). Στα συνέδρια αυτά δεν ακολουθήθηκαν συγκεκριμένες μέθοδοι αξιολόγησης. Παρόλα αυτά, το παιχνίδι αντιμετωπίστηκε με ενθουσιασμό και περισσότερες «ιδέες για επέκταση», παρά με παράπονα.

Σε μια απόπειρα πιο μεθοδευμένης αξιολόγησης, το παιχνίδι παρουσιάστηκε σε έξι παιδιά ηλικίας 7 έως 11 ετών στο εργαστήριο HCIL (Human-Computer Interaction Lab) στο University of Maryland, στα πλαίσια της τακτικής ερευνητικής δραστηριότητας «Kidsteam», στην οποία τα παιδιά αντιμετωπίζονται ως ισάξιοι σχεδιαστές. Η μεθοδολογία που ακολουθήθηκε, είναι μία εξ αυτών που σχεδιάστηκαν στο ίδιο το εργαστήριο και αποκαλείται «sticky notes» και στην οποία τα παιδιά καλούνται να γράψουν σε μικρά χαρτάκια τι τους άρεσε, τι δεν τους άρεσε, και κάποιες ιδέες σχετικά με το σχεδιασμό του παιχνιδιού. Τα χαρτάκια μετά συλλέγονται από τους ενηλίκους και οργανώνονται ανάλογα, με κατηγοριοποίηση η οποία προκύπτει από τη θεματολογία των απαντήσεων.

Τα αποτελέσματα της αξιολόγησης έδειξαν πως τα παιδιά ήταν ενθουσιασμένα με την ιδέα της κάμερας και δεν τα ενόχλησε το ότι δεν ήξεραν την ιστορία. Τα μοναδικά σημεία δυσαρέσκειας αφορούσαν την έλλειψη ανταμοιβής μετά τη λήξη του παιχνιδιού, την αναμονή μέχρι να απαντήσει ο αντίπαλος και διάφορα άλλα στοιχεία της διεπαφής. Τα παιδιά είχαν επίσης αρκετές ιδέες που αφορούν τη βελτίωση του παιχνιδιού, όπως συνεργατικές τακτικές ή προσωποποιημένες επιλογές.



Εικόνα 6: Συγκεντρωτικός πίνακας αποτελεσμάτων από την αξιολόγηση των παιδιών

8. Μελλοντικές επεκτάσεις

Σύμφωνα με τις εσωτερικές και εξωτερικές αξιολογήσεις, προτείνουμε κάποιες πιθανές μελλοντικές επεκτάσεις. Αρχικά, το σύστημά μας θα μπορούσε να επεκταθεί, έτσι ώστε η ιστορία και οι εικόνες να αποτελούν είσοδο στο σύστημα και οι παίκτες να μπορούν να επιλέξουν παραπάνω από μια ιστορία. Μια σημαντική επέκταση θα ήταν η εμφάνιση και των δύο παικτών σε κάθε οθόνη, καθότι αυτό θα βελτίωνε σημαντικά την εμπειρία χρήσης κατά τη διάρκεια της αναμονής. Τέλος, το παιχνίδι θα βελτιωνόταν σημαντικά αν η ανταπόκριση του συστήματος στην εμφάνιση των καρτών βελτιωνόταν και αν υπήρχαν περισσότερα μέσα ανατροφοδότησης προς τους παίκτες και περισσότεροι τρόποι ανταμοιβής.

9. Συμπεράσματα

Καθ' όλη τη διάρκεια της ανάπτυξης του παιχνιδιού έγινε κατανοητό το πόσο σημαντικός είναι ο προσεκτικός σχεδιασμός ενός παιχνιδιού, που να συνδυάζει εύχρηστη και κατανοητή διεπαφή, να προκαλεί το ενδιαφέρον του παίκτη και να έχει ταυτόχρονα εκπαιδευτικό χαρακτήρα. Από τα σχόλια που έκαναν άτομα που έπαιξαν το παιχνίδι, έγινε εμφανές το πόσο σημαντική είναι η γρήγορη απόκριση του παιχνιδιού. Παρατηρώντας τις αντιδράσεις και τον τρόπο που συμπεριφέρονταν τα άτομα που έπαιξαν το παιχνίδι, είδαμε ότι πολλοί δεν έδιναν σημασία στις συμβουλές που παρείχε το παιχνίδι για να τους διευκολύνει ώστε να διαλέξουν τη σωστή κάρτα. Επίσης, κάποιοι από τους χρήστες δεν διάβαζαν προσεκτικά τα ονόματα των καρτών, αλλά αυτό συνέβη μόνο υπό συνθήκες πίεσης.

Αναφορές

1. Νικολαΐδης, Δημήτρης. «Επαυξημένη Πραγματικότητα: Πολλαπλασιάζοντας τις δυνατότητες των αισθήσεων». Περισκόπιο της Επιστήμης, Μάρτιος 2003

2. Nielsen, Jacob. «Usability Engineering». Morgan Kaufmann Publishers, 1993

3. Ρούσσου, Μαρία. «Ο Ρόλος της Διαδραστικότητας στη Διαμόρφωση της Άτυπης Εκπαιδευτικής Εμπειρίας». Στο Αλεξάνδρα Μπούνια, Νίκη Νικονάνου, & Μαρία Οικονόμου (Επιμ.), Η Τεχνολογία στην Υπηρεσία της Πολιτισμικής Κληρονομιάς (σελ. 251-261). Αθήνα: Καλειδοσκόπιο. 2008

4. Malone, T.W. & Lepper, M.R. «Making Learning Fun: a Taxonomy of Intrinsic Motivations for Learning». Aptitude, Learning, and Instruction: Cognitive and Affective Process Analyses, R. Snow and M. Farr, eds., Lawrence Erlbaum Associates, Hillsdale, NJ. 1987 http://education.calumet.purdue.edu/vockell/edPsybook/Edpsy5/edpsy5_intrinsic.htm [Jun. 23, 2011]

5. «FLARManager: Augmented Reality in Flash» http://words.transmote.com/wp/flarmanager/ [Jun 23, 2011]

Μέθοδοι Διάσπασης Συμμετριών για Προβλήματα Ικανοποίησης Περιορισμών

Καλλιρρόη Δογάνη
kallirroi@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Σε αρκετά προβλήματα ικανοποίησης περιορισμών είναι πιθανό να επισκεφθούμε ισοδύναμες καταστάσεις πάνω από μια φορά λόγω των συμμετριών. Σε αυτή την εργασία παρουσιάζονται οι σημαντικότερες μεθόδοι διάσπασης συμμετριών, οι οποίες έχουν σαν στόχο να μην εξερευνηθούν ποτέ δύο καταστάσεις που είναι συμμετρικά ισοδύναμες. Επίσης, επεκτείναμε τον επιλυτή Naxos [3], ένα γενικό επιλυτή προβλημάτων ικανοποίησης περιορισμών που είναι ευρύτερα διαθέσιμος σε χρήστες-προγραμματιστές, εισάγοντας σε αυτόν δύο βασικές τεχνικές διάσπασης συμμετριών. Τέλος, υπάρχουν μετρήσεις απόδοσης για τις συγκεκριμένες δύο μεθόδους και συμπεράσματα.

Λέξεις-Κλειδιά: Προγραμματισμός με περιορισμούς, Διάσπαση συμμετριών, Ισοδύναμες καταστάσεις, SBDS, SBDD

Επιβλέποντες: Σταματόπουλος Παναγιώτης, Επίκουρος Καθηγητής
Ποθητός Νικόλαος, Υποψήφιος Διδάκτορας

1. Εισαγωγή

Οι συμμετρίες εμφανίζονται στη ζωή μας με διάφορες μορφές και σε διάφορους τομείς. Οι θεωρητικοί φυσικοί ισχυρίζονται ότι σχεδόν όλοι οι νόμοι της φύσης προέρχονται από συμμετρίες. Οι βιολόγοι και χημικοί μελετούν τις συμμετρίες μέσα από τα μακρομόρια, τις χημικές δομές, την κρυσταλλογραφία κ.α. Σε όλα αυτά έρχονται να προστεθούν οι συμμετρίες που περιγράφονται στην αρχιτεκτονική, αλλά και σε διάφορες μορφές τέχνης, όπως η ζωγραφική, η αγγειοπλαστική, η μουσική κτλ.

Σε αυτή την εργασία θα επικεντρωθούμε σε συμμετρίες που εμφανίζονται στον προγραμματισμό, και συγκεκριμένα σε προβλήματα ικανοποίησης περιορισμών στον κλάδο της τεχνητής νοημοσύνης. Σε αυτές τις περιπτώσεις η ύπαρξη συμμετρίας δεν είναι πάντοτε επιθυμητή, καθώς οδηγεί σε ισοδύναμες καταστάσεις προσθέτοντας κόστος σε χρόνο και χώρο. Όταν στόχος μας είναι τα αποτελέσματα που θα λάβουμε να είναι διακριτά μεταξύ τους (όχι συμμετρικά ισοδύναμα), τότε θα πρέπει με κάποιο τρόπο να διασπάσουμε τις συμμετρίες. Αυτό είναι και το βασικό αντικείμενο μελέτης μας. Θα περιγράψουμε *μεθόδους διάσπασης συμμετριών* (symmetry breaking) σε προβλήματα ικανοποίησης περιορισμών.

Οι μέθοδοι στις οποίες θα αναφερθούμε διαφέρουν ως προς το είδος των συμμετριών που αντιμετωπίζουν, ως προς τη χρονική στιγμή εφαρμογής τους και ως προς το ποσοστό συμμετρίας που διασπούν. Πολλές φορές δεν μας ενδιαφέρει να σπάσουμε κάθε συμμετρία, παρά μόνο τόσες όσες χρειάζεται ώστε να μειώσουμε τα χρονικά και χωρικά κόστη σε ένα ικανοποιητικό πλαίσιο.

Στην Παράγραφο 2 γίνεται μία παρουσίαση των Προβλημάτων Ικανοποίηση Περιορισμών. Στην Παράγραφο 3 περιγράφονται οι σημαντικότερες έννοιες πάνω στις Συμμετρίες, οι οποίες αποτελούν βάση για όσα θα αναλύσουμε στην Παράγραφο 4 για τις διάφορες Μεθόδους Διάσπασης Συμμετριών. Στην Παράγραφο 5 π εριγράφουμε την υλοποίηση των μεθόδων SBDS και SBDD στον επιλυτή Naxos. Στην Παράγραφο 6 περιέχονται τα αποτελέσματα των μετρήσεων απόδοσης των δύο αυτών μεθόδων και, τέλος, στην Παράγραφο 7 κλείνουμε με κάποια συμπεράσματα που αφορούν την απόδοση αυτών των μεθόδων και τη χρησιμότητά τους.

2. Προβλήματα ικανοποίησης περιορισμών

2.1 Ορισμός

Τυπικά, ένα *πρόβλημα ικανοποίησης περιορισμών* (constraint satisfaction problem - CSP) ορίζεται από [4]:

- ένα σύνολο μεταβλητών $X = \{ X_1, X_2, \dots, X_n \}$
- για κάθε μεταβλητή X_i ένα μη κενό *πεδίο* (domain) D_i των δυνατών *τιμών* της
- ένα σύνολο περιορισμών $C = \{ C_1, C_2, \dots, C_n \}$. Κάθε περιορισμός δρα σε ένα σύνολο $S \subseteq X$ και καθορίζει τους επιτρεπτούς συνδυασμούς τιμών που μπορούν να πάρουν οι μεταβλητές που ανήκουν στο S.

Μία *κατάσταση* (state) του προβλήματος ορίζεται με ανάθεση τιμών σε κάποιες από τις μεταβλητές, $\{ X_i = \upsilon_i, X_j = \upsilon_j, \dots \}$. Αν μία ανάθεση δεν παραβιάζει κανέναν από τους περιορισμούς, τότε ονομάζεται *συνεπής* (consistent). Μία ανάθεση που περιλαμβάνει όλες τις μεταβλητές ονομάζεται *πλήρης*. Λύση του προβλήματος ικανοποίησης περιορισμών είναι μία πλήρης, συνεπής ανάθεση.

2.2 Το πρόβλημα των N βασιλισσών

Το πρόβλημα των N βασιλισσών έγκειται στην τοποθέτηση N αριθμού βασιλισσών σε διαφορετικές θέσεις μιας σκακιέρας N x N , με τέτοιο τρόπο ώστε καμία βασίλισσα να μην απειλεί κάποια άλλη, δηλαδή να μην βρίσκεται σε ίδια γραμμή, στήλη ή διαγώνιο. Ας το μοντελοποιήσουμε ως πρόβλημα ικανοποίησης περιορισμών. Θα έχουμε N μεταβλητές, μία για κάθε γραμμή της σκακιέρας και ουσιαστικά για κάθε βασίλισσα:

$$X = \{ X_1, X_2, \dots, X_n \}$$

Το πεδίο τιμών κάθε μεταβλητής είναι όλες οι πιθανές στήλες στις οποίες μπορεί να τοποθετηθεί η βασίλισσα της συγκεκριμένης γραμμής:

$$\forall i \quad D_i = \{ 1, 2, 3, \dots, N \}$$

Αν προσέξουμε, ο τρόπος που μοντελοποιήσαμε το πρόβλημα δεν επιτρέπει δυο βασίλισσες στην ίδια γραμμή, αφού κάθε γραμμή σημαίνει διαφορετική βασίλισσα. Άρα οι περιορισμοί που πρέπει να βάλουμε είναι να μην υπάρχουν βασίλισσες στην ίδια στήλη ή στην ίδια διαγώνιο. Για να μην είναι στην ίδια στήλη βάζουμε τον περιορισμό:

$$X_i \neq X_j, \ \forall\, i, j = 1, \ldots, N \text{ και } i < j$$

Για να μην υπάρχουν βασίλισσες στην ίδια διαγώνιο βάζουμε τους περιορισμούς:

$$X_i + i \neq X_j + j \text{ και } X_i - i \neq X_j - j, \ \forall\, i, j = 1, \ldots, N \text{ και } i < j$$

3. Συμμετρίες

3.1 Συμμετρίες στο πρόβλημα των N βασιλισσών

Στο πρόβλημα των N βασιλισσών αν δεν εξαλείψουμε τις συμμετρίες, για κάθε λύση θα εμφανίζονται και όλες οι συμμετρικές της στο σύνολο των λύσεων. Ποιες είναι όμως οι συμμετρίες στο δεδομένο πρόβλημα;

Υπάρχουν συνολικά οχτώ συμμετρίες στο πρόβλημα, συμπεριλαμβανομένης και της ταυτοτικής συμμετρίας (identity symmetry) η οποία αφήνει αφήνει κάθε σημείο εκεί που βρίσκεται. Μπορούμε να περιστρέψουμε το σκάκι κατά 90, 180 και 270 μοίρες με κατεύθυνση ίδια με των δεικτών του ρολογιού. Επίσης, υπάρχουν οι συμμετρίες για τις αντανακλάσεις ως προς τον κάθετο άξονα, τον οριζόντιο και ως προς τους δυο διαγώνιους άξονες.

3.2 Ορισμοί συμμετρίας

Υπάρχουν δύο ειδών ορισμοί της συμμετρίας: αυτοί που ορίζουν ότι η συμμετρία εντοπίζεται στο σύνολο λύσεων και αυτοί που ορίζουν ότι η συμμετρία μπορεί να αναγνωριστεί στην κατάσταση του προβλήματος, χωρίς όμως να το λύνουν. Αυτές οι δυο περιπτώσεις αναφέρονται αντίστοιχα ως συμμετρία λύσεων και συμμετρία προβλήματος.

Δύο βασικά είδη συμμετρίας λύσεων είναι η συμμετρία μεταβλητών και η συμμετρία τιμών. Η πρώτη ουσιαστικά αναφέρεται σε μεταθέσεις μεταβλητών, ενώ η δεύτερη διατηρεί τη διάταξη των μεταβλητών και μεταθέτει τις τιμές τους. Οι πιο επίσημοι ορισμοί είναι οι εξής:

Ορισμός 1 (Συμμετρία Μεταβλητών):

Η συμμετρία μεταβλητών είναι μια αμφιμονοσήμαντη συνάρτηση θ που αντιστοιχίζει μεταξύ τους τους δείκτες των μεταβλητών έτσι ώστε αν η $X_1 = d_1 , \ldots , X_n = d_n$ αποτελεί λύση του προβλήματος, τότε και η $X_{\theta(1)} = d_1 , \ldots , X_{\theta(n)} = d_n$ είναι επίσης λύση.

Ορισμός 2 (Συμμετρία Τιμών):

Η συμμετρία τιμών είναι μια αμφιμονοσήμαντη συνάρτηση σ που αντιστοιχίζει μεταξύ τους τις τιμές των μεταβλητών έτσι ώστε αν η $X_1 = d_1 , \ldots , X_n = d_n$ αποτελεί λύση του προβλήματος, τότε και η $X_1 = \sigma(d_1) , \ldots , X_n = \sigma(d_n)$ είναι επίσης λύση.

Παρακάτω ακολουθούν οι ορισμοί της συμμετρίας λύσεων και της συμμετρίας προβλήματος όπως περιγράφονται στο [6]:

Ορισμός 3 (Συμμετρία λύσεων):

Η συμμετρία λύσεων είναι μια μετάθεση του συνόλου των ζευγαριών <μεταβλητή, τιμή> η οποία διατηρεί το σύνολο των λύσεων.

Ορισμός 4 (Συμμετρία προβλήματος):

Η συμμετρία προβλήματος είναι μια μετάθεση του συνόλου των ζευγαριών <μεταβλητή, τιμή> η οποία διατηρεί το σύνολο των περιορισμών.

4. Αλγόριθμοι διάσπασης συμμετριών

Οι αλγόριθμοι διάσπασης συμμετριών (symmetry breaking) χωρίζονται σε τρεις βασικές κατηγορίες: αναδιατύπωση του προβλήματος, προσθήκη περιορισμών πριν την αναζήτηση και προσθήκη περιορισμών κατά τη διάρκεια της αναζήτησης (dynamic symmetry breaking).

4.1 Αναδιατύπωση του προβλήματος

Η μοντελοποίηση του προβλήματος επιδρά σε μεγάλο βαθμό στο πόσο αποτελεσματικά θα λυθεί το πρόβλημα (κόστος σε χώρο και χρόνο). Η αναδιατύπωση του προβλήματος με μια διαφορετική μοντελοποίηση παίζει βασικό ρόλο για τη διάσπαση των συμμετριών. Διαφορετικά μοντέλα του ίδιου προβλήματος μπορεί να έχουν διαφορετικές συμμετρίες από τις οποίες κάποιες μπορεί να είναι πιο εύκολο να τις χειριστούν από ό,τι κάποιες άλλες. Σε εξειδικευμένες περιπτώσεις, άλλη μοντελοποίηση μπορεί να εξαφανίζει εντελώς τις συμμετρίες, ενώ σε άλλες απλά να μειώνει τη συμμετρία σε σημαντικό βαθμό.

4.2 Προσθήκη περιορισμών πριν την αναζήτηση

Σε αυτή την κατηγορία οι μέθοδοι διάσπασης συμμετριών περιλαμβάνουν την προσθήκη κάποιων περιορισμών πριν αρχίσει η αναζήτηση για να εξαφανίστουν οι συμμετρίες. Οι νέοι αυτοί περιορισμοί θα μας επιτρέψουν να έχουμε στην καλύτερη περίπτωση μόνο μια λύση από κάθε σύνολο ισοδύναμων συμμετρικών λύσεων.

Μια τέτοια μέθοδος είναι η «Lex-Leader» [2], η οποία αφορά συμμετρίες μεταβλητών. Η βασική της ιδέα είναι ότι διαλέγουμε μια συγκεκριμένη διάταξη για τις μεταβλητές και προσθέτουμε περιορισμούς για να εξασφαλίσουμε ότι η τελική λύση θα είναι λεξικογραφικά μικρότερη από οποιαδήποτε συμμετρική αναδιάταξη των μεταβλητών. Για παράδειγμα, αν έχουμε μια συμμετρία που αντιστρέφει τη σειρά των μεταβλητών και αντιστοιχίζει το X_i στο X_{n-i+1}, τότε προσθέτουμε τον παρακάτω περιορισμό που διατηρεί τη λεξικογραφική διάταξη:

$$[X_1, \dots, X_n] \leq_{lex} [X_n, \dots, X_1]$$

όπου $\leq_{lex}$ συμβολίζει τη σχέση λεξικογραφικής διάταξης.

4.3 Δυναμικές μέθοδοι διάσπασης συμμετριών

Οι δυναμικές μέθοδοι διάσπασης συμμετριών προσθέτουν περιορισμούς κατά τη διάρκεια της αναζήτησης. Οι αλγόριθμοι SBDS και SBDD είναι δύο βασικές μέθοδοι που ανήκουν σε αυτή την κατηγορία και είναι αυτές που υλοποιήσαμε στον Naxos Solver στα πλαίσια της πτυχιακής.

SBDS:

Η μέθοδος SBDS (Symmetry Breaking During Search [7, 10, 11]) στηρίζεται στο εξής: αφού έχουμε οπισθοχωρήσει από κάποια ανάθεση τιμής σε μεταβλητή, θα αφαιρέσουμε από το πεδίο τιμών της μεταβλητής όχι μόνο τη συγκεκριμένη τιμή, αλλά αν εξακολουθεί να υπάρχει συμμετρία, θα αφαιρέσουμε και τη συμμετρική ισοδύναμή της για τη συγκεκριμένη συμμετρία. Η μέθοδος αυτή αποτελεί μια δυναμική τεχνική, καθώς δεν μπορούμε να προσθέσουμε περιορισμούς μέχρι να ληφθεί μια απόφαση στη διάρκεια της αναζήτησης.

Έστω ότι βρισκόμαστε σε ένα σημείο στην αναζήτηση, όπου A είναι η μερική ανάθεση που έχει γίνει μέχρι στιγμής και θέλουμε να την επεκτείνουμε με την απόφαση. Για κάθε συμμετρία g του προβλήματος προσθέτουμε τον περιορισμό:

$$A \;\&\; A^{g} \;\;\&\; var \neq val \Rightarrow (var \neq val)^{g}$$

(Με A^{g} συμβολίζουμε τις αναθέσεις που θα πάρουμε αν στις αναθέσεις A εφαρμόσουμε τη συμμετρία g).

Αυτός ο περιορισμός ουσιαστικά σημαίνει πως αν ισχύουν κάποιες αναθέσεις και ταυτόχρονα ισχύουν και οι συμμετρικές τους ως προς τη συμμετρία g (δηλαδή ισχύει A^{g}), αν αφαιρέσω από τη μεταβλητή val την τιμή var θα πρέπει να αφαιρέσω και το συμμετρικό του $val = var$.

SBDD:

Η μέθοδος SBDD (Symmetry Breaking via Dominance Detection [8, 9]) ελέγχει κάθε κόμβο του δέντρου αναζήτησης για να δει αν είναι συμμετρικά ισοδύναμος με κάποιο άλλο κόμβο που έχει ήδη εξερευνηθεί και σε αυτή την περίπτωση δεν επεκτείνει άλλο τον κόμβο. Με άλλα λόγια, χτίζοντας το δέντρο προσπαθούμε να εντοπίσουμε μονοπάτια που αποτελούν συμμετρικά ισοδύναμα ήδη εξερευνημένων μονοπατιών και να τα αποκόψουμε.

Στη συνέχεια θα μας χρειαστούν και οι παρακάτω ορισμοί:

Ορισμός 5 (No-good):

Ο κόμβος ν είναι ένας no-good του n αν υπάρχει πρόγονος n_a του n, ο ν είναι το αριστερό παιδί του n_a και ο ν δεν είναι πρόγονος του n.

Ορίζουμε ως $\delta(n)$ το σύνολο των αποφάσεων στο μονοπάτι από τη ρίζα μέχρι τον κόμβο no-good. Γράφουμε $\Delta(S)$ για το σύνολο των ζευγαριών $v_i = \alpha_i$ για όλες τις μεταβλητές v_i που στο πεδίο τιμών τους έχει μείνει μόνο μια τιμή.

Ορισμός 6 (Κυριαρχία):

Λέμε πως ένας κόμβος n κυριαρχείται αν υπάρχει ένας no-good κόμβος ν του n και μια συμμετρία g έτσι ώστε $(\delta(\nu))^{g} \subseteq \Delta(n)$. Τότε ο ν κυριαρχεί πάνω στον n.

5. Υλοποίηση των αλγορίθμων SBDS και SBDD στον επιλυτή Naxos

Ο επιλυτής προβλημάτων ικανοποίησης περιορισμών Naxos κατασκευάστηκε από τον Ν. Ποθητό στα πλαίσια της δικής του πτυχιακής εργασίας [1]. Πρόκειται για μία βιβλιοθήκη που μπορεί να χρησιμοποιηθεί για να λυθούν προβλήματα μέσω του προγραμματισμού με περιορισμούς.

Η αναζήτηση στον Naxos υλοποιείται μέσα από τους στόχους. Η συνήθης λειτουργία ενός στόχου είναι να προκαλεί την ανάθεση μίας τιμής σε μία μεταβλητή ή την αφαίρεση κάποιας τιμής από το πεδίο της. Αν κατά τη διάρκεια της αναζήτησης αφαιρεθούν όλες οι τιμές από το πεδίο κάποιας μεταβλητής τότε ο επιλυτής αναλαμβάνει να ακυρώσει τους στόχους που οδήγησαν στην αποτυχία και να επαναφέρει το γράφο περιορισμών στην προηγούμενη κατάστασή του. Επίσης, ένας στόχος έχει τη δυνατότητα να επιλέξει την επόμενη μεταβλητή στην οποία θα γίνει ανάθεση, ή, κατά τον τερματισμό της εκτέλεσης του, να δημιουργήσει έναν άλλο στόχο.

5.1 Symmetry Breaking During Search (SBDS)

Το πλεονέκτημα με το μηχανισμό στόχων, που αναφέραμε παραπάνω, είναι ότι μπορούμε να κατευθύνουμε την αναζήτηση εκεί που θέλουμε χωρίς να χρειαστεί να επέμβουμε στον τρόπο που εκτελείται, ο οποίος είναι σταθερός ανεξάρτητα με το τι επιδιώκουμε από αυτή. Για να εκμεταλλευτούμε αυτό το γεγονός, αλλά και για να ακολουθήσουμε το πνεύμα του Naxos Solver, στηρίξαμε την υλοποίηση του SBDS σε στόχους.

Φτιάξαμε τους δικούς μας στόχους, οι οποίοι είναι υπεύθυνοι για τη διάσπαση συμμετριών. Είδαμε ότι ο SBDS όταν αποφασίζει να αφαιρεθεί μια τιμή από κάποια μεταβλητή, ελέγχει αν υπάρχουν συνθήκες συμμετρίας ώστε να αφαιρεθεί και η συμμετρική της. Ουσιαστικά προσθέτει υποσυνθήκη περιορισμούς. Το δύσκολο κομμάτι στη δική μας υλοποίηση ήταν ότι δεν μπορούσαμε να προσθέσουμε τέτοιους περιορισμούς, καθώς η υλοποίηση του Naxos δεν επιτρέπει νέους περιορισμούς από τη στιγμή που ξεκινάει η αναζήτηση. Αυτό σημαίνει πως αν η συνθήκη του περιορισμού $A_i \ \& \ A_i^g \ \& \ var_i \neq val \Rightarrow (var_i \neq val)^g$ δεν ήταν ικανοποιήσιμη εκείνη τη στιγμή που ξεκινούσε να ισχύει ο περιορισμός, ώστε να θέσουμε κατευθείαν σαν στόχο την αφαίρεση της συμμετρικής τιμής, θα έπρεπε με κάποιο τρόπο να αποθηκεύουμε τη συνθήκη και να ελέγχουμε μήπως πραγματοποιηθεί πιο κάτω στο δέντρο αναζήτησης. Αυτό το κάναμε έχοντας έναν δισδιάστατο πίνακα όπου στον 1ο πίνακα κρατάγαμε ποιες τιμές έχουν αφαιρεθεί από την 1η μεταβλητή και χρειάζονται έλεγχο, ομοίως στο 2ο πίνακα ποιες τιμές έχουν αφαιρεθεί από τη 2η μεταβλητή κ.ο.κ. Το A_i δείχνει ποιες αναθέσεις τιμών έχουν γίνει μέχρι την ανάθεση της μεταβλητής i. Στις περιπτώσεις που οι μεταβλητές δεν παίρνουν τιμή με τη σειρά, καθώς έχουμε πει πως ο SBDS σέβεται τη σειρά επιλογής μεταβλητών και τιμών των διάφορων ευρετικών, θα πρέπει να κρατάμε κάπου τη σειρά με την οποία πήραν τιμή οι μεταβλητές. Όταν τελικά εντοπίζαμε πιο κάτω στο δέντρο να ισχύει η συνθήκη, απλά αφαιρούσαμε κατευθείαν τη συμμετρική τιμή. Υπήρχαν και άλλα θέματα που αντιμετωπίσαμε κατά την υλοποίηση του SBDS στο Naxos, π.χ.:

- τι γίνεται στις περιπτώσεις που εξαιτίας της διάδοσης περιορισμών μένει μόνο μια τιμή σε κάθε μεταβλητή και ο επιλυτής εμφανίζει αυτές τις τιμές σα λύση πριν προλάβουμε να ελέγξουμε αν υπάρχει συμμετρία
- όταν γίνεται διάδοση περιορισμών και αφαιρούνται τιμές από μεταβλητές, θα πρέπει να αφαιρεθούν και οι συμμετρικές τους σε συνθήκες συμμετρίας;

5.2 Symmetry Breaking via Dominance Detection (SBDD)

Η υλοποίηση του SBDD είναι και αυτή βασισμένη σε στόχους. Πριν θέσουμε σα στόχο την ανάθεση κάποιας τιμής σε μεταβλητή, καλούμε μια συνάρτηση για έλεγχο κυριαρχίας, δηλαδή εξετάζουμε αν ο κόμβος στον οποίο βρισκόμαστε είναι ισοδύναμος με κάποιον που έχουμε ήδη εξερευνήσει. Αν ισχύει αυτό, τότε απλά ο στόχος που επιστρέφουμε είναι να αφαιρεθεί αυτή η τιμή από τη μεταβλητή, διακόπτοντας το συγκεκριμένο μονοπάτι, και να δοκιμάσουμε κάποια άλλη τιμή για αυτή τη μεταβλητή.

Το ιδιαίτερο σε αυτό τον αλγόριθμο ήταν με ποιο τρόπο θα κρατάμε τους κόμβους που έχουμε ήδη επισκεφθεί, οι οποίοι χρειάζονται για να εφαρμόσουμε στη συνέχεια τον ελέγχο κυριαρχίας. Το να αποθηκεύουμε όλους τους κόμβους δεν είναι καθόλου αποδοτικό από άποψη χώρου. Αυτοί που χρειάζεται τελικά να αποθηκεύουμε είναι οι κόμβοι no-good, δηλαδή οι ρίζες πλήρως εξερευνημένων υποδέντρων.

Στον επιλυτή Naxos το συγκεκριμένο θέμα ήταν αρκετά απλό να υλοποιηθεί λόγω των μεταβλητών NsIntVar *(Ακέραιες περιορισμένες μεταβλητές πεπερασμένων πεδίων)*, που υπακούουν στους κανόνες της οπισθοχώρησης. Για να κρατάμε τους no-good κόμβους φτιάξαμε ένα δικό μας στόχο, ο οποίος αφού αφαιρούσε κάποια τιμή από κάποια μεταβλητή, την αποθήκευε σε έναν πίνακα με μεταβλητές τύπου NsIntVar.

6. Μετρήσεις απόδοσης των SBDS και SBDD

Εφαρμόσαμε τις μεθόδους SBDS και SBDD σε δύο προβλήματα ικανοποίησης περιορισμών που εμφανίζουν συμμετρία. Τα αποτελέσματα ήταν παρόμοια και στα δύο. Εμείς εδώ θα εξετάσουμε μόνο το πρόβλημα των N Βασιλισσών. Θα δούμε πόσο μειώθηκαν οι λύσεις με τον αποκλεισμό των συμμετρικών, θα κάνουμε διάφορες μετρήσεις σχετικά με το χρόνο και τους κόμβους και θα συγκρίνουμε τα αποτελέσματα με το απλό Naxos Solver που δε χρησιμοποιεί κάποια μέθοδο διάσπασης συμμετριών. Επίσης, θα κάνουμε και μια σύγκριση μεταξύ των δύο μεθόδων που υλοποιήσαμε για να δούμε αν παρουσιάζουν κάποια διαφοροποίηση στα αποτελέσματα.

Ο Πίνακας 1 μας δείχνει τον αριθμό των λύσεων για διάφορα N στην περίπτωση που δεν εφαρμόζεται κάποιος αλγόριθμος διάσπασης συμμετριών και στην περίπτωση που εφαρμόσαμε κάποιον από τους SBDS και SBDD (ο αριθμός των μη συμμετρικών λύσεων είναι ο ίδιος και στις δύο μεθόδους). Παρατηρούμε ότι οι λύσεις μειώθηκαν σε μεγάλο βαθμό. Ειδικά για μεγάλα N βλέπουμε ότι καταφέραμε να αποφύγουμε κάποια εκατομμύρια συμμετρικές λύσεις, οι οποίες βέβαια αυξάνονται εκθετικά όσο αυξάνεται και το N.

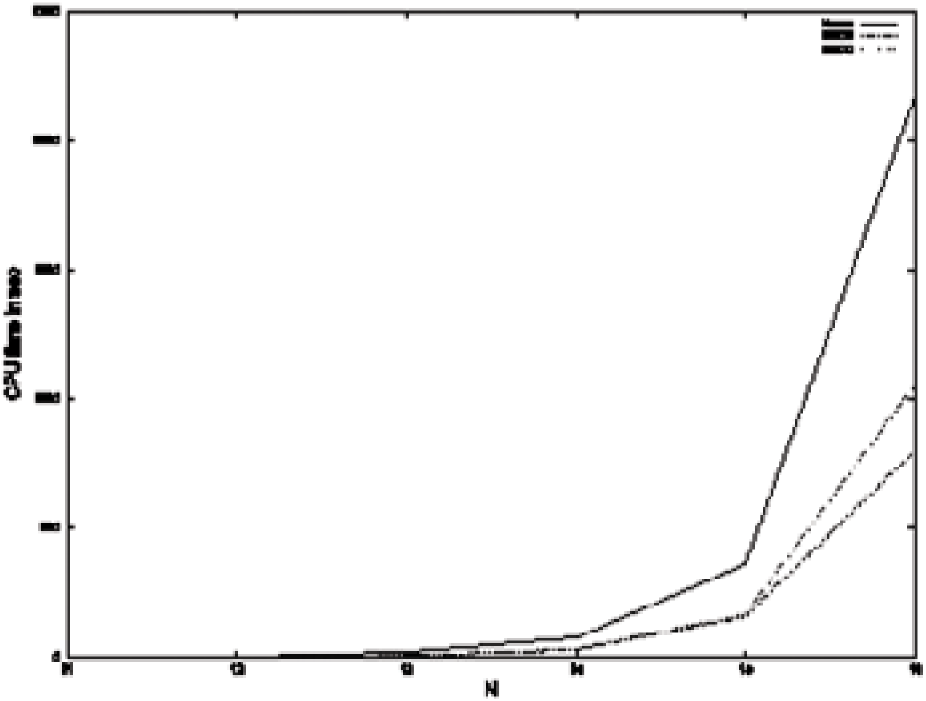
N	Naxos Solver	SBDS/SBDD
4	2	1
5	10	2
6	4	1
7	40	6
8	92	12
9	352	46

N	Naxos Solver	SBDS/SBDD
10	724	92
11	2680	341
12	14200	1787
13	73712	9233
14	365596	45752
15	2279184	285053
16	14772512	1846955

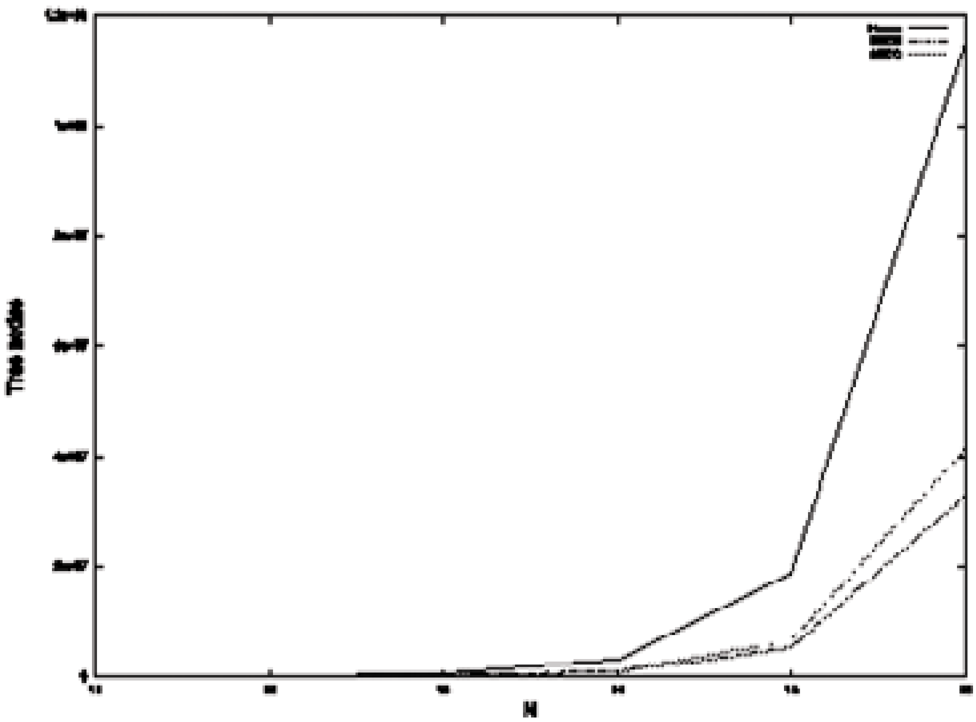
Πίνακας 1: Αριθμός λύσεων στο πρόβλημα των N βασιλισσών για διάφορες τιμές του N στην περίπτωση που δεν έχουμε διάσπαση συμμετριών και, αντίστοιχα, στην περίπτωση που έχουμε.

Στο Σχήμα 1 βλέπουμε το σχεδιάγραμμα χρόνου για τις δύο μεθόδους και για το απλό Naxos. Οι διαφορές αρχίζουν και γίνονται έντονα αισθητές για $N \geq 13$. Παρατηρούμε ότι κερδίζουμε όντως και σε χρόνο. Και οι δύο μέθοδοι που υλοποιήσαμε έχουν μικρότερο κόστος σε χρόνο, κάτι που ήταν αναμενόμενο καθώς αφαιρώντας συμμετρικά ισοδύναμα μονοπάτια μειώνεται ο χώρος αναζήτησης. Όσο αυξάνεται το N, τόσο πιο αισθητή γίνεται και η διαφορά. Από $N = 15$ παρατηρούμε μια διαφορά και στο χρόνο των δύο αλγορίθμων. Για $N = 16$ το απλό Naxos κάνει περίπου 32 λεπτά, ο SBDD 17 λεπτά και ο SBDS 13 λεπτά. Αν σκεφτούμε πώς δουλεύουν οι δύο μέθοδοι θα μπορέσουμε να καταλάβουμε πού οφείλεται αυτή η διαφορά. Ο SBDS εντοπίζει συμμετρικά ισοδύναμους κόμβους και τους αφαιρεί πριν τους επισκεφτεί. Αντίθετα, ο SBDD πρώτα επισκέπτεται τον κόμβο και αν δει ότι είναι συμμετρικά ισοδύναμος με κάποιον που έχει ήδη εξερευνηθεί, τότε αποκόπτει το μονοπάτι. Άρα αφού επισκέπτεται λιγότερους κόμβους είναι λογικό για μεγάλα N να υπάρχει διαφορά στο χρόνο.

Το Σχήμα 2 που δείχνει τον αριθμό των κόμβων μας επιβεβαιώνει όσα είπαμε για τους κόμβους. Ο SBDS φαίνεται να έχει σταθερά λιγότερους κόμβους. Όμως και οι δύο μέθοδοι διάσπασης συμμετριών επισκέπτονται εμφανώς λιγότερους κόμβους από ό,τι αν δεν εφαρμόζαμε κάποια μέθοδο. Όπως ήδη αναφέραμε, ο χώρος αναζήτησης είναι πολύ μικρότερος.



Σχήμα 1: Σχεδιάγραμμα Χρόνου (sec) –N για το απλό Naxos, τον SBDS και τον SBDD για το πρόβλημα των βασιλισσών. Χρησιμοποιήθηκε επεξεργαστής Intel i5, 3.3 Hz, 4GB μν



Σχήμα 2: Σχεδιάγραμμα Αριθμός κόμβων – N για το απλό Naxos, για τον SBDS και για τον SBDD για το πρόβλημα των βασιλισσών.

7. Συμπεράσματα και μελλοντικές κατευθύνσεις

Στα πλαίσια της πτυχιακής μελετήσαμε διάφορους αλγόριθμους διάσπασης συμμετριών και επικεντρωθήκαμε κυρίως στους SBDS και SBDD, τους οποίους υλοποιήσαμε και στον επιλυτή Naxos. Από τις μετρήσεις που κάναμε για τις δύο αυτές μεθόδους είδαμε ότι μειώνουν σε μεγάλο βαθμό το χώρο αναζήτησης, καθώς και το κόστος σε χρόνο. Από τις ίδιες μετρήσεις διαπιστώσαμε ότι για μεγάλες τιμές εισόδων (δηλαδή περισσότερες συμμετρικές λύσεις), ο SBDS φαίνεται να είναι καλύτερος από τον SBDD, καθώς έχει καλύτερα αποτελέσματα σε χώρο και χρόνο. Ωστόσο, μπορούμε να συμπεράνουμε ότι οι δύο μέθοδοι δεν απέχουν τόσο πολύ μεταξύ τους. Και οι δύο προσπαθούν να εντοπίσουν συμμετρικά ισοδύναμους κόμβους και να τους αποκόψουν. Η μόνη ουσιαστική διαφορά, η οποία προκαλεί και τις διαφοροποιήσεις στα αποτελέσματα, είναι ότι ο SBDD περιμένει να φτάσει σε ένα κόμβο για να ελέγξει αν είναι συμμετρικά ισοδύναμος με κάποιον άλλο, ενώ ο SBDS τον εντοπίζει και τον αποκόπτει νωρίτερα. Αν εξαιρέσουμε αυτό το λεπτό σημείο, θα μπορούσαμε να πούμε ότι μια υλοποίηση του SBDS είναι ουσιαστικά μια υλοποίηση του SBDD και αντίστροφα.

Υπάρχουν διάφορα θέματα τα οποία δεν τα εξετάσαμε. Για παράδειγμα, όταν ο αριθμός των συμμετριών είναι μεγάλος οι SBDS και SBDD διατηρούν την ίδια συμπεριφορά; Σε αυτή την περίπτωση είναι πολύ πιθανό ο SBDD να είναι πιο αποδοτικός λόγω της απλότητάς του.

Κάτι άλλο που θα μπορούσε να μελετηθεί είναι ο συνδυασμός διάφορων μεθόδων διάσπασης συμμετριών. Αν το πρόβλημα είναι κατάλληλο, θα ήταν καλύτερο να προσθέσουμε κάποιους περιορισμούς που να σπάνε κάποιες συμμετρίες πριν ξεκινήσουμε την αναζήτηση με τους SBDS και SBDD.

Αναφορές

1. Ν. Ποθητός: Αυτόματη κατασκευή ωρολογίων προγραμμάτων μέσω προγραμματισμού με πε-ριορισμούς. Πτυχιακή εργασία, Τμήμα Πληροφορικής και Τηλεπικοινωνιών, ΕΚΠΑ,2005. http://www.di.uoa.gr/~pothitos/skoptiko.

2. J. Crawford, M. L. Ginsberg, E. Luks, and A. Roy: Symmetry-breaking predicates for search problems. In L. C. Aiello, J. Doyle, and S. Shapiro, editors, KR'96: Principles of Knowledge Representation and Reasoning, Morgan Kaufmann, San Francisco, California,

1996.

3. Ν. Ποθητός: Naxos Solver: Εγχειρίδιο Χρήσης, Διαθέσιμο στη διεύθυνση http://www.di.uoa.gr/_pothitos/naxos.

4. S. Russell and P. Norvig: Artificial Intelligence A Modern Approach, Prentice Hall, 2003.

5. P. van Beek, editor: Principles and Practice of Constraint Programming - CP 2005, 11th In-ternational Conference, CP 2005, Sitges, Spain, October 1-5, 2005, Proceedings, volume 3709 of Lecture Notes in Computer Science, 2005. Springer. ISBN 3-540-29238-1.

6. D. A. Cohen, P. Jeavons, C. Jefferson, K. E. Petrie, and B. M. Smith: Symmetry definitions for constraint satisfaction problems. In van Beek [5], pages 17–31. ISBN 3-540-29238-1.

7. R. Backofen and S. Will: Excluding symmetries in concurrent constraint programming. In Workshop on Modeling and Computing with Concurrent Constraint Programming held in conjunction with CP 98, 1998.

8. F. Focacci and M. Milano: Global cut framework for removing symmetries. In T. Walsh Principles and Practice of Constraint Programming - CP 2001, volume Lecture Notes in Computer Science 2239, pages 77–92. Springer, 2001.

9. T. Fahle, S. Schamberger, and M. Sellmann: Symmetry breaking. In T. Walsh, editor, Principles and Practice of Constraint Programming - CP 2001, volume Lecture Notes in Computer Science 2239, pages 93–107. Springer, 2001.

10. R. Backofen and S. Will: Excluding symmetries in constraint-based search. In J. Jaffar, edi-tor, Principles and Practice of Constraint Programming - CP ’99, volume Lecture Notes in Computer Science 1713, pages 73–87. Springer, 1999.

11. I. P. Gent and B. M. Smith: Symmetry breaking in constraint programming. In Proceeding of European Conference on Artificial Intelligence - ECAI 2000, pages 599–603. IOS press, 2000.

Εξαγωγή Ζευγών Ερώτησης – Απάντησης από Forum και Αυτόματη Απάντηση Νέων Ερωτήσεων

Μιχαήλ Ν. Ζερβός
std04079@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Η εργασία αυτή ασχολείται με τα ζεύγη δημοσιεύσεων ερώτησης - απάντησης που εμφανίζονται σε online fora. Παρουσιάζεται ένα σύστημα με δύο λειτουργίες: α) την εξαγωγή γνώσης, με την μορφή ερωταποκρίσεων, από ένα forum και β) την παρακολούθηση του forum για εμφάνιση νέων ερωτήσεων και αυτόματη απάντηση αυτών.

Λέξεις-Κλειδιά: Αναγνώριση ερώτησης, Ζεύγη ερωτήσεων – απαντήσεων, Εξόρυξη γνώσης, Αυτόματη απάντηση ερωτήσεων, Παρόμοιες ερωτήσεις.

Επιβλέποντες:
Παναγιώτης Σταματόπουλος, Επίκουρος Καθηγητής
Ιζαμπώ Καράλη, Επίκουρη Καθηγήτρια

1. Εισαγωγή

Οι online περιοχές συζήτησης (forum) αποτελούν σημαντικό κομμάτι του παγκόσμιου ιστού. Ο αριθμός των μηνυμάτων που δημοσιεύονται σε διάφορα fora είναι τεράστιος. Σκοπός της εργασίας αυτής είναι να εκμεταλλευτεί την υπάρχουσα γνώση ενός forum και να παρέχει άμεσα και αυτοματοποιημένα απαντήσεις σε νέες ερωτήσεις που δημοσιεύονται. Το πρόβλημα μπορεί να χωριστεί σε δύο φάσεις. Αρχικά πρέπει να αναγνωριστούν και να εξαχθούν τα ζεύγη ερωτήσεων – απαντήσεων που υπάρχουν ήδη δημοσιευμένα. Στην συνέχεια, πρέπει να παρακολουθείται το forum και μόλις εμφανιστεί μια καινούρια ερώτηση που έχει ήδη απαντηθεί, να γίνεται αυτόματα η κατάλληλη δημοσίευση.

Οι συγγραφείς της εργασίας [1] παρουσιάζουν μια μέθοδο για την αναγνώριση ερωτήσεων και των απαντήσεων τους σε online fora. Για την αναγνώριση των ερωτήσεων βασίζονται στα Labeled Sequential Patterns (LSP) [2]. Για τη σύνδεση ερωτήσεων και απαντήσεων προτείνουν έναν αλγόριθμο διάδοσης σκορ πάνω σε ένα γράφο των δημοσιεύσεων. Τα αρχικά σκορ υπολογίζονται με βάση το μοντέλο Query Likelihood ή το KL-Divergence. Οι εργασίες [3], [4] αντιμετωπίζουν αντίστοιχα προβλήματα, όπως αναγνώριση ερωτήσεων – απαντήσεων σε μηνύματα ηλεκτρονικού ταχυδρομείου. Και στις τρεις αυτές εργασίες, γίνεται χρήση ενός SVM ταξινομητή για την αντιστοίχηση ερωτήσεων – απαντήσεων.

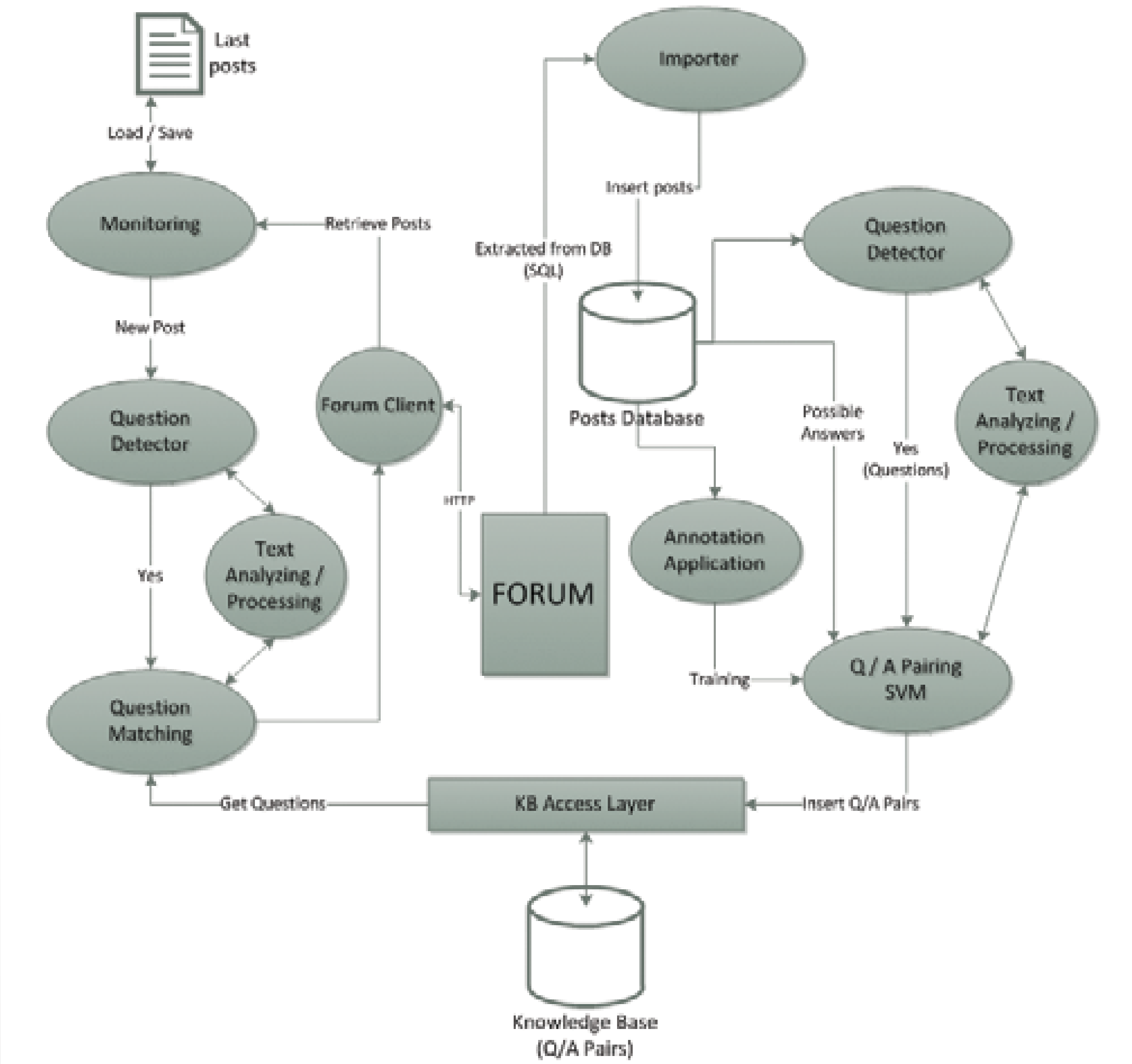
Το σύστημα που παρουσιάζεται στην παρούσα εργασία δημιουργήθηκε με σκοπό την εφαρμογή του στο forum του μαθήματος «Εισαγωγή στον Προγραμματισμό» του Τμήματος. Πρόκειται για το πρώτο (στο βαθμό που είναι δυνατό να γνωρίζω) ολοκληρωμένο σύστημα αυτόματης απάντησης ερωτήσεων σε forum, καθώς και το πρώτο σύστημα αναγνώρισης ζευγών ερώτησης – απάντησης στα Ελληνικά. Τα αποτελέσματα που προκύπτουν είναι ικανοποιητικά και αποδεικνύουν ότι μπορεί να υπάρξει πρακτική εφαρμογή του συστήματος.

Στο κεφάλαιο 2 υπάρχει μια γενική περιγραφή του συστήματος. Στα κεφάλαια 3, 4, 5 και 6 αναλύονται τα βασικότερα τμήματα του συστήματος. Τέλος στο κεφάλαιο 7 παρουσιάζονται τα αποτελέσματα των πειραμάτων.

2. Δομή συστήματος

Η εργασία αυτή χωρίζεται σε δύο βασικά μέρη:

1. Στην εξαγωγή ερωταποκρίσεων από τις υπάρχουσες δημοσιεύσεις και δημιουργία της βάσης γνώσης (δεξί μισό στο Διάγραμμα 1).
2. Στην αντιμετώπιση νέων ερωτήσεων που εμφανίζονται στο forum (αριστερό μισό στο Διάγραμμα 1)



Διάγραμμα 1: Διάγραμμα λειτουργίας

Ο μόνος τρόπος επικοινωνίας αυτών των δύο μερών είναι η βάση γνώσης. Κατά την εξαγωγή γνώσης, η βάση γεμίζει με τα ζεύγη ερωτήσεων – απαντήσεων. Όταν εμφανιστούν νέες δημοσιεύσεις / ερωτήσεις τότε το πρόγραμμα αντιμετώπισης τους χρησιμοποιεί την βάση για να εντοπίσει μια παρόμοια ερώτηση και να δημοσιεύσει την αντίστοιχη απάντηση. Κάθε ένα από τα υποσυστήματα που φαίνονται στο Διάγραμμα 1 παρουσιάζεται στα επόμενα κεφάλαια.

3. Ανάλυση και επεξεργασία κειμένου

Το υποσύστημα ανάλυσης και επεξεργασίας κειμένου παρέχει μια σειρά από λειτουργίες όπως: αφαίρεση τονισμού, διαγραφή stopwords, αφαίρεση ή/και αντικατάσταση του κώδικα μορφοποίησης BBCode που χρησιμοποιείται στο forum, καθώς και ειδικών χαρακτήρων HTML (πχ √) και emoticons. Έκτος από αυτά, αναλαμβάνει και ορισμένες ακόμα λειτουργίες που αναλύονται παρακάτω.

3.1 Stemmer

Το Stemming είναι η διαδικασία μετατροπής μια λέξης στη ρίζα της, ανεξαρτήτως πτώσης ή χρόνου. Για παράδειγμα οι λέξεις «τηλέφωνα», «τηλεφωνώ», «τηλέφωνο», «τηλεφώνησαν» θα αντιστοιχούν στην ίδια ρίζα «τηλεφων». Έτσι μπορεί να μειωθεί δραστικά ο αριθμός των λέξεων στο λεξικό και η διαστασιμότητα (dimensionality) της αναπαράστασης του κειμένου. Για τις ανάγκες της εργασίας χρησιμοποιήθηκε μια τροποποιημένη έκδοση του stemmer που περιγράφεται στο [5].

3.2 Ομοιότητα συνημιτόνου σε συλλογή εγγράφων

Για την αναπαράσταση των δημοσιεύσεων επιλέχτηκε το μοντέλο διανυσματικού χώρου (Vector Space Model). Κάθε stemmed όρος t_i που εμφανίζεται στις δημοσιεύσεις εισάγεται στο λεξικό $T = (t_1 , t_2 , \dots , t_N)$. Έτσι μια δημοσίευση d_j αναπαριστάται από ένα διάνυσμα N διαστάσεων όπου κάθε διάσταση αντιστοιχεί σε έναν όρο του λεξικού: $d_j = (w_{(1,j)}, w_{(2,j)}, \dots ,w_{(N,j)})$. Το βάρος $w_{(1,j)}$ είναι το TF-IDF (Term Frequency – Inverse Document Frequency) σκορ:

$$w_{(1,j)} = tfidf\,(i,j) = tf\,(i,j) * idf\,(i)$$

όπου $tf\,(i,j)$ είναι το πλήθος των εμφανίσεων $tc\,(i,j)$ του όρου t_i κανονικοποιημένο ως προς το μέγεθος του κειμένου και $idf\,(i)$ είναι η αντίστροφη συχνότητα εγγράφου για τον όρο t_i. Το IDF είναι ένα μέτρο της σημαντικότητας του όρου t_i στη συλλογή εγγράφων C.

$$tf(i,j)=\frac{tc\ (i,j)}{\sum_k tc\ (k,j)}=\frac{tc\ (i,j)}{\left|d_j\right|}\ ,\ idf(i)=\log\frac{|C|}{1+\left|\left\{j\ :\ t_i\in d_j\right\}\right|}$$

Το μέτρο της ομοιότητας που χρησιμοποιείται είναι η ομοιότητα συνημιτόνου. Κάθε δημοσίευση αποτελεί ένα έγγραφο d και μια καινούρια δημοσίευση αποτελεί την επερώτηση (query) q. Η ομοιότητα συνημιτόνου υπολογίζει την γωνία μεταξύ των διανυσμάτων q και d.

$$cos\_sim(\bar{q},\bar{d})=\frac{<\bar{q}\cdot\bar{d}>}{\|\bar{q}\| * \|\bar{d}\|}=\hat{q}\cdot\hat{d}$$

Όπου $q\hat{}$, $d\hat{}$ τα αντίστοιχα κανονικοποιημένα διανύσματα:

$$\hat{q}=\frac{\bar{q}}{\|\bar{q}\|},\ \hat{d}=\frac{\bar{d}}{\|\bar{d}\|}$$

Το να υπολογίζει κανείς κάθε φορά το $cos\_sim\ (q,\ d)$ μεταξύ του query και κάθε εγγράφου d με αυτό τον τρόπο είναι πάρα πολύ ακριβό. Για τον ταχύτερο υπολογισμό της ομοιότητας, τα διανύσματα $\dot{d}_i$ κάθε εγγράφου μια συλλογής κειμένων εισάγονται ως στήλες σε ένα πίνακα A. Έστω ότι υπάρχουν N όροι και M έγγραφα. Τότε ο πίνακας A έχει μέγεθος $N\times M$. Το στοιχείο $(i,\ j)$ του πίνακα A αποτελεί το TF-IDF score του όρου t_i στο έγγραφο d_j. Ο πίνακας αυτός υπολογίζεται μια φορά κατά την εκκίνηση του συστήματος.

$$\bar{s}=\hat{q}\ *A=\left(q_1,q_2,\ldots,q_n\right)*\begin{bmatrix} d_{1,1} & d_{2,1} & \ldots & d_{m,1} \\ \ldots & & & \\ d_{1,n} & d_{2,n} & \ldots & d_{m,n} \end{bmatrix}$$

Πολλαπλασιάζοντας το διάνυσμα $\hat{q}$ με τον A παίρνουμε ένα διάνυσμα $\bar{s}$, το οποίο αποτελείται από την ομοιότητα του query με κάθε έγγραφο της συλλογής.

