**Κοινωνικά Δίκτυα και Ανάκτηση Πληροφορίας**

*Θεωρητικό Project*

****

*Προσωπικά Στοιχεία Ομάδας*

Βελαώρα Μαρία

Στεφοπούλου Γεωργία

Τα papers που μας ανατέθηκαν:

1. Random Graphs with Prescribed 𝐾-Core Sequences: A New Null Model for Network Analysis
2. Hypergraph Ego-networks and Their Temporal Evolution
3. Detecting Strong Ties Using Network Motifs
4. Predicting Reciprocity in Social Networks

Η έρευνα μας καλύπτει *εις πλάτος* το αρχικό υλικό πρώτα μέσω της κατανόησης του και έπειτα μέσω της εύρεσης παρόμοιων δημοσιευμάτων στην σχετική περιοχή.

Εύρεση εξωτερικού υλικού παράλληλο στο κάθε θέμα, έγινε με την βοήθεια *Google* και *Google Scholars .*

Στο τέλος της αναφοράς υπάρχουν link για όλα τα επιπλέον κείμενα.

**1. Random Graphs with Prescribed 𝐾-Core Sequences: A New Null Model for Network Analysis**

Στο πρώτο κείμενο οι επιστήμονες προσπαθούν να μας γνωρίσουν σε ένα *μηδενικό νέο μοντέλο για την ανάλυση δικτύων*. Πρώτα αφού έκαναν έρευνα γύρω από τα ήδη υπάρχοντα δίκτυα μεγάλης κλίμακας, αναρωτήθηκαν αν είναι δυνατόν η δομή που υπάρχει σε αυτά να προκύψει και σε άλλα τυχαία δίκτυα με παρόμοια χαρακτηριστικά. Όμως αυτό εμποδίστηκε από μια μακροχρόνια πρόκληση του τομέα που αφορά τη σχετική έλλειψη λογικών ιδανικών μοντέλων. Για αυτό το λόγο οι ερευνητές προτείνουν μια νέα οικογένεια μηδενικών μοντέλων δικτύου που λειτουργούν με τη διάσπαση κ πυρήνων. Με αυτή την πρόταση μας επιτρέπεται να έχουμε έναν τυχαίο γράφο G που δειγματοληπτείται σχεδόν ομοιόμορφα, μεταξύ όλων των γραφών με την ίδια ακολουθία αριθμών πυρήνων με το G, αλλά και να συγκρίνουμε τα παρατηρούμενα δίκτυα G με τυχαίους γράφους που έχουν τους ίδιους αριθμούς πυρήνων.

Στον τομέα της ανάλυσης δικτύων, τα τυχαία γραφήματα είναι εδώ και πολύ καιρό απαραίτητα, και μια από τις βασικές εφαρμογές τους ήταν ως μηδενικά μοντέλα. Χρησιμοποιώντας ισχυρά μηδενικά μοντέλα, μπορούμε να προσδιορίσουμε αν ένας τυχαίος γράφος με τις ίδιες ιδιότητες θα εμφάνιζε το ίδιο φαινόμενο με το παρατηρούμενο φαινόμενο του δικτύου. Υπό το πρίσμα των μειονεκτημάτων του μοντέλου διαμόρφωσης, οι ερευνητές έχουν αναζητήσει εναλλακτικά μηδενικά μοντέλα στα οποία δειγματοληπτούμε ομοιόμορφα ή ωτο-ομοιόμορφα σε διάφορες ομάδες γραφημάτων που χαρακτηρίζονται από τα χαρακτηριστικά ενός συγκεκριμένου πραγματικού δικτύου. Οι Stanton και Pinar παρουσιάζουν τον τρόπο δειγματοληψίας από γραφήματα που μοιάζουν πολύ με ένα παρατηρούμενο δίκτυο τόσο ως προς τα ζεύγη βαθμών (di, j) όσο και ως προς την ακολουθία βαθμών. Επομένως, και οι επιστήμονες επιμένουν με το βασικό ερώτημα: μπορεί να παρασκευαστεί ένα μηδενικό μοντέλο με δειγματοληψία από μια οικογένεια γραφημάτων που ταιριάζουν με τα χαρακτηριστικά του γραφήματος g με τέτοιον τρόπο ώστε τα προκύπτοντα τυχαία δείγματα να παρουσιάζουν μη τετριμμένα πλούσια τοπική δομή και δομή κοινότητας; Η απάντηση είναι ναι και αυτή η εργασία μας δίνει μια νέα προσέγγιση στο ερώτημα, παίρνοντας ομοιόμορφα δείγματα από γραφήματα και δείχνοντας ότι μπορούν να οδηγήσουν σε δυνητικά διαφορετικά συμπεράσματα όταν χρησιμοποιούνται ως μηδενικά μοντέλα. Ψάχνοντας λοιπόν μια εμπεριστατωμένη απάντηση στο ερώτημα, η έρευνα περνά από διάφορες «φάσεις». Αναφορικά με την δειγματοληψία, αναφέρουμε την μέθοδο της αλυσίδας Markov. Μια αλυσίδα Markov είναι η συλλογή όλων των γραφημάτων με μια συγκεκριμένη ακολουθία βασικών τιμών και οι μεταβάσεις της είναι μια συλλογή τροποποιήσεων του γραφήματος που διατηρούν τις βασικές τιμές ανέπαφες. Έτσι, μια αλυσίδα Markov που είναι αξιόπιστα ομοιόμορφη σε όλα αυτά τα δίκτυα παράγεται με την εφαρμογή τυχαίας αλλαγής σε έναν υποκείμενο γράφο. Είναι σημαντικό να προσδιοριστεί αν ορισμένες συχνότητες δευτερευόντων υπογραφών σε ένα δεδομένο γράφημα είναι αισθητά μεγαλύτερες, μικρότερες ή ίδιες με εκείνες που παρατηρούνται σε ένα τυχαίο γράφημα. Παρατηρείται πώς μια σύγκριση με τυχαίους γράφους με την ίδια ακολουθία βαθμών με τον G θα μπορούσε ενδεχομένως να δώσει διαφορετικά αποτελέσματα. Αποδεικνύεται ότι οι τυχαίοι περίπατοι στην αλυσίδα Markov θα καταλήξουν σε μια στάσιμη κατανομή που είναι ομοιόμορφη σε γραφήματα με δεδομένο βαθμό. Συνεχίζουν αναφέροντας κι άλλες συναφείς εργασίες, όπως τις έρευνες των Sala κ.ά. και Drobyshevskiy   
και Turdakov, που σχετίζονται αφού χρησιμοποιούνται ως μηδενικά μοντέλα, το μοντέλο διαμόρφωσης, το μοντέλο Chang-Lu, ο δειγματολήπτης των Pinar και Stanton.

Το δεύτερο «κεφάλαιο» αναφέρεται στην δημιοργία τυχαίων γραφημάτων με δεδομένη ακολουθία βασικών τιμών, όπου οι γράφοι είναι μη κατευθυνόμενοι και απλοί, χωρίς αυτό-βρόχους ή παράλληλες ακμές. Ορίζεται ένα σύνολο κινήσεων, και εκτελούμε την αλυσίδα Markov από ένα αυθαίρετο σημείο εκκίνησης.

Αναλύοντας το πρόβλημα υλοποίησης: Δεδομένης μιας ακολουθίας c, πώς μπορούμε να προσδιορίσουμε αποτελεσματικά αν υπάρχει ένας γράφος που έχει αυτή την ακολουθία ως βασική τιμή του; Οι Erdos και Gallai έλυσαν το ανάλογο πρόβλημα για ακολουθίες βαθμών. Ένας γράφος ονομάζεται 𝑑-κανονικός εάν όλοι οι βαθμοί των κόμβων του είναι ίσοι.

Στο επόμενο «βήμα» ορίζουμε ένα σύνολο κινήσεων στον προ ολίγου ορισμένο χώρο καταστάσεων και παρέχουμε τρόπους μετασχηματισμού ενός δεδομένου γραφήματος σε Sc σε άλλα γραφήματα Sc. Αφού καταφέρουμε αυτόν τον στόχο, σειρά έχει η συνδεσιμότητα του χώρου καταστάσεων για την οποία χρησιμοποιούμε αυτό το σύνολο. Δείχνουμε ότι το μη κατευθυνόμενο γράφημα Hc είναι συνδεδεμένο και ακολουθουμε τα εξης βήματα: Συνδέονται όλες οι ακμές με τον κορυφαίο πυρήνα, εκείνος μετατρέπεται σε c-ομοιόμορφο γράφημα, ο c- ομοιόμορφος πυρήνας μετατρέπεται σε άλλον και τελικά συνδιάζονται οι υποδιαδρομές.

Στην Τρίτη ενότητα, εξετάζουμε τους υπολογιστικούς παράγοντες που εμπλέκονται στην εκτέλεση της αλυσίδας Markov. Υπολογίζονται οι βασικοί αριθμοί ενός γραφήματος, που παρέχονται ως SparseMatrix, και κατασκευάζεται ένας πίνακας βασικών τιμών, ταξινομημένος από τη μεγαλύτερη προς τη μικρότερη. Ο αριθμός των εφικτών μεταβάσεων από αυτόν τον γράφο σε άλλους γράφους εκτιμάται στη συνέχεια ως ανώτερο όριο.

Στο τέταρτο μέρος, αποδεικνύεται ότι το μοντέλο διαμόρφωσης, το οποίο καθορίζει τους βαθμούς των κόμβων σε ένα δίκτυο γραφημάτων, μπορεί να αντικατασταθεί από το μοντέλο μηδενικής τιμής πυρήνα. Με διάφορους σημαντικούς τρόπους, το μοντέλο διαμόρφωσης και το μοντέλο βασικών αξιών οδηγούν σε διαφορετικά αποτελέσματα. Το μοντέλο διαμόρφωσης μπορεί να κάνει κάποιους μικροσκοπικούς υπογράφους να φαίνονται πιο συνηθισμένοι στα πραγματικά δεδομένα του δικτύου, καθορίζοντας μόνο έναν μικρό αριθμό βαθμών κόμβων, γεγονός που εξαφανίζει την πλειονότητα της τοπικής δομής. Για το σκοπό αυτό, λαμβάνουμε 50 δείγματα από έναν δειγματολήπτη για απλούς γράφους με ετικέτες κορυφών σε ένα μοντέλο διαμόρφωσης αλυσίδας Markov. Στη συνέχεια, τα στατιστικά στοιχεία του παραγόμενου γράφου συγκρίνονται με την έξοδο του μοντέλου διαμόρφωσης. Αυτά τα δείγματα έχουν συνολικά σημαντικά λιγότερα τρίγωνα από τα αντίστοιχα σύνολα δεδομένων. Οι υπογράφοι με μοτίβο είναι εκείνοι που εμφανίζονται αισθητά συχνότερα ή λιγότερο συχνά από ό,τι στο μηδενικό. Το μοντέλο διαμόρφωσης χρησιμοποιείται παραδοσιακά ως το μηδενικό μοντέλο. Εδώ, η ρύθμιση και το δικό μας μοντέλο k-core θεωρούνται και τα δύο ως μηδενικά μοντέλα. Οποιοδήποτε μοντέλο τυχαίου γράφου που τροποποιεί ένα συγκεκριμένο συστατικό της δομής του πραγματικού δικτύου μπορεί να χρησιμοποιηθεί για τη δημιουργία του SRP. Οι θετικές τιμές SRP θεωρούνται απόδειξη ότι ο υπογράφος είναι ένα αξιοσημείωτα άφθονο μοτίβο στα δεδομένα του δικτύου. Το βασικό σημείο είναι αν έχουμε τα ίδια αποτελέσματα όταν χρησιμοποιούμε ένα μηδενικό μοντέλο με βάση τον πυρήνα σε αντίθεση με ένα μοντέλο που καταργεί την τοπική δομή.

