ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Μηχανισμοί πρόβλεψης προσήμων σε προσημασμένα μοντέλα κοινωνικών δικτύων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Δημήτριος Σ. Κανελλόπουλος

Επιβλέπων: κα. Βαρβαρίγου Θεοδώρα
Καθηγήτρια

Αθήνα, Οκτώβριος 2015

Αθήνα, Οκτώβριος 2015

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Μηχανισμοί πρόβλεψης προσήμων σε προσημασμένα μοντέλα κοινωνικών δικτύων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Δημήτριος Σ. Κανελλόπουλος

Επιβλέπων: κα. Βαρβαρίγου Θεοδώρα
Καθηγήτρια

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 29 Οκτωβρίου 2015.

.....	.....	.....
Θεοδώρα Βαρβαρίγου	Εμμανουήλ Βαρβαρίγος	Βασίλειος Λούμος
Καθηγήτρια	Καηγητής	Καθηγητής

Αθήνα, Οκτώβριος 2015

.....

Δημήτριος Σ. Κανελλόπουλος

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

Copyright © Δημήτριος Σ. Κανελλόπουλος, 2015

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περιεχόμενα

Ευχαριστίες 9
Περίληψη 10
Abstract 11
1. Εισαγωγή 12
1.1 Γενική ανασκόπηση 12
1.2 Στόχος παρούσας εργασίας 14
1.3 Διάρθρωση κειμένου 14
2.Πρόβλημα- θέμα προς μελέτη 16
2.1 Τα Κοινωνικά δίκτυα από την εμφάνισή τους μέχρι σήμερα. 16
2.2 Σχέσεις μεταξύ μελών 18
2.3 Μηχανές προτάσεων (Recommender Systems, RSs). 19
2.4 Ευστάθεια 20
2.5 Πρόβλεψη για τις σχέσεις μεταξύ μελών 22
2.5.1 Βασικές αρχές πρόβλεψης 23
2.5.2 Σημασία πρόβλεψης 23
3. Αντικείμενο υπό διερεύνηση 25
3.1 Σχετικό ερευνητικό έργο 25
3.2 Προτεινόμενη προσέγγιση 25
4.Μεθοδολογία- Υλικό 27
4.1 Μοντελοποίηση 27
4.2 Μεθοδολογίες 29
4.2.1 Πρόβλεψη βασισμένη στην δομική ευστάθεια 29
4.2.2 Πρόβλεψη βασισμένη στην αναγνώριση συχνών υπο-γράφων 30
4.2.3 Σύγκριση των δύο μεθόδων 31
4.3 Εργαλεία υλοποίησης 31
4.3.1 Slashdot 31
4.3.2 Java 33
4.3.3 Netbeans 36
5.Δεδομένα που συλλέχθηκαν 38

5.1 Μέθοδος βασισμένη στην δομική ευστάθεια 38
5.2 Μέθοδος βασισμένη στη συχνή εμφάνιση υπο-γράφων 38
5.3 Πρόταση βελτίωσης 38
6. Συμπεράσματα και μελλοντικά ζητήματα 46
6.1 Συμπεράσματα 46
6.2 Μελλοντικά θέματα προς έρευνα 46
Γλωσσάριο 48
Βιβλιογραφία 49

Περιεχόμενα εικόνων

Εικόνα 1: Το internet of things 13

Εικόνα 2: Λογότυπο Sixdegrees 15

Εικόνα 3: Τα λογότυπα των πιο δημοφιλών online social networks (OSNs)- Facebook,linked in, Flickr, MySpace, Pinterest, Wikipedia, Epinions, twitter, Bebo,tumbler 15

Εικόνα 4: Τα πιο δημοφιλή κοινωνικά δίκτυα τον Αύγουστο του 2015 16

Εικόνα 5: Το λογότυπο του κοινωνικού δικτύου Facebook 16

Εικόνα 6: Το πλήθος των ενεργών χρηστών του Facebook από το 2008 μέχρι σήμερα 17

Εικόνα 7: Η εξέλιξη της φιλίας μιας τριάδας: ενός παντρεμένου ζευγαριού και ενός φίλου τους 20

Εικόνα 8: Στις μη ισορροπημένες τριάδες χρειάζεται ένα βήμα αλλαγής σχέσης ώστε να αποκατασταθεί η ισορροπία.....21

Εικόνα 9: Screenshots από το αρχείο δεδομένων του κοινωνικού δικτύου Slashdot (43).....27

Εικόνα 10: Ως ego-graph ορίζεται ο υπο-γράφος που περιλαμβάνει μόνο τους κοινούς γείτονες του A και του B 28

Εικόνα 11: Ενδεικτικός γράφος που εξετάζεται. Ο κόμβος C είναι ένας κοινός γειτονικός κόμβος των A και B.....29

Εικόνα 12: Λογότυπο της ιστοσελίδας Slashdot.com 30

Εικόνα 13: Χάρτης των σχέσεων των χρηστών δείγματος στο Slashdot Zoo. Το δίκτυο περιέχει 77.985 χρήστες και 510.157 σχέσεις 31

Εικόνα 14: Screenshot από το δημοφιλές site slashdot.com 32

Εικόνα 15: Λογότυπο γλώσσας προγραμματισμού Java 32

Εικόνα 16: Λογότυπο εφαρμογής Netbeans 35

Εικόνα 17: Screenshot του γραφικού περιβάλλοντος προγραμματισμού Netbeans IDE 35

Εικόνα 18: Ποσοστά επιτυχίας μεθόδων βιβλιογραφίας στο δείγμα που δίνει ασφαλή πρόβλεψη 39

Εικόνα 19: Ποσοστά επιτυχίας των μεθόδων της βιβλιογραφίας 40

Εικόνα 20: Ποσοστά επιτυχίας πρόβλεψης σε ολόκληρο το δείγμα 40

Εικόνα 21: Διάγραμμα που απεικονίζει την αναλογία followers και friends 41

Εικόνα 22: Η διασπορά του μεγέθους των ego-graphs.....42

Εικόνα 23: Πλήθος ego-graphs σε συνάρτηση με το μέγεθος τους (100%).....43

Εικόνα 24: Πλήθος ego-graphs σε συνάρτηση με το μέγεθος τους (<1000).....43

Εικόνα 25: Πλήθος ego-graphs σε συνάρτηση με το μέγεθος τους (< 100).....44

Ευχαριστίες

Με την ευκαιρία της ολοκλήρωσης αυτής της διπλωματικής εργασίας, θα ήθελα να ευχαριστήσω θερμά την Καθηγήτρια Ε.Μ.Π. κ. Θεοδώρα Βαρβαρίγου τόσο για τις συμβουλές της σχετικά με την διπλωματική μου όσο και για την γενικότερη πολύτιμη καθοδήγησή της. Επίσης, θα ήθελα να ευχαριστήσω ιδιαίτερα τον Δρ. Κωνσταντίνο Τσερπέ για την συνεχή ενθάρρυνση και καθοδήγηση που μου παρείχε κατά την διάρκεια της εκπόνησης της εργασίας.

Περίληψη

Σε αυτήν την εργασία γίνεται μελέτη των σχέσεων των χρηστών σε ένα σύγχρονο κοινωνικό δίκτυο. Δεδομένης της αυξημένης χρήσης τους, από ολοένα και περισσότερους χρήστες είναι ένα ενδιαφέρον πεδίο προς διερεύνηση. Επιπλέον η δραστηριοποίηση των πιο δημοφιλών δικτύων σε νέους τομείς όπως η διαφημίσεις και οι ηλεκτρονικές αγορές, καθιστούν την μελέτη των ιδιοτήτων τους ακόμα πιο ενδιαφέρουσα και σημαντική.

Για τον σκοπό αυτό επιλέξαμε ένα δίκτυο που έχει μια ιδιαιτερότητα σε σχέση με τα άλλα. Επιτρέπει στους χρήστες του να βαθμολογούνται θετικά και αρνητικά μεταξύ τους. Στην συνέχεια αναπτύξαμε ένα πρόγραμμα προσομοίωσης που μας επιτρέπει να προβλέψουμε την κάθε σχέση μεταξύ όλων των χρηστών.

Οι μέθοδοι που εφαρμόστηκαν βασίζονται σε παλιότερη βιβλιογραφία. Η πρώτη μέθοδος σχετίζεται με την δομική ευστάθεια των δικτύων και η δεύτερη βασίζεται στην πρόβλεψη με την μέθοδο εύρεσης συχνών υπογράφων.

Συγκρίναμε τα αποτελέσματα των δυο μεθόδων και περιγράψαμε τα πλεονεκτήματα και τα μειονεκτήματα της κάθε μεθόδου. Στην συνέχεια της εργασίας μας προσπαθήσαμε να βελτιώσουμε την απόδοση ενός συστήματος πρόβλεψης σχέσεων προτείνοντας ένα συνδυασμό των 2 μεθόδων με αποτελεσματικό τρόπο. Στο τέλος συγκεντρώνουμε όλα τα αποτελέσματα και καταλήγουμε ως προς την απόδοση του συστήματος που προτάθηκε.

<u>Λέξεις κλειδιά</u>

κοινωνικά δίκτυα, σύστημα πρόβλεψης σχέσης, δομική ευστάθεια, συχνοί υπογράφοι

Abstract

In this dissertation, we study the connections between users of online social networks. Due to the rapid growth of the social networks these past years the number of the total registered and daily active users and is remarkably high. Thus it has become an interesting field of research. Another factor contributing to the importance of this field is the recent attempt to make online shopping possible inside the network which results in creating new paths for the social networks in general.

We chose to examine a network which has a unique characteristic in comparison to other similar sites. It gives its users the ability to mark others users not only friends but also foes. We studied this network and developed a program to simulate and predict the connections, positive or negative, between the users.

The fundamental methods used for the prediction are based on previous research on this field. We made use of the structural balance theory for edge sign perdition and then we used a method based on frequently discovered subgraphs.

After extracting the results from the two above mentioned methods we attempted to compared them and analyze their pros and cons. Subsequently we proposed a system that will utilize both methods in order increase the accuracy of the predictions made. Finally we collected all the data and concluded on the throughput if this system.

Keywords

Social networks, edge sign prediction, structural balance, frequent subgraph discovery

1. Εισαγωγή

1.1 Γενική ανασκόπηση

Το διαδίκτυο αποτελεί, πλέον αναπόσπαστο κομμάτι της καθημερινότητάς μας. Πριν από μερικά χρόνια για να το περιγράψει κανείς, αρκούσε να πει πως είναι μια απέραντη πηγή πληροφοριών. Τα τελευταία χρόνια όμως έχει εξελιχθεί με τέτοιο τρόπο, ώστε η παραπάνω περιγραφή δεν αρκεί.

Το διαδίκτυο στην σημερινή του μορφή, είναι ένα δίκτυο αποτελούμενο από δίκτυα *(1)*.Είναι μια απέραντη πηγή εφαρμογών και υπηρεσιών (internet of things). Αποτελεί πλέον, ένα δυναμικό εργαλείο που προσφέρει εφαρμογές, το πλήθος και η ποικιλία των οποίων είναι αχανής. Ενδεικτικά υπάρχουν εφαρμογές που θεωρούνται σήμερα αυτονόητες (mail), εφαρμογές που γίνονται σιγά-σιγά απαραίτητες στον σύγχρονο άνθρωπο (e-banking, e-shops) και εφαρμογές που προσπαθούν να ανταγωνιστούν και να αντικαταστήσουν κλασικά προϊόντα (web streaming TV, news channels, VOIP). Έχει γίνει πολύτιμο εργαλείο για πολλά επαγγέλματα και μάλιστα τα τελευταία χρόνια έχει δημιουργήσει καινούρια (Video Bloggers, News Channel Journalists). Ζούμε σε μια εποχή που το διαδίκτυο εξαπλώνεται και δημιουργεί συνεχώς νέες εφαρμογές.

Στο πλαίσιο των υπηρεσιών του διαδικτύου, οι χρήστες λαμβάνουν πολυάριθμες πληροφορίες άμεσα και έμμεσα. Άμεσα, όταν κάνουν μια στοχευμένη αναζήτηση, η εφαρμογή που χρησιμοποιούν αποφασίζει ποια αποτελέσματα θα εμφανίσει και με ποια σειρά. Έμμεσα δέχονται πληροφορίες με την μορφή διαφημίσεων και προτάσεων, που στοχεύουν να εμφανίσουν στον χρήστη το κατάλληλο ποιόν και την κατάλληλη είδηση αντίστοιχα. Κρίνοντας από τα δεδομένα του χρήστη που είναι διαθέσιμα (ποιους χρήστες εμπιστεύεται, ποιους δικτυακούς τόπους επισκέπτεται συχνότερα κλπ.) αποφασίζουν ποιο προϊόν και ποια είδηση είναι πιο πιθανό να ενδιαφέρει τον χρήστη.

Είναι λοιπόν φανερό ότι πραγματοποιείται μια συνεχής διάδοση πληροφοριών προς τους χρήστες του διαδικτύου, η οποία είναι πολύ σημαντική. Αν οι προτάσεις είναι οι σωστές, επωφελούνται οι ίδιοι οι χρήστες εφόσον βρίσκουν εύκολα και γρήγορα αξιόπιστες πληροφορίες. Φυσικά στην περίπτωση των προϊόντων, επωφελούνται και οι επιχειρήσεις που που προβάλλονται στο σωστό κοινό που θα ενδιαφερθεί να προμηθευτεί τα προϊόντα τους. Στην αντίθετη περίπτωση της λανθασμένης πρότασης, ο χρήστης παραπληροφορείτε και δεν βρίσκει την σωστό και αξιόπιστο δεδομένο που αναζητά.

Εδώ είναι σημαντικό να αναφέρουμε μία μεγάλη αλλαγή που παρατηρείται στο διαδίκτυο τα τελευταία χρόνια και ονομάζεται Internet of Things, γνωστό επίσης και ως internet of objects. Σε γενικές γραμμές, το IoT αναφέρεται στη δικτυακή διασύνδεση των καθημερινών αντικειμένων τα οποία συχνά είναι εξοπλισμένα με “νοημοσύνη”.

Ο όρος Internet of Things επινοήθηκε για πρώτη φορά από τον Kevin Ashton, το 1999 *(2,3,4)*. Ωστόσο, την τελευταία δεκαετία, ο όρος τείνει να περιλαμβάνει όλο και περισσότερες εφαρμογές όπως σε υπηρεσίες, υγείας, , επιχειρήσεις κοινής ωφέλειας, μέσα μεταφοράς, κλπ. *(2,3,5)*. Παρά το γεγονός ότι ο ορισμός των «πραγμάτων» έχει αλλάξει, καθώς η τεχνολογία

εξελίσσεται, ο κύριος στόχος, του να μεταδίνει πληροφορίες σε ένα υπολογιστικό σύστημα χωρίς ανθρώπινη παρέμβαση, παραμένει ο ίδιος. Πρόκειται για μία επαναστατική εξέλιξη του σύγχρονου διαδικτύου σε ένα δίκτυο διασυνδεόμενων αντικειμένων και συσκευών που θα επικοινωνούν τόσο με τον ίδιο τον άνθρωπο όσο και με άλλες συσκευές. Το δίκτυο αυτό των αντικειμένων όχι μόνο θα αντλεί πληροφορίες από το περιβάλλον και θα αλληλοεπιδρά με τον φυσικό κόσμο αλλά επιπλέον θα χρησιμοποιεί τα υφιστάμενα πρότυπα του Διαδικτύου για την παροχή υπηρεσιών που αφορά την μεταφορά πληροφοριών , analytics, εφαρμογών και επικοινωνιών.

Εικ. 1 Το internet of things

Ζούμε σε μια εποχή που το internet αλλάζει μορφή. Στο εγγύς μέλλον διαφαίνεται ότι θα πραγματοποιηθούν πολλές αλλαγές στον τρόπο που οργανώνεται το internet ενώ γενικότερα όπως φαίνεται και από τα παραπάνω κινούμαστε προς μια κατεύθυνση όπου οι εφαρμογές τείνουν να γίνουν πιο «έξυπνες» προς όφελος του ανθρώπου και της διευκόλυνσης της διακίνησης πληροφοριών με την γενικότερη έννοια.

Μια εφαρμογή με ραγδαία εξάπλωση τα τελευταία χρόνια είναι τα κοινωνικά δίκτυα (social networks). Ορίζουμε ως κοινωνικά δίκτυα τους ιστοτόπους που δίνουν την υπηρεσία σε άτομα, να δημιουργήσουν ένα δημόσιο ή ημι-δημόσιο λογαριασμό (profile account) μέσα σε ένα δεδομένο σύστημα, να προσδιορίσει μια λίστα από χρήστες με τους οποίους θα μοιράζονται μια σχέση, και να βλέπουν και να αναζητούν την λίστα σχέσεων που διατηρούν άλλοι χρήστες μέσα στο ίδιο σύστημα. Το είδος και η ονοματολογία των σχέσεων ποικίλει ανάλογα με την περίπτωση.*(6)*

Πρακτικά, τα online κοινωνικά δίκτυα είναι εφαρμογές που χτίζουν εικονικά κοινωνικά δίκτυα και εικονικές κοινωνικές σχέσεις μεταξύ ανθρώπων που μοιράζονται τα ίδια ενδιαφέροντα και πιθανώς πραγματικές κοινωνικές σχέσεις. Κάθε κοινωνικό δίκτυο αναπαριστά έναν χρήστη με μια ιστοσελίδα που αποτελείται από κείμενο, συνδέσμους προς άλλους διαδικτυακούς τόπους,

και από μια πληθώρα αρχείων πολυμέσων. Επιτρέπουν στους χρήστες να συνδιαλέγονται μεταξύ τους και να δημιουργούν λίστες χρηστών, με τους οποίους μοιράζονται σχέσεις. Οι περισσότερες εφαρμογές είναι web based και δίνουν στους χρήστες τους την δυνατότητα να ανταλλάσσουν μηνύματα άμεσα.

Οι χρήστες των κοινωνικών δικτύων, τα χρησιμοποιούν με διαφορετικούς τρόπους κάθε φορά. Ένας από τις πιο απλούς είναι η συνομιλία με άλλους χρήστες. Σε άλλες περιπτώσεις, κοινωνικά δίκτυα χρησιμοποιούνται από χρήστες για να αξιολογήσουν προϊόντα τεχνολογίας και να αποφασίσουν μια αγορά, για να εντοπίσουν μια θέση εργασίας κ.α.