4. Αναγνώριση ερωτήσεων

Η διαδικασία της ταξινόμησης μιας δημοσίευσης ως ερώτηση προς αποθήκευση γίνεται σε δύο στάδια. Αρχικά, η δημοσίευση ταξινομείται ως ερωτηματική αν περιέχει τουλάχιστον μια πρόταση που είναι ερώτηση. Στην συνέχεια, εκτιμάται η «σημαντικότητα» της δημοσίευσης / ερώτησης, για να αποφασιστεί αν εισαχθεί στην βάση γνώσης. Για παράδειγμα και οι δύο ακόλουθες δημοσιεύσεις ταξινομήθηκαν αρχικά ως ερωτηματικές. Όμως μόνο η πρώτη έχει αξία να αποθηκευτεί στη βάση γνώσης.

Δημοσίευση 1: *«Τι είναι το segmentation fault;»*

Δημοσίευση 2: *«Δηλαδή μπορώ να έρθω αργότερα;»*

Το πρώτο στάδιο του αλγορίθμου βασίζεται σε έναν συνδυασμό ευριστικών συναρτήσεων και σε μια σειρά από patterns ερωτηματικών προτάσεων που προσπαθεί να ταιριάξει. Η δημοσίευση χωρίζεται σε προτάσεις και κάθε πρόταση χωρίζεται σε φράσεις. Για την αναγνώριση χρησιμοποιούνται η πρώτη και η τελευταία φράση. Οι ενδιάμεσες φράσεις συνήθως περιέχουν επεξηγήσεις και μπορούμε να τις παραλείψουμε. Οι ευριστικές συναρτήσεις που χρησιμοποιούνται είναι οι εξής:

- Ύπαρξη ερωτηματικού στο τέλος της πρότασης
- Αρχή της φράσης με ερωτηματική λέξη κλειδί (π.χ. πότε)
- Αρχή της φράσης με μια λιγότερο πιθανή ερωτηματική λέξη

Επίσης ελέγχεται αν η φράση ταιριάζει σε ένα από τα είκοσι (20) patterns που δημιουργήθηκαν για το σκοπό αυτό. Ένα παράδειγμα αποτελεί το:

```
"^ΜΠΟΡΕΙ\s+(ΚΑΠΟΙΟΣ|ΚΑΝΕΙΣ)\s+(ΝΑ)\s"
```

Συνδυάζοντας τα αποτελέσματα των ευριστικών και το κατά πόσο βρέθηκε ένα pattern, υπολογίζεται μια τιμή (confidence) που εκφράζει κατά πόσο η πρόταση αποτελεί ερώτηση.

Στο δεύτερο στάδιο, ο αλγόριθμος αποκλείει ερωτήσεις που δεν έχουν ιδιαίτερη αξία για να αποθηκευτούν. Στο στάδιο αυτό ελέγχονται οι εξής παράμετροι:

- Χρήστης που έκανε την δημοσίευση
- Μέγεθος δημοσίευσης

- Αριθμός παραθέσεων (quote) στην δημοσίευση
- Ερώτηση μέσα σε εισαγωγικά ή παρενθέσεις

Συνήθως, οι ερωτήσεις γίνονται από τους φοιτητές και όχι από καθηγητές ή συνεργάτες μαθήματος. Επίσης, έχει παρατηρηθεί ότι οι «καλές» ερωτήσεις δεν είναι πολύ μεγάλες σε μέγεθος, ούτε πολύ μικρές (π.χ. «Σωστά;»). Ακόμα, οι ερωτηματικές δημοσιεύσεις σπάνια κάνουν quote κάποια άλλη δημοσίευση. Τέλος, οι ερωτήσεις μέσα σε εισαγωγικά, συνήθως έχουν αντίστοιχη σημασία με τα quotes, ενώ αυτές που βρίσκονται μέσα σε παρενθέσεις δεν αποτελούν το βασικό μήνυμα της δημοσίευσης, οπότε αγνοούνται. Με βάση αυτά, προσαρμόζεται ανάλογα το confidence της ερώτησης. Όσες ερωτήσεις δεν αποκλείστηκαν, αποθηκεύονται στην βάση δεδομένων. Πριν αναζητηθούν οι απαντήσεις τους, μπορεί να υπάρξει ένας ενδιάμεσος χειροκίνητος έλεγχος, μέσω της διεπαφής που δημιουργήθηκε.

5. Εύρεση ζευγών ερωτήσεων – απαντήσεων

Για την αναγνώριση της απάντησης σε μια δεδομένη ερώτηση υλοποιήθηκε μια μέθοδος που βασίζεται στις μηχανές διανυσμάτων υποστήριξης (SVM). Ως υποψήφιες απαντήσεις σε μια ερωτηματική δημοσίευση, θεωρούνται οι 10 επόμενες δημοσιεύσεις μετά την ερώτηση. Για μια δεδομένη ερωτηματική δημοσίευση Q και μια υποψήφια απάντηση A, τα χαρακτηριστικά που χρησιμοποιούνται είναι τα εξής:

1. Αν η Q γίνεται quote στην A
2. Είδος χρήστη (φοιτητής ή όχι) που δημοσίευσε την A
3. Πλήθος posts που μεσολαβούν μεταξύ A και Q
4. Ομοιότητα συνημιτόνου μεταξύ Q και A
5. Αν η A περιέχει code tag

Η επιλογή του $2^{ου}$ χαρακτηριστικού έγινε με βάση το ότι οι καθηγητές και οι συνεργάτες του μαθήματος συνηθίζουν να απαντούν, παρά να κάνουν ερωτήσεις. Επίσης, το γεγονός ότι μια δημοσίευση περιέχει κομμάτι κώδικα, μπορεί να είναι μια ένδειξη για το αν είναι απάντηση ($5^{ο}$ χαρακτηριστικό).

Ο πυρήνας που επιλέχτηκε για τον SVM, είναι ο Radial Basis Function:

$K(\overline{x},\overline{y})=e^{-\gamma\|\overline{x}-\overline{y}\|^2}, \gamma>0.$ Η αποτελεσματικότητα του SVM βασίζεται κατά πολύ στην επιλογή των σωστών παραμέτρων [6]. Για τον προσδιορισμό του βέλτιστου συνδυασμού, δοκιμάζονται διάφορες τιμές τους από ένα πλέγμα. Κάθε συνδυασμός ελέγχεται με N-fold cross validation στο σύνολο εκπαίδευσης. Ο συνδυασμός παραμέτρων που δίνει το καλύτερο αποτέλεσμα χρησιμοποιείται για την εκπαίδευση του μοντέλου.

6. Αντιμετώπιση νέων ερωτήσεων

Τα υποσυστήματα του προγράμματος απάντησης νέων ερωτήσεων φαίνονται στο αριστερό μισό του διαγράμματος (Διάγραμμα 1). Ο Forum Client αναλαμβάνει την επικοινωνία με το forum μέσω του πρωτοκόλλου HTTP, μιμούμενος τη λειτουργία ενός web browser. Το υποσύστημα παρακολούθησης (Forum Monitoring) ελέγχει το forum για νέες δημοσιεύσεις ανά τακτά χρονικά διαστήματα. Ο Question Classifier είναι υπεύθυνος για την ταξινόμηση ενός κειμένου ως ερώτηση ή όχι και είναι ο ίδιος με αυτόν που αναλύθηκε στο κεφάλαιο 4. Το βασικότερο υποσύστημα είναι το Question Matching. Αυτό έχει ως σκοπό να ταιριάξει μια καινούρια ερώτηση με κάποια από τις υπάρχουσες στη βάση γνώσης, υπολογίζοντας την ομοιότητα συνημιτόνου, όπως παρουσιάστηκε στο κεφάλαιο 3.2.

7. Αποτελέσματα

Από το forum του μαθήματος «Εισαγωγή στον προγραμματισμό», έγιναν annotate 1215 δημοσιεύσεις από 36 διαφορετικά νήματα με εφαρμογή που δημιουργήθηκε για αυτό το σκοπό. Για τα πειράματα χρησιμοποιήθηκε υπολογιστής με επεξεργαστή Intel Core i7 Q720 @ 1.6 GHz με 6GB RAM.

7.1 Αναγνώριση ερωτήσεων

Κάθε δημοσίευση που περιείχε τουλάχιστον μια ερώτηση (ευθεία ή πλάγια) είχε χαρακτηριστεί ως ερώτηση. Από τις 1215 δημοσιεύσεις, οι 455 περιείχαν τουλάχιστον μια ερώτηση. Οι δημοσιεύσεις περιείχαν κατά μέσο όρο 3.06 προτάσεις. Ο αλγόριθμος που δοκιμάζεται είναι αυτός που αποφασίζει αν μια δημοσίευση περιέχει τουλάχιστον μια ερώτηση. Στον πίνακα που ακολουθεί εμφανίζονται τα αποτελέσματα για διάφορες τιμές κατωφλιού.

Η μέθοδος αναγνώρισης ερωτήσεων που προτείνεται δίνει πολύ καλά αποτελέσματα, ελαφρώς χειρότερα από της μεθόδου [1]. Μια ευθεία σύγκριση των μεθόδων δεν θα ήταν σωστή αφού τα αποτελέσματα είναι σε διαφορετικά σύνολα δεδομένων και σε διαφορετικές γλώσσες. Ο μέσος χρόνος απόφασης ανά πρόταση είναι 0.37ms. Χωρίς τη προεπεξεργασία, ο μέσος χρόνος απόφασης ανά πρόταση 0.13ms. Για τιμή κατωφλιού 0.7, ο αλγόριθμος αναγνώρισης ερωτηματικών δημοσιεύσεων επέστρεψε τις 416 από τις 455. Ο αλγόριθμος επιλογής «καλών» ερωτήσεων κράτησε τις 140 από αυτές. Έτσι τελικά, αν κάποιος θέλει να ελέγξει τις ερωτήσεις πριν αυτές μπουν στην βάση γνώσης, χρειάζεται να διαβάσει μόνο το 11.6% των αρχικών δημοσιεύσεων. Σε σύγκριση με το [4], ο αλγόριθμος της παρούσας εργασίας κάνει παραπλήσιο χρόνο ανά πρόταση.

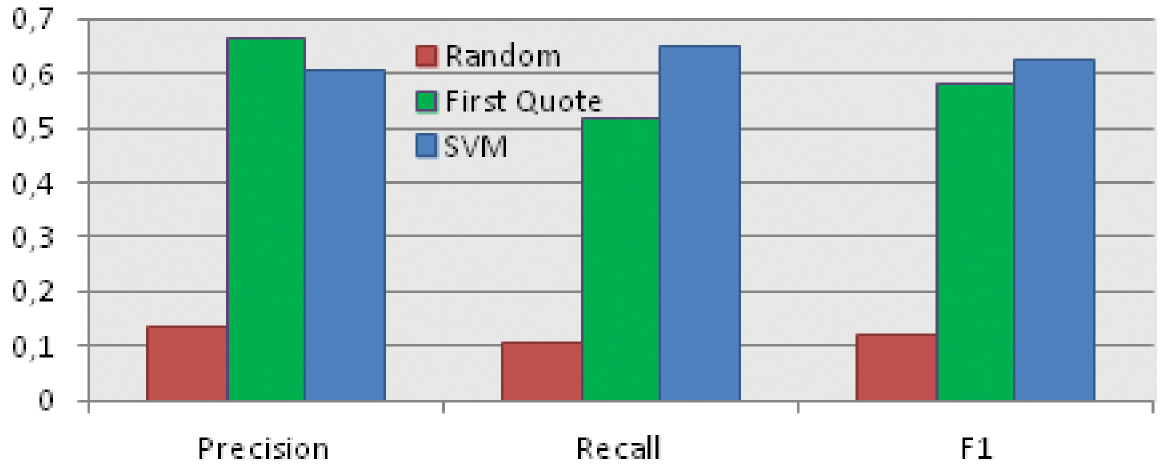
Κατώφλι	Precision	Recall	F_1
0	0.8516	**0.9208**	0.8848
0.5	0.891	0.9164	0.9035
0.7	0.8965	0.9142	**0.9053**
0.9	**0.968**	0.2747	0.428

Πίνακας 1: Αποτελέσματα αναγνώρισης ερωτήσεων

7.2 Αναγνώριση ζευγών ερώτησης – απάντησης

Για την αξιολόγηση του SVM αναγνώρισης ζευγών ερώτησης – απάντησης, χρησιμοποιήθηκε το annotated σύνολο δεδομένων που περιγράφηκε προηγουμένως. Από το σύνολο των 1215 ερωτήσεων, σημειώθηκαν απαντήσεις για τις 225 από αυτές. Ορισμένες ερωτήσεις έχουν πολλαπλές απαντήσεις, έτσι δημιουργήθηκαν 289 ζεύγη ερωτήσεων - απαντήσεων, τα οποία αποθηκεύτηκαν ως θετικά δείγματα. Για αρνητικά δείγματα χρησιμοποιήθηκαν οι 10 επόμενες δημοσιεύσεις (υποψήφιες απαντήσεις που δεν είναι σωστές) από κάθε ερώτηση. Έτσι, δημιουργήθηκε ένα σύνολο με 289 θετικά δείγματα και 1845 αρνητικά. Τα πειράματα έγιναν σε ερωτήσεις για τις οποίες υπήρχε τουλάχιστον μια σωστή απάντηση. Τα αποτελέσματα κάθε αλγορίθμου φαίνονται στο Διάγραμμα 2. Για λόγους σύγκρισης δημιουργήθηκαν δύο απλοί αλγόριθμοι:

- Random: Επιλογή τυχαίας απάντησης από τις 10 υποψήφιες
- First Quote: Επιλογή της πρώτης υποψήφιας απάντησης που παραθέτει την ερώτηση



Διάγραμμα 2: Αποτελέσματα αναγνώρισης ζευγών ερώτησης - απάντησης

Αναφορές

1. G. Cong, L. Wang, C.Y. Lin, Y.I. Song, and Y. Sun, «Finding question-answer pairs from online forums» in Proceedings of the 31st International Conference on Research and Development in Information Retrieval (SIGIR 08), 2008, p. 467-474.

2. J. Pei et al., «PrefixSpan: Mining sequential patterns efficiently by prefix-projected pattern growth» in Proceedings of the 10th International Conference on Computer Communication Networks (ICCCN 01), 2001, p. 0215.

3. L. Shrestha and K. McKeown, «Detection of question-answer pairs in email conversations» in Proceedings of the 20th International Conference on Computational Linguistics (COLING 04), 2004, p. 889.

4. H. Kwong and N. Yorke-Smith, «Detection of imperative and declarative question-answer pairs in email conversations», in Proceedings of the 21st International Jont Conference on Artifical intelligence (IJCAI 09), Pasadena, California, USA, 2009.

5. G. Ntais, «Development of a Stemmer for the Greek Language», Stockholm University - Royal Institute of Technology, Stockholm, 2006.

6. C. Hsu, C. Chang, and C. Lin, A practical guide to support vector classification, 2003.

Μέθοδοι και Εργαλεία Προσομοίωσης Γονιδιακών Κυκλωμάτων στη Συστημική Βιολογία

Ελένη Ε. Καραμασιώτη
std06115@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Μελετήθηκαν μέθοδοι και εργαλεία μοντελοποίησης και προσομοίωσης της συμπεριφοράς πολύπλοκων γονιδιακών κυκλωμάτων, μέσω των εργαλείων iBioSim και SPiM που βασίζεται στον π-λογισμό. Σχετικά με το SPiM, παρουσιάζεται μια εναλλακτική προσέγγιση των γονιδιακών κυκλωμάτων βασισμένη στην αναλογία ενός γονιδίου με μια λογική πύλη. Δημιουργήθηκε μια βιβλιοθήκη πρότυπων υλοποιήσεων γονιδιακών πυλών σε SPiM, όπου ενσωματώθηκαν σαν δομικά στοιχεία και πρότυπες υλοποιήσεις γενετικών δικτύων για τις λειτουργίες βασικών λογικών πυλών NAND και NOR. Παρά τις ορθογώνιες προσεγγίσεις των δύο εργαλείων, επιτύχαμε εξομοίωση των αποτελεσμάτων προσομοίωσης του ίδιου κυκλώματος μέσω αντιστοιχίας των παραμέτρων του iBioSim σε αυτές που χρησιμοποιήθηκαν στις περιγραφές γενετικών πυλών του SPiM.

Λέξεις-Κλειδιά: Συστημική Βιολογία, Συνθετική Βιολογία, Προσομοίωση, iBioSim, SPiM, Γενετικά κυκλώματα

Επιβλέπων: Ηλίας Μανωλάκος, Αναπληρωτής Καθηγητής

1. Εισαγωγή

Η Μοριακή Βιολογία έχει αποκαλύψει τα τελευταία χρόνια μεγάλο πλήθος πληροφοριών, αυτό όμως δεν αρκεί για την πλήρη κατανόηση της δυναμικής των βιολογικών συστημάτων. Παρότι οι εξελίξεις στις πειραματικές μεθόδους αναμφισβήτητα θα συνεχιστούν με ταχύ ρυθμό, περισσότερη γνώση σχετικά με τη λειτουργία και δυναμική συμπεριφορά των βιολογικών συστημάτων θα προκύψει από ένα συνδυασμό πειραματικών και υπολογιστικών μεθόδων.

Οι μηχανισμοί με τους οποίους ελέγχεται η έκφραση της γενετικής πληροφορίας στους επιμέρους κυτταρικούς τύπους, παραμένουν σε ένα βαθμό άγνωστοι, αφού μέσω σύνθετων γονιδιακών αλληλεπιδράσεων οδηγούν στη δημιουργία πολύπλοκων δικτύων. Σε τέτοιου είδους γονιδιακά ρυθμιστικά δίκτυα, μια συνδυαστική προσέγγιση πειραματικών και υπολογιστικών μεθόδων, μπορεί μέσω επαναληπτικών διαδικασιών να βοηθήσει ουσιαστικά, δημιουργώντας μοντέλα που θα περιγράφουν τη συμπεριφορά του επιθυμητού γονιδιακού δικτύου.

Στην παρούσα εργασία, αναλύονται οι μέθοδοι και τα εργαλεία με τις οποίες μπορούν να επιτευχθούν η μοντελοποίηση και η στοχαστική προσομοίωση της συμπεριφοράς τέτοιων δικτύων. Γίνεται παρουσίαση των δύο εργαλείων που θα χρησιμοποιηθούν στα πλαίσια της εργασίας iBioSim και SPiM, καθώς επίσης και του π-λογισμού, στον οποίο βασίζει τη λειτουργία του το εργαλείο SPiM. Σε ότι αφορά τη μέθοδο του SPiM, παρουσιάζεται μια διαφορετική προσέγγιση των γονιδιακών δικτύων, που στηρίζεται στον παραλληλισμό ενός γονιδίου με μια λογική πύλη. Δημιουργήθηκε μια βιβλιοθήκη πρότυπων υλοποιήσεων με βάση τις οποίες ως δομικά στοιχεία μπορεί να επιτευχθεί η μοντελοποίηση και η προσομοίωση οποιουδήποτε γονιδιακού δικτύου. Η βιβλιοθήκη αυτή εμπλουτίστηκε με πρότυπες υλοποιήσεις και για τις βασικές λογικές πύλες NAND και NOR. Επίσης, προκειμένου να επιτευχθεί στατιστική ομοιότητα ανάμεσα στα αποτελέσματα προσομοίωσης των δύο εργαλείων, προτείνεται μια υβριστική μεθοδολογία αντιστοίχισης των παραμέτρων που χρησιμοποιούνται στα μοντέλα πυλών SPiM σε ρυθμιζόμενες παραμέτρους που χρησιμοποιεί το ανωτέρω εργαλείο iBioSim. Για να αποδειχθεί η αποτελεσματικότητα αυτής της μεθόδου έγινε μοντελοποίηση και προσομοίωση και με τα δύο εργαλεία για μια μεγάλη γκάμα γονιδιακών δικτύων, κάνοντας χρήση των υλοποιημένων βασικών στοιχείων της βιβλιοθήκης πρότυπων υλοποιήσεων και εφαρμόζοντας την προτεινόμενη αντιστοίχιση των παραμέτρων.

2. Δίκτυα ρύθμισης έκφρασης γονιδίων

Παρότι το γονιδιακό υλικό ενός οργανισμού είναι το ίδιο σε κάθε κύτταρο, δεν παράγονται σε κάθε κύτταρο όλες οι πρωτεΐνες σε κάθε χρονική στιγμή, κάτι που καθιστά απαραίτητη την ύπαρξη κάποιου μηχανισμού ελέγχου της έκφρασης κάθε γονιδίου. Η ικανότητα της RNA πολυμεράσης να προσδεθεί στον υποκινητή και να ξεκινήσει τη μεταγραφή του γονιδίου, μπορεί είτε να βοηθηθεί (ενεργοποίηση) είτε να εμποδιστεί (καταστολή) από άλλες πρωτεΐνες, τους μεταγραφικούς παράγοντες. Με τον τρόπο αυτό καταλήγουμε σε γονιδιακά δίκτυα, όπου το προϊόν της έκφρασης του ενός μπορεί να οδηγήσει στην ενεργοποίηση ή την καταστολή ενός άλλου.

Πέρα από τα πειραματικά εργαλεία, υπολογιστικές μέθοδοι για τη μοντελοποίηση και την προσομοίωση των διεργασιών γονιδιακής ρύθμισης, είναι απαραίτητες. Με τη χρήση τέτοιων μεθόδων, η δομή ενός συστήματος γονιδιακής ρύθμισης μπορεί να περιγραφεί με σαφή τρόπο ενώ προβλέψεις για τη συμπεριφορά του μπορούν να γίνονται συστηματικά.

Μπορούμε να επιτύχουμε μοντελοποίηση ενός γονιδιακού δικτύου μέσω διαφορετικών μεθόδων, καθεμία από τις οποίες προσφέρει διαφορετικού είδους πληροφορίες για το μοντέλο και διαφορετικό βάθος μελέτης. Δύο από τις πιο δημοφιλείς μεθόδους είναι αυτή που απεικονίζει το δίκτυο μέσω βελών ενεργοποίησης ή καταστολής ανάμεσα στα γονίδια και η πλέον κλασσική μέθοδος των χημικών αντιδράσεων. Πρόκειται για δύο μεθόδους που χρησιμοποιούν διαφορετικό επίπεδο αφαίρεσης για το δίκτυο που μελετάνε προσφέροντας έτσι γενικότερες (δίκτυο αλληλεπιδράσεων) ή ειδικότερες (δίκτυο αντιδράσεων) πληροφορίες για το σύστημα.

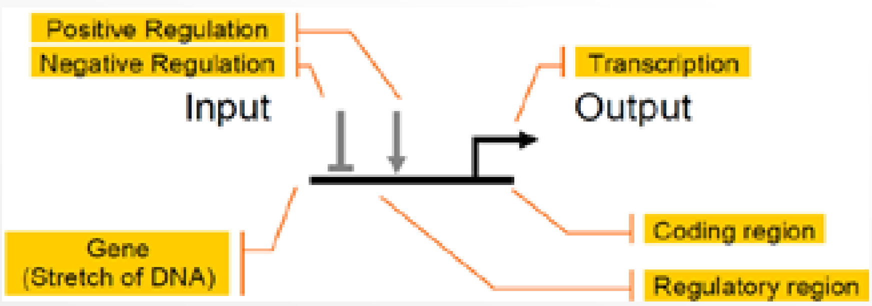
3. Εργαλεία μοντελοποίησης και προσομοίωσης βιολογικών μοντέλων

Η ανάγκη για χειρισμό του τεράστιου όγκου των βιολογικών δεδομένων, αλλά και για την εύκολη ανταλλαγή δεδομένων και μοντέλων ανάμεσα στα μέλη της επιστημονικής κοινότητας, οδήγησε στη δημιουργία των πρώτων εργαλείων μοντελοποίησης και προσομοίωσης βιολογικών συστημάτων. Η δημιουργία τέτοιων εργαλείων, που είχαν τη δυνατότητα να προσομοιώνουν βιολογικά συστήματα με την πάροδο του χρόνου επιτρέποντας στο χρήστη να δημιουργήσει τις συνθήκες προσομοίωσης του συστήματος κατα βούληση, οδήγησε εκτός των άλλων στη μείωση του κόστους και του χρόνου που απαιτούνταν για τη διενέργεια πειραμάτων στο εργαστήριο.

Το πρώτο από τα δύο εργαλεία που χρησιμοποιήθηκαν στα πλαίσια αυτής της εργασίας, είναι το iBioSim. Το εργαλείο αυτό δίνει τη δυνατότητα για μοντελοποίηση, ντετερμινιστική και στοχαστική προσομοίωση βιολογικών συστημάτων αλλά και συμπερασμό αιτιωδών γονιδιακών δικτύων με βάση συγκεκριμένα δεδομένα. Το δεύτερο εργαλείο που χρησιμοποιήθηκε είναι το SPiM, το οποίο περιλαμβάνει μια γλώσσα προγραμματισμού, βασισμένη στο μαθηματικό φορμαλισμό του π-λογισμού, για το σχεδιασμό και την προσομοίωση υπολογιστικών μοντέλων για βιολογικές διαδικασίες. Βασική ιδέα του φορμαλισμού του π-λογισμού είναι πως οι οντότητες επικοινωνούν μεταξύ τους μέσω καναλιών επικοινωνίας, επιτρέποντας έτσι σε ένα σύστημα χημικών αντιδράσεων να ξαναγραφεί ως μια συλλογή από διαδικασίες, οι οποίες αλληλεπιδρούν μεταξύ τους μέσω αποστολής μηνυμάτων (?a : είσοδος στο κανάλι a, !a : έξοδος από το κανάλι a). Βασικό πλεονέκτημα της μοντελοποίησης μέσω διεργασιών π-λογισμού έναντι της μοντελοποίησης μέσω αντιδράσεων, είναι η απλότητα των μοντέλων που προκύπτουν, κάτι που είναι εμφανές κυρίως σε μεγαλύτερα συστήματα.

4. Απεικόνιση δικτύων αλληλεπίδρασης γονιδίων με τη μορφή γονιδιακών πυλών

Στην ενότητα αυτή παρουσιάζεται μια προσέγγιση μοντελοποίησης γονιδιακών κυκλωμάτων που στηρίζεται στην ιδέα πως το κάθε γονίδιο χαρακτηρίζεται από τη ρυθμιστική του είσοδο (ενεργοποίηση/καταστολή) και έξοδο και έτσι ένα ολόκληρο δίκτυο αλληλεπίδρασης γονιδίων μπορεί να θεωρηθεί ως μια σύνθεση από γονίδια-πύλες. Σύμφωνα με τη μέθοδο αυτή, δεν μοντελοποιούνται οι αντιδράσεις ολόκληρου του συστήματος αλλά κάθε τμήμα (γονίδιο) του συστήματος ξεχωριστά, κάτι που διευκολύνει σημαντικά μια πιθανή αλλαγή ή επέκταση του μοντέλου.



Εικόνα 1: Το γονίδιο ως πύλη [5]

Σε αυτή τη βάση και χρησιμοποιώντας τη γλώσσα του εργαλείου SPiM, δημιουργήσαμε μια σειρά από πρότυπες υλοποιήσεις για τα βασικά στοιχεία που μπορεί να περιλαμβάνει ένα γονιδιακό κύκλωμα. Παρουσιάζονται οι αντίστοιχες SPiM διεργασίες [3].

- Γονίδιο-πύλη χωρίς ρυθμιστική είσοδο (nullp)

```
nullp( cst:float, p:proc()) =
delay@cst; (p() | nullp(p))        (* constitutive production *)
```

- Γονίδιο-πύλη με αρνητική ρύθμιση έκφρασης (negp)

```
negp(a:chan(inh:float,cst:float,rbnd:float),p:proc()) =
do ?a*rbnd; delay@inh; negp(a,(inh,cst,rbnd), p)
                                  (* repression – no production *)
or delay@cst; (p() | negp(a,(inh,cst,rbnd),p))
                                     (* constitutive production *)
```

- Γονίδιο-πύλη με θετική ρύθμιση έκφρασης (posp)

```
posp(a:chan,(cst:float,acp:float,abnd:float),p:proc()) =
do ?a*abnd; delay@acp; (p() | posp(a,(cst,acp,abnd),p))
                                        (* activated production *)
or delay@cst; (p() | posp(a,(cst,acp,abnd),p))
                                     (* constitutive production *)
```

- Απλός μεταγραφικός παράγοντας (tr)

```
tr(dk:float, b:chan() ) =
do !b; tr(dk,b)                 (* binding to some promoter site*)
or delay@dk                                   (* degradation *)
```

- Μεταγραφικός παράγοντας που μπορεί να κατασταλεί (rtr)

```
rtr(dk:float, b:chan(), r:chan()) =
do !b; rtr(dk,b,r)             (* binding to some promoter site *)
or !r                                     (* being repressed *)
or delay@dk                                   (* degradation *)
```

- Καταστολέας μεταγραφικών παραγόντων (rep)

```
rep(r:chan(), rbnd:float ) =
?r*rbnd; rep(r,rbnd)             (*binding to some transcription
                                          factor and repress it*)
```

5. Συγκριτική μελέτη προσομοίωσης δικτύων αλληλεπίδρασης γονιδίων μέσω iBioSim και SPiM

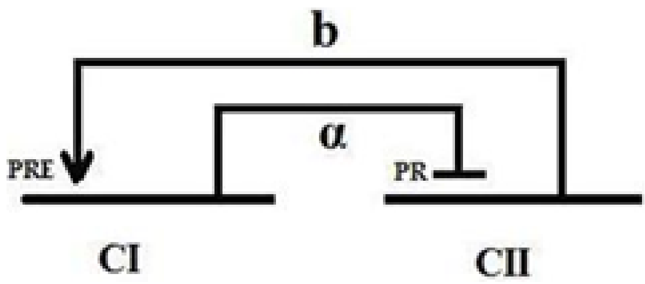
Βασική προϋπόθεση για να επιτύχουμε συμφωνία αποτελεσμάτων όταν προσεγγίζουμε ένα σύστημα με τα δύο προαναφερθέντα εργαλεία, είναι να εξασφαλίζουμε τις ίδιες συνθήκες, αντιστοιχίζοντας τις τιμές των παραμέτρων των δύο εργαλείων, στο βαθμό όπου αυτό είναι εφικτό.

Ερμηνεύοντας τη σημασία καθεμίας από τις παραμέτρους τόσο για το iBioSim (έκδοση 1.5) όσο και για το SPiM, προέκυψαν ζεύγη από παραμέτρους που επηρεάζουν με τον ίδιο τρόπο το δίκτυο που μελετάμε. Στο σημείο αυτό θα πρέπει να σημειωθεί, πως παρότι το iBioSim χρησιμοποιεί συγκεκριμένες παραμέτρους που επιλέχθηκαν κατά την υλοποίηση του εργαλείου, οι παράμετροι που αναφέρονται στην εικόνα 2, για το εργαλείο SPiM, αποτελούν παραμέτρους που δημιουργήσαμε και εισάγαμε ως παραμέτρους στα στοιχεία της βιβλιοθήκης πρότυπων υλοποιήσεων, ύστερα από ενδελεχή μελέτη της φύσης των κυκλωμάτων που μοντελοποιούνται.

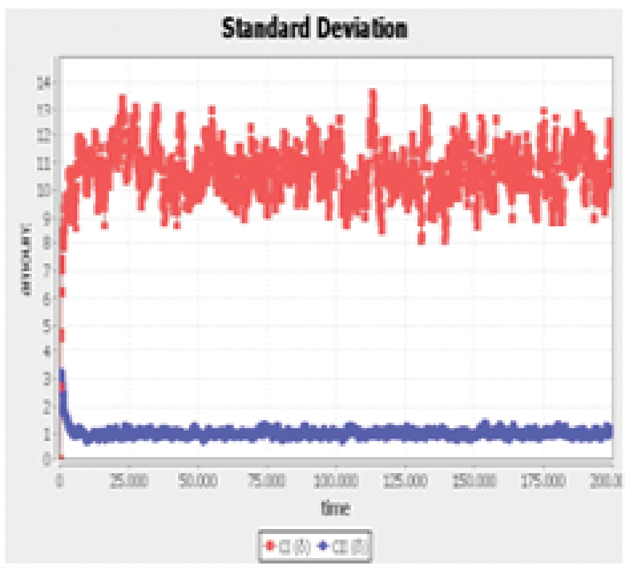
iBioSim	SPiM
Degradation Rate (kd)	Degradation Rate (dk)
Activated Production Rate (ka)	Activated Production Rate (acp)
Basal Production Rate (kb)	Constitutive Production Rate (cst)
Open Complex Production Rate (ko)	Constitutive Production Rate (cst)
Repression binding equillibrium (Kr)	Repression Rate (rbnd)
Activation binding equillibrium (Ka)	Activation Rate (abnd)

Εικόνα 2: Αντιστοίχιση παραμέτρων των εργαλείων iBioSim και SPiM

Για να αξιολογήσουμε τα αποτελέσματα της προτεινόμενης αντιστοίχισης παραμέτρων, υλοποιήσαμε μια σειρά γενετικών κυκλωμάτων τόσο με το iBioSim όσο και με το SPiM, εφαρμόζοντας την προτεινόμενη μέθοδο. Προτού παρουσιάσουμε τα αποτελέσματα που προέκυψαν για ένα από αυτά τα γενετικά κυκλώματα, θα πρέπει να θέσουμε κάποιον περιορισμό για να διακρίνουμε τις περιπτώσεις που έχουμε επαρκή ομοιότητα στα αποτελέσματα των δύο προσεγγίσεων από αυτές που πιθανώς να μην έχουμε. Για το σκοπό αυτό, θεωρήσαμε πως τα αποτελέσματα των δύο εργαλείων παρουσιάζουν ουσιαστικές διαφορές όταν η καμπύλη που προκύπτει από μια εκτέλεση του SPiM είναι κατά απόλυτη τιμή μεγαλύτερη ή ίση από τη μέση τιμή των αποτελεσμάτων από 100 προσομοιώσεις στο iBioSim, αυξημένη κατά το μισό της τυπικής απόκλισης των προσομοιώσεων του iBioSim. Το γενετικό κύκλωμα που θα παρουσιάσουμε αποτελεί τμήμα του δικτύου απόφασης του βακτηριοφάγου λ.

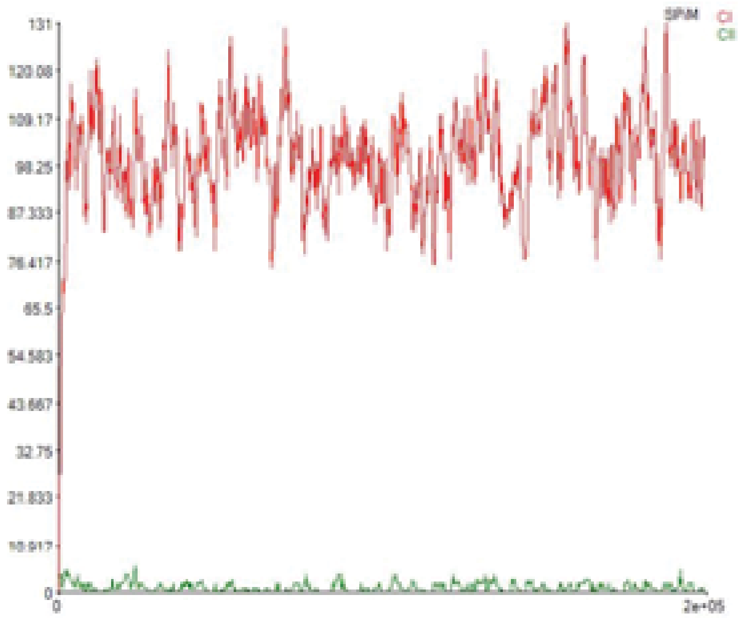


Εικόνα 3: Μοντέλο πυλών για το τμήμα CI CII του δικτύου απόφασης του βακτηριοφάγου λ



Εικόνα 4: Μέσος όρος αποτελεσμάτων απο 100 προσομοιώσεις του μοντέλου της εικόνας 3, στο iBioSim για χρόνο 200.000 δευτερολέπτων

Εικόνα 5: Τυπική απόκλιση αποτελεσμάτων προσομοίωσης του μοντέλου της εικόνας 3 με το iBioSim



Εικόνα 6: Αποτελέσματα προσομοίωσης του μοντέλου της εικόνας 3 για χρόνο 200.000 δευτερολέπτων με το SPiM

Παρατηρώντας τα αποτελέσματα των δύο εργαλείων και λαμβάνοντας υπόψη τον περιορισμό ομοιότητας που έχουμε θέσει, μπορούμε να ισχυριστούμε πως τα δύο εργαλεία παράγουν όμοια αποτελέσματα τόσο ποιοτικά όσο και ποσοτικά.

6. Υλοποίηση λογικών πυλών με γενετικές πύλες

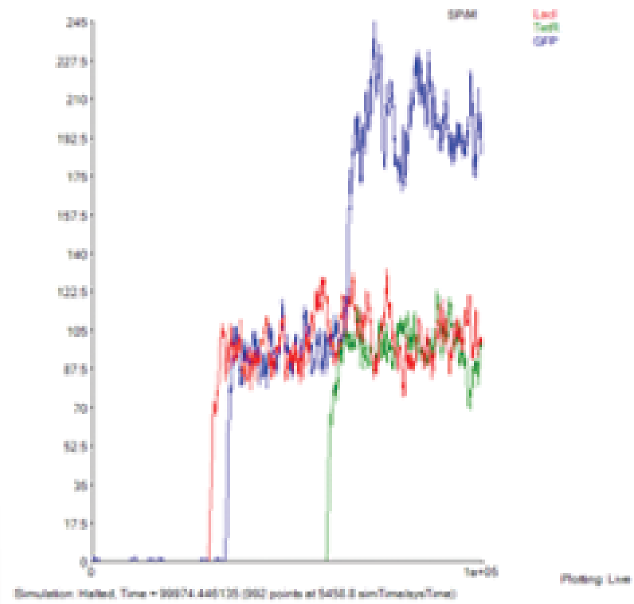
Στην ενότητα αυτή μελετήσαμε τρόπους με τους οποίους μπορούμε να υλοποιήσουμε γονιδιακά κυκλώματα που θα συμπεριφέρονται με τον ίδιο τρόπο που λειτουργούν ηλεκτρονικές λογικές πύλες. Με τον τρόπο αυτό η συμπεριφορά μιας λογικής πύλης μπορεί να υλοποιηθεί από κυκλώματα γενετικών πυλών. Στα πλαίσια αυτής μας της μελέτης, εμπλουτίσαμε τη βιβλιοθήκη πρότυπων υλοποιήσεων, με SPiM περιγραφές των λογικών πυλών NAND και NOR.

```
(* nand gate *)
let nandp(a:chan, b:chan, (inh1:float, cst1:float, rbnd1:float),
(inh2:float, cst2:float, rbnd2:float),
p:proc()) =
( negp(a,(inh1,cst1,rbnd1), p) | negp(b,(inh2,cst2,rbnd2), p) )
        (* Parallel execution of two negative gates with the same
                                product but different promoters*)
```

```
(* nor gate *)
let norp(a:chan, b:chan, (inh:float, cst:float, rbnd:float),
p:proc()) =
do ?a*rbnd; delay@inh; norp(a,b,(inh,cst,rbnd), p)
                                    (* Be repressed by input1 *)
or ?b*rbnd; delay@inh; norp(a,b,(inh,cst,rbnd), p)
                                    (* Be repressed by input2 *)
or delay@cst; (p() | norp(a,b,(inh,cst,rbnd),p))
                                 (* Or produce constitutively *)
```

Χρησιμοποιώντας τις παραπάνω πρότυπες υλοποιήσεις σε συνδυασμό με όσες παρουσιάστηκαν σε προηγούμενες ενότητες, υλοποιήσαμε γενετικά κυκλώματα που συμπεριφέρονται όπως οι βασικές λογικές πύλες NAND, NOR, AND και OR. Στις εικόνες 7 και 9 παρουσιάζονται τα αποτελέσματα προσομοίωσης του γενετικού αυτού κυκλώματος τόσο με το iBioSim όσο και με το SPiM. Για να επιτευχθεί συμφωνία των αποτελεσμάτων εφαρμόστηκε η μέθοδος που προτάθηκε στην προηγούμενη ενότητα.

Εικόνα 7: Μέσος όρος αποτελεσμάτων από 100 προσομοιώσεις του κυκλώματος που αποτελεί τη γονιδιακή υλοποίηση μιας πύλης OR, στο iBioSim για χρόνο 100.000 δευτερολέπτων



Εικόνα 8: Αποτελέσματα προσομοίωσης του κυκλώματος που αποτελεί τη γονιδιακή υλοποίηση μιας πύλης OR, στο SPiM για χρόνο 100.000 δευτερολέπτων

7. Συμπεράσματα – μελλοντικές κατευθύνσεις της έρευνας

Μέσω της παρούσας εργασίας, καταφέραμε όχι μόνο να δημιουργήσουμε μια γέφυρα επικοινωνίας ανάμεσα στα εργαλεία iBioSim και SPiM, αλλά και να διευκολύνουμε τη μοντελοποίηση κυκλωμάτων γονιδιακής ρύθμισης, παρέχοντας στον χρήστη μια βιβλιοθήκη υλοποιήσεων την οποία θα μπορεί να χρησιμοποιήσει για τη μοντελοποίηση οποιουδήποτε γενετικού κυκλώματος, χωρίς να γνωρίζει τις λεπτομέρειες της υλοποίησης, παρά μόνο δίνοντας τις απαιτούμενες εισόδους, εξόδους και παραμέτρους.

Συνεχίζοντας και μετά το πέρας αυτής της πτυχιακής εργασίας στην ίδια κατεύθυνση, θα μπορούσε ως μελλοντικό πλάνο, να αυτοματοποιηθεί σε μεγαλύτερο βαθμό η διαδικασία προσομοίωσης ρυθμιστικών γονιδιακών δικτύων μέσω του SPiM, με τη δημιουργία κάποιου εργαλείου το οποίο παίρνοντας ως είσοδο δεδομένα σχετικά με τα γονίδια που περιλαμβάνονται στο δίκτυο, και τιμές για τις απαιτούμενες παραμέτρους, και χρησιμοποιώντας τα υλοποιημένα στοιχεία της βιβλιοθήκης θα παράγει όλον τον απαραίτητο SPiM κώδικα για την προσομοίωση του δικτύου μέσω του εργαλείου.

Αναφορές

1. Chris J. Myers, «Engineering Genetic Circuits», London, UK : Taylor and Francis, 2010

2. Andrew Phillips, “The SPiM Language”, 2007

3. Ralf Blossey, Luca Cardelli, and Andrew Phillips, “A compositional approach to the stochastic dynamics of gene networks”, Transactions in Computational Systems Biology, 3939:99–122, January 2006. doi: 10.1007/11539452_3

4. R. Blossey, L. Cardelli, and A. Phillips, “Compositionality, stochasticity and cooperativity in dynamic models of gene regulation”. HFPS Journal, Vol. 2, No. 1, pp 17-28, February 2008, doi: 10.2976/1.2804749

5. Luca Cardelli. “Abstract Machines of Systems Biology”, Transactions in Computational Systems Biology, 145–168, 2005, doi:10.1007/11599128_10

6. Luca Cardelli. “Artificial Biochemistry”, Algorithmic Bioprocesses, Springer 2009, part 7, pp 429-462

EucleiDIs :
Μια Γεωμετρική Βιβλιοθήκη για Απλούς Ανθρώπους

Γεώργιος Κολοβέντζος

g.koloventzos@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Στην πτυχιακή που παρουσιάζεται ασχολούμαστε με το να δημιουργήσουμε μια βιβλιοθήκη για την κατασκευή γεωμετρικών σχημάτων και την αντιμετώπιση προβλημάτων της υπολογιστικής γεωμετρίας. Η δημιουργία μιας μη εξαρτημένης βιβλιοθήκης απο το λειτουργικό σύστημα και η εύκολη εγκατάσταση σε υπολογιστές είναι οι προτεραιότητές μας. Εδώ παρουσιάζονται τα βήματα και οι αποφάσεις για τη δημιουργία αυτής της βιβλιοθήκης καθώς και το αποτέλεσμα.

Λέξεις-Κλειδιά: Δισδιάστατος χώρος, Υπολογιστικη γεωμετρία, Ακρίβεια, Σημεία.

Επιβλέπων: Ιωάννης Ζ. Εμίρης, Καθηγητής
Σε συνεργασία με τον Δρ. Χριστόδουλο Φραγουδάκη

1. Εισαγωγή

Η βιβλιοθήκη που αναπτύξαμε εγκαθίσταται εύκολα σε οποιοδήποτε υπολογιστή, δεν εξαρτάται από κάποιο συγκεκριμένο λειτουργικό σύστημα (cross platform application), επιτρέπει την εύκολη δημιουργία/αλλαγή σημείων – πολυγώνων (τα σχήματα, με τα οποία η υπολογιστική γεωμετρία ασχολείται περισσότερο) από το χρήστη και περιέχει συναρτήσεις υπολογισμού τομών δισδιάστατων γεωμετρικών αντικειμένων και προσανατολισμού τριών σημείων. Όλα αυτά δεν θα μπορούσαν να γίνουν εύκολα χωρίς γραφικό περιβάλλον, καθώς όλα τα γεωμετρικά αντικείμενα θα ήταν σε μη οπτική μορφή.

Με τη βοήθεια της VPython η γραφική απεικόνιση διάφορων γεωμετρικών εννοιών έγινε εύκολη υπόθεση. Ο χρήστης, με τη βοήθεια του ποντικιού, μπορεί να δημιουργεί σημεία και κάθε άλλο γεωμετρικό αντικείμενο που χρειάζεται. Η εισαγωγή σημείων γίνεται είτε με το ποντίκι είτε προγραμματιστικά. Ο εύκολος χειρισμός της βιβλιοθήκης είναι πιθανό να της επιτρέψει να κατακτήσει και χώρους εκτός του Πανεπιστήμιου. Αν και η VPython υποστηρίζει και τις 3 διαστάσεις, έχουμε υλοποιήσει έννοιες μόνο στις 2, αντικαθιστώντας την 3η διάσταση των σημείων με μηδέν (0). Το παρόν κείμενο οφείλει αρκετά στα ερεθίσματα της CGAL [4] και τις αποκτηθείσες γνώσεις σε python από σχετική βιβλιογραφία [5].

2. Επιλογή γλώσσας υλοποίησης

2.1 Python

Η επιλογή της Python ως γλώσσα υλοποίησης έγινε καθώς είναι μια από τις πιο διαδομένες γλώσσες στην υλοποίηση από των πιο απλών μέχρι των πιο πολύπλοκων προγραμμάτων, όπως οι πολυνηματικοί διακομιστές. Η Python δεν εξαρτάται από το λειτουργικό σύστημα και αυτό την κάνει ιδιαίτερα χρήσιμη, καθώς η βιβλιοθήκη γίνεται τελείως μετακινήσιμη (portable) σε όλα τα μηχανήματα. Επίσης, είναι πιο εύχρηστη, λόγω της σύνταξής της.

Η Python χωρίς ερωτηματικά (;), σύμβολα δολαρίου ($) και αγκύλες κάνει τον κώδικα πιο ευανάγνωστο στους κοινούς ανθρώπους ενώ, γενικά, δεν έχει πολλά σύμβολα που να μην γίνονται αντιληπτά στο μέσο άνθρωπο. Επιπρόσθετα, οι λέξεις για τις λογικές πράξεις δεν εκφράζονται με τα συνηθισμένα σύμβολα που χρησιμοποιούν άλλες γλώσσες (λογικό και: &&, λογικό ή: ||,λογική άρνηση: !, κτλ), αλλά με τις αντίστοιχες λέξεις τους στην αγγλική γλώσσα (and, or, not, κλπ). Με αυτόν τον τρόπο, η μετατροπή ενός ψευδοκώδικα αλγορίθμου σε Python είναι σαν τη μετάφραση ενός κειμένου. Τέλος, όντας μια γλώσσα υψηλού επίπεδου (high-level), καθιστά τον προγραμματισμό πιο εύκολο, καθώς δεν υπάρχει η ανάγκη ενασχόλησης με τη δέσμευση και αποδέσμευση μνήμης.

2.2 VPython [1]

Η βιβλιοθήκη VPython είναι ένα εύχρηστο εργαλείο που προσφέρει πολλές δυνατότητες. Είναι ήδη συμβατή με τη δημιουργία αντικειμένων τριών διαστάσεων και εισάγει πλήθος αντικειμένων, όπως σφαίρα, γραμμή, δακτύλιος, πυραμίδα, βέλος, ελατήριο και πολλά άλλα. Χρησιμοποιώντας μερικά από αυτά τα αντικείμενα δημιουργήσαμε όλα τα απαραίτητα αντικείμενα για τη βιβλιοθήκη μας. Η δυνατότητα απεικόνισης των αντικείμενων, αλλαγής χρώματος και εισαγωγής κειμένου στα ήδη απεικονισμένα αντικείμενα, την κάνει ένα εύχρηστο περιβάλλον για να μπορέσει κάποιος να διδάξει σχήματα και αντικείμενα σε φοιτητές, με λίγο κόπο και χωρίς ιδιαίτερες προγραμματιστικές δυσκολίες. Η βιβλιοθήκη αυτή έχει ήδη δοκιμαστεί στη διδασκαλία της Φυσικής (μέσω δημιουργίας πειραματικών διατάξεων με τα αντικείμενά της) αλλά και της Χημείας, καθώς μπορεί κάποιος να δημιουργήσει τρισδιάστατα μοντέλα χημικών στοιχείων και ενώσεων για την καλύτερη κατανόηση των διαφόρων μορφών τους στο χώρο.

3. Προβλήματα της Python

3.1 Υπερφόρτωση συναρτήσεων

Η Python είναι μια από τις πιο διαδεδομένες γλώσσες προγραμματισμού. Ο προγραμματισμός σε αυτήν παρουσιάζει διαφορές από τη C++. Η Python χρησιμοποιείται και ως διαδικαστική, αλλά και ως αντικειμενοστραφής γλώσσα, διαφέρει όμως ως προς τη C++ στην υπερφόρτωση των συναρτήσεων. Η Python προσφέρει υπερφόρτωση για τα μαθηματικά σύμβολα (==, >=,<=,!=,<,>), ώστε να μπορούν να χρησιμοποιηθούν για την ταξινόμηση μιας λίστας ή για τον έλεγχο των στοιχείων της κλάσης. Όμως, οι συναρτήσεις που φτιάχνει ο προγραμματιστής δεν υπερφορτώνονται με ανάλογο της C++ τρόπο.

Αναπτύξαμε τη συνάρτηση isinstance που παίρνει 2 ορίσματα. Το πρώτο όρισμα είναι η μεταβλητή που θέλουμε να κατηγοριοποιήσουμε και το δεύτερο είναι το όνομα του τύπου ή της κλάσης που θέλουμε να είναι η μεταβλητή μας. Αυτό επιτρέπει να δημιουργείται το αντικείμενο μόνο με τη συνάρτηση __init__ κάθε κλάσης, ανάλογα με το τι ορίσματα δέχεται η συνάρτηση.

3.2 Αναπαράσταση σημείων μεγάλης ακριβείας

Ένα μεγάλο ζήτημα για τις υπάρχουσες βιβλιοθήκες δημιουργίας σημείων είναι η ακρίβεια. Αν προσπαθήσουμε να δημιουργήσουμε σημεία μεγάλης ακριβείας, τότε όλες οι βιβλιοθήκες έχουν ζήτημα ορθότητας όταν κάνουν πράξεις με αυτά τα σημεία. Ιδιαίτερα το κατηγόρημα orientation (προσανατολισμός) σημείων αρχίζει να παράγει λανθασμένα αποτελέσματα [6].

Η βιβλιοθήκη Decimal που συνοδεύει τη Python, προσφέρει σημαντικές δυνατότητες όπως:

- Ορισμός της ακριβείας του σημείου
- Υπερφορτωμένες συναρτήσεις για όλες τις πράξεις
- Μεγάλη ακρίβεια στις πράξεις

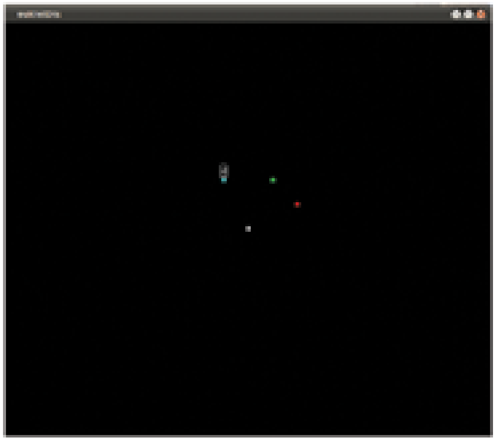
Χρειάστηκε να αντιμετωπιστούν διάφορα θέματα με την χρήση της Decimal, καθώς n βιβλιοθήκη ακολουθεί τη λογική «παίρνεις ό,τι πληκτρολογείς». Κατά την εκτέλεση αριθμητικών πράξεων διαφορετικής ακρίβειας αντιμετωπίσαμε το ακόλουθο πρόβλημα: αν ο χρήστης έβαζε σημεία με ακρίβεια μικρότερη της προεπιλεγμένης, το αποτέλεσμα ήταν ένας αριθμός με την προεπιλεγμένη ακρίβεια. Αυτό πολλές φορές οδηγούσε στη εμφάνιση ενός σημείου με δύο διαφορετικές ακρίβειες, με αποτέλεσμα όταν γινόταν σύγκριση ισότητας να αποτυγχάνει. Το ζήτημα αυτό λύθηκε με την εισαγωγή των απαραίτητων μηδενικών (0), ώστε να έχουν όλα τα σημεία την ίδια ακρίβεια

4. Βασικές έννοιες της βιβλιοθήκης

Με τον όρο βασικές έννοιες της βιβλιοθήκης, εννοούμε τις κλάσεις που δημιουργήθηκαν για τα επιμέρους αντικείμενα που προσφέρει η βιβλιοθήκη. Όλα τα αντικείμενα αφορούν τον δισδιάστατο χώρο (παραδείγματα κώδικα συνοδευόμενα από αντίστοιχα στιγμιότυπα οθόνης: *http://test-georgios-koloventzos.web.cern.ch/test-georgios-koloventzos/examples.php*).

4.1 Σημείο (Point_2)

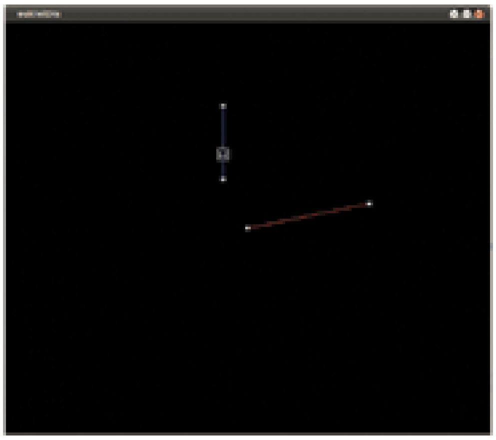
Το σημείο είναι η βασικότερη κλάση της βιβλιοθήκης. Αρχικοποιείται με τις συντεταγμένες του σημείου σε 2 μεταβλητές decimal, μια λίστα με τις συντεταγμένες ως στοιχεία (πρώτα η τετμημένη και μετά η τεταγμένη), το χρώμα ως τύπος tuple και το αντικείμενο σφαίρα (sphere) της VPython. Επιπρόσθετα, αν εισάγουμε κείμενο για το σημείο, τότε δημιουργείται ένα αντικείμενο label της VPython. Το σημείο περιέχει, επίσης, υπερφορτωμένες συναρτήσεις για όλα τα μαθηματικά σύμβολα, ούτως ώστε να μπορούμε να κάνουμε ταξινομήσεις σε λίστες που περιέχουν σημεία. Οι συγκρίσεις γίνονται πρώτα ως προς την τετμημένη, και αν είναι ίσες, τότε συγκρίνεται και η τεταγμένη.



Εικόνα 1: Σημεία στο επίπεδο

4.2 Ευθύγραμμο τμήμα (Segment_2)

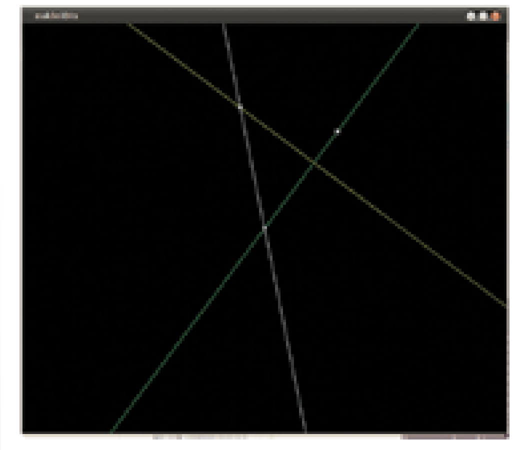
Το ευθύγραμμο τμήμα μπορούμε να πούμε ότι είναι η κλάση με τις περισσότερες χρήσεις. Αυτό συμβαίνει, διότι είναι η κλάση στην οποία απευθυνόμαστε όταν έχουμε να κάνουμε σύγκριση μεταξύ ανόμοιων πραγμάτων (ακόμα και όταν ζωγραφίζουμε τη γραμμή (σ.σ. βλέπε 4.3), τη δημιουργούμε με ένα ευθύγραμμο τμήμα που τα άκρα του είναι πολύ μακριά). Το ευθύγραμμο τμήμα περιέχει 2 σημεία, ως αρχή και τέλος, και μια καμπύλη (curve) της VPython για να δημιουργήσουμε την εικόνα του ευθύγραμμου τμήματος. Η κλάση του ευθύγραμμου τμήματος περιέχει πολλές συναρτήσεις, ώστε να δώσει διάφορες πληροφορίες, όπως το μήκος του στο τετράγωνο, το μέσον του τμήματος, τη γραμμή στην οποία βρίσκεται, αν ένα σημείο του είναι πάνω σε αυτό και αν είναι κάθετο ή οριζόντιο.



Εικόνα 2: Ευθύγραμμα τμήματα

4.3 Γραμμή (Line_2)

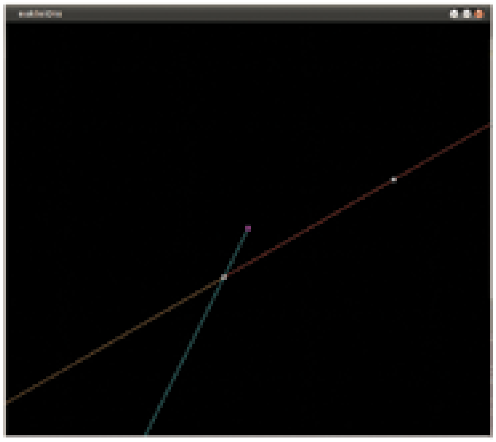
Η κλάση της γραμμής μας επιτρέπει να δημιουργήσουμε μια γραμμή και να κάνουμε πράξεις για τον προσανατολισμό σημείων με αυτή. Η κλάση περιέχει τους 3 συντελεστές της ευθείας (από την εξίσωση Ax + By +C =0) και τη γραμμή ως καμπύλη, ακριβώς όπως και το ευθύγραμμο τμήμα. Η ευθεία δημιουργείται από ένα ευθύγραμμο τμήμα με πολύ μεγάλα άκρα (παίρνουμε σημεία εκτός του αρχικού οπτικού πεδίου του προγραμματιστή). Έτσι, δίνεται η ψευδαίσθηση της ευθείας.



Εικόνα 3: Γραμμές στο επίπεδο

4.4 Ακτίνα (Ray_2)

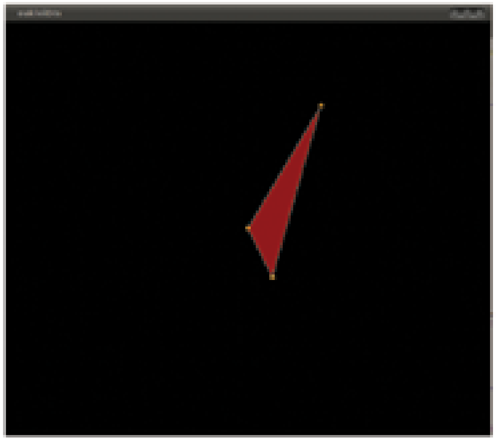
Η ακτίνα είναι ουσιαστικά μια γραμμή με σημείο αρχής. Έχει σχεδόν τις ίδιες ιδιότητες με τη γραμμή, αν και η ακτίνα χρειάζεται οπωσδήποτε ένα σημείο για την αρχή της. Η ακτίνα κρατά στην κλάση της το σημείο αρχής, τη διεύθυνση της, το σημείο τέλους της (για λογούς ευχέρειας και απεικόνισης) και την καμπύλη της, όπως όλες οι κλάσεις με γραμμές.



