



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ**  
**ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ**  
**ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ**

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

**«ΜΑΘΗΜΑΤΙΚΗ ΠΡΟΤΥΠΟΠΟΙΗΣΗ σε ΣΥΓΧΡΟΝΕΣ ΤΕΧΝΟΛΟΓΙΕΣ  
και την ΟΙΚΟΝΟΜΙΑ»**

*«Μαθηματική ανάλυση πολύπλοκων δικτύων επιστημόνων  
με βάση κοινές τους δημοσιεύσεις»*

Δημήτριος Νικολάου  
(Α.Μ. 09316027)

Δρ. Βασίλειος Κωνσταντούδης  
Ερευνητής ΕΚΕΦΕ «ΔΗΜΟΚΡΙΤΟΣ»  
(Επιβλέπων)

ΑΘΗΝΑ, Ιούλιος 2019

## ΕΥΧΑΡΙΣΤΙΕΣ

Πρωτίστως θα ήθελα να ευχαριστήσω τον Δρ. Βασίλειο Κωνσταντούδη, Ερευνητή του Εθνικού Κέντρου Έρευνας Φυσικών Επιστημών (ΕΚΕΦΕ) «Δημόκριτος», για την τιμή που μου έκανε να επιβλέψει την προσπάθεια και μέσω της καθοδήγησης και κρίσμης συμβολής του να καταστεί εφικτή η ολοκλήρωση της παρούσας εργασίας.

Επίσης, τον Δρ. Ιωάννη Λαζάρου, Ερευνητή του ιδίου ερευνητικού κέντρου, για τη βοήθειά του κατά την άντληση των δεδομένων που έτυχαν εν συνεχεία της κατάλληλης επεξεργασίας στο πλαίσιο εκπόνησης της παρούσας εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου για τη συνεχή στήριξη.

## ΠΕΡΙΛΗΨΗ

Τα τελευταία έτη έχει παρατηρηθεί ιδιαίτερο ερευνητικό ενδιαφέρον στη μελέτη δικτύων (γραφημάτων) που προκύπτουν από διάφορες κοινωνικές, τεχνολογικές, επιστημονικές κ.λπ. δραστηριότητες.

Μία από τις εφαρμογές της ανάλυσης δικτύων αποτελεί η μελέτη της επιστημονικής συνεργασίας και γενικότερα της επιστημονικής κοινότητας, καθώς η εξέλιξη της επιστήμης σε σημαντικό βαθμό έχει καταστήσει εδώ και αρκετά χρόνια αναγκαία τη μελέτη αυτής. Μέσω της ανάλυσης των στοιχείων αυτών δύναται να παραχθούν χρήσιμα συμπεράσματα τα οποία με τη σειρά τους θα μπορούσαν να συμβάλλουν σε περαιτέρω επιστημονική και τεχνολογική εξέλιξη.

Υπό το πρίσμα των ανωτέρω, αντικείμενο μελέτης της παρούσας εργασίας αποτέλεσε ο βαθμός συνεργασίας μεταξύ επιστημόνων και τα ιδιαίτερα χαρακτηριστικά αυτής, εκφραζόμενης μέσω της συν-συγγραφής (*co-authorship*), δηλαδή των κοινών δημοσιεύσεων, άρθρων, μελετών και επιστημονικών κειμένων δύο τουλάχιστον ερευνητών.

Προς τη κατεύθυνση αυτή, αξιοποιήθηκε η θεωρία γράφων (*graph theory*) και η μέθοδος ανάλυσης δικτύου (*network analysis*).

Ως εκ τούτου, μελετήθηκε η «αρχιτεκτονική» ενός δικτύου ερευνητών ενός εκ των πλέον διακριμένων Ινστιτούτων της Χώρας.

Συναφώς μελετήθηκαν οι σχέσεις και οι αλληλεπιδράσεις τόσο σε επίπεδο κόμβων ή όσο και σε επίπεδο υποομάδων/κοινοτήτων που ενυπάρχουν σε κάθε δίκτυο. Επίσης, εξετάσθηκε η εξέλιξη του δικτύου στο χρόνο (*network evolution*).

**Λέξεις κλειδιά:** ανάλυση δικτύου, πολύπλοκα δίκτυα, κοινωνικά δίκτυα, ανάλυση δεδομένων, θεωρία γράφων, συν-συγγραφή, επιστήμη της επιστήμης, πανεπιστημιακά δεδομένα

## ABSTRACT

In recent years, there has been particular research interest in the study of networks (graphs) resulting from various social, technological, scientific and other activities.

One of the applications of network analysis is the study of scientific cooperation and of the scientific community in general, as the evolution of science has made it's study necessary for many years now. By analyzing these elements, useful conclusions can be drawn which in turn could contribute to further scientific and technological developments.

In the light of the aforementioned, the subject of the present study was the degree of collaboration among scientists and its particular characteristics, as expressed by co-authorship. Co-authorship was studied by looking at joint publications, articles, studies and scientific papers of at least two researchers.

To this end, graph theory and network analysis were used.

Therefore, the "architecture" of a network of researchers of one of the most distinguished Institutes of the country was studied.

In addition, the relationships and interactions at both the node and the subgroup/community level, which are inherent in any network, were studied. Finally, the evolution of the network over time was examined.

**Keywords:** *network analysis, complex networks, social networks, data analysis, graph theory, co-authorship, science of science, scholarly data*

## ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

<b>ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ</b> .....	7
<b>1. ΕΙΣΑΓΩΓΗ</b> .....	8
1.1. Υπόβαθρο μελέτης.....	8
1.2. Δίκτυα & ανάλυση δικτύων .....	8
1.3. Αντικείμενο εργασίας .....	9
<b>2. ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ ΔΙΚΤΥΩΝ</b> .....	11
2.1. Ορισμός & δομικά στοιχεία δικτύων .....	11
2.2. Κατηγορίες δικτύων .....	12
2.3. Ποσοστικά χαρακτηριστικά δικτύων .....	13
2.4. Μέτρα κεντρικότητας (centrality measurements).....	17
2.5. Εύρεση κοινοτήτων .....	19
2.6. Μοντέλα δικτύων .....	21
2.6.1. Τυχαία δίκτυα ( <i>Erdos – Renyi model, ER</i> ).....	22
2.6.2. <i>Small World Network (Watts-Strogatz Model, SW)</i> .....	22
2.6.3. Τα τυχαία <i>scale-free Network (Barabasi-Albert, BA)</i> .....	24
2.6.4. Συγκριτική επισκόπηση κυρίαρχων μοντέλων δικτύων .....	25
<b>3. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ</b> .....	28
<b>4. ΔΕΔΟΜΕΝΑ</b> .....	33
4.1. Δεδομένα & Παραδοχές .....	33
4.2. Συλλογή δεδομένων.....	33
4.3. Βασικά χαρακτηριστικά & Ταξινόμηση δεδομένων .....	34
<b>5. ΜΕΘΟΔΟΛΟΓΙΑ – ΑΝΑΛΥΣΗ – ΑΠΟΤΕΛΕΣΜΑΤΑ</b> .....	36
5.1. Γενικά στοιχεία.....	36
5.2. Η εικόνα του ινστιτούτου – εξέλιξη (σύγκριση περιόδων).....	38
5.2.1. Κατανομή βάρθρωσης, <i>Degree distribution</i> .....	40
5.2.2. Μέτρα κεντρικότητας ( <i>centrality measurements</i> ).....	43
5.2.3. Συντελεστής Συσσωμάτωσης ( <i>Clustering coefficient</i> ) .....	49
5.2.4. Διαμέριση δικτύου – Εύρεση κοινοτήτων ( <i>modularity</i> ).....	50
5.3. Συγκριτική επισκόπηση των πρώην ινστιτούτων, κατά τις 2 περιόδους .....	54
5.3.1. Κύρια σύγκριση.....	54
5.3.2. Μία άλλη προσέγγιση (σύγκριση “intra” – “inter”) .....	55

5.4. Συγκριτική επισκόπηση των πρώην ινστιτούτων ως αυτοτελή δίκτυα, κατά τη β' περίοδο .....	57
5.4.1. Κατασκευή νέων δικτύων & παραδοχές .....	57
5.4.2. Σύγκριση των δικτύων $N_1, N_2, N_3 - N_{total}$ .....	58
5.4.3. Επιμέρους σύγκριση των δικτύων $N_1, N_2, N_3$ .....	59
<b>6. ΕΠΙΛΟΓΟΣ</b> .....	<b>60</b>
6.1. Σύνοψη & Συμπεράσματα .....	60
6.2. Ενδεχόμενες επεκτάσεις .....	62
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ</b> .....	<b>63</b>

## ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

ΕΛΛΗΝΟΓΛΩΣΣΗ	
ΕΚΕΦΕ	Εθνικό Κέντρο Έρευνας Φυσικών Επιστημών
INN	Ινστιτούτου Νανοεπιστήμης & Νανοτεχνολογίας
ΞΕΝΟΓΛΩΣΣΗ	
BA	<i>Barabasi-Albert</i>
ER	<i>Erdos-Renyi</i>
IMEL	<i>Institute of Microelectronics</i>
IMS	<i>Institute of Materials Science</i>
IPC	<i>Institute of Physical Chemistry</i>
SOS	<i>Science of Science</i>
WS	<i>Watts-Strogatz</i>

## 1. ΕΙΣΑΓΩΓΗ

### 1.1. Υπόβαθρο μελέτης

Η εξέλιξη της επιστήμης σε σημαντικό βαθμό έχει καταστήσει εδώ και αρκετά χρόνια αναγκαία τη μελέτη αυτής. Ουσιαστικά πρόκειται για την «Επιστήμη της Επιστήμης» (*“science of science” – SOS*), έναν ταχέως αναπτυσσόμενο τομέα που αποσκοπεί στην κατανόηση, στην ποσοτικοποίηση και στην πρόβλεψη της επιστημονικής έρευνας και των παραγόμενων αποτελεσμάτων. Πρόκειται για οριζόντια διαδικασία καθώς αφορά όλους τους επιστημονικούς κλάδους, με ποικίλη θεματολογία όπως η επίδραση των επιστημονικών δημοσιεύσεων, των ερευνητών, των περιοδικών και των πανεπιστημίων, η μοντελοποίηση της επιστημονικής συνεργασίας, η πρόβλεψη της μελλοντικής εξέλιξης της επιστήμης κ.λπ.

Ένα από τα αντικείμενα μελέτης της «Επιστήμης της Επιστήμης» (*“science of science” – SOS*) αποτελεί η συνέργεια που αναπτύσσεται μεταξύ των επιστημόνων. Η διεπιστημονική φύση της επιστήμης, η πολυπλοκότητα και το κόστος είναι ορισμένοι από τους λόγους που ωθούν τους ερευνητές στη συνεργασία, η οποία με τη σειρά της δύναται να αυξήσει την ακαδημαϊκή παραγωγικότητα, την ποιότητα της εργασίας και την επιστημονική ανάπτυξη.

Νέα ευρήματα δείχνουν μια τάση αυξανόμενης συνεργασίας στην έρευνα, που αποτελεί ένα από τα κριτήρια για την αξιολόγηση της ποιότητας της εργασίας των ερευνητών και των ερευνητικών ομάδων, και κατ' επέκταση ένας αποτελεσματικός τρόπος για την πρόοδο της επιστήμης και τεχνολογίας. Κατά συνέπεια, η συνεργασία μεταξύ των επιστημόνων φαίνεται να συσχετίζεται θετικά τόσο με την υψηλότερη ποιότητα της εργασίας τους όσο και με το υψηλότερο επίπεδο επιστημονικής ανάπτυξης.

Η επιστημονική συνεργασία λαμβάνει διάφορες εκφάνσεις, όπως μία απλή ανταλλαγή απόψεων μεταξύ 2 τουλάχιστον ερευνητών έως την από κοινού έρευνα σε ένα εργαστήριο και την συν-συγγραφή άρθρων, μελετών κ.λπ.

Η συν-συγγραφή (*co-authorship*) αποτελεί χαρακτηριστική περίπτωση και ένας από τους σημαντικότερους δείκτες επιστημονικής συνεργασίας. Ως συν-συγγραφή θα μπορούσε να θεωρηθεί η συνεργασία δύο ή περισσότερων ερευνητών για την εκπόνηση μίας επιστημονικής μελέτης.

Από την άλλη, η επιστημονική συνεργασία και η συν-συγγραφή συνιστούν ένα κοινωνιολογικό φαινόμενο, καθώς πέραν του δικτύου ανταλλαγής γνώσεων, συμπαραγωγής νεάς γνώσης και επιστήμης, ανακλύπει ένα (κοινωνικό) δίκτυο μεταξύ των ερευνητών. Η δομή των εν λόγω δικτύων αναδεικνύει κάποια ιδιαίτερα χαρακτηριστικά της επιστημονικής/ερευνητικής κοινότητας, τα οποία φαίνεται να ακολουθούν τους νόμους των φυσικών δικτύων.

### 1.2. Δίκτυα & ανάλυση δικτύων

Ένα δίκτυο αποτελείται από ένα πεπερασμένο σύνολο διακεκριμένων οντοτήτων (*entities*) που συνδέονται με μια σχέση, συνιστώντας ένα σύστημα από πομπούς -



δέκτες που επικοινωνούν μεταξύ τους. Γνωστά δίκτυα στον φυσικό κόσμο είναι το δίκτυο των υπολογιστών, όπου ανταλλάσσουν πληροφορία με την μορφή *bits*, το δίκτυο των φίλων που επικοινωνούν μεταξύ τους με διάφορους τρόπους, το δίκτυο των ηθοποιών, των επιστημόνων, το δίκτυο μεταφορών, το δίκτυο διανομής ηλεκτρικής ενέργειας, νερού, κλπ.

Περαιτέρω, η ανάλυση δικτύων συνδέεται με τη θεωρία γράφων και αποτελεί μία μέθοδος για την εξέταση της αλληλεπίδρασης μεταξύ των οντοτήτων που δημιουργούν το δίκτυο. Σημειώνεται δε ότι, ένα δίκτυο έχει ευρύτερη έννοια και σημασία πέραν των οντοτήτων και των σχέσεων μεταξύ αυτών δεδομένου ότι έχει δομή και λειτουργία.

Εν προκειμένω το δίκτυο συν-συγγραφέων συνιστά ένα τύπο κοινωνικού δικτύου, στο οποίο οι συγγραφείς είναι οι πρωταγωνιστές του, και οι μεταξύ τους δεσμοί είναι οι κοινές τους δημοσιεύσεις.

### 1.3. Αντικείμενο εργασίας

Λαμβάνοντας υπόψη τόσο τη σπουδαιότητα της ανάπτυξης συνεργειών εντός της ακαδημαϊκής κοινότητας όσο και τη χρησιμότητα της ανάλυσης δικτύων, η παρούσα εργασία σκοπεί στην εξέταση του φαινομένου της συν-συγγραφής (*co-authorship*)<sup>1</sup> - υπό το πρίσμα της ανάλυσης δικτύων- εντός του Ινστιτούτου Νανοεπιστήμης και Νανοτεχνολογίας (εφεξής INN) του Εθνικού Κέντρου Έρευνας Φυσικών Επιστημών (ΕΚΕΦΕ) «Δημόκριτος».

Το χρονικό διάστημα που εξετάσθηκε αφορά στη 10ετία 2008-2017, το οποίο επιπρόσθετα χωρίστηκε σε 2 χρονικές περιόδους, ώστε να εξετασθούν τυχόν μεταβολές στο δίκτυο. Ομοίως, χωρίστηκε και το δίκτυο σε επιμέρους ομάδες, στη βάση της προηγούμενης διοικητικής δομής του Ινστιτούτου<sup>2</sup>, ούτως ώστε να εξετασθεί η αλληλεπίδρασή τους. Σημειώνεται ότι δεν συμπεριλαμβάνονται οι συνδημοσιεύσεις ερευνητών του ως άνω Ινστιτούτου με ερευνητές άλλων ινστιτούτων, ερευνητικών κέντρων, πανεπιστημίων κ.λπ., που σαφώς υφίστανται.

Η ανάλυση πραγματοποιείται σε 3 επίπεδα: α) κόμβου, β) υποομάδας/κοινότητας και γ) δικτύου υπό 2 οπτικές γωνίες. Υπολογίζονται επιλεγμένα από τη θεωρία μέτρα ανάλυσης που χαρακτηρίζουν τον κάθε κόμβο, αλλά και το δίκτυο ή τις υποομάδες/κοινότητες που ενυπάρχουν, ενώ παράλληλα οπτικοποιούνται σε γραφήματα. Επίσης εξετάζονται διάφορα στοιχεία ως προς τη δομή του δικτύου, όπως η συνοχή και η πυκνότητά του.

Με τον τρόπο αυτό επιχειρείται να αποτυπωθεί η εικόνα η δομή του δικτύου, τα βασικά χαρακτηριστικά αυτής. Επίσης εξετάζονται συγκριτικά τυχόν μεταβολές που επήλθαν τόσο με την πάροδο των ετών όσο ενδεχομενως και μέσω διοικητικών

---

<sup>1</sup> Ως μία από τις σημαντικότερες εκφάνσεις/μορφές επιστημονικής συνεργασίας.

<sup>2</sup> Ειδικότερα των Ινστιτούτων Επιστήμης Υλικών, Μικροηλεκτρονικής και Φυσικοχημείας (Ν. 4310/2014).

μεταβολών. Επίσης, το υπό μελέτη δίκτυο αντιπαραβάλλεται με τα κυρίαρχα μοντέλα δικτύων και εντοπίζονται ομοιότητες, συγκλίσεις και διαφορές.

Ολοκληρώνοντας το κεφάλαιο 1 (εισαγωγή), το περιεχόμενο των επόμενων κεφαλαίων έχει περιληπτικά ως εξής:

- Στο κεφάλαιο 2 αναπτύσσεται το θεωρητικό υπόβαθρο των δικτύων και η μέθοδος της ανάλυσης δικτύων, η οποία εφαρμόστηκε στα δεδομένα της εργασίας. Συγκεκριμένα, περιγράφονται τα δομικά στοιχεία των δικτύων, τα ποσοτικά χαρακτηριστικά και οι κατηγορίες αυτών, καθώς και τα κυριότερα μοντέλα δικτύων.
- Στο κεφάλαιο 3 παρατίθενται οι σημαντικότερες έρευνες και μελέτες αναφορικά με κοινές δημοσιεύσεις ερευνητών (συν-συγγραφή, *co-authorship*), με χρήση της μεθόδου ανάλυσης δικτύων, καθώς και τα βασικά ευρήματα αυτών.
- Στο κεφάλαιο 4 αναφέρεται ο τρόπος διαχείρισης των στοιχείων που αναλύθηκαν στη συνέχεια (μεθοδολογικές παραδοχές-προσδιορισμός των δεδομένων, καθώς επίσης συλλογή και ταξινόμηση αυτών).
- Στο κεφάλαιο 5 εφαρμόζεται η μέθοδος ανάλυσης δικτύων στα δεδομένα της εργασίας, εξάγονται ποσοτικά στοιχεία περιγραφής του δικτύου και επιχειρείται, κατά περίπτωση σύμφωνα με τις τιθέμενες θεματικές και επίπεδα ανάλυσης, μια προσπάθεια ερμηνείας των αποτελεσμάτων, υπό το πρίσμα των συναφών μελετών (κεφ. 3) και γενικότερα της βιβλιογραφίας περί δικτύων.
- Στο κεφάλαιο 6 αναφέρονται συνοπτικά και συγκεντρωτικά τα αποτελέσματα της ανάλυσης και πιθανές/ενδεχόμενες επεκτάσεις.

Τέλος παρατίθεται η σχετική βιβλιογραφία.

## 2. ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ ΔΙΚΤΥΩΝ

Στο κεφάλαιο αυτό παρατίθενται τα δομικά στοιχεία ενός δικτύου, πιθανές κατηγοριοποιήσεις βάσει διαφόρων κριτηρίων, καθώς και τα ειδικότερα ποσοτικά χαρακτηριστικά του (μετρικές κ.λπ.). Επιπρόσθετα περιγράφονται τα 3 κυριότερα μοντέλα δικτύων με επιμέρους συγκρίσεις.

### 2.1. Ορισμός & δομικά στοιχεία δικτύων

Ένας απλός ορισμός των δικτύων θα μπορούσε να είναι ο εξής: «Μια θεωρητική κατασκευή που περιγράφει τις σχέσεις μεταξύ ομοειδών ή ομάδων ομοειδών στοιχείων αναπαριστώντας τις με σχηματικό τρόπο».

Βασική έννοια στα δίκτυα είναι αυτή των «σχέσεων». Απομονωμένα στοιχεία δεν συνιστούν δίκτυο αλλά ένα απλό σύνολο. Αυτό που ορίζει ένα δίκτυο είναι το πλέγμα των σχέσεων ή των «συνδέσεων» μεταξύ τους. Τα δίκτυα είναι στην ουσία γραφικές αναπαραστάσεις πολύπλοκων συστημάτων που βοηθούν στη καλύτερη κατανόηση τους. Από τις συγκοινωνίες, τις τηλεπικοινωνίες και τις κοινωνικές σχέσεις, ως τις οικονομικές συναλλαγές και το διαδίκτυο (εξ ορισμού), δίκτυα ενυπάρχουν παντού.

Από μαθηματική άποψη, τα δίκτυα είναι γράφοι (*graphs*) στους οποίους αναπαρίστανται οι σχέσεις μεταξύ των στοιχείων ενός συνόλου.

Ειδικότερα, ως δίκτυο ορίζεται το σύστημα  $G$  που αποτελείται από ένα σύνολο  $N$  κόμβων  $n_i, i=1, \dots, N$ , που συνδέονται μεταξύ τους με συνδέσμους  $c_{ij}$ :  $G=\{n_i, c_{ij} \mid i, j=1, \dots, N$ .