Είναι σίγουρο πως τα μέσα κοινωνικής δικτύωσης είναι πολύ σημαντικά στην ζωή των ανθρώπων που έχουν εντάξει το διαδίκτυο στην καθημερινότητά τους.

1.2 Στόχος παρούσας εργασίας

Στόχος της παρούσας εργασίας είναι να μελετήσει τις σχέσεις που αναπτύσσουν οι χρήστες μεταξύ τους μέσα σε ένα online κοινωνικό δίκτυο. Θα κάνουμε μια προσομοίωση η οποία θα προσπαθεί να προβλέψει μια δεδομένη σχέση, κρίνοντας από το υποδίκτυο των άμεσων σχέσεων που συνορεύουν με αυτήν. Θα εφαρμόσουμε έναν αριθμό προσομοιώσεων που θα διαφέρουν μεταξύ τους ως προς τον τρόπο πρόβλεψης και από την συλλογή και επεξεργασία των δεδομένων θα επιχειρήσουμε να συγκρίνουμε την αποτελεσματικότητα των διαφόρων τεχνικών που χρησιμοποιήθηκαν.

1.3 Διάρθρωση κειμένου

Στην επόμενη ενότητα της εργασίας αυτής (2η Πρόβλημα- θέμα προς μελέτη) θα αναφερθούμε στα κοινωνικά δίκτυα γενικότερα και την σημασία αυτών στην καθημερινή ζωή μας, καθώς και την σημασία των σχέσεων που αναπτύσσουν μεταξύ τους οι χρήστες τους. Για τον σκοπό της μελέτης αναλύεται η διαδικασία πρόβλεψης και η σημασία αυτής. Γίνεται επίσης αναφορά στις δύο κύριες μεθόδους στις οποίες βασίζεται η πρόβλεψη.

Έπειτα στην 3η ενότητα (Αντικείμενο υπό διερεύνηση) θα κάνουμε μια αναφορά στις προηγούμενες ερευνητικές εργασίες που έχουν ασχοληθεί με το ζήτημά μας και τις οποίες χρησιμοποιήσαμε σαν θεωρητική βάση για την μελέτη μας.

Στην 4η ενότητα (Μεθοδολογία – Υλικό) εξηγούμε την μεθοδολογία της δικής μας εργασίας και την συλλογιστική πορεία που ακολουθήσαμε. Επιπλέον θα αναλύσουμε τα εργαλεία που χρησιμοποιήσαμε για τον σκοπό της εργασίας.

Στη 5η ενότητα (Δεδομένα που συλλέχθηκαν), συλλέγουμε τα δεδομένα από τα πειράματά μας και εξηγούμε την διαδικασία που ακολουθήσαμε.

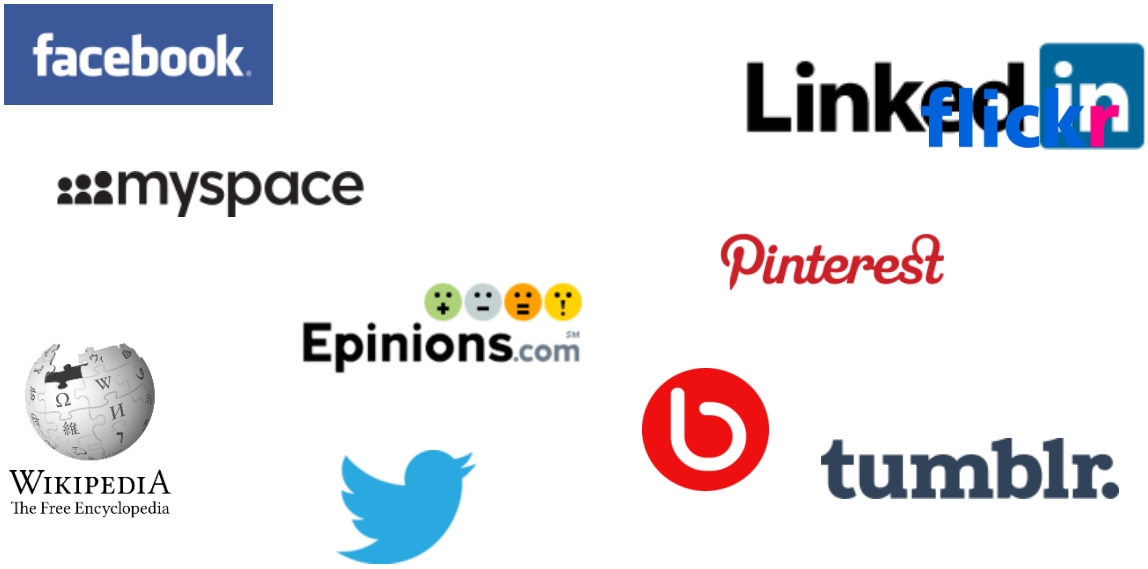
Τέλος στην 6η ενότητα (Συμπεράσματα και μελλοντικά ζητήματα) καταλήγουμε στα συμπεράσματά μας ως προς το σύστημα που προτείναμε και αναφέρουμε τις ευκαιρίες που δίνονται για περεταίρω μελέτη πάνω στο συγκεκριμένο αντικείμενο.

2.Πρόβλημα- θέμα προς μελέτη

2.1 Τα Κοινωνικά δίκτυα από την εμφάνισή τους μέχρι σήμερα.

Εικ. 2 Λογότυπο Sixdegrees

Το πρώτο κοινωνικό δίκτυο δημιουργήθηκε το 1997. Πρόκειται για το "SixDegrees.com" που έδινε στους χρήστες του την δυνατότητα να δημιουργούν δικό τους λογαριασμό (profile account) να απαριθμούν και να καταγράφουν τους φίλους τους καθώς και, από τις αρχές του 1998, να περιηγούνται στις καταγραφές των φίλων τους. Κάθε μια από αυτές τις εφαρμογές βεβαίως προϋπήρχε ,σε κάποια μορφή, και πριν την δημιουργία του SixDegrees. *(6)*



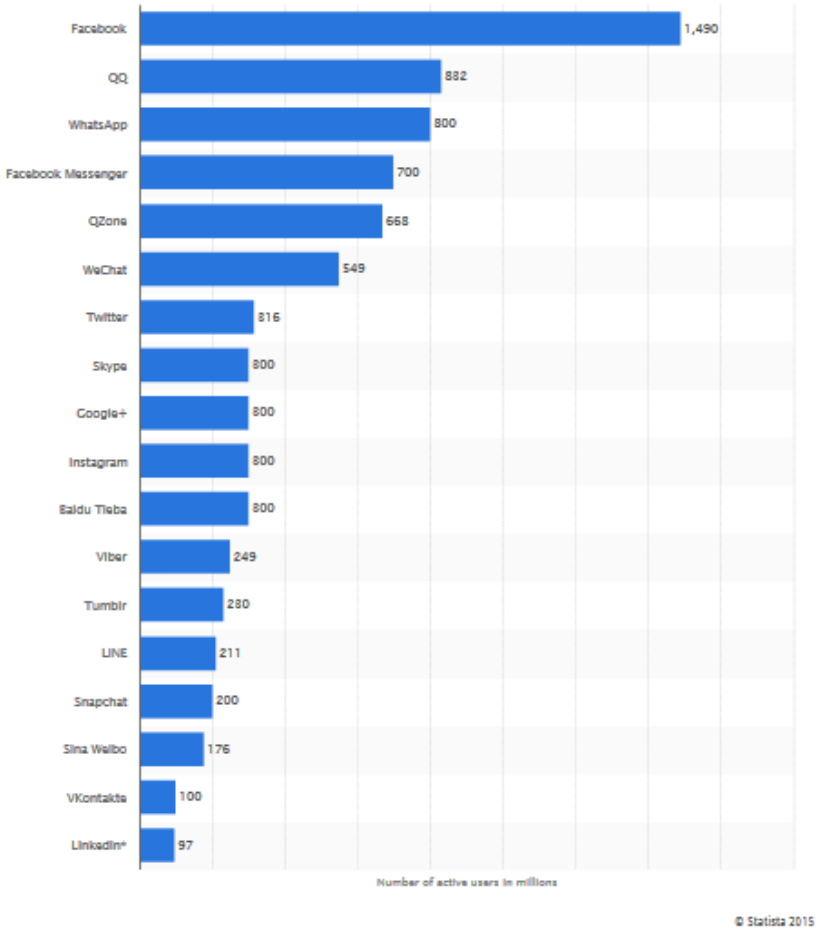
Εικ. 3 Τα λογότιπα των πιο δημοφιλών online social networks (OSNs)- facebook,linked in, flickr, myspace, pinterest, wikipedia, Epinions, twitter, Bebo,tumbler

Από τότε που δημιουργήθηκαν για πρώτη φορά, τα κοινωνικά δίκτυα, όπως MySpace, Facebook, Cyworld και Bebo έχουν προσελκύσει εκατομμύρια χρηστών, πολλοί από τους οποίους έχουν βάλει την χρήση αυτών των ιστοτόπων στην καθημερινή τους ζωή *(7)*. Οι χρήστες τους, ανταλλάσσουν μηνύματα μεταξύ τους δημοσιεύουν πληροφορίες, αξιολογούν δημοσιεύσεις άλλων χρηστών κοκ. Η εξάπλωση είναι τόσο μεγάλη, ώστε πολλές επιχειρήσεις δημιουργούν λογαριασμούς για να διαφημιστούν και μάλιστα με επιτυχία.

Υπάρχουν πολλές παραλλαγές κοινωνικών δικτύων, ανάλογα με το αν δίνουν υπηρεσίες συνδεσιμότητας μέσω κινητών τηλεφώνων, δυνατότητα διαμοιρασμού φωτογραφιών, βίντεο,

και δημοσίευσης άρθρων. Με την ευρεία έννοια, τα κοινωνικά δίκτυα είναι «ομαδο-κεντρικά», δηλαδή προτρέπουν τους χρήστες να δημιουργούν κύκλους γνωριμιών και να μοιράζουν ιδέες, ειδήσεις, αρχεία πολυμέσων μέσα στον κύκλο τους. Αυτό το χαρακτηριστικό τα διαχωρίζει από τις υπηρεσίες στιγμιαίων μηνυμάτων (Instant Messaging – IM) και από τους χώρους συζητήσεων (forums).

Παρακάτω απεικονίζονται τα πιο δημοφιλή κοινωνικά δίκτυα τον Αύγουστο του 2015. Πιο δημοφιλές αυτήν την στιγμή είναι το δίκτυο του Facebook που ξεπερνά το 1 δισεκατομμύριο ενεργούς χρήστες. Δέκατο είναι το twitter με πάνω από 316 εκατομμύρια ενεργούς χρήστες κατά μήνα. *(8)*

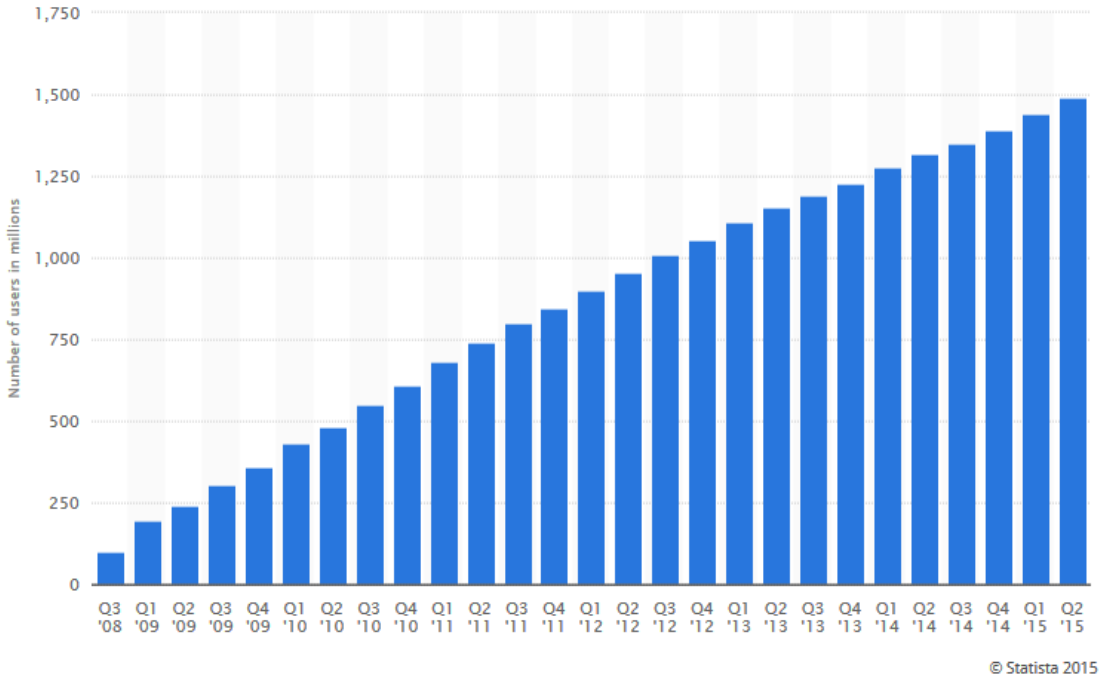


Εικ. 4 Τα πιο δημοφιλή κοινωνικά δίκτυα τον Αύγουστο του 2015

Εικ. 5 Το λογότυπο του κοινώνικού δικτύου facebook

Το κοινωνικό δίκτυο Facebook δημιουργήθηκε τον Φεβρουάριο του 2004 από τους Mark Zuckerberg ,Eduardo Saverin, Andrew McCollum, Dustin Moskovitz και Chris Hughes*(9, 10)*.

Στατιστικές παρουσιάζουν ότι το 2ο τρίμηνο του 2015 το Facebook είχε 1.49 δισεκατομμύρια ενεργούς χρήστες, θεωρώντας ως ενεργούς χρήστες εκείνους που συνδέθηκαν στον λογαριασμό τους τις τελευταίες 30 ημέρες. Ο αντίστοιχος αριθμός ενεργών χρηστών το τρίτο τρίμηνο του 2012 ήταν 1 δισεκατομμύριο, όπως παριστάνεται και στο παρακάτω γράφημα *(11).*



Εικ. 6 Το πλήθος των ενεργών χρηστών του Facebook από το 2008 μέχρι σήμερα

2.2 Σχέσεις μεταξύ μελών

Μέσα στα online social networks (OSNs) ο κάθε χρήστης παροτρύνεται από το σύστημα να συνάψει σχέσεις με άλλους χρήστες. Οι σχέσεις αυτές μπορεί να έχουν είτε θετικό είτε και αρνητικό χαρακτήρα. Επιπλέον, μπορεί είτε να αντικατοπτρίζουν πραγματικές σχέσεις (ο χρήστης δηλαδή να συνδέεται με άτομα που γνωρίζει προσωπικά) είτε να βασίζονται σε κοινά ενδιαφέροντα χωρίς να υπάρχει προηγούμενη γνωριμία μεταξύ τους. Ορισμένα social networks προσφέρουν την δυνατότητα κατηγοριοποίησης των συνδέσεων ενός χρήστη ανάλογα με την προέλευση και την φύση της σχέσης(πχ linked in) Γενικά εκτός από το να υποδηλώνουν φιλία ή εχθρότητα, οι σχέσεις μεταξύ των μελών ενός δικτύου μπορεί να αφορούν παρουσία ή η έλλειψη εμπιστοσύνης, δηλαδή η αναγνώριση της εγκυρότητας της γνώμης ενός χρήστη από άλλους χρήστες του ίδιου μέσου *(12)*, υποστήριξη ή διαφωνία σχετικά με μια άποψης κ.α. *(13)*

Σε ορισμένες πλατφόρμες κοινωνικών δικτύων οι σχέσεις μεταξύ των μελών μπορεί να είναι είτε αμφίδρομες είτε μονόπλευρες. Ο χρήστης μπορεί για παράδειγμα να επιλέγει να ακολουθεί έναν άλλον χρήστη ο οποίος μπορεί είτε να τον ακολουθεί και εκείνος είτε όχι. Ο χρήστης επιλέγει πόσες σχέσεις θα συνάψει με άλλους χρήστες και όσες περισσότερες κάνει και διατηρεί θεωρείται ότι έχει μεγαλύτερη επιρροή καθώς οτιδήποτε δημοσιεύει έχει μεγαλύτερη προβολή

και ως εκ τούτου μεγαλύτερη πιθανότητα «αναδημοσίευσης». (κάτι που θα αναφερθεί και στην συνέχεια στην θεωρεία κοινωνικής θέσης)

Σε κάποιες ιστοσελίδες κοινωνικής δικτύωσης, όπως το Slashdot και Epinions οι χρήστες έχουν την δυνατότητα να βαθμολογούν άλλους χρήστες ή αντικείμενα μέσα στη σελίδα ή να ονοματίζουν τους άλλους χρήστες ως φίλους ή εχθρούς *(14, 15).* Αυτό το είδος δικτύου που σχηματίζεται αναπαρίσταται εύκολα με την χρήση γράφων. Οι βαθμολογίες χρηστών αναπαρίστανται με βάρη στις ακμές των γράφων *(14,15,16).* Αν θέλουμε να αναπαραστήσουμε την κατάσταση «φίλος- εχθρός» χρησιμοποιούνται ως βάρη το +1 και το -1 αντίστοιχα. Για παράδειγμα, ο διαχωρισμός φίλος ή εχθρός συναντάται σε ιστοσελίδες κριτικής προϊόντων τεχνολογίας ως ένδειξη εμπιστοσύνης ή μη *(14,17,18)* της γνώμης κάποιου χρήστη. Τέλος πάνω σε αυτό το σύστημα βαθμολόγησης των χρηστών βασίζεται και η δυνατότητα παρότρυνσης για σύνδεση με νέους χρήστες που συναντάται στις περισσότερες ιστοσελίδες κοινωνικής δικτύωσης *(14, 19-23).*

Από τα ανωτέρω προκύπτει άμεσα ότι οι σχέσεις μεταξύ των χρηστών είναι σημαντικές για την μετάδοση πληροφορίας διαμέσου των κοινωνικών κύκλων και κατ' επέκταση για την ομαλή λειτουργία του ιστότοπου κοινωνικής δικτύωσης. Το λογισμικό που χειρίζεται τα δεδομένα του δικτύου είναι υπεύθυνο για να δίνει προτάσεις στους χρήστες ώστε αφενός να συμβάλει στην εύκολη και γρήγορη μετάδοση πληροφορίας διαμέσων των χρηστών του και αφετέρου να προβάλει στον κάθε χρήστη ξεχωριστά πληροφορίες τέτοιες ώστε να κρατάει το ενδιαφέρον του χρήστη αμείωτο. Το βασικό στοιχείο με το οποίο αποφασίζει το κάθε λογισμικό για το ποια πληροφορία θα προβάλει στους χρήστες του κάθε χρονική στιγμή είναι η σχέσεις που έχουν συνάψει οι χρήστες μεταξύ τους.