Εικόνα 4: Ακτίνες στο επίπεδο

4.5 Τρίγωνο (Triangle_2)

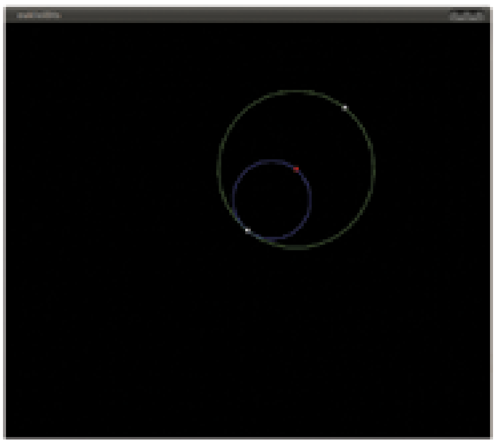
Η κλάση τρίγωνο χρησιμοποιεί 3 στιγμιότυπα της κλάσης ευθύγραμμο τμήμα, ώστε να δημιουργήσει την οπτικοποίηση του τρίγωνου. Έτσι, κάθε ευθύγραμμο τμήμα και κάθε σημείο έχει τις ιδιότητες της κλάσης του, και όλα μαζί δημιουργούν το τρίγωνο. Εδώ χρησιμοποιείται και η κλάση της VPython με την ονομασία face (έδρα) και με αυτόν τον τρόπο, μπορούμε να χρωματίσουμε το εσωτερικό του τριγώνου. Έχουν δημιουργηθεί συναρτήσεις για το χρωματισμό μόνο των σημείων ή μόνο των ευθύγραμμων τμημάτων. Επίσης, οι συναρτήσεις που χρησιμοποιούνται για να αποφανθούμε αν 2 τρίγωνα είναι ίσα, έχουν υλοποιηθεί ούτως ώστε, αν δοθούν τα σημεία με κυκλικό συνδυασμό, η κλάση να μην έχει πρόβλημα να συμπεράνει ότι πρόκειται για ιδία τρίγωνα.



Εικόνα 5: Τρίγωνο στο επίπεδο

4.6 Κύκλος (Circle_2)

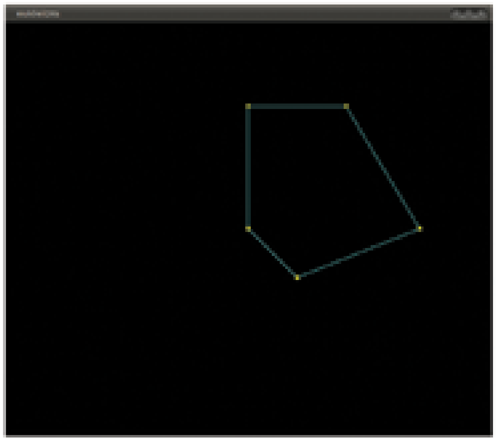
Αν και γενικά η Υπολογιστική Γεωμετρία δεν ασχολείται ιδιαίτερα με κύκλους, η υλοποίηση μιας κλάσης για αυτό το γεωμετρικό σχήμα ίσως να βοηθήσει σε κάποιον αλγόριθμο αργότερα, όπως για παράδειγμα στη δημιουργία του διαγράμματος Voronoi. Ο κύκλος έχει σαν κύρια χαρακτηριστικά το κέντρο, ως ένα σημείο (Point_2), και το μήκος της ακτίνας στο τετράγωνο (συνηθίζεται τέτοιες τιμές να είναι στο τετράγωνο, καθώς η εύρεση του μήκους είναι υπολογιστικά ακριβή και στους περισσότερους τύπους (όγκου, κλπ) χρειάζεται να είναι στο τετράγωνο). Η κλάση της VPython για την οπτικοποίηση του κύκλου είναι ο δακτύλιος (ring).



Εικόνα 6: Κύκλοι στο επίπεδο

4.7 Πολύγωνο (Polygon_2)

Η Υπολογιστική Γεωμετρία ασχολείται κατά κόρον με τα πολύγωνα. Τα κυρία χαρακτηριστικά της κλάσης, εκτός από τις 2 λίστες, είναι και το αν είναι απλό ή κυρτό πολύγωνο, καθώς και ο προσανατολισμός των σημείων του πολυγώνου. Αυτά αρχικά δεν υπολογίζονται. Αν ο χρήστης τα ζητήσει, τότε υπολογίζονται και αποθηκεύονται για τις επόμενες χρήσεις. Η κλάση αυτή έχει πολλές συναρτήσεις για να αλλάζουμε το περιεχόμενο των 2 λιστών, ούτως ώστε να μπορούμε να αλλάζουμε το πολύγωνο για να ικανοποιεί τις ανάγκες του προγραμματιστή. Έχει προβλεφθεί κάθε επανάληψη που γίνεται στα σημεία ή στα ευθύγραμμα τμήματα να έχει τη συνάρτηση xrange αντί για range της Python, ώστε να μην υπάρξουν προβλήματα με τη μνήμη σε πολύ μεγάλα πολύγωνα ή σε μη ισχυρούς υπολογιστές. Η συνάρτηση του εμβαδού υλοποιεί τον κλασικό αλγόριθμο για την εύρεση εμβαδού πολυγώνου (χωρίς να έχει προβλήματα αν είναι κυρτό ή μη) με δεδομένο μόνο τα σημεία του.



Εικόνα 7: Πολύγωνο στο επίπεδο

5. Βασικές και μη, συναρτήσεις της βιβλιοθήκης

5.1 Συναρτήσεις για τη χειραγώγηση σημείων

Υπάρχουν 2 συναρτήσεις χειραγώγησης σημείων. Όλες οι συναρτήσεις χρειάζονται καλύτερη υλοποίηση, καθώς δεν εκμεταλλεύονται όλες τις παραμέτρους που μπορούν να δοθούν. Ήδη έχουν υλοποιηθεί πολύ καλές επιλογές στην ομάδα από άλλους συναδέλφους, που θα εισαχθούν στη βιβλιοθήκη στο προσεχές μέλλον.

5.1.1 Η συνάρτηση getVisualPoints

Η συνάρτηση αυτή δεν παίρνει ορίσματα και ουσιαστικά ετοιμάζει τη βιβλιοθήκη για να δεχθεί την είσοδο από το ποντίκι, ούτως ώστε να απεικονίσει τα σημεία. Ο τρόπος για να σταματήσει ο χρήστης την είσοδο είναι να πατήσει το πλήκτρο backspace. Επιλέχτηκε αυτός ο τρόπος γιατί όλα τα κουμπιά του ποντικιού έχουν μία χρήση στις επόμενες εκδόσεις της VPython. Η συνάρτηση κρατάει τα σημεία σε μία λίστα καθολικής εμβελείας για να μπορούμε να την προσπελαύνουμε από οποιαδήποτε σημείο της βιβλιοθήκης.

5.1.2 Η συνάρτηση getPolygon

Όπως και η «αδελφική» της συνάρτηση, getVisualPoints, δεν παίρνει ορίσματα. Η είσοδος των σημείων σταματάει, και αυτή, με το πλήκτρο backspace. Η συνάρτηση ενώνει τα σημεία, καθώς ο χρήστης τα εισάγει (με τη σειρά που τα δημιουργεί ο χρήστης), με Segment_2. Στο τέλος της εισόδου η συνάρτηση επιστρέφει τις 2 λίστες (μία με τα σημεία και μία με τα ευθύγραμμα τμήματα) για να μπορέσει ο χρήστης να χρησιμοποιήσει όπως θέλει τις 2 λίστες ξεχωριστά. Αυτές μπορούν να εισαχθούν κατευθείαν σαν ορίσματα της κλάσης Polygon_2 για την αρχικοποίηση της.

5.2 Συναρτήσεις χρησίμων λειτουργιών για την υπολογιστική γεωμετρία

5.2.1 Συνάρτηση προσανατολισμού (orientation) [2]

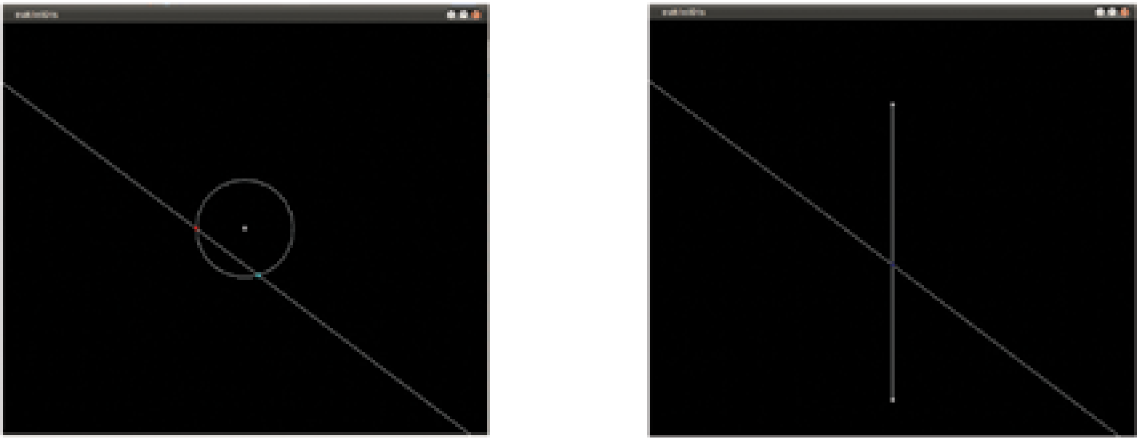
Η συνάρτηση (κατηγόρημα) προσανατολισμού είναι μια συνάρτηση που, μέσα στις συναρτήσεις των κλάσεων, χρησιμοποιείται συνέχεια. Μας δίνει τη πληροφορία αν τα σημεία είναι προσανατολισμένα με τη φορά του ρολογιού, την αντίστροφη φορά, ή είναι συνευθειακά. Αυτό, γιατί μέσα στη βιβλιοθήκη είναι προγραμματιστικά καλύτερο να γίνονται συγκρίσεις με αριθμούς, παρά συγκρίσεις με αλφαριθμητικά.

Από μαθηματικής άποψης, η συνάρτηση βρίσκει τον προσανατολισμό με την αφαίρεση των συντεταγμένων των σημείων μεταξύ τους. Με αυτόν τον τρόπο, μια από τις κυριότερες συναρτήσεις και ερωτήσεις της υπολογιστικής γεωμετρίας δεν χρειάζεται τεράστια υπολογιστική ισχύ. Η συνάρτηση επιστρέφει αλφαριθμητικό για να μπορεί να εκτυπωθεί κατευθείαν, αν χρειαστεί. Στις κλάσεις αποθηκεύεται ως αριθμός για να υπάρχει ευκολία για λιγότερη χρονική και χωρική πολυπλοκότητα σε πράξεις μεταξύ προσανατολισμών μεταξύ κλάσεων.

5.2.2 Συνάρτηση τομής (intersection)

Η συνάρτηση τομής είναι η μεγαλύτερη συνάρτηση της βιβλιοθήκης από την πλευρά κώδικα, διότι πρέπει να καλύψει όλους τους συνδυασμούς ανάμεσα στις κλάσεις. Στην υλοποίηση χρησιμοποιήσαμε το τέχνασμα για την υπερφόρτωση συναρτήσεων (βλ. 3.1) στις κλάσεις, ώστε να μπορέσουμε να φτιάξουμε αυτήν τη συνάρτηση με περισσή αντικειμενοστρέφεια. Έτσι, η συνάρτηση παίρνει ως δεδομένα τα 2 στιγμιότυπα κλάσεων και το αν θέλουμε να εμφανίσουμε την τομή τους. Όταν έχουμε τομές ανάμεσα σε γραμμές, ευθύγραμμα τμήματα και ακτίνες, τότε όλα μετατρέπονται σε γραμμές και γίνεται η τομή. Αν επιστρέψει κάτι η τομή των γραμμών, τότε γίνονται συγκρίσεις για να δούμε αν το σημείο είναι μέσα στην ακτίνα ή στο ευθύγραμμο τμήμα. Έτσι, αντί να γραφούμε τον κώδικα συνέχεια, έχουμε μια πιο αντικειμενοστραφή προσέγγιση.

Το ίδιο γίνεται και με τις τομές αυτών των κλάσεων με τον κύκλο και το πολύγωνο. Δίνουμε βάρος στην τομή της γραμμής με όλα τα αλλά αντικείμενα για καλύτερη υλοποίηση. Η συνάρτηση δεν επιστρέφει το ίδιο αντικείμενο κάθε φορά. Αυτό έγινε διότι ανάμεσα σε 2 γραμμές μπορεί να έχουμε είτε ένα σημείο είτε μια γραμμή. Όμως, όταν έχουμε τομή μιας γραμμής με έναν κύκλο μπορεί να έχουμε 2 σημεία, ένα σημείο ή και τίποτα. Ακόμα χειρότερα είναι, όταν έχουμε να κάνουμε με τομή πολυγώνου με γραμμή, ευθύγραμμο τμήμα ή ακτίνα. Εκεί μπορεί να είναι ευθύγραμμα τμήματα, σημεία ή και τίποτα. Έτσι, θα πρέπει στην επομένη έκδοση, η συνάρτηση να επιστρέφει λίστα με μηδενικό αριθμό, αν δεν υπάρχουν σημεία τομής ή με όλα τα σημεία, ευθύγραμμα τμήματα, κλπ σε μια λίστα. Αυτή τη στιγμή, αυτό το κάνει μόνο για το πολύγωνο.



Εικόνα 8: Τομή Κύκλου με Γραμμή, Τομή Γραμμής και Ευθύγραμμου τμήματος

6. Συμπεράσματα

Η βιβλιοθήκη αυτή είναι η αρχή μιας πολύ μεγάλης δουλειάς. Η αρχή της, να είναι εύκολη για τους μη εξειδικευμένους χρήστες, είναι μια μεγάλη επιτυχία της. Μπορεί να σχεδιάσει πολύγωνα σε πολύ λίγο χρόνο με μία εντολή και χωρίς ιδιαίτερες δυσκολίες για τον χρήστη. Αν και η μετακίνηση σημείων δεν είναι στο καλύτερό της στάδιο, θα δώσει μια μεγάλη βοήθεια στη διδακτική της Υπολογιστικής Γεωμετρίας και γενικότερα της γεωμετρίας και στην ενασχόληση των λιγότερο μυημένων με τον προγραμματισμό. Μια απαραίτητη αναβάθμιση της βιβλιοθήκης είναι η συμβατότητα με την Python 3 (θα πρέπει να αλλάξουν κάποια πράγματα ιδιαίτερα στην εκτύπωση των κλάσεων). Η βιβλιοθήκη μπορεί να επεκταθεί στο χώρο των 3 διαστάσεων, αλλά και να δημιουργηθούν οι κλάσεις και οι συναρτήσεις για τις 3 διαστάσεις. Η βιβλιοθήκη δύναται να εξελιχθεί ως εργαλείο για κάθε δάσκαλο που θέλει να διδάξει γεωμετρία, υπολογιστική γεωμετρία και οποιοδήποτε μάθημα έχει να κάνει γενικά με τη δημιουργία αντικειμένων στον χώρο. Η Python είναι εύκολη σαν γλώσσα και όλο και περισσότερα πανεπιστήμια την εντάσσουν στο πρόγραμμα σπουδών τους, ιδιαίτερα στο μάθημα της «Εισαγωγής στον προγραμματισμό». Ακόμα και παιδιά Λυκείου είναι εύκολο να φτιάξουν έναν αλγόριθμο σε μια ψευδογλώσσα και με πολύ λίγη προσπάθεια να τον μετατρέψουν σε κώδικα Python για τη βιβλιοθήκη. Η βιβλιοθήκη έχει πολύ μεγάλες προοπτικές για να γίνει ένα βοήθημα για το μαθητή και τον καθηγητή στην εκμάθηση και την κατανόηση αλγορίθμων, σχημάτων και γενικώς του χώρου που μας περιβάλλει.

Αναφορές

1. Scherer, D.; Dubois, P.; Sherwood, B.; , «VPython: 3D interactive scientific graphics for students,» Computing in Science & Engineering , vol.2, no.5, pp.56-62, Sep/Oct 2000

2. I. Z. Εμίρης, Υπολογιστική Γεωμετρία: Μια σύγχρονη αλγοριθμική προσέγγιση, Κλειδάριθμος, 2008

3. J. O'Rource, Computational geometry in C 2nd edition, Campridge University Press, 2008

4. CGAL Reference and Users Manual, CGAL Consortium, 2001

5. W.J. Chun, Core Python Programming 2nd edition, Prentice Hall, 2007

6. Classroom Examples of Robustness Problems in Geometric Computations. Lutz Kettner, Kurt Mehlhorn, Sylvain Pion, Stefan Schirra, and Chee Yap, Computational Geometry: Theory and Applications, vol.40, no.1, pp.61-78, May 2008

Μέθοδοι
Διανομής Χρόνου και Συχνότητας

Βασιλική Κ. Πάρρυ
std06157@di.uoa.gr

Εβίσα Χ. Τσολάκου
std06148@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Σε αυτή την εργασία παρουσιάζονται οι πηγές και οι μέθοδοι διάδοσης προτύπων χρόνου και συχνοτήτων στη μικροκυματική ζώνη συχνοτήτων και κυρίως στην οπτική. Αρχικά περιγράφεται ο τρόπος λειτουργίας και τα είδη των ατομικών ρολογιών, τα οποία αποτελούν τη σημαντικότερη πηγή προτύπων. Το κυρίως σκέλος της εργασίας αφορά τη διανομή αυτών των προτύπων με διάφορες μεθόδους που διατηρούν την εξαίρετη σταθερότητα και ακρίβειά τους, οι δύο βασικές κατηγορίες των οποίων είναι η μετάδοση μέσω οπτικών ινών και η μετάδοση μέσω δορυφόρου. Έμφαση δίνεται στην πρώτη κατηγορία, όπου και αναλύονται οι τρεις τεχνικές μετάδοσης μέσω ίνας που χρησιμοποιούνται σήμερα. Τέλος, αναφέρονται οι εφαρμογές των προτύπων και των μεθόδων διανομής τους, καθώς και οι οργανισμοί που εμπλέκονται στη σχετική έρευνα.

Λέξεις-Κλειδιά: Αναφορά συχνότητας, Οπτικές ίνες, Ατομικά ρολόγια, Οπτικό φέρον, Laser, Σταθερότητα, Αβεβαιότητα, Σύγκριση συχνοτήτων, Συγχρονισμός, Frequency comb, GPS, Telecommunication fiber

Επιβλέπων: Δημήτριος Συβρίδης, Καθηγητής

1. Εισαγωγή

Στη σημερινή εποχή, τα ατομικά ρολόγια παρέχουν πρότυπα συχνοτήτων ή χρόνου που χαρακτηρίζονται από πολύ μεγάλη σταθερότητα και ακρίβεια. Για να χρησιμοποιηθούν σε μία πληθώρα από εφαρμογές, είναι απαραίτητη η ανάπτυξη τεχνικών διανομής και μετάδοσής τους, οι οποίες διατηρούν τη σταθερότητά τους και εισάγουν όσο το δυνατόν λιγότερο θόρυβο. Επί του παρόντος, πραγματοποιούνται συνεχώς πειράματα και έρευνες για νέες μεθόδους και τεχνικές μεταφοράς, μία από τις οποίες είναι η μετάδοση μέσω οπτικών ινών. Η ανάγκη για μεταφορά και μετάδοση γίνεται πιο κατανοητή και από το γεγονός ότι τα ατομικά ρολόγια και γενικά οι πηγές τόσο σταθερών προτύπων είναι πολύ ακριβοί μηχανισμοί και μεταφέρονται δύσκολα. Έτσι, μόνο ορισμένα εργαστήρια διαθέτουν τέτοιες πηγές, ενώ οι υπόλοιπες εγκαταστάσεις πρέπει να προμηθευτούν τα πρότυπα από αυτά.

Τα πρότυπα συχνοτήτων είναι χρήσιμα ή και απαραίτητα για ένα μεγάλο αριθμό εφαρμογών. Τέτοιες είναι το GPS, ο παγκόσμιος χρόνος UTC, πειράματα που εξετάζουν θεμελιώδεις θεωρίες και σταθερές της φυσικής, ο συγχρονισμός τηλεσκοπίων ή άλλων μηχανισμών και οι επιταχυντές σωματιδίων. Τα πρότυπα χρησιμοποιούνται γενικά από κλάδους της επιστήμης όπως η αστρονομία, η μετρολογία κ.α.

Στην παρούσα εργασία γίνεται μία ανασκόπηση των προτύπων συχνοτήτων/ χρόνου και των σύγχρονων μεθόδων μεταφοράς τους. Στο πρώτο κεφάλαιο, αφού δοθεί ένας σύντομος ορισμός για τα πρότυπα/αναφορές συχνότητας, παρουσιάζονται οι πηγές τους, δηλαδή τα διάφορα ήδη των ατομικών ρολογιών (μικροκυματικά, οπτικά, κβαντικά). Στο δεύτερο κεφάλαιο παρουσιάζεται το frequency comb, ο τρόπος λειτουργίας του και η χρησιμότητά του, αφού η κατανόηση της έννοιάς του απαιτείται από τα επόμενα κεφάλαια. Το τρίτο κεφάλαιο είναι το κυρίως κεφάλαιο της εργασίας, αφού αναλύει τις μεθόδους διανομής των προτύπων και χωρίζεται σε τρία μέρη: παράμετροι θορύβου, διανομή μέσω οπτικών ινών και διανομή μέσω δορυφόρων.

Το πρώτο μέρος αφορά τις γενικές παραμέτρους με τις οποίες μετράται ο θόρυβος μίας ζεύξης μεταφοράς και σε δεύτερη φάση πιο συγκεκριμένα το είδος του θορύβου που εισάγεται από μία οπτική ίνα και τη γενική αρχή ακύρωσής του.

Το δεύτερο μέρος περιγράφει τις παραμέτρους που επηρεάζουν τη μετάδοση μέσω οπτικής ίνας και αναλύει τις τρεις βασικές μεθόδους που χρησιμοποιούνται για μία τέτοια μετάδοση: μεταφορά μικροκυματικής συχνότητας με διαμόρφωση πλάτους ενός laser, απευθείας μεταφορά οπτικού φέροντος και μεταφορά ενός frequency comb. Τέλος, δίνονται περισσότερες λεπτομέρειες για κάποιες από τις εφαρμογές της μετάδοσης μέσω ίνας και παρουσιάζονται σύντομα οι κύριοι οργανισμοί που εμπλέκονται στην ανάπτυξη των παραπάνω μεθόδων και τη διεξαγωγή αντίστοιχων πειραμάτων.

Το τρίτο μέρος παρουσιάζει τις δύο κύριες μεθόδους μετάδοσης μέσω δορυφόρων, το GPS και το TWSTFT, τον τρόπο λειτουργίας τους και τα πλεονεκτήματα της κάθε μίας.

Τέλος, εξάγονται κάποια συμπεράσματα ως προς τις διάφορες μεθόδους μετάδοσης (οπτική ίνα ή δορυφορική), τα πλεονεκτήματα και μειονεκτήματά τους, και τις περιπτώσεις χρησιμότητάς τους.

2. Ατομικά ρολόγια

Ένα ατομικό ρολόι είναι ένα ρολόι που χρησιμοποιεί τη συχνότητα μιας ηλεκτρονικής μετάβασης των ηλεκτρονίων ενός ατόμου στη μικροκυματική, οπτική ή υπεριώδη περιοχή του ηλεκτρομαγνητικού φάσματος, ως πρότυπο συχνότητας για τη θεμελιώδη μονάδα με την οποία μετράει το χρόνο. Η αρχή λειτουργίας των ατομικών ρολογιών βασίζεται στην κβαντισμένη φύση των ατόμων. Ένα ατομικό ρολόι εκμεταλλεύεται την παραπάνω ιδιότητα των ατόμων για να δημιουργήσει μία αναφορά συχνότητας, η οποία εξαρτάται από τα άτομα που χρησιμοποιεί στο μηχανισμό του.

2.1 Ορισμός του δευτερολέπτου

Πριν τη χρήση ατομικών ρολογιών ο χρόνος είχε οριστεί με βάση την περιστροφή της γης γύρω από τον εαυτό της και την περιφορά της γύρω από τον ήλιο. Με την ανακάλυψη των ατομικών ρολογιών το δευτερόλεπτο μπορεί να οριστεί με πολύ μεγαλύτερη ακρίβεια, σε σχέση με τη φυσική συχνότητα διέγερσης ενός ατόμου, αφού αυτή η συχνότητα είναι πάρα πολύ σταθερή και δεν επηρεάζεται από τη θερμοκρασία, την ατμοσφαιρική πίεση κτλ. Το δευτερόλεπτο μπορεί να οριστεί ως η διάρκεια 9.192.631.770 επαναλήψεων ή περιόδων.

2.2 Ακρίβεια ατομικών ρολογιών

Αυτή η σχέση συχνότητας και χρόνου χρησιμοποιείται για να οριστεί η πάροδος του χρόνου με μεγάλη ακρίβεια. Τα ρολόγια, εν γένει, μετράνε την πάροδο του χρόνου με χρήση ενός ταλαντωτή που λειτουργεί σε μία συγκεκριμένη συχνότητα. Όσο μεγαλύτερη είναι η συχνότητα του ταλαντωτή, τόσο μεγαλύτερη είναι η ακρίβεια του ρολογιού. Τα μηχανικά ρολόγια που χρησιμοποιούσαν ένα εκκρεμές λειτουργούσαν σε συχνότητα μερικών Hz. Οι ταλαντωτές των ρολογιών που χρησιμοποιούν κρυστάλλους quartz λειτουργούν σε συχνότητες κάποιων MHz. Τα ατομικά ρολόγια χρησιμοποιούν ως «ταλαντωτή» τα ίδια τα άτομα ενός στοιχείου και λειτουργούν σε συχνότητες GHz, THz ή ακόμα και PHz. Είναι προφανές πόσο μεγαλύτερη ακρίβεια προσφέρουν σε σχέση με όλους τους παλαιότερους τύπους ρολογιών. Οι κυριότερες εφαρμογές των ατομικών ρολογιών συμπεριλαμβάνουν το Παγκόσμιο Σύστημα Θεσιθεσίας (Global Positioning System, GPS) και τη διεθνή αρχή της χρονομέτρησης, που είναι γνωστή ως Συγχρονισμένος Παγκόσμιος Χρόνος ή UTC. Τα ατομικά ρολόγια χρησιμοποιούνται επίσης σε πειράματα μεγάλης ακρίβειας, που εξετάζουν θεμελιώδεις θεωρίες, όπως η κβαντική ηλεκτροδυναμική και η Γενική Σχετικότητα, καθώς και σε επιστήμες όπως η ραδιοαστρονομία.

3. Frequency combs

3.1 Ορισμός και λειτουργία

Μέχρι πριν από λίγα χρόνια, ο μόνος τρόπος να συσχετιστούν οι οπτικές με τις μικροκυματικές συχνότητες ήταν μέσω μιας συσκευής που ονομάζεται αλυσίδα συχνοτήτων (frequency chain). Πρόσφατα όμως, έχει εμφανιστεί μια πολύ απλούστερη μέθοδος μέτρησης οπτικών συχνοτήτων ως αποτέλεσμα των εξελίξεων στην τεχνολογία laser femtosecond και μη γραμμικών οπτικών ινών. Είναι πλέον δυνατό να μετρηθεί με μία μοναδική συσκευή και με πρωτοφανή ακρίβεια η συχνότητα σχεδόν οποιασδήποτε σταθερής οπτικής πηγής. Η συσκευή με την οποία γίνεται αυτό είναι γνωστή ως optical frequency comb (OFC).

3.2 Εφαρμογές

Τα frequency combs έχουν απλοποιήσει πολύ σημαντικά τη μέτρηση οπτικών συχνοτήτων, δίνοντας την δυνατότητα σε δεκάδες εφαρμογές που χρησιμοποιούν οπτικές συχνότητες, όπως τα οπτικά ατομικά ρολόγια, να αναπτυχθούν. Συνδέοντας τις οπτικές με τις μικροκυματικές συχνότητες, τα frequency combs μπορούν να χρησιμοποιηθούν για να διαιρέσουν τις συχνότητες των οπτικών ρολογιών σε χαμηλότερες συχνότητες που μπορούν να μετρηθούν με χρήση μικροκυματικών προτύπων.

4. Διανομή μέσω οπτικών ινών

4.1 Μεταφορά μικροκυματικού φέροντος

Προς το παρόν, η διανομή ρολογιού για μακρινές αποστάσεις γίνεται μέσω των δορυφορικών μεθόδων GPS και TWSTFT, αλλά η αστάθεια των μεθόδων αυτών δεν επαρκεί για την μετάδοση των σύγχρονων οπτικών προτύπων. Έτσι, τα τελευταία χρόνια διερευνάται η μετάδοση προτύπων χρόνου και συχνότητας μέσω οπτικών ινών. Μία οπτική ίνα, είτε είναι μονότροπη είτε πολύτροπη, αποτελείται από τον κεντρικό πυρήνα (core), που περιβάλλεται από το μανδύα (cladding), που με τη σειρά του περιβάλλεται από ένα προστατευτικό εξωτερικό περίβλημα. Το φως καθοδηγείται μέσω του μανδύα κατά μήκος του πυρήνα με τη μέθοδο της ολικής εσωτερικής ανάκλασης. Κατά τη διάδοση μέσω οπτικών ινών, εμφανίζονται πολλά φαινόμενα που πρέπει να ληφθούν υπόψη και να αναλυθούν εξαιτίας της επίδρασης που έχουν στο σύστημα. Τέτοια φαινόμενα είναι αυτά της εξασθένησης, τα φαινόμενα της διασποράς, όπως η χρωματική διασπορά, η διασπορά τρόπων πόλωσης και διάδοσης και μη γραμμικά φαινόμενα, όπως τα φαινόμενα σκέδασης Raman και Brillouin.

Η πιο άμεση μέθοδος για διάδοση μιας μικροκυματικής συχνότητας μέσω οπτικών ινών είναι η διαμόρφωση πλάτους ενός laser συνεχούς κύματος στη συχνότητα αναφοράς και η μεταφορά αυτού του διαμορφωμένου φωτός μέσω της ίνας. Στη συνέχεια, η συχνότητα αναφοράς ανακτάται στο δέκτη με την ανίχνευση της συχνότητας διαμόρφωσης, πράγμα που επιτυγχάνεται με χρήση ενός φωτοανιχνευτή ο οποίος ανιχνεύει την ένταση του λαμβανόμενου φωτός και αποδιαμορφώνει τις πλευρικές ζώνες συχνοτήτων γύρω από τη φέρουσα συχνότητα. Η διαμόρφωση πλάτους εφαρμόζεται στο φως που παράγεται από ένα μονότροπο distributed feedback laser ημιαγωγού (που συνήθως εκπέμπει στα 1.3 ή 1.55 μm), με χρήση ένα διαμορφωτή έντασης Mach-Zehnder. Για να μετρηθεί ο θόρυβος του μεταφερόμενου μικροκυματικού σήματος, αρχικά δημιουργείται ένα τοπικό αντίγραφο του προτύπου σήματος μετατοπισμένο στη συχνότητα κατά κάποια KHz. Στη συνέχεια, αναμειγνύεται το σήμα ολοκληρωμένης διαδρομής (που έχει επιστραφεί από τον δέκτη πίσω στον πομπό) με αυτό το τοπικό αντίγραφο που δημιουργήθηκε και μετρούνται οι διακυμάνσεις στη συχνότητα του προκύπτοντος διακροτήματος. Ο θόρυβος είναι της τάξης του $2x10^{-13}$ για 1 δευτερόλεπτο, ενώ φτάνει το $2x10^{-14}$ ανάμεσα στα 100 και 1000 δευτερόλεπτα. Ένα άλλο ζήτημα που τίθεται λόγω της χρήσης συνηθισμένων ινών είναι αυτό της διασποράς τρόπων πόλωσης. Το φαινόμενο διορθώνεται με μίξη της πόλωσης (polarization scrambling).

4.2 Μεταφορά οπτικού φέροντος

Όπως προειπώθηκε, τα οπτικά ατομικά ρολόγια, με την ανώτερη σταθερότητα και ακρίβεια που παρουσιάζουν, απαιτούν δίκτυα μεταφοράς συχνοτήτων τεράστιας σταθερότητας για τη διανομή σήματος και το συγχρονισμό, μεγαλύτερης από αυτή που έχουν επιτύχει οι τεχνικές μεταφοράς μικροκυματικής συχνότητας. Αυτό μπορεί να επιτευχθεί με την απευθείας μεταφορά της οπτικής συχνότητας ενός laser συνεχούς κύματος, κάτι το οποίο έχει το πλεονέκτημα της υψηλής ανάλυσης για τη μέτρηση των διακυμάνσεων θορύβου φάσης στη ζεύξη. Επίσης, λόγω της αυξημένης σταθερότητας, η μετάδοση μπορεί να γίνει για μεγαλύτερες αποστάσεις. Η οπτική συχνότητα μπορεί να χρησιμοποιηθεί στο δέκτη για απευθείας σύγκριση με άλλες οπτικές συχνότητες ή για τη σταθεροποίηση ενός frequency comb μέσω του οποίου μπορεί να συσχετιστεί με μικροκυματικές συχνότητες. Η μεταφορά οπτικών συχνοτήτων μπορεί να γίνει απευθείας μέσω δικτύων οπτικών ινών, με την προϋπόθεση ότι ο θόρυβος φάσης, που προκύπτει κατά τη μεταφορά, μπορεί να ακυρωθεί αποτελεσματικά.

Η επιστροφή του σήματος από το δέκτη πίσω στον πομπό για την ενεργή ακύρωση του θορύβου μπορεί να γίνει με δύο τρόπους: είτε με απλή ανάκλαση του σήματος, είτε με τη λειτουργία ενός laser στο δέκτη, το οποίο ανιχνεύει το λαμβανόμενο σήμα, το ανακατασκευάζει και στέλνει το νέο σήμα πίσω στην ίνα. Μετά την ακύρωση του θορύβου, επιτυγχάνεται μία αστάθεια της τάξης του $3x10^{-15}$ για το πείραμα διαδρομής 6.9 km στο BRAN. Σε ένα πιο πρόσφατο (2008) πείραμα μετάδοσης οπτικού φέροντος μέσω σκοτεινής ίνας μήκους 86 km, επιτεύχθηκε αστάθεια $1.5x10^{-16}$ σε ένα δευτερόλεπτο. Παρά τους περιορισμούς, υπάρχει σημαντική βελτίωση σε σχέση με τα καλύτερα αποτελέσματα της σταθεροποιημένης μεταφοράς μικροκυματικού φέροντος. Αυτό οφείλεται στην επιλογή μιας πολύ μεγαλύτερης συχνότητας διαμόρφωσης, η οποία ουσιαστικά είναι η συχνότητα του οπτικού φορέα. Αν αυξηθεί το διαθέσιμο SNR, είτε με συγκόλληση σύντηξης των τμημάτων της ζεύξης, είτε με μικρή μετατόπιση συχνότητας του οπτικού φορέα στο δέκτη πριν την επιστροφή του, έτσι ώστε να αποφθεχθούν ανακλάσεις που επιδρούν αρνητικά σε αυτό, μπορεί να επιτευχθεί αστάθεια καλύτερη από

10^{-17} σε ένα δευτερόλεπτο.

4.3 Μεταφορά ενός frequency comb

Η μεταφορά μίας οπτικής συχνότητας προσφέρει πολύ μικρή ευελιξία, με αποτέλεσμα το ότι εάν ο χρήστης θέλει να μετρήσει μία συχνότητα που βρίσκεται σε απόσταση ακόμα και 100 GHz από τη μεταφερθείσα συχνότητα, πρέπει να χρησιμοποιήσει ένα optical frequency comb. Το ίδιο ισχύει για μέτρηση οπτικής συχνότητας εάν έχει μεταφερθεί μία μικροκυματική συχνότητα με διαμόρφωση πλάτους. Μία λύση στο παραπάνω πρόβλημα είναι να μεταφερθεί άμεσα ένα frequency comb, το οποίο μπορεί να σταθεροποιηθεί είτε σε μία οπτική, είτε σε μία μικροκυματική συχνότητα, αντί για τη μεταφορά της συχνότητας αυτής καθεαυτής.

Όπως και στις άλλες δύο μεθόδους, ο θόρυβος που προστίθεται από τη μεταφορά μέσω της οπτικής ίνας υπολογίζεται με τη σύγκριση μίας τοπικής αναφοράς στον πομπό και του σήματος που έχει μεταδοθεί, ενώ το αποτέλεσμα της σύγκρισης χρησιμοποιείται για να ρυθμιστεί μία διάταξη που τον ακυρώνει.

Τεχνική μετάδοσης	Αστάθεια σε 1 s
Μεταφορά μικροκυματικής συχνότητας με διαμόρφωση πλάτους ενός laser	1.3×10^{-15}
Άμεση μεταφορά οπτικού φέροντος	1.5×10^{-16}
Μεταφορά frequency comb	5×10^{-15}

Σύγκριση επιδόσεων των τριών τεχνικών για μετάδοση μέσω σκοτεινής ίνας μήκους 86km

4.4 Εφαρμογές μεταφοράς μέσω οπτικών ινών

Η μετάδοση προτύπων συχνοτήτων είναι απαραίτητη σε μία πληθώρα από εφαρμογές που βασίζονται στη σύγκριση συχνοτήτων ή στο συγχρονισμό, όπως η σύγκριση μεταξύ ατομικών ρολογιών που βρίσκονται σε διαφορετικά εργαστήρια, η ανάπτυξη πηγών ακτινών Χ νέας γενιάς σε εγκαταστάσεις επιταχυντή σωματιδίων για τη μελέτη πάρα πολύ γρήγορων φαινομένων σε κλάδους όπως η χημεία, η φυσική και η βιολογία, και ο συγχρονισμός τηλεσκόπιων, έτσι ώστε οι εικόνες που λαμβάνουν κάθε χρονική στιγμή να μπορούν να συνδεθούν με ακρίβεια προσομοιώνοντας ένα πολύ μεγαλύτερο τηλεσκόπιο. Όλες αυτές οι εφαρμογές απαιτούν μέγιστη βραχυπρόθεσμη ή μακροπρόθεσμη σταθερότητα και χαμηλό θόρυβο φάσης ή timing jitter.

5. Δορυφορική διανομή ρολογιού

5.1 GPS

Το GPS είναι μία μονόδρομη τεχνική σύμφωνα με την οποία σήματα μεταδίδονται από τους δορυφόρους (οι οποίοι βρίσκονται σε γνωστές τροχιές) μέσα από την ατμόσφαιρα προς σταθερούς επίγειους δέκτες. Οι συγκρίσεις γίνονται με πολλούς τρόπους που εξαρτώνται από το διαθέσιμο εξοπλισμό (ενός καναλιού (single channel), πολλών καναλιών (multi channel), P3 και PPP). Όλοι οι τρόποι παρέχουν την δυνατότητα πραγματικά παγκόσμιων συγκρίσεων, πράγμα που σημαίνει ότι ρολόγια που βρίσκονται σε οποιαδήποτε σημεία του πλανήτη μπορούν να συγκριθούν μεταξύ τους.

5.2 TWSTFT

Η μέθοδος TWSTFT χρησιμοποιεί δύο σήματα που μεταδίδονται ταυτόχρονα από τον ένα σταθμό στον άλλο μέσω του ίδιου μέσου, απλά προς διαφορετικές κατευθύνσεις. Αυτό της δίνει το πλεονέκτημα ότι οι καθυστερήσεις που οφείλονται στη μετάδοση αλληλοακυρώνονται σε μεγάλο βαθμό. Το TWSTFT, συνήθως χρησιμοποιεί γεωστατικούς τηλεπικοινωνιακούς δορυφόρους για την αναμετάδοση των σημάτων στο διάστημα. Αυτό όμως, σε συνδυασμό με το γεγονός ότι ο δορυφόρος πρέπει να έχει επαφή και με τους δύο επίγειους σταθμούς ταυτόχρονα, θέτει ένα περιορισμό στην απόσταση μεταξύ τους, αφού αυτή δεν μπορεί να είναι περισσότερη από περίπου 1000 km. Ένα πλεονέκτημα της μεθόδου είναι ότι τα αποτελέσματα της σύγκρισης των δύο ρολογιών είναι άμεσα διαθέσιμα στους δύο σταθμούς μετά την εγκατάσταση της ζεύξης μεταξύ τους, ενώ το GPS απαιτεί περαιτέρω επεξεργασία.

Με αυτές τις τεχνικές μετάδοσης, οι συχνότητες μπορούν να συγκριθούν με ακρίβεια 10^{-15} και οι κλίμακες χρόνου να συγχρονιστούν με αβεβαιότητα 1 ns.

6. Συμπεράσματα

Οι τρεις μέθοδοι μεταφοράς και σύγκρισης συχνοτήτων ή χρόνου (GPS, TWSTFT και οπτικές ίνες) έχουν η κάθε μία τα πλεονεκτήματα και τα μειονεκτήματά της. Η επιλογή της μεθόδου εξαρτάται κυρίως από την εφαρμογή.

Για παράδειγμα, εάν χρειάζεται αβεβαιότητα της τάξης των 5 ns, η μέθοδος GPS επαρκεί. Οι δέκτες του είναι φθηνοί σε σχέση με τις άλλες δύο μεθόδους, ενώ οι επιδόσεις του δεν εξαρτώνται από την εκάστοτε τοποθεσία των σταθμών και δίνει τη δυνατότητα σύγκρισης σταθμών που βρίσκονται οπουδήποτε στον πλανήτη. Επί του παρόντος, τα περισσότερα εργαστήρια που διατηρούν ένα τοπικό αντίτυπο του UTC χρησιμοποιούν το GPS, για να συνδεθούν με το διεθνές δίκτυο εργαστηρίων χρονισμού. Το TWSTFT παρέχει χαμηλότερη αβεβαιότητα και μεγαλύτερη σταθερότητα στη ζεύξη, ενώ οι άγνωστες καθυστερήσεις της μετάδοσης ακυρώνονται λόγω της συμμετρίας του. Περιορίζει, όμως, την απόσταση μεταξύ των επίγειων σταθμών και απαιτεί τη διαθεσιμότητα ενός γεωστατικού τηλεπικοινωνιακού δορυφόρου. Έτσι, η εγκαθίδρυση της σύνδεσης μεταξύ των σταθμών είναι πιο απαιτητική. Επίσης, ο αριθμός των ταυτόχρονων συγκρίσεων εξαρτάται από τον διαθέσιμο εξοπλισμό. Τέλος οι επίγειοι σταθμοί TWSTFT είναι ακριβοί και ο χρόνος χρήσης του δορυφόρου πρέπει να πληρώνεται.

Η ακρίβεια των μεθόδων GPS και TWSTFT είναι της τάξης του 10^{-15} σε μία ημέρα, που σημαίνει ότι οι αυτές οι δορυφορικές τεχνικές δεν επαρκούν για τη μετάδοση των σύγχρονων οπτικών προτύπων, αφού αυτά έχουν σταθερότητα της τάξης του 10^{-17} και βελτιώνονται συνεχώς. Η αναβάθμιση του δορυφορικού εξοπλισμού θα μπορούσε να βελτιώσει τις επιδόσεις τους, αλλά είναι πολύπλοκη και ακριβή. Επίσης, αυτά τα συστήματα μετάδοσης δεν επιτρέπουν βραχυπρόθεσμες συγκρίσεις σταθερότητας, αφού η ακρίβεια της τάξης των ns επιτυγχάνεται με χρήση ενός μακροπρόθεσμου μέσου όρου, και άρα οι βραχυπρόθεσμες διακυμάνσεις δεν μπορούν να μετρηθούν. Έτσι, οι δορυφορικές τεχνικές δεν παρέχουν τη βραχυπρόθεσμη σταθερότητα που απαιτείται από τις εφαρμογές συγχρονισμού. Επίσης, δεν είναι πρακτικές σε εφαρμογές που απαιτούν τη μετάδοση μίας αναφοράς συχνότητας μέσα στο ίδιο συγκρότημα.

Οι οπτικές ίνες από την άλλη, παρέχουν τις καλύτερες επιδόσεις λόγω του μεγαλύτερου εύρους ζώνης της μετάδοσης, όταν είναι διαθέσιμες σκοτεινές ίνες αποκλειστικής χρήσης. Η μετάδοση οπτικού φέροντος σε σκοτεινή ίνα, για παράδειγμα, έχει φτάσει αβεβαιότητα της τάξης του 10^{-16} σε ένα δευτερόλεπτο, δηλαδή πολύ καλύτερη από την ακρίβεια των δορυφορικών μεθόδων. Επίσης, όπως έχει αναφερθεί, οι οπτικές ίνες χαρακτηρίζονται από μικρές απώλειες, χαμηλή διασπορά και επεκτασιμότητα. Όμως, όταν η μετάδοση γίνεται για μεγάλες αποστάσεις, απαιτείται η χρήση επαναληπτών. Παρόλο που το υψηλό κόστος για πολύ μακρινή σύνδεση μέσω οπτικών ινών μπορεί να παρεμποδίσει εφαρμογές σε επίπεδο ηπείρων, δίκτυα που περιορίζονται σε πανεπιστημιακές περιοχές και δίκτυα για πειραματικούς σκοπούς ανάμεσα σε πόλεις μπορούν να υλοποιηθούν πολύ πιο εύκολα. Τα διαστημικά πειράματα απαιτούν επίσης επίγειες εγκαταστάσεις για το συγχρονισμό, άρα αποτελούν άλλη μία εφαρμογή για την οποία συνίσταται η χρήση οπτικών ινών.

Ίσως το μεγαλύτερο πλεονέκτημα της μετάδοσης μέσω οπτικών ινών, εκτός από τη μεγαλύτερη ακρίβεια που πετυχαίνει σε σχέση με τις δορυφορικές μεθόδους, είναι η δυνατότητα ανάπτυξης ενός δικτύου διανομής προτύπων μέσω των υποδομών του διαδικτύου. Εξάλλου, οι υποδομές των οπτικών ινών συνεχίζουν να αυξάνονται, πράγμα που σε συνδυασμό με την εκμετάλλευση των ήδη υπάρχουσων εγκαταστάσεων, τις καθιστά μία ελκυστική και αποδοτική, από πλευράς κόστους, λύση για ένα παγκόσμιο σύστημα μεταφοράς προτύπων. Τα πειράματα μεταφοράς οπτικού προτύπου, μέσω τηλεπικοινωνιακού δικτύου που μεταφέρει ταυτόχρονα ψηφιακά δεδομένα, έχουν πολύ θετικά αποτελέσματα. Η ακρίβεια που πετυχαίνεται ακόμα και μέσω των πολύ πιο θορυβωδών τηλεπικοινωνιακών ινών είναι και πάλι καλύτερη από την αντίστοιχη των δορυφορικών τεχνικών. Έτσι, μέσω του διαδικτύου, είναι δυνατόν να συνδεθούν εργαστήρια που βρίσκονται σε μεγάλες αποστάσεις οπουδήποτε στον πλανήτη, χωρίς τη βοήθεια δορυφόρου.

Τέλος, θα μπορούσαν να συνδυαστούν όλες οι παραπάνω τεχνικές μεταφοράς για ένα παγκόσμιο σύστημα διανομής προτύπων. Στις σύντομες αποστάσεις μερικών εκατοντάδων χιλιομέτρων, τα εργαστήρια μπορούν να συνδέονται μέσω οπτικών ινών, ενώ αυτά τα συγκροτήματα θα συνδέονται μεταξύ τους με διηπειρωτικές ζεύξεις TWSTFT και GPS.

Αναφορές

1. D. Piester, M. Rost, M. Fujieda, T. Feldmann, A. Bauch, "Remote atomic clock synchronization via satellites and optical fibers", Advances in Radio Science, (2011)

2. Giuseppe Marra, Radan Slavík, Helen S. Margolis, Stephen N. Lea, Periklis Petropoulos, David J. Richardson, and Patrick Gill, "High-resolution microwave frequency transfer over an 86-km-long optical fiber network using a mode-locked laser", Optics Letters, Vol. 36,

Issue 4, pp. 511-513 (2011)

3. A. Pape, O. Terra, J. Friebe, M. Riedmann, T. Wübbena, E. M. Rasel, K. Predehl, T. Legero, B. Lipphardt, H. Schnatz, and G. Grosche , "Long-distance remote comparison of ultrastable optical frequencies with 10-15 instability in fractions of a second", Optics Express, Vol. 18, Issue 20, pp. 21477-21483 (2010)

4. H. S. Margolis, G. Marra, V. Tsatourian, S. N. Lea, D. T. Reid and P. Gill, "Optical frequency combs and applications at NPL" , (2010)

5. Olivier Lopez, Adil Haboucha, Fabien Kéfélian, Haifeng Jiang, Bruno Chanteau, Vincent Roncin, Christian Chardonnet, Anne Amy-Klein, and Giorgio Santarelli, "Cascaded multiplexed optical link on a telecommunication network for frequency dissemination", Optics Express, Vol. 18, Issue 16, pp. 16849-16857 (2010)

6. Seth M. Foreman, Kevin W. Holman, Darren D. Hudson, David J. Jones, and Jun Ye, "Remote transfer of ultrastable frequency references via fiber networks", Rev. Sci. Instrum. 78, 021101 (2007)

7. Seth M. Foreman, Andrew D. Ludlow, Marcio H. G. de Miranda, Jason E. Stalnaker, Scott A. Diddams, and Jun Ye, "Coherent Optical Phase Transfer over a 32-km Fiber with 1 s Instability at 10-17", Phys. Rev. Lett. 99, 153601 (2007)

8. R. L. Tjoelker, M. Calhoun, P. F. Kuhnle, J. Lauf, and R. L. Sydnor, "Master Clock and Time Distribution System for the NASA Deep Space Network: Part I", IPN Progress Report 42-167, (2006)

9. Jun Ye, Jin-Long Peng, R. Jason Jones, Kevin W. Holman, John L. Hall, David J. Jones, Scott A. Diddams, John Kitching, Sebastien Bize, James C. Bergquist, Leo W. Hollberg, Lennart Robertsson, and Long-Sheng Ma, "Delivery of high-stability optical and microwave frequency standards over an optical fiber network", JOSA B, Vol. 20, Issue 7, pp. 1459-1467 (2003)

10. Diddams SA, Jones DJ, Ye J, Cundiff ST, Hall JL, Ranka JK, Windeler RS, Holzwarth R, Udem T, Hansch TW, "Direct link between microwave and optical frequencies with a 300 THz femtosecond laser comb", Phys Rev Lett. 29;84(22):5102-5 (2000)

Δημιουργία Εκπαιδευόμενης Μηχανής Παραγωγής Προβλέψεων Κειμένου

Παναγιώτης Σ. Σάκκος
panos.sakkos@gmail.com

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Η παρούσα εκπόνηση έχει σκοπό τη δημιουργία μίας μηχανής παραγωγής προβλέψεων, η οποία δέχεται ως είσοδο ακολουθίες γραμμάτων και παράγει στην έξοδό της έναν πιθανοτικό χώρο. Αυτή η μηχανή διαθέτει το χαρακτηριστικό ότι μπορεί να παράγει, εξατομικευμένες στο χρήστη της, προβλέψεις, καθώς και να εντοπίσει και να μάθει λέξεις που δεν γνωρίζει. Η υλοποίηση που θα παρουσιαστεί, στοχεύει στην αγγλική γλώσσα. Απώτερος σκοπός της είναι να χρησιμοποιηθεί από ένα λογισμικό πληκτρολόγιο για συσκευές αφής και τα αποτελέσματά της να αξιοποιηθούν οπτικά από το λογισμικό πληκτρολόγιο. Χρησιμοποιώντας τα αποτελέσματα της μηχανής αυτής, δημιουργήσαμε ένα λογισμικό πληκτρολόγιο που μεταβάλει τα μεγέθη των εικονικών πλήκτρων, με αποτέλεσμα να διευκολύνει τη γραπτή επικοινωνία των χρηστών απο τις συσκευές αφής τους.

Λέξεις-Κλειδιά: Πρόβλεψη, Πιθανότητες, Εκπαίδευση, Εξατομίκευση

Επιβλέπων: Δημήτριος Γουνόπουλος, Αναπληρωτής Καθηγητής

1. Εισαγωγή

Το πρόβλημα που καλούμαστε να αντιμετωπίσουμε είναι η καταγραφή της γραπτής επικοινωνίας ενός χρήστη, με σκοπό να προβλέψουμε επιτυχώς ποιό γράμμα θα πληκτρολογήσει, σε βάθος ενός γράμματος. Αυτό, πέρα από την καταγραφή της γραπτής επικοινωνίας του χρήστη, απαιτεί και μία αρχική γνώση, προκειμένου να βρισκόμαστε σε θέση να δώσουμε προβλέψεις χωρίς να έχουμε καταγράψει κάποια γραπτή του ενέργεια. Τα κεφάλαια που συντελούν το παρόν κείμενο, ακολουθούν τα βήματα που ακολουθήθηκαν, προκειμένου να αναπτυχθεί η μηχανή των προβλέψεων που πραγματεύεται η εργασία.

Αρχικά, έπρεπε να βρεθεί ένα σύνολο λέξεων της αγγλικής γλώσσας και να διαμορφωθεί σε μία απλή μορφή, η οποία θα ήταν ένα αρχείο κειμένου που θα περιείχε μόνο τις λέξεις διαχωρισμένες με κάποιο διαχωριστή. Στη συνέχεια, σχεδιάστηκε και αναπτύχθηκε η μηχανή και η διεπαφή που τη χρησιμοποιούσε προκειμένου να μπορεί να δοκιμαστεί. Επιπλέον, θα παρουσιαστεί μία υλοποίηση εισόδου για συσκευές αφής, η οποία εκμεταλλεύεται τη μηχανή που θα περιγραφεί. Η διεπαφή αυτή είναι ένα δυναμικό λογισμικό πληκτρολόγιο, το οποίο αποφασίζει ποιά πλήκτρα να συρρικνώσει με βάση τις προβλέψεις της μηχανής προβλέψεων. Με αυτό τον τρόπο, διευκολύνει το χρήστη κατά την πληκτρολόγηση και υποσυνείδητα τον οδηγεί εύκολα και γρήγορα στη λέξη που θέλει να πληκτρολογήσει. Επίσης, υπάρχει μηχανισμός ακύρωσης της πρόβλεψης, σε περίπτωση που το πλήκτρο που επιθυμεί να πληκτρολογήσει έχει συρρικνωθεί λόγω κακής πρόβλεψης.

Αυτή η διεπαφή είναι μία από τις πολλές που μπορούν να υπάρξουν και να αξιοποιήσουν τις προβλέψεις της μηχανής προβλέψεών μας. Η γλώσσα προγραμματισμού που επιλέχθηκε για να υλοποιηθεί οποιαδήποτε αυτοματοποιημένη βοηθητική λειτουργία, καθώς και η μηχανή και το κέλυφος διεπαφής της, είναι η C# [1]. Πρόκειται για μία αντικειμενοστραφής γλώσσα, σχεδιασμένη από τη Microsoft, η οποία αυτή τη στιγμή είναι η πιο ισχυρή γλώσσα γενικού σκοπού. Για να τρέξει ο κώδικας σε Unix-like λειτουργικά συστήματα απαιτείται να έχει εγκαταστασθεί το Mono [2] framework. Η άδεια του πηγαίου κώδικα είναι ακαδημαϊκή και διαθέσιμη στο διαδίκτυο και συγκεκριμένα στο Github. Η ανάπτυξη του λογισμικού πληκτρολογίου πραγματοποιήθηκε σε γλώσσα Java και περιβάλλον ανάπτυξης Eclipse [3] με την προσθήκη του Android [4] Software Development Kit και μπορεί να τρέξει σε οποιαδήποτε Android συσκευή που διαθέτει Android 2.1 ή νεότερη έκδοση του λειτουργικού. Αξίζει να σημειωθεί πως η συρρίκνωση των πλήκτρων είναι δυναμική, ώστε να μπορεί να συμπεριφέρεται σωστά σε οποιοδήποτε μέγεθος οθόνης και έχει σχεδιαστεί έτσι ώστε να συμπεριφέρεται ανάλογα και σε οριζόντιο προσανατολισμό της οθόνης. Να σημειωθεί, ότι όταν θα αναφέρεται ο όρος «ακολουθία», θα πρόκειται για λέξη και όχι για ολόκληρη είσοδο που έχει πληκτρολογηθεί. Αυτό πρακτικά σημαίνει πως κάθε φορά που πληκτρολογείται στην είσοδο ένας διαχωριστής λέξης (‘ ‘, ‘,’, ‘.’, ‘!’ κ.α.) τότε η ακολουθία γίνεται κενή. Επίσης, θα παρουσιάσουμε μόνο την περίπτωση της γραμμικής εισόδου. Αυτό σημαίνει πως δεν θα εξετάσουμε τις περιπτώσεις όπου ο χρήστης επιθυμεί να διαγράψει ένα γράμμα ή να μεταφερθεί σε κάποια παλιότερα πληκτρολογημένη ακολουθία γραμμάτων. Παρόλ’ αυτα, αυτές οι περιπτώσεις είναι απαραίτητο να καλυφθούν από τη διεπαφή λογισμικού πληκτρολογίου, η οποία θα αξιοποιήσει τη μηχανή προβλέψεων. Στη διεπαφή που θα παρουσιάσουμε, έχουν καλυφθεί οι περιπτώσεις αυτές επιτυχώς.

2. Αξιοποίηση της μηχανής σε περιβάλλον αφής

Η υλοποίηση η οποία παρουσιάστηκε, καλύπτει, σε πολύ μεγάλο βαθμό, το σενάριο του να χρησιμοποιηθεί από κάποιο πληκτρολόγιο λογισμικού το οποίο εκτελείται σε φορητή συσκευή. Διαθέτει τη δυνατότητα εκπαίδευσης και διατηρεί την κατάστασή του όταν αυτό δεν εκτελείται. Επομένως, με αυτό τον τρόπο ο χρήστης θα είναι σε θέση να λαμβάνει συνεπή εξατομικευμένα απότελέσματα μετά από μία επανεκίνηση της συσκευής του. Επίσης, τα αποτελέσματα παράγονται σε σταθερό χρόνο, αφού το μόνο που κάνει είναι να τρέξει τη φόρμουλα των αξιολογήσεων για κάθε λεξικογραφικό υποδέντρο. Το γεγονός ότι η έξοδος υπολογίζεται σε σταθερό χρόνο είναι εξαιρετικά σημαντικό, διότι αν είχε μεγαλύτερη χρονική πολυπλοκότητα (έστω και πολυωνυμική), θα καθιστούσε αδύνατη την αξιοποίησή της από φορητές συσκευές. Αυτό συμβαίνει διότι τέτοιες συσκευές, προκειμένου να έχουν μεγαλύτερη αυτονομία, διαθέτουν αργούς επεξεργαστές αρχιτεκτονικής ARM [9], επομένως, μη σταθερή ή ακόμα και γραμμική πολυπλοκότητα είναι πολύ πιθανό να σήμαινε χρόνο εκτέλεσης αντιληπτό από το χρήστη. Αυτό, πρακτικά, θα καθιστούσε την εφαρμογή μας άχρηστη. Συμπερασματικά λοιπόν, οι προσθήκες πρέπει να προσανατολιστούν στην όσο το δυνατόν καλύτερη εξατομίκευση που μπορεί να γίνει για τον εκάστοτε χρήστη.