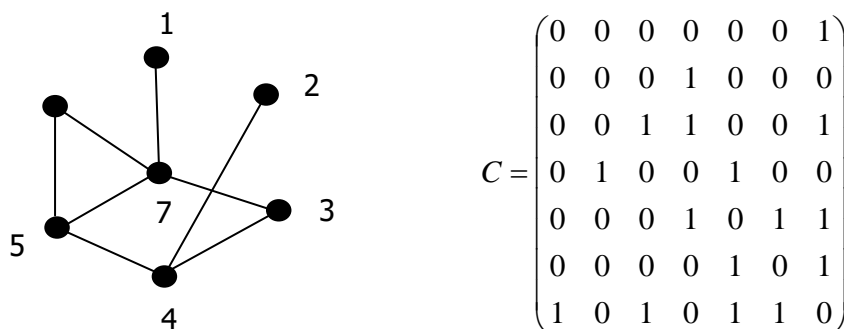
Τα βασικά χαρακτηριστικά στοιχεία του δικτύου είναι τα ακόλουθα:

- **Κόμβοι (*Nodes/vertices/points*)  $i=1, \dots, N$**  : Βασικά δομικά στοιχεία του δικτύου.
- **Σύνδεσμοι (*Edges/Links/lines*)** : εκφράζουν την σύνδεση μεταξύ 2 κόμβων και ειδικότερα:
  - ✓  $c_{ij}=1$  όταν οι κόμβοι  $i$  και  $j$  συνδέονται (μία φορά) και
  - ✓  $c_{ij}=0$  όταν οι κόμβοι  $i$  και  $j$  δεν συνδέονται.
- **Βαθμός/μέγεθος δικτύου  $N^3$**  : ορίζεται ως ο συνολικός αριθμός κόμβων που απαρτίζουν το δίκτυο.
- **Ισχύς Συνδέσμου,  $c_{ij}$**  : Εκφράζει την ύπαρξη και σχετική ισχύ του συνδέσμου ανάμεσα στους κόμβους  $i$  και  $j$ .

Για την ανάλυση των ποσοτικών χαρακτηριστικών ενός δικτύου αυτό απαιτείται να λάβει μία από τις δύο τυπικές του μορφές, που είναι πίνακας ή λίστα συνδεσιμότητας/γεινιάσης. Ειδικότερα:

<sup>3</sup> Ένα άλλο μέτρο του μεγέθους του δικτύου είναι η διάμετρος, η οποία λαμβάνει υπόψη όχι μόνο τον αριθμό των κόμβων αλλά και το βαθμό διασύνδεσής τους (βλ. παρακάτω).

- Ως προς τον **πίνακα συνδεσιμότητας ή γειτνίασης  $C_{ij}$  (*adjacency matrix*)** : Είναι ένας  $N \times N$  πίνακας του οποίου τα στοιχεία είναι τα βάρη των συνδέσμων  $c_{ij}$ .



Εικόνα 1: Απεικόνιση ενός δικτύου και του πίνακα συνδεσιμότητάς του  
(πηγή: Constantoudis)

- Ως προς την **λίστα συνδεσιμότητας ή γειτνίασης (*adjacency list*)**: Σε αυτή την περίπτωση, οι σχέσεις μεταξύ των κόμβων (τα βάρη των συνδέσμων,  $c_{ij}$ ) παρατίθενται με τη μορφή ζευγών κόμβων σε γειτονικές στήλες, το μήκος των οποίων ισούται με το άθροισμα των συνδέσμων και των μη συνδεδεμένων κόμβων.<sup>4</sup>

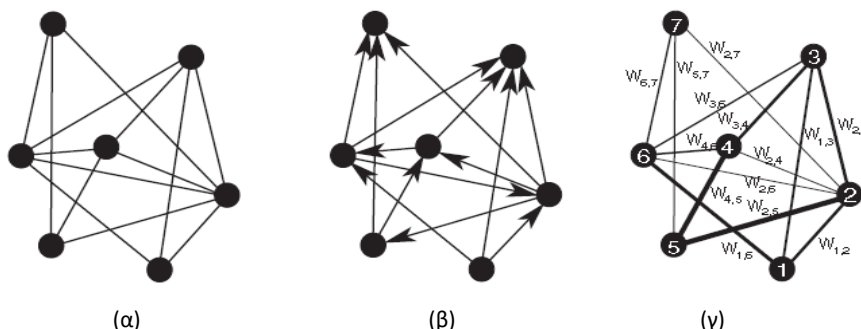
## 2.2. Κατηγορίες δικτύων

Τα δίκτυα δύναται να κατηγοριοποιηθούν, βάσει διαφόρων κριτηρίων. Δύο εκ των σημαντικότερων ταξινομήσεων είναι οι παρακάτω:

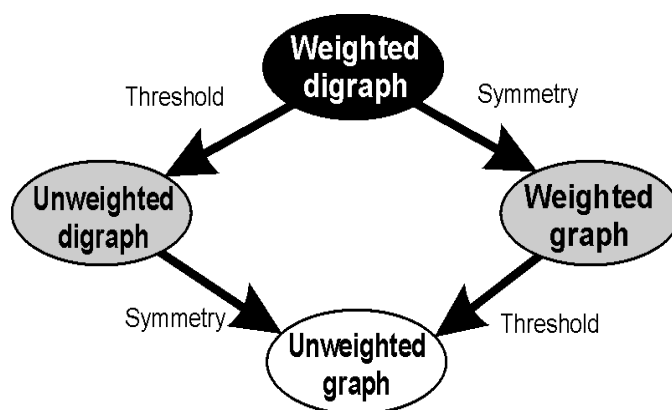
- Βάσει του χαρακτήρα των συνδέσμων (δυναδικός ή συνεχής) :**
  - ✓ **Σταθμισμένα (*weighted*)**: Δίκτυα στα οποία οι σύνδεσμοι  $c_{ij}$  μπορούν να πάρουν τιμές στο διάστημα  $[0,1]$  ανάλογα με το ποσοστό “φορτίου” που μεταφέρει ο σύνδεσμος στο δίκτυο. Στα σταθμισμένα δίκτυα τα στοιχεία του πίνακα συνδεσιμότητας παίρνουν μόνο πραγματικές τιμές στο διάστημα  $[0,1]$ .
  - ✓ **Αστάθμιστα (*non-weighted*)**: Δίκτυα στα οποία οι τιμές των συνδέσμων είναι δυαδικές, δηλαδή, είναι 1 εάν υπάρχει σύνδεσμος και μηδέν εάν δεν υπάρχει. Στα αστάθμιστα δίκτυα τα στοιχεία του πίνακα συνδεσιμότητας παίρνουν μόνο τις τιμές 0 ή 1.
- Βάσει της κατευθυντικότητας των συνδέσμων :**
  - ✓ **Κατευθυνόμενα (*directed*)**: Δίκτυα στα οποία ισχύει  $c_{ij} \neq c_{ji}$ . Στα κατευθυνόμενα δίκτυα ο πίνακας συνδεσιμότητας είναι ασύμμετρος.
  - ✓ **Μη κατευθυνόμενα (*undirected*)**: Δίκτυα στα οποία ισχύει  $c_{ij} = c_{ji}$ . Στα μη κατευθυνόμενα δίκτυα ο πίνακας συνδεσιμότητας είναι συμμετρικός γύρω από τη διαγώνιο.

<sup>4</sup> Οι λίστες γειτνίασης δεν χρησιμοποιούνται συνήθως για «πυκνά» δίκτυα όπου ο αριθμός των συνδέσμων ξεπερνά το τετράγωνο του αριθμού των κόμβων. Στην περίπτωση αυτή προτιμώνται οι πίνακες συνδεσιμότητας ή γειτνίασης.

**Σημείωση:** Μία πρόσθετη κατηγορία αφορά στα **μεικτά (mixed)** δίκτυα, όπου ορισμένες σχέσεις των κόμβων είναι προσανατολισμένες (βλ. «κατευθυνόμενα») και άλλες όχι (βλ. «μη κατευθυνόμενα»).



Εικόνα 2: Απεικόνιση 3 δικτύων, ως εξής: α) μη κατευθυνόμενο αστάθμιστο, β) κατευθυνόμενο αστάθμιστο και γ) μη κατευθυνόμενο σταθμισμένο (πηγή: Boccaletti, 2006)



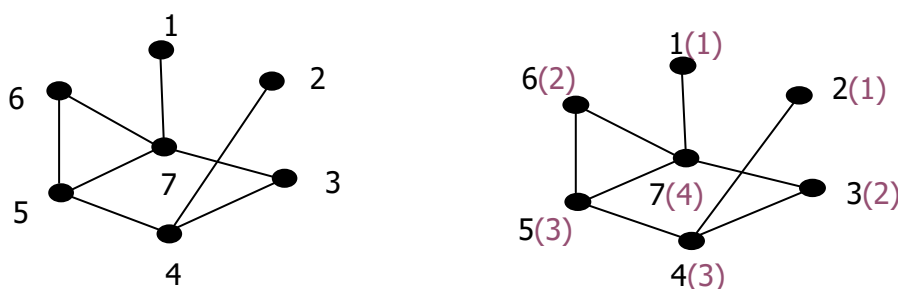
Εικόνα 3: Οι 4 κύριοι τύποι δικτύων και οι μετασχηματισμοί τους (όλοι οι τύποι δικτύων μπορούν να εξαχθούν από το σταθμισμένο διάγραμμα μέσω κατάλληλων μετασχηματισμών) (πηγή: Costa, 2008)

Περαιτέρω, τα δίκτυα δύναται να κατηγοριοποιηθούν βάσει του πλήθους των τύπων των κόμβων του δικτύου (π.χ. δίκτυα με έναν-κοινό τύπο κόμβων – *monopartite*, δίκτυα με δύο τύπους κόμβων – *bipartite*, δίκτυα με περισσότερους/πολλαπλούς τύπους κόμβων – *multilayer*) ή βάσει του πλήθους των συνδέσμων των κόμβων (μία ή περισσότερες συνδέσεις – *multiplex*).

### 2.3. Ποσοτικά χαρακτηριστικά δικτύων

**Βαθμός Κόμβου, Degree:  $k_i$**  : Συνολική ισχύς συνδέσμων που φτάνουν (ή φεύγουν) από τον κόμβο  $i$ .

$$k_i = \sum_j c_{ij} \equiv \sum_{j=1}^N c_{ij}$$



Εικόνα 4: Απεικόνιση ενός δικτύου και των βαθμών αυτού, ανά κόμβο (πηγή: Constantoudis)

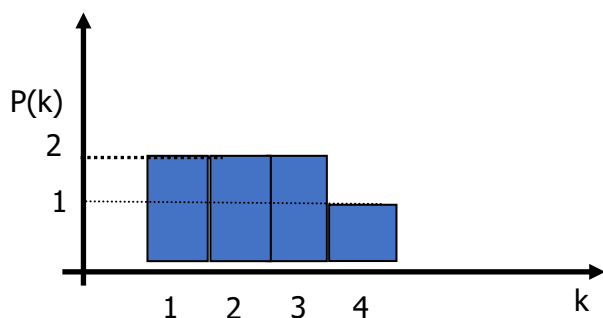
Σημειώνεται δε, ότι σε περίπτωση ?? κατευθυνόμενου δικτύου, για κάθε κόμβο έχουμε τον βαθμό του σε σχέση με τις εισερχόμενες και τις εξερχόμενες συνδέσεις, ως εξής:

$$k_i^{in} = \sum_j c_{ij} \qquad k_i^{out} = \sum_j c_{ji}$$

ενώ ο συνολικός βαθμός του κόμβου  $i$  είναι το άθροισμα των προαναφερομένων.

**Κατανομή βάθμωσης, Degree distribution,  $P(k)$ :** Είναι μία ιδιότητα που χαρακτηρίζει συνολικά το δίκτυο και εκφράζει την πιθανότητα να βρούμε έναν κόμβο που να έχει βαθμό μεταξύ  $(k, k+\delta k)$ .

$$P(k) = \frac{1}{N} \sum_{i=1}^N \delta_{k,k_i}$$



Εικόνα 5: Διαγραμματική απεικόνιση κατανομής βάθμωσης (πηγή: Constantoudis)

**Πυκνότητα δικτύου (Density):** Πέρα από το συνολικό μέγεθος ενός δικτύου, όπως αναφέρθηκε παραπάνω, ιδιαίτερη σημασία έχει και ο βαθμός στον οποίο τα στοιχεία του δικτύου είναι διασυνδεδεμένα μεταξύ τους. Ένα δίκτυο μεγάλου μεγέθους (μεγάλου αριθμού κόμβων) μπορεί να έχει ελάχιστες συνδέσεις, μεταξύ των στοιχείων του, την ίδια στιγμή που ένα πολύ μικρότερο μπορεί να είναι πολύ πυκνότερο.

Την έννοια αυτή της πυκνότητας του δικτύου την ορίζουμε ποσοτικά ως το πηλίκο των πραγματικών συνδέσεων που υπάρχουν στο δίκτυο διά του αριθμού των συνδέσεων που θα είχε ένα πλήρως διασυνδεδεμένο (*fully-connected*) δίκτυο ίσου μεγέθους. Ένα πλήρως διασυνδεδεμένο δίκτυο είναι ένα δίκτυο στο οποίο κάθε στοιχείο συνδέεται άμεσα με καθένα από τα υπόλοιπα.

Το σύνολο τιμών που μπορεί να πάρει η πυκνότητα είναι από 0 μέχρι 1. Στην περίπτωση που όλοι οι κόμβοι συνδέονται άμεσα μεταξύ τους, τότε η πυκνότητα θα ισούται με τη μονάδα. Γενικά, όσο μεγαλύτερο είναι το δίκτυο, τόσο η πυκνότητα θα μειώνεται, αφού οι εν δυνάμει δεσμοί θα αυξάνονται.

**Απόσταση κόμβων, distance/shortest path length:** Ορίζεται ως ο ελάχιστος αριθμός συνδέσεων για να φτάσουμε από τον κόμβο  $i$  στον κόμβο  $j$ . Πρόκειται δηλαδή για το συντομότερο μονοπάτι μεταξύ δύο κόμβων.

$$d_{ij} = \min_{\ell \leq N} \ell \cdot \theta[(c^\ell)_{ij} - \frac{1}{2}]$$

Σε μικρά δίκτυα η συντομότερη διαδρομή είναι εύκολο να υπολογιστεί εποπτικά. Σε μεγαλύτερα, ωστόσο, δίκτυα χρειάζεται κανείς να καταφύγει σε πιο συστηματικές μεθόδους για τον υπολογισμό της.

Μια αποτελεσματική τέτοια μέθοδος είναι η Αναζήτηση μέσω Πλάτους (*Breadth-first search*) (Zhou & Hansen, 2006), η οποία υπολογίζει την απόσταση μεταξύ δύο κόμβων A και B, ξεκινώντας από τον κόμβο A και προχωρώντας μέσω διαδοχικών ελέγχων των άμεσων γειτόνων ως το B.

Το μήκος του διανύσματος της συντομότερης διαδρομής μας δίνει ένα βαθμό της εγγύτητας μεταξύ των δύο στοιχείων στο δίκτυο. Μια μικρού μήκους συντομότερη διαδρομή είναι ενδεικτική μιας στενότερης σχέσης μεταξύ των κόμβων.

Σημειώνεται ότι, εάν το δίκτυο είναι σταθμισμένο, λαμβάνεται υπ' όψιν και το βάρος των συνδέσεων ανάμεσα στους κόμβους που συνδέουν τους  $i$  και  $j$ .

Περαιτέρω, ορίζεται και η **μέση (“ελάχιστη”) απόσταση κόμβων (average path length)** ως:

$$\bar{\ell} = \frac{\text{sum over all possible distances}}{\text{total nr. of possible links}} = \frac{\sum_{i \neq j} d_{ij}}{N(N-1)}$$

**Διάμετρος, Diameter, d:** Η διάμετρος είναι ένα ακόμα μέτρο του μεγέθους ενός δικτύου, που σχετίζεται όμως όχι μόνο με το πλήθος των κόμβων, αλλά λαμβάνει υπόψη του και τη διασυνδεσιμότητα.

Ορίζεται ως η μέγιστη απόσταση ανάμεσα στους κόμβους του δικτύου. Εάν το δίκτυο αποτελείται από ασύνδετα υποδίκτυα, τότε ορίζεται ως η μέγιστη απόσταση μεταξύ των κόμβων του μεγαλύτερου υποδικτύου.<sup>5</sup>

$$d = \max_{(i,j)} d_{ij}$$

**Συντελεστής Συσσωμάτωσης, Clustering coefficient,  $C_i$**  : Δείχνει πόσο ισχυρή είναι η συνδεσιμότητα γύρω από τον κόμβο  $i$ .

Σχετίζεται με το βαθμό στον οποίο οι κόμβοι του δικτύου τείνουν να σχηματίζουν τοπικά υποδίκτυα. Σχηματικά μπορεί κανείς να συλλάβει την έννοια του συντελεστή συσσωμάτωσης για τον κόμβο  $i$  ως τον αριθμό των τριγωνικών σχέσεων που αναπτύσσονται μεταξύ του  $i$  και των κόμβων με τους οποίους ο  $i$  συνδέεται άμεσα.

Επίσης, σχετίζεται άμεσα με την έννοια της κλίκας (*clique*), η οποία περιγράφει ένα δίκτυο (ή ένα τμήμα δικτύου) του οποίου τα στοιχεία είναι πλήρως και άμεσα διασυνδεδεμένα μεταξύ τους.

Ποσοτικά υπολογίζεται με πολλούς διαφορετικούς τρόπους, οι οποίοι δίνουν αρκετά διαφορετικά αποτελέσματα μεταξύ τους, εκφράζοντας ωστόσο όλοι τους τις τοπικές συσπειρώσεις στο δίκτυο. Η επιλογή του συντελεστή συσσωμάτωσης εξαρτάται κάθε φορά από την εφαρμογή.

1<sup>ος</sup>) Έστω ότι ο κόμβος  $i$  έχει  $E_i$  συνδέσεις με μη-μηδενική ισχύ. Έστω επίσης ότι ανήκει σε γειτονιά η οποία περιέχει  $n_i$  κόμβους με τους οποίους ο κόμβος  $i$  έχει τη δυνατότητα να συνδεθεί, και τους ονομάζουμε “γείτονες” του κόμβου  $i$ . Να σημειωθεί ότι μπορεί ο κόμβος  $i$  να έχει την δυνατότητα να συνδέεται με όλους τους υπόλοιπους κόμβους του δικτύου και τότε  $n_i = N$ . Τότε ο δυνατός αριθμός συνδέσεων μέσα στη γειτονιά του  $i$  είναι  $[n_i(n_i-1)/2]$ , και ο συντελεστής συσσωμάτωσης γύρω από τον κόμβο  $i$  ορίζεται ως:

$$C_i = \frac{\text{αριθμός υπαρκτών συνδέσεων}}{\text{αριθμός δυνατών συνδέσεων σε όλο την γειτονιά}} = \frac{2 E_i}{n_i(n_i-1)}$$

2<sup>ος</sup>) Υπολογίζεται ο λόγος του αριθμού των συνδέσεων  $E_i$  διαιρεμένος με τον αριθμό των δυνατών συνδέσεων που μπορεί να έχει ο κόμβος  $i$ . Στην πιο γενική περίπτωση που ο κόμβος  $i$  μπορεί να συνδέεται με όλους τους άλλους κόμβους ( $N-1$ ) του συστήματος, τότε ο συντελεστής συσσωμάτωσης είναι:

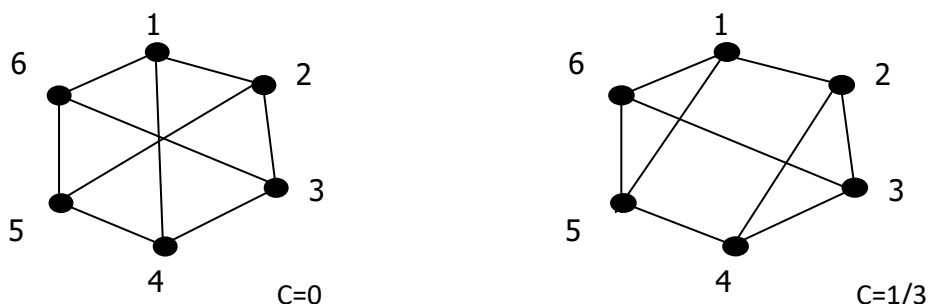
$$C_i = \frac{\text{αριθμός υπαρκτών συνδέσεων}}{\text{αριθμός δυνατών συνδέσεων του } i \text{ με όλους τους υπόλοιπους κόμβους}} = \frac{E_i}{N-1}$$

<sup>5</sup> Είναι έτσι το ανάλογο της διαμέτρου ενός γεωμετρικού σχήματος και αποτελεί μια καλή προσέγγιση για το βαθμό «διασποράς» των κόμβων ενός δικτύου. Η ερμηνεία της ωστόσο, χρειάζεται προσοχή καθώς σε περιπτώσεις ακραίων, πολύ απομακρυσμένων κόμβων η τιμή της μπορεί να μην είναι αντιπροσωπευτική.



3<sup>ος</sup>) Βασίζεται στον υπολογισμό των τριαδικών συνδέσεων (εκφράζει το ποσοστό ύπαρξης τριαδικών σχέσεων γύρω από έναν κόμβο). Εάν  $\xi_{ij}$  είναι η ισχύς σύνδεσης των κόμβων  $i$  και  $j$ , τότε

$$C_i = \frac{\sum_k \xi_{ij} \xi_{jk} \xi_{ki}}{\sum_k \xi_{ij} \xi_{ki}}$$



Εικόνα 6: Απεικόνιση δικτύων με διαφορετικό συντελεστή συσσωμάτωσης  
(πηγή: Constantoudis)

Σημειώνεται ότι ο συντελεστής συσσωμάτωσης συγκεκριμένου κόμβου εξαρτάται από το βαθμό του. Η σχέση με την οποία τα δύο μεγέθη συσχετίζονται είναι πολύ σημαντική για το είδος του δικτύου.