2.3 Μηχανές προτάσεων (Recommender Systems, RSs).

Οι μηχανές προτάσεων(Recommender Systems ,RSs) είναι εργαλεία και τεχνικές λογισμικού που παρέχουν συστάσεις για αντικείμενα που θεωρείται ότι μπορεί να είναι χρήσιμα ή ενδιαφέροντα σε έναν χρήστη. *(24,25, 26, 27).* Οι συστάσεις αυτές μπορεί να είναι διαφόρων ειδών όπως για παράδειγμα προτάσεις για πιθανές αγορές, μουσικές επιλογές ή είδη διαδικτυακών ειδήσεων που μπορεί να ενδιαφέρουν τον χρήστη.

Τα κοινωνικά δίκτυα φροντίζουν να διακινούν πληροφορίες προς τους χρήστες σε μορφή προτάσεων ώστε να τους προσελκύουν και να τους κεντρίζουν το ενδιαφέρον. Για αυτόν τον σκοπό επεξεργάζονται τα δεδομένα και τις πληροφορίες που χαρακτηρίζουν τους χρήστες και τους φίλους τους (πχ. Τις δημοσιεύσεις που έχουν κάνει ή σχολιάσει, τις σελίδες που έχουν επισκεφθεί και τις αναρτήσεις που έχουν δηλώσει ότι τους ενδιαφέρουν) ώστε οι προτάσεις που θα τους κάνουν να έχουν αυξημένη πιθανότητα να είναι αρεστές και κατ' επέκταση αποδεκτές από αυτούς.

Καθώς οι προτάσεις είναι συνήθως προσωποποιημένες, διαφορετικοί χρήστες η ομάδες χρηστών λαμβάνουν διαφορετικές συστάσεις. Εκτός από τις προσωποποιημένες υπάρχουν και μη προσωποποιημένες προτάσεις. Αυτές είναι προφανώς πιο απλές στην υλοποίησή τους και τις συναντάμε συνήθως σε εφημερίδες, περιοδικά κ.α. *(24).*

Η δημιουργία μηχανών προτάσεων ξεκίνησε από μια απλή παρατήρηση: τα άτομα συχνά βασίζονται σε συστάσεις άλλων κατά την λήψη καθημερινών απλών αποφάσεων *(24, 26, 28*).

Οι μηχανές προτάσεων ξεχώρισαν σαν ανεξάρτητος ερευνητικός τομέας στα μέσα της δεκαετίας του 90 *(24, 26, 28, 29, 30).* Τα τελευταία χρόνια το ενδιαφέρον για αυτό το είδος λογισμικού έχει αυξηθεί ραγδαία και χρησιμοποιούνται κατακόρον από τις πιο δημοφιλείς διαδικτυακές εφαρμογές όπως Amazon.com, YouTube, Netflix, Yahoo, TripAdvisor and IMDb και πολλές άλλες.

Για παράδειγμα το «YouTube.com» χρησιμοποιεί συνεχώς μηχανές προτάσεων για τον κάθε χρήση κατ' αρχήν βάση της γεωγραφικής του τοποθεσίας και στην συνέχεια, βάση του ιστορικού των προηγούμενων αναζητήσεων του, καθώς «σχηματίζει» άποψη για τα ενδιαφέροντα και τις προτιμήσεις του. Αναζητά χρήστες που έχουν παρακολουθήσει παρόμοιο υλικό, χρήστες δηλαδή με παρόμοια ενδιαφέροντα, και με βάση το ιστορικό εκείνων προτείνει υλικό στον αρχικό χρήστη.

Οι χρήστες επηρεάζονται με διαφορετικό τρόπο από τους φίλους και με διαφορετικό από τους « εχθρούς τους» κάτι που υποδεικνύει ότι οι άνθρωποι χρησιμοποιούν αυτά τα δίκτυα για διαφορετικούς σκοπούς .

Τα συστήματα πρότασης (recommender systems) βασίζονται σε συλλογικό φιλτράρισμα που αυτοματοποιεί αποτελεσματικά την διαδικασία υποδηλώνοντας αντικείμενα που είναι δημοφιλή σε άτομα με παρόμοια ενδιαφέροντα, χωρίς αυτά τα άτομα να είναι κατ' ανάγκη φίλοι *(29, 31, 32).*

Σε γενικές γραμμές, πάντως, οι συστάσεις φαίνεται να είναι πιο αξιόπιστες, όταν προέρχονται από «αξιόπιστους» ανθρώπους (trusted humans) και όχι από αλγόριθμους ηλεκτρονικών υπολογιστών *(31,33).*

2.4 Ευστάθεια

Βασική προϋπόθεση της επιτυχίας μιας υπηρεσίας κοινωνικής δικτύωσης είναι το δίκτυο το οποίο δημιουργούν οι χρήστες της να είναι ευσταθές. Όπως και στα πραγματικά κοινωνικά δίκτυα, οι σχέσεις που συνάπτουν μεταξύ τους οι άνθρωποι πρέπει να είναι τέτοιες ώστε να μην οδηγείται το δίκτυο σε σχισμό.

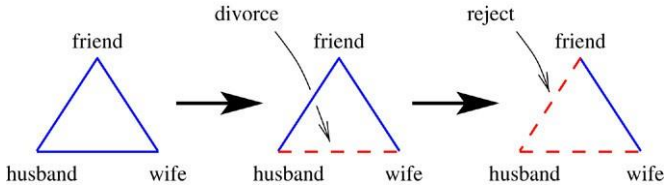
Η πρώτη αναφορά που έγινε αναφορικά με το πρόβλημα της ευστάθειας ήταν από τον Fritz Heider (1946) πάνω στην Θεωρία Δομικής Ευστάθειας (Structural Balance Theory) *(34).* Έπειτα οι Cartwright and Harary, συνδύασαν την θεωρία του Heider με τους γράφους *(35).* Επειδή οι γράφοι είναι ικανοί να περιγράψουν τις σχέσεις μεταξύ των χρηστών στα κοινωνικά δίκτυα, η θεωρία αυτή είναι πολύ χρήσιμη.

Μια προσπάθεια να ορισθούν οι κανόνες που περιγράφουν την ευστάθεια έγινε και από τους T. Antal, P.L. Krapivsky, S. Redner. Η ουσία της εξήγησης που δόθηκε συνοψίζεται στις εξής φράσεις:

- Ο φίλος του φίλου μου, όπως και ο εχθρός του εχθρού μου, είναι φίλος μου.

- Ο φίλος του εχθρού μου, όπως και ο εχθρός του φίλου μου είναι εχθρός μου.

Αυτό περιγράφει την Δομική ευστάθεια και μπορεί να γίνει κατανοητό με ένα απλό παράδειγμα από την καθημερινή ζωή, όπως δείχνει η εικόνα 1. *(36)*



Εικ. 7. Η εξέλιξη της φιλίας μιας τριάδας: ενός παντρεμένου ζευγαριού και ενός φίλου τους. Έπειτα από το διαζύγιο η τριάδα είναι σε κατάσταση αστάθειας, αλλά η ισορροπία αποκαθίσταται έπειτα από άλλη μια αλλαγή μεταξύ των σχέσεων.
Οι μπλε γραμμές αντιπροσωπεύουν φιλία και οι διακεκομμένες κόκκινες μη-φιλία. (36)

Ένα δίκτυο είναι ισορροπημένο , αν κάθε τριάδα εξισορροπείται *(34,36, 37).*Ένας φαινομενικά πιο γενικός ορισμός ενός ισορροπημένου δικτύου είναι να απαιτεί ότι κάθε κλειστό κύκλωμα, κύκλος, είναι ισορροπημένο, δηλαδη:Q`2path s` = 1. Ο Cartwright και ο Harary έδειξαν *(35, 36)* ότι ένας ορισμός ισορροπίας βασισμένος στον κύκλο είναι ισοδύναμος με έναν ορισμό που βασίζεται στην τριάδα με μια τριάδα για πλήρης γράφους (complete graphs). Αυτό το αποτέλεσμα μπορεί να αναδιατυπωθεί ως εξής: Εάν εντοπίσουμε έναν μη ισορροπημένο κύκλο οποιουδήποτε μήκους σε έναν πλήρη γράφο, πρέπει να υπάρχει και μια μη ισορροπημένη τριάδα.

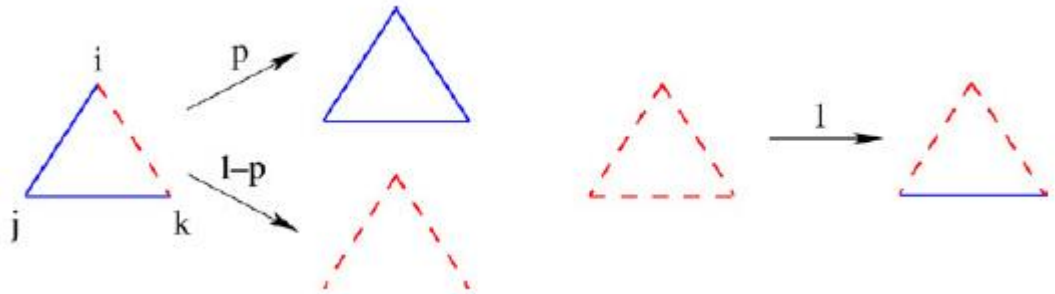
Οι Cartwright και Harary *(35, 36)* έδειξαν ότι σε έναν πλήρη γράφο οι ισορροπημένες κοινωνίες είναι εξαιρετικά απλές: είτε όλα τα άτομα είναι αμοιβαία φίλοι («ουτοπία»), ή το δίκτυο διχάζεται σε δύο αμοιβαία ανταγωνιστικές, αλλά εσωτερικά φιλικές κλίκες - μια «διπολική» κατάσταση. Για να καταλάβουμε πώς ένα δίκτυο φθάνει σε μια ισορροπημένη κατάσταση πρέπει να προχωρήσουμε πέρα από στατικές περιγραφές για να διερευνήσουμε τον τρόπο με τον οποίο μια αρχικά μη ισορροπημένη κοινωνία μετατρέπεται σε ισορροπημένη μέσω της κοινωνικής δυναμικής.

Είναι σημαντικό να αναφερθεί ότι τα κοινωνικά δίκτυα είναι δυναμικά, κατ' αναλογία με τις κοινωνικές σχέσεις στον « πραγματικό» κόσμο. Μια σχέση μπορεί να μεταπέσει από την κατηγορία της φιλίας στην αποξένωση ή στην διαγραφή και αντιστρόφως. Νέες φιλίες σχηματίζονται ενώ παλιοί φίλοι μπορεί να τείνουν να απομακρύνονται. *(36).*

Για να μπορέσουμε να κατανοήσουμε τον παραπάνω μηχανισμό αλλαγής σχέσης θα μελετήσουμε αλλαγές στη δομική μονάδα του κοινωνικού δικτύου που είναι η τριάδα. Αν μία τριάδα δεν είναι ισορροπημένη κάποια από τις σχέσεις της θα αλλάξει στο κοντινό μέλλον ώστε να μεταπέσει σε κατάσταση ισορροπίας. Η μελέτη των παραδειγμάτων αλλαγής κατάστασης σε τριάδες και κατά συνέπεια η επινόηση ενός κανόνα τον οποίο ακολουθούν οι ισορροπημένες τριάδες είναι σημαντική για να μπορούμε να κρίνουμε την ισορροπία σε επίπεδο ολόκληρου του δικτύου.

Μελετώντας παραδείγματα καταλήγουμε στον παρακάτω κανόνα για ισορροπημένες τριάδες:

Μία τριάδα σχέσεων φιλίας και εχθρότητας είναι ισορροπημένη όταν το πλήθος των θετικών σχέσεων είναι περιττό. Αν ο κανόνας αυτός δεν ισχύει σε κάποια ή κάποιες τριάδες σχέσεων είναι πολύ πιθανό να αλλάξει κάποια σχέση ώστε να ισχύσει ο παραπάνω κανόνας *(36).*



Εικ. 8 Στις μη ισορροπημένες τριάδες χρειάζεται ένα βήμα αλλαγής σχέσης ώστε να αποκατασταθεί η ισορροπία. Οι μπλε γραμμές αντιπροσωπεύουν φιλικές σχέσεις και οι κόκκινες διακεκομμένες εχθρικές (36).

Η παραπάνω θεωρία δίνει μια εξήγηση για ισορροπημένα δίκτυα η οποία ονομάζεται δομική ευστάθεια.

Η δομική ευστάθεια είναι σημαντική για τις μηχανές προτάσεων (recommendation engines).Αυτό συμβαίνει διότι ο κανόνας ισορροπίας της δομικής ευστάθειας μας δίνει μια εικόνα για τις μεγάλες ομάδες- κλάσεις χρηστών που φαίνεται να σχηματίζονται στο κοινωνικό δίκτυο. Ουσιαστικά αν αναπαραστήσουμε γραφικά ένα κοινωνικό δίκτυο θα παρατηρήσουμε ότι οι χρήστες ανήκουν σε μεγάλες ομάδες οι οποίες έχουν το χαρακτηριστικό να έχουν μέλη με αποκλειστικά φιλικές σχέσεις μεταξύ τους και που έχουν εχθρικές σχέσεις με μέλη άλλων ομάδων.(Όπως αναφέρθηκε ήδη) Μία ακόμα παρατήρηση είναι πως οι αρνητικές σχέσεις είναι πολύ λιγότερες από τις φιλικές , κάτι που κάνει τις ομάδες να είναι πιο ευδιάκριτες, και αυτό οφείλεται στην ψυχολογία των ατόμων που προτιμούν να μην δίνουν έμφαση στις εχθρικές σχέσεις αλλά να σχηματίζουν κυρίως φιλικές.

Έχοντας αυτό σαν δεδομένο, και φυσικά το γεγονός ότι, όπως προαναφέρθηκε, άτομα που συνδέονται με φιλικές σχέσεις είναι πιο πιθανό να έχουν και παρόμοια ενδιαφέροντα, οι μηχανές προτάσεων μπορούν να κρίνουν με μεγάλη ακρίβεια ποιες προτάσεις θα κάνουν και σε ποιους χρήστες.

2.5 Πρόβλεψη για τις σχέσεις μεταξύ μελών

Στα πλαίσια της καλής λειτουργίας των online κοινωνικών δικτύων είναι πολύ χρήσιμο να μπορούμε να κάνουμε πρόβλεψη για τις σχέσεις που θα μπορούσε να έχει ένα μέλος με άλλα. Φυσικά υποθέτουμε ότι το εν λόγω μέλος έχει ήδη συνάψει κάποιες σχέσεις(φιλία- εχθρότητα) τις οποίες θα χρησιμοποιήσουμε ως δεδομένο για να αποφασίσουμε για το πρόσημο της πιθανής σχέσης του με τα άλλα μέλη.

2.5.1 Βασικές αρχές πρόβλεψης

Γενικότερα για την επίλυση του προβλήματος της πρόβλεψης των σχέσεων χρησιμοποιούμε σαν δεδομένα τα βάρη των ακμών του δικτύου. Η πιο απλή περίπτωση προσημασμένου γράφου είναι αυτή στην οποία τα βάρη μπορεί να έχουν μόνο δύο διαφορετικές τιμές, + 1 ή – 1 (δηλαδή συμπάθεια ή αντιπάθεια) *(14,38,12,13)* όπως συναντάται πχ. στο Slashdot. Σε κάποιους ιστοτόπους χρησιμοποιείται ένα μεγαλύτερο σύνολο τιμών, για παράδειγμα ακεραίων από 0 έως 10. Όμως από άποψη υπολογιστικής πολυπλοκότητας είναι πολύ πιο εύκολο να υπολογίσουμε την προηγούμενη περίπτωση και να κάνουμε γενίκευση με πολύ μεγάλη ακρίβεια και για ιστοτόπους που δεν έχουν μόνο δυαδικές τιμές.

Για την πρόβλεψη των σχέσεων αυτών χρησιμοποιούνται δύο βασικοί θεωρητικοί πυρήνες. Ο πρώτος είναι η δομική ευστάθεια *(14,35,39)* των δικτύων που αναφέρθηκε παραπάνω ενώ ο δεύτερος είναι η πρόβλεψη που βασίζεται στην κοινωνική θέση (social status theory) *(12,14,40).*

Η θεωρία κοινωνικής θέσης βασίζεται σε κατευθυνόμενα δίκτυα. Στηρίζεται στην παραδοχή πως οποιαδήποτε κατευθυνόμενη κοινωνική σχέση, θετική ή αρνητική, υποδεικνύει πως ο χρήστης ο οποίος συνάπτει την σχέση παραδέχεται την υψηλότερη ή χαμηλότερη αντίστοιχα κοινωνική θέση του άλλου. Ως υψηλή ή χαμηλή κοινωνική θέση εννοείται η υψηλή ή χαμηλή, αντίστοιχα, δημοτικότητα του χρήστη στο συγκεκριμένο δίκτυο. Για παράδειγμα μια θετική βαθμολόγηση από ένα χρήστη Α προς ένα χρήστη Β υποδηλώνει ότι ο χρήστης Α αναγνωρίζει τον Β ως πιο δημοφιλή και ως εκ τούτου με υψηλότερη κοινωνική θέση *(14).*

2.5.2 Σημασία πρόβλεψης

Η σημασία της δυνατότητας πρόβλεψης των σχέσεων ανάμεσα στους χρήστες ενός κοινωνικού δικτύου και η προσπάθεια για συνεχή βελτίωση και εξέλιξη των τεχνικών που την επιτυγχάνουν θεωρείται ότι είναι πολύ σημαντική για την βιωσιμότητα, δημοτικότητα και μελλοντική εξέλιξη του.