Σε αυτό το μέρος, όπου η ακολουθία βασικών τιμών δημιουργείται από ένα σύνολο δεδομένων του πραγματικού κόσμου, εξετάζουμε την κατανομή του αριθμού των ακμών σε τυχαία δείγματα που προκύπτουν από την τεχνική μας. Παρατηρούμε ότι δεν είναι ακριβώς οι ίδιες όταν συγκρίνουμε την ακολουθία βαθμών των τυχαίων δειγμάτων με τα αρχικά δεδομένα.

Στην τελική διερεύνηση, για καλύτερη κατανόηση χρησιμοποιούμε ως παράδειγμα ένα σύνολο-περιβάλλον δικηγόρων. Το σύνολο δεδομένων για τους δικηγόρους έχει διάφορα χαρακτηριστικά για κάθε κόμβο και μετράμε την ποικιλία του δικτύου για την ιδιότητα στην εταιρεία (εταίρος ή συνεργάτης), την τοποθεσία του γραφείου, το φύλο, τη νομική σχολή και άλλα χαρακτηριστικά . Η ανομοιογένεια είναι θετική εάν υπάρχει τάση εμφάνισης ακμών μεταξύ κόμβων που μοιράζονται το ίδιο χαρακτηριστικό.

Συμπερασματικά, παρόλο που οι ερευνητές βρέθηκαν αντιμέτωποι με δύσκολα ερωτήματα κατάφεραν με ποικίλα παραδείγματα και ενδελεχή έρευνα να παρουσιάσουν μια νέα οικογένεια τυχαίων γραφημάτων που μπορούν να λειτουργήσουν ως μια κατηγορία μηδενικών μοντέλων στην ανάλυση δικτύων.

*Στη συνέχεια παραθέτουμε υλικό σχετικό με το αρχικό μας.*

* **The Giant k-core of a Random Graph with a Specified Degree Sequence**

Daniel Fernholz Department of Computer Sciences University of Texas at Austin fernholz@cs.utexas.edu Vijaya Ramachandran Department of Computer Sciences University of Texas at Austin vlr@cs.utexas.edu November 26, 2003

Ο k-πυρήνας ενός γράφου είναι ο μέγιστος επαγόμενος υπογράφος με ελάχιστο βαθμό k. Σε αυτή την εργασία, εντοπίζονται συνθήκες υπο τις οποίες ο πυρήνας k ενός τυχαίου γράφου με συγκεκριμένο βαθμό ακολουθίας περιέχει σχεδόν σίγουρα ένα σταθερό κλάσμα των κορυφών του. Το πρόβλημα αυτό έχει μελετηθεί νωρίτερα από τους Pittel, Spencer και Wormald για την ειδική περίπτωση ενός τυχαίου γράφου που προέρχεται από το Gn,m.

Οι συγγραφείς, ξεκινάνε με τα προκατασκευαστικά. Μια ακολουθία D = {d1, d2, ... , dn} είναι γραφική αν το σύνολο των γραφικών παραστάσεων D (δηλαδή τέτοιων ώστε ο βαθμός της i-οστής κορυφής να είναι di) είναι μη κενό. G(D) είναι ένας τυχαίος γράφος με ασυμπτωτική ακολουθία βαθμών D, όπου D είναι μια ακολουθία γράφων με xed ακολουθίες βαθμών. Εδώ εξετάζεται αν ένας γράφος D αποτελεί μέρος μιας αραιής ομαλής ασυμπτωτικής ακολουθίας βαθμών ή αν περιέχει βρόχους ή/και πολλαπλές ακμές. Χρησιμοποιείται το CM μοντέλο του Bollobas για να υπολογιστεί ποιος τύπος γράφου είναι πιθανότερο να εµφανιστεί σε µια συγκεκριµένη περίπτωση.

Στο δέυτερο μέρος προσδιορίζονται οι συνθήκες υπο τις οποίες ένας τυχαίος γράφος με ακολουθία βαθμών D έχει έναν γιγαντιαίο k-πυρήνα. Σε όλο το κομμάτι, η ανάλυση υποστηρίζεται από ποικίλα θεωρήματα και αποδείξεις. Ένα τυχαίο γράφημα με ακολουθία βαθμών D συμπεριφέρεται τοπικά σαν μια διαδικασία με βάση την κατανομή υπολειμμάτων fig. Οι Pittel, Spencer και Wormald χρησιμοποίησαν το μοντέλο γραφήματος Gn,m για να αποδείξουν ότι ένας γιγαντιαίος k-πυρήνας σε ένα τυχαίο γράφημα σχετίζεται με την πιθανότητα ενός ενημερωμένου δέντρου διακλάδωσης.

Στην συνέχεια, επαληθεύουν τις συνθήκες των προαναφερόμενων θεωρημάτων. Μπορεί όμως να υπάρξει δυσκολία στο να βρεθούν xed points.

* **Random graph models of social networks**

M. E. J. Newman\*†, D. J. Watts‡, and S. H. Strogatz§\*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501;‡Department of Sociology, Columbia University, 1180 Amsterdam Avenue,New York, NY 10027; and§Department of Theoretical and Applied Mechanics, Cornell University, Ithaca, NY 14853-1503

Σε αυτό το άρθρο περιγράφονται ορισμένα νέα ακριβώς επιλύσιμα μοντέλα της δομής των κοινωνικών δικτύων, βασισμένα σε τυχαίους γράφους με αυθαίρετες κατανομές βαθμού. ∆ίνονται µοντέλα τόσο για απλά µονοµερή δίκτυα, όπως τα δίκτυα γνωριµίας, όσο και για διµερή δίκτυα, όπως τα δίκτυα σύνδεσης. Οι ερευνητές συγκρίνουν τις προβλέψεις των µοντέλων µας µε δεδοµένα για ορισµένα πραγµατικά κοινωνικά δίκτυα και διαπιστώνεται ότι σε ορισµένες περιπτώσεις, τα µοντέλα συµφωνούν αξιοσηµείωτα µε τα δεδοµένα. Κι άλλες μελέτες έχουν δείξει ότι ο αριθμός των δεσμών που έχουν οι φορείς με άλλους φορείς, οι λεγόμενοι "βαθμοί" τους, είναι ιδιαίτερα παραμορφωμένοι. Αυτή η διαστρέβλωση θα μπορούσε να έχει αντίκτυπο στον τρόπο με τον οποίο λειτουργούν οι κοινότητες, συμπεριλαμβανομένου του τρόπου με τον οποίο οι πληροφορίες διακινούνται μέσα σε αυτές. Σε αυτό το άρθρο περιγράφονται ορισμένα νέα μοντέλα κοινωνικών δικτύων που μας επιτρέπουν να διερευνήσουμε άμεσα τις επιπτώσεις των διαφορετικών κατανομών βαθμών. Πρόσφατες εργασίες σχετικά με τα κοινωνικά δίκτυα έχουν επικεντρωθεί σε τρία διακριτικά χαρακτηριστικά της δομής των δικτύων. Το φαινόμενο του "μικρού κόσμου" μας κάνει να αναρωτιόμαστε πώς δύο άνθρωποι μπορούν να έχουν μια σύντομη διαδρομή σύνδεσης γνωριμιών σε ένα δίκτυο που έχει άλλη κοινωνική δομή, όπως νησιωτικές κοινότητες ή γεωγραφικά και πολιτιστικά εμπόδια. Σε πολλά δίκτυα του πραγματικού κόσμου η πιθανότητα σύνδεσης μεταξύ δύο συντελεστών είναι πολύ μεγαλύτερη αν έχουν έναν άλλο κοινό γνωστό ή περισσότερους. Αυτός ο συντελεστής ομαδοποίησης μπορεί να πάρει τιμές από μερικά ποσοστά έως 40 ή 50%, και άλλες μελέτες έχουν δείξει παρόμοια αποτελέσματα για άλλα δίκτυα. Η συνεργασία µεταξύ επιστηµόνων και ηθοποιών έχει ένα εφάπαξ κόστος, το χρόνο και την προσπάθεια που καταβάλλεται για τη συγγραφή ενός άρθρου ή τη δηµιουργία µιας ταινίας, αλλά ο δεσµός υπάρχει επ' αόριστον στη συνέχεια. Η διατήρηση δεσµών στο τρίτο δίκτυο, το δίκτυο των διευθυντών εταιρειών, έχει σηµαντικό κόστος. Το 1959, οι Erdos και Renyi εισήγαγαν την έννοια του τυχαίου γράφου, που είναι ένα από τα απλούστερα μοντέλα δικτύου που υπάρχουν. Ωστόσο, ως μοντέλο ενός δικτύου του πραγματικού κόσμου, ο τυχαίος γράφος έχει ορισμένες σοβαρές ελλείψεις, συμπεριλαμβανομένης μιας εξαιρετικά λανθασμένης κατανομής βαθμών. Μπορεί να φτιαχτεί ένα μοντέλο δικτύου με την ίδια κατανομή βαθμών χρησιμοποιώντας έναν αλγόριθμο παρόμοιο με αυτόν των Molloy και Reed. Ο αλγόριθμος παράγει ένα γράφημα με ακριβώς τον επιθυμητό βαθμό degreek αλλά το οποίο είναι κατά τα άλλα τυχαίο.

Έπειτα ορίζεται υποκεφάλαιο με τίτλο «ακριβή αποτελέσματα». Σχεδόν όλα τα δίκτυα που απαντώνται στην κοινωνία και τη φύση φαίνεται να αναπτύσσονται μέσα στην περιοχή στην οποία υπάρχει το γιγαντιαίο στοιχείο. Ωστόσο, τα δίκτυα χωρίς προφανή γιγαντιαία συνιστώσα είναι σπάνια- είναι πιθανό ότι οι ερευνητές σπάνια σκέφτονται να εξετάσουν μια δικτυακή αναπαράσταση ενός συστήματος που δεν είναι σε μεγάλο βαθμό διασυνδεδεμένο.

Τα δίκτυα σύνδεσης είναι δίκτυα στα οποία οι φορείς συνδέονται μεταξύ τους με την κοινή συμμετοχή σε ομάδες ή συλλόγους. Οι πληροφορίες που περιέχουν αναπαρίστανται πληρέστερα ως γράφημα που αποτελείται από δύο είδη κορυφών. Μπορούμε να μοντελοποιήσουμε μαθηματικά τέτοια δίκτυα χρησιμοποιώντας μηχανήματα παρόμοια με αυτά που παρουσιάστηκαν στα Ακριβή Αποτελέσματα.

Τελικά, καταλήγουμε στο ότι, σε αυτό το άρθρο, γνωρίσαμε μια κατηγορία πρότυπων δικτύων που αποτελούν γενικεύσεις του πολύ μελετημένου τυχαίου γράφου των Erdos και Renyi. Οι επιστήμονες έδειξαν ότι μπορεί να υπάρξει μια μετάβαση φάσης κατά την οποία σχηματίζεται μια γιγαντιαία συνιστώσα συνδεδεμένων κορυφών. Εφάρμοσαν τη θεωρία γραφημάτων στην περίπτωση των διμερών τυχαίων γραφημάτων, τα οποία χρησιμεύουν ως μοντέλα για τα δίκτυα διασύνδεσης. Οι προβλέψεις για τις τυπικές αποστάσεις vertex-vertex και τους συντελεστές ομαδοποίησης διαφέρουν από τα εμπειρικά δεδομένα σε ορισμένες περιπτώσεις. Υπονοείται εν τέλει ότι οι αποκλίσεις υποδηλώνουν μη τυχαία κοινωνικά φαινόμενα στη διαμόρφωση του δικτύου.