**Μέσος Συντελεστής Συσσωμάτωσης, *Average clustering coefficient*:** Πρόκειται για μια γενικευμένη έκφραση του  $C$  (συντελεστής συσσωμάτωσης) που αναφέρθηκε παραπάνω και αντιστοιχεί στη μέση τιμή των  $C$  για όλους τους κόμβους  $N$  ενός δικτύου.

$$\bar{C} = \frac{1}{N} \sum_i C_i$$

Συνιστά ένα μέτρο της συνδεσιμότητας (*connectedness*) ενός δικτύου. Ένα δίκτυο στο οποίο ο μέσος συντελεστής συσσωμάτωσης είναι 0, αποτελεί ένα σύνολο πλήρως ασύνδετων κόμβων. Από την άλλη εάν είναι 1, το δίκτυο είναι πλήρως συνδεδεμένο με όλους τους κόμβους να συνδέονται μεταξύ τους με όλους τους δυνατούς τρόπους, να συνθέτουν δηλαδή μια κλίκα.

#### 2.4. Μέτρα κεντρικότητας (*centrality measurements*)

Η σημαντικότητα ενός κόμβου δεν καθορίζεται μονοσήμαντα, αλλά εξαρτάται από την ερμηνεία που δίνεται στην έννοια κεντρικότητα (*centrality*). Τα μέτρα της κεντρικότητας αναγνωρίζουν του προεξέχοντες κόμβους, δηλαδή αυτούς που συσχετίζονται πιο ενεργά με τα υπόλοιπα μέλη του δικτύου.

Τα πιο σημαντικά μέτρα της κεντρικότητας είναι τα ακόλουθα:

- Βαθμική Κεντρικότητα, *Degree Centrality*
- Κεντρικότητα εγγυτήτας, *Closeness Centrality*
- Διαμεσότητα, *Betweenness Centrality*
- Ποιότητα συνδέσεων, *Eigenvector centrality*

Ακολουθεί η απαιτούμενη για τις ανάγκες της παρούσας εργασίας περιγραφή των 3 πρώτων μετρικών, οι οποίες θα αξιοποιηθούν σχετικώς στα δεδομένα της παρούσας εργασίας.

#### **α) Βαθμική Κεντρικότητα, *Degree Centrality***

Η βαθμική κεντρικότητα (*degree centrality*) ενός κόμβου  $v$  είναι ανάλογη με το βαθμό του κόμβου αυτού. Γενικά, η βαθμική κεντρικότητα ορίζεται ως εξής:

$$D = \frac{\text{deg}(v)}{n-1}$$

#### **β) Κεντρικότητα εγγυτήτας, *Closeness Centrality***

Σε συνδεδεμένα γραφήματα υπάρχει μία μετρική φυσικής απόστασης ανάμεσα σε όλα τα ζευγάρια κόμβων που ορίζεται από το μήκος των ελαχίστων μονοπατιών του. Το πόσο μακριά βρίσκεται ένας κόμβος  $s$  ενός γράφου ορίζεται σαν το άθροισμα των αποστάσεων από όλους τους άλλους κόμβους.

Ως εκ τούτου, υπάρχει η δυνατότητα να οριστεί το κατά πόσο κοντά είναι ένας κόμβος σαν το αντίστροφο του προαναφερθέντος μέτρου. Επομένως, όσο πιο κεντρικός είναι ένας κόμβος βάσει του δείκτη αυτού, τόσο μικρότερη είναι η συνολική του απόσταση από όλους τους άλλους κόμβους. Προκύπτει, δηλαδή, ότι αν  $d_{ij}$  η απόσταση της κορυφής  $i$  από την κορυφή  $j$ , τότε:

$$l_i = \frac{1}{n-1} \sum_{i \neq j} d_{ij}$$

όπου  $l_i$  είναι η μέση απόσταση η οποία είναι μικρή για τις κεντρικές κορυφές και μεγάλη για τις απομακρυσμένες κορυφές.

Αν δεν υπάρχει μονοπάτι μεταξύ δύο κορυφών τότε αυθαίρετα τίθεται ως απόσταση το πλήθος όλων των κορυφών. Η αντίστροφη ποσότητα  $C_i = 1/l_i$  είναι η κεντρικότητα εγγυτήτας.

Το εν λόγω μέτρο εγγυτήτας είναι μεν δηλωτικό της θέσης του κόμβου στο δίκτυο από άποψη αποστάσεων, αλλά δεν αποδίδει τις διαφορές που μπορεί να υπάρχουν μεταξύ ενός κόμβου, που απλά βρίσκεται κοντά στο κέντρο του δικτύου από γεωμετρική άποψη, και των κόμβων που είναι και κεντρικοί από πλευράς διασυνδεσιμότητας. Ένα μέτρο που επιτρέπει να ποσοτικοποιηθεί αυτή τη έννοια είναι η διακεντρικότητα (*betweenness centrality*).

**γ) Διαμεσότητα, *Betweenness Centrality***

Το μέτρο *Betweenness Centrality* είναι ένας δείκτης που αναφέρεται σε έναν κόμβο του γράφου και είναι αυτός που μετρά πόσες φορές ο κόμβος λειτουργεί σαν «γέφυρα» του ελάχιστου μονοπατιού που ενώνει 2 κόμβους του γράφου. Κόμβοι που έχουν μεγάλη πιθανότητα να βρεθούν ανάμεσα σε ένα ελαχίστου μήκους μονοπάτι μεταξύ δύο τυχαίων επιλεγόμενων κόμβων έχουν υψηλό *betweenness centrality*.

Για τον υπολογισμό του *betweenness centrality* ενός κόμβου  $u$  του γράφου  $G=(V,E)$  μπορεί να χρησιμοποιηθεί ο παρακάτω αλγόριθμος:

- Για κάθε ζεύγος κόμβων  $(s,t)$  υπολόγισε τα ελαχίστου μήκους μονοπάτια.
- Για κάθε ζεύγος κόμβων  $(s,t)$  μέτρα πόσα ελαχίστου μήκους μονοπάτια διέρχονται από τον κόμβο  $u$ .
- Βρες το πλήθος των κόμβων για τους οποίους ισχύει το 2.

Σύμφωνα με τον παραπάνω αλγόριθμο προκύπτει ο τύπος (*Brandes, Ulrik 2001*):

$$C_B(u) = \sum_{s \neq u \neq t} \frac{\sigma_{st}(u)}{\sigma_{st}}$$

όπου  $\sigma_{st}$  είναι ο αριθμός των ελάχιστων μονοπατιών που υπάρχουν από τον κόμβο  $s$  στον κόμβο  $t$  και  $\sigma_{st}(u)$  είναι ο αριθμός των ελάχιστων μονοπατιών από τον κόμβο  $s$  στον κόμβο  $t$  που διέρχονται όμως από τον κόμβο  $u$ . Ο παραπάνω δείκτης, με τον τρόπο που ορίστηκε, μπορεί να κανονικοποιηθεί διαιρώντας τον με τον αριθμό των ζευγών των κόμβων του γράφου  $G-\{u\}$ . Για μη κατευθυνόμενους γράφους ο αριθμός ισούται  $(n-1)(n-2)/2$ .

Στο σημείο αυτό, είναι αναγκαίο να επισημανθεί, ότι για τον υπολογισμό των *closeness* και *betweenness centrality* είναι απαραίτητος ο υπολογισμός των ελάχιστων μονοπατιών για κάθε ζεύγος κόμβων του γραφήματος. Αυτός ο υπολογισμός, όταν ο γράφος έχει μεγάλο αριθμό κόμβων και ακμών, είναι αδύνατος λόγω περιορισμένων υπολογιστικών πόρων.

### 2.5. Εύρεση κοινοτήτων

Ποικίλες μέθοδοι έχουν αναπτυχθεί στη θεωρία γράφων ως προς την εύρεση κοινοτήτων. Πρόκειται κατ' ουσίαν για την προσπάθεια διαχωρισμού/διαμέρισης ενός δικτύου σε ομάδες/κοινότητες εντός των οποίων οι κόμβοι παρουσιάζουν υψηλή πυκνότητα (ταυτόχρονα οι κόμβοι διαφορετικών κοινοτήτων παρουσιάζουν - συγκριτικά- χαμηλή/χαμηλότερη διασύνδεση). Η εν λόγω προσπάθεια έχει προσελκύσει ιδιαίτερο ενδιαφέρον τα τελευταία έτη, δεδομένου ότι έχει αυξηθεί η διαθεσιμότητα μεγάλων συνόλων δεδομένων δικτύων που απαιτούν την διαμέριση αυτών για την κατά περίπτωση μελέτη τους.

Οι μέθοδοι αυτοί βασίζονται συνήθως στην μεγιστοποίηση μίας μετρικής ή στην τυχαία διάλυση του δικτύου ή ακόμη και στη διάσπασή του.

Ένα από τα μέτρα που έχουν αναπτυχθεί για την εύρεση κοινοτήτων είναι γνωστό ως “modularity”. Πρόκειται για μετρική δομής του δικτύου, η οποία σχεδιάστηκε ακριβώς για τον σκοπό αυτό, δηλαδή τον καθορισμό της δυνατότητας διαίρεσης ενός δικτύου σε κοινότητες.

Ένα δίκτυο που έχει υψηλή τιμή *modularity* παρουσιάζει αυξημένη συνεκτικότητα εντός των συστάδων και αραιή σύνδεση μεταξύ κόμβων διαφορετικών συστάδων. Η μετρική στην γενική περίπτωση των  $c$  κοινοτήτων, ως προς τα σταθμισμένα δίκτυα (*weighted networks*) – όπως αυτά της παρούσας εργασίας (βλ. παρακάτω), ορίζεται ως εξής:

$$Q = \frac{1}{2m} \sum_{i,j} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j).$$

Η τιμή  $k_i k_j / 2m$  περιγράφει τον μέγιστο αριθμό συνδέσμων μεταξύ των 2 κόμβων, ενώ η τιμή  $A_{ij}$  τον πραγματικό αριθμό συνδέσμων.

Μία συσταδοποίηση θεωρείται καλή όταν η συνάρτηση  $Q$  παίρνει μεγάλη τιμή. Οι κανονικοποιημένες τιμές της συνάρτησης βρίσκονται στο διάστημα  $-1$  έως  $+1$ . Όταν  $Q = 0$  τότε όλοι οι κόμβοι έχουν αποδοθεί στην ίδια κοινότητα ενώ όταν  $Q < 0$  τότε κάθε κόμβος αποτελεί και μία κοινότητα.

Γενικά, αν για κάποια διαμέριση του γράφου η μετρική έχει αρνητική τιμή τότε το δίκτυο δεν μπορεί να χωριστεί σε κοινότητες. Ωστόσο, παρά το γεγονός πως ο ορισμός της *modularity* είναι διαισθητικά απλός, η υλοποίησή της παρουσιάζει προβλήματα (π.χ. αποτυγχάνει να εντοπίσει μικρές κοινότητες σε μεγάλους γράφους ακόμα κι αν αυτές είναι καλά ορισμένες). Στη συνέχεια παρουσιάζεται ο αλγόριθμος που χρησιμοποιείται από το πρόγραμμα ανοικτού κώδικα *gerphi*, το οποίο εφαρμόστηκε στα δεδομένα της παρούσας εργασίας (βλ. παρακάτω).

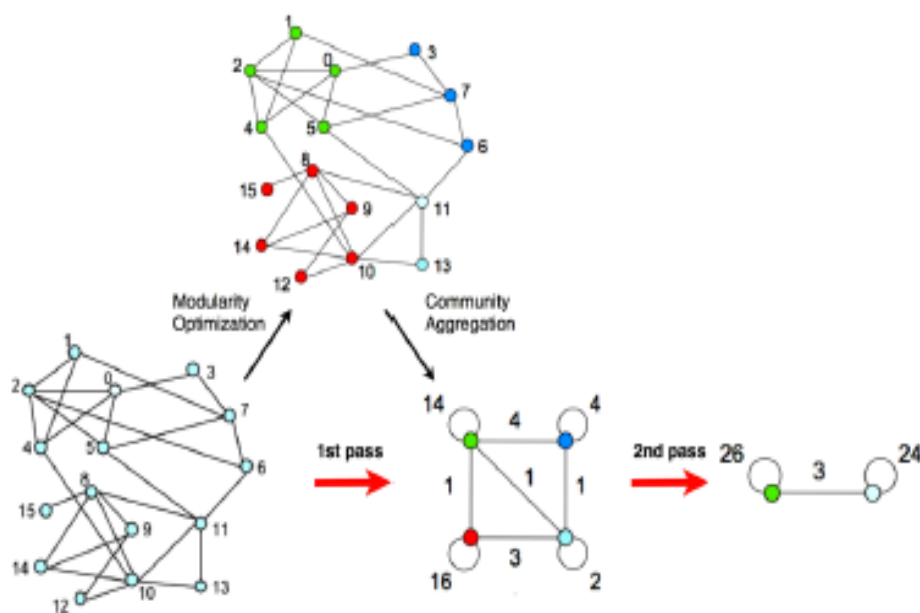
Ο αλγόριθμος αυτός (*V. Blondel, 2008*) φαίνεται να βρίσκει σε γρήγορο χρόνο διαμερίσεις (επιτυγχάνοντας υψηλή τιμή *modularity*) και να είναι αποδοτικός για δίκτυα με εκατομμύρια κόμβους.

Η διαδικασία που ακολουθείται χωρίζεται σε 2 επαναλαμβανόμενες φάσεις. Αρχικά, για ένα δίκτυο  $N$  κόμβων κάθε κόμβος αποτελεί μία διαφορετική κοινότητα. Κάθε κόμβος  $i$  μετακινείται ιδεατά σε κάθε κοινότητα των  $j$  γειτόνων του. Για κάθε μετακίνηση υπολογίζεται η νέα τιμή *modularity* και ο κόμβος μετακινείται στην κοινότητα που αποδίδει την μεγαλύτερη τιμή. Σε περίπτωση μη βελτίωσης ο κόμβος παραμένει στην αρχική του κοινότητα. Η πρώτη φάση υλοποιείται αφού η διαδικασία έχει επαναληφθεί για κάθε κόμβο και δεν παρατηρείται επιπλέον βελτίωση. Έτσι, ο αλγόριθμος επιτυγχάνει ένα τοπικό μέγιστο της τιμής *modularity*. Ο υπολογισμός της βελτίωσης της μετρικής προκύπτει γρήγορα από τον ακόλουθο τύπο:

$$\Delta Q = \left[ \frac{\sum_{in} + 2k_{i,in}}{2m} - \left( \frac{\sum_{tot} + k_i}{2m} \right)^2 \right] - \left[ \frac{\sum_{in}}{2m} - \left( \frac{\sum_{tot}}{2m} \right)^2 - \left( \frac{k_i}{2m} \right)^2 \right]$$

όπου  $\Sigma_{in}$  είναι το άθροισμα των βαρών των συνδέσμων εντός της κοινότητας,  $\Sigma_{tot}$  το άθροισμα των βαρών των συνδέσμων που προσπίπτουν στους κόμβους της κοινότητας,  $k_i$  είναι το άθροισμα των βαρών των συνδέσμων που προσπίπτουν στον κόμβο  $i$ ,  $k_{i,in}$  το άθροισμα των βαρών των συνδέσμων από τον κόμβο  $i$  προς τους κόμβους της κοινότητας και  $m$  είναι το άθροισμα όλων των βαρών των συνδέσμων του δικτύου.

Στη β' φάση δημιουργείται ένα νέο δίκτυο οι κόμβοι του οποίου συνιστούν τις κοινότητες που προέκυψαν κατά την α' φάση. Οι νέες σύνδεσμοι δίνονται από το άθροισμα των βαρών των συνδέσμων μεταξύ των κόμβων των κοινοτήτων. Στη συνέχεια, επαναλαμβάνεται η α' φάση.



Εικόνα 7: Οπτικοποίηση των βημάτων του αλγορίθμου. Η διαδικασία αποτελείται από 2 φάσεις: Α' φάση, όπου η modularity βελτιστοποιείται επιτρέποντας μόνο τοπικές αλλαγές των κοινοτήτων, Β' φάση, οι κοινότητες που βρέθηκαν συγκεντρώνονται προκειμένου να κατασκευαστεί ένα νέο δίκτυο κοινοτήτων. Η διαδικασία επαναλαμβάνεται έως ότου μία περαιτέρω αύξηση της modularity να μην είναι δυνατή. (πηγή: Vincent D Blondel, 2008)

## 2.6. Μοντέλα δικτύων

Τα κυρίαρχα μοντέλα που συνιστούν ταυτόχρονα ορόσημα στην μελέτη της ανάλυσης δικτύων είναι τα ακόλουθα:

- Τυχαία δίκτυα (*Erdos-Renyi model, ER*)
- Small World Network (*Watts-Strogatz Model, WS*)
- Τα τυχαία scale-free Network (*Barabasi-Albert Model, BA*).

Ακολουθεί η αναγκαία για την εργασία, περιγραφή αυτών.

### 2.6.1. Τυχαία δίκτυα (*Erdos – Renyi model, ER*)

Ένα τυχαίο γράφημα μπορεί να προκύψει με τον εξής τρόπο:

Ξεκινώντας με ένα σύνολο  $V=\{1,2,\dots,n\}$   $n$  κόμβων και προσθέτοντας τυχαία ακμές που συνδέουν τους κόμβους αυτούς.

Ο μέγιστος αριθμός ακμών που μπορούν να προστεθούν είναι ίσος με  $N = \binom{n}{2}$

Διαφορετικά τυχαία γραφήματα παρουσιάζουν διαφορετικές κατανομές βαθμών στους κόμβους. Τα τυχαία γραφήματα μελετώνται συνήθως βάσει του μοντέλου των *Erdos-Renyi*  $G(n,M)$ .<sup>6</sup>

Σε ένα τυχαίο δίκτυο *Erdos-Renyi*, το μέγεθος του δικτύου θα είναι  $N$  και η πυκνότητά του θα εξαρτάται από την τιμή της παραμέτρου  $p$  (όσο χαμηλότερη, τόσο πυκνότερο το δίκτυο). Μια βασική στατιστική ιδιότητα που μας ενδιαφέρει είναι η κατανομή των βαθμών των κόμβων του. Με βάση τον τρόπο που σχηματίστηκε το δίκτυο, μπορούμε να υπολογίσουμε την πιθανότητα που έχει ο κάθε κόμβος να έχει βαθμό  $d$ .

Η πιθανότητα αυτή προσεγγίζεται πολύ ικανοποιητικά από την κατανομή *Poisson*. Ως εκ τούτου, τα τυχαία δίκτυα έχουν μια πολύ χαρακτηριστική κατανομή βαθμού κόμβων σε σχέση με τη μέση τιμή (όπως διαμορφώνεται από την παράμετρο  $p$ ). Η κατανομή αυτή προβλέπει ότι ο μεγαλύτερος αριθμός των κόμβων θα έχει βαθμό  $d$  κοντά στην μέση τιμή των βαθμών του δικτύου.

Περαιτέρω, ο μέσος συντελεστής συσπείρωσης φαίνεται να φθίνει αντιστρόφως ανάλογα με τον αριθμό των κόμβων για μεγάλες τιμές του  $N$ , κάτι που δεν συμβαίνει στα φυσικά δίκτυα, όπου παραμένει σταθερός και σχετικά μεγάλος.

Το μοντέλο *ER* απέχει πολύ από το να περιγράφει δίκτυα του πραγματικού κόσμου, γιατί η κατανομή κόμβων δεν είναι είναι νόμος δύναμης, αλλά νόμος *Poisson*, με εκθετικές ουρές που χαρακτηρίζουν τις κατανομές πεπερασμένης εμβέλειας.

### 2.6.2. *Small World Network (Watts-Strogatz Model, SW)*

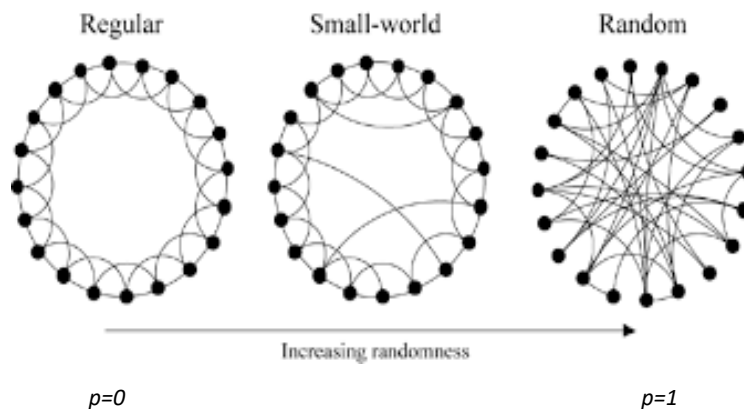
Μία σημαντική ιδιότητα που παρατηρείται σε πολλά πραγματικά δίκτυα είναι ο συνδυασμός μικρού μέσου μήκους μονοπατιού και μεγάλου συντελεστή συσσωμάτωσης. Αυτός ο συνδυασμός καλείται φαινόμενο του μικροκόσμου (*small-world*) και για την προσομοίωσή του έχουν δημιουργηθεί διάφορα μοντέλα. Το πιο γνωστό μοντέλο είναι αυτό των *Watts-Strogatz (1998)*.

Για την παρεμβολή μεταξύ τυχαίων δικτύων και δικτύων που παρουσιάζουν το *small world* φαινόμενο η παρακάτω διαδικασία επανασύνδεσης ακμών θεωρείται τυχαία.

Ξεκινώντας από ένα κυκλικό γράφο με  $n$  κόμβους, όπου κάθε κόμβος έχει  $k$  ακμές και συνδέεται με τους  $k/2$  πλησιέστερους γείτονες αριστερά του και τους  $k/2$

<sup>6</sup> Αλλά και του μοντέλου που προτάθηκε από τον Gilbert  $G(n,p)$ .