Μέσω της πρόβλεψης των σχέσεων των χρηστών εξυπηρετούνται 2 βασικές ανάγκες του δικτύου.

i. Το να διατηρεί το ενδιαφέρων των μέλλον του αμείωτο και σε υψηλά επίπεδα
ii. Για κοινωνικά δίκτυα που εμπεριέχουν διαφήμιση προϊόντων ή υπηρεσιών ,το να πραγματοποιεί στοχευμένη και περισσότερο αποτελεσματική διαφήμιση αυξάνοντας έτσι τα έσοδα του.

Όσον αφορά το πρώτο σκέλος, το σύστημα πρόβλεψης σχέσεων κρίνει τους χρήστες οι οποίοι θεωρεί ότι είναι πιο πιθανό να μοιράζονται τις ίδιες απόψεις και ενδιαφέροντα και τους προτείνει για νέους φίλους. Επιπλέον επιλέγει συζητήσεις (από αυτές που ο χρήστης δεν έχει παρακολουθήσει μέχρι στιγμής) που συμμετέχουν φίλοι ή πιθανοί φίλοι του χρήστη και του τις προτείνει. Κατ’ αυτόν τον τρόπο ο ιστότοπος κράτα αμείωτο το ενδιαφέρον του χρήστη και διατηρεί την επισκεψημότητα site σε υψηλά επίπεδα.

Σχετικά με το δεύτερο σκέλος, το σύστημα πρόβλεψης συμβάλει είτε άμεσα είτε έμμεσα στην αύξηση των εσόδων του εν λόγω ιστότοπου που προέρχονται από διαφημίσεις.

Χρησιμοποιώντας τις προβλέψεις και αναγνωρίζοντας τους χρήστες που θεωρεί ότι μοιράζονται παρόμοια ενδιαφέροντα επιτυγχάνει το να πραγματοποιεί στοχευμένη διαφήμιση προϊόντων ή υπηρεσιών καθώς ομαδοποιεί τα μέλη με με αντίστοιχα ενδιαφέροντα και τους προβάλει διαφημίσεις παρόμοιου περιεχομένου.

Εκτός από τις δυο βασικές ανάγκες που αναφέραμε υπάρχουν άλλοι δύο τομείς που μπορεί να επωφεληθούν από την μηχανή πρόβλεψης των σχέσεων.

Υπάρχει μια κατηγορία ιστοτόπων που εμπορικά διακινούν υπηρεσίες και προϊόντα οι ίδιες (πχ. Spotify, Apple Music (iTunes)). Σε αυτή την κατηγορία ιστοτόπων, η μηχανή πρόβλεψης των σχέσεων μπορεί να δίνει προτάσεις στους χρήστες με βάση τις αγορές άλλων συμβατών χρηστών. Αρχικά συλλέγει την πληροφορία από τις αγορές όλων των χρηστών, και έπειτα κρίνει για τον κάθε χρήστη με βάση τις προτιμήσεις των φίλων του και των πιθανών φίλων του ποια προϊόντα είναι πιθανό να τον ενδιαφέρουν. Έτσι βελτιώνει τις προτάσεις που κάνει στον κάθε χρήστη και αυξάνει τα έσοδά του από τις υπηρεσίες και προϊόντα που προσφέρει.

Μια άλλη ανάγκη που εξυπηρετεί η πρόβλεψη σχέσεων χρηστών είναι ο περιορισμός του φαινομένου παραπληροφόρησης (trolling). Ένα παράδειγμα για την συγκεκριμένη ανάγκη είναι ο ιστότοπος “slashdot.com”. Ο ιστότοπος αυτός αποτελεί ένα παράδειγμα κοινωνικού δικτύου όπου συναντάται το φαινόμενο της παραπληροφόρησης σε αρκετά μεγάλο βαθμό (trolling). Υπάρχει μια ομάδα χρηστών οι οποίοι εκούσια απαντούν σε ερωτήσεις και υποστηρίζουν απόψεις με αποκλειστικό στόχο την πόλωση της συζήτησης, την προσβολή άλλων χρηστών ή την γενικότερη παραπληροφόρηση τους *(41)*. Έτσι φυσικά δεν προάγει τον ορθό διάλογο και την διακίνηση σωστών πληροφοριών που είναι και ο στόχος την υπηρεσίας. Αυτός είναι ένας από τους λόγους που το εν λόγω site έδωσε από το 2002 (εφαρμογή Slashdot Zoo) την δυνατότητα στους χρήστες του να επισημαίνουν ως εχθρούς εκτός από φίλους άλλους χρήστες. Σαν αποτέλεσμα το site πλέον έχει έναν τρόπο να μπορεί να αποφασίσει ποια σχόλια είναι εποικοδομητικά για την συζήτηση και ποια όχι, ανάλογα με το αν αυτός που τα δημοσιεύει έχει μαρκαριστεί πολλές φορές ως εχθρός (foe) ή όχι. Στην συγκεκριμένη περίπτωση λοιπόν η μελέτη των σχέσεων μεταξύ των χρηστών και κατ’ επέκταση ή πρόβλεψη τους βοηθά στην έγκαιρη αναγνώριση χρηστών που έχουν ως στόχο την παραπληροφόρηση (trolls) και στον περιορισμό και την απομόνωση τους στο δίκτυο*(41)*.

3. Αντικείμενο υπό διερεύνηση

3.1 Σχετικό ερευνητικό έργο

Με το ζήτημα της πρόβλεψης προσήμου σε προσημασμένο κοινωνικό δίκτυο ασχολήθηκαν πρόσφατα οι A. Papaoikonomou, M.Kardara, K.Tserpes, D.Varvarigou στην δημοσίευση με τίτλο Edge sign prediction in social network via frequent subgraph Discovery *(42)* η οποία χρησιμοποιήθηκε και ως θεωρητική βάση της εργασίας αυτής.

Στην συγκεκριμένη δημοσίευση αρχικά γίνεται αναπαράσταση του κοινωνικού δικτύου με έναν κατευθυνόμενο γράφο, όπου οι κόμβοι αναπαριστούν τους χρήστες και οι ακμές τις κατευθυνόμενες σχέσεις. Τα κοινωνικά δίκτυα από τα οποία έγινε εξαγωγή γράφων είναι το Slashdot, το Wikipedia και το Epinions. Επιδιώκεται μια προσομοίωση, η οποία εκτελείται για κάθε σχέση και προσπαθεί να προβλέψει το είδος της σχέσης αυτής (edge sign prediction).

Η προσομοίωση αυτή αρχικά σαρώνει όλον τον γράφο αναγνωρίζοντας συγκεκριμένα μοτίβα τριάδων ή τετράδων και εν συνεχεία συνδέει κάποια από αυτά με θετικές ακμές και κάποια άλλα με αρνητικές.

Σαν δεδομένα του προβλήματος χρησιμοποιούνται οι κοινοί γείτονες-κόμβοι και των δύο προσκείμενων κόμβων της ακμής. Στην συνέχεια λαμβάνεται ο υπο-γράφος που σχηματίζεται από την παραπάνω θεώρηση και ένα πρόγραμμα αναζήτησης ψάχνει να βρει ποιο είναι το τρίγωνο κόμβων που συναντάται συχνότερα.

Αν η συχνότητα εμφάνισης του τριγώνου αυτού είναι πάνω από έναν ουδό (theshold) τότε μπορεί να μας δώσει ασφαλές συμπέρασμα για το είδος της σχέσης. Επιπρόσθετα επιχειρείται και μία προσομοίωση χρησιμοποιώντας τετράδες γειτονικών χρηστών, αντί για τριάδες, και συμπεραίνεται ότι τα αποτελέσματα φαίνεται να έχουν ελαφρώς χαμηλότερη ακρίβεια πρόβλεψης σε σχέση με την ακρίβεια που επιτυγχάνεται όταν χρησιμοποιούνται τρίγωνα κόμβων.

Η μεθοδολογία αυτή δεν δίνει 100% ασφαλές συμπέρασμα και για αυτό χρησιμοποιείται διαφορετικός ουδός (threshold) για κάθε προσομοίωση και η ακρίβεια ποικίλει ανάλογα με αυτόν. Επίσης κάτι που πρέπει να αναφερθεί είναι ότι υπάρχει ένα σύνολο σχέσεων για τις οποίες η προσομοίωση αδυνατεί να δώσει ασφαλές συμπέρασμα καθώς βρίσκονται κάτω από τον ουδό.

3.2 Προτεινόμενη προσέγγιση

Με βάση το παραπάνω θεωρητικό μοντέλο, στην παρούσα εργασία, δημιουργήσαμε δύο προσομοιώσεις πρόβλεψης σχέσεων οπού η κάθε μία βασίζεται σε διαφορετική μέθοδο.

Οι μέθοδοι που επεξεργαστήκαμε είναι:

Α. η δομική ευστάθεια

Β. η αναγνώριση συχνών υπο-γράφων και πρόβλεψη με βάση αυτούς.

Γ. Έπειτα θα επιχειρήσουμε να συνδυάσουμε τις 2 μεθόδους και να εξετάσουμε αν κατ' αυτόν τον τρόπο μπορούμε να επιτύχουμε υψηλότερη ακρίβεια.

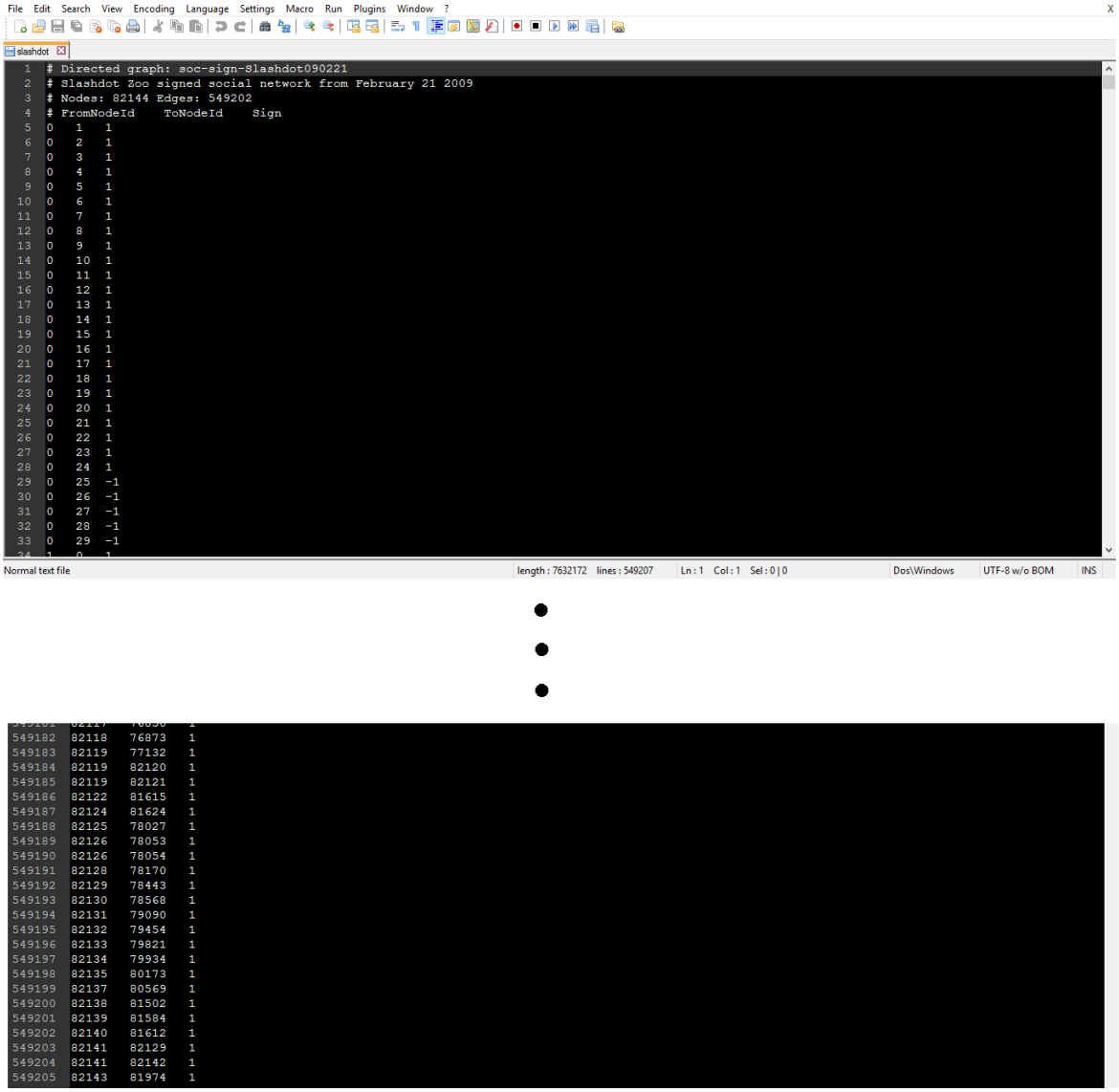
4.Μεθοδολογία- Υλικό

4.1 Μοντελοποίηση

Σε ένα online κοινωνικό δίκτυο, όπως το χρησιμοποιούν οι χρήστες, προβάλλονται τα ονόματα των χρηστών (πραγματικό όνομα ή alias) και ο καθένας από αυτούς έχει συνήθως την δυνατότητα να βλέπει συγκεντρωτικά σε μια σελίδα με ποιους από αυτούς έχει συνάψει σχέσεις και τι είδους σχέσεων ακριβώς.

Στην εργασία αυτή, όπως θα αναφερθεί εκτενέστερα στην συνέχεια, χρησιμοποιούμε το κοινωνικό δίκτυο Slashdot του οποίου οι σχέσεις είναι θετικές ή αρνητικές (εκφράζουν φιλία ή εχθρότητα). Στα πλαίσια του προγράμματος προσομοίωσης που δημιουργήσαμε θεωρούμε πως ο κάθε χρήστης έχει ένα αύξοντα ακαριαίο αριθμό που τον ορίζει και αντιστοιχεί στο πόσο πρόσφατα εγγράφηκε στο συγκεκριμένο δίκτυο (οι πιο παλαιοί χρήστες έχουν μικρότερους αριθμούς και οι πιο πρόσφατοι μεγαλύτερους). Τέλος η φιλιά αντιστοιχεί στο +1 ενώ η αντιπάθεια στο -1, χρησιμοποιούμε δηλαδή ένα δυαδικό σύστημα.

Στο αρχείο που περιέχει τα δεδομένα που θα επεξεργαστούμε κάθε σειρά αντιπροσωπεύει μια σχέση. Ο πρώτος ακαριαίος αριθμός στην σειρά αντιπροσωπεύει το όνομα του χρήστη από τον οποίο ξεκινά η σχέση, ο δεύτερος ακαριαίος αντιπροσωπεύει τον χρήστη στον οποίο καταλήγει η σχέση και τέλος ο τελευταίος ακαριαίος είναι +1 για φιλία και -1 για εχθρότητα όπως αναφέρθηκε παραπάνω.



Εικ. 9 Screenshots από το αρχείο δεδομένων του κοινωνικού δικτύου Slashdot (43)

Όπως φαίνεται και στην παραπάνω εικόνα το αρχείο δεδομένων μας περιγράφει 549.205 σχέσεις ανάμεσα σε 82.143 κόμβους.

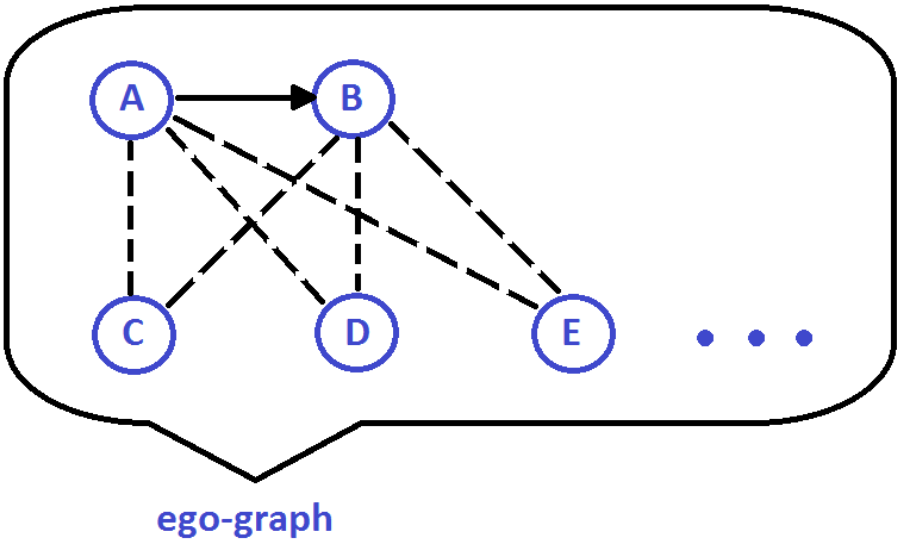
Για χάριν απλότητας στην προσομοίωση που δημιουργήσαμε για κάθε σχέση θα χρησιμοποιούμε τους εξής συμβολισμούς

i. Ο χρήστης- κόμβος από τον οποίο ξεκινάει η σχέση αναφέρεται ως κόμβος A
ii. Ο κόμβος στον οποίο καταλήγει η σχέση αναφέρεται ως κόμβος B
iii. Το πρόσημο της σχέσης περιγράφεται με +/- 1

4.2 Μεθοδολογίες

Πριν προχωρήσουμε στην μεθοδολογία που χρησιμοποιήθηκε πρέπει να ορίσουμε την έννοια του ego-graph.

Δεδομένου ενός γράφου G μίας ακμής e η οποία ανήκει στο σύνολο ακμών του γράφου (E) ο ego-graph είναι ο υπο-γράφος του G του οποίου το σύνολο κόμβων αποτελείται από τους προσκείμενους κόμβους της ακμής e και των κοινών τους γειτόνων. Η ακμή e ονομάζεται γεννήτρια ακμή του ego-graph (42).