* **A guide to null models for animal social network analysis**

Damien R. Farine\*,1,2,31Department of Collective Behaviour, Max Planck Institute for Ornithology, 78457 Konstanz, Germany;2Chair of Biodiversityand Collective Behaviour, Department of Biology, University of Konstanz, 78457 Konstanz, Germany; and3Department ofZoology, Edward Grey Institute of Field Ornithology, Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

Το κείμενο αυτό παρεκλίνει λίγο από τις έννοιες που έχουμε μελετήσει, στον κορμό τους όμως θεωρούμε μπορούν να σχετιστούν με την έρευνα που κάνουμε. Ξεκινώντας περιλιπτικά για το θέμα της μελέτης, ενα βασικό εργαλείο στην εργαλειοθήκη της ανάλυσης κοινωνικών δικτύων είναι τα μηδενικά μοντέλα. Ωστόσο, η εφαρμογή τους στον έλεγχο υποθέσεων δεν είναι στην φάση της σύνταξης του paper πολύ διαδεδομένη. Επιπλέον, υπάρχουν πολυάριθμες μέθοδοι για τη δημιουργία μηδενικών μοντέλων, καθεμία από τις οποίες έχει σχετικά πλεονεκτήματα και μειονεκτήματα και συχνά ελέγχει διάφορες υποθέσεις.

Στην παρούσα μελέτη, επισημαίνεται η αναγκαιότητα των μηδενικών μοντέλων για τον αξιόπιστο έλεγχο υποθέσεων σε έρευνες για τα κοινωνικά δίκτυα ζώων. Εξετάζουν την αποτελεσματικότητα διαφόρων στατιστικών δοκιμών και μηδενικών μοντέλων χρησιμοποιώντας προσομοιωμένα δεδομένα με γνωστές μεροληψίες παρατήρησης. Ο ερευνητής αποδεικνύει πώς οι μεταθέσεις των ακατέργαστων δεδομένων παρατήρησης (ή "προ-δικτύου") λαμβάνουν με αξιοπιστία υπόψη την υποκείμενη δομή στο κοινωνικό δίκτυο που παράγεται, γεγονός που μπορεί να μειώσει τόσο τα ποσοστά σφάλματος τύπου Ι όσο και τα ποσοστά σφάλματος τύπου ΙΙ. Ωστόσο, επειδή είναι δύσκολο να εκτελεστούν για ορισμένους τύπους δεδομένων, ιδίως για εκείνα που περιλαμβάνουν εστιακή παρακολούθηση και παρακολούθηση με GPS, οι μεταθέσεις των δεδομένων προ-δικτύου παραμένουν μάλλον ασυνήθιστες στην ανάλυση των κοινωνικών δικτύων των ζώων. Παρέχεται κώδικας R που εφαρμόζει κάθε μορφή μηδενικού μοντέλου στο ίδιο σύνολο προσομοιωμένων δεδομένων και περιγράφω απλές διαδικασίες που μπορούν να αναπτυχθούν γρήγορα σε πολλούς τύπους δεδομένων. Για τη χρήση εμπειρικών δεδομένων για τον έλεγχο των υποθέσεων, ο κώδικας R είναι απλό να τροποποιηθεί. Οι ερευνητές θα επωφεληθούν από την ευρεία υιοθέτηση των τεχνικών αντιμετάθεσης δεδομένων πριν από το δίκτυο, δεδομένου ότι θα καταστήσει δυνατό τον αξιόπιστο έλεγχο υποθέσεων.

Τα μηδενικά μοντέλα χρησιμοποιούνται για τη δημιουργία των προτύπων που αναμένονται από τα δεδομένα εν απουσία της διαδικασίας ενδιαφέροντος . Είναι ρουτίνες που δημιουργούν σύνολα δεδομένων με τα οποία μπορεί να συγκριθεί το παρατηρούμενο σύνολο δεδομένων. Τα πιο συνηθισμένα που χρησιμοποιούνται με τα κοινωνικά δίκτυα είναι οι έλεγχοι μετάθεσης, όπου τα ίδια τα παρατηρούμενα δεδομένα ανακατεύονται για να δημιουργηθούν τα τυχαία σύνολα δεδομένων. Όμως, η ανακατανομή μπορεί να διατηρήσει πτυχές των δεδομένων σταθερές, ενώ επιτρέπει σε άλλες να αλλάξουν. Τα μηδενικά μοντέλα λαμβάνουν υπόψη τους μη κοινωνικούς παράγοντες που επηρεάζουν τη συνύπαρξη των ατόμων έτσι ώστε να μπορούμε να εξάγουμε το σήμα των κοινωνικών παραγόντων που δομούν το κοινωνικό δίκτυο.

Τα μοντέλα αυτά είναι απαραίτητα για την ανάλυση κοινωνικών δικτύων ζώων. Τα κοινωνικά δεδομένα από τη φύση τους δεν είναι ανεξάρτητα. Πρέπει να υπάρχει ένα σημείο αναφοράς για σύγκριση όταν προσδιορίζεται κατά πόσον η συνολική δομή ενός δικτύου είναι μη τυχαία. Στην έρευνα των κοινωνικών δικτύων σε όλα τα επίπεδα, τα μηδενικά μοντέλα παρέχουν μια ισχυρή μέθοδο για την αξιολόγηση ιδεών. Μας επιτρέπουν έτσι να συγκρίνουμε δίκτυα, να εντοπίσουμε ανταγωνιστικές θεωρίες και ενδεχομένως να ελέγξουμε συγκεκριμένες υποθέσεις.

Συνεχίζει η έρευνα, με τον σχεδιασμό των μηδενικών μοντέλων τα βήματα του και μια προσομοίωση των μοντέλων. Κατά την διαδικασία αυτή είναι σημαντικό να αναρωτηθούμε κάποια πράγματα: Τι θα μπορούσε να συμβεί από τύχη; Και πως θα έδειχναν τα δεδομένα αν η διαδικασία ενδιαφέροντος ήταν παρούσα ή απούσα. Για την καλύτερη εξήγηση του σχεδιασμού φροντίζουν οι συγγραφείς να συνδυάσουν την έρευνα τους και να διευκολύνουν την κατανόηση από τους αναγνώστες. Στο επόμενο «υποκεφάλαιο» αναλύεται η κατασκευή ενός τεστ μεταθέσεων. Υπάρχουν πολλοί διαφορετικοί τρόποι με τους οποίους μπορούν να υλοποιηθούν οι μεταθέσεις. Αυτοί μπορούν να ταξινομηθούν σε δύο κατηγορίες: μεταθέσεις δικτύου και μεταθέσεις δεδομένων πριν από το δίκτυο. Σε αυτό το άρθρο, εξηγούν τον πιο συνηθισμένο έλεγχο μεταθέσεων (μεταθέσεις κόμβων). Για τις μεταθέσεις δεδομένων προ-δικτύων, ο αλγόριθμος ανταλλαγής μπορεί να περιορίσει τις ανταλλαγές να συμβαίνουν μεταξύ ατόμων που παρατηρούνται στο ίδιο σημείο ή/και εντός της ίδιας περιόδου. Με τον τρόπο αυτό, ο αλγόριθμος μπορεί να ελέγξει παράγοντες όπως η περιοχή κατοικίας, οι μεροληψίες της δειγματοληψίας παρατήρησης ή οι υποκείμενες διαφορές συμπεριφοράς που μπορεί να καλύπτουν την υπό εξέταση υπόθεση.

Προχωράνε, αναφορικά με την μετάθεση δεδομένων εστιακής παρατήρησης. Οι εστιακές παρατηρήσεις αποτελούν έναν ιδιαίτερα δύσκολο τύπο δεδομένων για την κατασκευή μηδενικών μοντέλων. Τα δεδομένα αυτά είναι δομημένα με βάση το εστιακό άτομο και η τυχαιοποίηση των δεδομένων με τη χρήση παραδοσιακών μεθόδων θα έσπαγε αυτό το χαρακτηριστικό. Οι προ-δικτυακές μεταθέσεις δεδομένων μπορούν να χρησιμοποιηθούν για την εκτέλεση μεταθέσεων ροής εστιακών δεδομένων. Σε μια πρόσφατη μελέτη σχετικά με τις νυσταγμένες σαύρες, το κοινωνικό δίκτυο κατασκευάστηκε από σημεία GPS που συλλέγονταν κάθε 10 λεπτά. Ωστόσο, αυτή η προσέγγιση εξακολουθούσε να ανταλλάσσει ολόκληρα κομμάτια δεδομένων μεταξύ των ατόμων. Αντ' αυτού, οι αρχές της ανταλλαγής, όταν παρατηρήθηκαν άτομα, (η βασική αρχή των Bejder, Fletcher & Brager (1998) ήταν μια πιο ισχυρή προσέγγιση.

Έπειτα, προτείνονται κι άλλες μέθοδοι για τα τεστ της ανάλυσης των κοινωνικών δικτύων, όπως QAP, MRQAP, ERGMs, και αμέσως μετά παρέχεται ένας «οδηγός» χρήσης μηδενικών μοντέλων αναλογα με την χωρική κατανομή, την δομή δικτύων κ.α.

Ολοκληρώνεται, με αναφορά στην σύγκριση δικτύων, μια χρήσιμη διαδικασία. Παραδειγματίζουν αναφέροντας, πως οι μεταθέσεις μπορούν να ελέγξουν αν τα δίκτυα διαφέρουν περισσότερο ή λιγότερο από το αναμενόμενο κατά τύχη. Στην περίπτωση αυτή, το παρατηρούμενο στατιστικό είναι ένα μέτρο της διαφοράς (π.χ. η διαφορά στο μέσο βαθμό).

Τελικά, τα μοντέλα αντιμετάθεσης δεδομένων πρo-δικτύου περιγράφηκαν για πρώτη φορά πριν από σχεδόν 20 χρόνια, αλλά δεν είναι ακόμη τόσο διαδεδομένα όσο θα έπρεπε. Τα μηδενικά μοντέλα αποτελούν σημαντικό μέρος του τεστ υποθέσεων στα κοινωνικά δίκτυα ζώων και απαιτείται προσεκτική εξέταση κατά την επιλογή του μηδενικού μοντέλου που θα χρησιμοποιηθεί.

**2. Hypergraph Ego-networks and Their Temporal Evolution**

Σε αυτό το άρθρο εξετάζεται μια *διαφορετική μορφή γράφων, οι υπερ-γράφοι*. Οι υπεργράφοι είναι μια γενίκευση των παραδοσιακών γραφημάτων στα οποία κάθε ακμή μπορεί να συνδέει οποιονδήποτε αριθμό κόμβων. Η μελέτη των υπεργραφών έχει αποφέρει μοναδικές γνώσεις σχετικά με τα δομικά χαρακτηριστικά των προσομοιωμένων συστημάτων. Η εντοπισμένη δομή αυτών των συστημάτων και ο τρόπος με τον οποίο μεταβάλλονται με την πάροδο του χρόνου είναι λιγότερο καλά κατανοητά. Παρακάτω, μελετάται αυτό το δίκτυο ego και διαμορφώνεται ένα αντίστοιχο μοντέλο ανακατασκευής τους.