πλησιέστερους γείτονες δεξιά του, κάθε ακμή επανασυνδέεται τυχαία με πιθανότητα  $p$ . Αυτή η κατασκευή δίνει τη δυνατότητα να συντονισμού του δίκτυο μεταξύ της κανονικότητας ( $p=0$ ) και της απόλυτης διαταραχής ( $p=1$ ). Επιπροσθέτως, παρουσιάζει ιδιαίτερο ενδιαφέρον η περίπτωση στην οποία  $0 \leq p \leq 1$ .



Εικόνα 8: Απεικόνιση 3 δικτύων με χρήση της σταθεράς αποκοπής-επανασύνδεσης ( $p$ ) (πηγή Watts, 1998)

Ως εκ τούτου, το μοντέλο  $WS$  χρησιμοποιεί την σταθερά αποκοπής-επανασύνδεσης  $p$  ως παράμετρο τάξης, καθώς:

- για  $p=0$  έχουμε τον ντετερμινιστικό οργανωμένο δίκτυο με τους  $N$  ταλαντωτές που συνδέονται ο καθένας με  $K$  γείτονες τοποθετημένους συμμετρικά γύρω του, ενώ
- για  $p=1$  έχουμε ένα πλήρως στοχαστικό δίκτυο, τύπου  $ER$  (βλ. παραπάνω).

Έτσι η παράμετρος  $p$  μας οδηγεί από την οργάνωση στην αταξία, οπότε για τις ενδιάμεσες τιμές περιμένουμε να έχουμε μη τετριμμένη συμπεριφορά.

Για το μοντέλο  $WS$  έχει δειχτεί ότι υπάρχει ένας κρίσιμος αριθμός κόμβων  $N^*(p)$  που εξαρτάται από την σταθερά τάξης  $p$ , ώστε για την μέση απόσταση κόμβων  $\langle l \rangle$  να ισχύει:

- $\langle l \rangle \sim N$ , when  $N < N^*$   $N^* = N^*(p)$
- $\langle l \rangle \sim \ln N$ , when  $N > N^*$

Δηλαδή, για μεγάλα μεγέθη συστήματος η μέση απόσταση δεν μεγαλώνει ανάλογα με το μέγεθος του συστήματος αλλά με τον λογάριθμο του μεγέθους. Αυτή είναι και η έννοια του δικτύου μικρού κόσμου.

Τα υπόλοιπα μεγέθη, συντελεστής συσσωμάτωσης και κατανομή κόμβων δεν αλλάζουν δραστικά για μικρές τιμές του  $p$ , σε σχέση με το στοχαστικό δίκτυο Erdos-Renyi. Μάλιστα για  $p=0$  ο συντελεστής συσσωμάτωσης είναι σταθερός κι εξαρτάται από την τοπολογία του κάθε κόμβου, κι όχι από το  $N$ .

Μάλιστα, εάν ο συντελεστής συσσωμάτωσης για την περίπτωση δακτυλίου είναι  $C_{ring}$ , τότε ο συντελεστής συσσωμάτωσης για το μοντέλο Watts-Strogatz είναι:

$$C_{WS} = C_{ring} (1-p)^3 \approx C_{ring}, \text{ for small } p$$

Ο συντελεστής  $(1-p)^3$  εξασφαλίζει ότι και οι 3 συνδέσεις κάθε τριγωνικής σχέσης είναι ενεργές.

Το μοντέλο WS προσομοιάζει τα φυσικά δίκτυα στο ότι και τα δύο :

- είναι δίκτυα «μικρού κόσμου», δηλαδή έχουν λογαριθμική σχέση ανάμεσα στην ελάχιστη απόσταση κόμβων και στο μέγεθος του συστήματος και
- έχουν μεγάλο συντελεστή συσσωμάτωσης

Αντίθετα το μοντέλο WS διαφέρει από τα φυσικά δίκτυα στο ότι η κατανομή κόμβων ακολουθεί νόμο πεπερασμένης εμβέλειας, καθώς για τιμές του  $p$  κοντά στο 0 προσεγγίζεται από την κατανομή  $\delta(k-\langle k \rangle)$ , η οποία πλαταίνει προς την κατανομή Poisson όταν οι τιμές του  $p \rightarrow 1$ . Οι κατανομές αυτές διαφέρουν μη-τετριμμένα από τους νόμους δύναμης που παρατηρούνται στα φυσικά δίκτυα.

### 2.6.3. Τα τυχαία *scale-free Network* (Barabasi-Albert, BA)

Σε έναν μικρό αριθμό κόμβων, προσθέτουμε σε κάθε χρονικό βήμα έναν νέο κόμβο με ακμές, που συνδέουν τον κόμβο αυτό με κάποιους από τους υπάρχοντες κόμβους. Η πιθανότητα  $P$  ένας νέος κόμβος να συνδεθεί με τον κόμβο  $i$  εξαρτάται από τον βαθμό  $k_i$  αυτού του κόμβου. Η μέθοδος αυτή ονομάστηκε από τους Barabasi-Albert ως «επιλεκτική συνδεσιμότητα» (preferential attachment). Συναφώς, ισχύει ότι:

$$P(k_i) = \frac{k_i}{\sum_j k_j}$$

Μετά από  $t$  χρονικά βήματα θα προκύψει ένα δίκτυο. Το εν λόγω δίκτυο εξελίσσεται σε μία αμετάβλητης κλίμακας κατάσταση με την πιθανότητα ένας κόμβος να έχει βαθμό  $k$ , να ακολουθεί εκθετική κατανομή με εκθέτη  $\gamma$ .

Η εκθετική κατανομή των περισσότερων πραγματικών δικτύων αφορά και στο βαθμό των κόμβων, περιγράφει πολλά διαφορετικά μεγέθη και είδη δυναμικών συστημάτων. Συνεπώς, αναμένεται ότι το καλύτερο μοντέλο προσομοίωσης ενός δικτύου θα παρέχει κατανομή που τα κύρια χαρακτηριστικά της θα είναι η ανεξαρτησία της ως προς το χρόνο. Η κατανομή  $P(k)$  είναι ανεξάρτητη του χρόνου και του μεγέθους του συστήματος το οποίο αναδεικνύει πως παρά το γεγονός ότι το δίκτυο αυξάνεται συνεχώς, φαίνεται να συγκλίνει σε μία *scale-free* στατική κατάσταση.

Με την εκθετική κατανομή στο δίκτυο να κλιμακώνεται, εξάγεται ότι η ανάπτυξη και η μέθοδος *preferential attachment* παίζουν εξέχοντα ρόλο στην εξέλιξη του δικτύου.

Ένεκα της «επιλεκτικής συνδεσιμότητας», ένας κόμβος που αποκτά περισσότερες ακμές από έναν άλλο, θα αυξήσει στην πορεία την συνδεσιμότητά του με



υψηλότερο ρυθμό. Κατά αυτόν τον τρόπο, μία αρχική διαφορά της συνδεσιμότητας μεταξύ 2 κόμβων θα αυξάνεται με την πάροδο των χρονικών βημάτων.

Παλιοί κόμβοι του δικτύου φαίνεται να αυξάνουν την συνδεσιμότητά τους «εις βάρος» των νέων κόμβων του δικτύου. Αυτό έχει ως αποτέλεσμα μερικοί κόμβοι μετά από μία μεγάλη περίοδο στο χρόνο να έχουν μεγάλο βαθμό. Κατά αυτόν τον τρόπο εμφανίζεται και το φαινόμενο γνωστό ως “rich get richer” (ένας κόμβος μεγάλου βαθμού αυξάνει περισσότερο το βαθμό του σε σχέση με άλλους κόμβους που έχουν -συγκριτικά- μικρότερο βαθμό). Το φαινόμενο αυτό απαντάται συνήθως στα δίκτυα του πραγματικού κόσμου.

Περαιτέρω, σημειώνονται και οι ακόλουθες ιδιότητες του δικτύου BA:

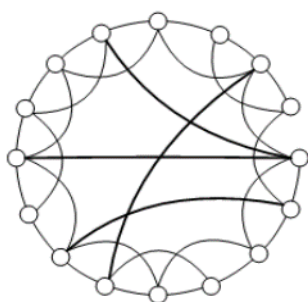
- Η μέση απόσταση μεταξύ κόμβων έχει λογαριθμική σχέση με τον αριθμό των κόμβων  $\langle l \rangle \approx \ln(N+c)$  BA-11
- Ο συντελεστής συσσωμάτωσης C είναι μεγαλύτερος από των WS.<sup>7</sup>

#### 2.6.4. Συγκριτική επισκόπηση κυρίαρχων μοντέλων δικτύων

Έχοντας παραθέσει τα βασικά χαρακτηριστικά του κάθε μοντέλου, κρίνεται σκόπιμη μια επιμέρους σύγκριση μεταξύ αυτών, ούτως ώστε να αναδειχθούν κοινά σημεία, προσεγγίσεις/συγκλίσεις και διαφορές.

Καταρχάς, σημειώνονται συνοπτικά τα κύρια χαρακτηριστικά κάθε ενός μοντέλου:

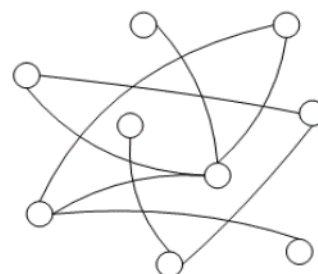
- **Στοχαστικό δίκτυο, ER:** όλοι οι κόμβοι είναι περίπου ισοδύναμοι, με τις συνδέσεις τυχαία κατανομημένες.
- **Δίκτυο πεπερασμένου – «μικρού» κόσμου, WS:** Είναι οργανωμένο και ντετερμινιστικό, αλλά φέρει έναν μικρό αριθμό τυχαίων συνδέσεων μακράς εμβέλειας.
- **Δίκτυο ελεύθερης κλίμακας, BA:** Δημιουργείται ένας μικρός αριθμός κόμβων που φέρουν μεγάλη συνδεσιμότητα.



(α) Smal-World Network (SWN)



(β) Scale-Free Network (SFN)



(γ) Random Network (RN)

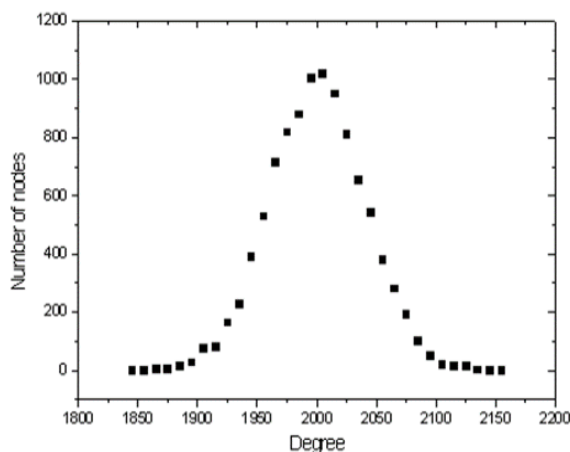
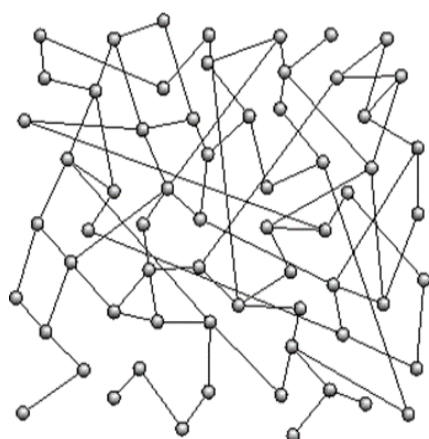
Εικόνα 9: Απεικόνιση δικτύων των κυρίαρχων μοντέλων  
(πηγή: Provata)

<sup>7</sup> Πρέπει να αναφερθεί ότι πρόκειται για ευρήματα κυρίως βασισμένα σε αριθμητικά αποτελέσματα (δεν υφίστανται αξιόπιστοι αναλυτικοί υπολογισμοί).

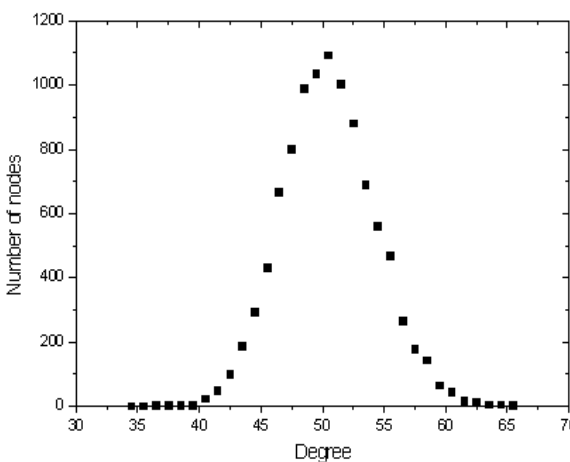
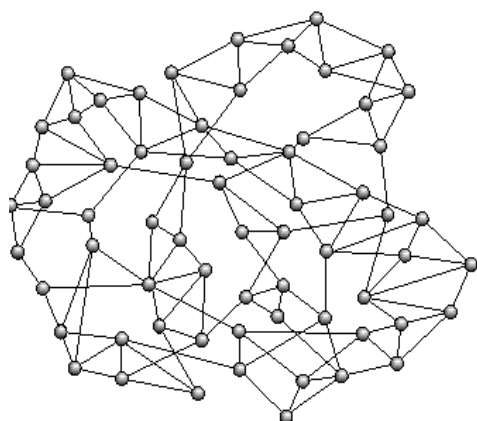
Περαιτέρω, παρατίθεται συγκριτικός πίνακας με ορισμένα βασικά ποσοτικά χαρακτηριστικά των παραπάνω δικτύων καθώς και απεικονίσεις αυτών, με έμφαση στη κατανομή μέσης βάρθρωσης (average degree distribution).

Measurement	Erdős-Rényi	Watts-Strogatz	Barabási-Albert
Degree distribution	$P(k) = \frac{e^{-\langle k \rangle} \langle k \rangle^k}{k!}$	$P(k) = \sum_{i=1}^{\min(k-\kappa, \kappa)} \binom{\kappa}{i} (1-p)^i p^{\kappa-i} \frac{(p\kappa)^{k-\kappa-i}}{(k-\kappa-i)!} e^{-p\kappa}$	$P(k) \sim k^{-3}$
Average vertex degree	$\langle k \rangle = p(N-1)$	$\langle k \rangle = 2\kappa^*$	$\langle k \rangle = 2m$
Clustering coefficient	$C = p$	$C(p) \sim \frac{3(\kappa-1)}{2(2\kappa-1)} (1-p)^3$	$C \sim N^{-0.75}$
Average path length	$\ell \sim \frac{\ln N}{\ln \langle k \rangle}$	$\ell(N, p) \sim p^\tau f(Np^\tau)^*$	$\ell \sim \frac{\log N}{\log \langle k \rangle}$

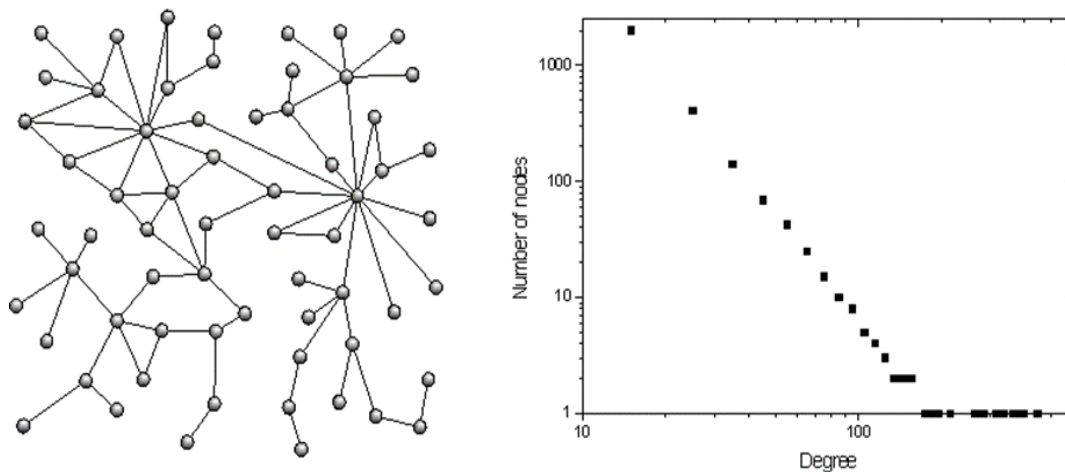
Πίνακας 1: Βασικά ποσοτικά χαρακτηριστικά των 3 κυρίαρχων μοντέλων (Costa, 2008)



Εικόνα 10: Απεικόνιση ενός τυχαίου δικτύου και της κατανομής μέσης βάρθρωσης (Costa, 2008)



Εικόνα 11: Απεικόνιση ενός δικτύου πεπερασμένου-«μικρού» κόσμου και της κατανομής μέσης βάρθρωσης (Costa, 2008)



Εικόνα 12: Απεικόνιση ενός δικτύου ελεύθερης κλίμακας και της κατανομής μέσης βάρθρωσης (Costa, 2008)

### 3. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

Κατά τη διάρκεια της δεκαετίας του 1990 (και ενδεχομένως νωρίτερα), ένας αριθμός συγγραφέων επεσήμανε τη δυνητική χρησιμότητα των δεδομένων συν-συγγραφής και σε ορισμένες περιπτώσεις πραγματοποίησε στατιστικές αναλύσεις μικρής κλίμακας αξιοποιώντας συαυφή στοιχεία.<sup>8</sup>

Ωστόσο, η σε βάθος μελέτη αυτών των δικτύων κατέστη εφικτή με την έλευση των ηλεκτρονικών βιβλιογραφιών. Γύρω στο 2000, αρκετοί ερευνητές ξεκίνησαν την κατασκευή δικτύων μεγάλης κλίμακας παρουσιάζοντας χαρακτηριστικά της έρευνας σε διάφορα γνωστικά πεδία όπως στα μαθηματικά, στη βιολογία, στη φυσική, στην επιστήμη των υπολογιστών και στη νευροεπιστήμη. Για το σκοπό αυτό αξιοποίησαν διάφορες βάσεις δεδομένων του επιστημονικού κόσμου<sup>9</sup>. Ενδεικτικά αναφέρονται οι ακόλουθες (παρατίθενται με αλφαβητική σειρά):

- *American Physical Society (APS)*
- *arXiv. arXiv*
- *Google Scholar*
- *MEDLINE/PubMed*
- *Microsoft Academic*
- *Scopus*
- *Web of Science (WoS)*

Μία από τις σημαντικότερες δημοσιεύσεις επί του θέματος είναι του *Newman (2004)* ο οποίος ανέλυσε και συνέκρινε 3 δίκτυα επιστημόνων διαφορετικών γνωστικών πεδίων (βιοιατρικής, φυσικής και μαθηματικών), μελετώντας τα βασικότερα ποσοτικά χαρακτηριστικά αυτών και αναγνωρίζοντας συγκεκριμένα μοτίβα. Σημειώνεται ότι, πρόκειται για δίκτυα με ένα τύπο κόμβων (*monopartite networks*), εν προκειμένω τους ερευνητές/επιστήμονες. Δύο επιστήμονες θεωρήθηκαν ότι συνδέονται μεταξύ τους εάν είχαν συγγράψει κάποια μελέτη.

Στον ακόλουθο πίνακα αποτυπώνονται κάποια βασικά στατιστικά στοιχεία, όπως ο αριθμός των συγγραφέων, των δημοσιεύσεων κ.λπ. ανά επιστημονικό πεδίο/δίκτυο καθώς και ποσοτικά στοιχεία που εξήχθησαν μέσω της ανάλυσης δικτύων (π.χ. διάμετρος δικτύων κ.λπ.), ενώ αναδεικνύονται κοινά σημεία και επιμέρους διαφορές μεταξύ αυτών.

---

<sup>8</sup> Π.χ. ως προς τη συχνότητα εμφάνισης κοινών δημοσιεύσεων/μελετών από συγκεκριμένους συγγραφείς και ως προς τους συγγραφείς σε συγκεκριμένα ινστιτούτα.

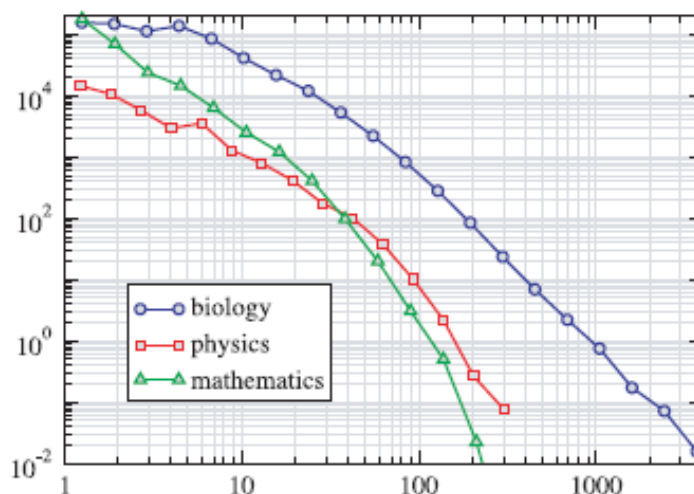
<sup>9</sup> Με στοιχεία αναφορικά με συγγραφείς, ερευνητές κάθε τομέα, δημοσιεύσεις, μελέτες, αναφορές κ.λπ.

	Biology	Physics	Mathematics
Number of authors	1,520,251	52,909	253,339
Number of papers	2,163,923	98,502	—
Papers per author	6.4	5.1	6.9
Authors per paper	3.75	2.53	1.45
Average collaborators	18.1	9.7	3.9
Largest component	92%	85%	82%
Average distance	4.6	5.9	7.6
Largest distance	24	20	27
Clustering coefficient	0.066	0.43	0.15
Assortativity	0.13	0.36	0.12

Πίνακας 2: Βασικά ποσοτικά χαρακτηριστικά των 3 δικτύων συσσυγγραφέων  
(πηγή: Newman, 2004)

Όπως διαπιστώθηκε τόσο η κατανομή των συν-συγγραφέων όσο και των δημοσιεύσεων τους ακολουθούν νόμο δύναμης (*power-law*).