Εικ. 10 Ως ego-graph ορίζεται ο υπο-γράφος που περιλαμβάνει μόνο τους κοινούς γείτονες του A και του B

4.2.1 Πρόβλεψη βασισμένη στην δομική ευστάθεια

Η πρώτη προσομοίωση που επιχειρήσαμε βασίζεται στην θεωρία της δομικής ευστάθειας. Το πρόγραμμα προσομοίωσης μας λοιπόν ανατρέχει όλο το δίκτυο (στην συγκεκριμένη περίπτωση το Slashdot) και αποθηκεύει τα δεδομένα σε αντικείμενα τα οποία μπορούμε να επεξεργαστούμε στην συνέχεια από τις μεταβλητές αντικειμένων για κάθε σχέση που μας δίνεται κάνουμε τα εξής βήματα:

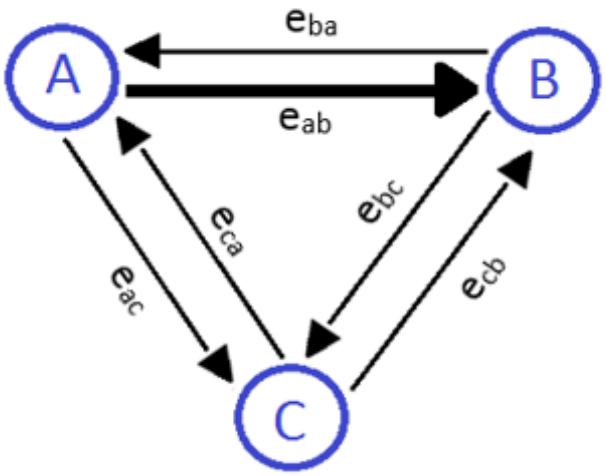
i. Αποθηκεύουμε την σχέση την οποία θέλουμε να προβλέψουμε
ii. Τρέχουμε δια μήκος όλων των γειτόνων του κόμβου A και εξετάζουμε πόσοι από αυτούς συνορεύουν ταυτόχρονα και με τον κόμβο B
iii. Συλλέγουμε όλους τους κοινούς γείτονες σε μια καινούρια δομή δεδομένων, λίστα.
iv. Ορίζουμε δύο μετρητές αρχικοποιημένους στο «0» οι οποίοι θα συγκρατούν τις προβλέψεις ο ένας για + και ο άλλος για – αντίστοιχα.
v. Για κάθε έναν από τους κοινούς γείτονες εξετάζουμε το τρίγωνο σχέσεων που σχηματίζεται και σύμφωνα με τους κανόνες που διέπουν την δομική ευστάθεια (ο φίλος

του φίλου μου είναι φίλος μου και ο φίλος του εχθρού μου είναι εχθρός μου και ο εχθρός του εχθρού μου είναι φίλος μου, και τα παράγωγα αυτών) προσπαθούμε να αποφανθούμε για το πρόσημο της σχέσης που ξεκινάει από τον κόμβο A και καταλήγει στον κόμβο B.

vi. Όταν ολοκληρώσουμε την διαδικασία v. για όλους τους κοινούς γειτονικούς κόμβους συγκρίνουμε τους μετρητές και ανάλογα με το ποιος μετρητής είναι μεγαλύτερος δίνουμε την αντίστοιχη πρόβλεψη.

vii. Σε περιπτώσεις που οι σχέσεις δίνουν ego-graph που είναι κενός (περιέχει μόνο τους κόμβους A και B) δεν δίνουμε πρόβλεψη. Επίσης σε περίπτωση που η πρόβλεψη για +1 είναι ίση με την πρόβλεψη για -1 πάλι δεν μπορούμε να δώσουμε ασφαλή πρόβλεψη. Τέλος επαναλαμβάνουμε το πείραμα i- vi ζητώντας οι προβλέψεις για + και – να έχουν τουλάχιστον διαφορά n, και παρατηρούμε για ποιο n η πρόβλεψη είναι πιο ακριβής.(στα δικά μας πειράματα το βέλτιστο n είναι 4)

Αναφορικά με τους κοινούς γείτονες των κόμβων A και B το τρίγωνο σχέσεων το οποίο επεξεργαζόμαστε σύμφωνα με την θεωρία της δομικής ευστάθειας φαίνεται στην παρακάτω εικόνα. Κάποιες από τις σχέσεις (ακμές γράφου) ενδέχεται να λείπουν διότι δεν μιλάμε για πλήρη γράφο, όμως σίγουρα υπάρχει μια από τις e_{ac} και e_{ca} και αντίστοιχα μια εκ των e_{bc} και e_{cb}.



Εικ. 11 Ενδεικτικός γράφος που εξετάζεται. Ο κόμβος C είναι ένας κοινός γειτονικός κόμβος των A και B.

4.2.2 Πρόβλεψη βασισμένη στην αναγνώριση συχνών υπο-γράφων

Στην δεύτερη προσομοίωση η οποία βασίζεται στην αναγνώριση συχνών υπο-γράφων και χρησιμοποιείται και στην δημοσίευση edge sign prediction in social networks via frequent subgraph discovery των A. Papaoikonomou, M.Kardara, K.Tserpes, D.Varvarigou *(42)* επιχειρήσαμε να κάνουμε προσομοίωση ακολουθώντας τα εξής βήματα:

i. Διατρέχουμε όλες τις σχέσεις οι οποίες είναι αποθηκευμένες στο αρχείο εισόδου και αποθηκεύουμε τις σχέσεις σε μεταβλητές
ii. Για κάθε σχέση δημιουργούμε τον αντίστοιχο ego-graph
iii. Για κάθε ego-graph του προηγούμενου βήματος ανατρέχουμε όλους τους κόμβους που τον αποτελούν. Για κάθε έναν από αυτούς σχηματίζεται ένα τρίγωνο σχέσης με τους κόμβους A και B.
iv. Αναγνωρίζουμε τα ποιο συχνά εμφανιζόμενα μοτίβα τριαδικών σχέσεων τα οποία έπειτα συνδέουμε με σχέση θετική ή αρνητική ανάλογα με το πρόσημο της ακμής που παράγει τον κάθε ego-graph.
v. Όταν τελειώσει αυτή η διαδικασία για όλους τους ego-graphs έχουμε δημιουργήσει ένα σύνολο συχνά εμφανιζόμενων τριάδων σχέσεων οι οποίες σχετίζονται άμεσα με μια θετική ή αρνητική σχέση κατά περίπτωση.
vi. Με βάση αυτό το δεδομένο μπορούμε να ανατρέξουμε σε οποιαδήποτε σχέση και να την προβλέψουμε ως εξής:
Σχηματίζοντας τον υπο-γράφο που αποτελείται από τους κοινούς γείτονες των κόμβων που είναι προσκείμενοι στην εν λόγω σχέση και εξετάζοντας αν κάποιες από τις πιο συχνές τριάδες ξεχωρίζει σε αυτήν την περίπτωση.
vii. Τέλος και σε αυτήν την τεχνική υπάρχουν περιπτώσεις στις οποίες, παρόμοια με την μέθοδο 4.2.1 δεν μπορούμε να δώσουμε ασφαλή πρόβλεψη.

4.2.3 Σύγκριση των δύο μεθόδων

Μετά την περάτωση των δύο προσομοιώσεων θα συγκρίνουμε τα αποτελέσματα εξάγοντας τα πλεονεκτήματα και τα μειονεκτήματα της κάθε μεθόδου. Θα προσπαθήσουμε επιπλέον να βελτιώσουμε τον βαθμό επιτυχίας της πρόβλεψης μας συνδυάζοντας τα πλεονεκτήματα και των δυο αυτών προσεγγίσεων.

Παρακάτω θα αναφερθούμε εκτενώς στην αποδοτικότητα της κάθε μεθόδου εξετάζοντας την ακρίβεια τους και συνυπολογίζοντας την πολυπλοκότητα της κάθε μιας.

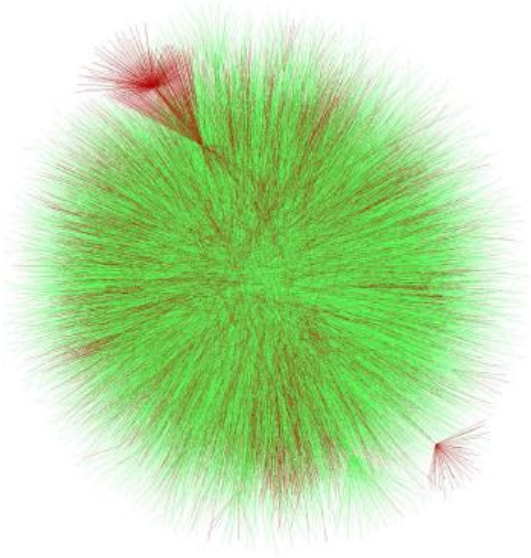
4.3 Εργαλεία υλοποίησης

Εικ. 12 Λογότυπο της ιστοσελίδας Slashdot.com

4.3.1 Slashdot

Το Slashdot.com είναι μια ιστοσελίδα τεχνολογικών ειδήσεων που ιδρύθηκε το 1997 από τους Rob Malda και Jeff Bates. Εκδίδει ειδήσεις, γραμμένες από τους συντάκτες ή ειδήσεις που υποβάλλονται από τους χρήστες του ενώ παράλληλα επιτρέπει στους χρήστες να τις σχολιάζουν. Το 2002, το Slashdot προσέθεσε το χαρακτηριστικό Slashdot Zoo, πρόκειται για μια νέα υπηρεσία του site η οποία επιτρέπει στους χρήστες να επισημαίνουν άλλους χρήστες ως εχθρούς έκτος από φίλους(friend or foe). Αποτελεί ένα πρώιμο παράδειγμα μιας OnLine πλατφόρμας κοινωνικής δικτύωσης η οποία σε αντίθεση με τα ποιο δημοφιλή social networks επιτρέπει στα μέλη της να αξιολογήσουν άλλα μέλη αρνητικά.

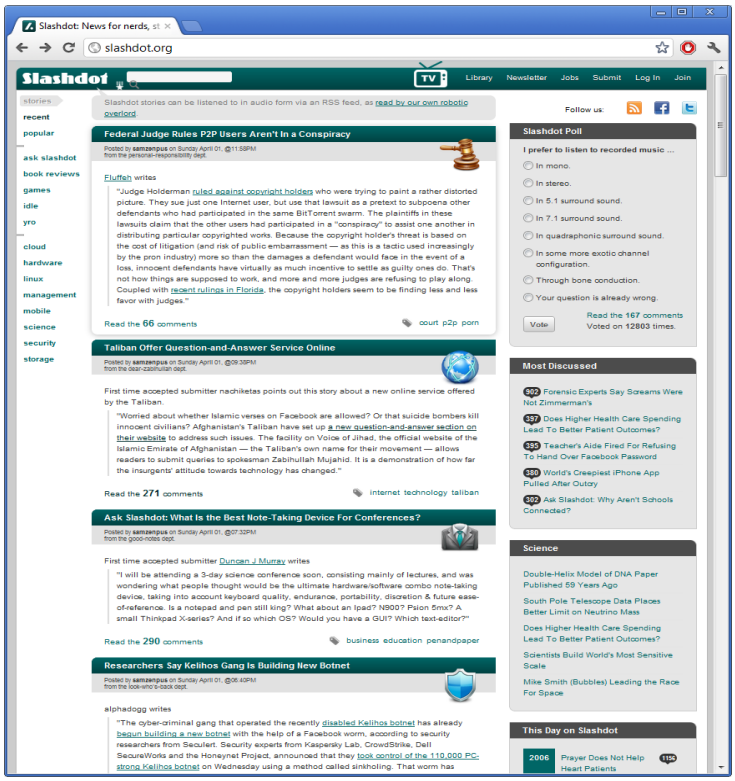
Γενικά είναι γνωστό για το ότι περιέχει στο δίκτυο του κάποια μέλη με μεγάλη δημοτικότητα και κάποια άλλα με πολύ χαμηλή ενώ παράλληλα, όπως έχει αναφερθεί ήδη εμφανίζει το φαινόμενο trolling σε αρκετά μεγάλη κλίμακα. Ως εκ του του, η ύπαρξη της υπηρεσίας Slashdot Zoo είναι πολύ χρήσιμη στον συγκεκριμένο ιστότοπο καθώς συμβάλει στην αναγνώριση των «επιβλαβών χρηστών « (trolls) και στον περιορισμό τους *(41).*



Εικ. 13

Χάρτης των σχέσεων των χρηστών δείγματος στο Slashdot Zoo. Το δίκτυο περιέχει 77.985 χρήστες και 510.157 σχέσεις. Με πράσινο απεικονίζονται οι σχέσεις «φιλίας» ενώ με κόκκινο οι «εχθρικές". Η κατεύθυνση των σχέσεων δεν εμφανίζεται στον χάρτη αυτόν. Το γράφημα έχει ως κέντρο τον χρήστη CmdrTaco ο οποίος είναι ο ιδρυτής του site (41)

Αυτό γίνεται φανερό και στην ανωτέρω εικόνα όπου φαίνεται ξεκάθαρα η έντονη θετική πόλωση γύρω από ορισμένους χρήστες καθώς και η έντονη αρνητική πόλωση γύρω από άλλους χρήστες.



Εικ. 14 Screenshot από το δημοφιλές site slashdot.com

Το λογισμικό που χρησιμοποιεί ο ιστότοπος Slashdot.com ονομάζεται Slash (http://www.slashcode.com/) και σήμερα χρησιμοποιείται επίσης σε άλλες ιστοσελίδες όπως Barrapunto, Slashdot Ιαπωνία, και τη χρήση Perl οι οποίες επίσης χρησιμοποιούν το χαρακτηριστικό Zoo. *(41)*

4.3.2 Java

Η γλώσσα προγραμματισμού που χρησιμοποιήθηκε και για τα δύο προγράμματα προσομοίωσης είναι η Java.

Η Java δημιουργήθηκε από τους James Gosling, Patrick Naughton, Chris Warth, Ed Frank και Mike Sheridan στην εταιρία sun Micro systems το 1991.Ηαρχική της ονομασία ήταν oak μετονομάστηκε σε Java το 1995. Μετά την μετονομασία της εκδόθηκε με το κωδικό όνομα Java 1.0. Παρότι η Java έχει γίνει κυρίαρχή στο internet σήμερα δεν ήταν αυτό η αιτία της γέννησής της. Το βασικό κίνητρο για την δημιουργία της ήταν η ανάγκη για μια γλώσσα που δεν θα εξαρτιόταν από καμία συγκεκριμένη πλατφόρμα και θα μπορούσε να χρησιμοποιηθεί για την δημιουργία λογισμικού που θα ενσωματωνόταν σε διάφορες συσκευές του εμπορίου όπως τοστιέρες, φούρνους μικροκυμάτων και

Εικ. 15 Λογότυπο γλώσσας προγραμματισμού Java

τηλεχειριστήρια *(44).*

Δεν είναι υπερβολή να πούμε ότι η αρχική έκδοση της Java έφερε επανάσταση στον προγραμματισμό. Αυτή η επανάσταση μετέτρεψε ριζικά το διαδίκτυο σε υψηλά διαλογικό περιβάλλον. Η Java είναι μια γλώσσα αντικειμενοστραφούς προγραμματισμού και πολλά από τα χαρακτηριστικά που την απαρτίζουν έχουν γίνει πολύ δημοφιλή και υιοθετούνται από γλώσσες προγραμματισμού που σχηματίζονται σήμερα. Η τρέχουσα έκδοση της Java σήμερα είναι η Java 8. Αυτή την στιγμή οι συσκευές οι οποίες είναι ικανές να τρέξουν εφαρμογές γραμμένες σε Java είναι πάνω από ένα δισεκατομμύριο , σύμφωνα με το site της sun microsystems. Το γεγονός ότι μπορεί να εγκαθίσταται σε πολλές συσκευές πολύ εύκολα συνέβαλε στο να είναι η πιο δημοφιλής γλώσσα προγραμματισμού σήμερα *(44).*

Ενδεικτικά αναφέρουμε παρακάτω τα κύρια χαρακτηριστικά της γλώσσας java τα οποία την ξεχωρίζουν από τις υπόλοιπες και ήταν χρήσιμα στην υλοποίηση των δύο προγραμμάτων προσομοίωσης αυτής της εργασίας *(44).*

A. Εικονική μηχανή της Java (JVM)

Η Java εκτελείται διαφορετικά σε σύγκριση με τις άλλες γλώσσες προγραμματισμού. Μια κλασσική γλώσσα απαιτεί έναν compiler ο οποίος εγκαθίσταται σε ένα υπολογιστικό σύστημα και είναι υπεύθυνος για να τρέξει τα προγράμματα. Όμως επειδή υπάρχουν πολλά διαφορετικά είδη υπολογιστικών συστημάτων για μία κλασσική γλώσσα προγραμματισμού θα έπρεπε να γραφτεί σε κώδικα διαφορετικός compiler για κάθε διαφορετικό επεξεργαστή που κυκλοφορεί. Το πρόβλημα αυτό η Java το λύνει με την εικονική μηχανή. Η εικονική μηχανή είναι ένα γενικό πρόγραμμα το οποίο εγκαθίσταται σε όλα τα υπολογιστικά συστήματα και τα προγράμματαJava είναι γραμμένα με τον ίδιο τρόπο για όλους τους τύπους επεξεργαστών. Ενδεικτικά αναφέρουμε ότι ο εκτελέσιμος κώδικας των προγραμμάτων Java ονομάζεται Byte (*44).*

B. Ασφάλεια.

Εξαιτίας της ιδιαιτερότητας της Java που απαιτεί εικονική μηχανή καθίσταται εξαιρετικά ασφαλής. Ο κύριος λόγος είναι ότι τα προγράμματα της (κώδικας Byte) μεταγλωττίζονται μόνο όταν χρειάζονται κατά την εκτέλεση και κατά κομμάτια. Ο υπόλοιπος κώδικας κάθε φορά απλώς ερμηνεύεται. Είναι αξιοσημείωτο πως αυτή η προέγκριση προσφέρει και μι σημαντική αύξηση της απόδοσης. Ένας επιπλέον παράγων που συμβάλει στην ασφάλεια είναι και η αντικειμενοστραεφεια της γλώσσας που θα αναφερθεί παρακάτω *(44).*