Αρχικά, η δομή των υπεργράφων μπορεί να ταξινομηθεί σε δύο κατηγορίες: στατική ή χρονική και παγκόσμια ή τοπική. Οι Kook κ.a. εξέτασαν πρόσφατα τα συνολικά μοτίβα που συναντώνται σε προσωρινούς υπεργράφους του πραγματικού κόσμου. Λίγα είναι γνωστά για την τοπική, χρονική δομή των υπεργράφων ή των συστημάτων ανώτερης τάξης και για τον τρόπο με τον οποίο αλληλεπιδρούν μεταξύ τους. Αυτές οι αλληλεπιδράσεις μοντελοποιούνται συνήθως με τη χρήση ego-δικτύων, ένα δίκτυο αλληλεπιδράσεων ανά ζεύγη μεταξύ γειτόνων ενός κόμβου. Με την καλύτερη κατανόηση της χρονικής φύσης των ομαδικών αλληλεπιδράσεων, μπορούμε να αποκτήσουμε νέες γνώσεις σχετικά με το πώς αναπτύσσονται αυτά τα δίκτυα σε διάφορους τομείς. Ένα πρόβλημα αναφοράς για τα μοντέλα που επιδιώκουν να προβλέψουν την τοπική χρονική δομή των υπεργράφων είναι η χρονική ανακατασκευή των υπεργραφικών δικτύων. Περιγράφουμε ένα μοντέλο το οποίο, προσθέτοντας δομικά χαρακτηριστικά που ανακαλύφθηκαν σε υπεργραφικά ego-δίκτυα σε διάφορους τομείς, ξεπερνά αυτό το βασικό όριο. Χρησιμοποιούμε τρία σύνολα δεδομένων για να δείξουμε πόσο καλά το μοντέλο μας αποτυπώνει τη δομή υψηλότερης τάξης σε τρεις διαφορετικούς τομείς. Στο παρόν άρθρο θα εξεταστούν μόνο υπεργραφικά δίκτυα ego-networks με τουλάχιστον 20 υπερδιαστολές ή περισσότερες. Αυτό γίνεται έτσι ώστε πριν διατυπώσουμε οποιεσδήποτε προβλέψεις, το μοντέλο μας να μπορεί να παρατηρήσει έναν επαρκή αριθμό αλληλεπιδράσεων υψηλότερης τάξης. Η προσέγγισή μας γενικεύεται σε όλα τα σύνολα δεδομένων χωρίς αλλαγές στις ρυθμίσεις μάθησης και αποδίδει αξιοσημείωτα καλά με λίγα μόνο βασικά συνδυαστικά χαρακτηριστικά.

Οι συγγραφείς συνεχίζουν την μελέτη, παραθέτωντας κάποιους ορισμούς για την αναφερθείσα ορολογία. Επεξηγούν:

* τους **υπεργράφους** :Οι υπεργράφοι είναι γενικεύσεις των παραδοσιακών γραφημάτων όπου μια ακμή μπορεί να συνδέει οποιονδήποτε αριθμό κόμβων. Αναφερόμαστε σε αυτές τις ακμές ως υπερδιαστολές. Ένας υπεργράφος G = (V, E) αποτελείται από ένα σύνολο κόμβων V και υποσύνολα του V γνωστά ως E.
* την **Simplex:** εδώ συζητώνται οι υπεργράφοι στους οποίους κάθε υπερχείλος συνδέεται με μια συγκεκριμένη χρονική περίοδο. Ο αριθμός των κόμβων σε μια Simplex χρησιμεύει ως το μέγεθός της, ενώ το S συμβολίζει το σύνολο των απλοτήτων
* τον **υπεργράφο ego-δίκτυο**: Το υπεργράφημα ego-δίκτυο E ενός κόμβου u είναι το σύνολο των απλοτήτων που αναπαριστούν τις αλληλεπιδράσεις μεταξύ των γειτόνων του u. Υπάρχουν τρεις φυσικοί, διακριτοί τύποι ego-δικτύων: αστέρι, ακτινωτά και συνεπτυγμένα. Για λόγους συντομίας, θα αναφερόμαστε σε αυτά ως επί το πλείστον ως δίκτυα εγωισμού.
* το **Star Ego δίκτυο**: Ο πιο απλός τύπος Ego δικτύου. S είναι το σύνολο όλων των απλοποιημένων κόμβων και u είναι ο κόμβος του χρήστη. Όμως μόνο δομές με υπερεγέφυρες σε τουλάχιστον τρεις κόμβους παρουσιάζουν στην πραγματικότητα τον πλούτο αυτής της δομής.
* το **Radial Ego δίκτυο:** Το radial ego-δίκτυο αποτελείται από όλες τις απλότητες, όπου κάθε κόμβος της simplex είναι είτε ο κόμβος του χρήστη είτε ένας άλλος κόμβος. Τα radial ego-δικτυα είναι σε θέση να συμπεριλάβουν τις αλληλεπιδράσεις μεταξύ των αλλαγών, καθώς ορισμένες από αυτές τις απλότητες μπορεί να μην περιλαμβάνουν καθόλου τους κόμβους του χρήστη.
* το **Contracted Ego** **δικτυο** : Το contracted ego-δίκτυο είναι η τομή κάθε simplex στο S με το σύνολο A(u), όπου A είναι το σύνολο όλων των αλλαγών. Το C(u) περιλαμβάνει υποσύνολα απλοτήτων που αλληλεπιδρούν με υποκατεστημένους κόμβους εκτός του ego-δικτύου του u.

Η επόμενη ενότητα, κάνει αναφορά στα δεδομένα που χρησιμοποιούνται στην εργασία. Αναφέρονται 3 σύνολα δεδομένων που το καθένα αποτελείται από ένα σύνολο απλοποιήσεων με χρονοσήμανση. Με αυτό μπορούμε να κατασκευάσουμε Star Ego-δίκτυα, Radial Ego-δίκτυα και Contracted Ego-δίκτυα. Μας επιτρέπεται δηλαδή, χρησιμοποιώντας τα να αναλύσουμε την δομή Ego-δικτύων κάθε χρήστη (δίκτυα με μήκος τουλάχιστον 20 ή μεγαλύτερο για το coauth-DBLP και 10 ή για το email-Avocado και το threads-ask-ubuntu).

Προχωρώντας, εξετάζεται η εξέλιξη των δικτύων ego-networks που βρίσκονται στα παραπάνω και διαπιστώνεται ότι κάθε δίκτυο ego-network μπορεί να θεωρηθεί ως μια ένωση υποδικτύων σε σχήμα αστεριού, ένα για κάθε alter, και ότι η ανάλυση της χρονικής τους δομής μπορεί να μας δώσει πληροφορίες για την εξέλιξή τους.

Συνεχίζουν τις βασικές παρατηρήσεις ανα κατηγορίες.

Πρώτα αναφορικά με το μέγεθος της διασταύρωσης, διαπιστώνουμε ότι οι συνεχόμενες απλοποιήσεις σε ένα δίκτυο "εγώ" τείνουν να έχουν σχετικά μεγάλες διασταυρώσεις μεταξύ τους, γεγονός που υποδηλώνει ότι οι χρονικά γειτονικές αλληλεπιδράσεις υψηλότερης τάξης στα δίκτυα του πραγματικού κόσμου είναι παρόμοιες. Αυτό ισχύει για όλα τα ego-δίκτυα στα coauth-DBLP και email-Avocado καθώς και για τυχαία ανακατεμένα δίκτυα. Αποδείχτηκε ότι οι πρώτες αλληλεπιδράσεις στη διάρκεια ζωής ενός υπεργραφικού ego-δικτύου είναι παρόμοιες με εκείνες σε έναν κλασικό γράφο. Αυτά τα αποτελέσματα είναι σημαντικά για κάθε μοντέλο που προσπαθεί να συλλάβει τη χρονική δομή των ego-δικτύων, καθώς είναι και τα δύο σημαντικά για την κατανόηση του τρόπου με τον οποίο εξελίσσονται τα δίκτυα.

Στην συνέχεια αναλύονται τα alter-δίκτυα. Κάθε ego-δίκτυο οποιουδήποτε τύπου μπορεί να θεωρηθεί ως μια ένωση των αλληλεπιδράσεων του alter του. Η εξάπλωση του alter-δικτύου είναι ο μέσος χρόνος διάταξης μεταξύ δύο γειτονικών απλοτήτων σε ένα alter-δίκτυο. Αναλύεται αυτή η δομή χρησιμοποιώντας το coauth-DBLP και το email-Avocado για να μελετηθεί η εξέλιξη των ego-δικτύων. Για μεγάλα alter-δίκτυα, τα οποία ορίζεται ότι έχουν μέγεθος τουλάχιστον 10, μετράμε επίσης την εξάπλωση σε διάφορα τμήματα του alter-δικτύου. Οι συγγραφείς παρατήρησαν ότι το μεσαίο τρίτο είναι κατά μέσο όρο πολύ πιο πυκνό από το αρχικό και το τελικό τρίτο. Οι alter με υψηλό βαθμό κατά μέσο όρο τείνουν να εμφανίζονται νωρίτερα σε αυτά τα δίκτυα από ό,τι οι alter με χαμηλό βαθμό.

Το επόμενο «υποκεφάλαιο» προσεγγίζει φιλοσοφικά το θέμα μας. Σε ένα ego-δίκτυο, ο κόμβος χρήστη u κατέχει μια προνομιακή θέση στο κέντρο του δικτύου, αλλά πολλές από τις αλληλεπιδράσεις ανώτερης τάξης στα ego-δίκτυα δεν περιλαμβάνουν τον u. Αυτή η κατάσταση σημαίνει ότι το σύστημα που παρατηρούμε μπορεί να υπήρχε πριν από εμάς που ήμασταν εκεί για να το παρατηρήσουμε.

Τέλος, το ποσοστό καινοτομίας. Η καινοτομία μιας Simplex είναι ο αριθμός των κόμβων της Simplex που δεν έχουν περιληφθεί ποτέ σε προηγούμενες απλότητες. Στο εξής μετράμε τον ρυθμό με τον οποίο εισέρχονται νέοι κόμβοι στα δίκτυα εγωισμών για να κατανοήσουμε πότε και πόσο συχνά φτάνουν συνήθως νέοι κόμβοι.

Η μελέτη συνεχίζεται, με την λύση στο πρόβλημα της χρονικής ανακατασκευής των ego-δικτύων με υπεργράφους και την λεπτομερή ανάλυση του τρόπου αξιολόγησης και κατανόησης των μηχανισμών που διέπουν την εξέλιξη τους.

Πρώτα, η εργασία μάθησης. Οι επιστήμονες θα εκπαιδεύουν και θα δοκιμάζουν το μοντέλο σε ένα δείγμα 50000 δικτύων ego-networks μήκους 20 έως 50. Για το email-Avocado και το threads-ask-ubuntu θα πάρουν δείγμα περίπου 1000 και 1800 αντίστοιχα. Θα κατασκευάσουν ένα σύνολο χαρακτηριστικών από κάθε ego-δίκτυο. Αυτά θα περιλαμβάνουν χαρακτηριστικά με βάση την πυκνότητα διασταύρωσης, τη μέση εξάπλωση alter-δικτύων και τον αριθμό των προηγούμενων απλοτήτων. Εκπαιδευσαν ακόμη νευρωνικά δίκτυα για τα coauth-DBLP, email-Avocado και threads-ask-ubuntu.

Κάθε δίκτυο είχε αρχικό ρυθμό μάθησης 0,001 με ελάχιστο μέγεθος παρτίδας 200. Για κάθε μοντέλο, πραγματοποιήθηκε διασταυρούμενη επικύρωση 10 φορές και θα αναφέρεται η μέση ακρίβεια ταξινόμησης.