Στο ακόλουθο διάγραμμα εμφανίζονται χαρακτηριστικά οι κατανομές βάρθρωσης μεταξύ των 3 δικτύων, οι οποίες είναι παρόμοιες, αν και η κατανομή για τη βιοϊατρική (κύκλοι) έχει μεγαλύτερη ουρά, αντανακλώντας τον υψηλότερο μέσο αριθμό συνεργατών που έχουν τα άτομα στον τομέα αυτό. Σε κάθε περίπτωση, η κατανομή είναι *fat-tailed*, όπως η κατανομή του αριθμού των δημοσιεύσεων (βλ. παραπάνω), με έναν μικρό αριθμό επιστημόνων να έχει πολύ μεγάλο αριθμό συνεργατών, έως και χιλιάδες στην περίπτωση του δικτύου βιολογίας.



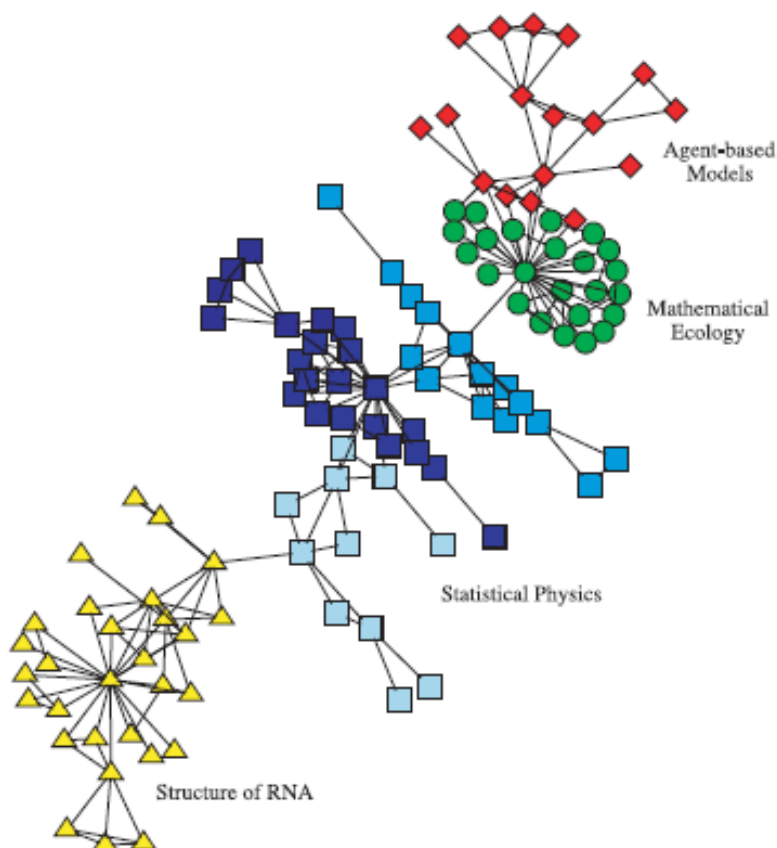
Εικόνα 13 Διάγραμμα της κατανομής αριθμών συνεργατών για τους επιστήμονες σε κάθε έναν από τους 3 τομείς που μελετήθηκαν  
(πηγή: Newman, 2004)

Οι επιμέρους διαφοροποιήσεις ασφαλώς σχετίζονται με εγγενείς διαφορές που ενυπάρχουν μεταξύ των 3 επιστημονικών πεδίων όπως τον τρόπο διεξαγωγής των ερευνών (π.χ. η βιολογία είναι κυρίως πειραματική, τα μαθηματικά είναι εξ ολοκλήρου θεωρητικά και η φυσική είναι ένας συνδυασμός των δύο).

Επιπρόσθετα, τα υποδίκτυα των επιμέρους γνωστικών πεδίων παρουσιάζαν τα χαρακτηριστικά των δικτύων του «μικρού κόσμου» (*small world*). Χαρακτηριστικά αναφέρεται η μικρή σχετικά τιμή της μέσης («ελάχιστης») απόστασης κόμβων, καθώς διαπιστώθηκε ότι κάθε επιστημόνας απείχε μόλις 5 ή 6 βήματα μακριά από έναν άλλο τυχαία επιλεγμένο επιστήμονα, όπως και η υψηλή συσταδοποίηση (2 επιστήμονες τείνουν να έχουν συνεργαστεί εάν έχουν έναν τρίτο κοινό συνεργάτη).

Τα δίκτυα συνεργασίας έχουν επίσης κοινοτική δομή. Μια κοινότητα ορίζεται ως μια υποομάδα κόμβων μέσα στους οποίους οι δεσμοί είναι πυκνοί, εν αντιθέσει με τους υπόλοιπους κόμβους – εξωτερικά της κοινότητας – με τους οποίους οι δεσμοί είναι πολύ πιο αραιωμένοι. Εν προκειμένω, ο εντοπισμός των κοινοτήτων στα δίκτυα συν-συγγραφιών πραγματοποιήθηκε μέσω της επαναλαμβανόμενης μείωσης των δεσμών με την υψηλή τιμή του δείκτη *betweenness centrality*.

Στη παρακάτω εικόνα εμφανίζεται ένα μικρό δίκτυο συνεργασίας επιστημόνων ενός ιδιωτικού ερευνητικού Ινστιτούτου<sup>10</sup>. Μέσω εφαρμογής της παραπάνω μεθόδου, το δίκτυο διασπάσθηκε σε 6 κοινότητες οι οποίες εν πολλοίς σχετίζονται με τα ερευνητικά πεδία των επιστημόνων (διαφορετικά σχήματα-χρώματα κόμβων).

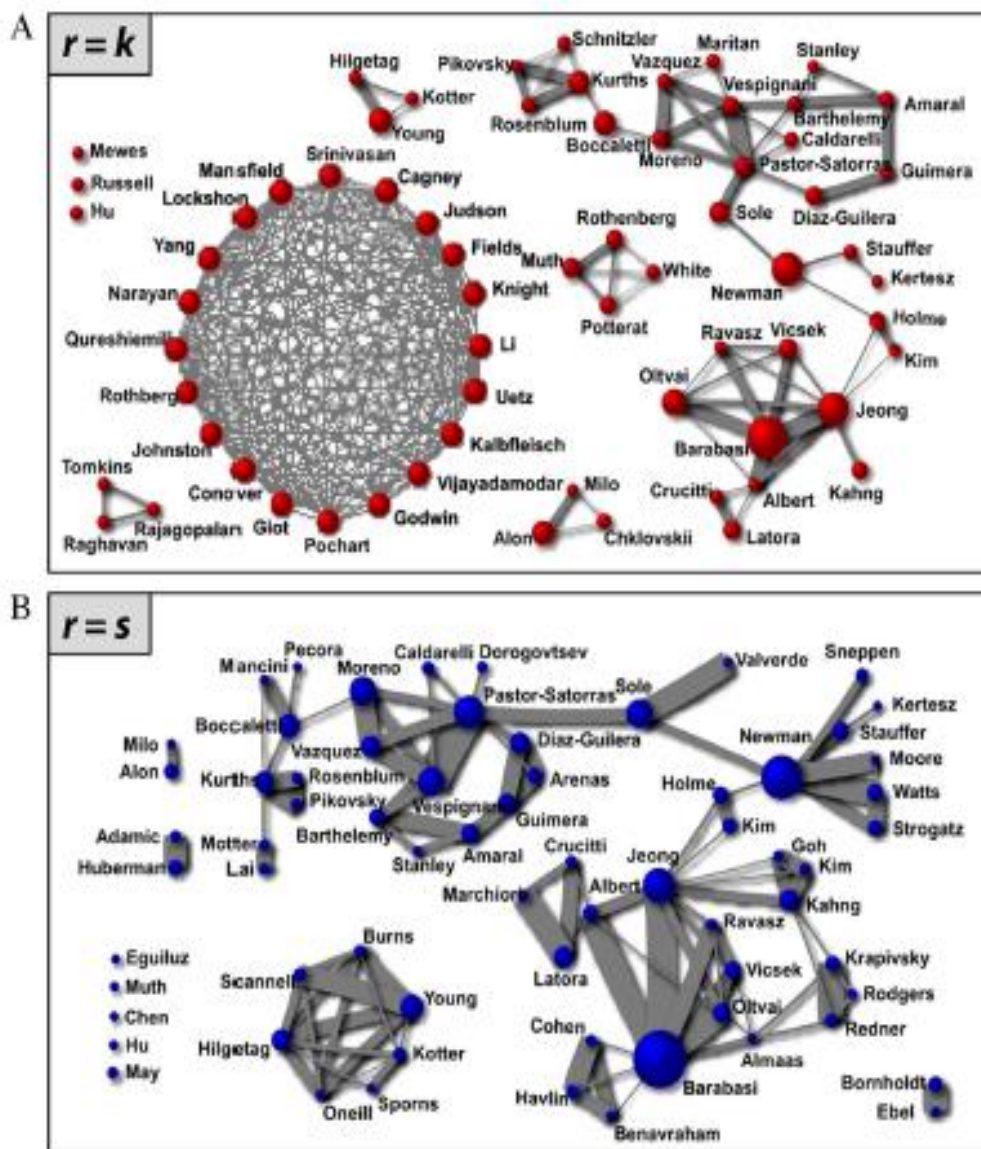


Εικόνα 14: Απεικόνιση ενός δικτύου συν-συγγραφής εντός ενός ερευνητικού Ινστιτούτου και ανάδειξη επιμέρους κοινοτήτων σε παραλληλία με τα επιμέρους γνωστικά πεδία των επιστημόνων (πηγή: Girvan, 2002)

<sup>10</sup> Ειδικότερα πρόκειται για το Santa Fe Institute (Santa Fe, New Mexico, USA).

Μια άλλη σημαντική ιδιότητα των δικτύων συνεργασίας έχει να κάνει με την ταξινομησιμότητα, δηλαδή με την προτίμηση των κόμβων ενός δικτύου να προσκολλώνται σε άλλους κόμβους με τους οποίους είναι όμοιοι κατά κάποιον τρόπο. Ειδικότερα οι υψηλά (ή χαμηλά) διασυνδεδεμένοι κόμβοι τείνουν να συνδέονται με άλλους υψηλά (ή χαμηλά) διασυνδεδεμένους κόμβους. Η τάση αυτή, η οποία αποκαλείται *assortative mixing* ή *assortativity*, παρατηρήθηκε -τοπικά- σε όλα τα παραπάνω δίκτυα (και στα 3 επιστημονικά πεδία). Στην ίδια συνάφεια αξίζει να αναφερθούν οι ακόλουθες μελέτες/ευρήματα έτερων ερευνητών:

- Το φαινόμενο γωνστό ως «*rich-club*» (μέσω της εστίασης στους εξαιρετικά συνδεδεμένους κόμβους), όταν δηλ. οι κόμβοι υψηλότερου βαθμού αλληλοσυνδέονται περισσότερο από κόμβους με χαμηλότερο βαθμό. Ειδικότερα το φαινόμενο αναφέρεται στην τάση των κόμβων με υψηλό βαθμό να σχηματίζουν στενά αλληλένδετες κοινότητες, οι οποίες είναι σε κάποιο βαθμό σύμφωνες με τη διαίσθηση ότι η «ελίτ» που σχηματίζεται από ισχυρούς επιστήμονες -από άποψη επιρροής- τείνει να σχηματίζει συνεργατικές ομάδες σε συγκεκριμένους τομείς.
- Ωστόσο, το φαινόμενο αυτό αμφισβητήθηκε αργότερα από τον *Opsahl* κ.α. οι οποίοι κατασκεύασαν ένα σταθμισμένο δίκτυο συνεργασίας από τα δεδομένα *arXiv* με βάση τις δημοσιεύσεις σχετικά με την επιστήμη δικτύων (το βάρος στις συνδέσεις μεταξύ 2 επιστημόνων αντιπροσώπευε τον αριθμό των άρθρων που συνυπέγραφαν οι ίδιοι). Κατά τον εντοπισμό του φαινομένου «*rich-club*» στο σταθμισμένο δίκτυο, οι ερευνητές υπολόγισαν το συνολικό βαθμό των βαρών βάσει της επιρροής των επιστημόνων αντί του αριθμού των δεσμών μεταξύ τους. Τα αποτελέσματα εμφάνισαν εντυπωσιακά διαφορετικές τάσεις για τα δίκτυα επιστημονικής συνεργασίας. Οι επιστήμονες με σημαντική επιρροή δημιουργούσαν πλέον μία μεγάλη κοινότητα μέσα στην οποία ο ένας συνεργάζονταν με τον άλλο. Αντ' αυτού, μεγιστοποιούσαν τους πόρους τους ανά συνεργασία, τείνοντας να συνεργάζονται εντατικά μεταξύ τους.



Εικ. 15. Υποσύνολα κόμβων με σημαντική επιρροή στο δίκτυο συνεργασίας δικτύων επιστήμης με βάση το βαθμό (Α) και την ισχύ (Β). Απεικονίζονται μόνο οι σύνδεσμοι μεταξύ των κόμβων με σημαντική επιρροή. Το μέγεθος των κόμβων είναι ανάλογο του «πλούτου» τους, ενώ το πλάτος των συνδέσμων είναι ανάλογο του βάρους τους.

(πηγή: Orsahl, 2008)



## 4. ΔΕΔΟΜΕΝΑ

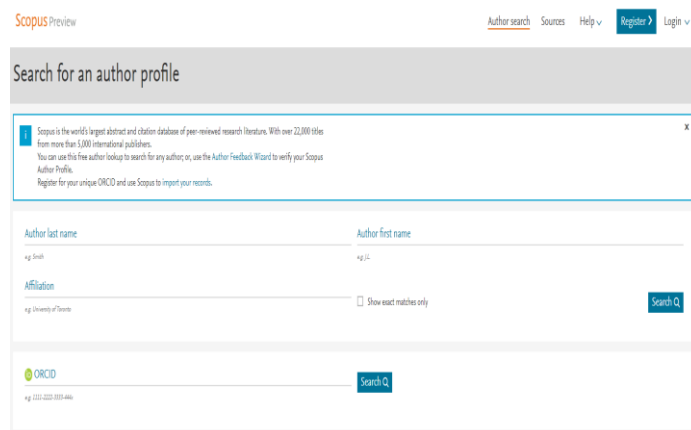
### 4.1. Δεδομένα & Παραδοχές

Τα στοιχεία τα οποία συλλέχθηκαν και εν συνεχεία επεξεργάστηκαν, στο πλαίσιο της παρούσας εργασίας, αφορούν σε κοινές δημοσιεύσεις ερευνητών του Ινστιτούτου Νανοεπιστήμης και Νανοτεχνολογίας (INN) του ΕΚΕΦΕ «Δημόκριτος». Ως κοινές δημοσιεύσεις θεωρήθηκαν οι δημοσιεύσεις μεταξύ τουλάχιστον 2 ερευνητών του ως άνω Ινστιτούτου. Κατά συνέπεια, δεν συμπεριλαμβάνονται κοινές δημοσιεύσεις ενός ερευνητή του ως άνω Ινστιτούτου, με ερευνητές άλλων Ινστιτούτων ή πανεπιστημίων.

Περαιτέρω, ως περίοδος μελέτης θεωρήθηκε το χρονικό διάστημα 2008-2017 (10ετία). Σημειώνεται ότι, για τις ανάγκες της παρούσας μελέτης (πραγματοποίηση συγκρίσεων – αναζήτηση τυχόν μεταβολών), το ως άνω χρονικό διάστημα, διασπάστηκε σε 2 επιμέρους χρονικές περιόδους και ειδικότερα τις 5ετίες 2008-2012 και 2013-2017. Επίσης ότι, στις ως άνω κοινές δημοσιεύσεις μεταξύ ερευνητών του παραπάνω Ινστιτούτου, καταχωρήθηκαν εκείνες που αφορούν σε αμφότερες τις περιόδους (διαχρονική παρουσία – συγγραφή ερευνητών). Κατά συνέπεια, δεν συμπεριλαμβάνονται κοινές δημοσιεύσεις μεταξύ ερευνητών του ως άνω Ινστιτούτου, εφόσον αυτές αφορούν σε μία μόνο χρονική περίοδο (2008-2012 ή 2013-2017).

### 4.2. Συλλογή δεδομένων

Όπως αναφέρθηκε νωρίτερα (κεφ. 3) στοιχεία αναφορικά με την πανεπιστημιακή/ακαδημαϊκή κοινότητα (άρθρα, μελέτες, παραπομπές, συγγραφείς, επιστημονικά πεδία κ.λπ.) είναι διαθέσιμα σε διάφορες βιβλιογραφικές βάσεις δεδομένων. Μία εξ αυτών είναι η Scopus, ευρέως διαδεδομένη στην πανεπιστημιακή κοινότητα, που περιλαμβάνει περιλήψεις και παραπομπές για ακαδημαϊκά άρθρα περιοδικών σε διάφορα επιστημονικά πεδία (π.χ. φυσικές επιστήμες, επιστήμες υγείας κ.λπ.), ενώ παρέχει τη δυνατότητα αναζήτησης αυτών με διάφορα κριτήρια (συγγραφέας/εκδότης, πανεπιστήμιο, επιστημονικό πεδίο, χρόνος δημοσίευσης κ.λπ.)



The image shows a screenshot of the Scopus search interface. At the top, there is a navigation bar with 'Scopus Preview', 'Author search', 'Sources', 'Help', 'Register', and 'Login'. Below this is a search bar with the text 'Search for an author profile'. A blue box contains a message: 'Scopus is the world's largest abstract and citation database of peer-reviewed research literature. With over 22,000 titles from more than 5,000 international publishers. You can use this free author lookup to search for any author or, use the Author Feedback Wizard to verify your Scopus Author Profile. Register for your unique ORCID and use Scopus to import your records.' Below the message are two search sections. The first section is for 'Author last name' and 'Author first name', with a 'Search' button. The second section is for 'ORCID', with a 'Search' button. There is also a checkbox for 'Show exact matches only'.

Εικόνα 16: Βιβλιογραφική βάση δεδομένων Scopus (αναζήτηση βάσει εκδότη/συγγραφέα)

Για την συλλογή των δεδομένων πραγματοποιήθηκε αναζήτηση στην προαναφερόμενη βιβλιογραφική βάση. Πέραν αυτού, το σύνολο το στοιχείων χορηγήθηκε/επιβεβαιώθηκε και μέσω του Ινστιτούτου Νανοεπιστήμης και Νανοτεχνολογίας, με τη συμπερίληψη της προέλευσης έκαστου ερευνητή (αρχικό ινστιτούτο – προ συγχώνευσης).

#### 4.3. Βασικά χαρακτηριστικά & Ταξινόμηση δεδομένων

Τα στοιχεία που συγκεντρώθηκαν, ταξινομήθηκαν και αναλύθηκαν στο πλαίσιο της παρούσας μελέτης, υπό το πρίσμα της θεωρίας γράφων, αφορούν στα εξής:

- Πλήθος ερευνητών του INN με κοινές δημοσιεύσεις, που συνιστούν τους κόμβους του δικτύου (*nodes/vertices/points*) και
- Πλήθος κοινών δημοσιεύσεων μεταξύ των ως άνω ερευνητών, που συνιστούν τις συνδέσεις του δικτύου (*edges/links/lines*).

Περαιτέρω, πρέπει να αναφερθούν τα εξής:

- Πρόκειται για δίκτυο με έναν τύπο κόμβων (*monopartite*), εν προκειμένω τους συγγραφείς/ερευνητές.
- Πρόκειται για μη κατευθυνόμενο δίκτυο (οι σχέσεις μεταξύ των κόμβων δεν είναι προσανατολισμένες – η κοινή δημοσίευση είναι αμοιβαία – συμμετρική σχέση).<sup>11</sup>
- Πρόκειται για σταθμισμένο δίκτυο, δεδομένου ότι καταγράφεται όχι μόνο η ύπαρξη ή μη κοινής δημοσίευσης αλλά και το πλήθος αυτών.<sup>12</sup>
- Τέλος, υφίσταται μόνο ένας τύπος συνδέσεων μεταξύ των κόμβων/ερευνητών<sup>13</sup>, αυτός της κοινής δημοσίευσης.

Στη βάση όσων προαναφέρθηκαν κατασκευάστηκε ο πίνακας συνδεσιμότητας (*adjacency matrix*) του δικτύου (ένας για κάθε χρονική περίοδο). Πρόκειται για συμμετρικό πίνακα, καθώς αφορά σε μη κατευθυνόμενο δίκτυο. Δεδομένου ότι το πλήθος των κόμβων που μελετήθηκε είναι 133 (βαθμός δικτύου N:133) προέκυψε ένας πίνακας διαστάσεων 133x133, τα στοιχεία του οποίου είναι τα βάρη των συνδέσεων των κόμβων (πλήθος κοινών δημοσιεύσεων μεταξύ των ερευνητών).

---

<sup>11</sup> Για την κατασκευή ενός κατευθυνόμενου δικτύου θα μπορούσε να ληφθεί υπόψη η συνεισφορά καθενός ερευνητή στην κοινή τους δημοσίευση. Ωστόσο, μία τέτοια θεώρηση ενέχει κινδύνους καθόσον δεν είναι εφικτό να καθορισθεί σε όλες τις περιπτώσεις (π.χ. προς τη κατεύθυνση αυτή, θα μπορούσε να αξιοποιηθεί η σειρά αναφοράς των συγγραφέων σε κάθε δημοσίευση, δεδομένου ότι κατά κανόνα παρατίθενται βάσει του βαθμού συμμετοχής τους, χωρίς ωστόσο αυτό να εφαρμόζεται απόλυτα).