C. Αντικειμενοστραφής προγραμματισμός.

Ο αντικειμενοστραφής προγραμματισμός είναι ένας ισχυρός τρόπος προσέγγισης του προγραμματισμού. Με πολύ απλά λόγια στην Java όλα είναι αντικείμενα. Το κάθε αντικείμενο έχει κάποια χαρακτηριστικά (μεταβλητές) και έχει κάποιες λειτουργίες (μέθοδοι). Αυτό

σημαίνει ότι τα ίδια τα αντικείμενα γνωρίζουν τι λειτουργίες πρέπει να επιτελέσουν. Ο αντικειμενοστραφής προγραμματισμός συνεπάγεται με τρία κύρια στοιχεία τα οποία θα αναλύσουμε παρακάτω (ενθυλάκωση, πολυμορφισμός, κληρονομικότητα) *(44).*

D. Ενθυλάκωση

Ενθυλάκωση είναι ένας μηχανισμός προγραμματισμού που συνδέει τον κώδικα και τα δεδομένα που χειρίζεται και έτσι διατηρείται ασφαλής από εξωτερικές μεταβολές και από κακή χρήση. Σε μια αντικειμενοστραφή γλώσσα ο κώδικας και τα δεδομένα συνδέονται έτσι ώστε να δημιουργείται ένα μαύρο κουτί που περιέχει τον εαυτό του. Μέσα στο κουτί βρίσκονται όλα τα απαραίτητα δεδομένα και ο απαραίτητος κώδικας. Όταν συνδέονται με αυτόν τον τρόπο, κώδικας και δεδομένα, δημιουργείται το αντικείμενο. Η βασική μονάδα ενθυλάκωσης της Java είναι η κλάση. Μια κλάση ορίζει την μορφή ενός αντικειμένου τα αντικείμενα είναι στιγμιότυπα μιας κλάσης. Με αυτόν τον τρόπο οι κλάσεις είναι σαν τα καλούπια για τα αντικείμενα που δημιουργούν *(44).*

E. Πολυμορφισμός

Πολυμορφισμός είναι το στοιχείο που επιτρέπει σε μια διεπαφή να προσπελάσει μια γενική κλάση ενεργειών. Ένα απλό παράδειγμα πολυμορφισμού μπορούμε να συναντήσουμε στο τιμόνι ενός αυτοκινήτου. Το τιμόνι, κατ' αναλογία η διεπαφή, είναι ίδιο σε πολλά διαφορετικά αυτοκίνητα (αυτόματης και χειροκίνητης πηδαλιούχησης) επομένως αφού μάθει κάποιος να χειρίζεται το τιμόνι μπορεί να χειριστεί το τιμόνι σε οποιοδήποτε αυτοκίνητο. Κατ' αναλογία ο πολυμορφισμός επιτρέπει μια γενικότερη γνώση μεθόδων ανεξάρτητα από το είδος των δεδομένων *(44).*

F. Κληρονομικότητα.

Κληρονομικότητα είναι η διαδικασία με την οποία ένα αντικείμενο μπορεί να αποκτήσει τις ιδιότητες άλλου αντικειμένου. Αυτό είναι σημαντικό επειδή υποστηρίζει την έννοια της ιεραρχικής κατηγοριοποίησης. Χρησιμοποιώντας την κληρονομικότητα ένα αντικείμενο πρέπει να ορίσει μόνο εκείνα τα χαρακτηριστικά που το καθιστούν μοναδικό μέσα στην κλάση. Μπορεί να κληρονομήσει τις γενικές ιδιότητες από τους «γονείς» του. Συμπερασματικά ένα αντικείμενο μπορεί να είναι στιγμιότυπο μιας πιο γενικής ομάδας και παράλληλα να έχει και χαρακτηριστικά που να το κάνουν μοναδικό (*44).*

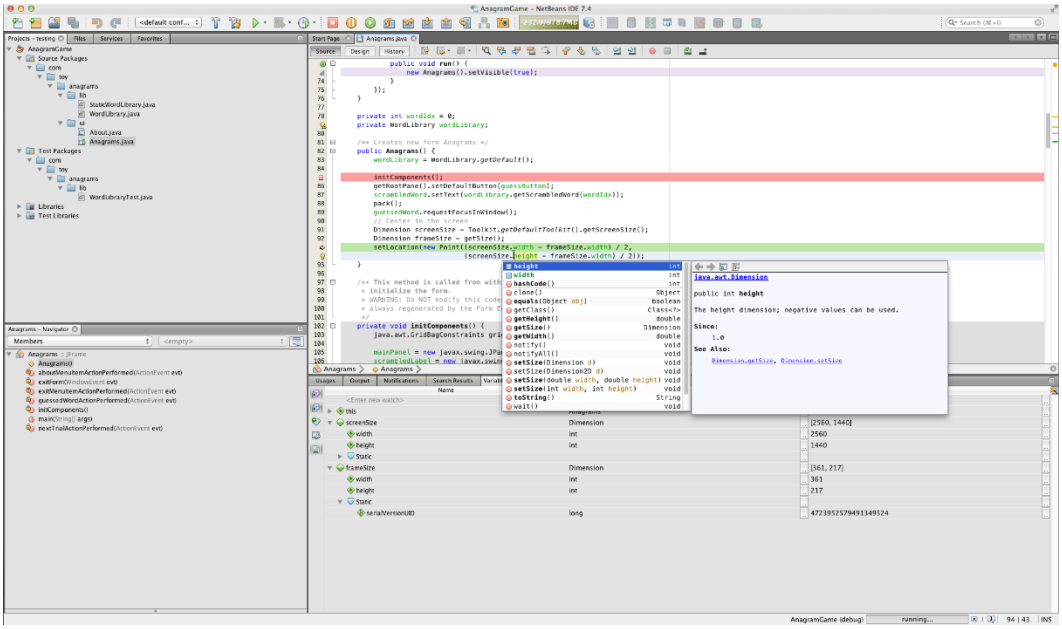
G. Πλούσιες βιβλιοθήκες.

Οι βιβλιοθήκες της Java είναι πολύ πλούσιες οργανώνονται σε πακέτα και εξαιτίας της δημοτικότητας της γλώσσας παρουσιάζεται υποστήριξη για όλων των ειδών τα προγράμματα *(44)*.

Εικ. 16 Λογότυπο εφαρμογής Netbeans

4.3.3 Netbeans

Το NetBeans είναι μια εφαρμογή ανάπτυξης λογισμικού την οποία χρησιμοποιήσαμε για να γράψουμε τις προσομοιώσεις. Η εφαρμογή αυτή είναι γραμμένη σε Java, υποστηρίζει εκτός από Java και άλλες δημοφιλείς γλώσσες όπως PHP, C/C++ και HTML5. Το NetBeans δημιουργήθηκε αρχικά το 1996, αργότερα αγοράστηκε από την sun microsystems και το 2010 εξαγοράστηκε από την Oracle (μαζί με την Sun). Το NetBeans προσφέρει πολλές ευκολίες προγραμματισμού , ειδικά στην JAVA,όπως για παράδειγμα το project matisse το οποίο επιτρέπει την δημιουργία γραφικών εφαρμογών Javaχωρίς την χρήση κώδικα *(45, 46, 47)*.



Εικ. 17 Screenshot του γραφικού περιβάλλοντος προγραμματισμού Netbeans IDE

Ένα από τα κύρια χαρακτηριστικά που καθιστούν το περιβάλλον προγραμματισμού του NetBeans δημοφιλές είναι η μεταφερσιμότητα των projects και των αρχείων. Υποστηρίζει επιπλέον την λειτουργία «drag and drop». Όλα αυτά το καθιστούν ιδιαίτερα εύχρηστο κατά το γράψιμο κώδικα, καθώς ο προγραμματιστής δεν χρειάζεται να πληκτρολογήσει μεγάλα κομμάτια κώδικα.

Πιο αναλυτικά αποτελείται από:

1. Επεξεργαστή πυγαίου κώδικα (Source code editor)

2. Μεταγλωττιστή (Compiler) ή Μεταφραστή (interpreter)

3. Αποσφαλματωτή (debugger)

Όλα αυτά συνδυάζονται σε ένα εύχρηστο γραφικό περιβάλλον το οποίο δίνει πολλές άλλες δυνατότητες όπως επιπρόσθετες εφαρμογές (add ons) *(45, 46, 47).*

5.Δεδομένα που συλλέχθηκαν

5.1 Μέθοδος βασισμένη στην δομική ευστάθεια

Σύμφωνα με ότι αναφέρθηκε στην ενότητα 4.2.1 εκτελέσαμε την προσομοίωση και από το δείγμα 82.143 κόμβων συνάπτουν μεταξύ τους 549.202 σχέσεις. Από τις σχέσεις αυτές το 40,9321% δεν μπορεί να προκύψει ασφαλής πρόβλεψη. Το υπόλοιπο 59,0238% η πρόβλεψη έχει ακρίβεια 83,19%.

Παρατηρούμε ότι αυτή η μέθοδος έχει το μειονέκτημα ότι δεν είναι πολύ ακριβής. Έχει χαμηλότερο ποσοστό ακρίβειας από την επόμενη μέθοδο και επιπλέον το δείγμα στο οποίο δεν μπορεί να δώσει ασφαλή πρόβλεψη είναι επίσης μεγαλύτερο συγκριτικά με εκείνο της επόμενης μεθόδου.

Στα πλεονεκτήματά της συγκαταλέγεται η απλότητα υλοποίησης (χαμηλή υπολογιστική πολυπλοκότητα).

5.2 Μέθοδος βασισμένη στη συχνή εμφάνιση υπο-γράφων

Στην μέθοδο την οποία περιγράψαμε στην ενότητα 4.2.2 από το ίδιο δείγμα το 65,0477% δίνει ασφαλή πρόβλεψη με επιτυχία 95,9037%.

Παρατηρούμε ότι η μέθοδος αυτή είναι εμφανώς πιο ακριβής από την προηγούμενη όμως όπως γίνεται φανερό και από την διαδικασία υλοποίησης της (4.2.2) έχει πιο πολύπλοκη υλοποίηση και η προσομοίωση της καταναλώνει περισσότερους υπολογιστικούς πόρους. Αυτό είναι επόμενο διότι όπως έγινε φανερό από την διαδικασία της μεγάλα κομμάτια του αρχείου εισόδου διαβάζονται αρκετές φορές. Παρ όλα αυτά αδυνατεί να δώσει ασφαλή προσέγγιση για όλο το σύνολο του δείγματος. Αυτό συμβαίνει διότι κάποιες από τις γεννήτριες ακμές δίνουν κενούς ή πολύ μικρούς ego-graphs.

5.3 Πρόταση βελτίωσης

Σε αυτό το σημείο, με σκοπό την βελτίωση των αποτελεσμάτων προτείνουμε το εξής: Να γίνει χρήση και των δύο προαναφερθέντων μεθόδων συνδυαστικά (hybrid) ώστε να προσπαθήσουμε κατ' αυτόν τον τρόπο να βελτιώσουμε την συνολική απόδοση του συστήματος πρόβλεψης πρόσημου (edge sign prediction)

Δημιουργούμε μια καινούρια προσομοίωση στην οποία εκτελούμε την πρόβλεψη με την βοήθεια της μεθόδου frequent sub-graph discovery (4.2.2) και έπειτα λαμβάνουμε το τμήμα του δείγματος στο οποία αυτή η μέθοδος δεν μπορεί να δώσει ασφαλή προσέγγιση (δηλαδή το 34,9523% του συνόλου) και σ' αυτό εφαρμόζουμε την μέθοδο η οποία βασίζεται στην δομική ευστάθεια (4.2.1). (Πείραμα Γ)

Πριν υλοποιήσουμε αυτήν την καινούρια προσεγγιστική μέθοδο εφαρμόζουμε δύο παραλλαγές που θα μας βοηθήσουν να συμπεράνουμε αν η πρόταση μας είναι πραγματικά πιο αποδοτική.

Για να το επιτύχουμε αυτό εφαρμόσαμε τα ακόλουθα.

- Επαναλάβαμε το ίδιο πείραμα με την διαφορά ότι το δείγμα που η μέθοδος 4.2.2 αδυνατεί να προβλέψει (34,9523%) το προσεγγίσαμε με μια τυχαία συνάρτηση (random). (Πείραμα Α)

- Ομοίως επαναλάβαμε το ίδιο πείραμα θεωρώντας ότι όλες οι προβλέψεις του υπολειπόμενου ποσοστού είναι θετικές, καθώς από τα δεδομένα του συγκεκριμένου κοινωνικού δικτύου (slashdot.com) υπολογίσαμε ότι το 77,4% των σχέσεων είναι θετικές. (Πείραμα Β)

Για να καταλήξουμε στο συμπέρασμα ότι η μέθοδος που προτείνουμε είναι αποδοτικότερη θα πρέπει το ποσοστό επιτυχίας της να είναι καλύτερο από το αντίστοιχο ποσοστό που θα προκύψει από την χρήση των δύο παραπάνω μεθόδων (1 και 2).

Πείραμα Α

Το πρώτο πείραμα συνδυάζει την μέθοδο της πρόβλεψης βασισμένη στην αναγνώριση συχνών υπο-γράφων (frequent sub-graph discovery) με τυχαία πρόβλεψη του υπολειπόμενου δείγματος. Από την υλοποίηση αυτής την μεθόδου προκύπτει ακρίβεια πρόβλεψης 78,3583% του συνόλου του δείγματος. Συγκεκριμένα για το τμήμα που η μέθοδος των υπογράφων μπορεί να δώσει ασφαλή πρόβλεψη (65,0477%) δίνει 95,9037% σωστή πρόβλεψη, όπως αναφέρθηκε και παραπάνω ενώ για το τμήμα που δεν μπορεί να προβλέψει με ασφάλεια η μέθοδος αυτή η τυχαία πρόβλεψη δίνει 55,8369% σωστή πρόβλεψη. Έτι συνολικά σε ολόκληρο το δείγμα προκύπτει ακρίβεια της τάξης των 78,3583%.

Πείραμα Β

Το δεύτερο πείραμα συνδυάζει την μέθοδο πρόβλεψης των υπο-γράφων με μόνιμη θετική πρόβλεψη (για το μέρος του δείγματος που η μέθοδος των υπογράφων δεν μπορεί να δώσει ασφαλές αποτέλεσμα). Από την εφαρμογή αυτής της προσομοίωσης προέκυψε ότι το ποσοστό επιτυχίας στο υπολειπόμενο δείγμα είναι 83,067% και συνολικά η ακρίβεια όλης την μεθόδου είναι 86,1491%.

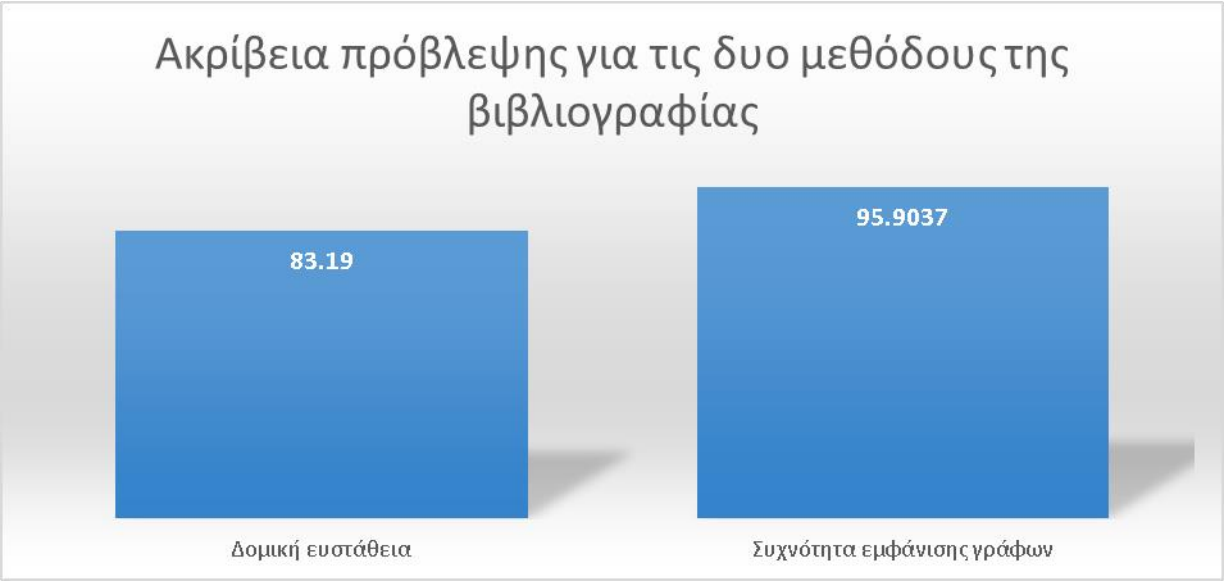
Σε αυτό το σημείο αξίζει να αναφέρουμε ότι η μέθοδος που βασίζεται στην συχνότητα υπογράφων είναι ντετερμινιστική οπότε σε όλα τα πειράματα μας έχει το ίδιο ακριβώς ποσοστό επιτυχίας.