Περνάμε μετά στον αλγόριθμο χρονικής ανακατασκευής. Η ακρίβεια του μοντέλου μετριέται από το πόσα ζεύγη στην προβλεπόμενη σειρά είναι εκτός σειράς σε σχέση με την πραγματική απάντηση. Παρέχεται ο ψευδοκώδικας του αλγορίθμου πρόβλεψης που περιγράφεται λεπτομερέστερα στον Αλγόριθμο. Ο αλγόριθμος απαιτεί μόνο τρεις παραμέτρους: ego-δίκτυο π0, εποπτευόμενο μοντέλο M και συνολικές επαναλήψεις T. Στα βήματα 1-3: Ανταλλάζεται κάθε ζεύγος απλοτήτων στο π και εφαρμόζεται το M σε κάθε ego-δίκτυο που προκύπτει. Για κάθε προκύπτουσα διάταξη, εάν η πιθανότητα να είναι σωστά ταξινομημένη η προκύπτουσα διάταξη είναι υψηλότερη από εκείνη της προηγούμενης διάταξης, θα αποθηκεύσουμε αυτή τη διάταξη. Αυτή η διαδικασία επαναλαμβάνεται Τ φορές μέχρι να έχουμε ένα σύνολο ταξινομήσεων με βελτιωμένη πιθανότητα, οπότε και επιλέγουμε την ταξιθέτηση με την υψηλότερη πιθανότητα να ταξινομηθεί σωστά.

Στην συνέχεια, διερωτάνται αν μπορούν να δοθούν θεωρητικές εγγυησεις για την ποιότητα διατάξεων των υπεργράφων που βρίσκονται με αυτό το είδος τοπικής αναζήτησης. Είναι δύσκολο να ειπωθεί κάτι επίσημο για την τοπική αναζήτηση με τη χρήση του εκπαιδευμένου μοντέλου. Αλλά μπορεί να επιτευχθεί κάποια εικόνα της δύναμης της τοπικής αναζήτησης αποδεικνύοντας ένα ανάλογο αποτέλεσμα για την τοπική αναζήτηση για τη μεγιστοποίηση του μέσου μεγέθους τομής. Υποθέτουμε ότι όλες οι απλότητες είναι μη κενές και ότι κάθε στοιχείο εμφανίζεται σε τουλάχιστον δύο απλότητες. Για κάθε τοπικό βέλτιστο, το μέσο μέγεθος τομής είναι τουλάχιστον 1/(2c 2d) φορές μεγαλύτερο από αυτό της λοβικά βέλτιστης λύσης. Δεδομένου ότι κάθε simplex έχει μέγεθος ≤ c, η βέλτιστη λύση (με τάξη π) έχει μέσο μέγεθος τομής το πολύ. Πλέον χρεώνουμε τον δείκτη i ∉ A στον δείκτη h ∉ B, και κάθε ένας μπορεί να χρεωθεί c(d - 1) φορές, οπότε |A| + |B| = m - 1. Επομένως, η βέλτιστη λύση έχει μέσο μέγεθος τομής το πολύ 2c\r 2d φορές μεγαλύτερο από αυτό της τοπικά βέλτιστης λύσης.

Μεγάλος όγκος εργασιών έχουν γίνει στο παρελθόν αναφορικά με την εξέλιξη των παγκοσμίων δυαδικών γράφων. Η εργασία αυτή επιχειρεί να κατανοήσει τη διαχρονική εξέλιξη των τοπικών υπεργραφικών εγω-δικτύων, εξάγοντας χαρακτηριστικά εγω-δικτύων υψηλότερης τάξης για την πρόβλεψη της εξέλιξης και όχι του τομέα του συστήματος. Στην εργασία αυτή, δείχνουμε ότι τα δυαδικά δίκτυα τείνουν να επεκτείνονται ταχύτατα στην αρχή της ζωής τους, αντί να προσθέτουν πολλούς κόμβους σε σύντομο χρονικό διάστημα.

Συνοψίζοντας, καταλήγουμε σε κάποια πορίσματα. Η εργασία αυτή, εξετάζει τη μελέτη των υπεργραφημάτων, μιας δομής που μπορεί να χρησιμοποιηθεί για τη μοντελοποίηση αλληλεπιδράσεων ανώτερης τάξης που αφορούν έναν κόμβο και τους μεταβλητούς του. Διαπιστώνεται ότι τα δίκτυα alter διαθέτουν ισχυρή χρονική εντοπιότητα, καθώς τείνουν να καταλαμβάνουν πολύ καθορισμένα σύνολα χρονικών σφραγίδων στο δίκτυο ego-network στο οποίο ανήκουν. Αυτές οι έννοιες της τοπικότητας αποτελούν το θεμέλιο για θεμελιώδη συνδυαστικά ερωτήματα σχετικά με πολυέδρες υψηλών διαστάσεων. Τελικά, προτείνεται μια μέθοδος βαθιάς μάθησης για να μάθουμε αν ένα δεδομένο υπεργραφικό ego-δίκτυο είναι σωστά ταξινομημένο σε οποιαδήποτε δεδομένη χρονική στιγμή.

*Στη συνέχεια παραθέτουμε υλικό παράλληλο ή συμπληρωματικό στο θέμa.*

* **Hypergraph patterns and collaboration structure**

Jonas L. Juul ∗ , Austin R. Benson† , and Jon Kleinberg†

Με αφορμή το δεύτερο κείμενο παίρνουμε την έννοια των υπεργραφικών ego δικτύων και τη συνδέουμε με το επίκαιρο θέμα του Covid-19.

Η εργασία ξεκινά κάνοντας λόγο στο πως οι επιδόσεις των ομάδων στον εργασιακό χώρο και σε άλλους τομείς, όπως η πολιτική και ο αθλητισμός, φέρονται να επηρεάζονται από αυτές τις αλληλεπιδράσεις μεταξύ των μελών της ομάδας. Εδώ εισάγει μία νέα οικογένεια δομικών προτύπων τα m-patterns που τυποποιούν συνδέσμους μεταξύ συνεργατών, ερευνώντας τη συχνότητα τέτοιων δομών στα δεδομένα χρησιμοποιώντας ένα απλό μηδενικό μοντέλο τυχαίου υπεργράφου. Επιπλέον εξετάζονται τα ποσοστά εμφάνισης διαφόρων δομών συνεργασίας στο μηδενικό μας μοντέλο και αποδεικνύουμε πώς αυτά τα ποσοστά εμφάνισης ποικίλλουν ανάλογα με το μέγεθος και την πυκνότητα των υπεργράφων .

Τέλος διαπιστώνεται ότι οι δομές των επιστημονικών συνεργασιών σε έγγραφα του COVID-19 σε ορισμένες περιπτώσεις είναι στατιστικά σημαντικά διαφορετικές από εκείνες των μη-COVID-19 εργασιών. Εξετάζοντας τον αριθμό των αναφορών για 4 διαφορετικά επιστημονικά πεδία, βρίσκουμε επίσης ενδείξεις ότι οι επαναλαμβανόμενες συνεργασίες είναι πιο επιτυχημένες για επιστημονικές δημοσιεύσεις με 2 συγγραφείς και λιγότερο επιτυχείς για επιστημονικές δημοσιεύσεις με 3 συγγραφείς σε σύγκριση με άλλες δομές συνεργασίας.

Μεμονωμένοι κόμβοι μπορεί περιστασιακά να συνυπάρχουν σε μια ποικιλία ρυθμίσεων. Είναι ζωτικής σημασίας να ερευνηθεί ποιοι κόμβοι είναι πιο πιθανό να αναπτύξουν αυτές τις συνδέσεις.

Σε αντίθεση με τις προβλέψεις του μοντέλου, ορισμένα σύνολα δεδομένων περιείχαν διάφορα μοτίβα που εκπροσωπούνταν κατά πολλές τάξεις μεγέθους. Η επαναλαμβανόμενη συνεργασία φάνηκε να είναι πιο συχνή στο σύνολο δεδομένων προτυπωμάτων τόσο για έργα 2 όσο και για έργα 3 συγγραφέων. Αυτό είναι ενδιαφέρον, δεδομένου ότι, αν η συνεργασία εμφανιζόταν τυχαία, μια τέτοια ανακάλυψη θα συνέβαινε μόνο σε εξαιρετικά πυκνά δίκτυα.

* **Hypergraph Data analysis with PAOHVis**

Paolo Buono, Miguel Ceriani and Maria Costabile University of Bari, via Orabona, 4, 70125 Bari, Italy

Το εν λόγω κείμενο, παίρνει ως φράση κλειδί την ανάλυση δεδομένων υπεργράφων και την συνδέει με τον αθλητισμό και την οπτικοποίηση των δεδομένων. Οι δραστηριότητες ανάλυσης δεδομένων αξιοποιούν γράφους για τη μοντελοποίηση πολύπλοκων σχέσεων μεταξύ δεδομένων. Από την μια, οι γράφοι είναι σε θέση να περιγράφουν πολυποίκιλα και πολύπλοκα δεδομένα και σχέσεις του πραγματικού κόσμου.Απο την άλλη, οι υπεργράφοι είναι κατάλληλοι για τη μοντελοποίηση δικτύων επιχειρηματικών εταίρων ή δικτύων συν-συγγραφής. Ο PAOH (Parallel Aggregated Οrdered Hypergraph), που έχει αναπτυχθεί τα τελευταία χρόνια, έχει εφαρμοστεί σε διάφορους τομείς, όπως τα ego-δίκτυα και οι ψηφιακές ανθρωπιστικές επιστήμες. Ένα εργαλείο που τον υλοποιεί είναι διαθέσιμο στο διαδίκτυο και ο πηγαίος κώδικάς του είναι ανοικτός και συντηρείται σε δημόσιο αποθετήριο.

Από θέμα «ορισμού» η PAOH εμπνεύστηκε αρχικά από το Biofabric, μια τεχνική που αντιμετωπίζει την επεκτασιμότητα της απεικόνισης του υπεργράφου, αλλά δεν λαμβάνει υπόψη την εξέλιξη της τοπολογίας του δικτύου με την πάροδο του χρόνου. Η τεχνική PAOH χαρακτηρίζεται από παράλληλα hyperedges που απεικονίζονται κατά μήκος ενός άξονα χρόνου. Στα σχήματα παρουσιάζεται η οπτικοποίηση PAOH της ίδιας συν-συγγραφής. Κάθε κόμβος τοποθετείται κάθετα, με αλφαβητική σειρά του ονόματος του συγγραφέα. Σε άλλο σχήμα παρουσιάζεται ένα ζεύγος 𝐺 = (𝑉, 𝓻) που μοιράζεται τις ίδιες κορυφές, αλλά με τοπολογία που μεταβάλλεται με την πάροδο του χρόνου.