<sup>12</sup> Σε διαφορετική περίπτωση θα αφορούσε σε αστάθμιστο δίκτυο, στο οποίο οι τιμές των συνδέσεων θα ήταν 0 ή 1 (υφίσταται ή όχι σχέση-σύνδεση-συνδημοσίευση).

<sup>13</sup> Σε διαφορετική περίπτωση θα αφορούσε σε πολλαπλές σχέσεις – συνδέσεις π.χ. μέσω των βιβλιογραφικών αναφορών/παραπομπών, βάσει του επιστημονικού πεδίου, του χώρου προέλευσης των ερευνητών κ.λπ.

Σημειώνεται ότι, για ευνόητους λόγους, οι ερευνητές κωδικοποιήθηκαν λαμβάνοντας έναν μοναδικό αριθμό από το 1 έως το 133. Επίσης ότι, στη διαγώνιο τέθηκε μηδενικός αριθμός (αφορούν σε έναν-τον ίδιο κάθε φορά ερευνητή).

Εικόνα 17: Αποτύπωση/απόσπασμα του κωδικοποιημένου πίνακα συνδεσιμότητας του δικτύου (αφορά στους πρώτους 50 κόμβους, πίνακας 50x50).

Τέλος αξίζει να αναφερθεί ότι για την α' χρονική περίοδο (2008-2012) το πλήθος των συνδέσμων είναι 500, ενώ τη β' χρονική περίοδο (2013-2017) είναι 657, καταγράφοντας αύξηση 31,4%. Στην ίδια συνάφεια, το άθροισμα των βαρών («φορτίων») των συνδέσμων ανά χρονική περίοδο παρουσιάζει παρόμοια αύξηση, της τάξεως 31,85% (βλ. πίνακα).

Βεβαίως, το πλήθος των κόμβων παραμένει σταθερό για αμφότερες τις περιόδους και ίσο με 133, βάσει υπόθεσης, δεδομένου ότι μελετήθηκε η μεταβολή των συνδέσμων των υφιστάμενων κατά την α' και β' χρονική περίοδο κόμβων.

	2008-2012	2013-2017	μεταβολή(%)
πλήθος κόμβων	133	133	-
πλήθος συνδέσμων	500	657	31,40%
άθροισμα βαρών συνδέσμων	1.824	2.405	31,85%

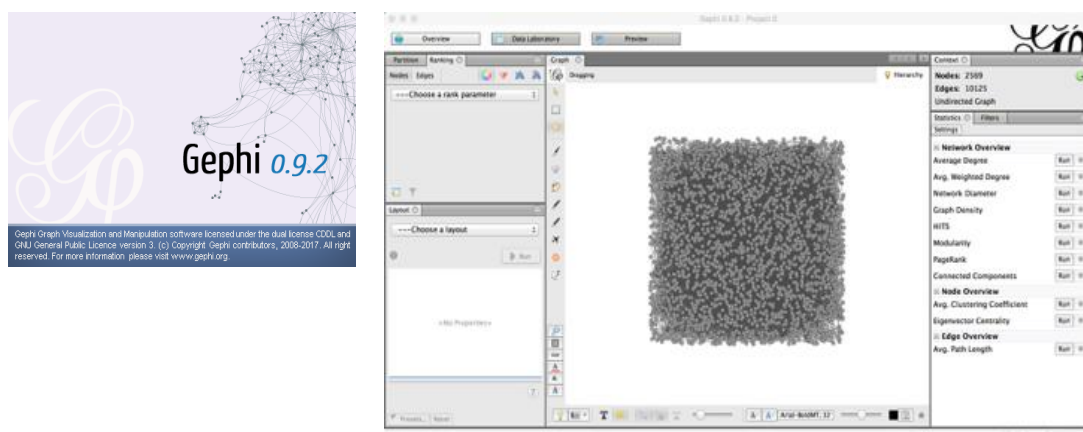
Πίνακας 3: Βασικά στοιχεία υπό μελέτη δικτύου (πλήθος κόμβων, συνδέσμων και άθροισμα βαρών)

## 5. ΜΕΘΟΔΟΛΟΓΙΑ – ΑΝΑΛΥΣΗ – ΑΠΟΤΕΛΕΣΜΑΤΑ

### 5.1. Γενικά στοιχεία

Η ανάλυση δικτύων δημιούργησε την ανάγκη της δημιουργίας προγραμμάτων για την επεξεργασία των δεδομένων τους (κόμβων, συνδέσμων κ.λπ.). Ένα δίκτυο μπορεί να αποτελείται από ελάχιστους κόμβους μέχρι ένα τεράστιο αριθμό κόμβων. Για το λόγο αυτό αναπτύχθηκαν διάφορα εργαλεία για την ανάλυση αυτών των δικτύων. Μία από τις λειτουργίες των εργαλείων αυτών είναι η οπτική αναπαράσταση των σχέσεων που δημιουργούνται μεταξύ των κόμβων του δικτύου. Τα εργαλεία αυτά δέχονται στην είσοδο τους διαφόρους τύπους δεδομένων που χρησιμοποιούνται για την αναγνώριση, ανάλυση, οπτικοποίηση και προσομοίωση των στοιχείων του δικτύου.

Πλήθος προγραμμάτων αναπτύχθηκε για το σκοπό αυτό.<sup>14</sup> Στην παρούσα εργασία το πρόγραμμα που αξιοποιήθηκε είναι το *Gephi*. Πρόκειται για πρόγραμμα ανοικτού κώδικα *gephi* (<http://gephi.org>).<sup>15</sup>



Εικόνα 18: Πρόγραμμα ανοικτού κώδικα *gephi*<sup>16</sup>

Τα δεδομένα που εισήχισαν σε 2 σύνολα/πακέτα (*datasets*) σε μορφή *excel*. Το ένα (*dataset*) αφορούσε αποκλειστικά κόμβους (*id*, *label*) όπου συμπεριελήφθηκε -σε πρόσθετη στήλη- για έκαστοτε ερευνητή, το ινστιτούτο προέλευσης (προ συγχώνευσης, 3 κατηγορίες). Το άλλο (*dataset*) πρόκειται για τον πίνακα

<sup>14</sup> Ενδεικτικά αναφέρονται τα ακόλουθα (με αλφαβητική σειρά): Gephi, InFlow, Jung, NetMiner, NodeXL, ORA, Pajek, Stanet, Stocnet, Ucinet κ.λπ.

<sup>15</sup> Εκτελείται σε περιβάλλον Windows, Mac OS X, και Linux.

<sup>16</sup> Η επιφάνεια εργασίας του προγράμματος διαιρείται σε 3 στήλες. Στη μέση βρίσκεται η οπτική αναπαράσταση του δικτύου, η οποία αναδιατάσσεται ανάλογα με τις εκάστοτε ρυθμίσεις στην αριστερή στήλη. Από τη δεξιά στήλη πραγματοποιούνται διάφορες στατιστικές επεξεργασίες (*metrics*) ή/και τίθενται φίλτρα για την αναδιαμόρφωση της οπτικοποίησης του δικτύου.

συνδεσιμότητας (*adjacency matrix*) ο οποίος μετασχηματίστηκε σε λίστα συνδεσιμότητας (*adjacency list*), για να εισαχθεί στο πρόγραμμα.

Προφανώς το 1<sup>ο</sup> σύνολο δεδομένων είναι κοινό για κάθε χρονική περίοδο (δηλ. 2008-2012 και 2013-2017), ενώ διαφοροποιείται η λίστα συνδεσιμότητας (μία για κάθε χρονική περίοδο).

Id	Label	timeset	institute	Source	Target	Type	Weight
1	1		IMEL	1	13	undirected	1
2	2		IPC	1	35	undirected	1
3	3		IMEL	1	80	undirected	9
4	4		IPC	2	24	undirected	4
5	5		IPC	2	50	undirected	4
6	6		IMS	2	92	undirected	2
7	7		IMS	2	112	undirected	2
8	8		IMS	3	8	undirected	2
9	9		IMEL	3	13	undirected	4
10	10		IMS	3	15	undirected	24
11	11		IMEL	3	18	undirected	2

Εικόνα 19: Απόσπασμα συνόλων δεδομένων (datasets) που εισήχθησαν στο *gerphi* προς επεξεργασία

Περαιτέρω σημειώνεται ότι, πέραν των χρονικών μεταβολών που εξετάστηκαν (εξέλιξη δικτύου-*network evolution*), το δίκτυο αναλύθηκε σε 3 επίπεδα, ως εξής:

- επίπεδο κόμβου/ερευνητή,
- επίπεδο κοινοτήτων (υποομάδων/υποσυνόλων εντός του δικτύου), βάσει:
  - ✓ της προέλευσης των ερευνητών, δηλ. της αρχικής διοικητικής διάρθρωσης του INN
  - ✓ της συνεργασίας που αναπτύχθηκε μεταξύ των ερευνητών, δηλ. της ύπαρξης συνεκτικών κοινοτήτων/υποομάδων εντός του δικτύου<sup>17</sup>, και
- επίπεδο δικτύου (σύνολο) – INN.

Άλλα επίπεδα ανάλυσης, τα οποία δεν αποτελούν αντικείμενο μελέτης της παρούσας, θα μπορούσαν να είναι τα ακόλουθα:

- η συνεργασία/κοινές δημοσιεύσεις σε επίπεδο ερευνητικού κέντρου («Δημόκριτος»), δηλ. πέραν του INN,
- η συνεργασία/κοινές δημοσιεύσεις σε επίπεδο χώρας, μεταξύ ομοειδών/συναφών ινστιτούτων, ερευνητικών κέντρων ή πανεπιστημίων, και τέλος
- η συνεργασία/κοινές δημοσιεύσεις σε ακόμα ευρύτερο επίπεδο (π.χ. ευρωπαϊκό ή παγκόσμιο).

(*πρόκειται κατ' ουσίαν για δίκτυα δικτύων*).

<sup>17</sup> Κοινότητες δηλ. εντός των οποίων οι κόμβοι παρουσιάζουν μεταξύ τους συχνούς/δυνατούς/έντονους δεσμούς.

Βάσει των παραπάνω, η παρούσα ενότητα διαρθρώνεται ως εξής:

- αρχικά παρατίθενται ποσοτικά στοιχεία (π.χ. κόμβοι, σύνδεσμοι, πυκνότητα, κατανομή βάρθρωσης, μέση απόσταση κόμβων κ.λπ.) για το σύνολο του ινστιτούτου, για κάθε μία χρονική περίοδο (συγκρίσεις – μεταβολές), καθώς επίσης και σχετικά διαγράμματα,
- εν συνεχεία, εξετάζονται οι σύνδεσμοι εντός των αρχικών ινστιτούτων αλλά και οι σύνδεσμοι μεταξύ διαφορετικών ινστιτούτων, διακρίνοντας τους σε 2 κατηγορίες, δηλ. *intra* (για συνδέσμους εντός του ίδιου ινστιτούτου) και *inter* (για συνδέσμους μεταξύ κόμβων/ερευνητών διαφορετικών ινστιτούτων), με συμπερίληψη των όποιων μεταβολών στο χρόνο.
- τέλος, εστιάζοντας στη β' χρονική περίοδο (2013-2017) πραγματοποιείται σύγκριση μεταξύ των επιμέρους ινστιτούτων, βάσει της αρχικής τους διάρθρωσης (ποσοτικά στοιχεία, όπως στην πρώτη περίπτωση, διαγράμματα κ.λπ.).

## 5.2. Η εικόνα του ινστιτούτου – εξέλιξη (σύγκριση περιόδων)

Όπως αναφέρθηκε, αρχικά καταγράφηκαν τα βασικά ποσοτικά στοιχεία για κάθε χρονική περίοδο, που αφορούν στο σύνολο του δικτύου, ως εμφανίζεται στον ακόλουθο πίνακα.

Ποσοτικά στοιχεία-μετρικές/χρονική περίοδος	2008-2012	2013-2017
Γενικές πληροφορίες		
▪ Κόμβοι ( <i>nodes</i> )	133	133
▪ Σύνδεσμοι ( <i>edges</i> )	500	657
Επισκόπηση του δικτύου ( <i>Network Overview</i> )		
▪ Μέση βάρθρωση ( <i>average degree</i> )	7,519	9,88
▪ Μέση σταθμισμένη βάρθρωση ( <i>average weighted degree</i> )	27,429	36,165
▪ Διάμετρος δικτύου ( <i>network diameter</i> )	6	6
▪ Πυκνότητα δικτύου ( <i>graph density</i> )	0,057	0,075
▪ Διαμέριση δικτύου ( <i>δείκτης modularity</i> )	0,671	0,628
✓ Πλήθος κοινοτήτων ( <i>Number of communities</i> )	13	13
Επίσκοπηση των κόμβων ( <i>Node Overview</i> )		
▪ Μέσος συντελεστής συσσωμάτωσης ( <i>average clustering coefficient</i> )	0,58	0,632
✓ Πλήθος τριαδικών σχέσεων ( <i>total triangles</i> )	735	<b>1533</b>
Επισκόπηση των συνδέσμων ( <i>Edge Overview</i> )		
▪ Μέση απόσταση κόμβων/μήκος μονοπατιού ( <i>average path length</i> )	2,966	2,693

Πίνακας 4: Βασικά ποσοτικά χαρακτηριστικά του δικτύου (σύγκριση περιόδων)

Συγκρίνοντας τα παραπάνω μεγέθη κατά τις 2 χρονικές περιόδους παρατηρείται ότι η αύξηση των συνδέσμων, όπως αναφέρθηκε παραπάνω, της τάξεως 31,4% συνοδεύεται από αυξήσεις/βελτίωση των βασικότερων ποσοστικών στοιχείων του δικτύου, όπως της μέσης βάρθρωσης (αντίστοιχα και της μέσης σταθμισμένης), της πυκνότητας του δικτύου, του μέσου συντελεστή συσσωμάτωσης (και των τριαδικών σχέσεων, όπου παρατηρείται περίπου διπλασιασμός αυτών). Παράλληλα παρατηρείται μείωση (βελτίωση) της μέσης απόστασης κόμβων, ενώ η διάμετρος του δικτύου παραμένει αμετάβλητη.

Ωστόσο, αξίζει να αναφερθεί ότι παρατηρείται μία μικρή μείωση της τιμής της modularity, με το πλήθος βέβαια των κοινοτήτων να παραμένει σταθερό.

Περαιτέρω, αξίζει να αναφερθεί τα εξής:

- Σύμφωνα με την έρευνα του Newman (βλ. κεφ. 3, αντίστοιχο πίνακα με α/α...) η μέση βάρθρωση των δικτύων των βιολόγων, των φυσικών και των μαθηματικών είναι 18,1, 9,7 και 3,9 αντίστοιχα. Συνεπώς το δίκτυο της παρούσας εργασίας, με τιμές μέσης βάρθρωσης 7,519 (2008-2012) και 9,88 (2013-2017) φαίνεται να προσομοιάζει με δίκτυο των φυσικών (δευτερευόντως των μαθηματικών), γεγονός που τελεί σε αντιστοιχία με τα επιμέρους ερευνητικά πεδία του INN (εντάσεως φυσικών, δευτερευόντως μαθηματικών και τέλος βιολόγων).
- Αντίστοιχα, ο μέσος συντελεστής συσσωμάτωσης του δικτύου της παρούσας εργασίας-INN (0,58-0,632) βρίσκεται πιο κοντά στην αντίστοιχη τιμή του δικτύου των φυσικών (0,43), εν συγκρίσει με τα δίκτυα των μαθηματικών (0,15) ή των βιολόγων (0,0066).
- Περαιτέρω, ως προς την μέση απόσταση κόμβων του δικτύου της παρούσας εργασίας-INN (2,966-2,693), αυτή φέρεται να βρίσκεται πιο κοντά στην αντίστοιχη τιμή του δικτύου των μαθηματικών (4,6), εν συγκρίσει με τα δίκτυα των φυσικών (5,9) ή των βιολόγων (7,6).

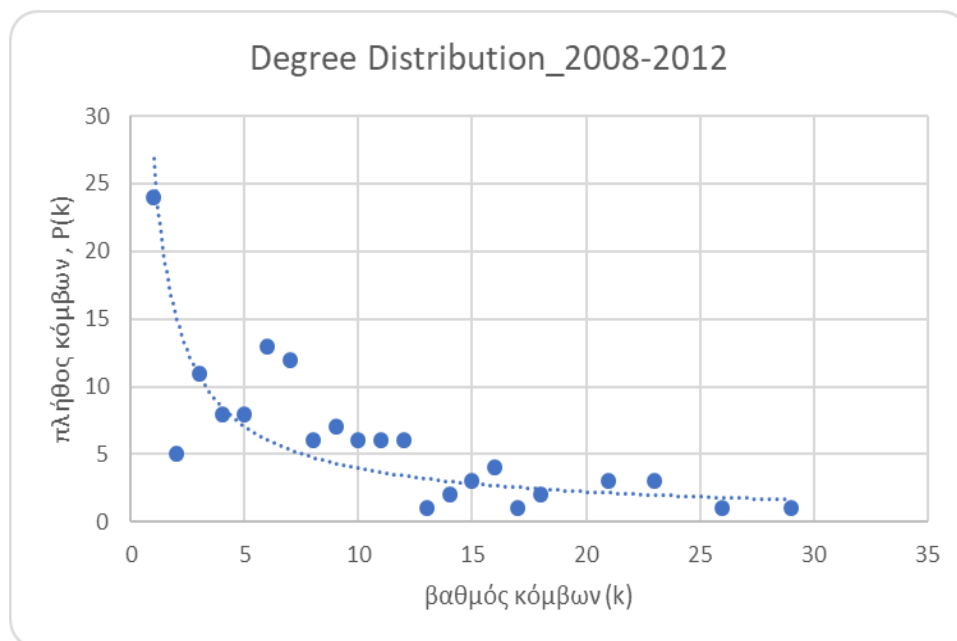
Οι παραπάνω ομοιότητες ή διαφορές στις τιμές, σχετίζονται σε μεγάλο βαθμό με τα ιδιαίτερα χαρακτηριστικά των επιμέρους επιστημονικών πεδίων, όπως αναφέρθηκε στο κεφ. 3 (π.χ. σε θέματα διεξαγωγής της έρευνας, συνεργασίας κ.λπ.).

Σε κάθε περίπτωση, πρέπει να σημειωθεί η επίδραση του μεγέθους του δικτύου στους παραπάνω δείκτες, δεδομένου ότι τα παραπάνω δίκτυα που μελετήθηκαν από τον Newman είναι πολύ μεγαλύτερα (δίκτυο βιολόγων: 1.520.251, δίκτυο μαθηματικών: 253.39 και δίκτυο φυσικών: 52.909) εν συγκρίσει με το δίκτυο που μελετάται στην παρούσα (133).

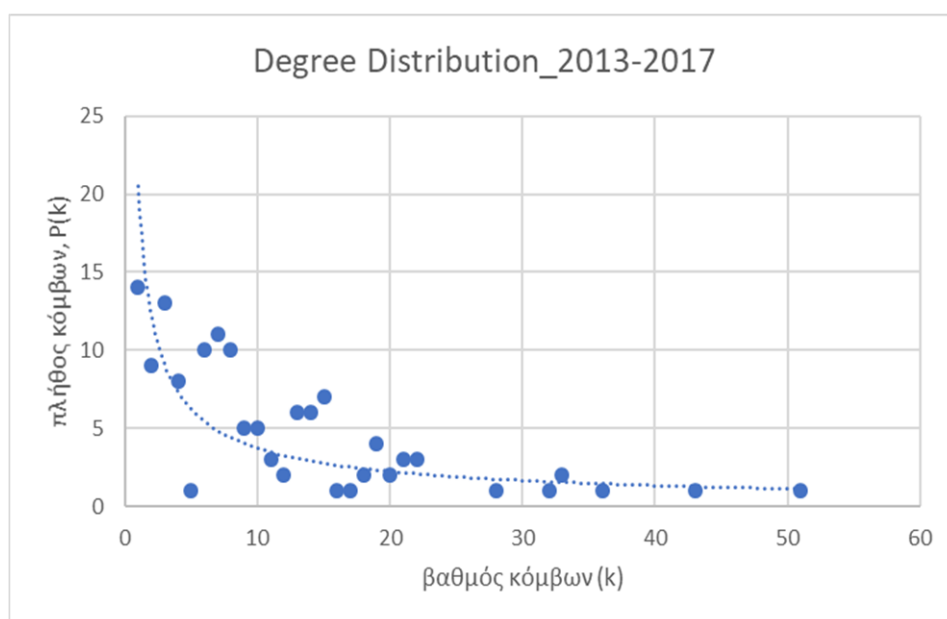
5.2.1. Κατανομή βάθμωσης, *Degree distribution*

Όπως αναφέρθηκε παραπάνω (βλ. κεφ. 2), η κατανομή βάθμωσης χαρακτηρίζει συνολικά το δίκτυο, ενώ η μορφή που λαμβάνει, σε συνδυασμό με άλλα ποσοτικά στοιχεία, δύναται να μας βοηθήσει στην ταξινόμηση/κατηγοριοποίηση του δικτύου (π.χ. *small world*, *scale free* κ.λπ.).

Ακολούθως παρατίθενται σχετικά διαγράμματα της κατανομής βάθμωσης του υπό μελέτη δικτύου για αμφότερες τις χρονικές περιόδους.