Πείραμα Γ

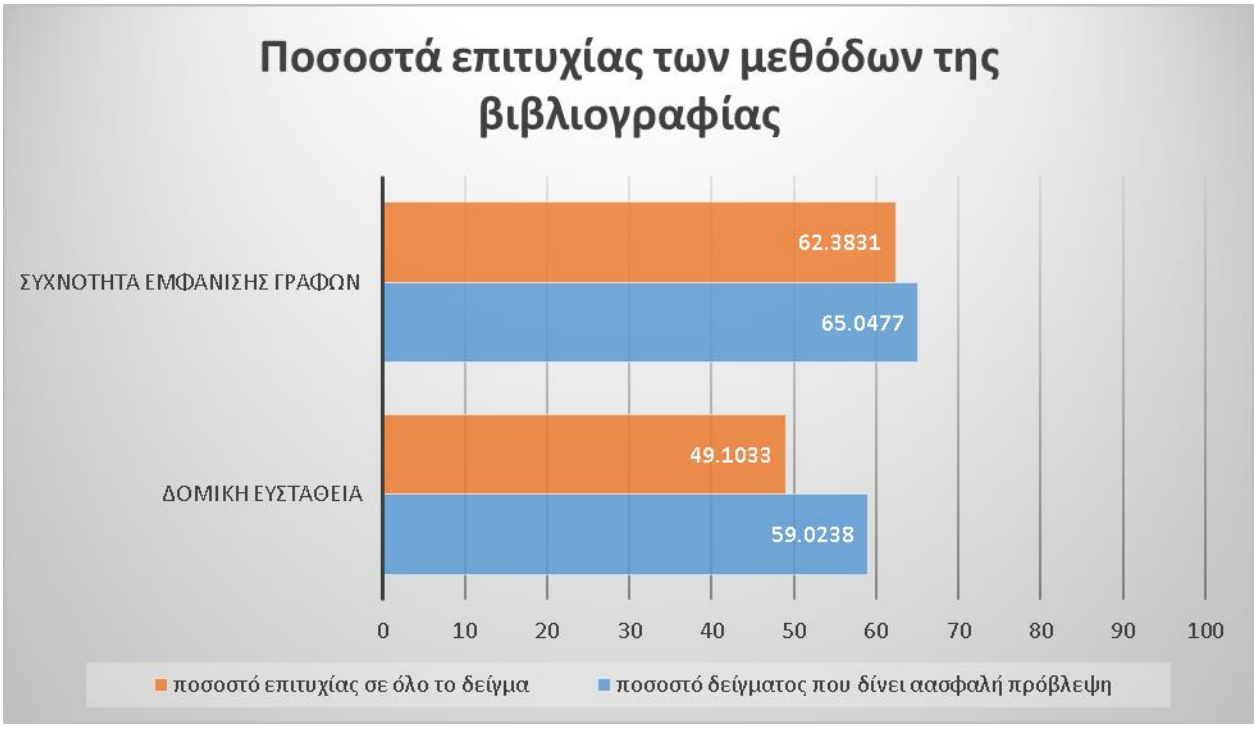
Στο τελευταίο πείραμα, όπως είπαμε, συνδυάζουμε τις δυο αρχικές μεθόδους που αναφέρουμε στην μεθοδολογία (4.2.1 και 4.2.2). Η ακρίβεια που προκύπτει είναι 79,6448%. Η ακρίβεια στο υπολειπόμενο δείγμα (μετρημένη με την μέθοδο της δομικής ευστάθειας) είναι 61,38%.

Τα αποτελέσματα όλων των πειραμάτων φαίνονται συγκεντρωτικά παρακάτω:

Ποσοστά (%)	Μέθοδοι βιβλιογραφίας		Συνδυασμοί μεθόδων (hybrid system)		
	Δομική ευστάθεια	Συχνότητα εμφάνισης γράφων	Πείραμα Α (Συχνότητα εμφάνισης γράφων και random)	Πείραμα Β (Συχνότητα εμφάνισης γράφων και θετική πρόβλεψη)	Πείραμα Γ (Συχνότητα εμφάνισης γράφων και δομική ευστάθεια)
Ποσοστό συνδέσεων στις οποίες μπορούμε να πάρουμε ασφαλή απόφαση.	59.0238	65.0477	**		
Ποσοστό επιτυχίας στο δείγμα με την ασφαλή απόφαση.	**83.19**	**95.9037**			
Συνολικό ποσοστό επιτυχίας.	**49.1033***	**62.3831***	**78.8583**	**86.1491**	**79.6448**
	* υποθέτουμε ότι το κομμάτι στο οποίο δεν γίνεται ασφαλή πρόβλεψη αποτυγχάνει.		** θεωρούμε ότι κάνουμε ασφαλή πρόβλεψη σε όλο το δείγμα, οπότε εξετάζουμε τη συνολική ακρίβεια της μεθόδου στο σύνολο του δείγματος.		



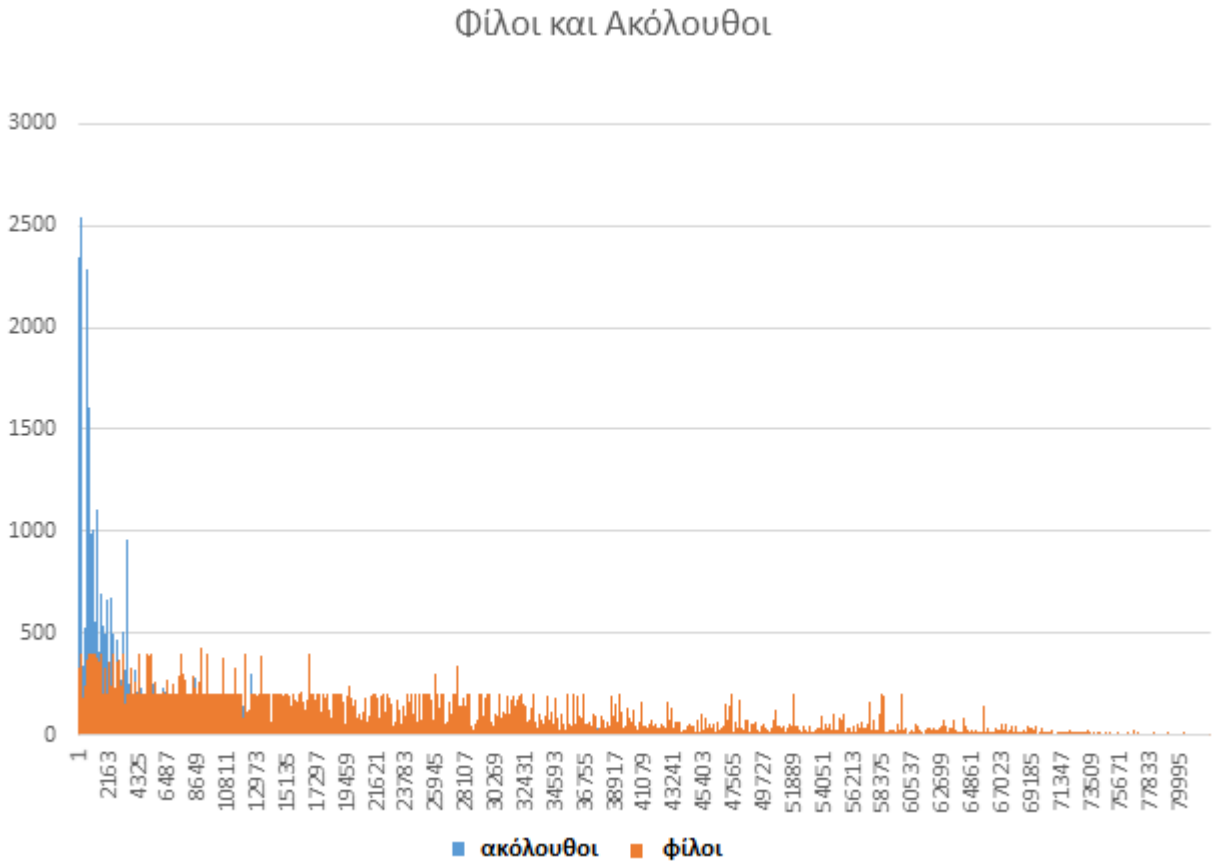
Εικ. 18 Ποσοστά επιτυχίας μεθόδων βιβλιογραφίας στο δείγμα που δίνει ασφαλή πρόβλεψη



Εικ. 19 Ποσοστά επιτυχίας των μεθόδων της βιβλιογραφίας

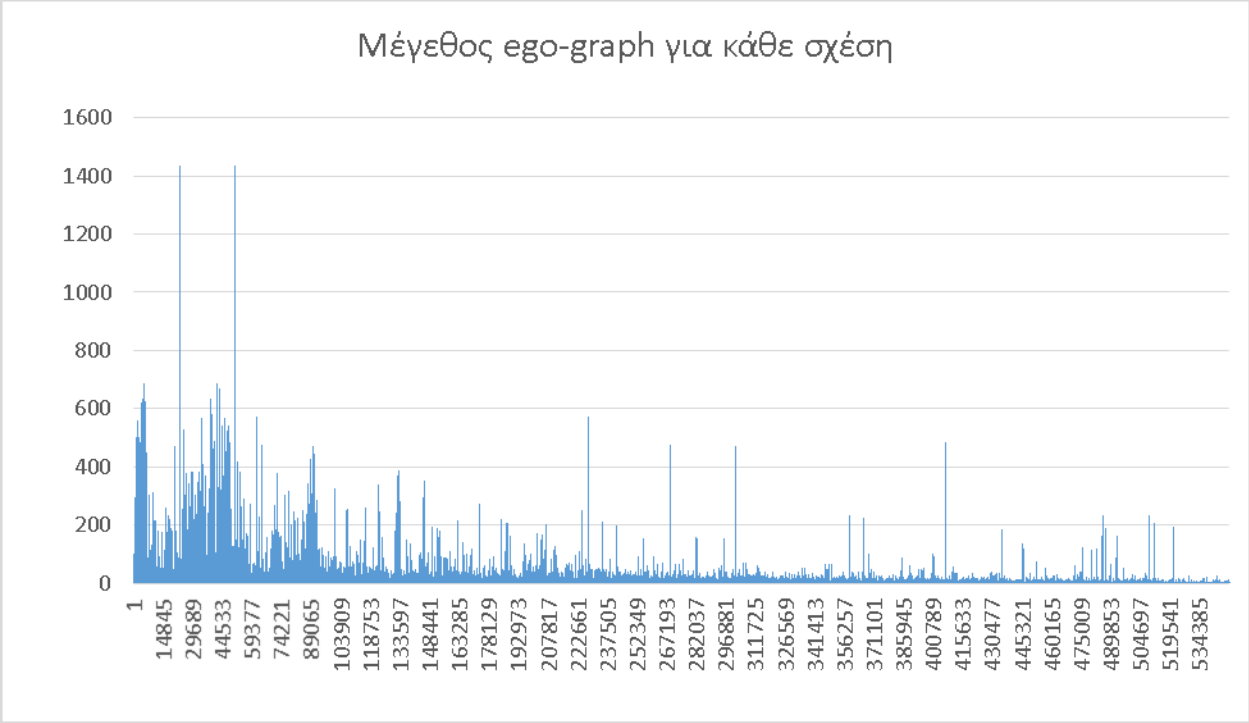
Εικ. 20 Ποσοστά επιτυχίας πρόβλεψης σε ολόκληρο το δείγμα

Τα παρακάτω γραφήματα παρουσιάζουν κάποια γενικότερα, συμπληρωματικά δεδομένα που συλλέχθηκαν από τις προσομοιώσεις μας, με αντίστοιχη επεξήγηση για το κάθε ένα.



Εικ. 21 Διάγραμμα που απεικονίζει την αναλογία followers και friends

Στο παραπάνω διάγραμμα με μπλέ χρώμα απεικονίζεται το πλήθος των χρηστών που έχουν συνάψει σχέση με τον χρήστη αναφοράς στον οριζόντιο άξονα (θετική ή αρνητική). Με πορτοκαλί χρώμα αντιπροσωπεύεται το πλήθος των χρηστών με τους οποίους έχει συνάψει σχέσεις ο ίδιος ο χρήστης αναφοράς. Συμπεραίνουμε ότι οι χρήστες που ξεκίνησαν να χρησιμοποιούν το Slashdot πρώτοι είναι πάρα πολύ δημοφιλείς και έχουν συνάψει σχέσεις μαζί τους πολλοί άλλοι χρήστες καθώς επίσης και το ότι οι νεότεροι χρήστες είναι λιγότερο δημοφιλείς σε σχέση με τους πιο παλαιούς.

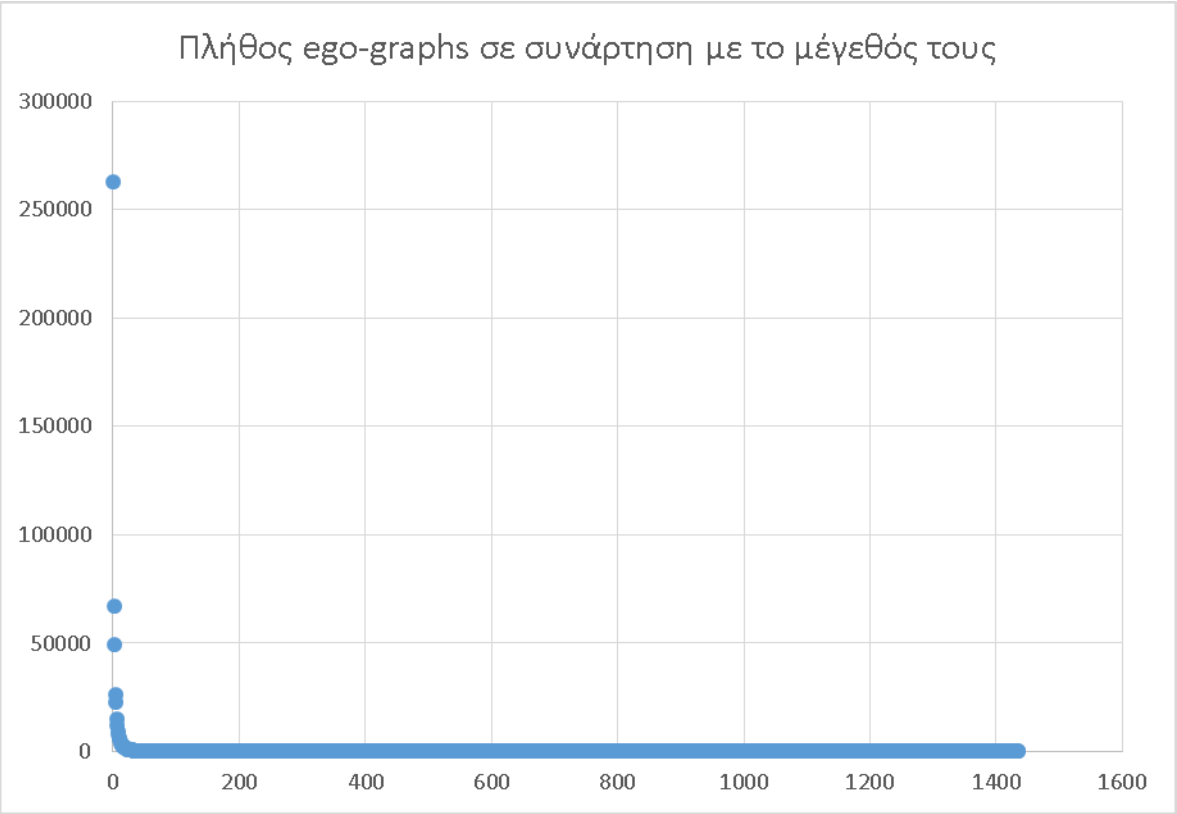


Εικ. 22 Η διασπορά του μεγέθους των ego-graphs.

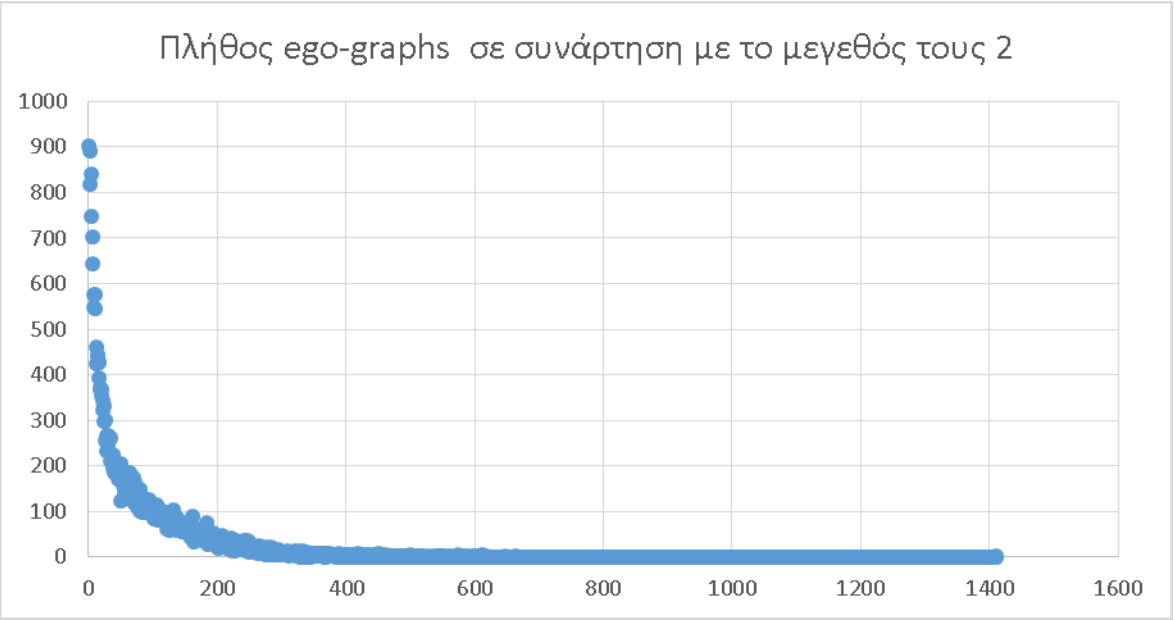
Στο παραπάνω διάγραμμα απεικονίζεται η διασπορά του μεγέθους των ego-graph,δηλαδή του πλήθους των κόμβων που περιέχουν σε συνάρτηση με τον αύξοντα αριθμό της σχέσης. Στο αρχείο υπάρχουν 549.202 σχέσεις. Συμπεραίνουμε πως οι πρώτες σχέσεις οι οποίες αντιστοιχούν σε παλαιότερους χρήστες (αφού η σειρά των σχέσεων είναι με αύξοντα αριθμό) έχουν πολύ μεγαλύτερους ego-graphs από τον μέσο όρο (8,5 κόμβοι ανά ego-graph).