Η PAOH υποστηρίζει τη μελέτη της απόδοσης παικτών και ομάδων σε πολλαπλά πλαίσια διαφορετικής λεπτομέρειας. Το σύνολο δεδομένων που χρησιμοποιήθηκε στην παρούσα εφαρμογή περιέχει δεδομένα από τις πρώτες 26 εβδομάδες της σεζόν 2019-2020 του ιταλικού πρωταθλήματος Major (Serie A). Υπάρχουν 20 ομάδες στη Serie A, 10 αγώνες κάθε εβδομάδα (κάθε ομάδα παίζει μία φορά την εβδομάδα) κατά τη διάρκεια 26 εβδομάδων, για συνολικά 260 αγώνες. Το PAOH μπορεί να χρησιμοποιηθεί για την προβολή όλων των φάσεων ενός ποδοσφαιρικού αγώνα. Μια φάση αγώνα αποτελείται από πολλές πάσες μεταξύ παικτών της ίδιας ομάδας. Στο PAOHVis, όλοι οι αγώνες που παίζονται σε μια εβδομάδα ομαδοποιούνται σε ένα Time Slot, το οποίο είναι ένας υπεργράφος. Οι επιλογές color by και group by μπορούν να χρησιμοποιηθούν για την ομαδοποίηση των παικτών σύμφωνα με ένα συγκεκριμένο χαρακτηριστικό. Για παράδειγμα, οι κορυφές στην Εικόνα 3 είναι χρωματισμένες ανά ομάδα και ομαδοποιημένες ανά ομάδα. Διατίθενται διάφορα κριτήρια ταξινόμησης, όπως το βαρυκεντρικό (οι πιο συνδεδεμένοι παίκτες είναι κατακόρυφα πιο κοντά) και ο βαθμός (οι παίκτες διατάσσονται κατακόρυφα από τους πιο συνδεδεμένους προς τους λιγότερο συνδεδεμένους, όπως φαίνεται στην Εικόνα 2). Ο συνολικός αριθμός των παικτών που συμμετέχουν σε φάσεις που οδηγούν σε γκολ είναι υψηλότερος στην Ίντερ από ό,τι στη Γιουβέντους . Εντός κάθε χρονικού πλαισίου, οι ακμές του οριζόντιου χώρου μπορούν να ταξινομηθούν σύμφωνα με το μήκος και τους αλγόριθμους ορθογένεσης. Το πέρασμα με το ποντίκι πάνω από μια υπερκεντροειδή ακμή αποκαλύπτει τις σχέσεις της με άλλες κορυφές που έχουν φιλτραριστεί από το σύνολο δεδομένων. Οι ξεθωριασμένες κουκκίδες αποκαλύπτουν ότι οι ορατές κορυφές έχουν σχέσεις μεταξύ τους που φιλτράρονται.

Καταλήγουμε πως, η χρήση του PAOHVis για την ανάλυση δυναμικών υπεργραφημάτων είναι ένα από τα πιο δημοφιλή εργαλεία στην κοινότητα του InfoVis. Έχει υιοθετηθεί σε διάφορους τομείς, όπως ο τομέας του αθλητισμού που αναφέρεται στην παρούσα εργασία. Η πρόκληση τώρα είναι η εύρεση κατάλληλων οπτικοποιήσεων των αποτελεσμάτων και η κατάλληλη αλληλεπίδραση με τον χρήστη.

**3. Detecting Strong Ties Using Network Motifs**

Το θέμα που πραγματεύεται η εργασία όπως ομολογεί και ο τίτλος είναι *η ανίχνευση ισχυρών δεσμών μεταξύ χρηστών* στα κοινωνικά δίκτυα. Οι ισχυρές σχέσεις των χρηστών στις πλατφόρμες κοινωνικής δικτύωσης είναι ιδιαιτέρως σημαντικές μιας και βοηθούν στην εξατομικευμένη εμπειρία που θα έχει κάθε χρήστης στις πλατφόρμες. Τι όμως εξ αρχής ορίζεται σαν «ισχυρή» σχέση; Το ερώτημα αυτό η εργασία απαντά μέσω ενός δείγματος ακμών που παρέχεται, όπως αναφέρεται στο κείμενο, που θεωρούνται ως ισχυροί δεσμοί . Με τον όρο ακμές εννοούνται οι χρήστες που συχνά συμμετέχουν σε πολλαπλά επικαλυπτόμενα δίκτυα μέσω λειτουργιών όπως η παρακολούθηση και η ανταλλαγή μηνυμάτων.

Δεδομένου λοιπόν ότι υπάρχουν κάποιοι ισχυροί δεσμοί ακμών, ανακάμπτει και μία φυσική εργασία της ανίχνευσης των δεσμών του υπόλοιπου δικτύου.

Με την έρευνα να λαμβάνει τόπο στο Twitter, το έργο αυτό γίνεται πιο δύσκολο, λόγω της περιορισμένης διαθεσιμότητας δεδομένων χαρακτηριστικών σχέσεων και των ελάχιστων δικτύων σύνδεσης.

Μέσω πειραματικής έρευνας αποδεικνύεται πως για τον εντοπισμό ισχυρών δεσμών υψηλής ακρίβειας είναι αναγκαία η χρήση δομικών χαρακτηριστικών, που προκύπτουν από την παρουσία μοτίβων μικρού δικτύου συνδυασμένα με δυνατούς και αδύναμους δεσμούς. Η χρήση των μοτίβων πρέπει να είναι συγκεκριμένη για παραμένει αναλλοίωτο το αποτέλεσμα. Τα μοτίβα πιο συγκεκριμένα, πρέπει να είναι μεγάλα σε μέγεθος για να μετριάζονται προβλήματα αραιότητας, με συνέπεια να μην είναι όλα τα μοτίβα χρήσιμα.

Η πειραματική έρευνα ενισχύει τα ευρήματα της με θεωρητική ανάλυση, περισσότερο για την αναγκαιότητα μεγάλων μοτίβων και πως αυτά οδηγούν στην αυξημένη απόδοση των μοντέλων.

Σε πολλές πλατφόρμες κοινωνικής δικτύωσης (social media) οι χρήστες συμμετέχουν σε πολλά δίκτυα ταυτόχρονα τα οποία μπορεί να διαφέρουν ως προς την χρήση τους.

Σε αυτά τα περιβάλλοντα, συναντάμε ένα ευρύ φάσμα περιπτώσεων στις οποίες ένα δίκτυο με μεγάλο αριθμό συνδέσμων συνυπάρχει με ένα επικαλυπτόμενο, αλλά πολύ πιο αραιό δίκτυο. Όπως και στην περίπτωση του γραφήματος ακολούθων και του γραφήματος του τηλεφωνικού καταλόγου στο Twitter.

Πληροφορίες σημαντικές για τον χρήστη χρησιμοποιούνται για τη βελτίωση της συνάφειας και της εξατομίκευσης στις προτάσεις περιεχομένου, για τη δημιουργία πιο εξατομικευμένων ειδοποιήσεων και άλλων εφαρμογών.

Στο άρθρο γίνεται αναφορά στο πως οι υπάρχουσες τεχνικές για την ανίχνευση ισχυρών δεσμών δεν είναι εξίσου δυνατές και αποδοτικές σε σχέση με αυτές που αναπτύσσονται στο κείμενο.

Ο τρόπος προσέγγισης σχετίζεται με διάφορες γραμμές έρευνας που αφορούν τη δομή του δικτύου, ιδιαίτερα στον τομέα των κοινωνικών δικτύων και των δικτύων πληροφοριών.

Δύο διαγράμματα φαίνεται να βοηθούν στην έρευνα, πρώτο το πυκνό γράφημα GL που αποτελείται κυρίως από ασθενείς δεσμούς για να συμπληρώσουμε τους ισχυρούς δεσμούς σε ένα δεύτερο αραιό γράφημα GS, χρησιμοποιώντας τη συχνότητα των μικρών υπογραφών.

Για τον εντοπισμό ισχυρών δεσμών για χρήστες που δεν είναι γνωστό αν αναπτύσσουν τέτοιου είδους δεσμούς, χρησιμοποιούνται τα δύο γραφήματα, το πυκνό GL που περιέχει όλες τις ακμές που διατίθενται και το επικαλυπτόμενο αραιό GS το οποίο περιέχει τους ισχυρούς δεσμούς.

Βασικό συστατικό για οποιαδήποτε εργασία πρόβλεψης είναι ένα κριτήριο με το οποίο θα αξιολογούνται και θα συγκρίνονται οι προβλέψεις. Ένα τέτοιο κριτήριο θα πρέπει να είναι μία ετικέτα βασικής αλήθειας. Στην εργασία προτείνεται μία τέτοια ετικέτα.

Στην έρευνα επίσης, όπως έγινε γνωστό χρησιμοποιείται ένα σύνολο δεδομένων του Twitter, συγκεκριμένα συλλέγεται ένα πλήρες στιγμιότυπο ενός γραφήματος που ονομάζει mutual follow graph (γράφημα κοινών ακολούθων) του 2015.

Τελικά παραθέτει μία αρκετά γενική μεθοδολογία, υψηλής ακρίβειας, που μπορεί να χρησιμοποιηθεί σε πολλές άλλες περιπτώσεις. Μέσω των γράφων GL, GS, χρησιμοποιώντας τη συχνότητα των μικρών υπογράφων και υπερβαίνοντας τις δομές των τυπικών ως τώρα προσεγγίσεων, καταφέρεται η εξέταση γραφημάτων πολλών κόμβων, με την εργασία να φέρνει και νέα ερωτήματα για μελλοντικές έρευνες.

*Στη συνέχεια παραθέτουμε υλικό σχετικό με την ανίχνευση ισχυρών δεσμών στα κοινωνικά δίκτυα.*

* **Social Features of Online Networks: The Strength of Intermediary Ties in Online Social Media**

Przemyslaw A. Grabowicz,José J. Ramasco ,Esteban Moro,Josep M. Pujol,Victor M. Eguiluz

Published: January 11, 2012

<https://doi.org/10.1371/journal.pone.0029358>

Ένας αυξανόμενος αριθμός κοινωνικών επαφών πραγματοποιείται μέσω της χρήσης των διαδικτυακών μέσων κοινωνικής δικτύωσης. Παρόλα αυτά, η κοινωνική χρησιμότητα των διαδικτυακών αλληλεπιδράσεων εξακολουθεί να αποτελεί αντικείμενο συζήτησης. Και αυτό το άρθρο μέσω μελέτης στο Twitter ανακαλύπτει ότι είναι κατανεμημένο σε ομάδες.

Η πλειονότητα των διαδικτυακών κοινωνικών δικτύων περιλαμβάνει μια ποικιλία αλληλεπιδράσεων μεταξύ χρηστών για να ανταποκρίνεται στα διαφορετικά επίπεδα συμμετοχής ή έντασης των σχέσεων των χρηστών.

Στην παραπάνω έρευνα ανακαλύπτεται μέσω μεθόδων ομαδοποίησης, ομάδες χρηστών του Twitter, που η παρουσία των ομάδων συσχετίζεται προφανώς με τη δραστηριότητα στο δίκτυο όσον αφορά τα μηνύματα που είναι γνωστά ως αναφορές και retweets. Οι αναφορές, οι οποίες προορίζονται για πιο εξατομικευμένα μηνύματα, συχνά επικεντρώνονται σε συνδέσμους μεταξύ παρόμοιων ομάδων ή στο εσωτερικό μεμονωμένων ομάδων. Όσο περισσότερες είναι οι αναφορές που ανταλλάσσονται και αν επιστρέφονται, τόσο ισχυρότερη είναι η επίδραση. Τα retweets, τα οποία συνδέονται με γεγονότα διάδοσης πληροφοριών, είναι πιο πιθανό να βρεθούν σε συνδέσμους που συνδέουν ομάδες, ιδίως σε εκείνους που συνδέουν ομάδες με μικρή αλληλοεπικάλυψη. Το πιο κρίσιμο είναι ότι είναι πιο πιθανό να βρεθούν σε συνδέσμους που σχετίζονται με άτομα που ενεργούν ως μεσάζοντες μεταξύ ομάδων.