Διάγραμμα 1: Κατανομή βάθμωσης του υπό μελέτη δικτύου, για την α' χρονική περίοδο (2008-2012)



Διάγραμμα 2: Κατανομή βάθμωσης του υπό μελέτη δικτύου, για την β' χρονική περίοδο (2013-2017)



Και στις δύο περιπτώσεις οι κατανομές κόμβων παρουσιάζουν «μακριές ουρές» τύπου νόμου δύναμης. Η μαθηματική έκφραση αυτών των νόμων είναι η εξής :

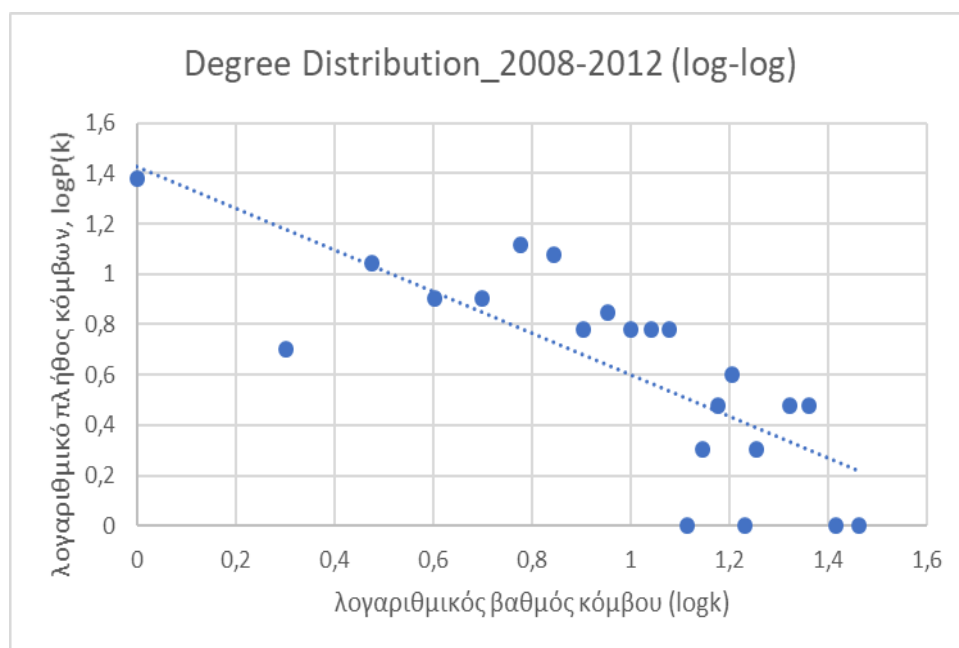
$$P(k)=k^{-\gamma}$$

όπου  $\gamma$  είναι ο χαρακτηριστικός εκθέτης της κατανομής.

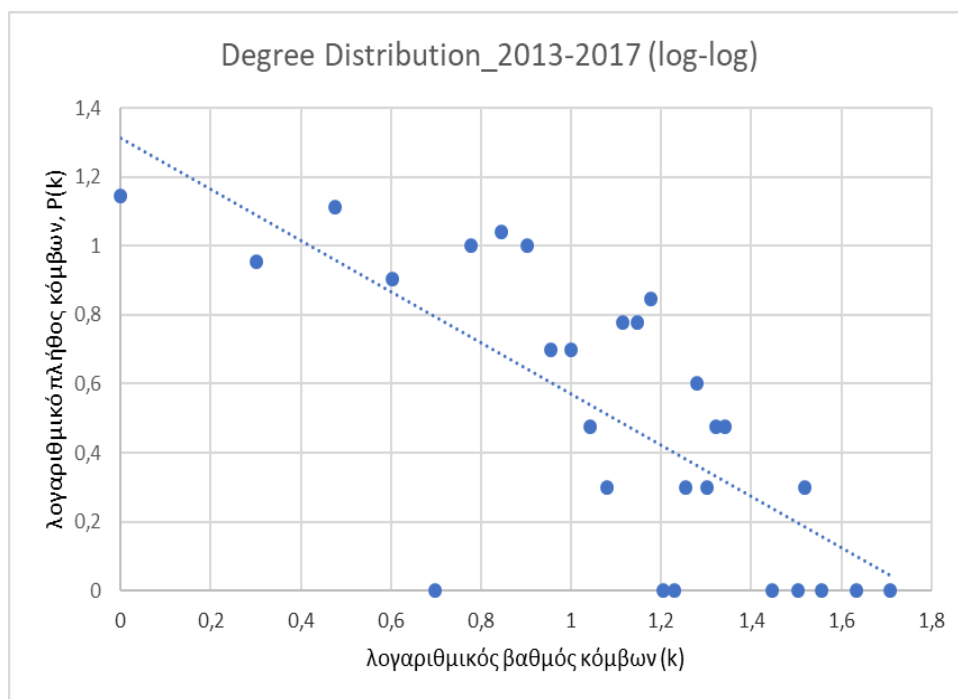
Η εμφάνιση του νόμου δύναμης συνδέεται με την ύπαρξη κόμβων μεγάλου βαθμού, γνωστά ως “hubs”, τα οποία φέρουν μεγάλο αριθμό συνδέσεων, με αποτέλεσμα τη δημιουργία ενός μικρού αριθμού κόμβων με μεγάλη συνδεσιμότητα. Από την άλλη, ένα μεγάλο μέρος των κόμβων φέρει μικρό αριθμό συνδέσεων. Τα παραπάνω συναντώνται στα δίκτυα του πραγματικού κόσμου και εκφράζονται κυρίως μέσω του μοντέλου των Barabasi-Albert (BA), δηλ. των δικτύων ελεύθερης κλίμακας (scale-free networks).

Επιπρόσθετα, βάσει του ως άνω μοντέλου (BA), οι νεοεισερχόμενοι κόμβοι ή/και κόμβοι με μικρό αριθμό συνδέσεων προτιμούν να συνδέονται με κόμβους που έχουν ήδη ισχυρή δικτύωση (προτιμητέα διασύνδεση-preferential attachment), γεγονός το παρατηρείται και στο υπό μελέτη δίκτυο (τα hubs της α' περιόδου αυξάνουν την δικτύωσή τους κατά τη β' περίοδο).

Για την παρουσίαση της κατανομής των βαθμών των κόμβων, θα ήταν προτιμότερο να χρησιμοποιηθεί η λογαριθμική κλίμακα, η οποία ενδείκνυται για περιπτώσεις όπου οι τιμές των δεδομένων έχουν μεγάλη διακύμανση, όπως στη προκειμένη περίπτωση.

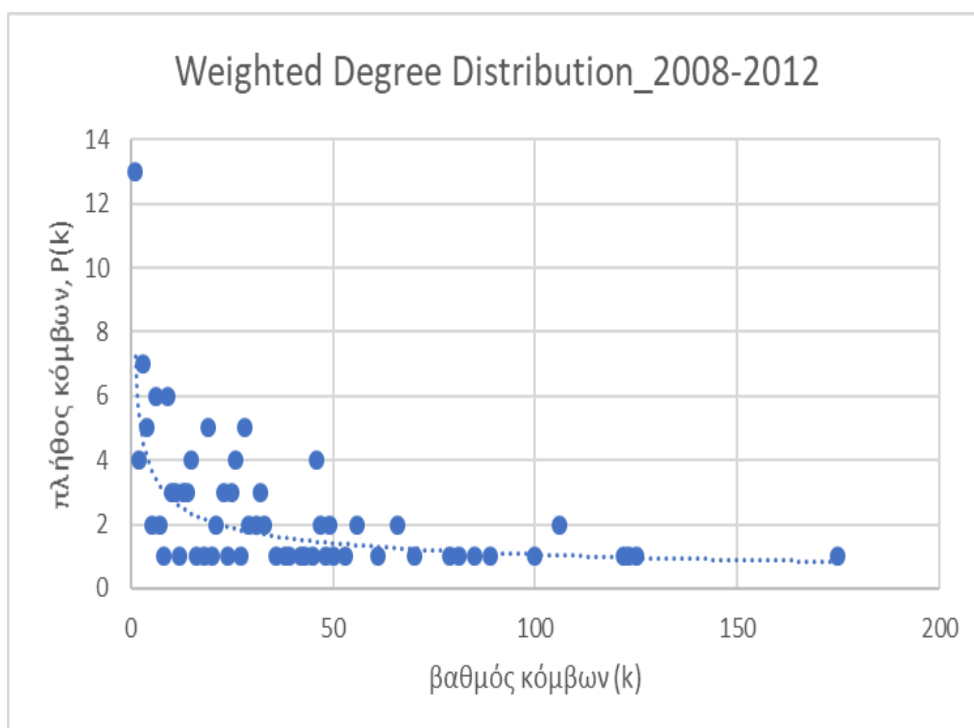


Διάγραμμα 3: Κατανομή βάθμωσης σε λογαριθμικούς άξονες του υπό μελέτη δικτύου, για την α' χρονική περίοδο (2008-2012)

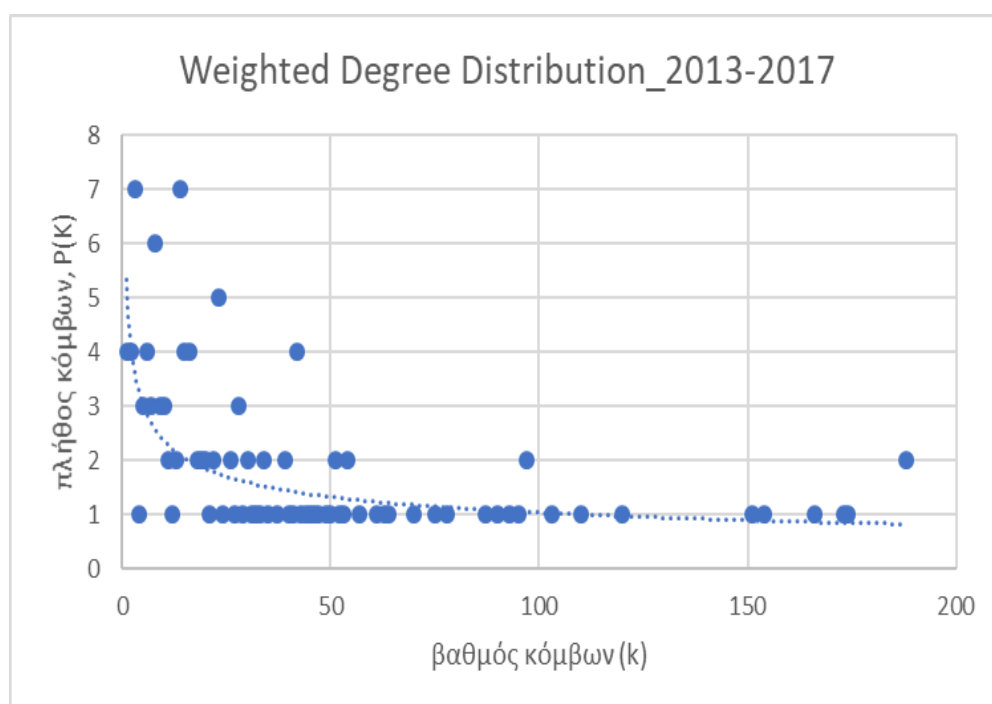


Διάγραμμα 4: Κατανομή βάρθρωσης σε λογαριθμικούς άξονες του υπό μελέτη δικτύου, για την β' χρονική περίοδο (2013-2017)

Επιπρόσθετα παρατίθενται αντίστοιχα διαγράμματα ως προς το σταθμισμένο δίκτυο, για κάθε χρονική περίοδο, τα οποία βρίσκονται σε συστοιχία με όσα προαναφέρθηκαν.



Διάγραμμα 5: Κατανομή σταθμισμένης βάρθρωσης του υπό μελέτη δικτύου, για την α' χρονική περίοδο (2008-2012)



Διάγραμμα 6: Κατανομή σταθμισμένης βάρωσης του υπό μελέτη δικτύου, για την β' χρονική περίοδο (2013-2017)

### 5.2.2. Μέτρα κεντρικότητας (*centrality measurements*)

Όπως έχει αναφερθεί τα μέτρα κεντρικότητας δίνουν μια ολοκληρωμένη εικόνα της σπουδαιότητας/κρίσιμότητας κάποιων κόμβων σε επίπεδο δικτύου. Ειδικότερα,

- μέσω της κεντρικότητας βαθμού (*degree*) αναδεικνύονται οι πιο ενεργοί κόμβοι, δηλ. οι κόμβοι με μεγάλο αριθμό συνδέσμων, με σημαντική δομικά θέση μέσα στο δίκτυο και αυξημένη πρόσβαση ή/και επιρροή σε άλλους,
- μέσω της κεντρικότητας μεσολάβησης (*betweenness*) αναδεικνύονται οι κρίσιμοι ενδιάμεσοι κόμβοι, οι οποίοι λειτουργούν ως δεσμοί ανάμεσα σε 2 υπο-ομάδες και
- μέσω της κεντρικότητας εγγύτητας (*closeness*) αναδεικνύονται οι κόμβοι που βρίσκονται κοντά σε έτερους κρίσιμους κόμβους.

Εφαρμόζοντας τα μέτρα κεντρικότητας και σταθμίζοντάς τα ως προς την κεντρικότητα βαθμού στο δίκτυο, προκύπτουν τα παρακάτω στοιχεία για τους πρώτους 20 κόμβους. (έχουν αξία τα τόσα πολλά δεκαδικά ψηφία?)

Κόμβος (Id)	Degree	Weighted Degree	Betweenness Centrality	Closeness Centrality
8	29	79	1326,198888	0,48249
3	26	123	956,85786	0,450909
24	23	56	1745,098121	0,508197
103	23	56	1002,863693	0,450909
25	23	48	886,698201	0,441281

Κόμβος (Id)	Degree	Weighted Degree	Betweenness Centrality	Closeness Centrality
67	21	70	862,229403	0,467925
43	21	106	656,535847	0,418919
95	21	106	257,860764	0,403909
30	18	31	758,153259	0,441281
69	18	66	459,102284	0,43662
82	17	50	247,449929	0,423208
92	16	125	242,618817	0,4
94	16	175	240,14103	0,398714
35	16	122	124,950758	0,369048
114	16	81	124,027417	0,371257
129	15	46	226,837427	0,397436
93	15	89	111,575421	0,357349
20	15	47	68,832652	0,3769
17	14	39	170,091595	0,409241
18	14	42	91,995074	0,392405

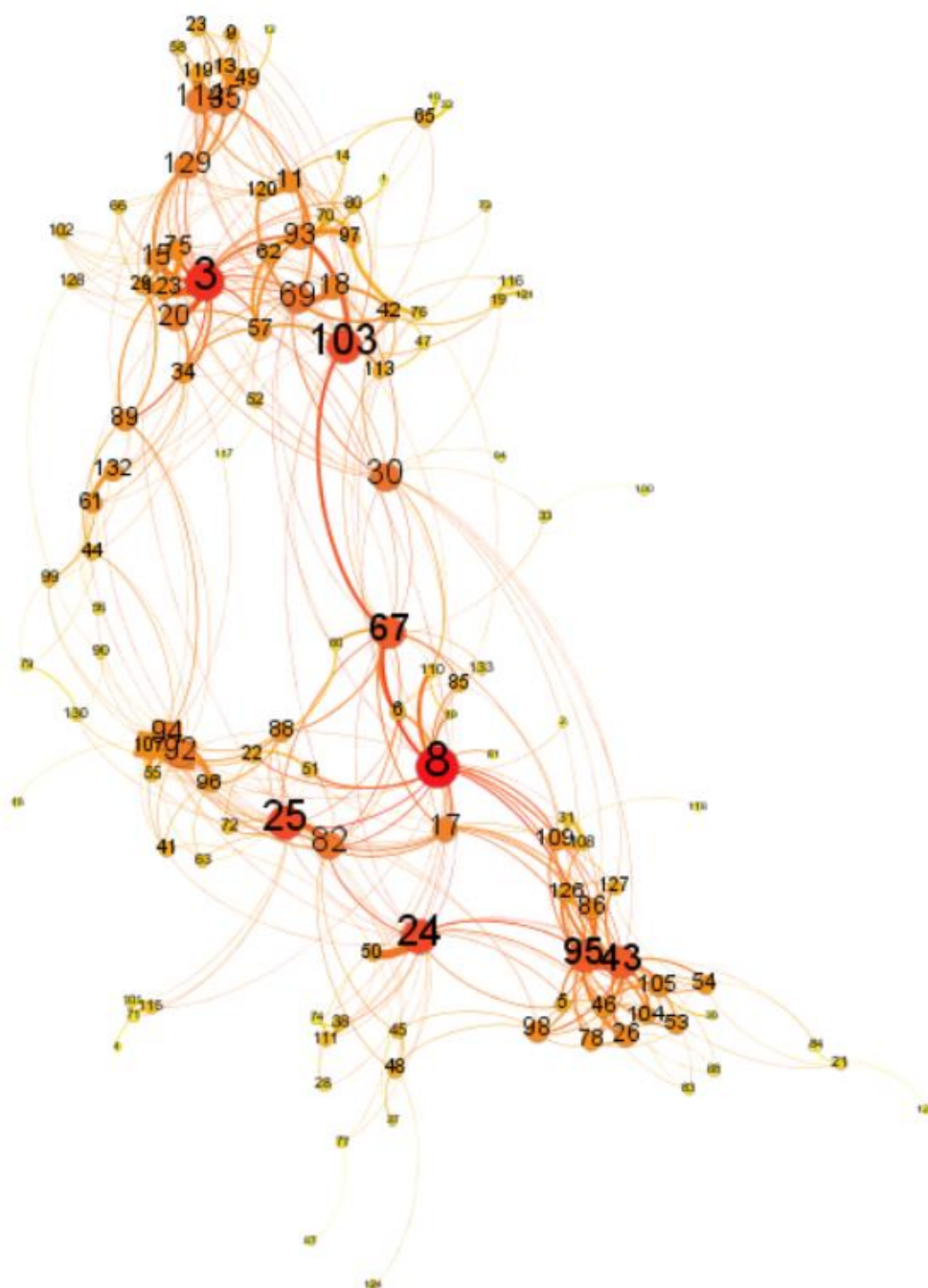
Πίνακας 5: Μέτρα κεντρικότητας (centrality measurements) του υπό μελέτη δικτύου, για την α' περίοδο (2008-2012) – Απόσπασμα σχετικού πίνακα (αποτύπωση των πρώτων 20-top20, βάσει degree

Αντιπαραβάλλοντας τα στοιχεία του πίνακα, παρατηρούμε ότι δεν υφίσταται η ίδια φθίνουσα σειρά σε όλα τα μεγέθη. Καταρχάς προκύπτουν διαφοροποιήσεις μεταξύ των μεγεθών *degree* και *weighted degree*, δηλ. ως προς την κεντρικότητα βαθμού αλλά και περαιτέρω της κεντρικότητας βαθμού με την κεντρικότητα μεσολάβησης ή/και εγγύτητας. Χαρακτηριστικά αναφέρονται τα ακόλουθα:

- υψηλότερο βαθμό κόμβων (*degree*) παρουσιάζει ο κόμβος με α/α 8,
- λαμβάνοντας υπόψη και τα βάρη των συνδέσμων (*weighted degree*), η μεγαλύτερη τιμή αντιστοιχεί στον κόμβο με α/α 94.
- ο κόμβος με την υψηλότερη τιμή κεντρικότητας μεσολάβησης (*betweenness centrality*) είναι ο με α/α 24.

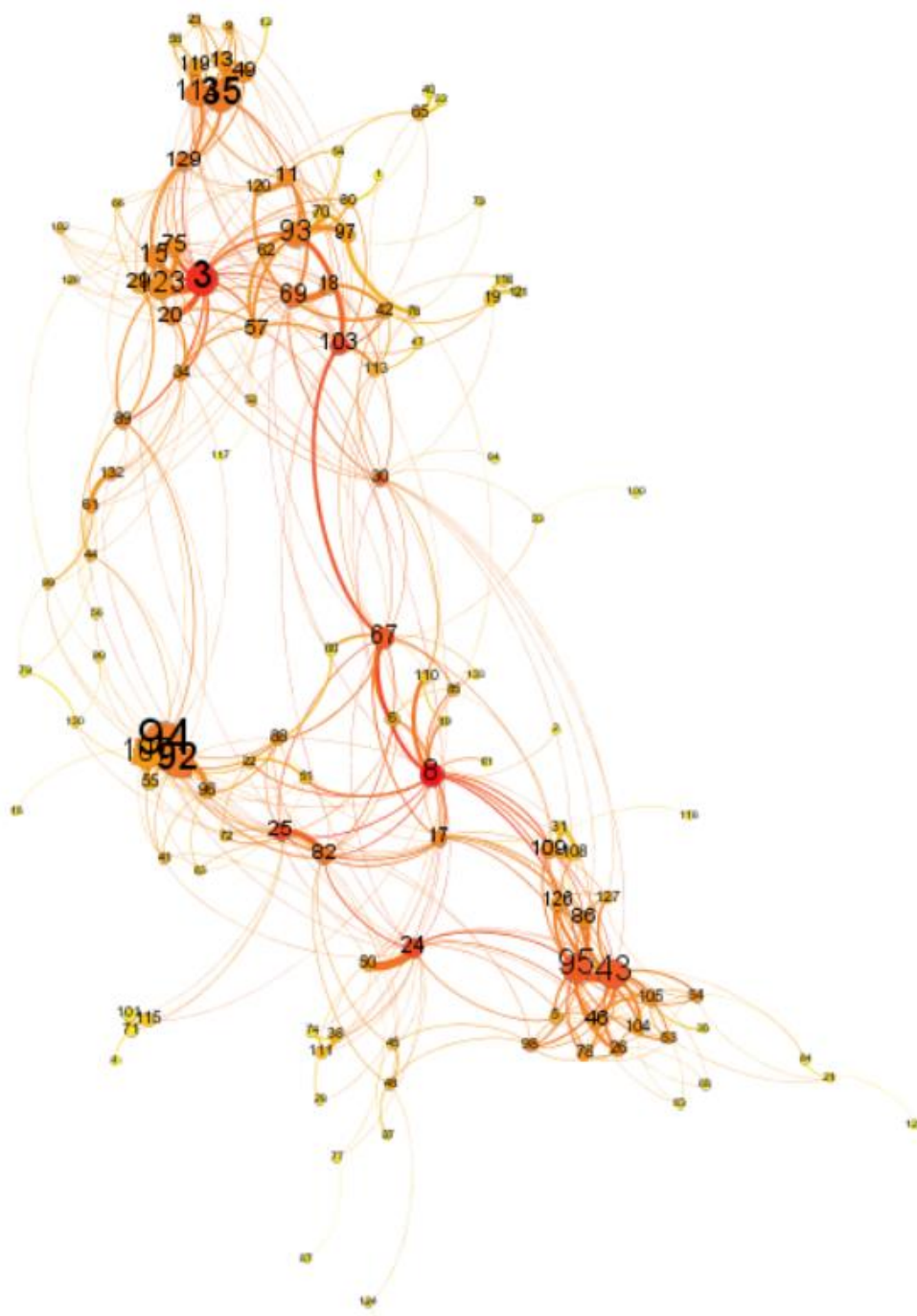
Στη συνέχεια παρουσιάζονται σχετικά γραφήματα, συγκριτικά αποτελέσματα της κεντρικότητας βαθμού (*degree & weighted degree*) και μεσολάβησης αντίστοιχα, προκειμένου να αναδειχθούν οι διαφορές που προκύπτουν μεταξύ των κόμβων. Οι κόμβοι, με την υψηλότερη εκάστοτε τιμή, αποτυπώνονται με μεγαλύτερο μέγεθος και εντονότερο χρώμα.