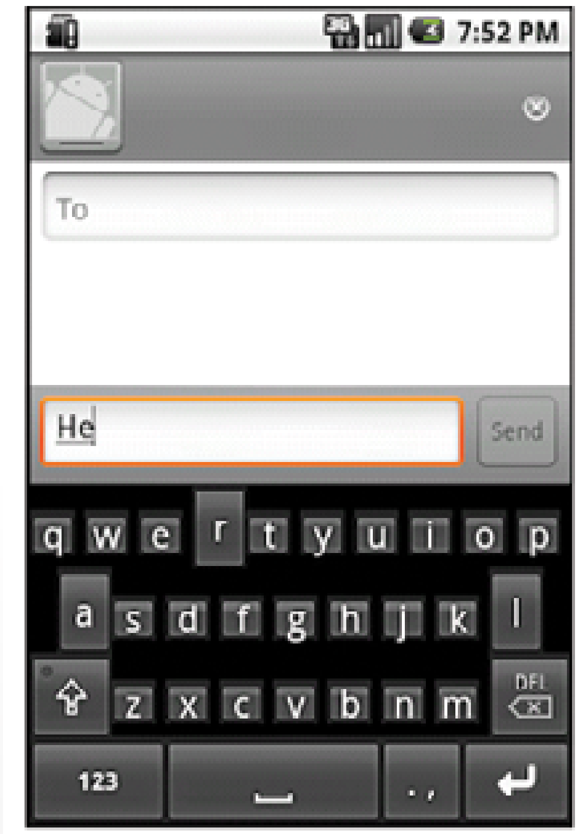
Δεδομένου ότι η παρούσα υλοποίηση προορίζεται για χρήση της σε φορητές συσκευές, οι οποίες κατα πάσα πιθανότητα θα είναι συνδεδεμένες στο διαδίκτυο, αυξάνει κατά πολύ τις πληροφορίες που θα μπορούσε να αξιοποιήσει η μηχανή, προκειμένου να παράγει ακόμα πιο ισχυρές προβλέψεις. Τα δύο πιο ενδιαφέροντα δεδομένα που οι φορητές αυτές συσκευές μπορούν να προσφέρουν, είναι ο χρόνος και ο χώρος. Οτιδήποτε πληκτρολογούμε κατά τη γραπτή μας επικοινωνία, έχει σχέση με το πότε και το πού το γράφουμε. Επίσης, όσον αφορά το χρόνο, είναι σημαντικό να μπορεί η μηχανή να προσαρμόζεται στο προσωρινό λεξιλόγιο του χρήστη. Ο κάθε άνθρωπος, στη γραπτή του επικοινωνία, χρησιμοποιεί ένα μικρό υποσύνολο των λέξεων της γλώσσας που χρησιμοποιεί. Επίσης, αυτό το γραπτό λεξιλόγιο, διαφέρει σημαντικά από άνθρωπο σε άνθρωπο. Επομένως, δεν είναι πρακτικό ο κόσμος γνώσης της μηχανής να είναι από την αρχή όλες αυτές οι εκατοντάδες χιλιάδες λέξεις που χρησιμοποιήθηκαν στην παρούσα υλοποίηση. Ο λόγος, πέρα από το ότι θα κάνει την εφαρμογή πολύ αργή κατά την εκκίνηση της, είναι πως ο χρήστης θα χρησιμοποιήσει περίπου 1000 με 5000 διαφορετικές λέξεις κατά τη γραπτή του επικοινωνία. Επομένως, η μηχανή πρέπει να είναι σε θέση να μάθει αυτές τις λέξεις και μόνο αυτές. Αυτό θα επιτευχθεί, ξεκινώντας τον κόσμο γνώσης της μηχανής με τις 1000 εώς 5000 πιο συχνά χρησιμοποιούμενες λέξεις της αγγλικής γλώσσας και στη συνέχεια, όταν ο αριθμός λέξεων της μηχανής φτάσει σε κάποιο οριακό σημείο, να διατηρούνται οι 1000 με 5000 πιο πρόσφατα χρησιμοποιημένες λέξεις και οι υπόλοιπες να διαγράφονται από τον κόσμο γνώσης της μηχανής. Με αυτό τον τρόπο επιλύουμε και ένα άλλο μείζων πρόβλημα των συστημάτων πρόβλεψης γραπτού λόγου.

Ένα συχνό φαινόμενο που παρατηρείται σε τέτοια συστήματα, είναι η αποθήκευση λάθος πληκτρολογήσεων του χρήστη. Αυτό έχει ως αποτέλεσμα να αποθηκεύονται μόνιμα και να προκαλούν θόρυβο στις προβλέψεις, ακόμα και να προκαλούν δυσάρεστη εντύπωση στο χρήστη όταν βλέπει πως το λογισμικό του προτείνει μία λάθος λέξη. Επομένως, βασιζόμενοι στο γεγονός ότι οι λάθος λέξεις πληκτρολογούνται λίγες φορές, το ξεκαθάρισμα, που θα γίνεται για να προσεγγίσουμε το προσωπικό σύνολο λέξεων του χρήστη, θα έχει ως απότέλεσμα να ξεχνιούνται από το σύνολο γνώσης της μηχανής και οι λάθος αυτές πληκτρολογήσεις, αφού δεν ανήκουν στο σύνολο λέξεων γραπτής επικοινωνίας του χρήστη. Αυτό φυσικά απαιτεί να κρατείται μια χρονοσφραγίδα κάθε φορά που πληκτρολογείται μία λέξη. Όσον αφορά τον γεωγραφικό χώρο στον οποίο πληκτρολογεί κάποια λέξη ο χρήστης, πρέπει να δωθεί προσοχή σε δύο σημεία που παρουσιάζουν δυσκολίες. Αφενός, πρέπει να λάβουμε υπόψην, πως η γεωγραφική τοποθεσία της συσκευής δεν είναι πάντα διαθέσιμη. Είτε ο χρήστης μπορεί να έχει απενεργοποιήσει αυτή τη λειτουργία, είτε η συσκευή να μην διαθέτει GPS [8] (Global Positioning System) ή συνεχή σύνδεση στο διαδίκτυο. Επομένως, θα πρέπει να ενεργοποιηθούν οι προβλέψεις που λαμβάνουν υπόψην το χώρο πληκτρολόγησης, μόνο αν κριθεί πως έχουν αποθηκευτεί αρκετές τοποθεσίες. Το δεύτερο λεπτό σημείο της γεωγραφικής καταγραφής, είναι πως πρέπει να απόφασιστεί το ποιά περιοχή θα θεωρείται ως χώρος για το χρήστη. Για παράδειγμα, για κάποιον που ταξιδεύει πολύ συχνά στο εξωτερικό λόγω των απαιτήσεων της εργασίας του, ως ένας χώρος είναι πολύ πιθανό να οριστεί ολόκληρη η χώρα διαμονής του. Ενώ, για κάποιον που εργάζεται εντός της χώρας όπου διαμένει, ως χώρος πολύ πιθανό να πρέπει να θεωρηθεί ο χώρος διαμονής του, ο χώρος εργασίας του και το εξοχικό του. Επομένως, θα πρέπει να μπορεί η μηχανή να ομαδοποιήσει τα γεωγραφικά σημεία που θα καταγράφει σε χώρους, όπου ο χρήστης όταν βρίσκεται μέσα τους πληκτρολογεί διαφορετικές λέξεις. Τέλος, αναπόφευκτο είναι πως θα πρέπει να υπάρξει υποστήριξη και για γλώσσες πέρα της Αγγλικής.

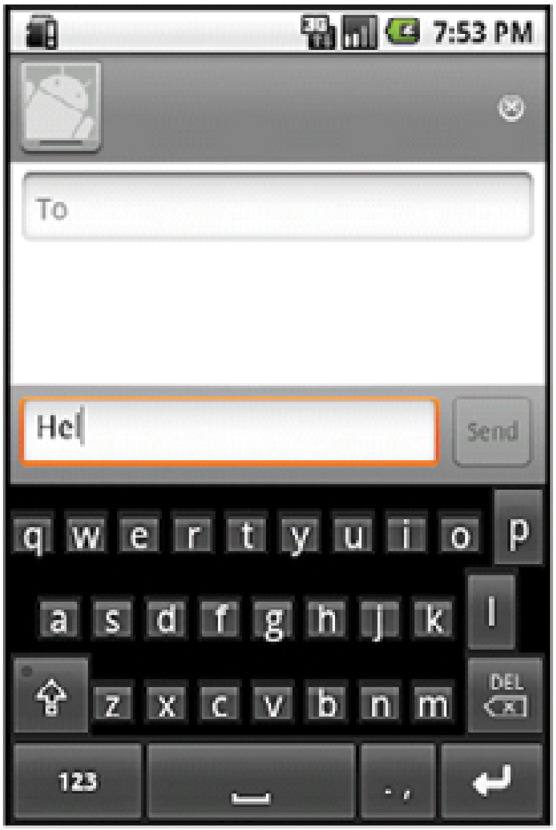
3. Παρουσίαση διεπαφής λογισμικού πληκτρολογίου

Αφού ολοκληρώσαμε τη μηχανή που παρουσιάσαμε, σχεδιάσαμε μία διεπαφή λογισμικού πληκτρολογίου, η οποία εκμεταλεύεται κατάλληλα τα αποτελέσματα της μηχανής μας. Η ιδέα μας ήταν να μικραίνουμε τα εικονικά πλήκτρα που βρίσκονται στην οθόνη του χρήστη, τα οποία ήταν χαμηλά στις αξιολογήσεις. Ακολουθούν οι εικόνες που καταγράψαμε από την οθόνη του κινητού πληκτρολογώντας τη λέξη «Hello»:

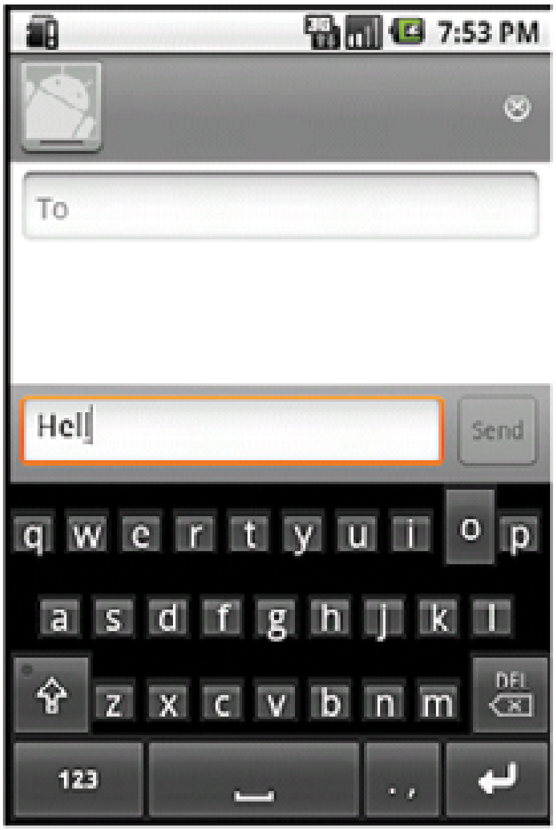
Εικόνα 1: WriteRight «H»



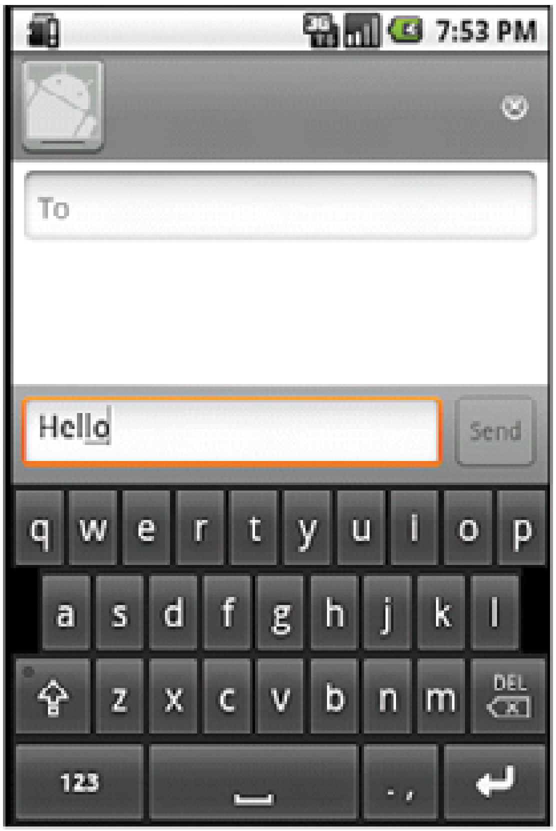
Εικόνα 2: WriteRight «He»



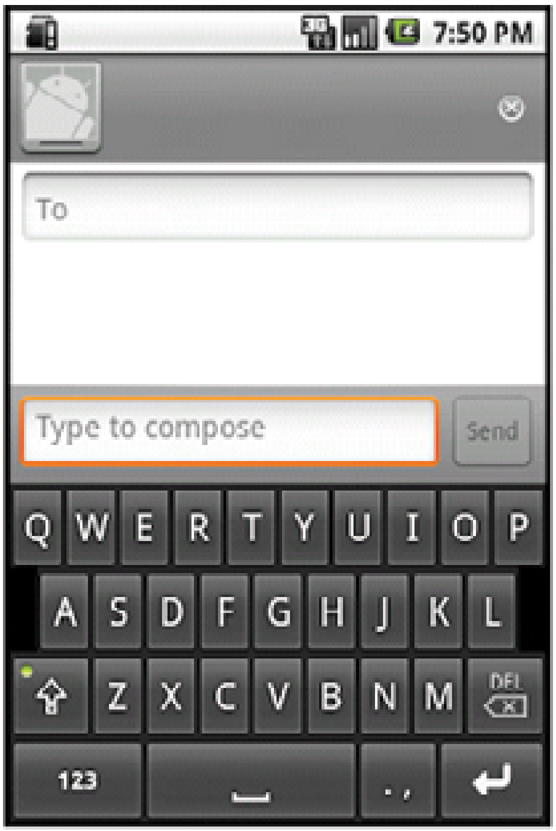
Εικόνα 3: WriteRight «Hel»



Εικόνα 4: WriteRight «Hell»

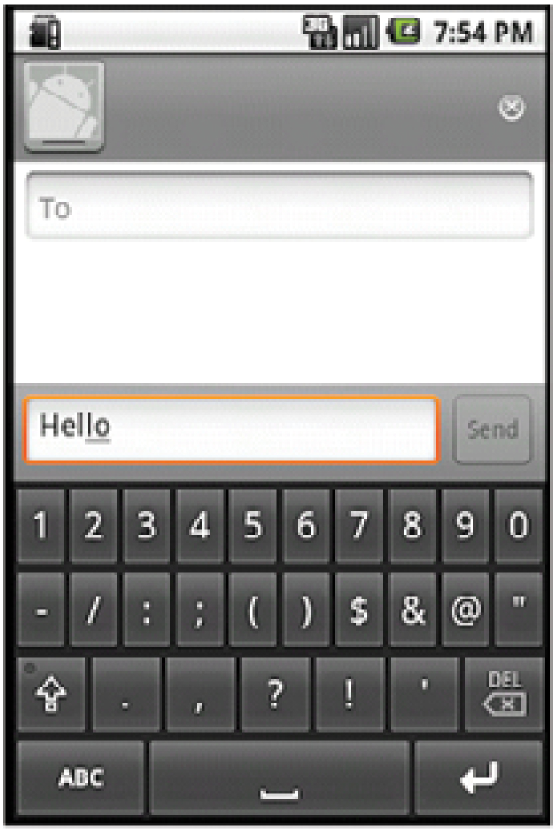


Εικόνα 5: WriteRight «Hello»

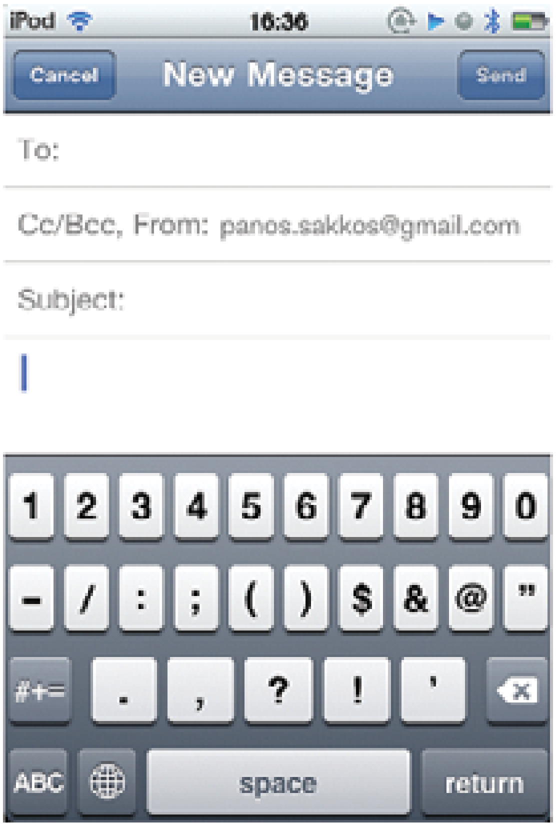


Εικόνα 6: WriteRight «Hello»

Πρόκειται για ένα πλήρες πληκτρολόγιο, το οποίο καλύπτει όλα τα γνωστά αλφαριθμητικά σύμβολα. Η σχεδίαση του πληκτρολογίου που παρουσιάζει τα πλήκτρα, που περιλαμβάνουν τους αριθμούς καθώς και τα υπόλοιπα μη αλφαβητικά σύμβολα, έγινε σύμφωνα με το πληκτρολόγιο του λειτουργικού iOS 5.



Εικόνα 7: Όψη πλήκτων αριθμών και συμβόλων στο WriteRight



Εικόνα 8: Όψη πλήκτων αριθμών και συμβόλων στο iOS

Προκειμένου να είναι έτοιμο ένα ολοκληρωμένο λογισμικό πληκτρολόγιο για να χρησιμοποιηθεί από τους χρήστες της πλατφόρμας Android, χρησιμοποιήσαμε δείγμα από τις 1000 πιο γνωστές λέξεις της αμερικάνικης γλώσσας. Επίσης, εφαρμόσαμε ένα αλγόριθμο παρόμοιο με τον LRU [10] (Least Recently Used) πάνω στη χρονοσφραγίδα των λέξεων (την οποία και κρατάμε και ανανεώνουμε κάθε φορά που πληκτρολογεί ο χρήστης μία λέξη) για να προσεγγίσουμε όσο το δυνατόν καλύτερα το προσωπικό λεξιλόγιο του χρήστη, αλλά και για να αποφύγουμε την αποθήκευση λάθος λέξεων, οι οποίες προκαλούν θόρυβο στα αποτελέσματά μας και προκαλούν δυσαρέσκεια στο χρήστη, όπως αναφέραμε και στο προηγούμενο κεφάλαιο. Το όριο, που ορίσαμε για να γίνεται αυτή ανανέωση της γνώσης, είναι οι 1500 λέξεις.

Η διεπαφή περιλαμβάνει ένα μηχανισμό με τον οποίο ο χρήστης μπορεί να ακυρώσει μία πρόβλεψη η οποία είναι λάθος. Αυτό σημαίνει πως, αν ο χρήστης θέλει να πληκτρολογήσει ένα γράμμα το οποίο η διεπαφή μας έχει σμικρύνει, μπορεί να ακυρώσει την πρόβλεψη. Ο μηχανισμός που σχεδιάστηκε και υλοποιήθηκε είναι να σύρει το δάκτυλό του πάνω στο πληκτρολόγιο δεξία ή αριστερά-διαγώνια. Τότε, η παραγωγή προβλέψεων ακυρώνεται και επαναφέρονται τα κανονικά μεγέθη όλων των πλήκτρων, έως ότου να πληκτρολογηθεί ένας διαχωριστής λέξης.

Το κριτήριο με το οποίο γίνεται η επιλογή για το ποιά πλήκτρα θα σμικρύνουμε, είναι δυναμικό. Πιο συγκεκριμένα, έχουμε μία μεταβλητή, η οποία υποδυκνύει το μέγιστο αριθμό των γραμμάτων που αφήνουμε στο κανονικό τους μέγεθος. Φυσικά, αν η μηχανή προβλέψεων μας επιστρέψει γράμματα τα οποία υπολόγισε πως έχουν πιθανότητα μηδέν να εμφανιστούν τότε τα σμικρύνουμε. Επίσης, υπάρχει ένα όριο σωστών προβλέψεων, αν αυτό ξεπεραστεί τότε το μηδενίζουμε και μειώνουμε τη μεταβλητή που φυλάει το μέγιστο αριθμό πλήκτων με μη μηδενική αξιολόγηση που θα αφήνουμε στο κανονικό τους μέγεθος. Σε περίπτωση λάθους πρόβλεψης (κάτι το οποίο εντοπίζεται αν ο χρήστης ακυρώσει την πρόβλεψή μας) μηδενίζουμε το μετρητή. Αυτό δίνει μία συμπεριφορά δυναμικής επιθετικότητας στη διεπαφή, σύμφωνα με την οποία όσο συνεχόμενες σωστές προβλέψεις παράγει, τόσο πιο πολλά πλήκτρα θα συρρικνώσει, αφού μπορεί να νιώθει μεγαλύτερη αυτοπεποίθηση για τα αποτελέσματα που παράγει.

Τέλος, προκειμένου να υπάρξει περισσότερος χώρος προς εκμετάλλευση στο πληκτρολόγιο, δεν τοποθετήσαμε κάποιο πλήκτρο το οποίο θα κρύβει το πληκτρολόγιο (όπως έχουν τα περισσότερα λογισμικά πληκτρολόγια), αλλά προσθέσαμε τη δυνατότητα, σύροντας ο χρήστης το δάκτυλό του πάνω στο πληκτρολόγιο προς τα κάτω, να κρύβεται το πληκτρολόγιο. Τέλος, αξίζει να σημειωθεί πως το πληκτρολόγιο αυτό, όταν εγκατασταθεί και ενεργοποιηθεί, τότε αντικαθιστά το πληκτρολόγιο του λειτουργικού. Επομένως, είναι διαθέσιμο σε οποιαδήποτε εφαρμογή τρέχει ο χρήστης και δέχεται είσοδο μέσω λογισμικού πληκτρολογίου. Μπορείτε να το βρείτε στο ψηφιακό κατάστημα εφαρμογών της Google.

4. Συμπεράσματα

Η πρόβλεψη εισόδου γραπτής επικοινωνίας είναι ένα ζήτημα που βρίσκεται στο κέντρο της εμπορικής όψης της επιστήμης των υπολογιστών. Οι εταιρείες πληροφορικής παγκοσμίως, που βρίσκονται στην αιχμή της τεχνολογίας και πραγματοποιούν τεράστια κέρδη σε κάθε οικονομικό τους τετράμηνο, προσπαθούν να κάνουν τη γραπτή επικοινωνία των φορητών συσκευών αφής όσο το δυνατόν πιο εύκολη και γρήγορη γίνεται. Η ύπαρξη μιας ιδανικής μηχανής προβλέψεων εισόδου γραπτής επικοινωνίας θα άνοιγε δρόμους για διεπαφές επαφής, όπου θα μείωναν δραματικά το χρόνο επικοινωνίας των χρηστών.

Παρόλ' αυτά, η κατασκευή μίας ιδανικής μηχανής πρόβλεψης δεν είναι δυνατή, δεδομένου πως πολλές φορές ακόμα και ο ίδιος ο χρήστης δεν έχει αποφασίσει τι θέλει να γράψει. Τα αποτελέσματα από τα πειράματα που παρουσιάστηκαν στην παρούσα εκπόνηση, μας δίνουν αισιοδοξία ώστε να πιστεύουμε σε μηχανές που αν εκπαιδευτούν, θα βρίσκονται σε θέση να προβλέπουν το επόμενο σύμβολο της ακολουθίας εισόδου μέσα στις δύο με τρεις πρώτες αξιολογήσεις. Πιο συγκεκριμένα, για τις εκτελέσεις που πραγματοποιήθηκαν, παρατηρούμε αφενός, πως ανάμεσα στις προβλέψεις που παρήγαγε η μηχανή, το σύμβολο που ακολουθούσε, βρισκόταν μέσα στις τέσσερεις με πέντε πρώτες προβλέψεις, πράγμα πολύ θετικό. Αυτό μας ενθαρρύνει, διότι σε περίπτωση που χρησιμοποιηθεί η μηχανή αυτή για πρόβλεψη κειμένου, θα αρκεί να θεωρηθούν ως πιθανά επόμενα γράμματα τα πέντε πρώτα σε αξιολόγηση.

Στη συνέχεια, ανάλογα το σχεδιασμό της διεπαφής θα λαμβάνεται υπόψη και η τιμή της αξιολόγησης του κάθε από τα πέντε αυτά γράμματα, διαφορετικά, μας αρκεί να γνωρίζουμε μόνο τα πέντε πρώτα πιο πιθανά γράμματα που θα ακολουθήσουν. Επίσης, όταν εντοπίστηκε μια λέξη η οποία δεν ήταν γνωστή, τη δεύτερη φορά που λάμβανε στην είσοδο την υπακολουθία αυτής της λέξης, συμπεριέλαβε στις προβλέψεις το καινούριο γράμμα που απότέλεσε τη λέξη "panos", ενώ στην πρώτη γραφή της λέξης, στην ακολουθία «pano» δεν είχε πρόβλεψη για το 's' ως επόμενο γράμμα. Αυτό πρακτικά σημαίνει πως δεν γνώριζε κάποια ακολουθία «panos», επομένως όταν την διδάχτηκε παρήγαγε μη μηδενική πιθανοτητα. Αυτή είναι μία συνιστώσα της επιθυμητής συμπεριφοράς που θέλαμε να πετύχουμε. Παρόλ' αυτα, η πιθανότητα αυτή ήταν πάρα πολύ μικρή, επειδή αφενός δεν υπήρχαν άλλες λέξεις που να ακολουθούν την ακολουθιά «panos», αφετέρου επειδή είχε πληκτρολογηθεί μόνο μία φορά, επομένως και είχε χαμηλή δημοφιλία.

Έκπληξη πρέπει να προκαλεί το γεγονός ότι παρόλο που οι παραπάνω εκτελέσεις πραγματοποιήθηκαν χωρίς κάποια προηγούμενη εκπαίδευση, οι προβλέψεις μόνο αποπροσανατολισμένες δεν μπορούν να χαρακτηριστούν. Το γράμμα που επρόκειτο να πληκτρολογηθεί, βρισκόταν σχεδόν πάντα μέσα στις πέντε ισχυρότερες προβλέψεις. Αυτό μας ενθαρρύνει πως με τακτική χρήση, το γράμμα που θα πληκτρολογηθεί θα βρίσκεται ακόμα υψηλότερα στην κατάταξη των προβλέψεων, λόγω της όλο και αυξανόμενης δημοφιλίας που θα παρουσίασει.

Αναφορές

1. Anders Hejlsberg, Scott Wiltamuth, Peter Golde, C# Language Specification, Bibliometrics, 2003

2. Mark Mamone, Introducing Development Tools and MonoDevelop, 2006

3. Jeff Mc Affer Jean-Michel Lemieux, Eclipse Rich Client Platform: Designing, Coding, and Packaging Java(TM) Applications, 2005

4. Sharon P. Hall, Clear Lake, Operating systems for mobile computing, 2009

5. George A. Miller, WordNet: a lexical database for English, 1995

6. Michael Widenius, Davis Axmark, Mysql Reference Manual, 2002

7. Edward Fredkin, Bolt Beranek and Newman, Trie memory, ACM Magazine Communications, vol. 3, no. 1, September 1960

8. John Spencer, Brian G. Frizzelle, Philip H. Page, John B. Vogler, Global Positioning System, 2008

Χαρτογράφηση Αναρτημένων Βίντεο στο YouTube

Γεώργιος Β. Σταμούλης
gstam@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Στην εργασία αυτή προσπαθήσαμε να χαρτογραφήσουμε τα βίντεο που υπάρχουν στο YouTube, με βάση πληροφορίες για τοποθεσίες που μπορεί να υπάρχουν στα δεδομένα που συνοδεύουν κάθε βίντεο στον ιστοχώρο αυτό. Κατασκευάσαμε έναν αλγόριθμο για αναζήτηση τοποθεσιών και ημερομηνιών στα δεδομένα των βίντεο, μέσα από ένα πλήθος τοποθεσιών που δημιουργήσαμε υπό μορφή λιστών, ανάλογα με το είδος της τοποθεσίας. Διαπιστώνουμε ότι ένα μεγάλο ποσοστό των βίντεο του YouTube διαθέτουν πληροφορία για τοποθεσία είτε στον τίτλο, είτε στα δεδομένα που το συνοδεύουν (περιγραφή, σχόλια κ.ά.). Στη συνέχεια, δημιουργήσαμε ένα λογισμικό που συνέδεσε τις τοποθεσίες με τις γεωγραφικές τους συντεταγμένες, ψάχνοντας για αυτές στο διαδίκτυο. Έχοντας, στο τέλος, μία βάση δεδομένων που έχει για κάθε βίντεο του YouTube τις τοποθεσίες με τις οποίες συνδέεται και τις συντεταγμένες αυτών, μπορέσαμε να χαρτογραφήσουμε την πληροφορία αυτή για τα βίντεο που αναζητούμε κάνοντας επερωτήσεις στη βάση μας.

Λέξεις-Κλειδιά: Χαρτογράφηση, YouTube, Βίντεο, Αναζήτηση τοποθεσίας, Αναζήτηση ημερομηνίας

Επιβλέπων: Δημήτριος Γουνόπουλος, Αναπληρωτής Καθηγητής

1. Εισαγωγή

Ο πρωταρχικός στόχος της εργασίας μας είναι να καταφέρουμε μέσα από τα δεδομένα που παρέχονται μαζί με κάθε βίντεο στη βάση δεδομένων του YouTube (http://www.youtube.com), να συνδέσουμε, αν γίνεται, κάποια βίντεο με τοποθεσίες όπως πόλεις, χώρες, αξιοθέατα, στάδια και γενικότερα οποιοδήποτε διακριτό είδος τοποθεσίας.

Στην παρούσα έκδοση του YouTube, ο χρήστης μπορεί να κάνει μία περιορισμένη θα λέγαμε αναζήτηση, μιας και αυτή αρκείται στο να βρίσκει λέξεις κλειδιά στους τίτλους των βίντεο και να παρουσιάζει τα αποτελέσματα. Πολλές φορές, όμως, ο χρήστης δεν ενδιαφέρεται τόσο για τις λέξεις κλειδιά, όσο για τα μεταδεδομένα που συνοδεύουν τα βίντεο με τις λέξεις αυτές. Πιο συγκεκριμένα, πολλές φορές ένας χρήστης ενδιαφέρεται για το πότε και πού έλαβε χώρα κάποιο γεγονός. Με τα τωρινά δεδομένα, για να αναζητήσει κάτι τέτοιο στο YouTube, ο μοναδικός τρόπος είναι να βάλει ως λέξεις κλειδιά το γεγονός που θέλει και να ψάξει ένα ένα τα βίντεο στα αποτελέσματα για να βρει την πληροφορία της ημερομηνίας και τοποθεσίας, αφού πολλές φορές τέτοιες πληροφορίες βρίσκονται στα δεδομένα κειμένου που συνοδεύουν τα βίντεο.

Δίνουμε τρία παραδείγματα από διαφορετικές κατηγορίες βίντεο, όπου παρουσιάζεται το ζήτημα αυτό:

Παράδειγμα 1: Αναζήτηση με λέξεις κλειδιά «data mining conference»

Στην αναζήτηση αυτή, το YouTube μας επιστρέφει 236 βίντεο με τις λέξεις κλειδί. Στα αποτελέσματα αυτά βρίσκουμε και κάποια βίντεο, όπως τα «Attend M2010 Data Mining Conference» και «Data Mining For Accounting Part 3», στα δεδομένα των οποίων βλέπουμε ότι τα συνέδρια αυτά έλαβαν χώρα στο Las Vegas και στο San Diego αντίστοιχα.

Παράδειγμα 2: Αναζήτηση με λέξεις κλειδιά «Olympic games»

Στην αναζήτηση αυτή, έχουμε 38.500 βίντεο στα αποτελέσματα. Σε πολλά βίντεο όπως στα «Sydney 2000 Olympic Games-4x100 Freestyle Relay-Australia Highlight», «1984 Los Angeles Olympic Games-Opening Ceremony» και «Opening Ceremony Olympic Games 2004 Athens» τόσο στους τίτλους, όσο και στα δεδομένα τους, βρίσκουμε πληροφορία χρονική και για την τοποθεσία με την οποία σχετίζονται.

Παράδειγμα 3: Αναζήτηση με λέξεις κλειδιά «Dream Theater live»

Στην αναζήτηση αυτή, έχουμε 13.900 αποτελέσματα μέσα στα οποία υπάρχουν και τα βίντεο «Octavarium (2006 live in Seoul)» και «Dream Theater-The spirits carries on (live new york 2000)» στα οποία έχουμε χρονική και τοπική πληροφορία στα δεδομένα.

Στα τρία αυτά παραδείγματα από τον επιστημονικό τομέα, τα αθλητικά και τη μουσική βλέπουμε ότι υπάρχουν χρονικές και γεωγραφικές αναφορές στα δεδομένα που συνοδεύουν τα βίντεο στην βάση του YouTube και συνεπώς είναι εφικτό να μπορέσουμε να συνδέσουμε τα βίντεο από τα αποτελέσματα για κάποιες λέξεις κλειδιά με χρονική και γεωγραφική πληροφορία. Η δημιουργία αυτής της πιο σύνθετης αναζήτησης και σύνδεσης των βίντεο με γεωγραφική και χρονική πληροφορία αποτελεί τον πυρήνα της εργασίας αυτής.

Στα πλαίσια της διαδικασίας αυτής, αρχικά προσπαθήσαμε να πετύχουμε τη σύνδεση αυτή με κάποια λογισμικά που χρησιμοποιούνται για ανάλυση κειμένου και στη συνέχεια, δημιουργήσαμε δικό μας λογισμικό για να αντιμετωπίσουμε τα προβλήματα που συναντήσαμε.

Αφού καταφέραμε να κάνουμε τη σύνδεση αυτή και μάλιστα σε ποσοστά που φτάνουν το 63% σε ορισμένες κατηγορίες βίντεο, δημιουργήσαμε λογισμικό που ψάχνει και συνδέει τις διάφορες τοποθεσίες με τις γεωγραφικές τους συντεταγμένες χρησιμοποιώντας δεδομένα από τον ιστοχώρο http://en.wikipedia.org και στη συνέχεια κάνοντας επερωτήσεις στη νέα βάση δεδομένων που έχουμε δημιουργήσει, κατασκευάζεται για κάθε επερώτηση ένα αρχείο javascript το οποίο οπτικοποιεί τα αποτελέσματα χρησιμοποιώντας το google maps api (http://code.google.com/apis/maps/index.html).

Με τον τρόπο αυτό, ο χρήστης μπορεί να κάνει επερωτήσεις προς τη βάση μας - για οτιδήποτε τον ενδιαφέρει και πιστεύει πως μπορεί να το βρει στην υπηρεσία του YouTube και συνδέεται με κάποιες τοποθεσίες - και χρησιμοποιώντας το λογισμικό μας να βλέπει όλα τα βίντεο της βάσης δεδομένων του YouTube που περιείχαν δεδομένα-τοποθεσίες και σχετίζονται με τα κλειδιά που αναζητά ο χρήστης, και πώς αυτά τοποθετούνται στον παγκόσμιο χάρτη.

Τέλος, ως επέκταση του λογισμικού, προσθέσαμε τη δυνατότητα αναζήτησης και ημερομηνιών, γεγονός που μας παρέχει τη δυνατότητα να βλέπουμε πώς κατανέμονται στον παγκόσμιο χάρτη όλα τα βίντεο της βάσης δεδομένων του YouTube που σχετίζονται με την ημερομηνία που αναζητούμε.

2. Αναζήτηση κατάλληλου λογισμικού

2.1 Παρουσίαση λογισμικών

Το πρώτο στάδιο της εργασίας μας ήταν να κάνουμε αναζήτηση τοποθεσιών μέσα στα δεδομένα που συνοδεύουν κάθε βίντεο. Για το σκοπό αυτό αρχικά δοκιμάσαμε να χρησιμοποιήσουμε συντακτικούς αναλυτές.

— The Stanford Parser [1]

Ο αναλυτής αυτός χρησιμοποιείται για να κάνει συντακτική ανάλυση της δομής κάθε πρότασης του κειμένου και να παράγει στατιστικά για τις λέξεις που χρησιμοποιούνται.

— Illinois Named Entity Tagger [6]

Το λογισμικό αυτό δεν παρουσιάζει τη συντακτική ανάλυση, αλλά στηρίζεται σε αυτή και σε κάποιες λίστες από τοποθεσίες που έχει στη διάθεσή του για να χαρακτηρίζει τις λέξεις ενός κειμένου σε διάφορες κατηγορίες (people/organizations/locations/miscellaneous). Το λογισμικό αυτό μπορεί να εκπαιδευτεί σε κάποια κείμενα για να παράγει καλύτερα αποτελέσματα, ανάλογα με τον χώρο από τον οποίο προέρχεται το κείμενο που θέλουμε να αναλύσουμε.

2.2 Λόγοι που οδήγησαν στη δημιουργία νέου λογισμικού

Στη δική μας περίπτωση, η χρήση των λογισμικών αυτών απορρίφθηκε για δυο λόγους. Πρώτον, παρατηρήθηκε ότι τα δεδομένα μας πολλές φορές δεν έχουν σωστή σύνταξη, λόγω του ότι χρησιμοποιούνται πολλές συντομογραφίες και γλώσσα argo και επίσης, διότι πολλές φορές οι χρήστες που δημιουργούν τα δεδομένα δεν γνωρίζουν καλά την αγγλική γλώσσα. Ο δεύτερος λόγος είναι ότι τα δεδομένα μας έχουν τεράστιο όγκο με αποτέλεσμα να αργεί σε απαγορευτικό βαθμό η συντακτική τους ανάλυση.

Ένα παράδειγμα χρήσης του δεύτερου προγράμματος που παρουσιάσαμε, φανερώνει το πρόβλημα που υπάρχει και αιτιολογεί την απόρριψη των λογισμικών αυτών για το δικό μας πρόβλημα.

Το παρακάτω κείμενο είναι η έξοδος του δεύτερου λογισμικού για ένα κομμάτι της βάσης μας, που αφορά σε ένα μόνο βίντεο:

«mn0frurrbvg | 1296601143000 5 2 | hatton tv - anthony crolla on british title challenge v john watson on feb 12th | shows | lightweight boxing british title john watson anthony crolla **liverpool** | Anthony million dollar crolla says it only took him ten seconds to decide he wants to fight john watson for the vacant british lightweight title at **liverpool** olympia on 12th february. full coverage and reaction can be found on http://www.hattonboxing.tv»

Όπως βλέπουμε υπάρχει δύο φορές η αναφορά στην πόλη Liverpool, αλλά λόγω της μη σωστής χρήσης της γραμματικής και του συντακτικού της αγγλικής γλώσσας, το πρόγραμμα δεν μπορεί να εντοπίσει τις αναφορές σε τοποθεσία. Λόγω του προβλήματος αυτού, δηλαδή ότι οι χρήστες του YouTube δεν χρησιμοποιούν πάντα σωστή σύνταξη και γραμματική, οδηγηθήκαμε στη δημιουργία νέου λογισμικού για να κάνουμε την αναζήτηση τοποθεσιών στα δεδομένα μας.

3. Δημιουργία και παρουσίαση νέου λογισμικού

3.1 Δημιουργία λιστών τοποθεσιών με συντεταγμένες

Για να μπορέσουμε να κάνουμε αναζήτηση τοποθεσιών στη βάση δεδομένων μας, πρέπει πρώτα να κατασκευάσουμε τις λίστες με τις τοποθεσίες που θα αναζητούμε. Στα πλαίσια της εργασίας αυτής, δημιουργήσαμε τέσσερις λίστες με τοποθεσίες που είναι οι εξής: χώρες, πόλεις, στάδια, αξιοθέατα. Τις λίστες αυτές τις κατασκευάσαμε από wikilists και από στατιστικές μελέτες για πληθυσμό πόλεων.

Η λίστα χωρών περιλαμβάνει όλες τις χώρες του κόσμου. Η λίστα των πόλεων περιλαμβάνει όλες τις πόλεις με πληθυσμό μεγαλύτερο από 100.000 κατοίκους. Η λίστα των σταδίων περιλαμβάνει τα πιο γνωστά στάδια κάθε χώρας, όπως επίσης και η λίστα των αξιοθέατων.

Στη συνέχεια, κατασκευάσαμε ένα λογισμικό που συνδέει κάθε τοποθεσία από τις λίστες μας με τις γεωγραφικές της συντεταγμένες, χρησιμοποιώντας τα δεδομένα του ιστοχώρου http://en.wikipedia.org.

Το λογισμικό αυτό χρησιμοποιεί τη μηχανή αναζήτησης που παρέχει ο ιστοχώρος και αναζητά την κάθε τοποθεσία στα δεδομένα του ιστοχώρου. Αν βρει την τοποθεσία και υπάρχουν οι γεωγραφικές συντεταγμένες της στη σελίδα με τα δεδομένα της τοποθεσίας, τότε κάνει τη σύνδεση. Διαφορετικά, προσπαθεί να βρει στα δεδομένα της σελίδας της τοποθεσίας, αν υπάρχει αναφορά σε ποια πόλη βρίσκεται η τοποθεσία, οπότε συνδέει την τοποθεσία αυτή με τις γεωγραφικές συντεταγμένες της πόλης αυτής. Τέλος, αν δεν υπάρχει σελίδα για τη συγκεκριμένη τοποθεσία, τότε αναζητά στα αποτελέσματα της αναζήτησης του wikipedia για να βρει αν η τοποθεσία μας συνδέεται με κάποια πόλη και αν ναι, τότε συνδέει την τοποθεσία με τις γεωγραφικές συντεταγμένες της πόλης αυτής.

Για παράδειγμα αν αναζητούμε την πόλη «Athens», τότε οδηγούμαστε στη σελίδα http://en.wikipedia.org/wiki/Athens όπου υπάρχει η πληροφορία «coordinates 37.966667, 23.716667». Αν τώρα ψάχνουμε το στάδιο «Abomoslem Mashhad», τότε στα αποτελέσματα της αναζήτησης υπάρχει ο σύνδεσμος για τη σελίδα http://en.wikipedia.org/wiki/F.C._Aboomoslem, στην οποία αναφέρεται ότι το στάδιο βρίσκεται στην πόλη «Mashhad», στη σελίδα της οποίας (http://en.wikipedia.org/wiki/Mashhad) βρίσκουμε τις γεωγραφικές συντεταγμένες «36.3, 59.6».

Με αυτό τον τρόπο, άμεσα ή έμμεσα, βρίσκουμε τις γεωγραφικές συντεταγμένες όλων των τοποθεσιών που έχουμε στις λίστες μας, για να μπορούμε τελικά να οπτικοποιήσουμε τα αποτελέσματα της εργασίας μας χρησιμοποιώντας το google maps api.

Στο σημείο αυτό, πρέπει να τονίσουμε το γεγονός ότι το λογισμικό μας μπορεί να εμπλουτιστεί με περισσότερες λίστες και τοποθεσίες, απλά δημιουργώντας μία νέα λίστα με τις τοποθεσίες που μας ενδιαφέρουν και έτσι να γίνουν πια ακριβή τα αποτελέσματα που θα πάρουμε στην οπτικοποίηση. Για παράδειγμα, μπορούμε να χρησιμοποιήσουμε όλες τις πόλεις του κόσμου, απλά προσθέτοντας μία λίστα με τις πόλεις του κόσμου με πληθυσμό μικρότερο των 100.000 ανθρώπων.

3.2 Αλγόριθμος αναζήτησης τοποθεσιών

3.2.1 Πρώτη προσέγγιση

Η τεχνική που σκεφτήκαμε να υλοποιήσουμε για την αναζήτηση των τοποθεσιών στα δεδομένα μας ήταν και η πιο απλή, να ψάχνουμε δηλαδή μία μία τις λέξεις των δεδομένων και να τις ελέγχουμε με τις τοποθεσίες μας. Αν βρεθεί μία ταύτιση, τότε συνδέουμε την τοποθεσία με τις γεωγραφικές της συντεταγμένες, με το α ναγνωριστικό του βίντεο στα δεδομένα του οποίου βρέθηκε η τοποθεσία.

Η ιδέα φαίνεται απλή, είναι όμως και η μόνη δυνατή επιλογή μας, αφού λόγω της φύσης των δεδομένων μας - που δεν υπάρχει σωστή σύνταξη και πολλές φορές σημεία στίξης, δεν τηρούνται γραμματικοί κανόνες και χρησιμοποιούνται πολλές συντομογραφίες και argo γλώσσα - θα ήταν πρακτικά αδύνατο να κάνουμε αναζήτηση με χρήση συντακτικών αναλυτών.

Αλγόριθμος σε ψευδοκώδικα.

```
create location lists
build indexes on the location lists
open data base file
while !EOF
    if word matches video ID then
        create new record in output
    else
        create buffer with the current video's data
        for each word in buffer
            if word appears in an index then
                write the location to the current record
            else if word mathes date format
                append date to the list of dates for current record
            end if
        end of for loop
        write list of dates to current record
    end if
end of while loop
```

Εικόνα 1: Αλγόριθμος προγράμματος σε ψευδοκώδικα

Η πρώτη υλοποίηση της τεχνικής αυτής ήταν και η πιο απλή προσέγγιση. Δημιουργήσαμε ένα πρόγραμμα που σειριακά ελέγχει ένα, ένα τα βίντεο της βάσης μας και δημιουργεί μια νέα εγγραφή σε ένα αρχείο για κάθε βίντεο. Αυτό επιτυγχάνεται σκανάροντας το αρχείο μας και όταν βρούμε αναγνωριστικό βίντεο δημιουργούμε μια νέα εγγραφή, αλλιώς δημιουργούμε έναν ενταμιευτή με τα δεδομένα του βίντεο και εξετάζουμε αν η κάθε λέξη σε αυτόν, ταυτοποιείται με κάποια τοποθεσία από τις λίστες μας. Αν ναι, τοποθετεί την τοποθεσία αυτή μαζί με τις γεωγραφικές της συντεταγμένες στην εγγραφή του βίντεο που εξετάζουμε. Όταν τελειώσει ο ενταμιευτής, προχωράμε στο επόμενο βίντεο της βάσης δεδομένων. Στην πρώτη αυτή προσέγγιση, ο έλεγχος για ταυτοποίηση μιας λέξης με τις τοποθεσίες στις λίστες μας, γίνεται σειριακά. Για να δείξουμε πως λειτουργεί ο αλγόριθμος, παραθέτουμε το παρακάτω κομμάτι της αρχικής μας βάσης και το αποτέλεσμα του αλγόριθμου για το υποσύνολο αυτό.

Όπως αναφέρουμε στην εισαγωγή, ο αλγόριθμος χρησιμοποιείται και για αναζήτηση χρονικής πληροφορίας. Το κομμάτι αυτό λειτουργεί με τον ίδιο τρόπο όπως και η αναζήτηση τοποθεσιών, με μια μικρή τροποποίηση. Κατά το σκανάρισμα της βάσης μας για αναζήτηση τοποθεσιών, ελέγχουμε για να βρούμε χρονική πληροφορία, είτε αυτή είναι σε μορφή αριθμών (π.χ. DD/MM/YY, MM/DD/YY κ.α.), είτε είναι με την αναλυτική μορφή (π.χ. 15 Aug 2010, September 13 2009 κ.α). Γενικά, ελέγχεται ένα σύνολο από τρόπους αναγραφής μιας ημερομηνίας που συναντώνται συχνότερα. Όταν βρεθεί στο κείμενο ενός βίντεο μία ημερομηνία, τότε αυτή την κρατάμε σε μία λίστα. Μόλις τελειώσουμε με τα δεδομένα του βίντεο, αυτό που κάνουμε είναι να προσαρτούμε τη λίστα με τις ημερομηνίες που βρήκαμε να συνδέονται με αυτό, στον τίτλο του βίντεο. Με αυτό τον τρόπο, όπως θα δούμε παρακάτω, μπορούμε να δημιουργούμε επερωτήσεις όχι με λέξεις κλειδιά, αλλά με ημερομηνίες.

Ο αλγόριθμος λειτουργεί και έχουμε σαν αποτέλεσμα τη νέα βάση δεδομένων, που αποτελείται από το αναγνωριστικό κάθε βίντεο μαζί με τις τοποθεσίες που βρέθηκαν στα δεδομένα τους. Μπορεί να φαίνεται ότι η υλοποίηση δεν είναι αποδοτική, αρκεί όμως να σκεφτούμε ότι ο αλγόριθμος αυτός αρκεί να τρέξει μία φορά στα αρχικά μας δεδομένα και μετά έχουμε τη νέα βάση δεδομένων μας που θα χρησιμοποιηθεί για επερωτήσεις, οπότε η απόδοση ισοσταθμίζεται με το γεγονός ότι έχουμε μία μόνο επανάληψη. Όμως, το γεγονός ότι η βάση δεδομένων του ιστοχώρου του YouTube είναι στην πραγματικότητα τεράστια σε μέγεθος (πολύ παραπάνω από τα 1.500.000 βίντεο που χρησιμοποιούμε εμείς στην εργασία μας), μας οδηγεί στο να χρησιμοποιήσουμε κάποιες τεχνικές βελτιστοποίησης απόδοσης του χρόνου εκτέλεσης του βασικού αλγόριθμου που μόλις παρουσιάσαμε.

3.2.2 Επιτάχυνση αλγορίθμου

Το βασικό κομμάτι του αλγόριθμου, ήταν όπως προαναφέραμε, η προσπάθεια για ταυτοποίηση κάθε λέξης των δεδομένων των βίντεο με τις τοποθεσίες στις λίστες μας. Για τη βελτίωση της ταχύτητας του αλγόριθμού μας, χρησιμοποιήσαμε δύο τεχνικές, τα νήματα και τα ευρετήρια.

Οι περισσότεροι επεξεργαστές στις μέρες μας υποστηρίζουν πολυνηματικές εφαρμογές. Για το λόγο αυτό, το πρώτο βήμα για την επιτάχυνση του αλγόριθμου ήταν να τον υλοποιήσουμε χρησιμοποιώντας νήματα. Η αλλαγή αυτή πετυχαίνεται με το να δημιουργούμε ένα νήμα για κάθε είδος λίστας τοποθεσιών που διαθέτουμε (χώρες, πόλεις, στάδια, αξιοθέατα) και στη συνέχεια να βάζουμε κάθε νήμα να τρέχει όλη τη βάση δεδομένων και να προσπαθεί να ταυτοποιήσει μόνο με τις τοποθεσίες της συγκεκριμένης λίστας.

Με τον τρόπο αυτό κάθε νήμα, όταν προσπαθεί να ταυτοποιήσει μία λέξη ως τοποθεσία, αρκείται στο να ελέγξει σειριακά τη μοναδική λίστα που του έχει ανατεθεί. Αναφέρουμε εδώ, ότι αν επιλέξουμε να εμπλουτίσουμε τις λίστες μας με άλλου τύπου τοποθεσίες, αρκεί να δημιουργήσουμε τα αντίστοιχα νήματα που θα επεξεργάζονται τις νέες αυτές λίστες, και έτσι δεν χρειάζεται να αλλάξουμε τίποτα στον αλγόριθμο. Τέλος, αφού τελειώσουν τα νήματα με την επεξεργασία, έχουμε πλέον τέσσερα αρχεία, όσα και οι λίστες μας, που περιέχουν τα αναγνωριστικά όλων των βίντεο μαζί με τις συνδέσεις με τοποθεσίες που μπορεί να έχουν από την κατηγορία τοποθεσιών που χρησιμοποιούσαν. Αρκεί να κάνουμε μία συγχώνευση και παίρνουμε τα πλήρη αποτελέσματα με τις συνδέσεις του κάθε βίντεο με όλες τις κατηγορίες τοποθεσιών. Η επέκταση αυτή του αλγόριθμου επέφερε βελτίωση της τάξης 50%, όσον αφορά στο χρόνο εκτέλεσης.

3.2.3 Προοπτικές βελτίωσης αλγορίθμου

Εκτός από την βελτίωση στην ποιότητα των αποτελεσμάτων που μπορεί να προκύψει, όπως έχουμε προαναφέρε,ι με τη δημιουργία νέων λιστών με τοποθεσίες, μία σημαντική ιδιότητα του αλγορίθμου που επιλέξαμε, είναι ότι μπορεί πολύ εύκολα να παραλληλοποιηθεί.

Ένα επιπλέον στάδιο για την περαιτέρω βελτίωση του αλγόριθμου λοιπόν, είναι να χωρίσουμε την αρχική βάση δεδομένων σε πολλές μικρότερες και στη συνέχεια να τρέχουμε σε κάθε έναν πυρήνα ή επεξεργαστή το βελτιωμένο αλγόριθμο που παρουσιάσαμε στην προηγούμενη ενότητα.

Με ένα πολύ απλό βήμα λοιπόν, βλέπουμε ότι μπορούμε να μετατρέψουμε τον αλγόριθμό μας σε κατανεμημένο αλγόριθμο με θεωρητικά άπειρες δυνατότητες στην κλιμάκωση, αφού το μόνο που πρέπει να κάνουμε είναι να χωρίζουμε σε περισσότερα κομμάτια τη βάση μας και να χρησιμοποιούμε περισσότερους επεξεργαστές. Το γεγονός αυτό, δηλαδή της εύκολης παραλληλοποίησης, είναι που κάνει τον αλγόριθμό μας, αν και απλό, πολύ αποδοτικό.

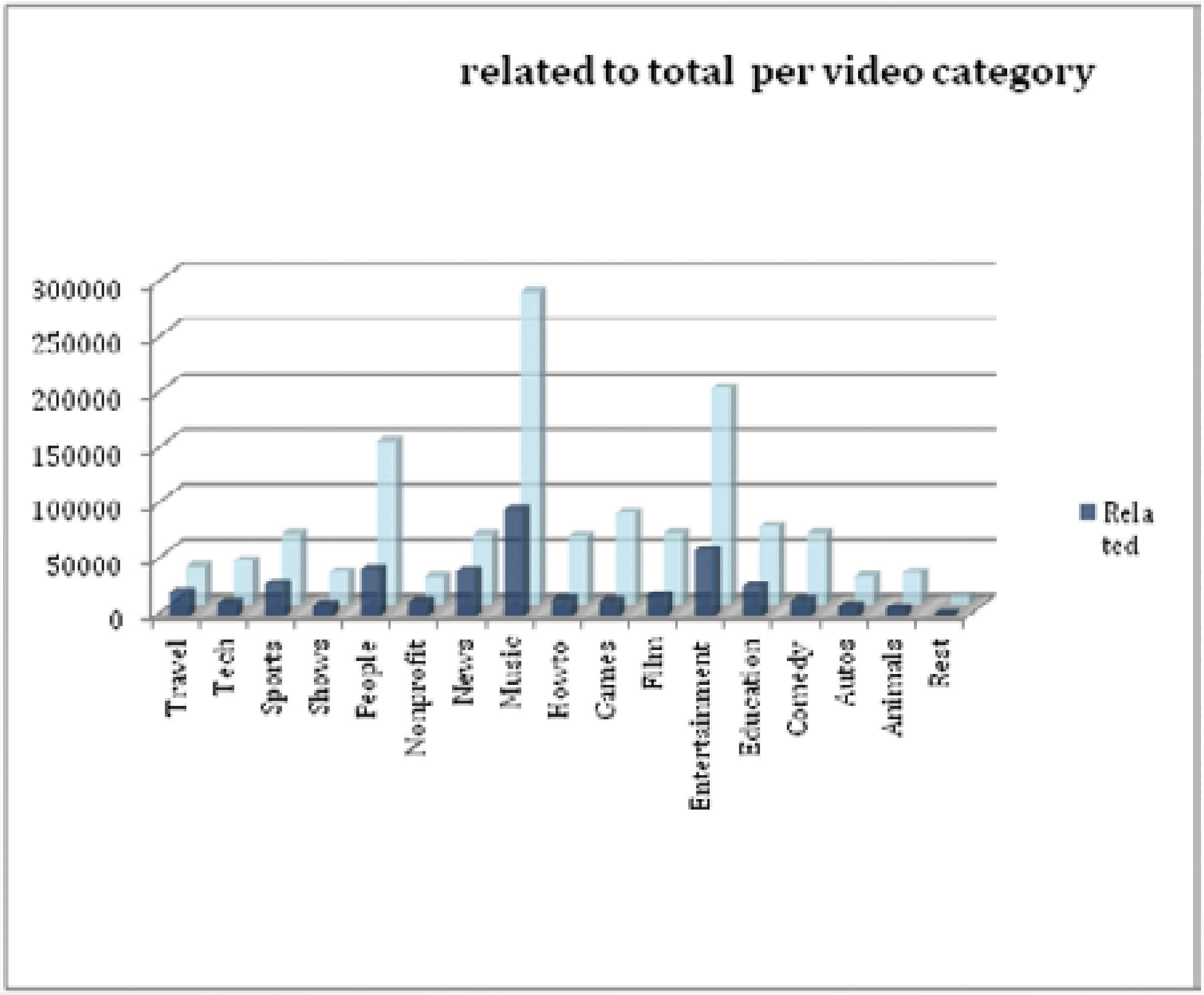
3.2.4 Παρουσίαση αποτελεσμάτων

Στο σημείο αυτό θα παρουσιάσουμε τα στατιστικά αποτελέσματα που πήραμε, χρησιμοποιώντας τον αλγόριθμο που υλοποιήσαμε πάνω στη βάση δεδομένων μας, η οποία περιέχει 1.275.434 βίντεο με τα δεδομένα τους. Ο αλγόριθμος υλοποιήθηκε και δοκιμάστηκε σε περιβάλλον MAC OSX 10.6.8 και τα χαρακτηριστικά του υπολογιστή είναι:

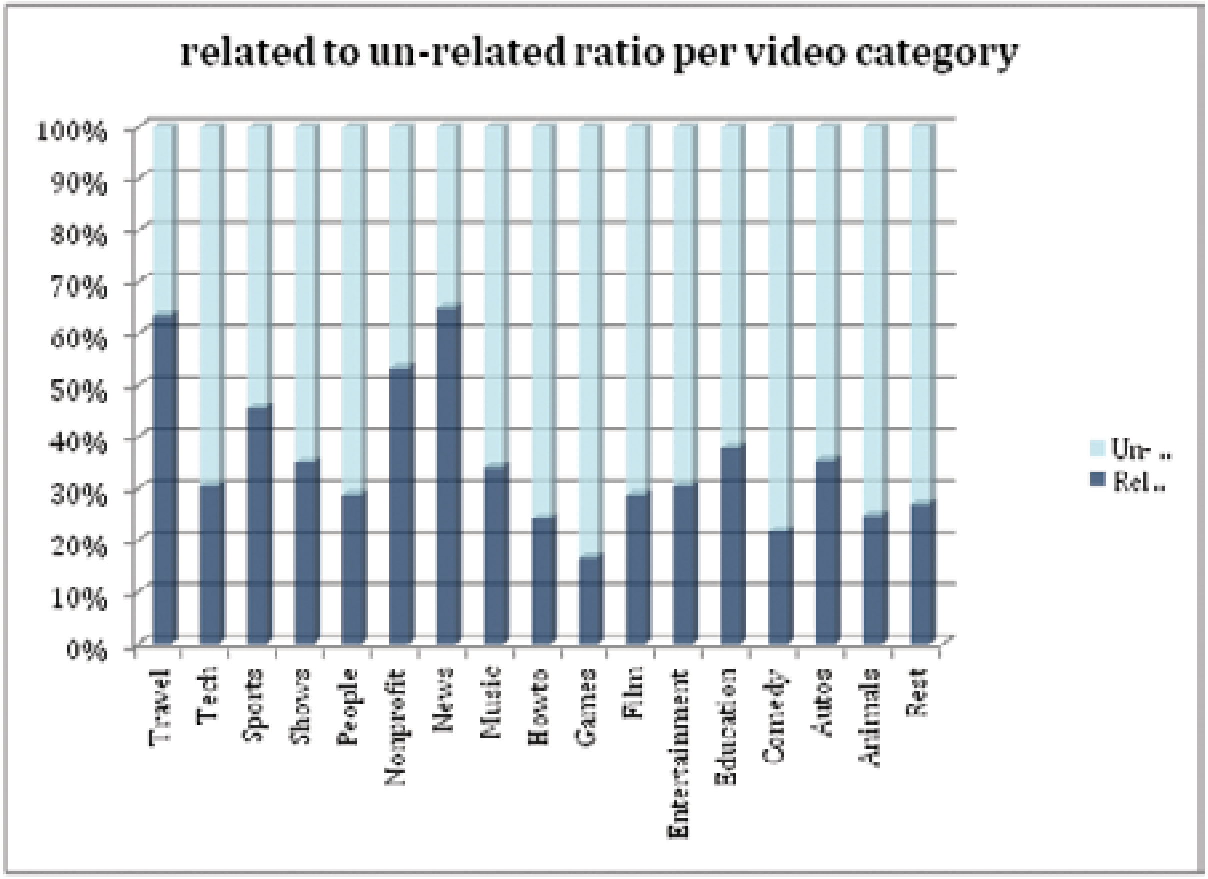
```
Processor      2.8 GHz Intel Core i7
Memory         4 GB 1067 MHz DDR3
Hard Disk      1 TB
```

Ο χρόνος εκτέλεσης του βελτιωμένου αλγόριθμου πάνω στη βάση μας ήταν 35′ 51′′.