* **‘’How well do we know each other?”: detecting ties strength in multidimensional social networks**

Luca Pappalardo KDD Laboratory ISTI-CNR and University of Pisa Via G. Moruzzi, 1, 56124 Pisa - Italy Email: lpappalardo@isti.cnr.it Telephone: +39 050 315 2934 Fax: +39 050 315 2040 Giulio Rossetti KDD Laboratory ISTI-CNR and University of Pisa Via G. Moruzzi, 1, 56124 Pisa - Italy Email: giulio.rossetti@isti.cnr.it Telephone: +39 050 315 2934 Fax: +39 050 315 2040 Dino Pedreschi KDD Laboratory University of Pisa Largo B. Pontecorvo 3, 56127 Pisa Italy Email: pedre@di.unipi.it Telephone: +39 050 2212 752 Fax: ++39 050 2212 752

Οι πλατφόρμες των μέσων κοινωνικής δικτύωσης θεωρούν όλους τους χρήστες εξίσου, είτε ως φίλους είτε ως ξένους με ελάχιστα έως καθόλου ενδιάμεσα, και δεν λαμβάνουν υπόψη τη δύναμη των δεσμών κατά τη δημιουργία και τη διατήρηση σχέσεων. Σε αυτό το κείμενο παρουσιάζεται ένα ολοκαίνουργιο, πολύπλευρο ορισμό της δύναμης δεσμού που χρησιμοποιεί τις πολυάριθμες κοινωνικές συνδέσεις που μπορούν να έχουν δύο άνθρωποι στο διαδίκτυο. Η σημασία της δύναμης δεσμού στην ανάπτυξη και διατήρηση αυτών των δεσμών δεν λαμβάνεται υπόψη από τα μέσα κοινωνικής δικτύωσης.

Η ισχύς ενός δεσμού είναι ένας συνδυασμός του χρόνου, της συναισθηματικής έντασης, της οικειότητας και της αμοιβαίας εξυπηρέτησης. Ενώ οι αδύναμοι δεσμοί βασίζονται σε λίγα κοινά διαθέσιμα μέσα, οι ισχυροί δεσμοί διαφοροποιούν την επικοινωνία μέσω πολλών διαφορετικών καναλιών.

Επεκτείνοντας το αρχικό μας κείμενο, το άρθρο αυτό αναπτύσσει μια μετρική δύναμης δεσμών για πολυδιάστατα δίκτυα. Ο αριθμός των αλληλεπιδράσεων μεταξύ των εμπλεκόμενων ατόμων είναι ο μόνος παράγοντας που επηρεάζει το πόσο ισχυρός είναι ένας δεσμός, ενώ τέλος και εδώ παρουσιάζεται χώρος για νέα ερωτήματα και μελλοντική έρευνα που θα επικεντρωθεί στον τρόπο χρήσης των δεδομένων που προσφέρει η ισχύς δεσμών για την αντιμετώπιση ζητημάτων όπως η πρόβλεψη συνδέσμων και η ανακάλυψη κοινοτήτων.

* **Community detection in social networks using user frequent pattern mining**

Seyed Ahmad Moosavi1 · Mehrdad Jalali1 · Negin Misaghian2 · Shahaboddin Shamshirband3 · Mohammad Hossein Anisi3 Received: 28 April 2015 / Revised: 17 February 2016 / Accepted: 27 May 2016 © Springer-Verlag London 2016

Το παραπάνω άρθρο στο πρόβλημα εντοπισμού ομάδων με παρόμοιους ‘’κόμβους’’ έρχεται να προτείνει μία μέθοδο με δύο κύριες συνεισφορές αρχικά εντοπίζοντας στα κοινωνικά δίκτυα μερικές μικροσκοπικές κοινότητες ανθρώπων με παρόμοια ενδιαφέροντα και συμπεριφορές και στη συνέχεια οι κοινότητες αυτές επεκτείνονται μέσω των κοινωνικών συνδέσεων.

Οι δυνατότητες επικοινωνίας και αλληλεπίδρασης μεταξύ ανθρώπων που βρίσκονται σε γεωγραφική απόσταση είναι τεράστιες στο Διαδίκτυο. Η αρχιτεκτονική των κοινωνικών δικτύων χρησιμεύει ως αξιόπιστος δείκτης των προθέσεων των χρηστών. Σε αυτό το κομμάτι, προτείνεται μια τεχνική για την εύρεση ομάδων σε ιστότοπους κοινωνικής δικτύωσης. Σε τέτοιους ιστότοπους, η ανακάλυψη κοινοτήτων μπορεί να απλοποιήσει άλλες εργασίες και εφαρμογές κοινωνικού υπολογισμού στα περισσότερα προγράμματα. Αυτή η προσέγγιση δίνει μεγάλη έμφαση στη φιλία των χρηστών και στα προσωπικά δεδομένα των χρηστών για την εύρεση κοινοτήτων. Στα κοινωνικά δίκτυα υπάρχουν τεράστιες βάσεις δεδομένων χρηστών. Σε πολλές εφαρμογές, η ανακάλυψη άγνωστων μοτίβων, όπως και στο αρχικό μας υλικό, σε αυτά τα δίκτυα μπορεί να είναι χρήσιμη. Οι συστάσεις φίλων, η τμηματοποίηση πελατών, η εξαγωγή συνδέσμων, η επισήμανση κορυφών και η ανάλυση κοινωνικής επιρροής είναι μερικές μόνο από τις εφαρμογές στις οποίες μπορεί να χρησιμοποιηθεί η αναγνώριση κοινοτήτων. Η ανάλυση των δικτύων υποβοηθείται από τον εντοπισμό μελών με επιρροή σε επιστημονικά κοινωνικά δίκτυα, δίκτυα ηλεκτρονικού ταχυδρομείου, ομάδες συζητήσεων κ.λπ. Ο προσδιορισμός κοινοτήτων σε ιστότοπους κοινωνικής δικτύωσης είναι ο στόχος της παρούσας μελέτης. Το πρόσθετο πλούσιο περιεχόµενο αυτών των δικτύων µπορεί να χρησιµοποιηθεί για την αύξηση του επιπέδου των κοινοτήτων.

Ένα από τα μεγαλύτερα ζητήματα στα κοινωνικά δίκτυα είναι πλέον η ανακάλυψη της κοινότητας. Στην παρούσα μελέτη, παρουσιάστηκε μια νέα προσέγγιση που βασίζεται στις κοινωνικές σχέσεις και τις προσωπικές προτιμήσεις των χρηστών. Η προτεινόμενη προσέγγιση είναι μια υβριδική προσέγγιση που λαμβάνει υπόψη τόσο τη δομή του γράφου όσο και το περιεχόμενό του για τη δημιουργία μιας κοινότητας.

**4. Predicting Reciprocity in Social Networks**

Κεντρικό θέμα σε αυτό το άρθρο είναι το πρόβλημα *πρόβλεψης της αμοιβαιότητας* (reciprocity), μία έννοια πολυσυζητημένη σε πολλές σύγχρονες εργασίες αυτού του επιστημονικού πεδίου.

Πως εξ αρχής όμως προκύπτει ένα τέτοιο πρόβλημα, αρχικά της αμοιβαιότητας στην επικοινωνία των χρηστών και έπειτα της πρόβλεψης αυτής ;

Θα έλεγε κανείς πως η αμφίδρομη επικοινωνία μεταξύ χρηστών στα κοινωνικά δίκτυα προκύπτει φυσικά, επόμενο λοιπόν είναι, με βάση χαρακτηριστικά που συλλέγονται από τους χρήστες, να επιχειρείται η προσέγγιση τρόπων επίτευξης της αμοιβαιότητας. Το άρθρο κάνει λόγο στον αναγκαίο προσδιορισμό καλών δεικτών αμοιβαιότητας, καθώς και την βοήθεια δέντρων απόφασης και μοντέλων παλινδρόμησης.

Την καθημερινή εξέλιξη της τεχνολογίας ακολουθούν βέβαια και οι εφαρμογές των social media που αποκτούν όλο και ένα περισσότερα χαρακτηριστικά και λειτουργίες. Κατά συνέπεια, περισσότερα μέσα επικοινωνίας των χρηστών.

Στην εργασία γίνεται χρήση του Twitter σαν κεντρικό παράδειγμα κοινωνικού δικτύου, στο οποίο οι χρήστες κοινοποιούν συνδέσμους (link) για σύνδεση με άλλους χρήστες και έναρξη συνομιλίας τους. Τέτοιου είδους διασυνδέσεις σε μεγάλες πλατφόρμες όπως το Twitter αποτελούν μέλη ενός δικτύου επικοινωνίας μεγαλύτερου βεληνεκούς.

Η αναγκαιότητα πρόβλεψης ενός link αμοιβαιότητας είναι μεγάλη και κύριος σκοπός του είναι περισσότερο η ταξινόμηση διαφορετικών προτιμήσεων και ενδιαφερόντων του χρήστη.

Όλα καταλήγουν στον στόχο συμμετρικών και αμφίδρομου ενδιαφέροντος σχέσεων μεταξύ χρηστών.

Στην πορεία της εργασίας νέοι στόχοι φέρνουν και νέες ερωτήσεις και την διαλεύκανση μας πάνω στο θέμα.

Ερωτήματα που πραγματεύονται στο αρχικό μας υλικό :

Τι καθιστά μία σχέση αμφίδρομη ;

Τα υποδίκτυα έχουν διαφορετική κατασκευαστική δομή στις αμφίδρομες και στις μονόδρομες σχέσεις ;

Εντοπίζουμε δύο κύριες προσεγγίσεις για την ανάλυση της αμοιβαιότητας. Στην πρώτη είναι η μελέτη του προβλήματος της πρόβλεψης της αμοιβαιότητας και τον προσανατολισμό μας σε ποια κατηγορία ανήκει μία σχέση δύο χρηστών, ανταποδοτική ή μη. Σε αυτή την προσέγγιση λαμβάνονται υπόψη τα δοθέντα χαρακτηριστικά των δύο. Η ανάλυση μας, εξάγει χαρακτηριστικά που υπόκεινται συχνά στην έννοια του status. Πιο επεξηγηματικά, άνθρωποι με κοινό status έχουν πιο συχνά αμοιβαία επικοινωνία. Προχωρώντας στην δεύτερη προσέγγιση, έχουμε την σύγκριση της δομής δύο υπογράφων, το ένα να αποτελείται μόνο από αμοιβαίους συνδέσμους και το άλλο να αποτελείται μόνο από μη αμοιβαίους συνδέσμους. Οι δύο δομές εμφανίζουν πολλές διαφορές όπως αυτή της ομαδοποιήσης.

Σε εργασίες με παρόμοιο υλικό η αμοιβαιότητα δεν θεωρείται παράγοντας στο θέμα της ετερογένειας των σχέσεων στα social media. Το πρόβλημα πρόβλεψης της αμοιβαιότητας σχετίζεται με άλλες εργασίες πρόβλεψης που ασχολούνται με συνδέσμους υποκείμενων δικτύων, πράγμα που διαφέρει από πολλές απόψεις. Μέσω της πρόβλεψης επιχειρείται να εντοπιστούν σύνδεσμοι που προς το παρόν τους λείπει ένα στιγμιότυπο δικτύου αλλά μπορεί να σχηματιστεί δυνητικά.

Βασική αντίθεση των δύο προβλέψεων είναι ότι ο σχηματισμός οποιουδήποτε συγκεκριμένου συνδέσμου είναι σπάνιος, ενώ η αμοιβαιότητα σε ένα ήδη υπάρχον link είναι συχνή.