Τέλος στα παρακάτω διαγράμματα φαίνεται το πλήθος των ego-graphs (άρα και το πλήθος των σχέσεων) σε συνάρτηση σε το μέγεθος των ego-graphs. Το πρώτο διάγραμμα αντιπροσωπεύει το 100% του δικτύου του Slashdot. Παρατηρούμε ότι η συντριπτική πλειοψηφία των ego-graphs έχουν πολύ μικρό μέγεθος (πολλοί από αυτούς έχουν μηδενικό).Για αυτόν τον λόγο αντιπαραθέτουμε άλλα δύο διαγράμματα που αποτελούμε μεγέθυνση του πρώτου. Το δεύτερο περιλαμβάνει το μέγεθος των ego-graphs που έχουν πλήθος μικρότερο από 1000 και το τρίτο μικρότερο από 100. Από τα διαγράμματα αυτά συμπεραίνει κανείς πως το Slashdot είναι σχετικά

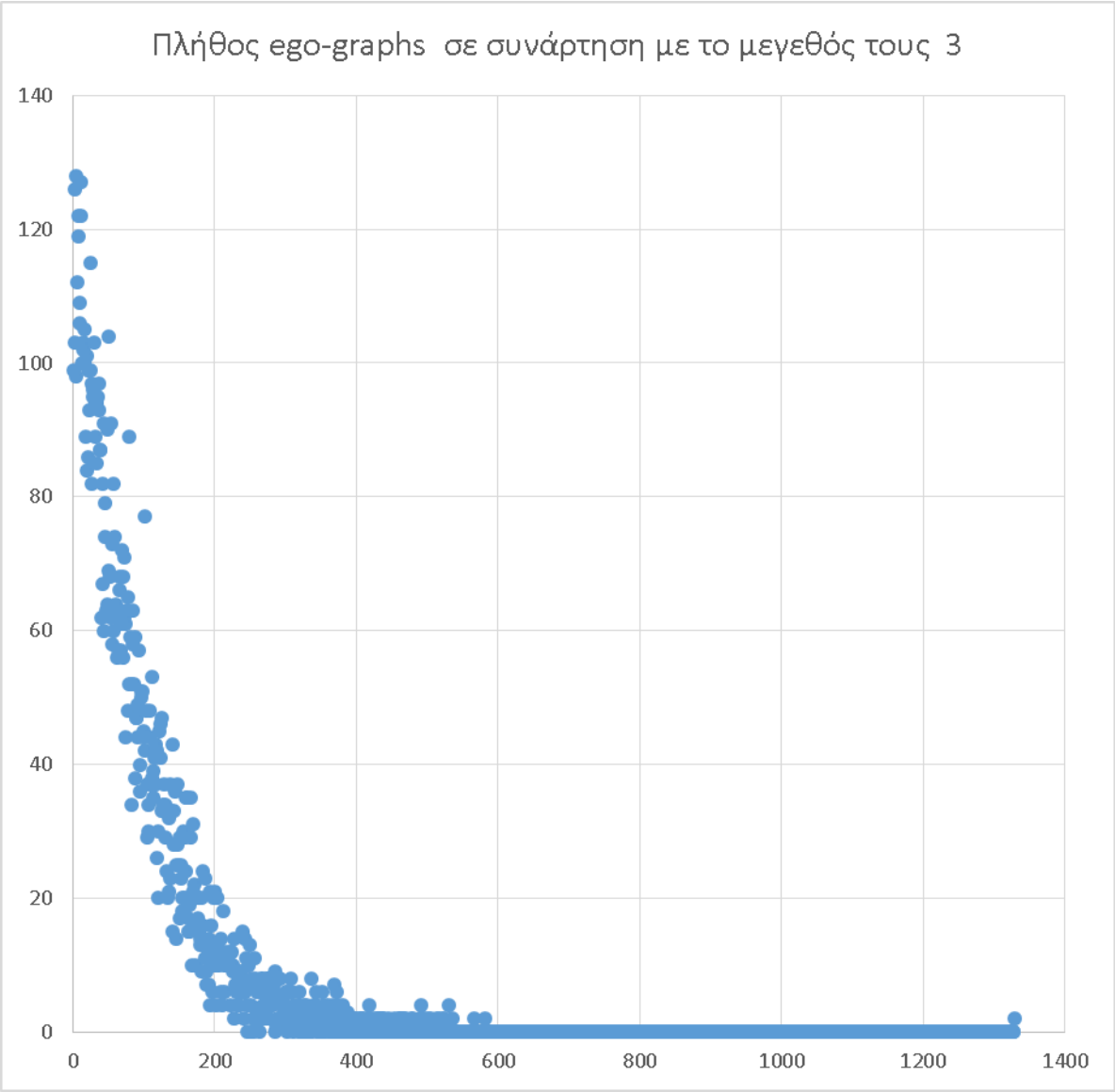
«αραιό» δίκτυο και πως οι περισσότεροι από τους νεότερους χρήστες του συνάπτουν πολύ λίγες σχέσεις μεταξύ τους.



Εικ. 23 Πλήθος ego-graphs σε συνάρτηση με το μέγεθος τους (100%)



Εικ. 24 Πλήθος ego-graphs σε συνάρτηση με το μέγεθος τους (<1000)



Εικ. 25 Πλήθος ego-graphs σε συνάρτηση με το μέγεθος τους (< 100)

6. Συμπεράσματα και μελλοντικά ζητήματα

6.1 Συμπεράσματα

Από το σύνολο των πειραμάτων μας παρατηρούμε ότι:

a) η μέθοδος της πρόβλεψης με χρήση των συχνών υπογράφων είναι η πιο ακριβής και εύχρηστη. Παρ' όλα αυτά αδυνατεί να δώσει πρόβλεψη για ένα κομμάτι του δικτύου, για τους λογούς που προαναφέρθηκαν.

b) Η μέθοδος που είναι βασισμένη στην δομική ευστάθεια είναι λιγότερο ακριβής από την μέθοδο με χρήση υπο-γράφων και το δείγμα για το οποίο δεν μπορεί να δώσει ασφαλές αποτέλεσμα είναι μεγαλύτερο. Όμως υπολογιστικά είναι πολύ πιο εύκολο να υλοποιηθεί.

c) Η συνδυαστική μέθοδος που προτείνουμε (πείραμα Γ) αν και υπολογίζει το σύνολο του δείγματος, δεν είναι η πιο ακριβής. Επιπρόσθετα είναι η πιο απαιτητική ως προς την χρήση υπολογιστικών πόρων.

d) Η μέθοδος που αντιστοιχεί στο πείραμα Β είναι απρόσμενα η πιο αποδοτική. Αυτό μπορεί να εξηγηθεί διότι στο συγκεκριμένο κοινωνικό δίκτυο το μεγαλύτερο ποσοστό των σχέσεων είναι θετικές (77,4%). Οπότε συμπεραίνοντας ότι το σύνολο του υπολειπόμενου δείγματος είναι θετικό επιτυγχάνουμε μεγαλύτερη ακρίβεια σε σχέση με την τυχαία-random (50-50) πρόβλεψη (πείραμα Α)καθώς στην συγκεκριμένη περίπτωση το πραγματικό αποτέλεσμα φαίνεται να απέχει αρκετά από το τυχαίο (50-50).

Συμπερασματικά προκύπτει ότι η ακρίβεια των μεθόδων πρόβλεψης εξαρτάται σε πολύ μεγάλο βαθμό από το είδος του κοινωνικού δικτύου το οποίο εξετάζουμε. Το Slashdot παρουσιάζει μεγάλη αρνητική πόλωση απέναντι σε συγκεκριμένους χρήστες αλλά συνολικά η πλειοψηφία των σχέσεων που συνάπτουν οι χρήστες του είναι θετικές. Ως εκ του το η μόνιμη πρόβλεψη για θετική σχέση (πείραμα Β) ήταν και η ακριβέστερη. Αυτό όμως το συμπέρασμα ενδέχεται να μην ισχύει σε κοινωνικά δίκτυα με διαφορετικά χαρακτηριστικά.

6.2 Μελλοντικά θέματα προς έρευνα

Για μελλοντική έρευνα προτείνεται να γίνει η προσομοίωση που υλοποιήσαμε για το Slashdot και για άλλα κοινωνικά δίκτυα (Wikipedia, Epinions κλπ.) έτσι ώστε:

Α. Να εξεταστεί η αποτελεσματικότητα της συνδυαστικής μεθόδου που προτείναμε.

Β. Να διερευνηθεί αν η απλούστερη μέθοδος με την θετική πρόβλεψη (πείραμα Β) η οποία στο δικό μας πείραμα φαίνεται να είναι η πιο αποδοτική, δίνει ικανοποιητικά αποτελέσματα και σε άλλα δίκτυα.

Επίσης ένα ενδιαφέρον πεδίο είναι η μελέτη της πρόβλεψης όταν η αξιολόγηση μεταξύ των χρηστών γίνεται σε δεκαδικό ή άλλο σύστημα, μεγαλύτερο του δυαδικού.

Γλωσσάριο

Slashdot: Κοινωνικό δίκτυο στο οποίο οι χρήστες συζητούν τεχνολογικά νέα.

C, C++: Η C είναι μια διαδικαστική γλώσσα προγραμματισμού γενικής χρήσης η οποία αναπτύχτηκε στις αρχές της δεκαετίας 1970 στα εργαστήρια Bell Labs για να χρησιμοποιηθεί για την ανάπτυξη του λειτουργικού συστήματος UNIX. Η C++ αναπτύχθηκε ως βελτίωση της C, και οι βελτιώσεις ξεκίνησαν με την προσθήκη κλάσεων, και ακολούθησαν, μεταξύ άλλων, εικονικές συναρτήσεις, υπερφόρτωση τελεστών, πολλαπλή κληρονομικότητα, πρότυπα, κτλ.

Java: Η Java είναι μια αντικειμενοστραφής γλώσσα προγραμματισμού που σχεδιάστηκε από την εταιρεία πληροφορικής Sun Microsystems.

IoT: Ακρώνυμο που σημαίνει Internet of Things. Το Internet of Things (IoT) είναι το δίκτυο των φυσικών αντικειμένων ή «πραγμάτων» ενσωματωμένων με ηλεκτρονικά κυκλώματα, λογισμικό, αισθητήρες και σύνδεση στο δίκτυο τα οποία επιτρέπουν σε αυτά τα αντικείμενα να συλλέγουν και να ανταλλάσσουν δεδομένα.

JVM: Η εικονική μηχανή της Java είναι μια αφηρημένη υπολογιστική μηχανή η οποία επιτρέπει σε ένα σύστημα να τρέξει προγράμματα java.

OSN: Το ακρωνύμιο αυτό σημαίνει Online Social Network. Πρόκειται για ένα κοινωνικό δίκτυο που στεγάζεται σε ένα site και επιτρέπει στους χρήστες του να αλληλοεπιδρούν μεταξύ τους, προσομοιώνοντας ένα κοινωνικό δίκτυο στον πραγματικό κόσμο.

Trolling: Ο όρος αυτός αναφέρεται σε σχόλια και ανακοινώσεις χρηστών που έχουν χαρακτήρα προσβλητικό και σκοπεύουν στην πόλωση της συζήτησης.

Netbeans: Εφαρμογή στην οποία ο χρήστης μπορεί να γράψει προγράμματα σε διάφορες γλώσσες και να χρησιμοποιήσει διάφορα βοηθήματα για τον σκοπό αυτό.

Βιβλιογραφία

1. RFC 1122, Requirements for Internet Hosts -- Communication Layers, 1.1.2 Architectural Assumptions, 1989
2. Feng Xia ,LaurenceT.Yang , Lizhe Wang and Alexey Vinel, (2012) International Journal Of Communication Systems Int. J. Commun. Syst. 2012; 25:1101–1102 Editorial Internet of Things,
3. Future Generation Computer Systems, Volume 29, Issue 7, September 2013, Pages 1645–1660 Including Special sections: Cyber-enabled Distributed Computing for Ubiquitous Cloud and Network Services & Cloud Computing and Scientific Applications — Big Data, Scalable Analytics, and Beyond
4. K. Ashton, (2009) That “Internet of Things” thing, RFiD Journal (2009)
5. H. Sundmaeker, P. Guillemin, P. Friess, S. Woelfflé, (2010) Vision and challenges for realising the Internet of Things, Cluster of European Research Projects on the Internet of Things—CERP IoT, 2010.
6. Danah M. Boyd, Nicole B. Ellison,(2007), Social Network Sites: Definition, History, and Scholarship, Danah
7. Danah M. Boyd, and Nicole B. Ellison, (2007) Journal of Computer-Mediated Communication, Volume 13, Issue 1, pages 210–230, October 2007, Social Network Sites: Definition, History, and Scholarship
8. Leading social networks worldwide as of August 2015, ranked by number of active users (in millions) http://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/)
9. Phillips, Sarah. "(2015) A brief history of Facebook". the Guardian
10. Carlson, Nicholas (2010). "At Last – The Full Story Of How Facebook Was Founded". Business Insider
11. Number of monthly active Facebook users worldwide as of 2nd quarter 2015 (in millions),http://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/)
12. R. V. Guha, R. Kumar, P. Raghavan, and A. Tomkins. (2004) Propagation of trust and distrust. In *Proc. 13th WWW*
13. Leskovec J, Huttenlocher D, Kleinberg J (2010) Predicting positive and negative links in online social networks In: Proceedings of the 19th international conference on World wide web. ACM, pp. 641–650.
14. Wang G-N, Gao H, Chen L, Mensah DNA,Fu Y (2015) Predicting Positive and Negative Relationships in Large Social Networks. PLoS ONE 10(6): e0129530.
15. Leskovec J, Huttenlocher D, Kleinberg J (2010) Signed networks in social media In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 1361–1370.
16. Grabowicz PA, Ramasco JJ, Eguíluz VM (2013) Dynamics in online social networks. Dyn Complex Networks 2: 3–17.
17. Li Y, Chen W, Wang Y, Zhang ZL (2013) Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships In: Proceedings of the

sixth ACM international conference on Web search and data mining. ACM, pp. 657–666. Available: http://dl.acm.org/citation.cfm?id=2433478
18. Scott J, Peter J (2011) The sage handbook of social network analysis. London: SAGE publications; 640 p.
19. Zhang QM, Lü L, Wang WQ, Zhou T (2013) Potential theory for directed networks.
20. Liao H, Zeng A, Xiao R, Ren Z-M, Chen D-B, Zhang Y-C (2014) Ranking Reputation and Quality in Online Rating Systems.
21. Yu M, Zhao K, Yen J, Kreager D (2013) Recommendation in Reciprocal and Bipartite Social Networks—A Case Study of Online Dating In: Social Computing, Behavioral-Cultural Modeling and Prediction. Springer; Berlin Heidelberg, pp. 231–239.
22. Lü L, Zhou T (2011) Link prediction in complex networks: A survey. Physica A 390: 1150–1170.
23. Liao H, Xiao R, Cimini G, Medo M (2014) Network-Driven Reputation in Online Scientific Communities.
24. Francesco Ricci and Lior Rokach and Bracha Shapira, Introduction to Recommender Systems Handbook, Recommender Systems Handbook, Springer, 2011, pp. 1-35
25. Resnick, P., Varian, H.R.: Recommender systems. Communications of the ACM 40 (3), 56– 58 (1997)
26. Mahmood, T., Ricci, F, (2009) Improving recommender systems with adaptive conversational strategies. In: C. Cattuto, G. Ruffo, F. Menczer (eds.) Hypertext, pp. 73–82. ACM (2009)
27. Burke, R.: Hybrid web recommender systems (2007) In: The Adaptive Web, pp. 377–408. Springer Berlin / Heidelberg
28. McSherry, F., Mironov, I.(2009) Differentially private recommender systems: building privacy into the net. In: KDD ’09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 627–636. ACM, New York, NY, USA
29. Goldberg, D., Nichols, D., Oki, B.M., Terry, D.(1992) Using collaborative filtering to weave an information tapestry. Commun. ACM 35 (12), 61–70
30. Anand, S.S., Mobasher, B (2005) Intelligent techniques for web personalization. In: Intelligent Techniques for Web Personalization, pp. 1–36. Springer
31. M. J. Brzozowski, T. Hogg, and G. Szabó. (2008) Friends and foes:ideological social networking. In *Proc. 26th CHI.*
32. Shardanand, U., and Maes, P. (1995) Social information filtering: algorithms for automating “word of mouth”. In Proc. CHI 1995, ACM Press, 210-217.
33. Sinha, R. R., and Swearingen, K. (2001) Comparing recommendations made by online systems and friends. In Proc. DELOS Workshop: Personalisation and Recommender Systems in Digital Libraries (2001)
34. Fritz Heider (1946), Attitudes And Cognitive Organization First published in The Journal of Psychology, 21, 107-112.).
35. D. Cartwright and F. Harary, (1956) “Structural Balance: A Generalization of Heider’s Theory,” Psychological Rev., vol. 63, no. 5, 1956, pp. 277–293
36. Social balance on networks (2006) The dynamics of friendship and enmity, T. Antal, P.L. Krapivsky, S. Redner, Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215, United States.

37. S. Wasserman, K. Faust, (1994) Social Network Analysis: Methods and Applications, Cambridge University Press, New York.
38. Ye J, Cheng H, Zhu Z, Chen M (2013) Predicting positive and negative links in signed social networks by transfer learning. In: Proceedings of the 22nd international conference on World Wide Web. International World Wide Web Conferences Steering Committee. ACM, pp. 1477–1488. Available: http://dl.acm.org/citation.cfm?id=2488517
39. Davis JA (1963) Structural balance, mechanical solidarity, and interpersonal relations. Am J Soc 68: 444–462.
40. Davis JA, Leinhardt S (1972) The Structure of Positive Interpersonal Relations in Small Groups. Sociological Theories in Progress 2: 218–251.
41. J. Kunegis, A. Lommatzsch, and C. Bauckhage. (2009) The Slashdot Zoo: Mining a social network with negative edges. In *Proc. 18th WWW*, pages 741–750.
42. A. Papaoikonomou, M.Kardara, K.Tserpes, D.Varvarigou (2014) Edge sign prediction in social network via frequent subgraph Discovery. IEEE Internet Computing 18(5): 36-43 (2014)
43. https://snap.stanford.edu/data/#signnets
44. Schildt.H, (2005) Java, a beginner's guide, third edition, California USA McGraw – Hill/Osborne
45. NetBeans Release Roadmap. Netbeans.org. Retrieved on 2013-07-18.
46. original Xelfi homepage. Retrieved 2008-05-17.
47. Happy Birthday NetBeans - interview with Jaroslav "Yarda" Tulach". Netbeans.org. Retrieved 2008-05-17.