Στη συνέχεια ακολουθούν οι συνδέσεις βίντεο με τουλάχιστον μία τοποθεσία, σε κάθε κατηγορία βίντεο χωριστά:



Σχήμα 1: Διάγραμμα συσχέτισης βίντεο με και χωρίς γεωγραφικές συσχετίσεις



Σχήμα 2: Διάγραμμα συσχέτισης βίντεο με και χωρίς γεωγραφικές συσχετίσεις, επί τοις εκατό

Όπως βλέπουμε στα διαγράμματα, τα ποσοστά σύνδεσης βίντεο με τουλάχιστον μία τοποθεσία κυμαίνονται από 15% μέχρι και 63%, ανάλογα με την κατηγορία στην οποία ανήκουν τα βίντεο. Τα ποσοστά αυτά είναι υψηλά σε σχέση με τα αναμενόμενα, γεγονός που μας οδηγεί στο συμπέρασμα ότι ένα μεγάλο ποσοστό, της τάξης του 30% όλων των βίντεο της βάσης μας, συσχετίζονται με τουλάχιστον μία τοποθεσία.

Αυτά τα αποτελέσματα είναι ενθαρρυντικά, στο να μπορούμε να πούμε ότι μία τέτοια εφαρμογή μπορεί να φανεί αρκετά χρήσιμη στην κοινότητα των χρηστών των υπηρεσιών του YouTube και για το λόγο αυτό προχωρήσαμε και υλοποιήσαμε την οπτικοποίηση των πιθανών επερωτήσεων στη βάση δεδομένων μας, χρησιμοποιώντας το google maps api.

3.3 Επερωτήσεις και προβολή αποτελεσμάτων

3.3.1 Δημιουργία επερωτήσεων

Έχοντας δημιουργήσει πλέον την καινούργια βάση δεδομένων με τα αναγνωριστικά των βίντεο και τις τοποθεσίες με τις οποίες συνδέονται, δημιουργούμε το τελικό αρχείο πάνω στο οποίο θα γίνονται οι επερωτήσεις, προσθέτοντας απλά σε κάθε εγγραφή τον τίτλο του βίντεο και την κατηγορία στην οποία ανήκει. Έτσι, έχουμε την τελική βάση δεδομένων μας που έχει την ακόλουθη μορφή:

```
ID: B29ju4b_HkE | TITLE:  Pantyhose party http://cocojambo7.
blogspot.com/    | CATEGORY:  People   | RELATIONS:
ID: EbM6-xKA8bo | TITLE:  Alphabet Soup     | CATEGORY:  Film   |
RELATIONS:    durham(T); 54.7761, -1.5733#
ID: DTKs2UTGmTo | TITLE:   Bo bia        | CATEGORY:   Howto    |
RELATIONS:    viet nam(C); 21.033, 105.85#
ID: pNi0Ixju75M | TITLE:  The 5 DeViLs - Le Grand Blonde ave une
chaussure noire      [13 february 2009]| CATEGORY:   Music    |
RELATIONS:     bologna(S); 44.5075, 11.35139#  buenos aires(T);
-34.60333,   -58.38167#livorno(T);   43.55,   10.317#bologna(T);
44.5075, 11.35139#modena(T); 44.65, 10.933#
```

Πάνω σε αυτό το αρχείο μπορούμε, εύκολα πλέον, να κάνουμε επερωτήσεις που να αφορούν στους τίτλους των βίντεο και να παίρνουμε σαν αποτέλεσμα τα βίντεο που έχουν στον τίτλο τους τα κλειδία αναζήτησης και τις τοποθεσίες με τις οποίες συνδέονται. Η λειτουργία που προσφέρουμε είναι παρόμοια με την υπηρεσία αναζήτησης που διαθέτει το YouTube, μιας και αυτή ψάχνει για τις λέξεις κλειδιά μόνο στους τίτλους των αρχείων. Η επέκταση, που προσφέρει η δική μας υλοποίηση, είναι ότι κάθε βίντεο που επιστρέφεται, συνοδεύεται από τις τοποθεσίες με τις οποίες έχει συνδεθεί και επιπλέον μπορούμε να κάνουμε και χρονική αναζήτηση.

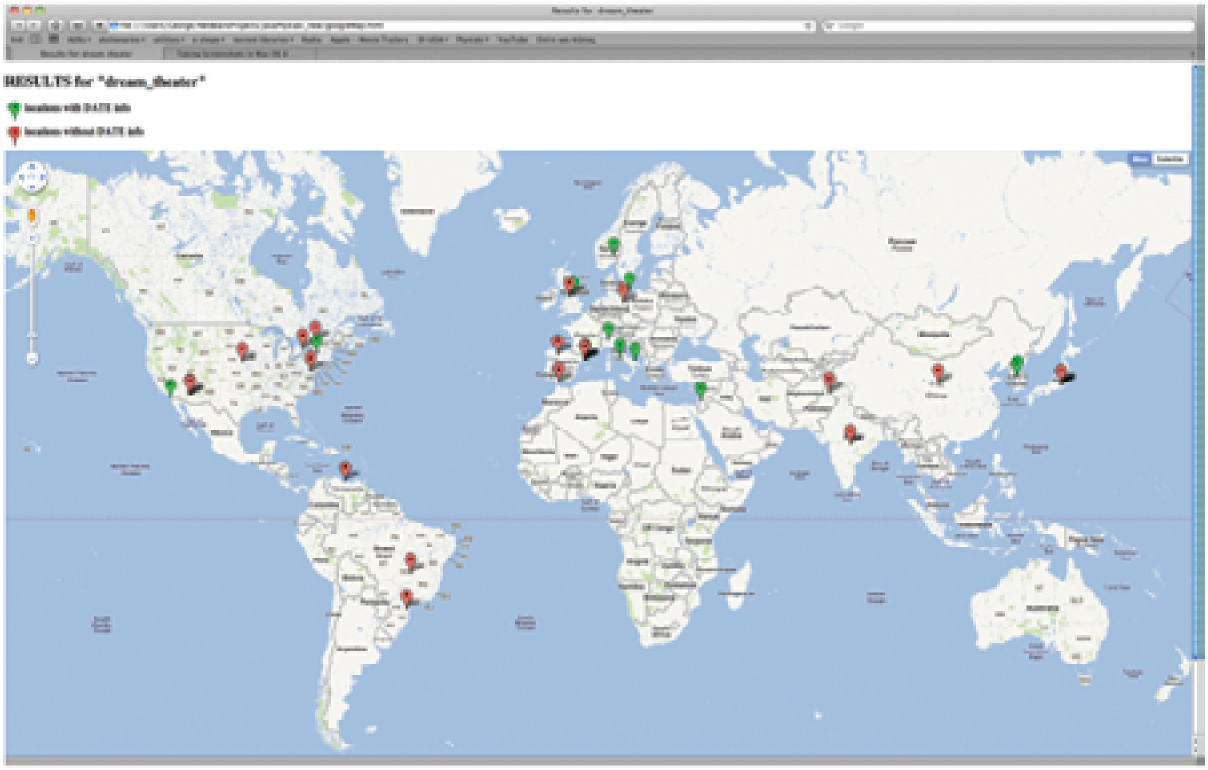
Κατά την υλοποίηση του αλγόριθμου, εκτός από τοποθεσίες, προσπαθήσαμε να εντοπίσουμε και ημερομηνίες στα δεδομένα του κάθε βίντεο (ελέγχονται οι περισσότερες γνωστές φόρμες για δήλωση ημερομηνιών, όπως π.χ. 12 Feb 2010, 12/02/10, 2/12/2010, February 12 2010). Με τον έλεχο αυτό, εκτός από τις τοποθεσίες, συνδέουμε κάθε τίτλο βίντεο και με τις ημερομηνίες με τις οποίες εμφανίζονται στα δεδομένα του. Όπως βλέπουμε στο τέταρτο παράδειγμα παραπάνω, ο τίτλος του βίντεο με κωδικό «pNiOIxju75M» έχει στο τέλος του μία λίστα που περιέχει την ημερομηνία «13 February 2009», η οποία είχε βρεθεί στα δεδομένα του βίντεο όπως περιγράψαμε. Αυτό το γεγονός, μας προσφέρει την δυνατότητα να μπορούμε να κάνουμε αναζήτηση, με βάση μία ημερομηνία και να παίρνουμε όλα τα βίντεο που συνδέονται με την ημερομηνία αυτή.

3.3.2 Οπτικοποίηση αποτελεσμάτων επερώτησης

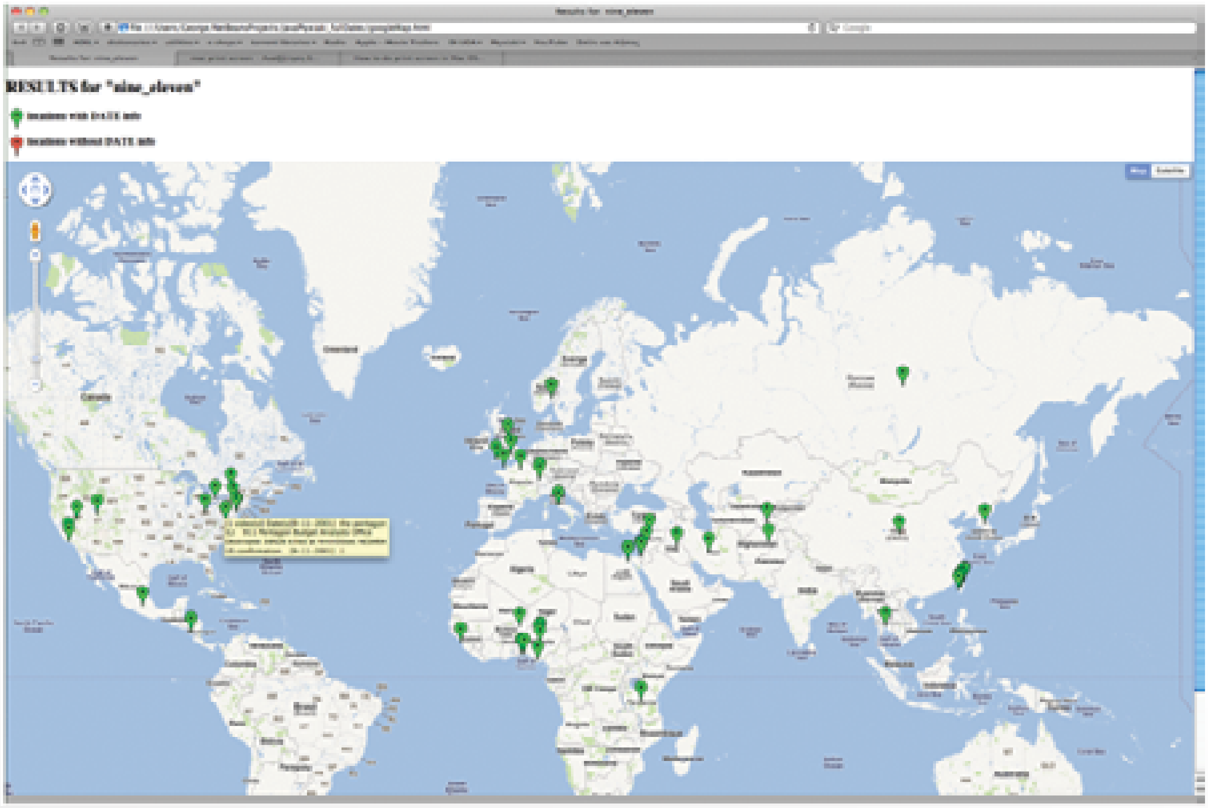
Το επόμενο βήμα στην παρουσίαση των αποτελεσμάτων είναι, εφόσον διαθέτουμε τις γεωγραφικές συντεταγμένες των τοποθεσιών με τις οποίες συνδέονται τα βίντεο, να οπτικοποιήσουμε τα αποτελέσματα μιας επερώτησης στον παγκόσμιο χάρτη. Το κομμάτι αυτό, αποτελεί και τον αρχικό στόχο της εργασίας, που ήταν να μπορούμε να παράγουμε οπτική απεικόνιση των βίντεο που μας ενδιαφέρουν, σε συσχέτιση με τις τοποθεσίες με τις οποίες συνδέονται.

Για να το επιτύχουμε αυτό, χρησιμοποιούμε το google maps api, ώστε να αναπαραστήσουμε τα βίντεο στον παγκόσμιο χάρτη, στα σημεία που αντιστοιχούν στις τοποθεσίες με τις οποίες συνδέονται. Κατασκευάσαμε, λοιπόν, ένα λογισμικό που παίρνει σαν είσοδο τα αποτελέσματα μιας επερώτησης και κατασκευάζει με βάση αυτά, ένα javascript το οποίο είναι σύμφωνο με τους κανόνες του google maps api και όταν το ανοίξουμε σε έναν φυλλομετρητή παίρνουμε ένα παγκόσμιο χάρτη με τα αποτελέσματα της επερώτησης. Οι ακίδες στο χάρτη έχουν δυο διαφορετικά χρώματα, πράσινο και κόκκινο. Οι πράσινες ακίδες αντιστοιχούν σε βίντεο με χρονική πληροφορία, ένω οι κόκκινες χωρίς. Ακόμα, όπως φαίνεται στην δεύτερη εικόνα παρακάτω, κάθε ακίδα περιέχει πληροφορίες που είναι οι εξής: πλήθος βίντεο που συνδέονται με την τοποθεσία, ημερομηνίες που συνδέονται με την τοποθεσία, όνομα της τοποθεσίας, τίτλοι των βίντεο που συνδέονται με την τοποθεσία.

Ακολουθούν δύο παραδείγματα με τα αποτελέσματα που πήραμε για μία επερώτηση, στην οποία αναζητούμε βίντεο με λέξεις κλειδιά «dream theater» και μια δεύτερη, που αναζητούμε βίντεο με κλειδί την ημερομηνία «09/11/2001» και έχουμε σαν αποτέλεσμα τους παρακάτω χάρτες:



Εικόνα 2: Χάρτης με αποτελέσματα για «dream theater»



Εικόνα 3: Χάρτης με αποτελέσματα για «09/11/2001»

4. Σύνοψη και συμπεράσματα

Στην εργασία αυτή, παρουσιάσαμε το πρόβλημα της επεξεργασίας μεγάλου όγκου δεδομένων για την εξαγωγή χρήσιμων πληροφοριών. Πιο συγκεκριμένα, προσπαθήσαμε να εξάγουμε από την πληθώρα δεδομένων που συνδέονται με ένα βίντεο στον Ιστοχώρο του YouTube, την πληροφορία που αφορά στο χρόνο και στον τόπο που εξελίσσεται το βίντεο, αν αυτό είναι δυνατό.

Χρησιμοποιήσαμε έναν αλγόριθμο αναζήτησης για να βρούμε τις τοποθεσίες και ημερομηνίες που υπάρχουν στα δεδομένα ενός βίντεο και στη συνέχεια, δημιουργήσαμε μια νέα βάση δεδομένων, στην οποία έχουμε πλέον κάθε βίντεο με το αναγνωριστικό του, τον τίτλο του, την κατηγορία στην οποία ανήκει και τις τοποθεσίες και ημερομηνίες με τις οποίες συνδέεται.

Κατόπιν, έχοντας τη νέα βάση δεδομένων, μπορούμε να κάνουμε επερωτήσεις με βάση λέξεις κλειδιά και ημερομηνίες, κάνοντας αναζήτηση στους τίτλους των βίντεο. Έχοντας τα αποτελέσματα κάθε επερώτησης, προχωρήσαμε στην οπτικοποίηση των αποτελεσμάτων αυτών με βάση τις τοποθεσίες με τις οποίες συνδέονται, χρησιμοποιώντας το google maps api.

Δημιουργήσαμε, με τον τρόπο αυτό, ένα λογισμικό που εμπλουτίζει την υπηρεσία αναζήτησης που παρέχει ο ιστοχώρος του YouTube, δίνοντας στο χρήστη αφενός τη δυνατότητα να αναζητά βίντεο και στα αποτελέσματα να μπορεί να δει και τις τοποθεσίες με τις οποίες συνδέονται τα βίντεο αυτά, και αφετέρου τη δυνατότητα αναζήτησης βίντεο, όχι με λέξεις κλειδιά που αφορούν στον τίτλο του, αλλά με βάση μία ημερομηνία.

Αναφορές

1. Dan Klein, Christopher D. Manning. Accurate Unlexicalized Parsing. Proceedings of the 41st Meeting of the Association for Computational Linguistics. 2003

2. David Crandall, Lars Backstrom, Daniel Huttenlocher and Jon Kleinberg. Mapping the World’s Photos, WWW 2009

3. David Nadeau, Satoshi Sekine. A survey of named entity recognition and classification. NERC 2007

4. George Toderici, Hrishikesh Aradhye, Marius Pasca, Luciano Sbaiz, Jay Yagnik. Finding

Meaning on YouTube: Tag Recommendation and Category Discovery. CVPR 2010

5. Klaus Zechner, Alex Waibel. A literature Survey on Information Extraction and Text Summarization. Directed Reading 1996.

6. Lev Ratinov, Dan Roth. Deisgn Challenges and Misconceptions in Named Entity Recognition. CoNLL 2009.

7. Raymond J. Mooney, Un Yong Nahm. Text Mining with Information Extraction. NSF 2005.

Σχεδίαση Αυτόνομου Ηλεκτρονικού Συστήματος Ελέγχου Φωτοδιόδων για Πειραματική Επαλήθευση Διαστημικής Πτήσης σε Σχηματισμό

Χρήστος Κ. Τσίγκανος
c.tsiganos@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Για την πειραματική μελέτη και επίδειξη στο εργαστήριο των δυνατοτήτων ενός Ηλιακού Στεμματογράφου, που αποτελείται από δύο διαστημόπλοια σε πτήση σε σχηματισμό (ASPIICS), στα πλαίσια του προγράμματος STARTIGER της ESA, απαιτείτο να σχεδιαστεί ένα αυτόνομο υποσύστημα ελέγχου φωτοδιόδων LED. Το σύστημα αυτό, έπρεπε να έχει την ικανότητα να υποστηρίζει δυναμικά, αρκετούς διαφορετικούς συνδυασμούς διατάξεων για διαφορετικά σενάρια χρήσης. Επίσης, αναπτύχθηκε μια τεχνική υψηλής συχνότητας, βασισμένη στη Διαμόρφωση Πλάτους Παλμού (PWM), λόγω της απαίτησης ύπαρξης δυνατότητας ακριβέστερης αυξομείωσης της έντασης του εξερχόμενου φωτός. Στην πειραματική αυτή διάταξη, συμπεριλήφθηκαν επίσης και άλλες χρήσιμες βοηθητικές λειτουργίες, όπως ο έλεγχος μηχανικών κλείστρων. Η υλοποίηση του υποσυστήματος, για χρήση σε καθαρό δωμάτιο, είναι βασισμένη σε ένα μικροελεγκτή και συμπληρώνεται με πλήρη, αποκλειστική διασύνδεση Ethernet στο σταθμό ελέγχου και ηλεκτρονικά μέρη σε μια πλακέτα κυκλωμάτων.

Λέξεις-Κλειδιά: Ευρωπαϊκός Οργανισμός Διαστήματος, PROBA-3, ASPIICS

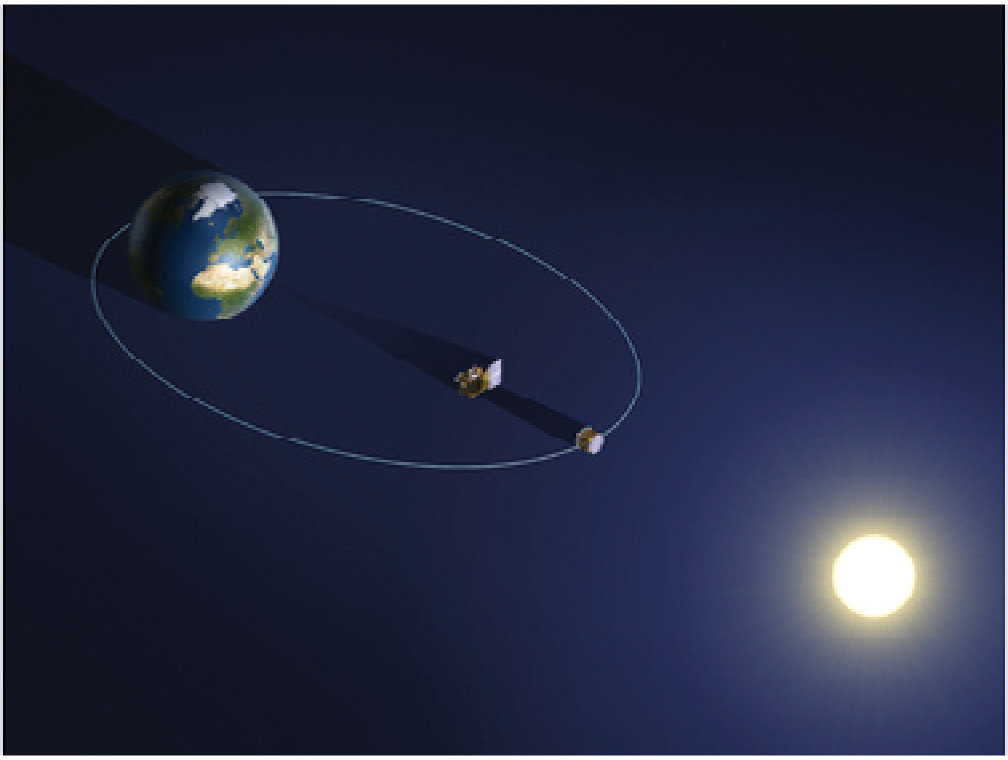
Επιβλέπων: Αντώνης Πασχάλης, Καθηγητής

1. Εισαγωγή

Τα τελευταία χρόνια, υπάρχει αυξανόμενο ενδιαφέρον για τη χρήση συνεργαζόμενων διαστημοπλοίων σε σχηματισμό. Αυτή η φιλοσοφία σχεδιασμού διαστημικών πτήσεων, προέκυψε από πληθώρα αποστολών και αποτελέσματα ερευνών ως ένα μέσο όχι μόνο μείωσης του κόστους και αύξησης της ευελιξίας σε διαστημικά προγράμματα, αλλά και για τα μοναδικά επιστημονικά πλεονεκτήματα που προσφέρει. Έτσι, υπάρχει ελπίδα να μειωθεί το μέγεθος και η πολυπλοκότητα μοναδιαίων διαστημικών αποστολών, υπέρ της χρήσης συστοιχίας μικρότερων και απλούστερων δορυφόρων, οι οποίοι συνεργάζονται μεταξύ τους για κάθε λειτουργία. Το σύστημα της πτήσης σε σχηματισμό είναι πλεονεκτικότερο, γιατί το μέγεθος του συνολικού οργάνου και ο προσανατολισμός του μπορούν να μεταβάλλονται ενόσω το σύστημα βρίσκεται σε τροχιά.

1.1 Η αποστολή PROBA-3

Η αποστολή PROBA-3 (PRoject for On-board Autonomy) θα προσφέρει μια ευκαιρία για την προτυποποίηση της μετρολογίας και των τεχνολογιών της πτήσης σε σχηματισμό, ενώ παράλληλα θα αναπτύξει και θα ελέγξει δυναμικά τη στρατηγική καθοδήγηση, την πλοήγηση και τους αλγόριθμους ελέγχου, που είναι αναγκαίοι για την πτήση σε σχηματισμό σε μελλοντικές διαστημικές επιστημονικές αποστολές. Θα μελετήσει, επίσης, ένα ευρύ φάσμα δυνατοτήτων και στρατηγικών ελέγχου, και θα αναπτύξει τα αντίστοιχα εργαλεία των σχετικών υποδομών. Το ζεύγος διαστημοπλοίων θα ακολουθεί μια αρκετά ελλειπτική τροχιά γύρω από τη Γη, ένα μέρος της οποίας θα είναι η οικονομικά ακριβή πτήση σε σχηματισμό, όπου είναι δυνατές οι παρατηρήσεις, και οι υπόλοιπες περίοδοι θα είναι ελεύθερης πτήσης. Τα δύο διαστημόπλοια θα πετούν σε σχηματισμό σε απόσταση 150m μεταξύ τους, με ακρίβεια μικρότερη του χιλιοστού. Το ένα διαστημόπλοιο θα προκαλεί τεχνητή έκλειψη του Ήλιου στο σύστημα αναφοράς του άλλου, και έτσι θα είναι δυνατή η παρατήρηση του Ηλιακού Στέμματος. Η υλοποίηση αυτής της αποστολής θα ανοίξει νέους ορίζοντες στη διαστημική τεχνολογία.



Σχήμα 1: Η τροχιά του PROBA-3.

1.2 Το επιστημονικό όργανο του PROBA-3: Ο Ηλιακός Στεμματογράφος ASPIICS *(Association de Satellites Pour l' Imagerie et l' Interferometrie de la Couronne Solaire)*

Το Ηλιακό Στέμμα, η εξωτερική ατμόσφαιρα του Ήλιου, είναι εξαιρετικά αραιό, αφού μόνο ένα στα ένα εκατομμύριο φωτόνια της φωτόσφαιρας σκεδάζονται απο το πλάσμα του, και εξαιρετικά θερμό με θερμοκρασίες μερικών εκατομμυρίων βαθμών. Αποτέλεσμα αυτών των υψηλών θερμοκρασιών είναι και η υπερηχητική εκτόνωση του πλάσματος του Στέμματος προς τον μεσοπλανητικό χώρο, μέσα από τις λεγόμενες στεμματικές οπές με τη μορφή του ηλιακού ανέμου Parker, που φθάνει με ταχύτητες 3 εκατομμυρίων χιλιομέτρων την ώρα στη Γη. Τα σχετικά ήσυχα φαινόμενα της ηλιακής φωτόσφαιρας (όπως οι ηλιακές κηλίδες), αλλά και τα εκρηκτικά και εξαιρετικά βίαια φαινόμενα που συνοδεύουν την εμφάνιση του μαγνητικού πεδίου στην ηλιακή ατμόσφαιρα, συνιστούν την ηλιακή δραστηριότητα.

Οι κύριοι επιστημονικοί στόχοι του προγράμματος PROBA-3 είναι η κατανόηση αφενός των φυσικών διεργασιών που συμβαίνουν στο ήρεμο Στέμμα και αφετέρου, των φυσικών διεργασιών που οδηγούν στις ισχυρές Στεμματικές Εκτινάξεις Μάζας και συνακόλουθα στη δημιουργία του Διαστημικού Καιρού, με όλες τις συνέπειές του για τη Γη και το Διάστημα. Η μελέτη του στέμματος με το ASPIICS θα απαντήσει σε συγκεκριμένες ερωτήσεις όπως:

- Πώς προσδιορίζεται η δομής της ηλιακής ατμόσφαιρας σε πολύ μικρή κλίμακα;
- Ποιος είναι ο μηχανισμός θέρμανσης του Ηλιακού Στέμματος και ο ρόλος των κυμάτων σε αυτό;
- Ποια είναι η προέλευση και φυσικές διαδικασίες που συνεισφέρουν στην επιτάχυνση του αργού και γρήγορου ηλιακού ανέμου;
- Από ποιές δομές του χαμηλού Στέμματος σχηματίζονται, επιταχύνονται και τελικά εκρήγνυνται οι Στεμματικές Εκτινάξεις Μάζας;
- Ποιά η σχέση των Στεμματικών Εκτινάξεων Μάζας με άλλες ενεργές περιοχές του ήλιου, όπως οι Εκλάμψεις και οι Προεξοχές;
- Πώς οι Στεμματικές Εκτινάξεις Μάζας δημιουργούν ισχυρά ωστικά κύματα στο χαμηλό Στέμμα;

Ο Στεμματογράφος ASPIICS που θα τοποθετηθεί στο PROBA 3 στοχεύει στο να παρατηρήσει το Ηλιακό Στέμμα στην απόσταση (1.03 - 3) ηλιακών ακτίνων από το ηλιακό κέντρο, με διακριτική ικανότητα εφάμιλλη αυτής που παρατηρούμε από τη Γη κατά τη διάρκεια των ολικών ηλιακών εκλείψεων.

2. Το πρόγραμμα STARTIGER της ESA

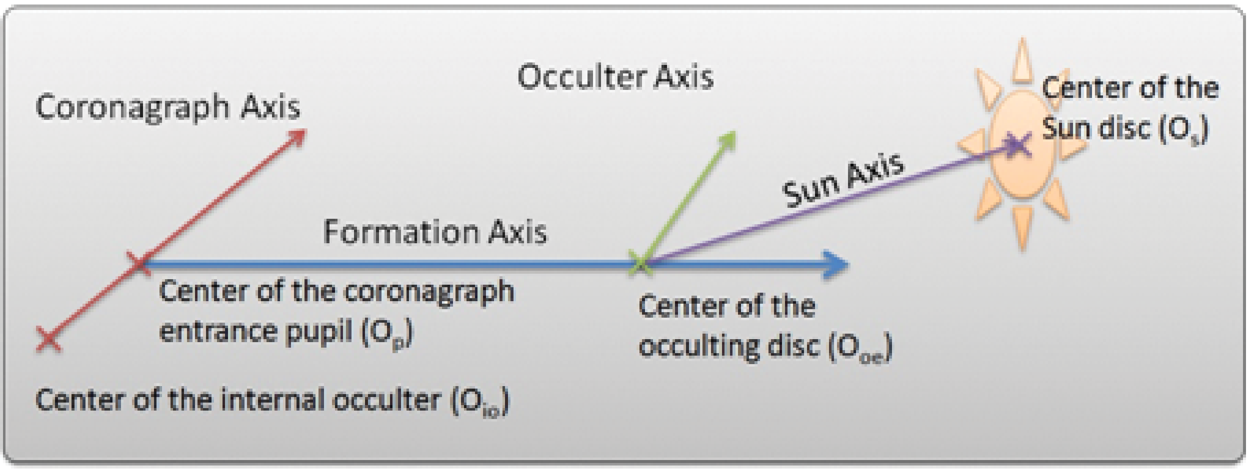
Το πρόγραμμα STARTIGER της ESA αποσκοπεί σε μια πρώτη επίδειξη της τεχνολογίας σε εργαστηριακές συνθήκες για όλα τα λεπτομερή συστήματα συγχρονισμού και μετρολογίας, των οπτικών υποσυστημάτων του ASPIICS.

2.1 Ευθυγράμμιση της πτήσης σε σχηματισμό

Η βασική ιδέα του στεμματογράφου, είναι να αποτελείται από ένα μεγάλο, εννοιολογικά ενιαίο, όργανο, που θα έχει ως συνιστώσες του δύο διαστημόπλοια, που θα βρίσκονται σε πτήση σε σχηματισμό. Ένα τέτοιο επιστημονικό όργανο πρέπει να επιτυγχάνει αφενός απόλυτη ευθυγράμμιση του σχηματισμού των δύο διαστημοπλοίων με τον ήλιο και αφετέρου, σχετική ευθυγράμμιση των δύο διαστημοπλοίων μεταξύ τους.

Σημειώνεται οτι αυτή η οργάνωση (Σχήμα 2), είναι πρωτόγνωρη στο χώρο της κατασκευής επιστημονικών οργάνων. Βεβαίως, το σύστημα των δύο δορυφόρων θα έχει ήδη το δικό του σύστημα μετρολογίας που θα παρακολουθεί το σχηματισμό (επικοινωνίες RF, οπτικοί αισθητήρες), από την πρώτη φάση του PROBA-3. Όμως, η αρχική, και περιοδική βαθμονόμηση (calibration) του συστήματος θέσης, όπως και ο λεπτομερέστερος καθορισμός του, θα γίνεται από το ίδιο το όργανο. Έτσι, στόχοι του STARTIGER επιπλέον, είναι:

- Ο προσδιορισμός, η ανάπτυξη και οι δοκιμές και η ορθή επιβεβαίωση των υποσυστημάτων ευθυγράμμισης και στόχευσης.
- Οι δοκιμές και η συνολική ορθή λειτουργία ολόκληρου του συστήματος εξομοίωσης σε μορφή breadboard, καθώς και εσωτερική ευθυγράμμιση του οργάνου.



Σχήμα 2: Οι άξονες ευθυγράμμισης του οργάνου:
Κόκκινο χρώμα: άξονας τηλεσκοπίου στεμματογράφου, Πράσινο χρώμα: άξονας εκλείπτη (occulter),
Μπλέ χρώμα: άξονας σχηματισμού, μεταξύ των δύο διαστημοπλοίων,
Μώβ χρώμα: άξονας στόχευσης ως προς τον ήλιο.

Η ευθυγράμμιση του οργάνου γίνεται αποκλειστικά από το ίδιο, και η επιβεβαίωση αυτής επιτυγχάνεται κβαντοποιώντας κατάλληλα τη θέση του εκλείπτη στο οπτικό πεδίο (Field of View, FoV) του στεμματογράφου. Η προτεινόμενη λύση αυτού, ο Αισθητήρας Θέσης Εκλείπτη (Occulter Position Sensor, OPS), βασίζεται σε αρκετές φωτεινές πηγές στο πίσω μέρος του εξωτερικού εκλείπτη, που φωτογραφούνται από οπτικά συστήματα του στεμματογράφου. Η απόλυτη στόχευση του συστήματος πτήσης γίνεται επιβεβαιώνοντας τη θέση της οπτικής κόρης (optical pupil) του οργάνου στη σκιά του κώνου που σχηματίζεται πίσω από τον εκλείπτη-δίσκο. Η προτεινόμενη λύση, ο Αισθητήρας Θέσης Σκιάς (Shadow Position Sensor, SPS), είναι βασισμένος στη μέτρηση της θέσης της παρασκιάς (penumbra) γύρω από την οπή εισόδου του στεμματογράφου. Η βέλτιστη στόχευση επομένως επιτυγχάνεται όταν τα επίπεδα παρασκιάς είναι ομοιόμορφα σταθμισμένα.

2.2 Υποσυστήματα

Το όργανο ASPIICS, όπως υλοποιήθηκε στο εργαστήριο για τις ανάγκες της επίδειξης αποτελείται από τα ακόλουθα υποσυστήματα:

Το υποσύστημα **TMA (Three-Mirror Anastigmat)**, το τηλεσκόπιο του οργάνου, είναι βασισμένο σε τρεις ασφαιρικούς καθρέπτες off-axis, ώστε να ελαχιστοποιηθεί το σκεδαζόμενο φώς με διάφραγμα οπτικής κόρης.

Το υποσύστημα **EOS (External Occulter Simulator)**, εξομοιώνει τον Ήλιο με παρόμοια καμπύλη φωτεινότητας. Παράγει επίσης, προκαθορισμένη ανά περίπτωση σκιά και παρασκιά, που μετράται στον Αισθητήρας Θέσης Σκιάς (SPS). Το φως, από τη φωτεινή πηγή κατευθύνεται μέσω οπτικών ινών σε ένα φίλτρο και σειρά φακών και στη συνέχεια προβάλλεται σε slide film, όπου δοκιμάζονται διαφορετικές ρυθμίσεις για εξομοίωση σκιών.

Το υποσύστημα **ASCS (Attitude S/C Simulator)**, αποτελείται απο ένα εξάποδο ρομπότ-πλατφόρμα και σκοπεύει στο να εξομοιώσει τη συμπεριφορά του ενός διαστημοπλοίου σε σχέση με το άλλο. Διαθέτει έξι βαθμούς ελευθερίας, και ικανότητα αλλαγής του κέντρου περιστροφής on-the-fly.

Η μονάδα **MCU (Metrology Control Unit)**, είναι ένα σύστημα βασισμένο σε έναν αυτοεστιαστή (autocollimator), για τη βασική ανάγκη ευθυγράμμισης δομικών στοιχείων (π.χ. της πλατφόρμας-τραπεζιού). Επίσης, χρησιμοποιείται για μελέτες δονήσεων εδάφους, ώστε να κανονικοποιηθούν οι μετρήσεις άλλων υποσυστημάτων σχετικά με αυτές.

Το βοηθητικό υποσύστημα **PCU (Photometric Control Unit)**, σχεδιάστηκε ώστε να προσφέρει ένα χάρτη δύο διαστάσεων της λαμπρότητας της προσομοιούμενης παρασκιάς (penumbra), μπροστά από τον Αισθητήρα Θέσης Σκιάς (SPS), για λόγους επαλήθευσης.

Το υποσύστημα **SPS (Shadow Position Sensor)**, ελέγχει το βαθμό της απόλυτης στόχευσης του οργάνου, ανιχνεύοντας τη μετατόπιση της παρασκιάς που σχηματίζεται από τον εξωτερικό εκλείπτη. Όταν ο σχηματισμός είναι βέλτιστος, δηλαδή ο άξονάς του είναι ευθυγράμμισμένος με το κέντρο του ήλιου, η φωτεινότητα που λαμβάνεται από οκτώ φωτο-αισθητήρες που βρίσκονται γύρω από την οπτική οπή, είναι ομοιόμορφα κατανεμημένη. Όταν η σκιά κινηθεί (μη ορθή στόχευση του σχηματισμού), το SPS εκτιμά/αποτιμά τη μετατόπιση μετρώντας και συγκρίνοντας τα λαμβανόμενα σήματα από τους αισθητήρες. Φυσικά, το υποσύστημα SPS απαρτίζεται από διάταξη οκτώ φωτοδιόδων γύρω από την οπτική οπή εισόδου του TMA.

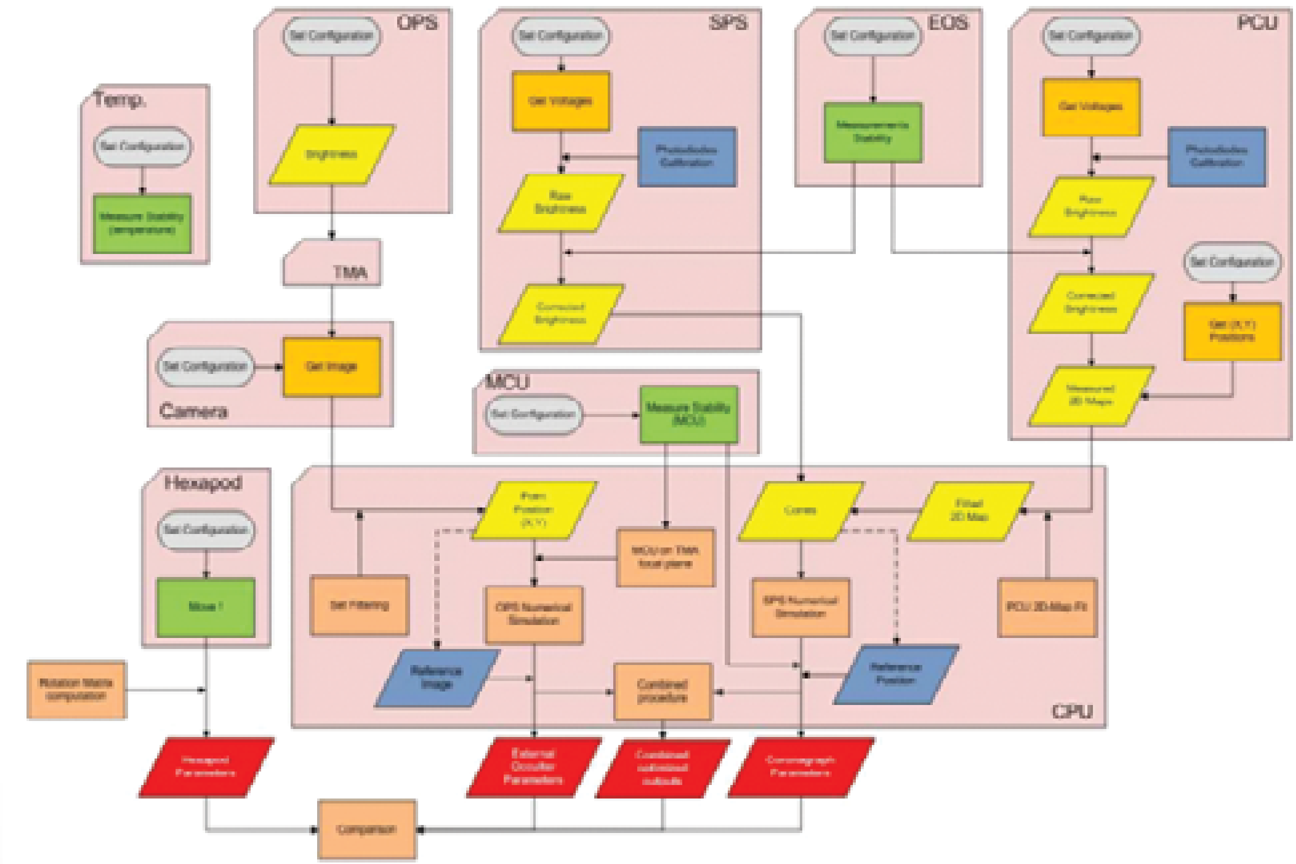
Το υποσύστημα **OPS (Occulter Position Sensor)**, προσφέρει πληροφορίες για τη θέση του εκλείπτη στο οπτικό πεδίο του στεμματογράφου. Το επιτυγχάνει αυτό με φωτοδιόδους LEDs, που βρίσκονται στο πίσω μέρος του εκλείπτη και απεστιάζονται από το CCD του τηλεσκοπίου. Έτσι, επιτυγχάνεται ευθυγράμμιση του άξονα του Σχηματισμού και του οπτικού άξονα του Στεμματογράφου. Το υποσύστημα OPS είναι υπεύθυνο για τις εξής λειτουργίες:

- Παροχή πηγών φωτός LED απέναντι από την οπή εισόδου του TMA
- Προσδιορισμού της ακριβής θέσης των φωτεινών πηγών μέσω του TMA και των φωτο-βαρυκεντρικών θέσεων τους
- Ακριβής ρύθμιση της φωτεινότητας των LEDs
- Εναλλαγή διαφορετικών διατάξεων LED
- Δυνατότητα αλλαγής εστίασης του τηλεσκοπίου TMA

Τέλος, η καρδιά του πειράματος, το **Λογισμικό Διοίκησης/Ελέγχου (Command-Control Software)**, που συνδυάζει υλικό και λογισμικό, είναι υπεύθυνο για όλες τις λογικές λειτουργίες του συστήματος. Συγκεκριμένα,

- Έλεγχος και συγχρονισμός όλων των εμπλεκόμενων τεχνικών υποσυστημάτων, όπως οι Κάρτες Λήψης Δεδομένων (Native Instruments 9205/9206), το autocollimator και το PCU για την ευθυγράμμιση, τη φωτογραφική κάμερα, του υποσυστήματος για τον έλεγχο των LEDs και το Hexapod.
- Υπηρεσίες δοκιμής (testing), με προκαθορισμένα προφίλ δοκιμής και ελέγχου σφαλμάτων
- Αποτελεσματικός χειρισμός του μεγάλου όγκου των δεδομένων
- Βασικές υπηρεσίες on-the-fly ανάλυσης δεδομένων, δυναμικού linking Matlab scripts και εξαγωγή των αποτελεσμάτων.

Το λογισμικό υλοποιήθηκε σε βάση μιας πολυεπίπεδης σπονδυλωτής προσέγγισης. Στο κατώτερο επίπεδο, βρίσκεται η υποστήριξη βιβλιοθηκών που διεκπεραιώνουν όλες τις επικοινωνίες με τις συσκευές υλικού. Στο ενδιάμεσο επίπεδο, το core application logic χειρίζεται τη λειτουργικότητα της κάθε συνιστώσας του συστήματος. Τέλος, στο ανώτατο επίπεδο, το περιβάλλον διεπαφής και η ανάλυση δεδομένων, παρέχουν μια απλή και αποτελεσματική διασύνδεση για το χειρισμό των υποκείμενων υποσυστημάτων. Στο Σχήμα 3 φαίνεται το λογικό διάγραμμα του λογισμικού, όπου διακρίνονται οι λειτουργίες των υποσυστημάτων και η ροή των πληροφοριών.



Σχήμα 3: Το λογισμικό Διοίκησης/Ελέγχου.

3. Σχεδιασμός Ηλεκτρονικού Συστήματος OPS

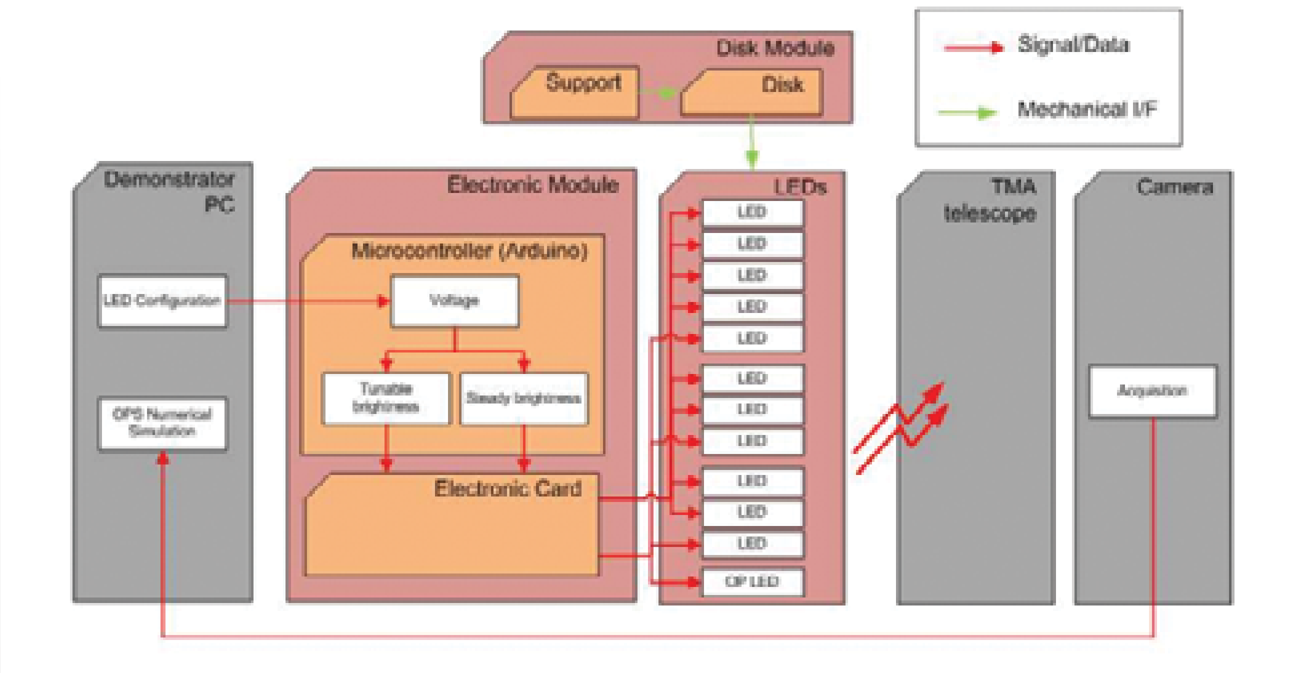
Για τις ανάγκες του ελέγχου των οπτικών διόδων (LEDs), σχεδιάστηκε και υ λοποιήθηκε το παρακάτω σύστημα, με τους εξής λειτουργικούς στόχους:

- Ικανότητα υποστήριξης διαφορετικών συνδυασμών διατάξεων LEDs
- Αυξομείωση της έντασης του εξερχόμενου φωτός

- Βοηθητικές λειτουργίες, όπως ο έλεγχος μηχανικών κλείστρων
- Πλήρης απομακρυσμένος έλεγχος μέσω Ethernet

Οριζόμενη από τις αρχές του γενικότερου λογισμικού Διοίκησης/Ελέγχου, η έμφαση δόθηκε στην ανάπτυξη μιας Ανεξάρτητης Αρχιτεκτονικής Συστήματος, διατηρώντας παράλληλα την επαναχρησιμοποίηση του κώδικα στο μέγιστο. Η ανάγκη για πλήρη ανεξαρτησία από το κυρίως σύστημα παρέχει τη δυνατότητα απομόνωσης ενός σφάλματος στο υλικό, ώστε να μην διαδοθεί.

Ως βάση ανάπτυξης του ηλεκτρονικού συστήματος, που φαίνεται στο Σχήμα 4, επελέγη ένα ανοιχτού κώδικα ενσωματωμένο σύστημα (Arduino), λόγω των πλεονεκτημάτων που προσφέρει, σαν γρήγορη, ευέλικτη και αποτελεσματική πλατφόρμα. Η ανοιχτή σχεδίαση του υλικού, βασισμένου σε ένα μικροελεγκτή της Atmel, επιτρέπει ταχεία διαδικασία ανάπτυξης με ακαδημαϊκά εργαλεία, ενώ κρατά άμεση πρόσβαση στις χαμηλού τύπου λειτουργίες του επεξεργαστή. Ένα πρωτόκολλο TCP/IP επιτρέπει τη βέλτιστη διασύνδεση με το λογισμικό Διοίκησης/ Ελέγχου. Με χρήση πακέτων, επιτρέπεται διαλειτουργικότητα στην ανταλλαγή πληροφοριών και στη μεταφορά εντολών στη συσκευή. Κάθε σύστημα λειτουργεί ως διακομιστής στο ιδιωτικό δίκτυο του πειράματος, και η χρήση αιτήσεων TCP μπορεί να επιφέρει αλλαγές, ή να ζητήσει διαγνωστικές πληροφορίες από το μικροελεγκτή. Χρησιμοποιώντας αυτό το ασύγχρονο μοντέλο Εξυπηρετητή/Πελάτη, κάθε πακέτο δικτύου ελέγχεται και ενεργοποιούνται οι κατάλληλες λειτουργίες.



Σχήμα 4: Το ηλεκτρονικό σύστημα OPS

Ολόκληρο το σύστημα βρίσκεται σε καθαρό δωμάτιο με απαίτηση να παραμένει στο ελάχιστο η ανθρώπινη δραστηριότητα, οπότε είναι αναγκαίος ο εκ των προτέρων λεπτομερής σχεδιασμός. Ο κεντρικός υπολογιστής Διοίκησης/Ελέγχου βρίσκεται στο απομονωμένο δωμάτιο ελέγχου, και συνδέεται αποκλειστικά δικτυακά με το εσωτερικό του χώρου του πειράματος.

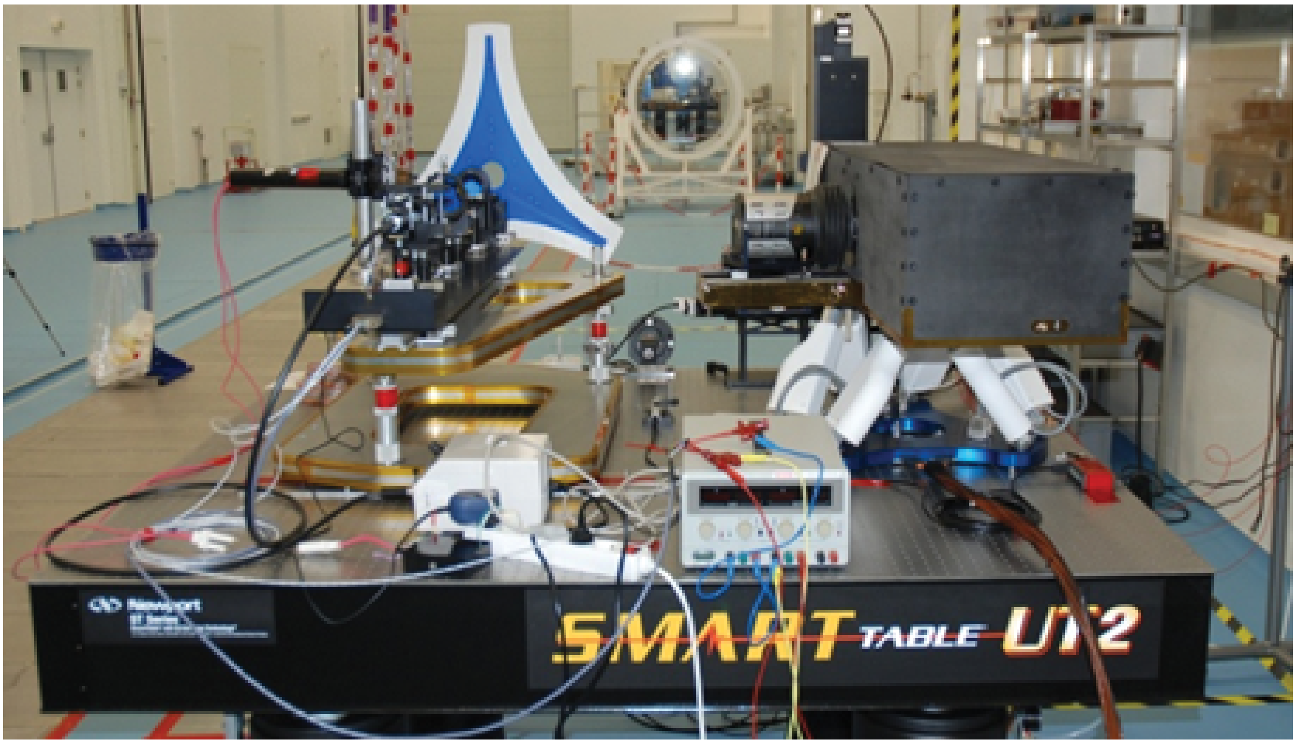
Με το ψηφιακό έλεγχο των αναλογικών κυκλωμάτων, το κόστος του συστήματος και η κατανάλωση ενέργειας μειώνονται δραστικά. Η διαμόρφωση πλάτους παλμού (PWM) είναι ένας τρόπος ψηφιακής κωδικοποίησης στάθμεων αναλογικών σημάτων. Ο κύκλος λειτουργίας ενός τετραγωνικού παλμού διαμορφώνεται για να κωδικοποιεί μία συγκεκριμένη στάθμη αναλογικού σήματος. Το σήμα PWM είναι ψηφιακό, δηλαδή σε κάθε χρονική στιγμή έχει μία από δύο καταστάσεις (High ή Low). Η τάση παρέχεται στο αναλογικό φορτίο με την έννοια των τετραγωνικών παλμών. Ο ενεργός χρόνος που είναι High είναι ο χρόνος κατά τη διάρκεια του οποίου εφαρμόζεται τάση στο φορτίο. Δεδομένου επαρκούς εύρους ζώνης, κάθε αναλογική τιμή της τάσης μπορεί να κωδικοποιηθεί με PWM. Πολλοί μικροελεγκτές, όπως και DSPs, περιλαμβάνουν ήδη on-chip ελεγκτές PWM, καθιστώντας εύκολη την εφαρμογή του. Ένα από τα πλεονεκτήματα της διαμόρφωσης PWM είναι ότι το σήμα παραμένει ψηφιακό σε όλο το κύκλωμα από τον επεξεργαστή μέχρι το ελεγχόμενο σύστημα, χωρίς να είναι απαραίτητη καμία μετατροπή του σήματος από ψηφιακό σε αναλογικό. Κρατώντας το σήμα ψηφιακό, τα φαινόμενα θορύβου ελαχιστοποιούνται.

Οι παρεχόμενες από την πλατφόρμα συχνότητες PWM λειτουργίας δεν ήταν κατάλληλα υψηλές, καθώς χρειαζόταν για τους λόγους του πειράματος δυνατότητες λεπτομερέστερης αυξομείωσης της έντασης του εξερχόμενου φωτός. Έτσι, χρησιμοποιώντας σαν σημείο αναφοράς το χρονιστή του chip, επιτυγχάνονται κατάλληλες (στατικά) συχνότητες PWM. Το chip Atmega328 διαθέτει τρεις χρονιστές. Με κατάλληλη χρήση, οι καταχωρητές χρονιστών του chip μπορούν να παράγουν διάφορες εξόδους PWM. Κάθε χρονιστής από τους Timer0, Timer1, Timer2, έχει δύο καταχωρητές σύγκρισης εξόδου, που ελέγχουν το πλάτος PWM για τις δύο του εξόδους: όταν ο χρονιστής φτάσει την τιμή σύγκρισης, η αντίστοιχη έξοδος ενεργοποιείται. Οι έξοδοι φυσικά, ανά δύο, έχουν την ίδια συχνότητα λειτουργίας, αλλά μπορούν να έχουν διαφορετικά duty cycles (που εξαρτώνται από τον αντίστοιχο καταχωρητή σύγκρισης). Το Atmega328 λειτουργεί στα 16MHz, επομένως η συχνότητα λειτουργίας κάθε χρονιστή θα είναι η συχνότητα λειτουργίας του chip διαιρεμένη με τον παράγοντα 1, 8, 64, 256 ή 1024.

Χρησιμοποιώντας τις υπάρχουσες δυνατότητες δικτύωσης, μηχανικά κλείστρα (σκίαστρα) οδηγούνται επίσης από το σύστημα. Κάθε κλείστρο, διαθέτει δύο πιθανές καταστάσεις: ανοιχτό και κλειστό. Έτσι, με γνώση της αρχικής φυσικής θέσης το σύστημα γνωρίζει την κατάσταση του κάθε κλείστρου, και είναι δυνατός ο έλεγχος τους από το λογισμικό Command/Control.

Τρεις κυκλιοειδείς συνδυασμοί LED σχεδιάστηκαν για τις ανάγκες του πειράματος, για δοκιμές σε αποστάσεις 15, 25 και 35 μέτρων, με αντίστοιχες διαμέτρους κύκλων. Ο μικρότερος δακτύλιος διαθέτει δύο παραπάνω φωτοδιόδους LED, που χρησιμοποιούνται για επαλήθευση ενός από τα μοντέλα αριθμητικής προσομοίωσης.

Στην Εικόνα 1 φαίνονται όλα τα υποσυστήματα του οργάνου ASPIICS, όπως αυτό υλοποιήθηκε στο εργαστήριο του LAM στα πλαίσια του STARTIGER.



Εικόνα 1: Η επίδειξη στο εργαστήριο του ASPIICS

Δημοσιεύσεις

Τα αποτελέσματα αυτής της πτυχιακής εργασίας συμπεριλήφθηκαν στην τεχνική αναφορά "Occulter Position Sensor Definition Document" του LAM (LAM-PJT-SFF-SPT-1046) και παρουσιάσθηκαν στα συνέδρια:

* *International Conference of Space Optics, Rhodes, Greece με τίτλο "The STARTIGER's Demonstrators: Toward a New Generation of Formation Flying Solar Coronagraphs"*
* *SPIE Space Telescopes and Instrumentation 2010: Optical, Infrared, and Millimeter Wave, Volume 7731, San Diego, USA, με τίτλο "Demonstrator of the Formation Flying Solar Coronagraph ASPIICS/PROBA-3".*

ΕΠΙΛΕΓΜΕΝΕΣ
ΔΙΠΛΩΜΑΤΙΚΕΣ ΕΡΓΑΣΙΕΣ

Συγκριτική Μελέτη Μηχανισμών Εκτίμησης Ελλιπούς Πληροφορίας σε Ασύρματα Δίκτυα Αισθητήρων

Αιμιλία Β. Αργυροπούλου

grad1046@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Στην παρούσα διπλωματική εργασία, επεξεργαζόμαστε τα δεδομένα που συλλέχθηκαν από ένα ασύρματο δίκτυο αισθητήρων. Σκοπός είναι να μελετήσουμε μεθοδολογίες πρόβλεψης ελλιπών τιμών και να συγκρίνουμε τα αποτελέσματα τους βάσει των ποσοστών επιτυχίας τους. Επίσης εξετάζουμε τεχνικές μείωσης του όγκου των δεδομένων προκειμένου να μειωθεί το υπολογιστικό και αποθηκευτικό κόστος.