Συνεχίζει με τις μεθόδους πρόβλεψης. Χαρακτηριστικά που καταγράφουν αν δύο χρήστες (στο άρθρο συχνά εμφανίζονται ως v και w) έχουν το ίδιο status ή τον ίδιο κοινωνικό κύκλο μελλοντικά θα βρεθεί χρήσιμο για την πρόβλεψη της αμοιβαιότητας. Αφιερώνεται μεγάλο κομμάτι στα διάφορα χαρακτηριστικά που χρησιμεύουν στην πρόβλεψη αμοιβαιότητας στα δίκτυα. Κάθε χαρακτηριστικό αντιστοιχεί σε μία μέθοδο που της εκχωρείται μία τιμή val(v,w).

Για κάθε χαρακτηριστικό εξετάζουμε την αξία του και εάν το εν λόγω άκρο είναι αμοιβαίο, χρησιμοποιούνται αυτά τα δεδομένα για μοντέλα εκπαίδευσης για την πρόβλεψη αμοιβαιότητας.

Τελικά η εργασία έχοντας διατυπώσει διάφορες παραλλαγές του προβλήματος της πρόβλεψης αμοιβαιότητας σε ένα διαδικτυακό κοινωνικό δίκτυο και έχοντας επιδείξει ικανότητα πρόβλεψης χρησιμοποιώντας κόμβων και οι γειτονιές των τοπικών δικτύων τους. Δεν παραλείπεται να μελετηθεί πως διαφέρουν τα αμοιβαία μοντέλα και μη ως προς το μέγεθος και την ομαδοιποιήση. Περαιτέρω ενδιαφέρουσες κατευθύνσεις που θα μπορούσε να πάρει το πρόβλημα είναι κάποια προσέγγιση που θα μπορούσε να χρησιμοποιήσει τα δομικά χαρακτηριστικά του συστήματος ή το γεγονός ότι οι χρονικές πληροφορίες είναι σημαντική πηγή δεδομένων για την ανάλυση της αμοιβαιότητας και την κατανόηση του τρόπου με τον οποίο αναπτύσσονται αμφίδρομες και μη σχέσεις. Η αμοιβαιότητα των συνδέσμων είναι τελικά, μία πτυχή του ευρύτερου ζητήματος της ετερογένειας στους τύπους των αλληλεπιδράσεων.

*Στη συνέχεια παραθέτουμε υλικό παράλληλο ή συμπληρωματικό στο θέμα της πρόβλεψης της αμοιβαιότητας.*

* **Predicting online extremism, content adopters, and interaction reciprocity**

Emilio Ferrara∗ , Wen-Qiang Wang∗ , Onur Varol† , Alessandro Flammini† and Aram Galstyan∗ ∗ Information Sciences Institute, University of Southern California, Marina del Rey, CA 90292, USA †School of Informatics and Computing, Indiana University, Bloomington, IN 47401, USA emiliofe@usc.edu, wenqianw@usc.edu, ovarol@indiana.edu, aflammin@indiana.edu, galstyan@isi.edu

Η παραπάνω εργασία παρουσιάζει ένα πλαίσιο μηχανικής μάθησης που αξιοποιεί ένα μοναδικό σύνολο δεδομένων, για τρεις εργασίες πρόβλεψης: αυτή του εντοπισμού εξτρεμιστών χρηστών, αυτή της εκτίμησης αν τακτικοί χρήστες χρησιμοποιούν εξτρεμιστικό περιεχόμενο και τέλος, το κομμάτι που μας ενδιαφέρει, αυτό της αμοιβαιότητας μεταξύ των χρηστών με εξτρεμιστές.

Συγκεκριμένα με την έρευνα να λαμβάνει τόπο στο Twitter, εκμεταλλευόμενη tweets και πιστοποιημένους χρήστες, η έρευνα εκτελεί τρεις εργασίες πρόβλεψης, μέσα σε αυτές και αυτή της αμοιβαιότητας. Η κάθε εργασία ρυθμίστηκε σε δύο σενάρια, ένα ανεξάρτητο του χρόνου σε προσομοιωμένη εργασία και ένα σε real-time προσομοίωση, με την ρύθμιση αυτή να αποδεικνύεται τελικά πολύ ελπιδοφόρα μιας και παρέχονται διάφορα σενάρια με καλή προγνωστική ισχύ.

Σημείο αναφοράς του άρθρου αποτελούν ζητήματα που σχετίζονται με τον διαδικτυακό εξτρεμισμό, την τρομοκρατική προπαγάνδα και τις εκστρατείες ριζοσπαστικοποίησης, με τα social media να έχουν κεντρικό ρόλο σε αυτές τις προσπάθειες. Από τις τρεις εργασίες μου αναφέρονται στο άρθρο, αυτή που ασχολείται με την αμοιβαιότητα είναι και η πιο απαιτητική από τις τρεις. Ουσιαστικά ελέγχει αν είναι σε θέση να προβλέψει εάν ένας τακτικός χρήστης που έρχεται σε επαφή με έναν εξτρεμιστικό λογαριασμό θα απαντήσει ή όχι.

Το άρθρο αυτό έρχεται και συμπληρώνει στο αρχικό μας υλικό με την ενσωμάτωση της χρονικής διάστασης, σαν μέσω καλύτερης απόδοσης.

Καταλήγει έπειτα από την εξέταση μεγάλου τυχαίου δείγματος στην παρατήρηση ότι τα χαρακτηριστικά των υπόλοιπων εργασιών είναι ίδια με αυτά της αμοιβαιότητας. Μέσα του crowd-sourcing εξάγονται εκατομμύρια tweets που αποδεικνύουν την αναγκαιότητα της πρόβλεψης αμοιβαιότητας.

* **Reciprocal versus Parasocial Relationships in Online Social Networks**

Author: Neil Zhenqiang Gong et al

Publication: Social Network Analysis and Mining

Publisher: Springer Nature

Date: Apr 11, 2014

Copyright © 2014, Springer-Verlag Wien

Σε αυτό το άρθρο κύριο ερευνητικό πρόβλημα του είναι η διερεύνηση των παραγόντων που επηρεάζουν τα μη αμοιβαία άκρα να γίνουν αμοιβαία και η πρόβλεψη εάν ένα μη αμοιβαίο πλεονέκτημα θα μετατραπεί σε αμοιβαίο. Γίνεται χρήση ενός συνόλου δεδομένων μεγάλης κλίμακας καθώς και ενός δημόσιου συνόλου δεδομένων κοινωνικού δικτύου.

Ξεκινά με την σύγκριση των δύο δομών, αμοιβαίων και μη, όπου τελικά φαίνεται ότι η αμοιβαιότητα φέρνει κοντά χρήστες με ίδιο degree, ενώ το μη αμοιβαίο φαίνεται να φέρνει κοντά χρήστες διαφορετικών βαθμών. Έπειτα γίνεται παρατήρηση πως η συμπεριφορά των χρηστών μαζί με χαρακτηριστικά κόμβων και ακμής, παίζουν καθοριστικό ρόλο στον σχηματισμό τελικά της αμοιβαιότητας.

Πραγματοποιούνται εκτενείς αξιολογήσεις με τα δύο σύνολα δεδομένων και αποδεικνύεται ότι η πρότασή ξεπερνά τις προηγούμενες προσεγγίσεις αμοιβαίας πρόβλεψης.

Λαμβάνοντας υπόψη μερικά στιγμιότυπα από κοινωνικά δίκτυα προβλέπετε πως τα μη αμοιβαία άκρα μελλοντικά θα γίνουν όλα αμοιβαία.

Σε αντίθεση με άλλες μελέτες που αντιμετωπίζουν την αμοιβαία πρόβλεψη ακμών είτε ως ένα εποπτευόμενο είτε ως ημιεποπτευόμενο πρόβλημα μάθησης, σε αυτή γίνεται γνωστό πως η πρόβλεψη αμοιβαίων ακμών μοντελοποιείται καλύτερα ως πρόβλημα ανίχνευσης ακραίων σημείων.

* **Reciprocity in evolving social networks**

Tackseung Jun · Rajiv Sethi

Published online: 30 August 2008

© Springer-Verlag 2008

Στην εργασία αυτή εξετάζεται η δυναμική της αμοιβαιότητας σε μικρών παγκόσμια δίκτυα, όταν τόσο οι στρατηγικές όσο και το ίδιο το δίκτυο εξελίσσονται.

H εξέλιξη των στρατηγικών στα κοινωνικά δίκτυα με τη μικρή περιουσία του κόσμου, όταν η αλληλεπίδραση παίρνει τη μορφή ενός διλήμματος κρατουμένων πολλών ατόμων. Η προσαρμογή της στρατηγικής γίνεται ως απόκριση στις διαφορικές αποδόσεις στη γειτονιά κάποιου, όπως είναι τυπικό στα εξελικτικά παιχνίδια, και υπόκειται σε περιστασιακή τυχαία μετάλλαξη. Το ίδιο το δίκτυο αλλάζει με την πάροδο του χρόνου καθώς οι σύνδεσμοι δημιουργούνται τυχαία

και σπασμένο. Διαπιστώνουμε ότι η μακροχρόνια συχνότητα συνεργασίας είναι υψηλότερη υπό τις ακόλουθες συνθήκες: η ακτίνα αλληλεπίδρασης δεν είναι ούτε πολύ μικρή ούτε πολύ μεγάλη, το δίκτυο είναι πιο ομαδοποιημένο και η μέση απόσταση μεταξύ των ατόμων είναι μικρότερη, ο ρυθμός μετάλλαξης της σκληρότητας είναι μικρότερος και ο ρυθμός προσαρμογής της σκληρότητας δεν είναι ούτε πολύ αργός ούτε πολύ γρήγορος.

Οι σύνδεσμοι με την σειρά που αναφέρθηκαν :

* *The Giant k-core of a Random Graph with a Specified Degree Sequence*

<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=0940870241d4801fe1999ddbdc4ed141e76f0266>

* *Random graph models of social networks*

<https://www.pnas.org/doi/epdf/10.1073/pnas.012582999>

* *A guide to null models for animal social network analysis*

<https://besjournals.onlinelibrary.wiley.com/doi/epdf/10.1111/2041-210X.12772>

* *Hypergraph patterns and collaboration structure* [*https://arxiv.org/pdf/2210.02163.pdf?fbclid=IwAR28RizlAifZueakc54C3KqZPEvMFYArrVr0cJWPila-6a\_cXVA9sYF1hCg*](https://arxiv.org/pdf/2210.02163.pdf?fbclid=IwAR28RizlAifZueakc54C3KqZPEvMFYArrVr0cJWPila-6a_cXVA9sYF1hCg)
* *Hypergraph Data analysis with PAOHVis*

[*https://ceur-ws.org/Vol-2994/paper15.pdf*](https://ceur-ws.org/Vol-2994/paper15.pdf)

* *Social Features of Online Networks: The Strength of Intermediary Ties in Online Social Media*

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0029358>

* *‘’How well do we know each other?”: detecting ties strength in multidimensional social networks*

*https://ieeexplore.ieee.org/abstract/document/6425622*

* *Community detection in social networks using user frequent pattern mining*

<https://123project.ir/wp-content/uploads/2018/07/294676_moosavi2016.pdf>

* *Predicting online extremism, content adopters, and interaction reciprocity*

<https://arxiv.org/pdf/1605.00659.pdf>

* *Reciprocal versus Parasocial Relationships in Online Social Networks*

<https://arxiv.org/pdf/1302.6309.pdf>

* *Reciprocity in evolving social networks*

<http://www.columbia.edu/cu/tract/projects/research-sources-for-contem/reciprocity_in_evolving.pdf>