Λέξεις-Κλειδιά: Κινητός υπολογισμός, Ελαττωματικοί αισθητήρες, Ανάλυση κύριων συνιστωσών, Προεκβολή, Κατηγοριοποίηση

Επιβλέπων: Ευστάθιος Χατζηευθυμιάδης, Επίκουρος Καθηγητής

1. Εισαγωγή

Ο κινητός υπολογισμός είναι ένας κλάδος που αναπτύσσεται ταχύτατα και έχει εφαρμογές σε πολλούς τομείς της ανθρώπινης δραστηριότητας. Τα ασύρματα δίκτυα αισθητήρων χρησιμοποιούνται κατά βάση στον κινητό υπολογισμό για τη συλλογή των πληροφοριών [1]. Ένα ασύρματο δίκτυο «έξυπνων» αισθητήρων επεξεργάζεται τα δεδομένα που συλλέγει [2]. Δεδομένου ότι η υπολογιστική ισχύ των αισθητήρων είναι περιορισμένη και ο όγκος των δεδομένων τεράστιος, επιβάλλεται να εφαρμοστούν τεχνικές μείωσης διαστάσεων των αρχικών δεδομένων [7]. Η μείωση των διαστάσεων είναι η διαδικασία κατά την οποία μειώνονται οι αρχικές μεταβλητές ενός συνόλου δεδομένων, χωρίς μεγάλη απώλεια πληροφορίας. Αυτή η διαδικασία μπορεί να γίνει είτε με την επιλογή γνωρισμάτων (feature selection), είτε με την εξαγωγή γνωρισμάτων (feature extraction). Οι μέθοδοι επιλογής γνωρισμάτων προσπαθούν να βρουν ένα υποσύνολο από τις αρχικές μεταβλητές που να μπορεί να αντιπροσωπεύσει το σύνολο των δεδομένων. Η εξαγωγή γνωρισμάτων προσπαθεί να προβάλει ένα σύνολο από διανύσματα υψηλής διάστασης σε ένα χώρο χαμηλότερης διάστασης. Η τεχνική που εξετάσαμε είναι η Ανάλυση Κύριων Συνιστωσών, που μειώνει στο μισό τις αρχικές διαστάσεις των δεδομένων, χρησιμοποιώντας την εξαγωγή γνωρισμάτων.

Πέρα από τη μείωση του υπολογιστικού και αποθηκευτικού κόστους, για να εξαχθούν συμπεράσματα που θα είναι χρήσιμα στους χρήστες, πρέπει να εφαρμοστούν μέθοδοι εξόρυξης γνώσης στις ροές δεδομένων που συλλέγονται [4], [5]. Ωστόσο υπάρχουν ορισμένα προβλήματα στην εφαρμογή αυτών των μεθόδων [3]. Τα προβλήματα που προκύπτουν προέρχονται κυρίως από περιττά δεδομένα που συλλέγονται από αισθητήρες, από ελλιπή δεδομένα και από την ανομοιογένεια που πιθανώς υπάρχει στη σημασιολογική ερμηνεία δεδομένων από την πλευρά των δικτύων [6]. Η εμφάνιση διαφόρων μορφών σφαλμάτων συνεπάγεται την ελλιπή πληροφόρηση προς τις εφαρμογές που καλύπτουν. Για την αποφυγή των προβλημάτων που μπορεί να προκύψουν από ελλιπείς τιμές, μελετούνται στατιστικές μέθοδοι όπως η προεκβολή/παρεμβολή, καθώς και αλγόριθμοι κατηγοριοποίησης που προβλέπουν ελλιπείς τιμές σε ένα δείγμα δεδομένων. Οι αλγόριθμοι που εξετάζονται είναι οι: Decision Tree, C4.5, M5P, Decision Stump και RepTree. Σκοπός είναι να συγκρίνουμε τις αποδόσεις των μεθοδολογιών και να διεξάγουμε συμπεράσματα για την αξιοπιστία τους, βάσει των ποσοστών επιτυχίας που έχουν στην πρόβλεψη των μετρήσεων. Επίσης, να υπολογίσουμε την απώλεια πληροφορίας που θα προκύψει, έπειτα από την εφαρμογή τεχνικών μείωσης δεδομένων.

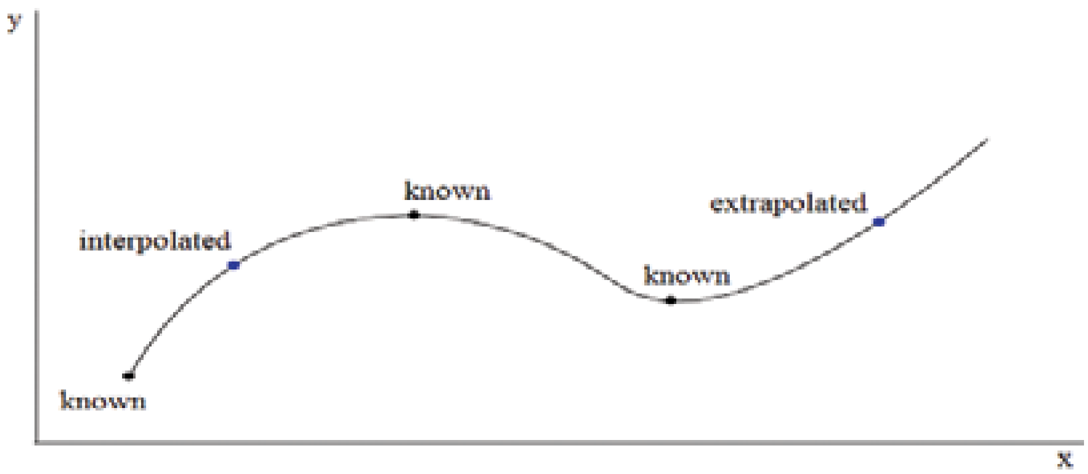
2. Ανάλυση κύριων συνιστωσών

Η ανάλυση σε κύριες συνιστώσες είναι μια τεχνική μείωσης του αρχικού όγκου του δείγματος. Χρησιμοποιείται όταν έχουμε ψηλά συσχετισμένες μεταβλητές, για τη μείωση των δεδομένων. Οι πρώτες κύριες συνιστώσες αρκούν για να αντικαταστήσουν την πλειονότητα των αρχικών δεδομένων. Η μείωση των διαστάσεων είναι επίσης χρήσιμη και όταν το πλήθος των παρατηρήσεων είναι μικρότερο από αυτό των μεταβλητών. Μια άλλη χρήση της ανάλυσης κύριων συνιστωσών είναι η οπτικοποίηση των δεδομένων καθώς είναι δύσκολο να απεικονιστούν όταν υπάρχουν παραπάνω από τρείς μεταβλητές. Μέσω των κύριων συνιστωσών απεικονίζεται η μεγαλύτερη μεταβλητότητα των δεδομένων σε δύο μόνο διαστάσεις. Έτσι, η πολυδιάστατη πληροφορία μπορεί να συγχωνευτεί σε μια πιο συμπαγή απεικόνιση. Τέλος, η ανίχνευση ακραίων τιμών (noise reduction) και η ομαδοποίηση των δεδομένων γίνεται πιο εύκολη με τη χρήση αυτής της τεχνικής [8].

Η Ανάλυση Κύριων Συνιστωσών είναι μια πολύ διαδεδομένη στατιστική τεχνική, που χρησιμοποιείται ευρέως σε προβλήματα που προκύπτουν από την επεξεργασία σήματος, αλλά και στην ανάλυση δεδομένων. Αυτή η μαθηματική διαδικασία χρησιμοποιεί τον ορθογώνιο μετασχηματισμό για να μετατρέψει ένα σύνολο από παρατηρήσεις, όπου το πλήθος των μεταβλητών του είναι πολύ μεγάλο και πιθανώς συσχετιζόμενο, σε ένα νέο σύνολο μη-συσχετισμένων μεταβλητών, που ονομάζονται κύριες συνιστώσες (principal components). Από τις νέες συνιστώσες, μόνο ένα μέρος αυτών θα χρησιμοποιηθεί για την εξαγωγή συμπερασμάτων. Η μελέτη δύο ή τριών μη-συσχετιζόμενων μεταβλητών είναι ευκολότερη από τη μελέτη του συνόλου των αρχικών μεταβλητών. Ο μετασχηματισμός γίνεται με τέτοιο τρόπο, ώστε η πρώτη κύρια συνιστώσα να προκύπτει από τα δεδομένα με τη μεγαλύτερη διακύμανση και κάθε επόμενη συνιστώσα να έχει την αμέσως επόμενη μεγαλύτερη δυνατή διακύμανση, με τον περιορισμό ότι οι κύριες συνιστώσες είναι μεταξύ τους ορθογώνιες (δηλαδή παραμένουν ασυσχέτιστες μεταξύ τους). Οι κύριες συνιστώσες είναι ανεξάρτητες (δεν συσχετίζονται), εάν το σύνολο των δεδομένων ακολουθεί κανονική κατανομή [9], [10].

3. Μεθοδολογίες παρεμβολής / προεκβολής

Σε περιπτώσεις που το δίκτυο αισθητήρων μπορεί να βρεθεί εκτός λειτουργίας και συνεπώς να αποτύχει να αποστείλει τις τιμές των μετρήσεών του, χρησιμοποιούνται οι μεθοδολογίες παρεμβολής και προεκβολής. Κατά τη διάρκεια συλλογής των δεδομένων από τους αισθητήρες μπορούμε να δημιουργήσουμε μια καμπύλη πάνω στην οποία θα εμφανίζονται όλα τα συλλεχθέντα σημεία. Παρεμβολή είναι η διαδικασία ανάκτησης τιμών από την καμπύλη, ανάμεσα στα ήδη γνωστά σημεία (δεδομένα). Αυτό επιτυγχάνεται μέσω της προσαρμογής της καμπύλης ή της χρήση της παλινδρομικής ανάλυσης. Προεκβολή (Extrapolation) είναι η διαδικασία απόκτησης μιας τιμής από ένα γράφημα ή μια γραφική παράσταση που εκτείνεται πέρα από τα συλλεχθέντα δεδομένα. Η διαδικασία της προεκβολής είναι παρόμοια με τη διαδικασία της παρεμβολής, αλλά τα αποτελέσματά της συνήθως υπόκεινται σε μεγαλύτερη αβεβαιότητα, γιατί δεν περιορίζονται σε ένα γνωστό εύρος τιμών.



Παρεμβολή (Interpolation) / Προεκβολή (Extrapolation)

Στα πειράματα που πραγματοποιήθηκαν, θέλαμε να προβλέψουμε τιμές σε περιπτώσεις που το δίκτυο αισθητήρων δεν εκπέμπει καθόλου τις μετρήσεις, οπότε μελετήσαμε τη μεθοδολογία της προεκβολής Cubic Spline [11].

4. Αλγόριθμοι κατηγοριοποίησης

Ένα σύνολο δεδομένων που συλλέγεται από αισθητήρες είναι πολύ μεγάλο σε όγκο, λόγω της συνεχούς ενημέρωσης του. Γι' αυτό το λόγο, πρέπει να δημιουργηθεί ένα μοντέλο πρόβλεψης τιμών που να αντιπροσωπεύει την πλειονότητα των δεδομένων και να μπορεί να ικανοποιήσει σε γενικό επίπεδο όλες τις μεμονωμένες ελλιπείς τιμές. Κατηγοριοποίηση των δεδομένων είναι η διαδικασία κατά την οποία εξετάζονται τα βασικότερα χαρακτηριστικά ενός συνόλου δεδομένων και βάσει αυτών κατηγοριοποιείται ένα νέο αντικείμενο σε μια προκαθορισμένη κατηγορία, σύμφωνα με το μοντέλο κατηγοριοποίησης. Για τη δημιουργία του μοντέλο κατηγοριοποίησης πρέπει πρώτα να δοθεί ένα σύνολο με δεδομένα εκμάθησης (training data) για την κάθε κατηγορία. Έπειτα, με τη χρήση αυτού του μοντέλου, θα μπορούν να κατηγοριοποιηθούν δεδομένα τα οποία δεν ανήκουν σε κάποια κατηγορία (άγνωστα/νέα αντικείμενα). Συνήθως, χρησιμοποιούνται δένδρα απόφασης (decision trees) ως μεθοδολογίες κατηγοριοποίησης. Στόχος είναι να δημιουργηθεί ένα μοντέλο που θα μπορεί να προβλέπει την τιμή μιας μεταβλητής, βασιζόμενο στην είσοδο δεδομένων των υπολοίπων μεταβλητών [12].

Οι αλγόριθμοι που μελετήσαμε είναι οι εξής:

- **ID3 / C4.5**

Η βασική ιδέα του αλγορίθμου είναι η κατασκευή ενός δένδρου από πάνω προς τα κάτω, εφαρμόζοντας «άπληστη» αναζήτηση στο σύνολο δεδομένων που δίνεται. Το χαρακτηριστικό γνώρισμα που θα επιλέξουμε να χρησιμοποιήσουμε για την κατηγοριοποίηση (classification) του συνόλου των δεδομένων, πρέπει να μεγιστοποιεί το κέρδος πληροφορίας (information gain). Ο αλγόριθμος C4.5 είναι μια επέκταση του αλγορίθμου ID3 και κατασκευάζει δένδρα απόφασης με τον ίδιο τρόπο. Ωστόσο, ο C4.5 έχει αρκετές βελτιστοποιήσεις σε σχέση με τον ID3. Αυτές οι βελτιώσεις περιλαμβάνουν μεθόδους για την αντιμετώπιση των αριθμητικών γνωρισμάτων, των ελλιπών τιμών, του «θορύβου» στα δεδομένα (με εφαρμογή τεχνική εκτίμησης ποσοστού λάθους πάνω στα ίδια τα δεδομένα εκπαίδευσης) και τη δημιουργία κανόνων από το δένδρο που δημιουργείται [12], [13].

- **M5P**

Ο αλγόριθμος M5P κατασκευάζει δένδρο όπου τα φύλλα του αποτελούνται από γραμμικά μοντέλα πολλών μεταβλητών. Το κριτήριο διαχωρισμού κάθε εσωτερικού κόμβου είναι η ελαχιστοποίηση της διακύμανσης των υποσυνόλων που θα προκύψουν [14].

- **RepTree**

Ο αλγόριθμος RepTree εφαρμόζεται σε αριθμητικά δεδομένα και δίνει ως αποτέλεσμα τη μέση τιμή των αντικειμένων του κάθε φύλλου. Χρησιμοποιεί τη μετρική του information gain και τη μείωση της διακύμανσης στους εσωτερικούς κόμβους [12].

- **Decision Stump**

Ο αλγόριθμος Decision Stump κατασκευάζει δένδρο απόφασης ενός επιπέδου (βάθους 1) [15]. Κάνει τις προβλέψεις βασιζόμενο στην τιμή μίας μόνο μεταβλητής εισόδου. Υπολογίζει ένα σκορ, για κάθε ένα γνώρισμα, καταμετρώντας πόσο καλά χωρίζει το σύνολο της τάξης.

5. Πειραματική αξιολόγηση

Για την εφαρμογή των πειραμάτων, χρησιμοποιήθηκαν δεδομένα που συλλέχθηκαν σε πραγματικό χρόνο από το πεδίο. Το δίκτυο αισθητήρων που χρησιμοποιήθηκε αποτελείται από τρία ζευγάρια αισθητήρων που μετρούν τη θερμοκρασία και την υγρασία του περιβάλλοντος και έναν τέταρτο αισθητήρα που μετρά την ταχύτητα του ανέμου τις ίδιες χρονικές στιγμές. Συνεπώς, η κάθε μία διανυσματική μέτρηση αποτελείται από επτά μεταβλητές της μορφής:

$$x = (\mathrm{temp\_1}, \mathrm{hum\_1}, \mathrm{temp\_2}, \mathrm{hum\_2}, \mathrm{temp\_3}, \mathrm{hum\_3}, \mathrm{wind\_speed})$$

Τα δεδομένα που έχουν συλλεχθεί και χρησιμοποιηθεί στα πειράματα, απεικονίζουν τις μετρήσεις για 387 διαφορετικές χρονικές στιγμές. Συνεπώς, έχουμε 387 διαφορετικές μετρήσεις για διακριτές χρονικές στιγμές.

Τα πειράματα που έγιναν καθώς και τα συμπεράσματα που εξήχθησαν, μπορούν να επεκταθούν και σε ένα γενικότερο περιβάλλον με πλήθος κόμβων αισθητήρων (άρα και διαφορετικό αριθμό μεταβλητών) καθώς και σε διαφορετικό είδος μετρήσεων.

5.1 Πειράματα προεκβολής

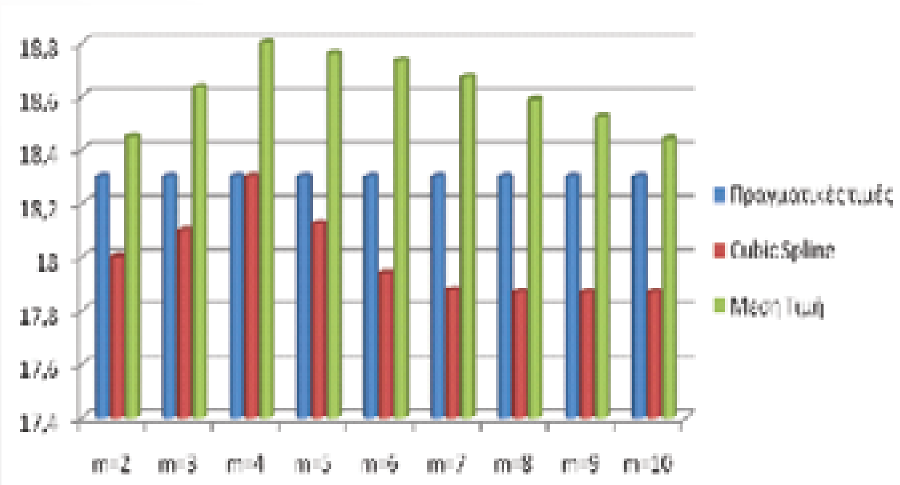
Για τον έλεγχο της ακρίβειας μιας πρόβλεψης που βασίζεται στην προεκβολή δεν λάβαμε υπόψιν όλες τις τιμές από την αρχή του dataset, αλλά μόνο τις m τελευταίες μετρήσεις, για να προβλέψουμε την αμέσως επόμενη τιμή του κάθε γνωρίσματος. Το μέγεθος του m ξεκίνησε από 2 και έφτασε μέχρι 10. Έπειτα από αρκετά πειράματα, παρατηρήθηκε ότι όταν το m έπαιρνε τιμές μεγαλύτερες του 10, οι τιμές που προέβλεπε συνέκλιναν, οπότε δεν υπήρχε λόγος να εξετάσουμε προβλέψεις για μεγαλύτερη τιμή m. Για κάθε τιμή του m, υπολογίσαμε την τιμή πρόβλεψης που προέκυπτε έπειτα από την εφαρμογή της μεθοδολογίας Cubic Spline καθώς και τη μέση τιμή του γνωρίσματος βάσει των προηγούμενων μετρήσεων.

Πιο συγκεκριμένα, σε κάθε σενάριο δόθηκαν διαδοχικά m μετρήσεις με σκοπό να δημιουργηθεί η αντίστοιχη συνάρτηση και έτσι να γίνει η πρόβλεψη για την 251η χρονική στιγμή. Ο σκοπός που δοκιμάστηκαν σενάρια με διαφορετικό πλήθος μετρήσεων είναι για να δούμε πώς το μέγεθος του παραθύρου (m) επηρεάζει την ακρίβεια της συνάρτησης που κατασκευάζεται και συνεπώς την ακρίβεια των αποτελεσμάτων. Επίσης, θέλαμε να συγκρίνουμε τα αποτελέσματα που προκύπτουν από την εφαρμογή της Cubic Spline με αυτά της μέσης τιμής της ίδιας της μεταβλητής, ώστε να ελέγξουμε ποιά μέθοδος είναι πιο αποτελεσματική και αξιόπιστη. Παρότι πραγματοποιήθηκαν προβλέψεις για όλες τις μεταβλητές ξεχωριστά, παρακάτω παρουσιάζονται τα αποτελέσματα για τις μετρήσεις της θερμοκρασίας του 1ου ζεύγους αισθητήρων. Τα συμπεράσματα που προκύπτουν είναι όμοια και για τις υπόλοιπες μεταβλητές.

Η πραγματική τιμή της μέτρηση της θερμοκρασίας του 1ου ζεύγους αισθητήρων την 251η χρονική στιγμή είναι 18,3. Ακολουθεί ο πίνακας με τις μετρήσεις των προβλέψεων.

Μέγεθος παραθύρου	Αποτέλεσμα Cubic Spline	Αποτέλεσμα Μέσης Τιμής
m=2	18,0000	18,4500
m=3	18,1000	18,6333
m=4	18,3000	18,8000
m=5	18,1250	18,7600
m=6	17,9400	18,7333
m=7	17,8786	18,6714
m=8	17,8679	18,5875
m=9	17,8672	18,5222
m=10	17,8674	18,4400

Πίνακας 1: Αποτελέσματα Πρόβλεψης 1ης Θερμοκρασίας



Σχήμα 1: Διαγραμματική Απεικόνιση Πρόβλεψης 1ης Θερμοκρασίας

Όπως παρατηρούμε, όσο το m παραμένει μικρό οι προβλέψεις της Cubic Spline είναι αρκετά ακριβείς. Πιο συγκεκριμένα, για m=4 η πρόβλεψη έχει 100% ακρίβεια και μέχρι το m = 6 η Cubic Spline δίνει πλησιέστερα αποτελέσματα στην πραγματική μέτρηση απ' ότι η μέση τιμή. Ωστόσο, και η μέση τιμή δίνει αρκετά ικανοποιητικά αποτελέσματα, κυρίως καθώς το m αυξάνεται. Δηλαδή, για m > 6 είναι προτιμότερο να χρησιμοποιήσουμε την πρόβλεψη της μέσης τιμής, η οποία φαίνεται ότι προσεγγίζει καλύτερα την πραγματική τιμή από την Cubic Spline που σταδιακά σταθεροποιείται περίπου στο 17,86.

Το γενικό συμπέρασμα από όλα τα πειράματα είναι πως καθώς αυξάνεται το m, η μέση τιμή βελτιώνει την απόδοσή της, ενώ το αντίθετο συμβαίνει με τη Cubic Spline. Για να επιτύχει καλύτερη πρόβλεψη η Cubic Spline από τη χρήση της μέσης τιμής, πρέπει το m να είναι πολύ μικρό, δηλαδή να λάβει υπόψιν του μόνο τις 2-3 τελευταίες μετρήσεις.

5.2 Πειράματα κατηγοριοποίησης

Στα πειράματα που εφαρμόστηκαν για τον έλεγχο των αλγορίθμων κατηγοριοποίησης, υλοποιήθηκαν τρία σενάρια. Σε κάθε ένα σενάριο δόθηκε ως είσοδο ένα διαφορετικό ποσοστό του αρχικού δείγματος, με στόχο να προβλέψουν ένα τμήμα από τις αρχικές μετρήσεις. Το πρώτο σενάριο χρησιμοποίησε περίπου το 1/20 του συνόλου ως δεδομένα εκπαίδευσης (20 μετρήσεις), το δεύτερο σενάριο το 1/5 του συνόλου ως δεδομένα εκπαίδευσης (75 μετρήσεις), ενώ το τρίτο σενάριο το 1/2 του συνόλου (190 μετρήσεις). Για κάθε σενάριο, το κομμάτι των δεδομένων που θα χρησιμοποιούταν για επαλήθευση, παρέμενε σταθερό και ήταν 50 μετρήσεις. Ουσιαστικά, το σύνολο εκπαίδευσης είναι ένα παράθυρο που κάθε φορά αυξάνεται για να προβλέψει τις επόμενες 50 μετρήσεις. Ο λόγος είναι για να δούμε τί επίδραση θα έχει αυτό στο δένδρο που δημιουργείται, και συνεπώς στις προβλέψεις.

Για τα δεδομένα που χρησιμοποιήθηκαν στην επαλήθευση των προβλέψεων, επιλέχθηκε η θερμοκρασία του τρίτου ζεύγους αισθητήρων (δηλαδή η πέμπτη παράμετρος των μετρήσεων) για την οποία θα γίνουν οι προβλέψεις. Θεωρήσαμε ότι οι τιμές αυτών των μετρήσεων δεν είχαν συλλεχθεί, οπότε αφαιρέθηκαν από τον πίνακα και έγινε η πρόβλεψη των ελλιπών τιμών με βάση τα δένδρα που είχαν δημιουργηθεί από τα δεδομένα εκπαίδευσης και με είσοδο στους αλγορίθμους τις υπόλοιπες μετρήσεις των αισθητήρων.

Ορισμένοι από τους αλγόριθμους που χρησιμοποιήθηκαν στα πειράματα δεν εφαρμόζονται σε συνεχή δεδομένα (π.χ. C4.5), συνεπώς έπρεπε να γίνει μια προεπεξεργασία των δεδομένων εισόδου, ώστε να μετατραπούν σε διακριτές οι τιμές της στήλης για πρόβλεψη. Βέβαια, υπάρχει και η περίπτωση κάποιων άλλων αλγορίθμων (όπως ο M5P) που εφαρμόζονται μόνο σε συνεχείς τιμές και όχι σε διακριτές. Σε αυτή την περίπτωση, χρησιμοποιήθηκε το αρχικό dataset που περιείχε τις συνεχείς τιμές, και αφού έγινε η πρόβλεψη, έπειτα έγινε η μετατροπή σε διακριτές τιμές, ώστε να γίνει η σύγκριση των αποτελεσμάτων. Το πλήθος των διακριτών τιμών που θα αντικαθιστούσαν τις πραγματικές τιμές (συνεχείς τιμές) επιλέχθηκε τυχαία. Επομένως, οι τιμές τις θερμοκρασίας αντικαταστάθηκαν με τις εξής διακριτές τιμές:

a : περιλαμβάνει θερμοκρασίες του εύρους (6 - 9,9) βαθμούς °C.

b : περιλαμβάνει θερμοκρασίες του εύρους (10 - 13,9) βαθμούς °C.

c : περιλαμβάνει θερμοκρασίες του εύρους (14 - 17,9) βαθμούς °C.

d : περιλαμβάνει θερμοκρασίες του εύρους (18 - 22,3) βαθμούς °C.

Σε δεύτερη φάση, τα παραπάνω σενάρια εκτελέστηκαν εκ νέου, αφού πρώτα εφαρμόστηκε η τεχνική PCA. Σκοπός ήταν να δούμε την απόκλιση των προβλέψεων όταν εφαρμόζονται τεχνικές μείωσης διαστάσεων. Για να μπορέσουμε να συγκρίνουμε την απόκλιση ανάμεσα στην τιμή πρόβλεψης του αρχικού μοντέλου και στην πρόβλεψη έπειτα από την εφαρμογή PCA, συγκρίναμε τα αποτελέσματα από τους αλγορίθμους που προβλέπουν πραγματικές τιμές (δηλαδή M5P, Decision Stump και RepTree). Σε ένα πραγματικό περιβάλλον, όπου θα υπάρχει συνεχής ροή μετρήσεων ως είσοδο, είναι αναπόφευκτη η χρήση τεχνικών μείωσης διαστάσεων. Συνεπώς, αυτό επηρεάζει την απόδοση των αλγορίθμων που εξετάζουμε.

Επίσης, όλα τα σενάρια εφαρμόστηκαν σε δύο διαφορετικά σύνολα δεδομένων (εκπαίδευσης και επαλήθευσης) για να εξετάσουμε αν η απόδοση των αλγορίθμων παραμένει σταθερή ή επηρεάζεται από το είδος των μετρήσεων.

Βασικός στόχος των πειραμάτων είναι να μετρηθεί η ακρίβεια των προβλέψεων και η απώλεια πληροφορίας έπειτα από την εφαρμογή PCA. Για να επιτευχθεί αυτό, συγκρίνουμε τα αποτελέσματα της πρόβλεψης με τις πραγματικές τιμές και υπολογίζουμε το ποσοστό επιτυχίας. Ύστερα, εφαρμόζουμε PCA και υπολογίζουμε το σχετικό σφάλμα και τη διακύμανση σε σχέση με την αρχική πρόβλεψη.

Πλήθος Δεδομένων Εκπαίδευσης	**Ποσοστό Επιτυχίας**					
	ID3 Χωρίς Pruning	ID3 Με Pruning	C4.5	M5P	Decision Stump	RepTree
1ο Δείγμα	**Προβλέψεις Μετρήσεων 251-300**					
20 (1ο Δένδρο)	54%	54%	64%	94%	52%	54%
75 (2ο Δένδρο)	82%	82%	100%	96%	36%	84%
190 (3ο Δένδρο)	98%	92%	92%	94%	36%	92%
2ο Δείγμα	**Προβλέψεις Μετρήσεων 338-387**					
20 (1ο Δένδρο)	42%	42%	42%	54%	60%	60%
75 (2ο Δένδρο)	90%	90%	92%	96%	10%	78%
190 (3ο Δένδρο)	92%	92%	92%	90%	84%	94%

Με βάση τα αποτελέσματα, παρατηρούμε ότι με τη χρήση 20 μετρήσεων, ως δεδομένα εκπαίδευσης, τα ποσοστά επιτυχίας ήταν αρκετά χαμηλά στην πλειονότητα των αλγορίθμων. Αυτό οφείλεται στο λεγόμενο underfitting, δηλαδή τα δένδρα δεν έχουν κατασκευαστεί με τα απαραίτητα δεδομένα για να καλύψουν τα αποτελέσματα από όλες τις δυνατές περιπτώσεις. Η χρήση 75 μετρήσεων ως δεδομένα εκπαίδευσης αυξάνει αρκετά το ποσοστό επιτυχίας και τέλος, η χρήση των 190 μετρήσεων σε ορισμένους αλγόριθμους αυξάνει σε μικρό βαθμό το ποσοστό επιτυχίας (π.χ. ID3 με/χωρίς pruning, RepTree, Decision Stump), ενώ στις περιπτώσεις των αλγορίθμων M5P και C.4.5. το ποσοστό επιτυχίας μειώνεται ελαφρώς, λόγω του overfitting.

Από τον παραπάνω πίνακα, παρατηρούμε ότι ο αλγόριθμος M5P διατηρεί υψηλά επίπεδα απόδοσης και στα δύο δείγματα δεδομένων. Ακολουθεί ο C4.5. Αξιοσημείωτο είναι το γεγονός ότι ο αλγόριθμος Decision Tree, με pruning και χωρίς, έχει τις περισσότερες φορές το ίδιο ποσοστό επιτυχίας. Τα αποτελέσματα του RepTree είναι εξίσου ικανοποιητικά. Ο αλγόριθμος Decision Stump έχει, με διαφορά, τη χειρότερη απόδοση από όλους τους άλλους και ιδιαίτερα στο δεύτερο δείγμα δεδομένων έχει περίεργη διακύμανση στα ποσστά επιτυχίας του. Συμπεραίνουμε λοιπόν ότι δεν είναι αξιόπιστος.

Η μεθοδολογία PCA εφαρμόστηκε στους αλγορίθμους πρόβλεψης πραγματικών τιμών (M5P, Decision Stump, RepTree). Φυσικά, μπορεί να εφαρμοστεί και στους αλγορίθμους πρόβλεψης διακριτών τιμών, αλλά ο σκοπός των πειραμάτων είναι να υπολογιστεί η απώλεια πληροφορίας έπειτα από την εφαρμογή τεχνικών μείωσης διαστάσεων. Για το λόγο αυτό, υπολογίσαμε τη μέση τιμή του σχετικού σφάλματος ανάμεσα στην αρχική πρόβλεψη και την πρόβλεψη με PCA με χρήση του τύπου:

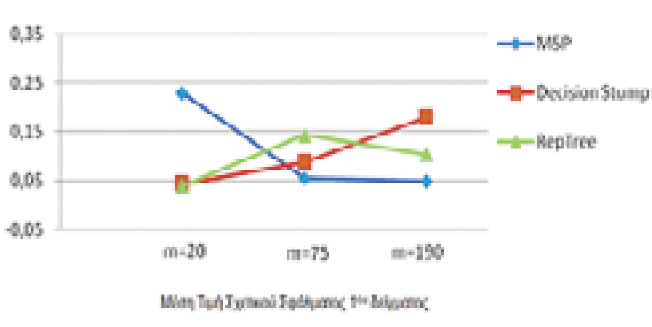
$$\frac{\left|p_i - p_j\right|}{\left|p_i\right|}$$, όπου p_i αρχική πρόβλεψη και p_j πρόβλεψη με pca.

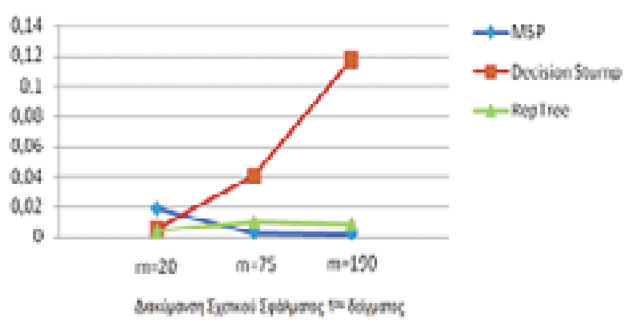
Ύστερα υπολογίσαμε τη διακύμανση του σχετικού σφάλματος βάσει του τύπου:

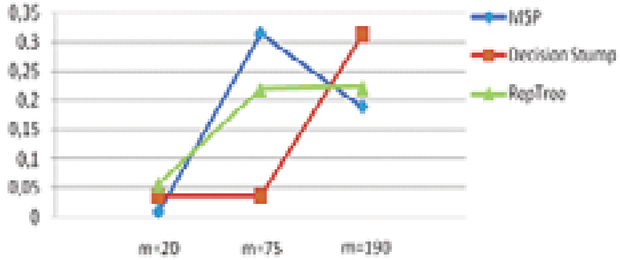
$$\frac{\sum\left(x - \overline{x}\right)^2}{n-1}$$, όπου $\overline{x}$ είναι η μέση τιμή του σχετικού σφάλματος και n είναι το μέγεθος του δείγματος.

Η χρήση της μεθοδολογίας PCA δημιούργησε έξι κύριες συνιστώσες που αντικαθιστούν τις έξι μεταβλητές εισόδου (temp_1, hum_1, temp_2, hum_2, hum_3, wind_speed). Για να μειωθεί το μέγεθος των αρχικών δεδομένων χρησιμοποιούμε μόνο τις τρείς πρώτες κύριες συνιστώσες, οι οποίες καλύπτουν πάνω από το 95% της αρχικής πληροφορίας.

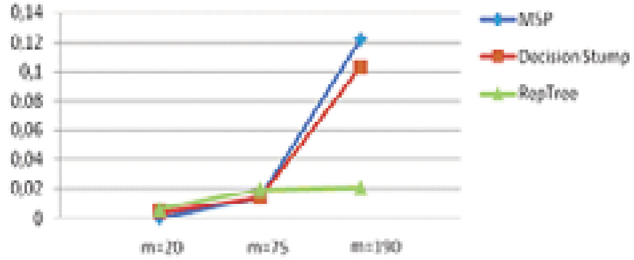
	M5P			Decision Stump			RepTree		
	m=20	m=75	m=190	m=20	m=75	m=190	m=20	m=75	m=190
1° Δείγμα	Προβλέψεις Μετρήσεων 251 – 300								
Μέση Τιμή Σχετικού Σφάλματος	0,228668	0,05469	0,04863	0,043291	0,08747	0,17949	0,039406	0,14224	0,10352
Διακύμανση Σχετικού Σφάλματος	0,018824	0,00292	0,00191	0,004781	0,04044	0,11726	0,004075	0,01003	0,00882
2° Δείγμα	Προβλέψεις Μετρήσεων 338-387								
Μέση Τιμή Σχετικού Σφάλματος	0,008488	0,31441	0,18857	0,036253	0,03634	0,31264	0,0565	0,219	0,22263
Διακύμανση Σχετικού Σφάλματος	0,000016	0,01371	0,12198	0,004009	0,01426	0,10289	0,006361	0,01915	0,02061







Μέση Τιμή Σχετικού Σφάλματος 2ου Δείγματος



Διακύμανση Σχετικού Σφάλματος 2ου Δείγματος

Όπως παρατηρούμε από το 1ο δείγμα, η μέση τιμή του σφάλματος για τον αλγόριθμο M5P ξεκινάει από μια υψηλή τιμή συγκριτικά με τους υπόλοιπους αλγορίθμους όταν το σύνολο δεδομένων εκπαίδευσης είναι μικρό, αλλά μειώνεται σημαντικά καθώς τα δεδομένα εκπαίδευσης αυξάνονται. Άρα, όσο περισσότερα δεδομένα εκπαίδευσης εισάγονται η απώλεια πληροφορίας με τη χρήση PCA μειώνεται σταδιακά. Δεν ισχύει το ίδιο και για τους άλλους δύο αλγορίθμους, όπου η απώλεια πληροφορίας αυξάνεται με την αύξηση των δεδομένων. Ειδικά στον αλγόριθμο Decision Stump αυξάνεται αρκετά το σφάλμα αλλά και η διακύμανση του, υποδεικνύοντας μεγάλες αποκλίσεις στην αρχική τιμή πρόβλεψης και στην πρόβλεψη που προκύπτει από την εφαρμογή PCA.

Όσον αφορά το 2ο δείγμα, το σφάλμα κυμαίνεται σε μεγαλύτερα επίπεδα. Ωστόσο, οι αλγόριθμοι Decision Stump και RepTree, διατηρούν την ίδια συμπεριφορά σε σχέση με το 1ο δείγμα. Ο M5P δεν έχει σταθερή πορεία καθώς το σφάλμα και η διακύμανση αυξομειώνονται ανάλογα με το πλήθος των δεδομένων εισόδου.

Είναι αναμενόμενο να αυξάνεται το σφάλμα και η διακύμανση καθώς αυξάνεται το m, γιατί τα δεδομένα που χρησιμοποιούνται για τη δημιουργία των κύριων συνιστωσών είναι περισσότερα και με μεγαλύτερες αποκλίσεις. Αυτό, συνεπάγεται μεγαλύτερη απώλεια πληροφορίας κατά το μετασχηματισμό. Ο ρυθμός αύξησης επηρεάζεται από την αναλογία των τιμών που πρέπει να προβλέψει ο κάθε αλγόριθμος σε σχέση με τις τιμές που έχουν δοθεί ως δεδομένα εκπαίδευσης. Για παράδειγμα, στο 2ο δείγμα όπου m=75, οι αλγόριθμοι M5P και RepTree παρουσιάζουν μια παράδοξη αύξηση. Αυτή οφείλεται στο ότι οι τιμές της θερμοκρασίας που πρέπει να υπολογιστούν είναι σχετικά χαμηλές, ενώ το εκπαιδευτικό σύνολο περιελάμβανε μικρό ποσοστό αντίστοιχων μετρήσεων.

Συνολικά, ο αλγόριθμος RepTree φαίνεται να είναι πιο σταθερός συγκριτικά με τους υπόλοιπους αλγόριθμους, γιατί διατηρεί χαμηλό σφάλμα και ακόμα χαμηλότερη διακύμανση. Στα τέσσερα από τα έξι πειράματα που έγιναν με χρήση PCA ο M5P είχε καλύτερα αποτελέσματα. Οι αλγόριθμοι M5P και Decision Stump παρουσιάζουν αυξομειώσεις στο σφάλμα και τη διακύμανση, αλλά σε γενικές γραμμές οι αποκλίσεις από τις αρχικές προβλέψεις δεν είναι υψηλές. Δηλαδή, όταν το σχετικό σφάλμα είναι περίπου 30% (περιπτώσεις M5P για m=75, Decision Stump για m=190) θεωρείται αρκετά υψηλό, αλλά δεν θα πρέπει να μας προκαλεί ανησυχία, γιατί αυτό το ποσοστό αντιστοιχεί σε μια διαφορά περίπου 3,5 βαθμών °C από την αρχική πρόβλεψη.

6. Συμπεράσματα

Η μεθοδολογία extrapolation έδειξε ότι χρειάστηκαν περίπου 10 μετρήσεις για να σταθεροποιηθεί η τιμή μιας πρόβλεψης. Όσο μικρότερο ήταν το δείγμα εκπαίδευσης, τόσο καλύτερες ήταν οι τιμές εκτίμησης. Σε σχέση με τις προβλέψεις που έγιναν με βάση τη μέση τιμή παρατηρήθηκε ότι η μεθοδολογία extrapolation μειονεκτεί.

Όσον αφορά τους αλγορίθμους κατηγοριοποίησης, τα δένδρα που κατασκευάστηκαν με χρήση pruning δεν διέφεραν στις προβλέψεις τους απ' ότι τα δένδρα χωρίς pruning. Στα πειράματα που χρησιμοποιήθηκε μεγαλύτερο ποσοστό δεδομένων ως δεδομένα εκπαίδευσης, τα αποτελέσματα των αλγορίθμων είχαν καλύτερη απόδοση τις περισσότερες φορές. Ωστόσο, σε αρκετές περιπτώσεις η αναλογία αύξησης του ποσοστού των δεδομένων εκπαίδευσης ως προς τη βελτίωση των αποτελεσμάτων ήταν δυσανάλογη. Η χρήση της τεχνικής μείωσης των αρχικών διαστάσεων PCA οδήγησε σε μικρή απώλεια πληροφορίας σε σχέση με τις αρχικές προβλέψεις. Ιδιαίτερα όταν τα δεδομένα εκπαίδευσης ήταν λίγα, το σφάλμα ανάμεσα στην αρχική πρόβλεψη και την πρόβλεψη με χρήση PCA ήταν σχεδόν αμελητέο.

Αναφορές

1. C. Siva Ram Murthy and B. S. Manoj, Ad Hoc Wireless Networks, Architectures and Protocols, Prentice-Hall, 2004

2. V. Tsetsos, N. Silvestros and S. Hadjiefthymiades, «Collaborative Sensing over Smart Sensors», 2nd Student Workshop on Wireless Sensor Netwotks, Athens, 2009.

3. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A Survey on Sensor Networks", IEEE Communications Magazine, vol. 40, no. 8, Aug. 2002, pp. 102–14.

4. E. Della Valle, S. Ceri, F. Harmelen and D. Fensel, "It's a Streaming World! Reasoning upon Rapidly Changing Information", IEEE Intelligent Systems , 24(6):83–89, 2009

5. J. Hong, E. Suh, J. Kim and S. Kim, "Context-Aware System for Proactive Personalized Service Based on Context History," Expert Systems with Applications, vol. 36, no. 4, 2009, pp. 7448–7457.

6. R. Dunia, S.J. Qin, T.F. Edgar and T.J. McAvoy, "Identification of Faulty Sensors Using Principal Component Analysis", AIChE Journal, October, vol. 42, no 10, 1996, pp. 2797-2812.

7. J. Himberg, J. Mäntyjärvi and P. Korpipää, «Using PCA and ICA for Exploratory Data Analysis in Situation Awareness», Proc. IEEE Conf. Multisensor Fusion and Integration for Intelligent Systems, IEEE CS Press, 2001, pp. 127-131.

8. J. Himberg, J. Mäntyjärvi and P. Korpipää, «Using PCA and ICA for Exploratory Data Analysis in Situation Awareness», Proc. IEEE Conf. on Multisensor Fusion and Integration for Intelligent Systems, IEEE CS Press, 2001, pp. 127-131.

9. R. Agarwal and A.R. Rao, "Data Reduction Techniques"; http://www.iasri.res.in/ebook/EB_SMAR/e-book_pdf%20files/Manual%20II/9-data_reduction.pdf [Προσπέλαση: 01/08/2011]

10. J. E. Jackson, A User's Guide to Principal Components, J. Wiley & sons, New York, 1991

11. R. L. Burden and D. J. Faires, Numerical Analysis, 8th Edition, Thomsom Brooks / Cole, Australia, 2005.

12. I. H. Witten and E. Frank: "Data Mining, Practical Machine Learning Tools and Techniques, Second Edition", Morgan Kaufmann Publishers, 2005.

13. J. R. Quinlan: "Induction of Decision Trees", Machine Learning, v.1 n.1, p.81-106, 1986.

14. J. R. Quinlan, "Learning with Continuous Classes", Proceedings of the 5th Australian Joint Conference on Artificial Intelligence, 1992, pp. 343–348

15. W. Iba and P. Langley, "Induction of One-Level Decision Trees", Proceedings of the 9th International Conference on Machine Learning, pp. 233-240, Morgan Kaufmann Publishers, 1992

Ανάπτυξη Αυτοματοποιημένου Διαγνωστικού Λογισμικού για Εξέταση Μαγνητικής Τομογραφίας Μαστών

Αντώνιος Δανελάκης
a.danelakis@gmail.com

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Οι παθολογίες μαστού παρουσιάζονται συχνά στις γυναίκες σε παγκόσμια κλίμακα. Η έγκαιρη διάγνωση επιτρέπει την ορθή θεραπεία. Για το λόγο αυτό, επιστρατεύονται ακτινοδιαγνωστικές απεικονιστικές μέθοδοι με τη Μαγνητική Τομογραφία (MRI) να είναι η πιο διαδεδομένη. Αυτό συμβαίνει, διότι παράγονται εικόνες υψηλής χωρικής διακριτικής ικανότητας και ευαισθησίας, χωρίς την επιβάρυνση του ασθενούς με ιοντίζουσα ακτινοβολία. Στα πλαίσια της παρούσας εργασίας, αναπτύχθηκε λογισμικό, το οποίο προσομοιώνει την κλινική μεθοδολογία των ακτινολόγων προκειμένου να εντοπίσει ύποπτες περιοχές του μαστού. Το παρόν λογισμικό διευκολύνει και επιταχύνει τη δουλειά των ακτινολόγων και λειτουργεί με το μικρότερο δυνατό ρίσκο διάγνωσης, όπως επιβάλλουν οι διεθνείς οδηγίες. Τέλος, η εν λόγω εφαρμογή, υποβλήθηκε σε κλινικές δοκιμές ανιχνεύοντας επιτυχώς όλες τις παθολογικές, βάση βιοψίας, περιπτώσεις.

Λέξεις-Κλειδιά: Μαγνητική τομογραφία, Παθολογία μαστού, Περιοχή ενδιαφέροντος, Καμπύλη φωτεινότητας, Χρωματικός Χάρτης

Επιβλέποντες: Θεοχάρης Θεοχάρης, Καθηγητής
Δρ. Δημήτριος Βεργανελάκης

1. Εισαγωγή

Οι παθήσεις μαστού είναι πολύ συχνά εμφανιζόμενες σε παγκόσμια κλίμακα, κυρίως στις γυναίκες λόγω ανατομίας. Ο αριθμός εμφανίσεων τους αυξάνει σταθερά, με τον καρκίνο να είναι ο κυριότερος εκφραστής τους [29]. Παρά ταύτα, η έγκαιρη διάγνωση μπορεί να οδηγήσει στην αποτελεσματική θεραπεία [28]. Για το λόγο αυτό, επιστρατεύονται ακτινοδιαγνωστικές απεικονιστικές μέθοδοι. Η χρήση της κλασικής μαστογραφίας και του υπερήχου (US) αποτελεί τη συνήθη μεθοδολογία. Η Μαγνητική Τομογραφία (MRI) λειτουργεί συμπληρωματικά στις παραπάνω μεθόδους [13]. Η MRI είναι πολύ διαδεδομένη, διότι παράγονται εικόνες υψηλής χωρικής διακριτικής ικανότητας, είναι πιο ευαίσθητη από το US, όπως και από κάθε άλλη απεικονιστική τεχνική [25, 26, 27] και δεν επιβαρύνει τον ασθενή με ιοντίζουσα ακτινοβολία.

Η συνεισφορά της MRI μαστών στον εντοπισμό παθολογιών του μαστού είναι κοινώς παραδεκτή και αδιαμφισβήτητη [23, 24]. Επίσης, με τις νέες τεχνολογίες που ενσωματώνονται, επιτυγχάνεται ακόμα μεγαλύτερη ευαισθησία στην παραγόμενη εικόνα [21, 22]. Δυστυχώς, πάντα υπάρχουν περιστατικά που η MRI δεν μπορεί να βοηθήσει [20], όμως πρόκειται για πολύ σπάνιες περιπτώσεις.

Η MRI παράγει ένα πλήθος παλμοσειρών (σαρώσεων) πριν και μετά την έγχυση σκιαγραφικού υλικού. Κάθε σάρωση αποτελείται από ένα πλήθος τομογραφικών εικόνων (τομές). Αν S είναι το πλήθος των παραγόμενων σαρώσεων και T το πλήθος των τομών κάθε σάρωσης, τότε S*T είναι το πλήθος των εικόνων που παράγονται. Ο ακτινολόγος επεξεργάζεται, μία προς μία, όλες τις εικόνες προκειμένου να διαγνώσει παθολογικές περιοχές του μαστού. Το τελευταίο, καθιστά τη δουλειά του ακτινολόγου κουραστική και χρονοβόρα. Ο σκοπός του αναπτυχθέντος λογισμικού είναι η αυτοματοποίηση της ακτινοδιαγνωστικής διαδικασίας. Η εφαρμογή δέχεται σαν είσοδο όλες τις σαρώσεις μιας τομής και παράγει στην έξοδο ένα σύνολο παθολογικά ύποπτων περιοχών. Με τον τρόπο αυτό, η δουλειά του ακτινολόγου καθίσταται ευκολότερη, γρηγορότερη και ακριβέστερη.

Η δομή της εργασίας έχει ως εξής. Στην παράγραφο 2 αναφέρονται τα τεχνικά χαρακτηριστικά των εικόνων MRI. Στην παράγραφο 3 αποτυπώνεται η μεθοδολογία ανίχνευσης ύποπτων περιοχών που χρησιμοποιούν οι ακτινολόγοι, ενώ στην παράγραφο 4 παρουσιάζεται η εφαρμογή και το πώς αυτή προσομοιώνει την εν λόγω μεθοδολογία. Στην παράγραφο 5 παρουσιάζονται τα αποτελέσματα κλινικών δοκιμών. Στην παράγραφο 6 συγκρίνεται η αναπτυχθείσα εφαρμογή με το εδραιωμένο λογισμικό της αγοράς. Η παράγραφος 7 αποτελεί τη σύνοψη και τα συμπεράσματα της παρούσας εργασίας. Τέλος, στη παράγραφο 7 παρουσιάζονται και κάποιες ενδιαφέρουσες μελλοντικές προκλήσεις.

2. Τεχνικά χαρακτηριστικά εικόνων MRI

Το πρωτόκολλο της MRI μαστογραφίας αποτελείται από πλήθος τύπων παλμοσειρών [13]. Οι παλμοσειρές λαμβάνονται πριν και μετά την έγχυση σκιαγραφικού υλικού [14, 15], το οποίο προσφέρει διαγνωστικά οφέλη [16, 17, 18] και είναι 8 σε σύνολο. Αυτού του τύπου οι παλμοσειρές αποτελούν τη δυναμική πληροφορία της MRI. Η MRI αποθηκεύει όλη την πληροφορία για κάθε τομή, σε αρχεία τύπου dicom [19].

Οι εικόνες που χρησιμοποιήθηκαν στην παρούσα εργασία είναι εικόνες δυναμικής πληροφορίας της MRI προβεβλημένες στο οβελιαίο επίπεδο προβολής. Είναι διαστάσεων 256 x 256 και κάθε εικονοστοιχείο έχει βάθος 8 (λαμβάνει τιμές φωτεινότητας στο ακέραιο διάστημα 28 = 256 στοιχείων [0, 255]). Η τιμή 0 αντιστοιχίζεται στο μαύρο χρώμα, ενώ η τιμή 255 στο λευκό. Ενδιάμεσες κλίμακες του γκρι αντιστοιχίζονται σε ενδιάμεσες τιμές [2].

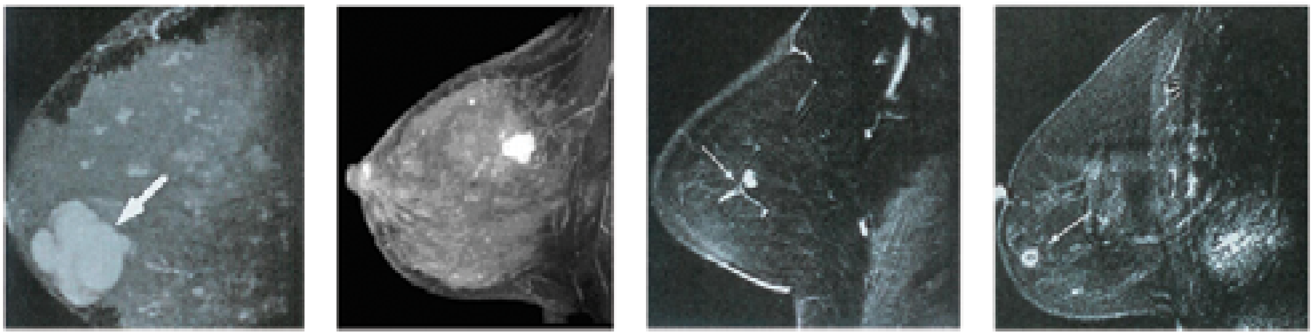
Το λογισμικό κανονικοποιεί τις τιμές φωτεινότητας στο διάστημα [0, 1], με το 0 να αντιστοιχίζεται στο μαύρο και το 1 στο λευκό χρώμα. Ενδιάμεσες τιμές απεικονίζουν ενδιάμεσες αποχρώσεις του γκρι. Η κανονικοποίηση γίνεται προκειμένου να αποφευχθούν σφάλματα στρογγύλευσης.

3. Κλινική μεθοδολογία ανίχνευσης παθολογιών μαστού

Η κλινική μεθοδολογία ανίχνευσης παθολογιών στο μαστό, χρησιμοποιώντας την MRI μαστού, χωρίζεται σε 3 βήματα [3]. Η μεθοδολογία αυτή πρέπει να εφαρμοστεί από τον ακτινολόγο σε όλες τις τομές κάθε MRI σάρωσης.

Βήμα 1: Φίλτρο φωτεινότητας

Κατά το βήμα αυτό, ο ακτινολόγος αναγνωρίζει με το μάτι ύποπτες περιοχές στο μαστό. Αυτές οι περιοχές, λόγω της αυξημένης ροής αίματος [12] και της χρήσης σκιαγραφικού υλικού παράγουν υψηλό σήμα εξόδου στην MRI, άρα χρώμα υψηλής τιμής φωτεινότητας στην εικόνα. Μερικά παραδείγματα παρουσιάζονται στην Εικόνα 1 [3]. Ο ειδικός εστιάζει σε αυτές ακριβώς τις περιοχές υψηλής φωτεινότητας τις οποίες και επιλέγει με το χέρι [11] (ανθρώπινο σφάλμα). Αυτές οι περιοχές θα ονομάζονται, υποψήφιες περιοχές παθολογίας.

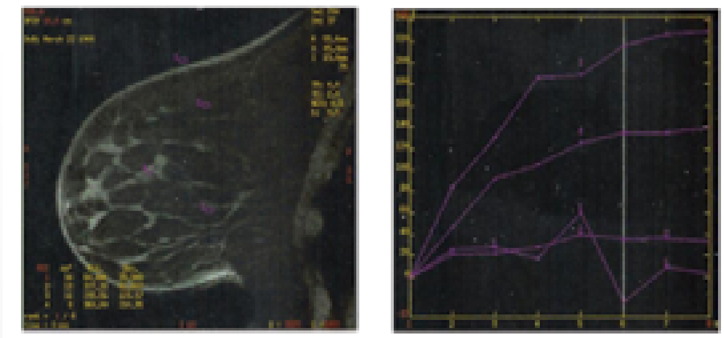


Εικόνα 1: Από αριστερά προς τα δεξιά: Ινοαδένωμα, Όγκος, Λεμφαδένας, Κύστες

Βήμα 2: Φίλτρο κλίσης καμπύλης φωτεινότητας

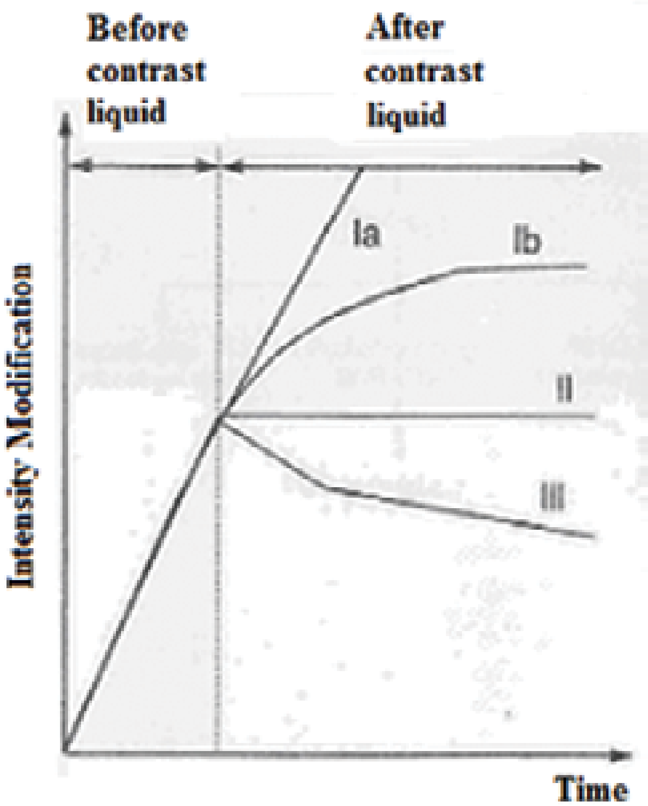
Η MRI παράγει ένα πλήθος σαρώσεων. Η πρώτη σάρωση παράγεται χωρίς χρήση σκιαγραφικού υλικού, ενώ οι υπόλοιπες 7 με χρήση σκιαγραφικού [3]. Η πρώτη σάρωση θα καλείται σάρωση αναφοράς. Κάθε σάρωση περιγράφει την κατάσταση του μαστού σε διαφορετικές χρονικές στιγμές.

Κατά το τρέχον βήμα, κατασκευάζονται οι καμπύλες φωτεινότητας (εικόνα 2) των υποψήφιων περιοχών παθολογίας. Η καμπύλη φωτεινότητας περιγράφει τη μεταβολή της φωτεινότητας της περιοχής, σε σχέση με τη φωτεινότητα στη σάρωση αναφοράς, στο χρόνο (ή κατά μήκος των σαρώσεων).



Εικόνα 2: Αριστερά: Υποψήφιες περιοχές, Δεξιά: Αντίστοιχες καμπύλες φωτεινότητας

Εν γένει, υπάρχουν 3 τύπων καμπύλες (εικόνα 3) [9]. Ο Τύπος Ι είναι μορφή καμπύλης σταθερής ενίσχυσης, ο Τύπος ΙΙ είναι μορφή στατικής καμπύλης, ενώ ο Τύπος ΙΙΙ είναι μορφή καμπύλης σταθερής εξασθένισης. Η πιθανότητα παθολογίας της υποψήφιας περιοχής παρουσιάζεται στον πίνακα 1.



Εικόνα 3: Τρεις τύποι καμπύλων φωτεινότητας

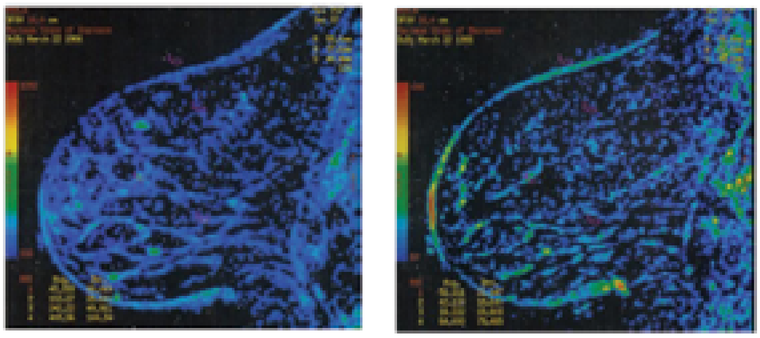
Τύπος Καμπύλης	Πιθανότητα Παθολογίας
Τύπος I (Ia, Ib)	8.9%
Τύπος II	33.6%
Τύπος III	57.4%

Πίνακας 1: Πιθανότητα παθολογίας ανάλογα τον τύπο καμπύλης φωτεινότητας [10]

Ο ακτινολόγος συγκεντρώνεται στις καμπύλες Τύπου III, το οποίο κλινικά μεταφράζεται σε 3 συνεχόμενες πτώσεις της τιμής της φωτεινότητας της καμπύλης, μεταξύ διαδοχικών σαρώσεων από την 3η σάρωση και μετά [8]. Οι υποψήφιες περιοχές παθολογίας, που αντιστοιχίζονται σε τέτοιου είδους καμπύλες, θα ονομάζονται ύποπτες περιοχές παθολογίας. Ιατρικά, τα παραπάνω μεταφράζονται στο ότι οι περιοχές αυτές μεταβολίζουν γρήγορα τις θρεπτικές ουσίες του σκιαγραφικού υλικού, που αποτελεί ιδιότητα των παθολογικών ιστών.

Βήμα 3: Φίλτρο χρωματικού φράγματος

Στο βήμα αυτό, ο ακτινολόγος χρησιμοποιεί τους χρωματικούς χάρτες κλίσης ανόδου/καθόδου (εικόνα 4). Αυτοί κατασκευάζονται από υπάρχοντα λογισμικά και αντιστοιχίζουν περιοχές του μαστού σε χρώματα, ανάλογα με την κλίση των καμπύλων φωτεινότητας τους. Πιο συγκεκριμένα, για το χάρτη κλίσεων ανόδου, όσο πιο θερμό είναι το χρώμα που αντιστοιχίζεται στην περιοχή, τόσο πιο γρήγορα ενισχύεται η φωτεινότητα της περιοχής κατά μήκος των σαρώσεων. Το κόκκινο χρησιμοποιείται στη μέγιστη ενίσχυση και το μπλε στη μηδενική. Αντίστοιχα ισχύουν και για το χάρτη κλίσεων καθόδου και την εξασθένιση της καμπύλης, με το κόκκινο να αντιστοιχεί στη μέγιστη εξασθένιση και το μπλε στη μηδενική. Άρα, ο ακτινολόγος επικεντρώνεται στις ύποπτες περιοχές που καλύπτονται με θερμά χρώματα στους χρωματικούς χάρτες. Αυτές οι περιοχές θα καλούνται επικίνδυνες περιοχές παθολογίας.



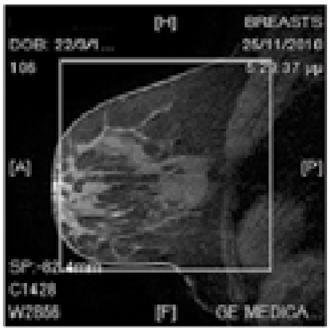
Εικόνα 4: Αριστερά: Χρωματικός χάρτης κλίσεων ανόδου
Δεξιά: Χρωματικός χάρτης κλίσεων καθόδου

4. Αυτοματοποίηση μεθοδολογίας ανίχνευσης

Τα τελευταία χρόνια, γίνεται μεγάλη προσπάθεια για την εκμετάλλευση των πανίσχυρων υπολογιστικών συστημάτων, τόσο για την ανάπτυξη της μηχανικής υποστήριξης ιατρικής απόφασης (Clinical Decision Support Systems) [4, 5], όσο και για την ανάπτυξη μηχανικής ιατρικής διάγνωσης (Computer Aided Design) [6]. Ειδικά η βελτίωση τεχνικών μηχανικής διάγνωσης, μπορεί να έχει επαναστατικές συνέπειες στη σύγχρονη ιατρική [7]. Στην παράγραφο που ακολουθεί θα αναλυθεί η εφαρμογή και το πώς αυτή προσομοιώνει τη διαδικασία που περιγράφηκε στην προηγούμενη παράγραφο.

Βήμα 0: Αρχικοποίηση

Αρχικά φορτώνονται τα 8 αρχεία dicom που αντιστοιχούν στις 8 σαρώσεις μιας συγκεκριμένης τομής. Η εφαρμογή, έπειτα, παρέχει την επιλογή της παραθύρωσης. Η παραθύρωση εκφράζεται μέσω μιας ακεραίας παραμέτρου w, που καθορίζει το πλήθος των εικονοστοιχείων που δεν θα ληφθούν υπόψη από την κορυφή, τον πάτο, τα δεξιά και τα αριστερά της μαστογραφικής εικόνας. Ένα παράδειγμα παρουσιάζεται στην Εικόνα 5. Με τη χρήση της παραθύρωσης, οι αλγόριθμοι περιορίζονται μόνο στο ωφέλιμο μέρος της εικόνας και ως εκ τούτου εκτελούνται ταχύτερα και ακριβέστερα.



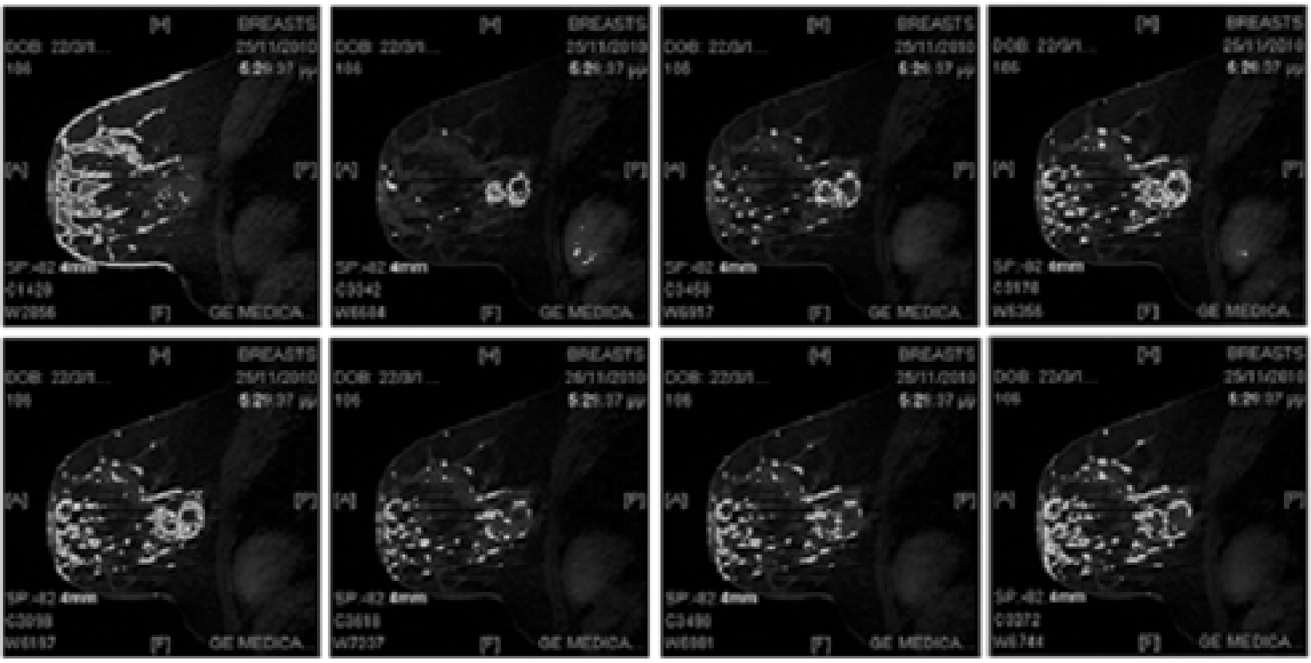
Εικόνα 5: Εφαρμογή παραθύρωσης με w = 45

Βήμα 1: Φίλτρο φωτεινότητας

Προκειμένου να προσομοιωθεί το φίλτρο φωτεινότητας, η εφαρμογή διαβάζει από το χρήστη μια αριθμητική τιμή, την prc, η οποία δηλώνει ένα φράγμα φωτεινότητας. Έπειτα, για την τρέχουσα τομή των 8 σαρώσεων περικλείονται από την εφαρμογή οι περιοχές, των οποίων η μέση φωτεινότητα υπερβαίνει την τιμή I_{thres}, όπου:

$$I_{thres} = I_{avg} + prc * I_{avg}$$

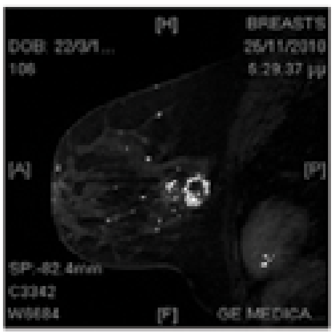
Η τιμή I_{avg} δηλώνει τη μέση φωτεινότητα της τρέχουσας MRI τομής. Οι εν λόγω περιοχές είναι οι υποψήφιες περιοχές παθολογίας και προβάλλονται στην οθόνη της εφαρμογής (Εικόνα 6). Ο χρήστης μπορεί να πειραματιστεί με την τιμή prc έως ότου να φτάσει σε οικείο οπτικό αποτέλεσμα. Με τον τρόπο αυτό, αποφεύγεται το ανθρώπινο σφάλμα της χειρωνακτικής επιλογής των υποψήφιων περιοχών.



Εικόνα 6: Υποψήφιες περιοχές παθολογίας για κάθε μια από τις 8 σαρώσεις της MRI (prc = 2.5)

Βήμα 2: Φίλτρο κλίσης καμπύλης φωτεινότητας

Έπειτα από την επιβεβαίωση των υποψήφιων περιοχών, ακολουθεί η προσομοίωση του φίλτρου κλίσης καμπύλης φωτεινότητας. Στο σημείο αυτό κατασκευάζονται οι καμπύλες φωτεινότητας όλων των υποψήφιων περιοχών παθολογίας και, ως ύποπτες περιοχές κρατούνται μόνο εκείνες που παρουσιάζουν 3 συνεχόμενες πτώσεις στην καμπύλη τους από την 3η σάρωση και έπειτα [3]. Οι εν λόγω περιοχές προβάλλονται στην οθόνη της εφαρμογής (Εικόνα 7). Η διαδικασία αυτή γίνεται αυτόματα, χωρίς να χρειάζεται ο ακτινολόγος να μελετήσει μία προς μία τις καμπύλες.

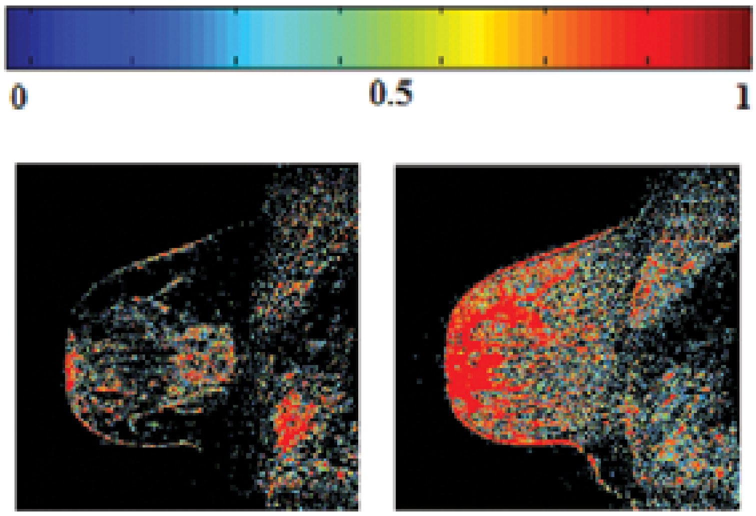


Εικόνα 7: Ύποπτες περιοχές παθολογίας

Βήμα 3: Φίλτρο χρωματικού φράγματος

Μετά την επιβεβαίωση των ύποπτων περιοχών παθολογίας, ακολουθεί η προσομοίωση του φίλτρου χρωματικού φράγματος. Η εφαρμογή κατασκευάζει τους χρωματικούς χάρτες κλίσεων ανόδου/καθόδου (Εικόνα 8 κάτω).

Για το χάρτη κλίσεων ανόδου αντιστοιχίζεται το κόκκινο χρώμα στη μέγιστη κλίση και το μπλε στην μηδενική. Τα ενδιάμεσα χρώματα αντιστοιχίζονται με τη μέθοδο της γραμμικής παρεμβολής [2] στην παλέτα ουράνιου τόξου (Εικόνα 8 πάνω). Τα αντίστοιχα ισχύουν για το χάρτη κλίσεων καθόδου.



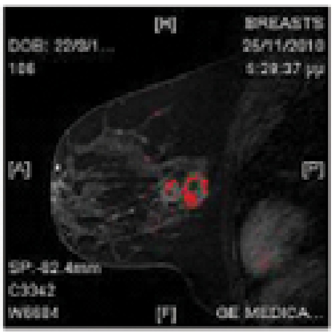
Εικόνα 8: Πάνω: Σύνδεση τιμών Tc και χρωμάτων μέσω της παλέτας ουράνιου τόξου
Κάτω αριστερά: Χρωματικός χάρτης κλίσεων ανόδου, Color map of increase slope
Κάτω δεξιά: Χρωματικός χάρτης κλίσεων καθόδου

Έπειτα, ο χρήστης εισάγει την αριθμητική τιμή T_c, η οποία εκφράζει το χρωματικό φράγμα. Ανήκει στο διάστημα [0, 1] και κάθε τιμή αντιστοιχίζεται σε ένα χρώμα σύμφωνα με την παλέτα ουράνιου τόξου (το 0 στο μπλε και το 1 στο κόκκινο). Για ενδιάμεσα χρώματα χρησιμοποιείται πάλι η γραμμική παρεμβολή.

Για κάθε ύποπτη περιοχή παθολογίας, ελέγχεται αν είναι αληθής ο παρακάτω τύπος:

$$T_c < T_{color\_map\_increase} \text{ AND } T_c < T_{color\_map\_decrease}$$

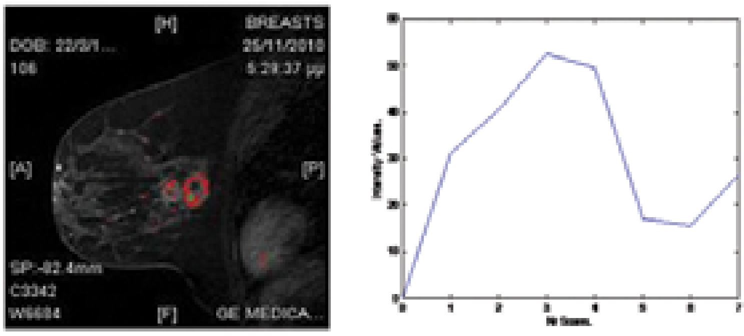
Όπου $T_{color\_map\_increase}$ είναι η βαθμωτή μεταβλητή στην οποία αντιστοιχίζεται ο χρωματικός μέσος όρος της τρέχουσας ύποπτης περιοχής παθολογίας στο χάρτη ανόδου. Αντίστοιχα η $T_{color\_map\_decrease}$ για το χάρτη καθόδου. Αν ο παραπάνω τύπος είναι αληθής, τότε η τρέχουσα ύποπτη περιοχή χαρακτηρίζεται ως επικίνδυνη περιοχή παθολογίας. Η διαδικασία γίνεται αυτόματα και δεν χρειάζεται ο ακτινολόγος να το κάνει με το μάτι. Τελικά, οι επικίνδυνες περιοχές παθολογίας, προβάλλονται στην οθόνη της εφαρμογής (Εικόνα 9).



Εικόνα 9: Επικίνδυνες περιοχές παθολογίας με T_c = 0.4

Βήμα 4: Έλεγχος καμπύλων φωτεινότητας

Τέλος, ο χρήστης μπορεί να επιλέξει μια επικίνδυνη περιοχή παθολογίας και να δει την καμπύλη φωτεινότητας που της αντιστοιχεί. Στο αριστερό τμήμα της Εικόνας 10 έχει επιλεγεί η επικίνδυνη περιοχή με το πράσινο περίβλημα και στο δεξί παρουσιάζεται η καμπύλη φωτεινότητάς της.

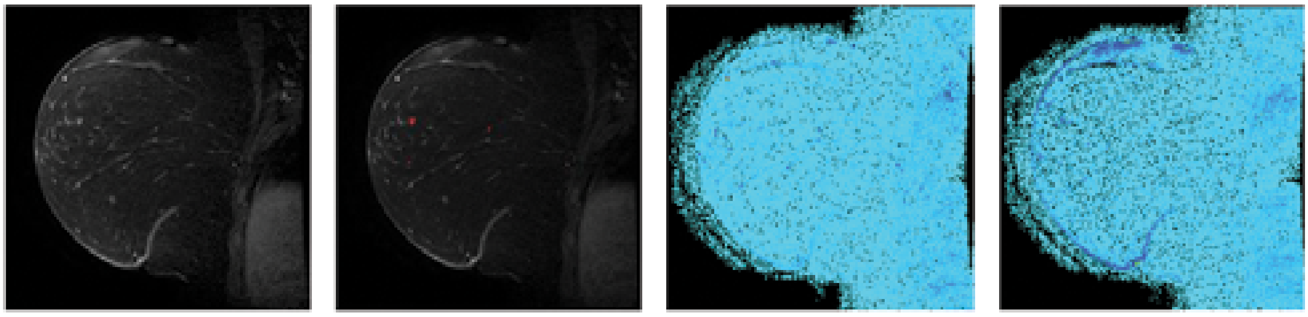


Εικόνα 10: Αριστερά: Επιλεγμένες επικίνδυνες περιοχές (πράσινος κύκλος)
Δεξιά: Αντίστοιχη καμπύλη φωτεινότητας

5. Αποτελέσματα κλινικών δοκιμών

1η Περίπτωση: Υγιής μαστός

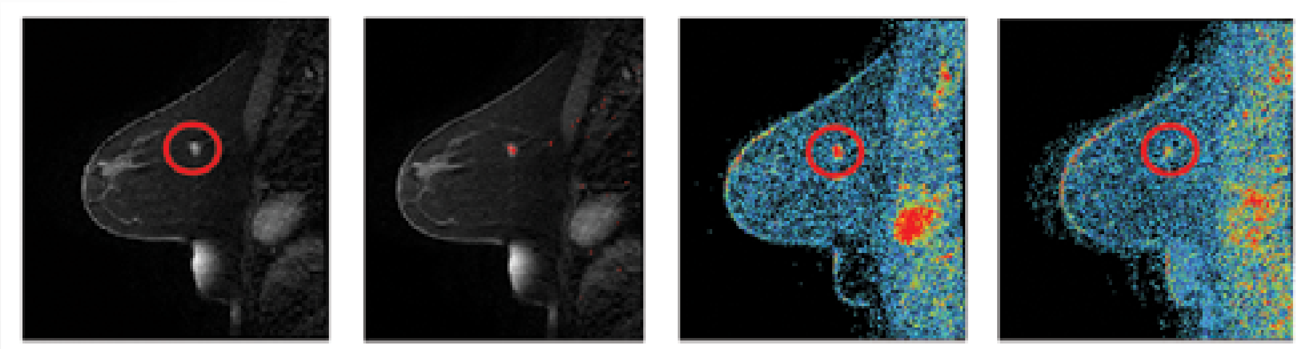
Στις εικόνες που ακολουθούν παρουσιάζονται οι επικίνδυνες περιοχές παθολογίας του μαστού, όπως προκύπτουν από την εφαρμογή (με κόκκινο χρώμα), καθώς και οι χρωματικοί χάρτες. Ο ακτινολόγος θα διαπιστώσει ότι οι περιοχές που δίνει η εφαρμογή είναι μόνο 2, και αυτές μικρές και αραιές, που αντιστοιχίζονται σε ψυχρά χρώματα στους χρωματικούς χάρτες, καταλήγοντας στο ορθό συμπέρασμα ότι πρόκειται για υγιή μαστό.



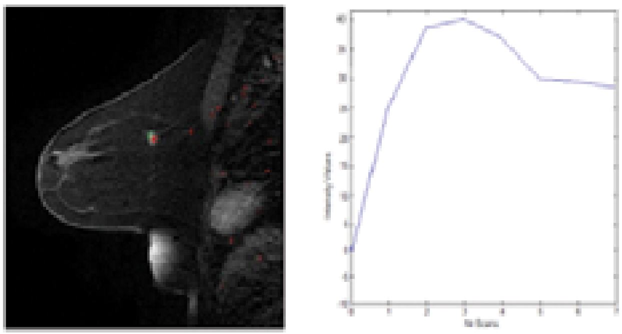
Εικόνα 11: Από αριστερά προς τα δεξιά: Αρχική MRI εικόνα. Επικίνδυνες περιοχές παθολογίας (w=40, prc=2.5, Tc=0). Χρωματικός χάρτης κλίσεων ανόδου. Χρωματικός Χάρτης κλίσεων καθόδου.

2η Περίπτωση: Μαστός με κακοήθεια

Στην περίπτωση αυτή, ο ακτινολόγος θα διαπιστώσει ότι υπάρχει μια μεγάλη επικίνδυνη περιοχή παθολογίας σε συγκεκριμένη θέση του μαστού που δεν δικαιολογεί την ύπαρξη της. Η αντίστοιχη περιοχή καλύπτεται από πολύ θερμά χρώματα στους χρωματικούς χάρτες, πράγμα που θα κάνει τον ακτινολόγο να καταλήξει στο συμπέρασμα ότι η εν λόγω περιοχή χρίζει περαιτέρω εξέτασης.



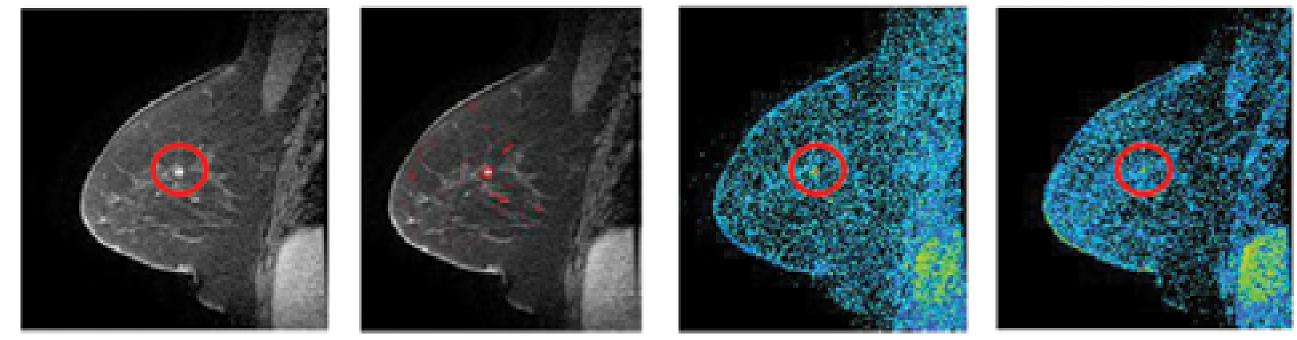
Εικόνα 12: Από αριστερά προς τα δεξιά: Αρχική MRI εικόνα (η κακοήθεια σημειώνεται με κόκκινο κύκλο). Επικίνδυνες περιοχές παθολογίας (w=70, prc=2.5, Tc=0.925). Χρωματικός χάρτης κλίσεων ανόδου. Χάρτης κλίσεων καθόδου.



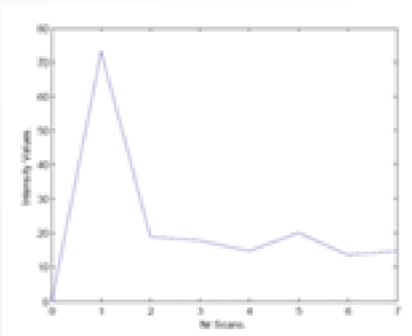
Εικόνα 13: Αριστερά: Επιλεγείσα επικίνδυνη περιοχή (πράσινος κύκλος)
Δεξιά: Αντίστοιχη καμπύλη φωτεινότητας.

<u>3η Περίπτωση: Μαστός με Καλοήθεια</u>

Και σε αυτή την περίπτωση, ο ακτινολόγος θα διαπιστώσει ότι υπάρχει μια μεγάλη επικίνδυνη περιοχή παθολογίας σε συγκεκριμένη θέση του μαστού που δεν δικαιολογεί την ύπαρξη της. Η αντίστοιχη περιοχή καλύπτεται από πολύ θερμά χρώματα στους χρωματικούς χάρτες πράγμα που θα κάνει τον ακτινολόγο να καταλήξει στο συμπέρασμα ότι η εν λόγω περιοχή χρίζει περαιτέρω εξέτασης.



Εικόνα 14: Από αριστερά προς τα δεξιά: Αρχική MRI εικόνα (η καλοήθεια σημειώνεται με κόκκινο κύκλο). Επικίνδυνες περιοχές παθολογίας (w=60, prc=2.5, Tc=0.4). Χρωματικός χάρτης κλίσεων ανόδου. Χρωματικός Χάρτης κλίσεων καθόδου.

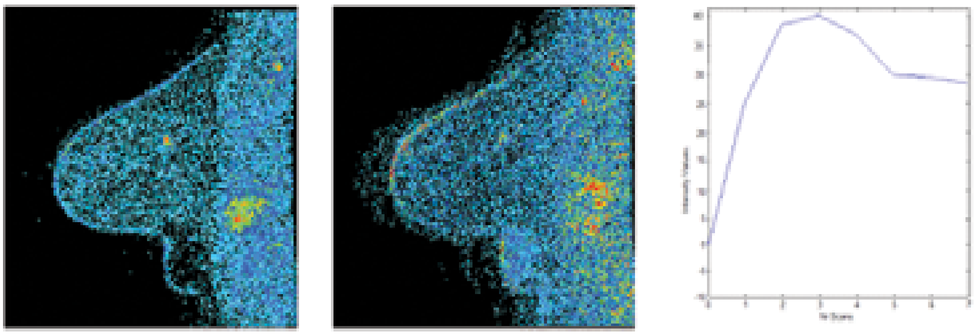


Εικόνα 15: Αντίστοιχη καμπύλη φωτεινότητας

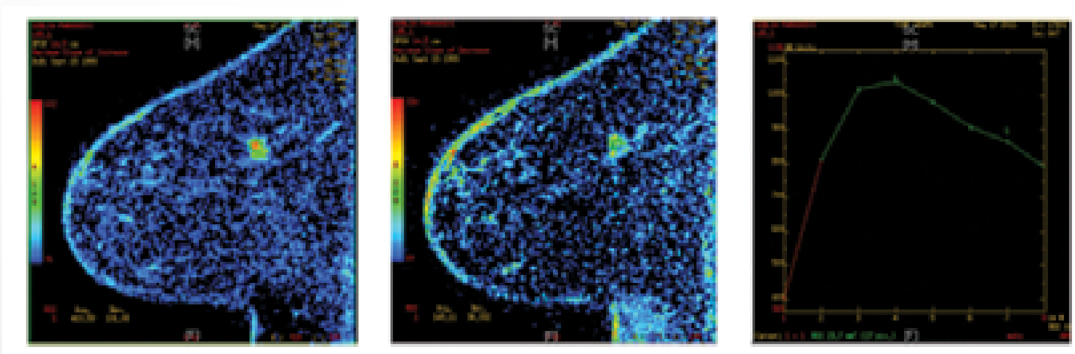
6. Σύγκριση με πιστοποιημένα λογισμικά

Στην παράγραφο αυτή συγκρίνονται τα αποτελέσματα της αναπτυχθείσας εφαρμογής, με τα αποτελέσματα εφαρμογών, πιστοποιημένων από τον (FDA). Η σύγκριση παρουσιάζεται στις Εικόνες 16 και 17. Οι καμπύλες φωτεινότητας αντιστοιχούν στην επιλεγμένη περιοχή του μαστού που παρουσιάζεται στην Εικόνα 18.

Η ελαφριά διαφορά στις καμπύλες φωτεινότητας οφείλεται σε 2 λόγους. Πρώτον, τα δύο λογισμικά δουλεύουν σε διαφορετικά διαστήματα φωτεινότητας. Το παρόν δουλεύει στο διάστημα [0, 1], ενώ τα πιστοποιημένα δουλεύουν στο [0, 255] [1]. Δεύτερον, το αναπτυχθέν λογισμικό περικλείει ακριβώς την ύποπτη περιοχή και μόνο αυτή. Αντίθετα, στα πιστοποιημένα λογισμικά, την ύποπτη περιοχή την επιλέγει ο χρήστης με το χέρι, επιλέγοντας και εικονοστοιχεία που δεν θα έπρεπε (λόγω της κίνησης του χεριού) τα οποία όμως, συνεισφέρουν στη δημιουργία της καμπύλης.



Εικόνα 16: Αποτελέσματα αναπτυχθέντος λογισμικού



Εικόνα 17: Αποτελέσματα λογισμικών της αγοράς

Εικόνα 18: Περιοχή αντιστοίχισης στις καμπύλες φωτεινότητας των εικόνων 16, 17 (πράσινος κύκλος)

7. Συμπεράσματα και μελλοντικές προκλήσεις

Παρουσιάστηκε αναπτυχθέν ιατρικό λογισμικό που προσομοιώνει την κλινική μεθοδολογία των ακτινολόγων για τον εντοπισμό ύποπτων περιοχών του μαστού στην MRI μαστού. Το λογισμικό διευκολύνει, επιταχύνει και κάνει ακριβέστερη τη δουλειά των ακτινολόγων. Όπως επιβάλλεται από τις διεθνείς οδηγίες, λειτουργεί με το μικρότερο δυνατό ρίσκο διάγνωσης. Τέλος, η εν λόγω εφαρμογή, υποβλήθηκε σε κλινικές δοκιμές ανιχνεύοντας επιτυχώς όλες τις παθολογικές, βάση βιοψίας, περιπτώσεις.

Η βελτίωση του αναπτυχθέντος λογισμικού, με χρήση στατιστικών και αναλυτικών μεθόδων, για την ελαχιστοποίηση των ψευδώς επικίνδυνων περιοχών, αποτελεί μια πολύ ενδιαφέρουσα πρόκληση για το μέλλον. Τέλος, σημαντική μελλοντική πρόκληση, αποτελεί η εκμετάλλευση μεθόδων Αναγνώρισης Προτύπων, προκειμένου να επιτευχθεί μηχανική διάγνωση υψηλής αξιοπιστίας.

Αναφορές

1. Edwin J.R. van Beek and Eric A. Hoffman, “Functional Imaging: CT and MRI”, Clin. Chest Med., 29(1): 195 – vii, 2008.

2. T. Theoxaris, G. Papaioannou, N. Platis and N. M. Patrikalakis, Graphics & Visualization:

Prinsiples & Algorithms, A. K. Peters, 2007.

3. Elizabeth A. Morris and Laura Liberman, Breast MRI: Diagnosis & Intervention, Springer, 2005.

4. Eta S. Berner, Clinical Decision Support Systems: State of the Art, AHRQ Publications, 2009.

5. J. Valach and F. Lehocki, "Clinical Decision Support Systems", MEDITECH - Proceedings of the ESF project conference, Bratislava, Slovakia, 26 May 2008.

6. Nicholas Petrick et al., "Breast Cancer Detection: Evaluation of a Mass-Detection Algorithm for Computer Aided Diagnosis - Experience in 263 Patients", Radiology, 224: 217 – 224, 2002.

7. Gromet Matthhew, "Comparison of Computer-Aided Detection to Double Reading of Screening Mammograms: Review of 231,221 Mammograms", AJR, 190: 854 – 859, 2008.

8. Schnall M.D., Blume J, Bluemke D.A., DeAngelis G.A., DeBruhl N., Harms S., Heywang-Kobrunner S.H., Hylton N., Kuhl C.K., Pisano E.D., Causer P., Schnitt S.J., Thickman D., Stelling C.B., Weatherall P.T., Lehman C. and Gatsonis C.A., "Diagnostic architectural and dynamic features at breast MR imaging: multicenter study", Radiology, 238: 42 – 53, 2006.

9. Kuhl C.K. and Schild H.H., "Dynamic image interpretation of MRI of the breast", J Magn Reson Imagin,.12: 965 – 974, 2000.

10. Kuhl C.K., Mielcarek P., Klaschik S., et al., "Are signal time course data useful for differential diagnosis of enhancing lesions in dynamic breast MR imaging?", Radiology, 211: 101 – 110, 1999.

11. Mussurakis S., Gibbs P. and Horsman A., "Primary breast abnormalities: selective pixel sampling on dynamic gadolinium enhanced MR images", Radiology, 206: 465 – 473, 1998.

12. Weidner N., Semple J.P.,Welch W.R. and Folkman J., "Tumor angiogenesis and metastasis—correlation in invasive breast carcinoma", N. Engl. J. Med., 324: 1 – 8, 1991.

13. Mann R.M., Kuhl C. K., Kinkel K. and Boetes M., "Breast MRI: Guidelines from the European Society of Breast Imaging", Eur. Radiol., 18: 1307 – 1318, 2008.

14. Heywang S.H., Wolf A., Pruss E., Hilbertz T., Eiermann W. and Permanetter W., "MR imaging of the breast with Gd-DTPA: Use and limitations", Radiology, 171: 95 – 103, 1989.

15. Kaiser W.A. and Zeitler E., "MR imaging of the breast: fast imaging sequences with and without Gd-DTPA. Preliminary observations", Radiology, 170: 681 – 686, 1989.

16. Heywang S.H., Fenzl G., Hahn D., Krischke I., Edmaier M., Eiermann W. and Bassermann R., "MR imaging of the breast: comparison with mammography and ultrasound", Journal of Comput. Assist. Tomogr., 10: 615 – 620, 1986.

17. Partain C.L., Kulkarni M.V., Price R.R., Fleischer A.C., Page D.L., Malcolm A.W., Winfield A.C. and James A.E., "Magnetic resonance imaging of the breast: functional T1 and three-dimensional imaging", Cardiovasc. Intervent. Radiol., 8: 292 – 299, 1986.

18. Heywang-Kobrunner S.H., Bick U., Bradley W.G. Jr, Bone B., Casselman J., Coulthard A.,

Fischer U., Muller-Schimpfle M., Oellinger H., Patt R., Teubner J., Friedrich M., Newstead G., Holland R., Schauer A., Sickles E.A., Tabar L., Waisman J. and Wernecke K.D., "International investigation of breast MRI: results of a multicentre study (11 sites) concerning diagnostic parameters for contrast-enhanced MRI based on 519 histopathologically correlated lesions", Eur. Radiol., 11: 531 – 546, 2001.

19. National Electrical Manufacturers Association, Digital Imaging and Communication in Medicine, 2009.

20. Shimauchi et al., "Breast Cancers not Detected at MRI: Review of False – Negative Lesions", American Journal of Roentgenology, 194: 1674 - 1679, 2009.

21. Haitham Elsamaloty et al., "Increasing Accuracy of Detection of Breast Cancer with 3-T MRI", American Journal of Roentgenology, 192: 1142 - 1148, 2009.

22. Van den Brink J.S.,Watanabe Y., Kuhl C.K., et al., "Implications of SENSE MR in routine clinical practice", Eur J Radiol.,46: 3 – 27, 2003.

23. Fischer U., Kopka L. and Grabbe E., "Breast carcinoma: effect of pre-operative contrast-enhanced MR imaging on the therapeutic approach", Radiology, 213: 881 – 888, 1999.

24. Rieber A., Zeitler H., Rosenthal H., et al., "MRI of breast cancer: influence of chemotherapy on sensitivity", Br J Radiol., 70: 452 – 458, 1997.

25. Francesko Sardanelli et al., "Sensitivity of MRI versus Mammography for Detecting Foci of Multifocal, Multicentric Breast Cancer in Fatty and Dense Breasts Using the Whole-Breast Pathologic Examination as a Gold Standard", American Journal of Roentgenology, 183: 1149 - 1157, 2004.

26. Peter L. Davis, Melinda J. Staiger, Kathleen B. Harris, Marie A. Ganott, Jolita Klementaviciene, Kenneth S. McCarty and Hector Tobon, "Breast cancer measurements with magnetic resonance imaging, ultrasonography, and mammography", Breast Cancer Research and Treatment, 37(1): 1-9, 1996.

27. Berg W.A., Gutierrez L., NessAiver M.S., Carter W.B., Bhargavan M., Lewis R.S. and Ioffe O.B., "Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer", Radiology, 233(3): 830 – 849, 2004.

28. Karellas and S. Vedantham, Breast cancer imaging: A perspective for the next decade, Department of Radiology, University of Massachusetts Medical School, October 2008.

29. John P. Fyssas, Breast and Pathologies: A manual for every woman, A. A. Livanis, 2006.

In Silico Πρόσδεση Εν Δυνάμει Αναστολέων της Φωσφορυλάσης του Γλυκογόνου με Χρήση του AutoDock για το Σχεδιασμό Αντιδιαβητικών Φαρμάκων

Χριστόφορος Γ. Ζαρκάδας
christoforos.zarkadas@gmail.com

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Ο σακχαρώδης διαβήτης είναι η τρίτη σπουδαιότερη, μη μεταδιδόμενη ασθένεια, παγκοσμίως και χαρακτηρίζεται από αυξημένη συγκέντρωση των επιπέδων της γλυκόζης στο αίμα πέρα από τα φυσιολογικά όρια. Το ένζυμο της φωσφορυλάσης του γλυκογόνου χρησιμοποιείται ως μοριακός στόχος για το σχεδιασμό νέων αναστολέων, εν δυνάμει αντιδιαβητικών φαρμάκων, με σκοπό τη ρύθμιση της γλυκαιμίας. Στην παρούσα εργασία έγινε αποτύπωση των προσεγγίσεων που εφαρμόζονται για την in silico μοριακή πρόσδεση, των πλέον διαδεδομένων λογισμικών που χρησιμοποιούνται γι' αυτήν, όπως και in silico πρόσδεση εν δυνάμει αναστολέων της φωσφορυλάσης του γλυκογόνου, με τη βοήθεια του λογισμικού AutoDock. Πραγματοποιήθηκε συγκριτική μελέτη των τεσσάρων αλγορίθμων του AutoDock, καθώς και σύγκριση του καλύτερου εξ αυτών με τα αποτελέσματα που προέκυψαν από α) τα πειραματικά δεδομένα) και β) την εφαρμογή του λογισμικού Glide για τους ίδιους προσδέτες.

Λέξεις-Κλειδιά: Φωσφορυλάση του γλυκογόνου, Μοριακή πρόσδεση, Σακχαρώδης διαβήτης, AutoDock, Κατευθυνόμενος - από τη δομή - σχεδιασμός φαρμάκων

Επιβλέποντες:
Ιωάννης Εμίρης, Καθηγητής
Δρ. Ευαγγελία Χρυσίνα

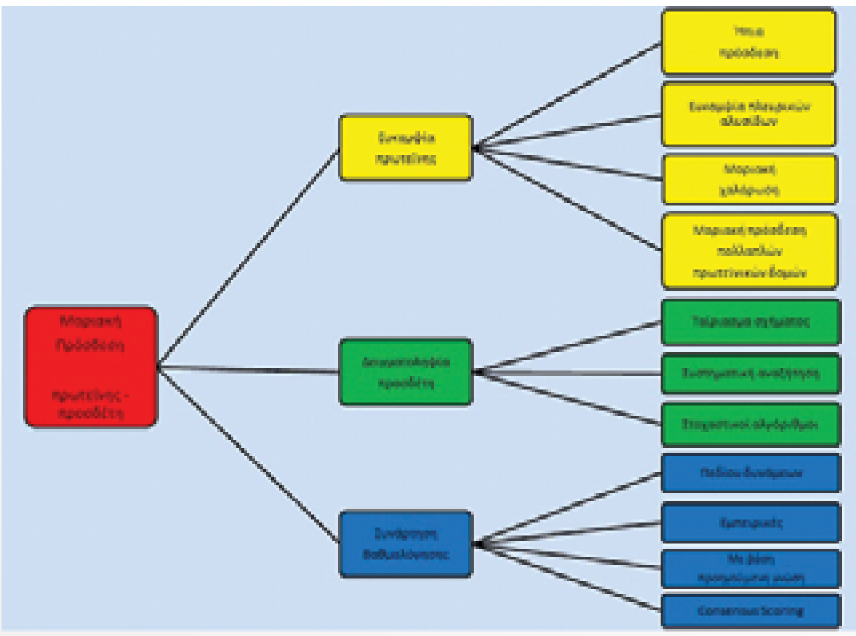
1. Μοριακή πρόσδεση

Οι διαδικασίες αυτές των in silico προσομοιώσεων σε επίπεδο ατόμων (μεθοδολογίες μοριακής πρόσδεσης) αφορούν στην εύρεση της ακριβούς διαμόρφωσης και προσανατολισμού του προσδέτη εντός μιας δοθείσας πρωτεΐνης-στόχου, της οποίας η δομή είτε είναι εκ των προτέρων γνωστή είτε μπορεί να υπολογιστεί. Οι υπολογιστικές μέθοδοι είναι επιθυμητές για την πρόβλεψη εν δυνάμει τρόπων πρόσδεσης, εξαιτίας των δυσκολιών και του οικονομικού κόστους των πειραματικών μεθόδων για τον καθορισμό της δομής των συμπλόκων. Κατά την προσομοίωση της μοριακής πρόσδεσης, χιλιάδες συνδυασμοί σύνδεσης δοκιμάζονται και εκτιμώνται. Η θέση και ο προσανατολισμός του προσδέτη με τη χαμηλότερη ενεργειακή βαθμολογία θεωρείται ως το καλύτερο ταίριασμα (best match), δηλαδή ο καλύτερος τρόπος πρόσδεσης.

Ένα πρόγραμμα μοριακής πρόσδεσης πρωτεΐνης-προσδέτη αποτελείται από δύο βασικές συνιστώσες: Τη δειγματοληψία (sampling) και τη βαθμολόγηση (scoring).

Η δειγματοληψία αναφέρεται στη δημιουργία εν δυνάμει προσανατολισμών/ διαμορφώσεων του προσδέτη κοντά σε ένα κέντρο σύνδεσης μιας πρωτεΐνης και μπορεί να χωριστεί σε δύο επιπλέον τμήματα: Τη δειγματοληψία του προσδέτη και την ευκαμψία της πρωτεΐνης.

Στο παρακάτω σχήμα φαίνεται μία λεπτομερέστερη κατηγοριοποίηση των μεθόδων που χρησιμοποιούνται στην προσομοίωση μοριακής πρόσδεσης.



Εικόνα 1: Κατηγοριοποίηση των μεθόδων που χρησιμοποιούνται στην προσομοίωση μοριακής πρόσδεσης [1]

2. Σακχαρώδης διαβήτης και Φωσφορυλάση του γλυκογόνου

Ο σακχαρώδης διαβήτης (ΣΔ) ανήκει στις λεγόμενες μη μεταδιδόμενες ασθένειες (non-communicable diseases, NCDs) και είναι μία νόσος που χαρακτηρίζεται από αύξηση της συγκέντρωσης της γλυκόζης στο αίμα πέρα από τα φυσιολογικά όρια με σοβαρές και χρόνιες επιπλοκές [2].

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (ΠΟΥ ή World Health Organization, WHO) είναι η τρίτη κυριότερη μη μεταδιδόμενη ασθένεια (μετά τις καρδιαγγειακές ασθένειες και τον καρκίνο).

Ο ΣΔ διακρίνεται σε δύο κατηγορίες: το σακχαρώδη διαβήτη τύπου 1 και το σακχαρώδη διαβήτη τύπου 2.

Ο σακχαρώδης διαβήτης τύπου 1 ή ινσουλινο-εξαρτώμενος σακχαρώδης διαβήτης (Insulin-Dependent Diabetes Mellitus, NIDDM) προκαλείται από αυτοάνοση καταστροφή των β-κυττάρων του παγκρέατος. Ο ασθενής είναι άμεσα εξαρτώμενος από εξωγενή χορήγηση της ινσουλίνης, ώστε να διατηρούνται φυσιολογικά τα επίπεδα γλυκόζης στο αίμα του.

Ο σακχαρώδης διαβήτης τύπου 2 ή μη ινσουλινο-εξαρτώμενος σακχαρώδης διαβήτης (Non-Insulin-Dependent Diabetes Mellitus, IDDM) είναι περισσότερο πολύπλοκος και χαρακτηρίζεται από διαταραχή στην έκκριση ινσουλίνης και ελάττωση της ευαισθησίας των κυττάρων σε αυτή (ινσουλινοαντοχή). Αφορά στο 90% των περιπτώσεων του ΣΔ και σχετίζεται με την παχυσαρκία, το οικογενειακό ιστορικό, την έλλειψη σωματικής άσκησης και ισορροπημένης διατροφής, την ηλικία και την εθνικότητα. Στο ΣΔ τύπου 2 η χορήγηση ινσουλίνης κρίνεται απαραίτητη μόνο στην περίπτωση που η δίαιτα και τα αντιδιαβητικά φάρμακα φανούν αναποτελεσματικά.

Ο ΣΔ χαρακτηρίζεται από αυξημένη συγκέντρωση των επιπέδων της γλυκόζης στο αίμα πέρα από τα φυσιολογικά όρια. Ο κύριος ρυθμιστής της συγκέντρωσης γλυκόζης στο αίμα είναι το ήπαρ, στο οποίο γίνεται η αποθήκευση της μεγαλύτερης ποσότητας του γλυκογόνου και η αποικοδόμησή του προς γλυκόζη (γλυκογονόλυση). Η φωσφορυλάση του γλυκογόνου (Glycogen Phosphorylase, GP) καταλύει το πρώτο στάδιο της αποικοδόμησης του γλυκογόνου προς 1-φωσφορική γλυκόζη, η οποία στη συνέχεια μετατρέπεται σε γλυκόζη και ελευθερώνεται στην κυκλοφορία του αίματος. Η αναστολή της δράσης του ενζύμου χρησιμοποιείται για τη ρύθμιση της γλυκαιμίας.

3. Πειραματικό μέρος

Ο κατευθυνόμενος - από τη δομή - σχεδιασμός φαρμάκων (Structure Based Drug Design, SBDD) αφορά στην ανακάλυψη νέων φαρμάκων με τη βοήθεια υπολογιστικών τεχνικών. Αρχικά, είχε δοθεί έμφαση στο μακρομοριακό στόχο και, πιο συγκεκριμένα, στα γνωστά κέντρα σύνδεσης όπως αυτά έχουν προκύψει από τις κρυσταλλογραφικές δομές. Για την προσομοίωση της μοριακής πρόσδεσης έχουν αναπτυχθεί πολλά λογισμικά, τα δημοφιλέστερα των οποίων είναι τα GOLD [3], [4], FlexX [5], [6], DOCK [7], ICM-Docking [8], Glide [9], [10], LigandFit [11] και AutoDock [12], [13].

Αντικείμενο της παρούσας εργασίας είναι η συγκριτική μελέτη της τρισδιάστατης δομής της GP ως συμπλόκου με ισχυρούς αναστολείς της δράσης του ενζύμου, όπως αυτή προκύπτει από κρυσταλλογραφικές μελέτες με χρήση ακτίνων X και την in silico μελέτη με το λογισμικό AutoDock.

Η επιλογή του λογισμικού AutoDock έγινε για τρεις λόγους: α) Είναι δωρεάν, β) είναι λογισμικό ανοιχτού κώδικα από το 2009 και γ) ήταν το δημοφιλέστερο λογισμικό για προσομοίωση μοριακής πρόσδεσης σύμφωνα με τη μέχρι τώρα νεότερη μελέτη [1]. Οι δομές των συμπλόκων που αναλύθηκαν προσδιορίστηκαν σε υψηλή ευκρίνεια από την Ομάδα Δομικής Βιολογίας του Ινστιτούτου Οργανικής και Φαρμακευτικής Χημείας του Εθνικού Ιδρύματος Ερευνών με χρήση συγχροτρονικής ακτινοβολίας.

Η μεν πρώτη αφορά στην ένωση σπιροϊσοξαζολίνη, κινητικές μελέτες της οποίας έδειξαν ότι αναστέλλει ισχυρά τη δράση του ενζύμου με σταθερά αναστολής K_i=0.63 μM και έχει κατατεθεί στην Protein Data Bank [14] με κωδικό 2QRP. Για την ένωση αυτή έχουν πραγματοποιηθεί παράλληλα με τις κρυσταλλογραφικές μελέτες και μελέτες μοντελοποίησης με χρήση του λογισμικού Schrödinger, τα αποτελέσματα των οποίων συνοψίζονται στο [15].

Η δεύτερη αφορά σε ένα ανάλογο γλυκόζης [16] και για την οποία επίσης έχουν πραγματοποιηθεί κρυσταλλογραφικές μελέτες. Ένα επιπρόσθετο και ενδιαφέρον στοιχείο της ένωσης αυτής είναι ότι παρουσιάζει ταυτομέρεια, δηλαδή μπορεί να εμφανίζεται σε τρεις ισοδύναμες μορφές (Ταυτομερές 1 ή T1, Ταυτομερές 2 ή T2 και Ταυτομερές 3 ή T3) που βρίσκονται σε ισορροπία μεταξύ τους και οι οποίες έχουν τον ίδιο χημικό τύπο, αλλά διαφέρουν ως προς τη θέση ενός ατόμου υδρογόνου και ενός διπλού δεσμού.

Τα υπό μελέτη σύμπλοκα προετοιμάστηκαν κατάλληλα για την προσομοίωση της μοριακής πρόσδεσης. Πιο συγκεκριμένα, τα σύμπλοκα χωρίστηκαν σε δύο διακριτά τμήματα, εκ των οποίων το ένα περιέχει τις συντεταγμένες των ατόμων του αναστολέα και το άλλο της πρωτεΐνης.

Εν συνεχεία, ορίστηκε το τρισδιάστατο πλέγμα, στο οποίο θα πραγματοποιηθεί η μοριακή πρόσδεση προκειμένου να συμπεριληφθούν τα αμινοξέα που συνθέτουν το καταλυτικό κέντρο του ενζύμου, καθώς και τον υπό μελέτη αναστολέα. Όσον αφορά στο πλέγμα, εξετάστηκαν δύο υποπεριπτώσεις: στη μεν πρώτη οι διαστάσεις του τρισδιάστατου πλέγματος ήταν 27 x 27 x 27 Å και επελέγησαν με βάση την προηγούμενη μελέτη που είχε πραγματοποιηθεί από τους Benltifa et al. ([15]) με το λογισμικό Schrödinger, ώστε να γίνουν απευθείας συγκρίσεις με το πρόγραμμα AutoDock. Στη δεύτερη περίπτωση το μέγεθος του τρισδιάστατου πλέγματος ήταν 17 x 17 x 21 Å, έτσι ώστε να εσωκλείονται σε αυτό τα αμινοξέα που σχηματίζουν εν δυνάμει απευθείας δεσμούς υδρογόνου ή/και Van der Waals αλληλεπιδράσεις με τον αναστολέα ή μέσω μορίων ύδατος, καθώς και τα αντίστοιχα μόρια ύδατος. Επιπλέον το μέγεθος του πλέγματος ήταν τέτοιο που να επιτρέπει στον αναστολέα να περιστραφεί ελεύθερα ως συμπαγές άκαμπτο σώμα εντός του πλέγματος. Η πυκνότητα του τρισδιάστατου πλέγματος επιλέχθηκε να είναι 1.000 Å και ως κέντρο του θεωρήθηκε το κέντρο του εκάστοτε αναστολέα, όπως αυτό προέκυψε από τα κρυσταλλογραφικά δεδομένα.

Έγινε προσθήκη των πολικών ατόμων υδρογόνου της πρωτεΐνης με το AutoDock και των ατόμων υδρογόνου του αναστολέα με την εφαρμογή Maestro του πακέτου λογισμικού Schrödinger, τα οποία απουσίαζαν αμφότερα από την κρυσταλλική δομή του συμπλόκου λόγω ευκρίνειας.

Επελέγησαν τα μόρια ύδατος, τα οποία θα περιλαμβάνονταν στο τρισδιάστατο πλέγμα, με βάση τις εξής περιπτώσεις: (α) κανένα μόριο ύδατος, (β) όλα τα μόρια ύδατος, (γ) τα μόρια ύδατος που σχηματίζουν απευθείας δεσμούς υδρογόνου με τον αναστολέα και (δ)–(ζ) τα μόρια ύδατος που βρίσκονται σε απόσταση μικρότερη ή ίση των 4.0, 8.0, 12.0 και 16.0 Å, αντίστοιχα, από τον αναστολέα ώστε να συμπεριλαμβάνονται όλες οι περιπτώσεις των μορίων ύδατος που σχηματίζουν απευθείας ή μέσω άλλων μορίων ύδατος αλληλεπιδράσεις με τον αναστολέα.

Εφαρμόστηκαν και οι τέσσερις αλγόριθμοι του AutoDock (Local Search (LS), Simulated Annealing (SA), Genetic Algorithm (GA) και Lamarckian Genetic Algorithm (LGA)) τόσο με τις εξ ορισμού παραμέτρους όσο και με διαφοροποίηση των τιμών των αλγορίθμων αυτών, ώστε να γίνονται περίπου 15 εκατομμύρια εκτιμήσεις για κάθε έναν από αυτούς.

Ο υποδοχέας μαζί με τα μόρια ύδατος που μπορεί να υπάρχουν σε κάθε περίπτωση θεωρείται άκαμπτος, ενώ ο αναστολέας ξεκινά από μία τυχαία θέση κάθε φορά εντός του τρισδιάστατου πλέγματος.

Ως RMSD των ποζών που προέκυψαν από την in silico μοριακή πρόσδεση ορίζεται το RMSD των συντεταγμένων των ατόμων του αναστολέα για την εκάστοτε πόζα ως προς την αντίστοιχη θέση των ατόμων στην κρυσταλλογραφική δομή και γίνεται ομαδοποίηση (clustering) των αποτελεσμάτων ως προς το RMSD, όπου το εύρος κάθε ομάδας είναι 0.5 Å, δηλαδή οι πόζες με RMSD 0.00-0.50 ανήκουν στο πρώτο cluster, οι πόζες με RMSD 0.51-1.00 ανήκουν στο δεύτερο cluster και ούτω καθεξής.

Οι παράμετροι που αποτέλεσαν κριτήριο για την αξιολόγηση των αποτελεσμάτων και δίνονται ως έξοδος του προγράμματος AutoDock είναι η ελάχιστη ελεύθερη ενέργεια πρόσδεσης (ΔG) του αναστολέα για τις παραμέτρους του κάθε αλγορίθμου και το RMSD της πόζας με την ελάχιστη ελεύθερη ενέργεια πρόσδεσης.

Οι παραπάνω παράμετροι υποβλήθηκαν σε περαιτέρω επεξεργασία για την εύρεση των μέσων τιμών των ενεργειών, των μέσων τιμών του RMSD, της σειράς ως προς το RMSD της πόζας με την ελάχιστη ελεύθερη ενέργεια πρόσδεσης, της μέσης αριθμητικής απόκλισης (mean deviation) των ελεύθερων ενεργειών πρόσδεσης ως προς την ελάχιστη ενέργεια πρόσδεσης που έχει υπολογιστεί για την εκάστοτε περίπτωση, της μέσης αριθμητικής απόκλισης του RMSD, της τυπικής απόκλισης (standard deviation) των ελεύθερων ενεργειών πρόσδεσης ως προς την ελάχιστη ενέργεια πρόσδεσης που έχει υπολογιστεί για την εκάστοτε περίπτωση, της τυπικής απόκλισης του RMSD, του ποσοστού των ποζών με αρνητική ελεύθερη ενέργεια πρόσδεσης, του εύρους του RMSD, του ποσοστού των ποζών με RMSD μικρότερο από 0.30 Å, που είναι το όριο του σφάλματος της θέσης των ατόμων όπως αυτό ορίζεται από την κρυσταλλογραφική δομή, του ποσοστού των ποζών που βρίσκονται στο πρώτο cluster και του μέσου πλήθους των clusters.

4. Πειραματικά αποτελέσματα και συμπεράσματα

4.1 Συγκριτική μελέτη των αλγορίθμων του AutoDock

Ο LGA φαίνεται πως είναι ο καλύτερος αλγόριθμος από τους τέσσερις, καθώς εμφανίζει τα βέλτιστα αποτελέσματα είτε σε όλες είτε στην πλειοψηφία των περιπτώσεων τόσο στις εξ ορισμού παραμέτρους όσο και στην περίπτωση, όπου το πλήθος των εκτιμήσεων των αλγορίθμων είναι παρεμφερές.

Το μέγεθος του πλέγματος παίζει ρόλο στην ποιότητα των αποτελεσμάτων, καθώς το μικρότερο εκ των δύο πλέγμα δίνει σαφώς καλύτερα αποτελέσματα στην περίπτωση όπου το πλήθος των εκτιμήσεων των αλγορίθμων είναι παρεμφερές, ενώ στην περίπτωση των εξ ορισμού παραμέτρων τα αποτελέσματα για τα δύο πλέγματα είναι σχεδόν ταυτόσημα.

Η αρχική θέση του αναστολέα, όπως αυτή προσδιορίζεται από το AutoDock, φαίνεται ότι δεν επηρεάζει τα αποτελέσματα που εξάγονται. Πιο συγκεκριμένα, όπως παρατηρήθηκε από την έξοδο του προγράμματος οι αρχικές θέσεις της ένωσης απείχαν κατά 5.629 Å έως 20.090 Å από αυτή της κρυσταλλογραφικής δομής, ενώ οι αντίστοιχες ενέργειες πρόσδεσης παρέμειναν σχεδόν αμετάβλητες.

Όσον αφορά στα μόρια ύδατος, παρατηρήθηκε πως η χαμηλότερη ενέργεια σύνδεσης προκύπτει στην περίπτωση που περιλαμβάνονται όλα τα μόρια ύδατος, ενώ τα αποτελέσματα στις περιπτώσεις που περιλαμβάνονται τα μόρια ύδατος σε απόσταση μικρότερη ή ίση των 4.0, 8.0, 12.0 και 16.0 Å, αντίστοιχα, από τον αναστολέα, ταυτίζονται με αυτά της περίπτωσης (γ), δηλαδή την περίπτωση στην οποία περιλαμβάνονται τα μόρια ύδατος που σχηματίζουν απευθείας δεσμούς υδρογόνου με τον αναστολέα.

4.2 Σύγκριση με την εφαρμογή Glide του πακέτου λογισμικού Schrödinger

Η σύγκριση των αποτελεσμάτων του Schrödinger και του αλγορίθμου LGA του AutoDock είναι περιορισμένη [17], καθώς ως έξοδος του πρώτου δίδεται η συνάρτηση βαθμολόγησης GlideScore, ενώ του δεύτερου είναι η ελεύθερη ενέργεια πρόσδεσης. Επιπλέον, δεν δίνεται ο χρόνος υπολογισμού της προσομοίωσης της μοριακής πρόσδεσης για τον αλγόριθμο του Glide. Το μοναδικό κοινό χαρακτηριστικό που έχουν στα αποτελέσματά τους είναι το RMSD των συντεταγμένων των ατόμων του αναστολέα ως προς την κρυσταλλογραφική δομή.

Όσον αφορά στο RMSD, το αποτέλεσμα στους Benltifa et al. [15] είναι 1.506 Å. Ο LGA παράγει μόνο στο 3.29% των αποτελεσμάτων του για το ίδιο πλήθος εκτιμήσεων πόζες με μεγαλύτερο RMSD. Συνυπολογίζοντας και την περίπτωση με τις εξ ορισμού παραμέτρους όπου ο LGA δεν παράγει ούτε μία πόζα με RMSD μεγαλύτερο του 1.506 Å, το ποσοστό αυτό γίνεται 1.64%. Όλες οι πόζες του LGA, 23 τον αριθμό, με RMSD μεγαλύτερο του 1.506 Å, έχουν ταυτόχρονα RMSD μεγαλύτερο του 13.00 Å. Επομένως, είναι πολύ εύκολο να τις εντοπίσει κάποιος και να τις αφαιρέσει από το σύνολο των λύσεων μιας και διαφέρουν τόσο πολύ από τις υπόλοιπες. Έτσι, το πλήθος των ποζών του LGA με RMSD μεγαλύτερο του 1.506 Å γίνεται 0. Από τα παραπάνω συμπεραίνεται ότι το AutoDock έχει πολύ καλύτερα αποτελέσματα από το Schrödinger για το κριτήριο αυτό σε ποσοστό τουλάχιστον 98.36%.

Όσον αφορά στην ελεύθερη ενέργεια πρόσδεσης που είχε ως έξοδο το Glide, αυτή υπολογίζεται περίπου σε –8.65 kcal/mol από το [15], η οποία αντιστοιχεί σε μία σταθερά αναστολής Ki = 0.455 μM, δηλαδή μικρότερη από αυτήν που προέκυψε από τα πειραματικά δεδομένα (0.63 μM). Στο AutoDock οι πόζες με RMSD μικρότερο του 1.506 Å έχουν χαμηλότερη ελεύθερη ενέργεια πρόσδεσης και άρα μικρότερη σταθερά αναστολής, δηλαδή χειρότερα αποτελέσματα. Για τις υπόλοιπες πόζες, η ελεύθερη ενέργεια πρόσδεσης είναι ιδιαίτερα αυξημένη δίδοντας μια σταθερά αναστολής περίπου 47–380 φορές μεγαλύτερη από αυτήν που έχει προσδιοριστεί από τα πειραματικά δεδομένα. Επομένως, για το κριτήριο αυτό το Schrödinger υπερέχει σε ποσοστό 100%.

Η σύγκριση των δύο λογισμικών έδειξε ότι το AutoDock παράγει πόζες που προσομοιάζουν πολύ περισσότερο από αυτές του Schrödinger σε αυτές που προσδιορίσθηκαν πειραματικά. Το Schrödinger, ωστόσο, έχει μία πολύ καλύτερη συνάρτηση βαθμολόγησης, η οποία μπορεί και προβλέπει ακριβέστερα την ελεύθερη ενέργεια πρόσδεσης σε σύγκριση με αυτήν που προκύπτει από την προσομοίωση της μοριακής πρόσδεσης του AutoDock.

4.3 LGA και ταυτομερή

Στην περίπτωση των ταυτομερών εφαρμόστηκε μόνο ο LGA αλγόριθμος, το μικρότερο από τα δύο πλέγματα και, όσον αφορά στα μόρια ύδατος, μελετήθηκαν οι περιπτώσεις όπου περιλαμβάνονται όλα τα μόρια ύδατος, κανένα μόρια ύδατος και τα μόρια ύδατος που σχηματίζουν απευθείας δεσμούς υδρογόνου με τον αναστολέα. Οι υπόλοιπες παράμετροι της μεθοδολογίας και του αλγορίθμου είναι ίδιες με αυτές που εφαρμόσθηκαν στην περίπτωση του LGA για την σπιροϊσοξαζολίνη.

Τα αποτελέσματα που προέκυψαν έδειξαν ότι το T2 φαίνεται να προσδένεται καλύτερα. Πιο συγκεκριμένα, έχει τα καλύτερα αποτελέσματα όσον αφορά στην ελάχιστη ελεύθερη ενέργεια πρόσδεσης, στη μέση ελεύθερη ενέργεια πρόσδεσης, στη μέση αριθμητική και τυπική απόκλιση της ελεύθερης ενέργειας πρόσδεσης. Όσον αφορά στα κριτήρια αυτά, το T1 έχει τα δεύτερα καλύτερα αποτελέσματα για τη μέση ελεύθερη ενέργεια πρόσδεσης, καθώς και τη μέση αριθμητική και τυπική απόκλιση της ελεύθερης ενέργειας πρόσδεσης, ενώ το T3 για την ελάχιστη ελεύθερη ενέργεια πρόσδεσης. Επίσης, σημειώνεται ότι και για τα τρία ταυτομερή η ελάχιστη ελεύθερη ενέργεια πρόσδεσης παρατηρείται για την περίπτωση που περιλαμβάνονται όλα τα μόρια ύδατος.

Στη σύγκριση που πραγματοποιήθηκε με βάση το RMSD, το ταυτομερές Τ1 εμφανίζει τα καλύτερα αποτελέσματα όσον αφορά στα κριτήρια του RMSD της πόζας με την ελάχιστη ελεύθερη ενέργεια πρόσδεσης, τη σειρά της πόζας με το ελάχιστο RMSD ως προς την ελεύθερη ενέργεια πρόσδεσης, καθώς και της τυπικής και μέσης απόκλισης του RMSD. Μάλιστα, και στις τρεις περιπτώσεις που αφορούν στα μόρια ύδατος, η πόζα του Τ1 με την ελάχιστη ενέργεια είναι και αυτή με το ελάχιστο RMSD. Αντίστοιχα, το Τ2 εμφανίζει τα καλύτερα αποτελέσματα για το μέσο RMSD, το μέσο πλήθος των clusters, το πλήθος των ποζών στο πρώτο cluster και, επίσης, ισοβαθμεί με το Τ1 στο πλήθος των ποζών με ευνοϊκή (αρνητική) ενέργεια πρόσδεσης. Αντίθετα, το ταυτομερές Τ3 δεν εμφανίζει τα καλύτερα αποτελέσματα σε καμία κατηγορία παρά μόνο έχει τα δεύτερα καλύτερα αποτελέσματα για τα κριτήρια του RMSD της πόζας με την ελάχιστη ελεύθερη ενέργεια πρόσδεσης, της σειράς της πόζας με το ελάχιστο RMSD ως προς την ελεύθερη ενέργεια πρόσδεσης και του πλήθους των ποζών στο πρώτο cluster. Άξιο αναφοράς, τέλος, είναι το γεγονός ότι δε βρέθηκε καμία πόζα που να έχει RMSD μικρότερο του 0.30 Å.

Λαμβάνοντας υπ' όψιν όλα τα κριτήρια που αφορούν την ελεύθερη ενέργεια πρόσδεσης και το RMSD, το Τ2 εμφανίζει τα καλύτερα αποτελέσματα για περισσότερα εξ αυτών σε σύγκριση με το Τ1 και το Τ3. Το Τ1 εμφανίζει τα καλύτερα αποτελέσματα για δύο κριτήρια, ενώ το Τ3 για κανένα. Αν συνυπολογιστεί και το γεγονός ότι το Τ1 εμφανίζει περισσότερες φορές τα δεύτερα καλύτερα αποτελέσματα σε σύγκριση με το Τ3, εξάγεται το συμπέρασμα ότι το Τ2 προσδένεται καλύτερα απ' όλα τα ταυτομερή και το Τ1 καλύτερα από το Τ3.

4.4 Μελλοντική εργασία

Για την ακριβέστερη πρόβλεψη της ελεύθερης ενέργειας πρόσδεσης, η αξιολόγηση των αποτελεσμάτων που προκύπτουν από τον LGA θα μπορούσε να πραγματοποιηθεί και με κάποια άλλη συνάρτηση ενεργειακής βαθμολόγησης, όπως για παράδειγμα το ChemScore.

Για τις δομές για τις οποίες δεν υπάρχουν κρυσταλλογραφικά δεδομένα, θα μπορούσε να γίνει πρόβλεψη της μοριακής πρόσδεσης με τη χρήση του LGA, αφού προηγηθεί προσθήκη των θεωρητικών μορίων ύδατος με κάποιο άλλο λογισμικό και στη συνέχεια γίνει η αξιολόγηση των δομών με τη συνάρτηση ενεργειακής βαθμολόγησης.

Μελλοντικά θα μπορούσαν να γίνουν περισσότερες μελέτες με τη βοήθεια του λογισμικού AutoDock, είτε με νέους αναστολείς για την εξαγωγή περισσότερων συμπερασμάτων ή/και την επιβεβαίωση/επανέλεγχο των ήδη υπαρχόντων είτε με τους ίδιους αναστολείς, αλλά με (επιπλέον) διαφοροποίηση των παραμέτρων, ώστε να εξαχθούν συμπεράσματα για το ρόλο τους.

Η ευκαμψία του υποδοχέα σε όλες τις υπολογιστικές μελέτες μοριακής πρόσδεσης είναι μια εξαιρετικά σπουδαία και ταυτόχρονα περίπλοκη παράμετρος, η οποία πρέπει να αξιολογείται κατάλληλα και να λαμβάνεται υπ΄όψιν στους υπολογισμούς. Στο πλαίσιο της παρούσας εργασίας αυτό δεν έγινε για λόγους σύγκρισης, ωστόσο θα παρουσίαζε ενδιαφέρον η μελέτη αλγορίθμων που λαμβάνουν υπ΄όψιν τους την ευκαμψία του υποδοχέα, όταν αυτοί χρησιμοποιούνται για την πρόβλεψη του τρόπου σύνδεσης ενώσεων, η δομή των οποίων έχει προσδιορισθεί τόσο με συγκρυστάλλωση όσο και με εμποτισμό του κρυστάλλου σε διάλυμα της προς μελέτη ένωσης.

Αναφορές

1. S. F. Sousa, P. A. Fernandes and M. J. Ramos, Protein-ligand docking: current status and future challenges, Proteins: Structure, Function, and Genetics, vol. 65, Issue 1, 2006, pp. 15-26.

2. P. Zimmet, K. G. Alberti and J. Shaw, Global and societal implications of the diabetes epidemic, IDF Diabetes Atlas fourth edition, 2009, pp. 782-787.

3. Gareth Jones, Peter Willett and Robert C. Glen, Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation, Journal of Molecular Biology, vol. 245, Issue 1, 1995, pp. 43-53.

4. Gareth Jones, Peter Willett and Robert C. Glen, A genetic algorithm for flexible molecular overlay and pharmacophore elucidation, Journal of Computer-Aided Molecular Design, vol. 11, no. 6, 1995, pp. 532-549.

5. Andrew R. Leach and Irwin D. Kuntz, Conformational analysis of flexible ligands in macromolecular receptor sites, Journal of Computational Chemistry, vol. 13, no. 6, 1992, pp. 730-748.

6. Frank H. Allen, John E. Davies, Jean J. Galloy, Owen Johnson, Olga Kennard, Clare F. Macrae, Eleanor M. Mitchell, Gary F. Mitchell, J. Michael Smith and David G. Watson, The development of versions 3 and 4 of the Cambridge Structural Database System, Journal of Chemical Information and Computer Sciences, vol. 31, Issue 2, 1991, pp. 43-53.

7. Irwin D. Kuntz, Jeffrey M. Blaney, Stuart J. Oatley, Robert Langridge and Thomas E.

Ferrin, A geometric approach to macromolecule-ligand interactions, Journal of Molecular Biology, vol. 161, Issue 2, 1982, pp. 269-288.

8. R. Abagyan, M. Totrov and D. Kuznetsov, ICM-A New Method for Protein Modeling and Design: Applications to Docking and Structure Prediction from the Distorted Native Conformation, Reviews in Computational Chemistry, vol. 15, Issue 5, 1994, pp. 488-506.

9. R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S Shenkin, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy, Journal of Medicinal Chemistry, vol. 47, no. 7, 2004, pp. 1739-1749.

10. Thomas A. Halgren, Robert B. Murphy, Richard A. Friesner, Hege S. Beard, Leah L. Frye, W. Thomas Pollard, and Jay L. Banks, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening, Journal of Medicinal Chemistry, vol. 47, no. 7, 2004, pp. 1750-1759.

11. C. M. Venkatachalam, X. Jiang, T. Oldfield and M. Waldman, LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites, Journal of Molecular Graphics and Modelling, vol. 21, Issue 4, 2003, pp. 289-307.

12. P. J. Goodford and Arthur J. Olson, Automated Docking of Substrates to Proteins by Simulated Annealing, Proteins: Structure, Function and Genetics, vol. 8, Issue 3, 1990, pp. 195-202.

13. Garrett M. Morris, David S. Goodsell, Robert S. Halliday, Ruth Huey, William E. Hart, Richard K. Belew, Arthur J. Olson, Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function, Journal of Computational Chemistry, vol. 19, no. 14, 1998, pp. 1639-1662.

14. Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov and Philip E. Bourne, The Protein Data Bank, Nucleic Acids Research, vol. 28, Issue 1, 2000, pp. 235-242.

15. Mahmoud Benltifa, Joseph M. Hayes, Sébastien Vidal, David Gueyrard, Peter G. Goekjian, Jean-Pierre Praly, Gregory Kizilis, Costas Tiraidis, Kyra-Melinda Alexacou, Evangelia D. Chrysina, Spyros E. Zographos, Demetres D. Leonidas, Georgios Archontis and Nikos G. Oikonomakos, Glucose-based spiro-isoxazolines: A new family of potent glycogen phosphorylase inhibitors, Bioorganic \& Medicinal Chemistry, vol. 17, Issue 20, 2009, pp. 7368-7380.

16. Crina Cismas, Joseph M. Hayes, Dimitris Sovantzis, Theodoros Hadjiloi, Michail Mamais, Eliada Lazoura, Paschalis Grammatopoulos, Dimitris Panagopoulos, Dimitris Stathis, George N. Oikonomakos, Spyros E. Zographos, Demetres D. Leonidas, Nikos G. Oikonomakos, Thanasis Gimisis, and Evangelia D. Chrysina, Synthesis, X-ray and computational studies of β-D-glucose pyrimidine derivatives as glycogen phosphorylase inhibitors: the role of ligand ionization/tautomeric states examined, unpublished results.

17. Jason C. Cole, Christopher W. Murray, J. Willem M. Nissink, Richard D. Taylor and Robin Taylor, Comparing Protein–Ligand Docking Programs Is Difficult, Proteins: Structure, Function and Genetics, vol. 60, Issue 3, 2005, pp. 325-332.

Περί Κοινωνιοκεντρικών Προσεγγίσεων στο Πρόβλημα Δρομολόγησης σε Ασύρματα Οπορτουνιστικά Δίκτυα

Παύλος Νικολόπουλος
p.nikolopoulos@di.uoa.gr

Θεράπων Παπαδημητρίου
theraponpap@di.uoa.gr

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Πανεπιστημιούπολη, Ιλίσια, 15784, Αθήνα, Ελλάς

Περίληψη

Εξετάζουμε πειραματικά τους παράγοντες που επηρεάζουν την κοινωνιοκεντρική προσέγγιση στη δρομολόγηση των ασύρματων οπορτουνιστικών δικτύων. Πέντε εναλλακτικά σχήματα δρομολόγησης που βασίζονται στη μετρική της «κεντρικότητας» των κόμβων, συγκρίνονται μεταξύ τους, με σημείο αναφοράς ένα ιδεατό σχήμα δρομολόγησης, που λαμβάνει υπόψη του τη μελλοντική ακολουθία των επαφών στο δίκτυο. Για το σκοπό αυτό, προτείνουμε μία μέθοδο που υπολογίζει τα βέλτιστα χωρο-χρονικά μονοπάτια μετάδοσης μηνυμάτων πάνω σε έναν πρωτότυπο γράφο, με χρονικά βάρη στις ακμές και κόμβους τις ίδιες τις επαφές.

Λέξεις-Κλειδιά: Δίκτυα Ανεκτικά σε Καθυστερήσεις (DTN), Κεντρικότητα, Γράφοι Χώρου-Χρόνου

Η εργασία δημοσιεύθηκε με τίτλο "How off-center are centrality metrics for opportunistic routing?" [12], στο ACM MobiCom workshop on Challenged Networks (CHANTS 2011), Las Vegas, Nevada-USA, το Σεπτέμβριο του 2011.

Επιβλέποντες:
Ιωάννης Σταυρακάκης, Καθηγητής
Δρ. Μερκούρης Καραλιόπουλος

1. Εισαγωγή

Στα οπορτουνιστικά δίκτυα, συχνά αναφερόμενα ως Delay Tolerant Networks (DTNs), η μόνιμη διασύνδεση όλων των κόμβων μεταξύ αποστολέα και παραλήπτη, κατά τη διάρκεια της μετάδοσης ενός μηνύματος, δεν αποτελεί προϋπόθεση. Συνεπώς, η δρομολόγηση έγκειται στην προώθηση ολόκληρων μηνυμάτων από κόμβο σε κόμβο, με αποτέλεσμα το πρόβλημα δρομολόγησης να εκφυλίζεται σε θετικές ή αρνητικές αποφάσεις προώθησης, κατά τη διάρκεια της συνάντησης-επαφής δύο ή περισσότερων κόμβων του δικτύου. Οι αρχικές προτάσεις δρομολόγησης ήταν ουσιαστικά παραλλαγές μίας «ελεγχόμενης πλημμύρας» μηνυμάτων μέσα στο δίκτυο, οι οποίες περιόριζαν τα προβλήματα του επιδημικού (epidemic) πρωτοκόλλου διάδοσης [11], θέτοντας άνω όριο στα παραγόμενα αντίγραφα (π.χ. [10]). Νεότερα σχήματα δρομολόγησης επεδίωξαν πιο στοχευμένες αποφάσεις προώθησης, υπολογίζοντας τη χρησιμότητα των κόμβων για αναμετάδοση του μηνύματος (relaying utility). Τα utilities λάμβαναν υπόψη τη συχνότητα των επαφών με τον κόμβο-προορισμό [8], ή πιο γενικά κοινωνικά χαρακτηριστικά, όπως τα συχνότερα επισκεπτόμενα μέρη ή τα κοινά ενδιαφέροντα του υποψήφιου αναμεταδότη με τον κόμβο-προορισμό [1].

Τελευταία, η κοινωνική πληροφορία έχει εισαχθεί πιο μεθοδικά στις συναρτήσεις των relaying utilities, μέσω της εκμετάλλευσης εννοιών και μετρικών της Κοινωνικής Ανάλυσης Δικτύων (Social Network Analysis-SNA). Χαρακτηριστικά παραδείγματα αποτελούν τα πρωτόκολλα SimBetTS [4] και BubbleRap [7], στα οποία οι μετρικές SNA υπολογίζονται πάνω σε κοινωνικούς γράφους που αποτυπώνουν την ακολουθία των επαφών μέσα σε καθορισμένα χρονικά παράθυρα. Και τα δύο πρωτόκολλα επιτυγχάνουν καλύτερη απόδοση έναντι των αρχικών προσεγγίσεων και αναδεικνύουν την «κεντρικότητα» (centrality) του εκάστοτε κόμβου, ως το σημαντικότερο κριτήριο για τις αποφάσεις προώθησης, ακόμη και όταν αυτή συνδυάζεται με άλλες.

Εντούτοις, η δρομολόγηση με βάση την κεντρικότητα εμφανίζει τρεις βασικές αδυναμίες, που μέχρι τώρα δεν έχουν μελετηθεί και αξιολογηθεί: α) οι τιμές κεντρικότητας των κόμβων δεν λαμβάνουν υπόψη τον προορισμό του μηνύματος, β) η απόδοση εξαρτάται άμεσα από τον τρόπο κατασκευής του γράφου επαφών πάνω στον οποίο υπολογίζονται οι μετρικές SNA [6] και γ) η κεντρικότητα μπορεί να προσεγγιστεί πρακτικά με υπολογισμό πάνω στο two-hops υποδίκτυο του κάθε κόμβου (ego network), παρέχοντας περιορισμένη εικόνα της χρησιμότητάς του στο συνολικό δίκτυο.

Στόχος της εργασίας μας είναι να αξιολογήσουμε το πληροφοριακό περιεχόμενο της κεντρικότητας των κόμβων, όταν αυτή χρησιμοποιείται ως κριτήριο δρομολόγησης σε οπορτουνιστικές τοπολογίες. Εξετάζουμε σε βάθος τη δημοφιλέστερη μετρική, Betweenness Centrality (BC), και μελετούμε πώς ο τύπος του γράφου επαφών – γράφος με βάρη (weighted) ή χωρίς (unweigthed) – και ο τρόπος που υπολογίζεται η μετρική – κοινωνιοκεντρικά (sociocentric BC) ή εγωκεντρικά (egocentric BC) – επηρεάζουν την απόδοση των αποφάσεων προώθησης. Επιπλέον, αξιολογούμε πώς η μετρική BC συγκρίνεται με την ισοδύναμή της Conditional BC (CBC) [9], η οποία διατηρεί τιμές κεντρικότητας ανά προορισμό, για κάθε κόμβο. Εξετάζουμε κυρίως τις εγγενείς αδυναμίες της δρομολόγησης με βάση την κεντρικότητα, επειδή θέτουν αυστηρά όρια στην απόδοση και δεν ενδιαφερόμαστε για τις ατέλειες συγκεκριμένων πρωτοκόλλων: Πόσο μπορούν οι αποφάσεις δρομολόγησης που βασίζονται στις μετρικές κεντρικότητας να πλησιάσουν το βέλτιστο; Πώς οι διάφορες εναλλακτικές για τον υπολογισμό της κεντρικότητας των κόμβων επιδρούν στην απόδοση της δρομολόγησης;

Μελετούμε θεωρητικά και πειραματικά τις εναλλακτικές λύσεις για τον υπολογισμό της κεντρικότητας των κόμβων. Πέντε τεχνικές, που προσεγγίζουν διαφορετικά τις τιμές της κεντρικότητας και τις χρησιμοποιούν ως κριτήριο των αποφάσεων προώθησης, συγκρίνονται μεταξύ τους και επιπλέον με δύο σχήματα που αντιπροσωπεύουν τις δύο ακραίες περιπτώσεις δρομολόγησης σε DTN: α) το απλό πρωτόκολλο προώθησης μηνυμάτων με βάση μία ορισμένη πιθανότητα και β) ένα ιδανικό σχήμα που υπολογίζει τα βέλτιστα χωρο-χρονικά μονοπάτια για τη μετάδοση ενός μηνύματος, έχοντας πλήρη γνώση των μελλοντικών επαφών. Εισάγουμε και αναλύουμε μία νέα μορφή γράφου επαφών, που επιτρέπει τη χρησιμοποίηση τυποποιημένων αλγορίθμων shortest-path σε δυναμικά οπορτουνιστικά δίκτυα. Τέλος, παρουσιάζουμε τα αποτελέσματα και συμπεράσματα της αξιολόγησής μας, τα οποία δεν συμβαδίζουν πάντα με τη διαισθητική λογική.

2. Εναλλακτικοί τρόποι υπολογισμού της κεντρικότητας

Destination-aware ή destination-unaware.

Όταν οι κόμβοι αποφασίζουν να προωθήσουν (ή όχι) ένα μήνυμα σε έναν κόμβο που συναντούν, θα πρέπει να λαμβάνουν υπόψη τη μεταβιβαστική χρησιμότητα του κόμβου ως προς τον προορισμό του μηνύματος. Ωστόσο, οι μετρικές κεντρικότητας συνήθως είναι μέσες τιμές ([4] [7]), που προκύπτουν από όλα τα πιθανά ζεύγη αποστολέων-παραληπτών του δικτύου. Έτσι, ένας κόμβος με υψηλή (χαμηλή) τιμή κεντρικότητας, μπορεί εύκολα να είναι μια άστοχη (εύστοχη) επιλογή αναμετάδοσης για ορισμένους κόμβους-παραλήπτες.

Unweighted ή weighted contact graphs.

Στις μέχρι τώρα μελέτες, η ακολουθία επαφών ενσωματώνεται σε ένα στατικό γράφο επαφών χωρίς βάρη, $G=(V,E)$, όπου V το σύνολο των κόμβων του δικτύου και $e \in E$ οι ακμές, που ενώνουν τα ζεύγη των κόμβων που έχουν πραγματοποιήσει $C_{\mathrm{e}} > thr_{\mathrm{e}}$ συναντήσεις, μέσα σε παράθυρο σταθερής χρονικής διάρκειας T. Η απόδοση των σχημάτων δρομολόγησης εξαρτάται άμεσα από τα thr_e και T [6]. Συνήθεις τιμές είναι thr_e=1, T = 6hrs (π.χ. [7]). Ωστόσο, μία ακολουθία επαφών μπορεί εναλλακτικά να αποτυπωθεί σε γράφο με βάρη, που δηλώνουν τη συχνότητα των συναντήσεων μεταξύ των κόμβων. Ο υπολογισμός των τιμών κεντρικότητας πάνω σ' αυτόν το γράφο, οδηγεί σε ένα διαφορετικό σύνολο τιμών κεντρικότητας για τους κόμβους της ίδιας ακολουθίας επαφών.

Egocentric ή sociocentric.

Ο υπολογισμός των μετρικών συνολικής κεντρικότητας (sociocentric) απαιτεί συνήθως πλήρη γνώση της τοπολογίας του δικτύου. Ακόμα και αν το προσεγγιστικό σφάλμα που προκύπτει από την ενσωμάτωση των επαφών σε στατικούς γράφους είναι αποδεκτό, οι μεμονωμένοι κόμβοι δεν μπορούν πάντα να γνωρίζουν τον πλήρη κοινωνικό γράφο που παράγεται από τις συναντήσεις όλων των κόμβων του δικτύου. Επομένως, ρεαλιστικά πρωτόκολλα δρομολόγησης, όπως τα [4] και [7], στηρίζονται σε εγωκεντρικές προσεγγίσεις (egocentric) της συνολικής κεντρικότητας, μέσω των ego δικτύων των κόμβων [5]. Εντούτοις, υπάρχουν λίγα στοιχεία σχετικά με το πόσο καλή είναι η συσχέτιση μεταξύ εγωκεντρικών και κοινωνιοκεντρικών τιμών και πότε η πλήρης γνώση του δικτύου απαιτείται πραγματικά.

3. Μεθοδολογία πειραμάτων

Πραγματοποιούμε πειράματα χρησιμοποιώντας πέντε γνωστές καταγραφές (traces) του προγράμματος Haggle [3], που αφορούν σε επαφές μέσω Bluetooth μεταξύ συσκευών iMotes, τις οποίες φέρουν χρήστες ρεαλιστικών πειραματικών δικτύων DTN. Κάθε επαφή Bluetooth θεωρείται ως ευκαιρία στην οποία οι κόμβοι μπορούν να ανταλλάξουν πληροφορία. Στον Πίνακα 1, φαίνονται τα βασικά παραμετρικά αυτών των πειραμάτων, ενώ περισσότερες πληροφορίες παρατίθενται στα [3] και [12].

Η μεθοδολογία πειραμάτων έχει ως εξής: Παράγουμε μηνύματα της μορφής $msg\ (s,\ d,\ t)$, όπου επιλέγονται τυχαία η πηγή s, ο προορισμός d και ο χρόνος δημιουργίας τους t και εξομοιώνουμε τα μονοπάτια μετάδοσής τους στο δίκτυο. Για το λόγο αυτό, εκτελούμε «replay» των traces, όπου οι κόμβοι υπολογίζουν τις τιμές BC και λαμβάνουν ανάλογες αποφάσεις προώθησης, για κάθε μήνυμα. Εξετάζουμε πέντε πιθανούς τρόπους διαφορετικού υπολογισμού των τιμών BC των κόμβων.

Διαμόρφωση	Intel	Cambridge	Infocom05	Content	Infocom06
Διάρκεια Πειράματος (ημέρες)	6	6	4	24	4
Διάρκεια Σάρωσης (sec)	5-10	5-10	5-10	5-10	5-10
Χρόνος μεταξύ Σαρώσεων (sec)	120	120	120	120-600	120
Κινητές Συσκευές	8	12	41	36	78
Σταθερές Συσκευές	1	0	0	18	20
Μ.Ο. επαφών/ ζεύγος/ημέρα	9.09	12.09	8.60	0.66	9.03
# Επαφών	2766	6732	28216	41330	227657

Πίνακας 1: Χαρακτηριστικά των πειραματικών δεδομένων

3.1 Εξομοίωση της δρομολόγησης με βάση την κεντρικότητα πάνω σε traces

Το πρώτο βήμα επεξεργασίας είναι η αποτύπωση των επαφών των traces σε ένα γράφο, πάνω στον οποίο υπολογίζονται οι τιμές κεντρικότητας των κόμβων.

Πρόκειται για μία αποτύπωση του δυναμικού οπορτουνιστικού δικτύου που συμπιέζει τη διάσταση του χρόνου. Για κάθε trace, παράγουμε δύο διαφορετικούς στατικούς γράφους, έναν χωρίς βάρη με thr_e=1 και έναν με βάρη, όπου το βάρος μιας ακμής μεταξύ δύο κόμβων ισούται με το αντίστροφο του πλήθους των επαφών τους.

Η ακολουθία μηνυμάτων msg (s, d, t) δρομολογείται στον προορισμό μέσω «άπληστων» (greedy) αποφάσεων προώθησης. Πιο συγκεκριμένα, εάν τη χρονική στιγμή t_j το μήνυμα είναι στον κόμβο u, ο u θα μεταβιβάσει το μήνυμα σε έναν άλλο κόμβο k , κατά την επόμενη συνάντησή τους, μόνο εάν $BC(k) > BC(u)$ (RANK [7]). Οι τιμές αυτές μπορούν να προέλθουν είτε egocentric είτε sociocentric και υπολογίζονται από το σύνολο των επαφών που συνέβησαν στο χρονικό παράθυρο | $t_j - T, t_j$ |, όπου T το χρονικό παράθυρο ενσωμάτωσης των επαφών. Δεδομένου ότι για τον υπολογισμό της sociocentric BC απαιτείται πλήρης γνώση των συναντήσεων σε όλο το δίκτυο, ενδεχόμενη μείωση της απόδοσης οφείλεται στην έλλειψη ισχυρού πληροφοριακού περιεχομένου της μετρικής και όχι της διαθέσιμης πληροφορίας για το δίκτυο.

Εξετάζουμε επίσης την Conditional BC (CBC) [9], όπου η κεντρικότητα ενός τυχαίου κόμβου υπολογίζεται σε συνάρτηση με ένα συγκεκριμένο προορισμό d:

$$CBC(u;d) = \sum_{s \in V,\, u \neq d} \frac{\sigma_{sd}(u)}{\sigma_{sd}} \qquad (1)$$

όπου σ_{sd} το πλήθος των shortest paths ανάμεσα στους κόμβους s και d και $\sigma_{sd}(u)$ τα shortest paths που περνούν από τον κόμβο u. Η $CBC(u;d)$ προκύπτει από όλα τα ζεύγη κόμβων (x, d), $\forall x \in V$, και όχι όλα τα πιθανά ζεύγη, όπως η $BC(u)$.

Η διαδικασία δρομολόγησης εξομοιώνεται πέντε φορές πάνω σε κάθε trace. Χρησιμοποιούμε τις συντμήσεις $soc(C)BC_{uw}$, $egoBC_{uw}$, $socBC_w$, $egoBC_w$, για τις sociocentric/egocentric $(C)BC$ τιμές που υπολογίζονται σε unweighted (uw)/weighted (w) γράφους, αντίστοιχα. Υπολογίζουμε την καθυστέρηση παράδοσης και το πλήθος των μεταβιβάσεων των μηνυμάτων msg (s, d, t) και τα συγκρίνουμε με δύο σχήματα δρομολόγησης, που χρησιμοποιούνται σαν σημείο αναφοράς: α) ένα πιθανοτικό σχήμα PF_p, όπου ένας κόμβος προωθεί με πιθανότητα p το μήνυμα, κάθε φορά που συναντά άλλον κόμβο και β) το βέλτιστο (opt) σχήμα, που περιγράφεται στην επόμενη Παράγραφο. Δεδομένου ότι η δρομολόγηση με βάση την κεντρικότητα αποτελεί greedy προσέγγιση στο πρόβλημα βελτιστοποίησης διαδρομής του μηνύματος, υπάρχει μία πιθανότητα, μη-αμελητέα, ένα μήνυμα να «παγιδεύεται» σε έναν κόμβο με μεγάλη τιμή κεντρικότητας, που δεν συναντά ποτέ τον κόμβο-προορισμό. Γι' αυτό το λόγο, για κάθε παραλλαγή δρομολόγησης, εκθέτουμε το σύνολο των παραδοθέντων μηνυμάτων, σαν ποσοστό του συνόλου των μηνυμάτων που δύνανται να παραδοθούν με το βέλτιστο σχήμα (παραδοτέα μηνύματα).

4. Υπολογισμός των βέλτιστων χωρο-χρονικών μονοπατιών

Τα βέλτιστα χωρο-χρονικά μονοπάτια μετάδοσης ορίζουν το άνω όριο της απόδοσης οποιασδήποτε τεχνικής δρομολόγησης. Περιγράφονται με τη βοήθεια χρονικά εξελισσόμενων γράφων [2] και διακρίνονται σε «ταχύτερα» και «κοντύτερα» μονοπάτια. Τα ταχύτερα μονοπάτια αντιστοιχούν στις συντομότερες διαδρομές (foremost journeys), υπό την έννοια της πιο έγκαιρης άφιξης του μηνύματος στον προορισμό, ενώ τα κοντύτερα μονοπάτια στις διαδρομές ελάχιστων μεταβιβάσεων (shortest journeys), υπό την έννοια του μικρότερου πλήθους προωθήσεων του μηνύματος. Για να ανακαλύψει πραγματικά και να χρησιμοποιήσει αυτές τις διαδρομές ένας κόμβος, θα πρέπει να έχει πλήρη γνώση της ακριβούς ακολουθίας των μελλοντικών επαφών, μεταξύ όλων των κόμβων του δικτύου.

Ωστόσο, ακόμη και στην ιδεατή αυτή περίπτωση, ο θεωρητικός υπολογισμός των βέλτιστων μονοπατιών παρουσιάζει προβλήματα. Η άμεση εφαρμογή κλασικών μεθόδων εύρεσης βέλτιστου μονοπατιού πάνω στο στατικό γράφο επαφών (με/ χωρίς βάρη), αποτυγχάνει να δώσει λύση, γιατί ο γράφος επαφών συμπιέζει τη διάσταση του χρόνου και δεν περιγράφει τη χρονική εξέλιξη του δικτύου. Έτσι, αναπτύξαμε μία διαφορετική αναπαράσταση γράφου, που εκμεταλλεύεται τη σχετική χρονική σειρά των επαφών των κόμβων και επιτρέπει τη χρήση τυποποιημένων αλγορίθμων shortest-path, π.χ., Dijkstra ή Depth First Search (DFS). Η διαδικασία παραγωγής του γράφου δέχεται σαν είσοδο την ακολουθία των επαφών, διαμορφωμένη όπως παρουσιάζεται στο Σχήμα 1 και ακολουθεί τα εξής δύο βήματα:

Από το αρχικό trace στην ακολουθία επαφών προώθησης (φιλτράρισμα).

Η ακολουθία επαφών φιλτράρεται, διατηρώντας μόνο εκείνες που μπορούν να οδηγήσουν σε προώθηση ενός συγκεκριμένου μηνύματος, εφεξής επαφές προώθησης (forwarding contacts). Εάν t_0 είναι ο χρόνος δημιουργίας ενός μηνύματος στην πηγή s για τον προορισμό d, τότε το πρώτο βήμα αποκλείει:

1. Τις επαφές πριν την πρώτη επαφή c_0 μετά το χρόνο t_0, που εμπλέκει τον κόμβο s.

2. Τις επαφές του υπόλοιπου trace, οι οποίες δεν αποτελούν ευκαιρία για μετάδοση του μηνύματος. Μια επαφή $c_j = (n_k\,,\, n_l\,,\, t_j)$ παρουσιάζει ευκαιρία, εφόσον ένας από τους δύο κόμβους, n_k και n_l , είχε και νωρίτερα την ευκαιρία να αποκτήσει το μήνυμα μέσω ενός ή περισσότερων επαφών προώθησης (βλ. Σχήμα 1).
3. Τις επαφές μετά από την πρώτη επαφή προώθησης c_d, που εμπλέκει τον κόμβο d.

Στη δημοσιευμένη εργασία μας [12], περιγράφεται αναλυτικά ότι η συνολική πολυπλοκότητα του πρώτου βήματος του trace είναι $O(Nlog|V|+N+\log N+|V|log|V|\ )=O(Nlog|V|)$, η οποία, για τυπικές τιμές πλήθους κόμβων, είναι της τάξεως του χρόνου που απαιτείται για να διαβαστούν οι καταχωρήσεις των επαφών.

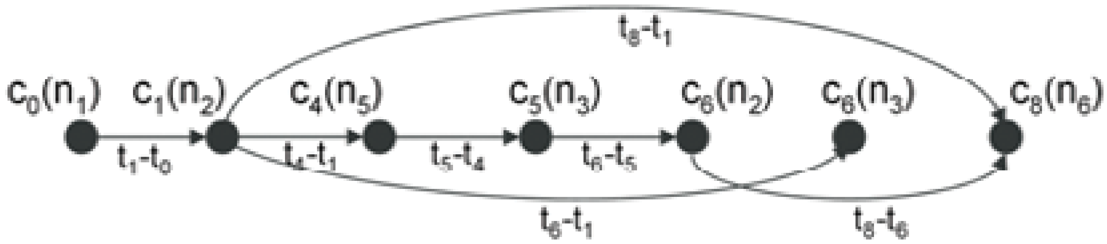
<u>Δημιουργία του γράφου επαφών προώθησης</u>.

Οι N_r καταχωρήσεις επαφών που απομένουν μετά το πρώτο βήμα, μετασχηματίζονται σε έναν κατευθυνόμενο γράφο επαφών προώθησης, όπου το σύνολο των κόμβων (vertices) V_c, αντιστοιχεί στις επαφές προώθησης, δηλαδή σε ζευγάρια κόμβων (node pairs) και όχι τους μεμονωμένους κόμβους του δικτύου (Σχήμα 2). Κάθε vertex έχει σε παρένθεση τον κόμβο στον οποίο μεταβιβάζεται το μήνυμα κατά την επαφή. Η ακολουθία των επαφών προώθησης διαβάζεται πάλι διαδοχικά, ενώ οι κόμβοι που συμβάλλουν στην προώθηση του μηνύματος, αποθηκεύονται σε μία μειωμένη λίστα, μεγέθους $O(|V|)$ (reduced forwarder list). Όταν ανακτάται μια επαφή που περιλαμβάνει τους κόμβους $(n_k\,,\, n_l)$: α) ένα νέο vertex προστίθεται εάν ακριβώς ένας από τους κόμβους που συναντώνται υπάρχει ήδη στη μειωμένη λίστα, β) δύο νέα vertices προστίθενται εάν και οι δύο κόμβοι που συναντώνται εμφανίζονται ήδη στη μειωμένη λίστα, γ) κανένα vertex δεν προστίθεται εάν η επαφή που περιλαμβάνει τους δύο κόμβους έχει συμβεί ξανά στο παρελθόν.

Επιπλέον, προσθέτουμε κατευθυνόμενες ακμές προς το νέο vertex (ή vertices) από όλα τα υπάρχοντα vertices, που έχουν σε παρένθεση έναν από τους κόμβους που συναντήθηκαν. Κατ’ αυτόν τον τρόπο, οι κατευθυνόμενες ακμές του γράφου δείχνουν πάντα προς τις πιο πρόσφατες επαφές. Ο γράφος G_c που δημιουργείται, είναι κατευθυνόμενος και ακυκλικός (DAG), με σύνολο vertex-επαφών $|V_c|\ =\ O(|V|^2)$ και σύνολο ακμών $|E_c|$=O($|V|^3$). Συνεπώς, ο αλγόριθμος Dijkstra μπορεί να παράγει τις ταχύτερες και κοντύτερες διαδρομές μηνυμάτων προς τον προορισμό d, σε χρόνο $O(|V_c\ |+|E_c|\ )$, αρκεί οι ακμές του γράφου να λάβουν τα σωστά βάρη.

Contact id	Involved nodes		Contact time	Add. Fields
C_1	n_1	n_2	t_1	...
C_2	n_3	n_4	t_2	...
C_3	n_4	n_5	t_3	...
C_4	n_2	n_5	t_4	...
C_5	n_5	n_3	t_5	...
C_6	n_3	n_2	t_6	...
C_7	n_4	n_6	t_7	...
C_8	n_2	n_6	t_8	...

Σχήμα 1: Ακολουθία επαφών C_i $(t_i > t_j, για\ i > j)$. Κάθε γραμμή δείχνει τους κόμβους που συναντώνται και το χρόνο συνάντησης. Οι έντονες εγγραφές δείχνουν τις επαφές προώθησης.



Σχήμα 2: Contact-vertex γράφος επαφών προώθησης. Τα βάρη χρησιμοποιούνται για τον υπολογισμό της ταχύτερης διαδρομής προς τον προορισμό n_6. Όταν αντικαθίστανται με τη μονάδα, χρησιμοποιούνται για τον υπολογισμό του μονοπατιού ελάχιστων μεταβιβάσεων.

4. Συμπεράσματα από τα αποτελέσματα των πειραμάτων

Εξετάσαμε πειραματικά τα σχήματα που βασίζονται στην κεντρικότητα των κόμβων για τη δρομολόγηση μηνυμάτων σε οπορτουνιστικά δίκτυα. Εστιάσαμε στις τρεις θεμελιώδεις εναλλακτικές υπολογισμού που περιγράφηκαν στην Παράγραφο 2. Χρησιμοποιώντας τη μεθοδολογία της Παραγράφου 3, προσομοιώσαμε τη δρομολόγηση 5.000 μηνυμάτων $msg(s,\ d,\ t)$, που δημιουργήθηκαν τυχαία μέσα στα υπό εξέταση traces. Οι τυπικές τιμές για το χρονικό παράθυρο και το κατώφλι ενσωμάτωσης των επαφών στο γράφο επαφών, ήταν $T = 6\text{hrs}$ και thr_e=1. Επιπλέον, προκειμένου να εξετάσουμε την εξάρτηση από το μέγεθος του χρονικού παραθύρου T, χρησιμοποιήσαμε δύο ακόμη τιμές: $T = 2\text{hrs}$ (προεπιλογή του BubbleRap) και το αθροιστικό παράθυρο (προεπιλογή του SimBetTS), όπου μία ακμή μεταξύ δύο κόμβων προστίθεται στο γράφο, αρκεί να υπάρχει τουλάχιστον μία επαφή αυτών στο παρελθόν. Τα αποτελέσματά μας υπήρξαν ποίκιλλα και συνοψίζονται εδώ με τη βοήθεια σχημάτων:

Γενικά συμπεράσματα.

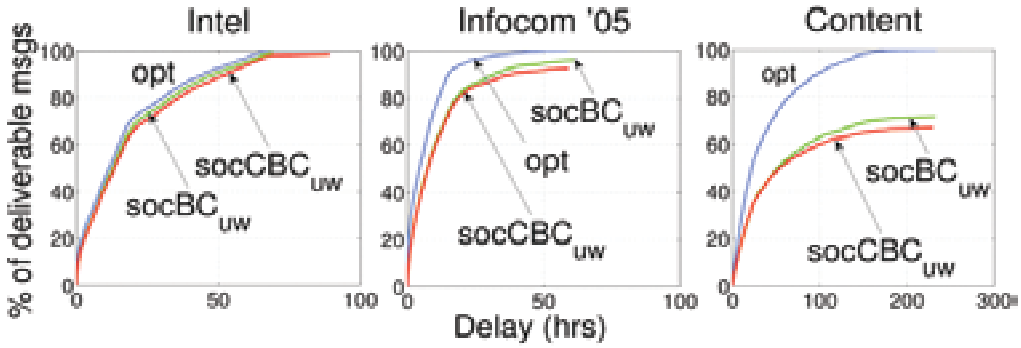
Οι greedy αποφάσεις προώθησης στα σχήματα δρομολόγησης που βασίζονται στη μετρική της κεντρικότητας, ανεξάρτητα από τον τρόπο υπολογισμού της, μπορεί να αποκλίνουν σημαντικά από τις βέλτιστες (opt) και να καταλήγουν σε ιδιαίτερα κακή απόδοση, τόσο από άποψη της καθυστέρησης παράδοσης όσο και των απαιτούμενων μεταβιβάσεων. Η απόδοση ποικίλλει από trace σε trace και εξαρτάται άμεσα από το βαθμό ανάμιξης των προφίλ κινητικότητας των χρηστών. Για παράδειγμα, στο πείραμα της Intel (Σχήμα 3), το οποίο περιλαμβάνει ένα μικρό αριθμό υπαλλήλων μίας εταιρίας, τα κοινωνιοκεντρικά σχήματα δρομολόγησης επιτυγχάνουν διαδρομές πολύ κοντά στις βέλτιστες. Οι κόμβοι κινούνται σε ένα φυσικά περιορισμένο χώρο και συναντώνται συχνά ο ένας με τον άλλον (Πίνακας 1), κάτι που επίσης απεικονίζεται στην αντίστοιχη πυκνότητα ακμών του γράφου των επαφών (Σχήμα 6). Αντίθετα, οι ίδιες τεχνικές αποδίδουν σαφώς χειρότερα στο πολύ αραιότερο Content trace, το οποίο περιλαμβάνει χρήστες που κινούνται γύρω από μία αστική περιοχή, για διάστημα σχεδόν ενός μήνα.

Έως και 30% των μηνυμάτων που θα μπορούσαν να φτάσουν στους προορισμούς τους, βάση των βέλτιστων χωρο-χρονικών μονοπατιών, παγιδεύονται μέσα σε λάθος διαδρομές και δεν φθάνουν ποτέ. Ακόμη όμως και όταν αυτό επιτυγχάνεται, απαιτούνται περισσότερες μεταβιβάσεις (μέχρι πέντε φορές) και καθυστέρηση μεγαλύτερη της μίας ημέρας. Αυτό είναι ξεκάθαρο τόσο στις κατανομές καθυστέρησης παράδοσης, όσο και στις κατανομές πλήθους μεταβιβάσεων.

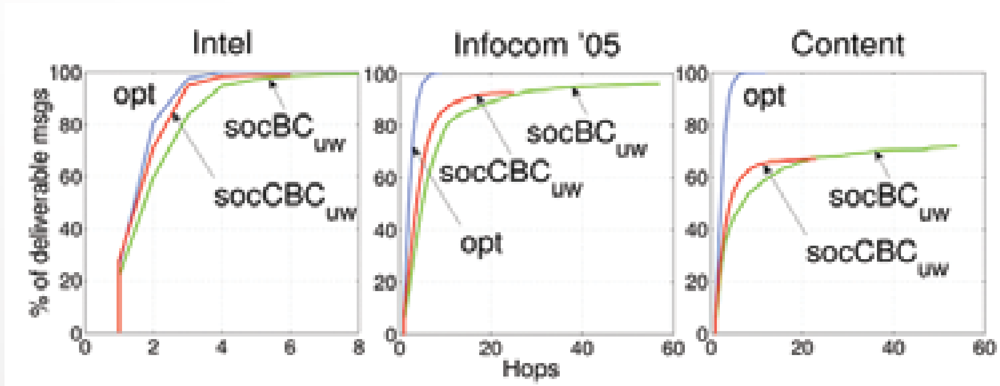
Σε ορισμένα traces, η απόκλιση από το βέλτιστο είναι τόσο μεγάλη που το πιθανοτικό σχήμα προώθησης μηνυμάτων, PF_p, το οποίο αγνοεί οποιαδήποτε πληροφορία σχετικά με τις μετρικές των χρηστών, παρουσιάζει συγκρίσιμη απόδοση ως προς την καθυστέρηση, με τα «πιο ενημερωμένα» κοινωνιοκεντρικά σχήματα. Ωστόσο, η ίδια αυτή απόδοση επιτυγχάνεται με σημαντικά λιγότερες μεταβιβάσεις έναντι των πιθανοτικών σχημάτων, με αποτέλεσμα μικρότερο ενεργειακό κόστος (Σχήμα 5α). Σημειώνεται τέλος, πως όταν η πιθανοτική προώθηση γίνει πιο επιθετική, μπορεί ακόμη και να ξεπερνά ορισμένα κοινωνιοκεντρικά σχήματα, σε βάρος πάντα περισσοτέρων μεταβιβάσεων. Για p=0.5 και p=0.8, τα μηνύματα διανύουν μέχρι πέντε φορές μακρύτερες διαδρομές (Σχήμα 5β).

Σύγκριση BC – CBC.

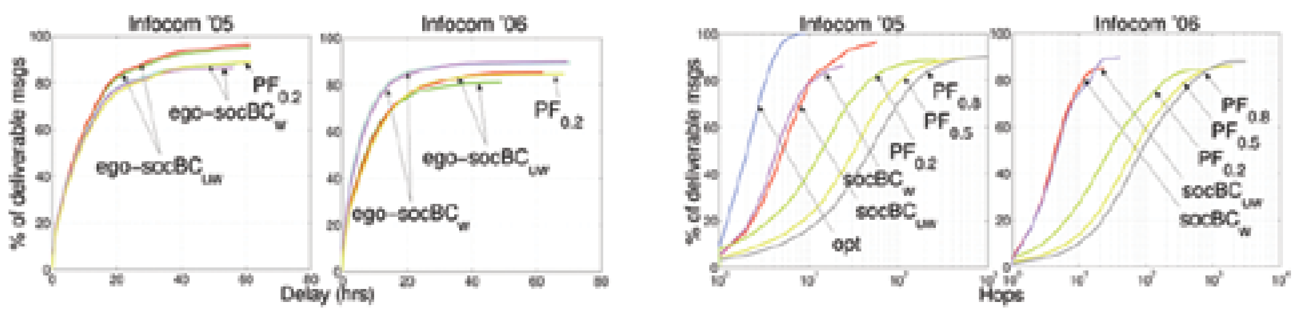
Η αντικατάσταση της μετρικής BC από τη μετρική CBC, η οποία λαμβάνει υπόψη τον προορισμό του μηνύματος, δεν βελτιώνει την καθυστέρηση ούτε την πιθανότητα παράδοσης των μηνυμάτων (Σχήμα 3). Αν και η χρήση του CBC υπονοεί δρομολόγηση μέσω των πιο κατάλληλων ως προς τον προορισμό κόμβων, υπάρχουν πολυάριθμες περιπτώσεις γράφων επαφών, όπου τελικά αποδεικνύεται προβληματική (π.χ. μηδενικές τιμές CBC για πολλούς κόμβους όταν υπάρχουν μη-συνδεδεμένα clusters στο γράφο). Ωστόσο, τα πειράματα έδειξαν ότι η προώθηση με βάση το BC, οδηγεί σε μακρύτερες διαδρομές (περισσότερα hops), έως και 50% σε σχέση με όταν χρησιμοποιείται το CBC (Σχήμα 4), δηλαδή, τα μηνύματα ταξιδεύουν πολύ περισσότερο στο δίκτυο. Κατά συνέπεια, η χρήση του CBC μπορεί να οδηγήσει σε σημαντική εξοικονόμηση ενέργειας-μπαταρίας.



Σχήμα 3: Κατανομή της καθυστέρησης παράδοσης μηνυμάτων για δρομολόγηση με βάση την κεντρικότητα. Destination-aware ($socCBC_{uw}$) και Destination-unaware ($socBC_{uw}$), T=6hrs.



Σχήμα 4: Κατανομή πλήθους μεταβιβάσεων μηνυμάτων για δρομολόγηση με βάση την κεντρικότητα. Destination-aware ($socCBC_{uw}$) και Destination-unaware ($socBC_{uw}$), T=6hrs.



Σχήμα 5: (α) Αριστερά: Κατανομή καθυστέρησης παράδοσης μηνυμάτων.
(β) Δεξιά: Κατανομή πλήθους μεταβιβάσεων. Οι μετρικές BC υπολογίζονται σε γράφους με βάρη και γράφους χωρίς βάρη. Συγκρίνονται με τα δύο σημεία αναφοράς, opt και $PF_{0.2}$. T = 6hrs.

Σύγκριση weighted και unweighted γράφων.

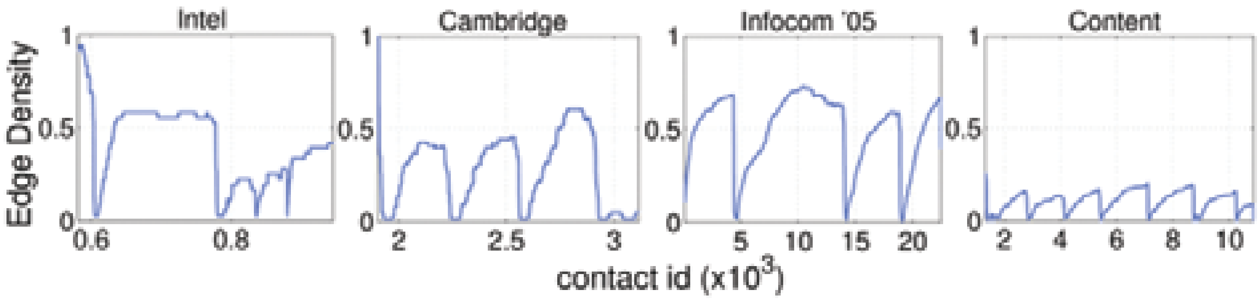
Ο υπολογισμός της μετρικής BC πάνω σε γράφους με βάρη, δεν συνεπάγεται σε κάθε περίπτωση trace βελτίωση της απόδοσης ως προς την καθυστέρηση παράδοσης. Για παράδειγμα, οδηγεί σε 10% περισσότερα παραδοθέντα μηνύματα, κατά τις πρώτες 15 ώρες του πειράματος Infocom'06 (Σχήμα 5α), αλλά υπολείπεται του σχήματος που αγνοεί τη συχνότητα των επαφών, $socBC_{uw}$, στο Infocom'05 (Σχήμα 5α).

Επίδραση του χρονικού παραθύρου ενσωμάτωσης επαφών.

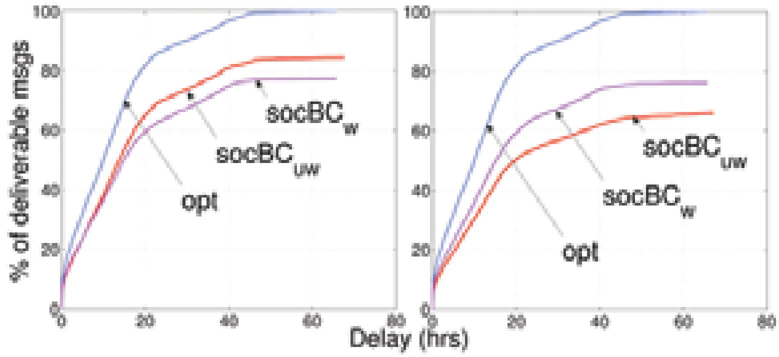
Τα μεγαλύτερα παράθυρα οδηγούν σε μεγαλύτερη πυκνότητα ακμών και η διακύμανση της μετρικής BC τείνει να εξαφανιστεί, με αποτέλεσμα τη μείωση της απόδοσης όλων των τεχνικών. Στην ακραία περίπτωση αθροιστικού παραθύρου ενσωμάτωσης επαφών, ο απλός γράφος χωρίς βάρη γίνεται clique και η δρομολόγηση που βασίζεται στην ιδέα της κεντρικότητας ακυρώνεται. Τα πειράματα έδειξαν (Σχήμα 7) ότι οι γράφοι με βάρη προσδιορίζουν καλύτερα τη χρονική διάσταση και συνιστώνται σε περιπτώσεις που χρησιμοποιείται αθροιστικό παράθυρο ενσωμάτωσης επαφών.

Socio-BC vs ego-BC.

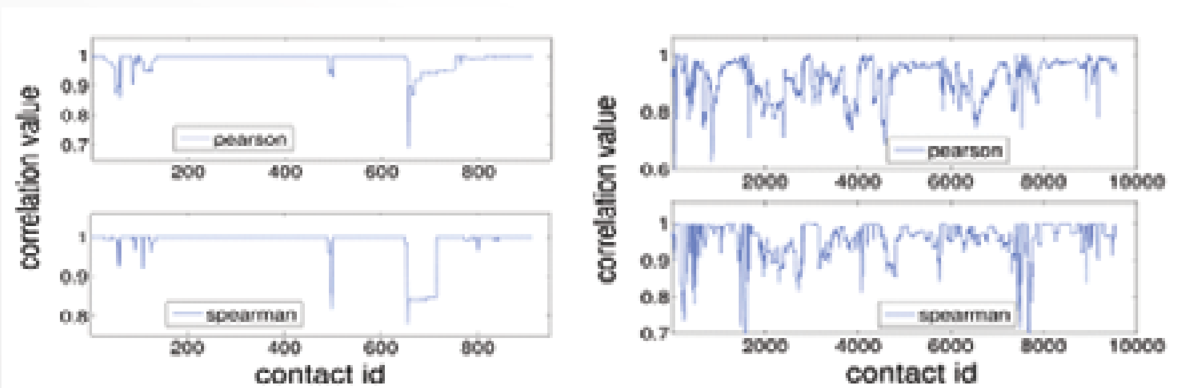
Η χρησιμοποίηση της εγωκεντρικής μετρικής BC μειώνει ελάχιστα την απόδοση, όταν υπολογίζεται σε γράφους χωρίς βάρη, ενώ ακόμα λιγότερο, όταν υπολογίζεται σε γράφους με βάρη (π.χ. Σχήμα 5). Συνεπώς, πρόσθετη πληροφόρηση, που προκύπτει από την ενσωμάτωση όλων των επαφών του δικτύου στο γράφο, δεν ωφελεί σημαντικά τις αποφάσεις προώθησης (μόνο 6% περισσότερα παραδοθέντα μηνύματα σε όλες τις περιπτώσεις). Αυτό το συμπέρασμα ενισχύεται περαιτέρω και από τον ισχυρά θετικό συσχετισμό μεταξύ της κοινωνιοκεντρικής και εγωκεντρικής BC σχεδόν σε όλα τα traces, τόσο σε απόλυτες τιμές (PEARSON), όσο και στην κατάταξη (Spearman) (βλ. Σχήμα 8 και 9).



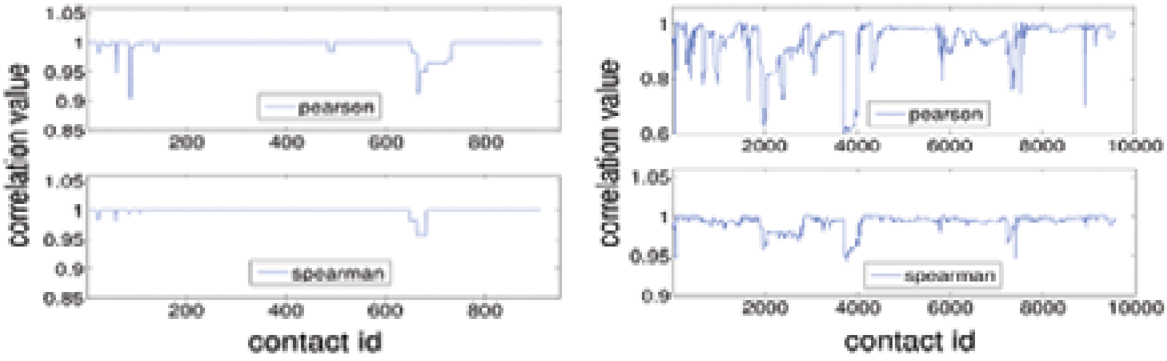
Σχήμα 6: Χρονική εξέλιξη (από επαφή σε επαφή) της πυκνότητας ακμών του γράφου.



Σχήμα 7: Cambridge dataset: Κατανομή καθυστέρησης παράδοσης μηνυμάτων για δύο διαφορετικά παράθυρα, T=6hrs (αριστερά) και αθροιστικό χρονικό παράθυρο (δεξιά).



Σχήμα 8: Συσχέτιση $socBC_{uw} - egoBC_{uw}$, για τα Intel (αριστερά) και Content (δεξιά) traces.



Σχήμα 9: Συσχέτιση $socBC_w - egoBC_w$, για τα Intel (αριστερά) και Content (δεξιά) traces.

Αναφορές

1. C. Boldrini et al. Hibop: a history based routing protocol for opportunistic networks. In IEEE WoWMoM, pages 1–12, 2007.

2. B. Bui-Xuan et al. Computing shortest, fastest, and foremost journeys in dynamic networks. Int'l Jnl of Foundations of Computer Science, 14(2), April 2003.

3. A. Chaintreau et al. Impact of human mobility on the design of opportunistic forwarding algorithms. In Proc. IEEE INFOCOM '06, pages 1–13, April 2006.

4. E. M. Daly and M. Haahr. Social network analysis for information flow in disconnected delay-tolerant manets. IEEE Trans. Mob. Comput., 8(5), May 2009.

5. M. Everett and al. Ego network betweenness. Social Networks, 27(1):31–38, 2005.

6. T. Hossmann et al. Know thy neighbor: Towards optimal mapping of contacts to social graphs for dtn routing. In IEEE Infocom'10, San Diego, USA, 2010.

7. P. Hui et al. Bubble rap: Social-based forwarding in delay tolerant networks. IEEE Trans. Mob. Comput., (To Appear) 2011.

8. A. Lindgren et al. Probabilistic routing in intermittently connected networks. Lecture Notes in Computer Science, 3126:239–254, Jan. 2004.

9. P. Pantazopoulos et al. Efficient social-aware content placement for opportunistic networks. In IFIP/IEEE WONS, Kranjska Gora, Slovenia, February, 3-5 2010.

10. T. Spyropoulos et al. Efficient routing in intermittently connected mobile networks: The Multiple-Copy case. IEEE/ACM Trans. Netw., 16(1):77–90, Feb. 2008.

11. A. Vahdat and D. Becker. Epidemic routing for partially connected ad hoc networks.

Technical report, Duke University, April 2000.

12. P. Nikolopoulos, T. Papadimitriou, P. Pantazopoulos, M. Karaliopoulos, I. Stavrakakis. How off-center are centrality metrics for opportunistic routing? ACM MobiCom workshop on Challenged Networks (CHANTS 2011), Las Vegas, USA, September 2011.