Κατά αυτόν τον τρόπο είναι πιο εμφανής ο ρόλος των διαφόρων κόμβων εντός του δικτύου, βάσει όσο προαναφέρθηκαν.



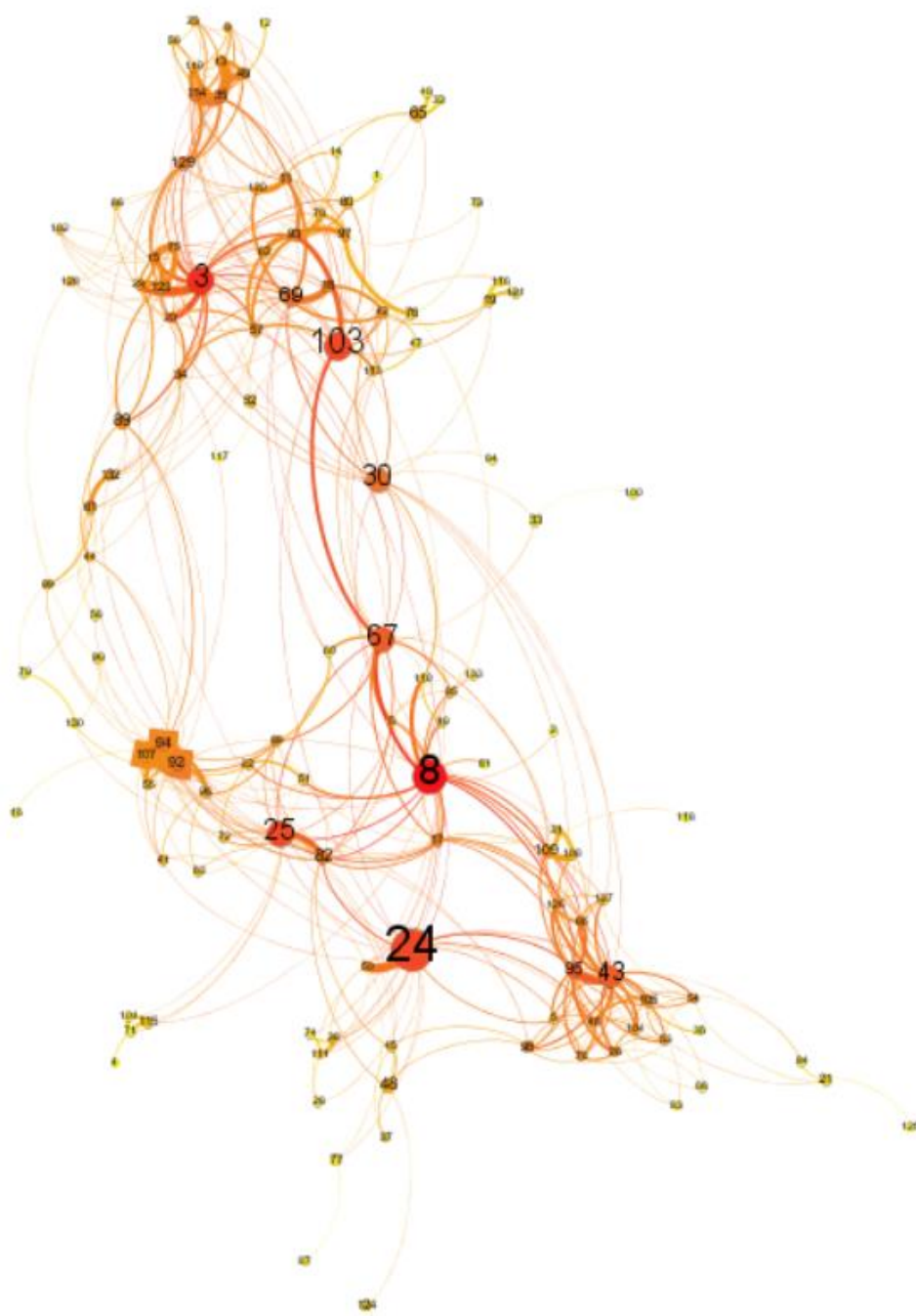
Διάγραμμα 7: Οπτικοποίηση του υπό μελέτη δικτύου της α' περιόδου, βάσει βαθμών κόμβων (degree centrality)

(ξεχωρίζουν οι κόμβοι με α/α 8, 3, 24, 103, 25, 67, 43 και 95)



Διάγραμμα 8: Οπτικοποίηση του υπό μελέτη δικτύου της α' περιόδου, βάσει σταθμισμένων βαθμών κόμβων (weighted degree centrality)

(ξεχωρίζουν οι κόμβοι με α/α 94, 92, 35, 3, 43 και 45)



Διάγραμμα 9: Οπτικοποίηση του υπό μελέτη δικτύου της α' περιόδου, βάσει κεντρικότητας μεσολάβησης (betweenness centrality)

(ξεχωρίζουν οι κόμβοι με α/α 24, 8, 103 και 3)

Σημειώνεται ότι δεν συμπεριλαμβάνεται αντίστοιχο διάγραμμα για *Closness Centrality*, δεδομένου ότι δεν παρέχει ιδιαίτερη πληρόφορηση.

Κατά τον ίδιο τρόπο εξήχθησαν αντίστοιχα διαγράμματα ως προς τις βασικές μετρικές κεντρικότητας, για την β' χρονική περίοδο. Παρατίθεται ο συγκεντρωτικός

πίνακας με τα σχετικά στοιχεία, για τους πρώτους 20 κόμβους, σταθμίζοντάς τα ως προς την κεντρικότητα βαθμού στο δίκτυο.

Κόμβος (Id)	Degree	Weighted Degree	Betweenness Centrality	Closness Centrality
8	51	174	1692,554508	0,572727
24	43	154	1111,247677	0,547826
92	36	188	1081,832456	0,518519
50	33	103	679,12621	0,522822
103	33	54	624,281728	0,504
95	32	151	650,210838	0,492188
3	28	173	658,756872	0,496063
15	22	166	155,69157	0,479087
35	22	97	393,133923	0,446809
86	22	42	69,335974	0,461538
123	21	188	130,101595	0,477273
29	21	97	206,995924	0,477273
43	21	57	39,523541	0,435986
19	20	70	90,085114	0,456522
17	20	44	228,999797	0,443662
102	19	120	89,155996	0,47191
46	19	75	88,046168	0,454874
98	19	40	44,079147	0,435986
30	19	39	142,039613	0,464945
20	18	95	119,486475	0,47191

Πίνακας 6: Μέτρα κεντρικότητας (centrality measurements) του υπό μελέτη δικτύου, για την β' περίοδο (2013-2017) – Απόσπασμα σχετικού πίνακα (αποτύπωση των πρώτων 20-top20, βάσει degree

Επιμέρους συγκρίσεις μεταξύ των δεικτών κεντρικότητας, παρατηρούνται και στον ως άνω πίνακα. Επί παραδείγματι, δεν υφίσταται ταύτιση ούτε ακολουθείται η ίδια φθίνουσα σειρά των δεικτών, κυρίως δε παρατηρείται διαφοροποίηση μεταξύ των δεικτών κεντρικότητας βαθμού (*degree & weighted degree*). Περαιτέρω οι δείκτες κεντρικότητας μεσολάβησης (*betweenness centrality*) και εγγύτητας (*closness centrality*) φαίνεται να συμβαδίζουν.

Επιπρόσθετα, δύναται να πραγματοποιηθούν και συγκρίσεις των ως άνω δεικτών για τις 2 χρονικές περιόδους, όπου επιβεβαιώνεται το γεγονός ότι τα hub της α' περιόδου ενισχύουν τη δικτύωσή τους. Χαρακτηριστικά αναφέρεται ο κόμβος με α/α 8, ο οποίος παραμένει 1<sup>ος</sup> στην κεντρικότητα βαθμού-*degree*, ενώ πλέον καταγράφει και την υψηλότερη τιμή κεντρικότητας μεσολάβησης (*betweenness centrality*). Παράλληλα καταγράφεται στην 3<sup>η</sup> θέση στην τιμή *weighted degree*, εν αντιθέσει με την α' περίοδο, όπου βρίσκονταν χαμηλότερα.



5.2.3. Συντελεστής Συσσωμάτωσης (*Clustering coefficient*)

Όπως αναφέρθηκε παραπάνω ο μέσος συντελεστής συσσωμάτωσης αυξήθηκε μεταξύ των 2 χρονικών περιόδων (0,58  $\rightarrow$  0,63). Μάλιστα οι τριαδικές σχέσεις μεταξύ των κόμβων (που προφανώς επίσης αυξήθηκαν), σχεδόν διπλασιάστηκαν.

Περαιτέρω, παρατίθενται οι 20 πρώτοι κόμβοι του δικτύου, βάσει του συντελεστή συσσωμάτωσης και των τριαδικών σχέσεων. Υπενθυμίζεται ότι το εν λόγω μέτρο δείχνει πόσο ισχυρή είναι η συνδεσιμότητα γύρω από ένα κόμβο.

Κόμβος (Id)	clustering	triangles
3	0,267692	87
95	0,357143	75
43	0,342857	72
8	0,172414	70
103	0,221344	56
69	0,339869	52
25	0,205534	52
35	0,425	51
82	0,367647	50
114	0,408333	49
24	0,193676	49
67	0,214286	45
20	0,419048	44
46	0,763636	42
15	0,538462	42
92	0,341667	41
94	0,341667	41
26	0,606061	40
75	0,590909	39
18	0,428571	39

α' χρονική περίοδος (2008-2012)

Κόμβος (Id)	clustering	triangles
8	0,218039	278
24	0,241417	218
95	0,358871	178
103	0,276515	146
50	0,265152	140
3	0,367725	139
92	0,195238	123
15	0,532468	123
123	0,566667	119
43	0,542857	114
86	0,47619	110
29	0,519048	109
102	0,625731	107
98	0,573099	98
19	0,5	95
46	0,54386	93
121	0,647059	88
20	0,542484	83
35	0,333333	77
78	0,503268	77

β' χρονική περίοδος (2013-2017)

Πίνακας 7: Στοιχεία συσπείρωσης για κάθε χρονική περίοδο

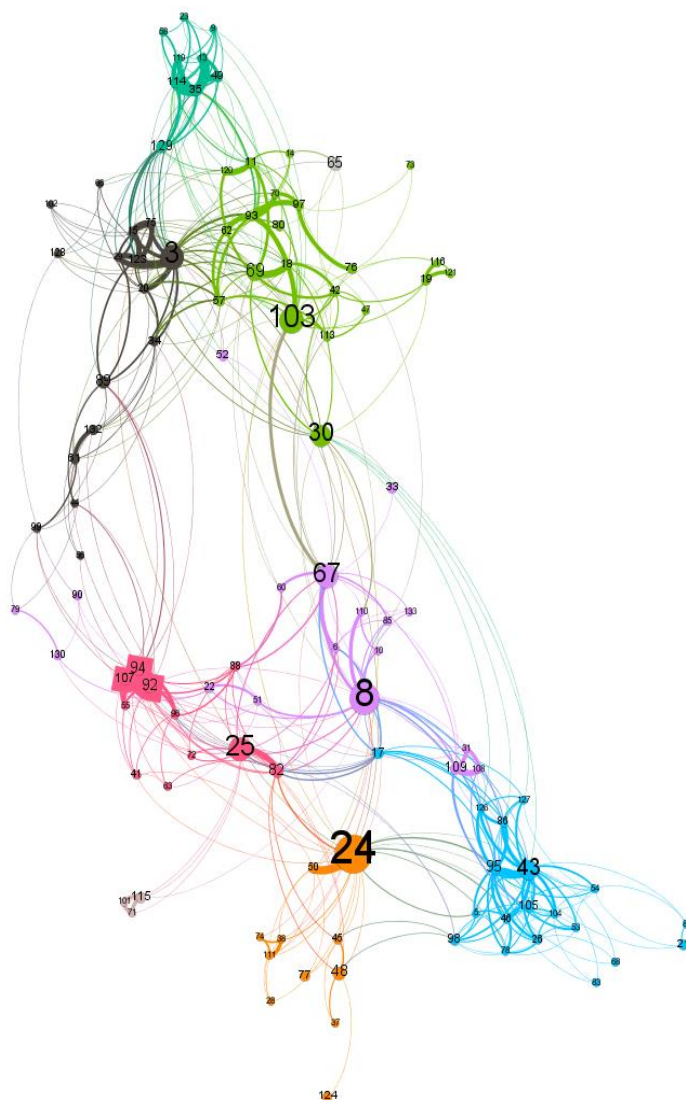
Από τη συνδυαστική επισκόπηση των παραπάνω πινάκων καθώς και άλλων που προηγήθηκαν ως προς τα μέτρα κεντρικότητας, γίνεται αντιληπτό ότι κρίσιμοι κόμβοι για το δίκτυο (π.χ. με υψηλές τιμές *degree*, *betweenness centrality* κ.λπ.) παροσουσιάζουν επίσης υψηλό αριθμό τριαδικών σχέσεων (π.χ. κόμβοι 3, 8, 24, 95 κ.λπ.). Ωστόσο δεν παρατηρείται πλήρη ταύτιση (σχετίζεται με την κατάσταση που επικρατεί σε επιμέρους «γειτονιές»).

Περαιτέρω συγκρίνοντας τα παραπάνω μεγέθη κατά τις 2 χρονικές περιόδους παρατηρούνται επιμέρους διαφοροποιήσεις ως προς την φθίνουσα σειρά εμφάνισης των κόμβων, βάσει του πλήθους των τριαδικών σχέσεων.

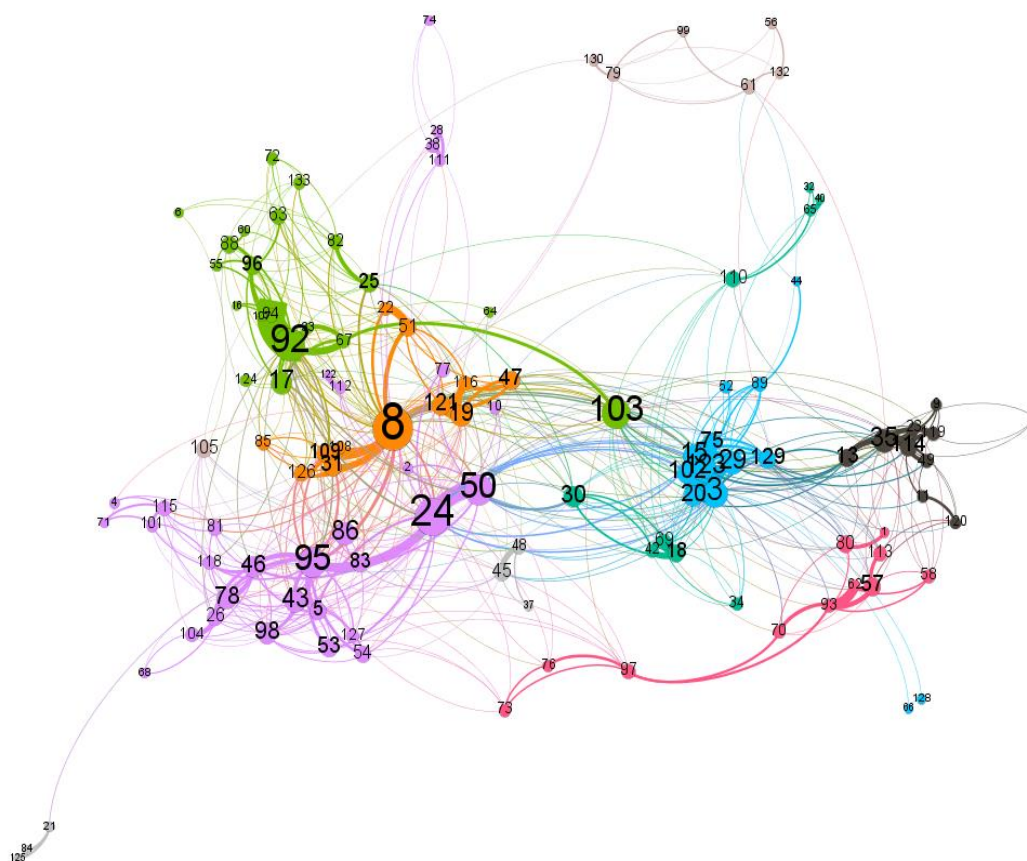
5.2.4. Διαμέριση δικτύου – Εύρεση κοινοτήτων (*modularity*)

Μία από τις σημαντικότερες εκδοχές στη μελέτη των δικτύων είναι η ανάλυση των συνεκτικών υποομάδων κόμβων (κοινοτήτες) που ενυπάρχουν στο πλαίσιο ενός δικτύου.

Όπως αναφέρθηκε, η εύρεση των κοινοτήτων στο υπό μελέτη δίκτυο, πραγματοποιήθηκε κάνοντας χρήση του δείκτη *modularity*, ενώ σε κάθε χρονική περίοδο αναγνωρίστηκαν 13 κοινότητες, όπως εμφανίζονται στα ακόλουθα γραφήματα.

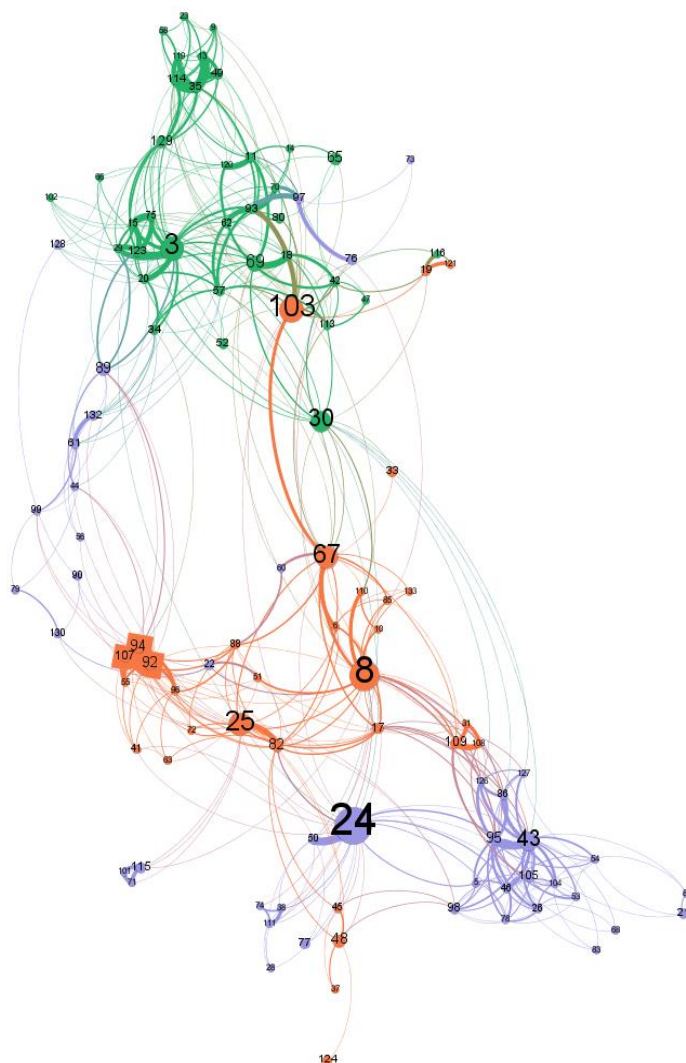


Διάγραμμα 10: Οπτικοποίηση του υπό μελέτη δικτύου της α' περιόδου (2008-2012), βάσει του δείκτη *modularity*



Διάγραμμα 11: Οπτικοποίηση του υπό μελέτη δικτύου της β' περιόδου (2013-2017), βάσει του δείκτη modularity

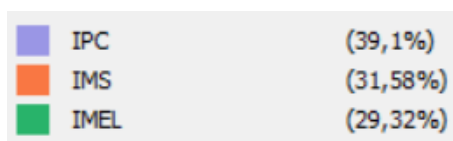
Δεδομένου και της επόμενης ενότητας που θα ακολουθήσει, κρίνεται σκόπιμο να παρουσιασθούν αντίστοιχα διαγράμματα, με κριτήριο διαμέρισης των κόμβων και αναγνώρισης κοινοτήτων, βάσει της προέλευσης/αρχικής κατανομής των ερευνητών σε 3 ισοτιμιά, τα οποία εν συνεχεία συγχωνεύτηκαν σε ένα (INN).



Διάγραμμα 12: Οπτικοποίηση του υπό μελέτη δικτύου της α' περιόδου (2008-2012), βάσει της προέλευσης/αρχικής κατανομής των ερευνητών

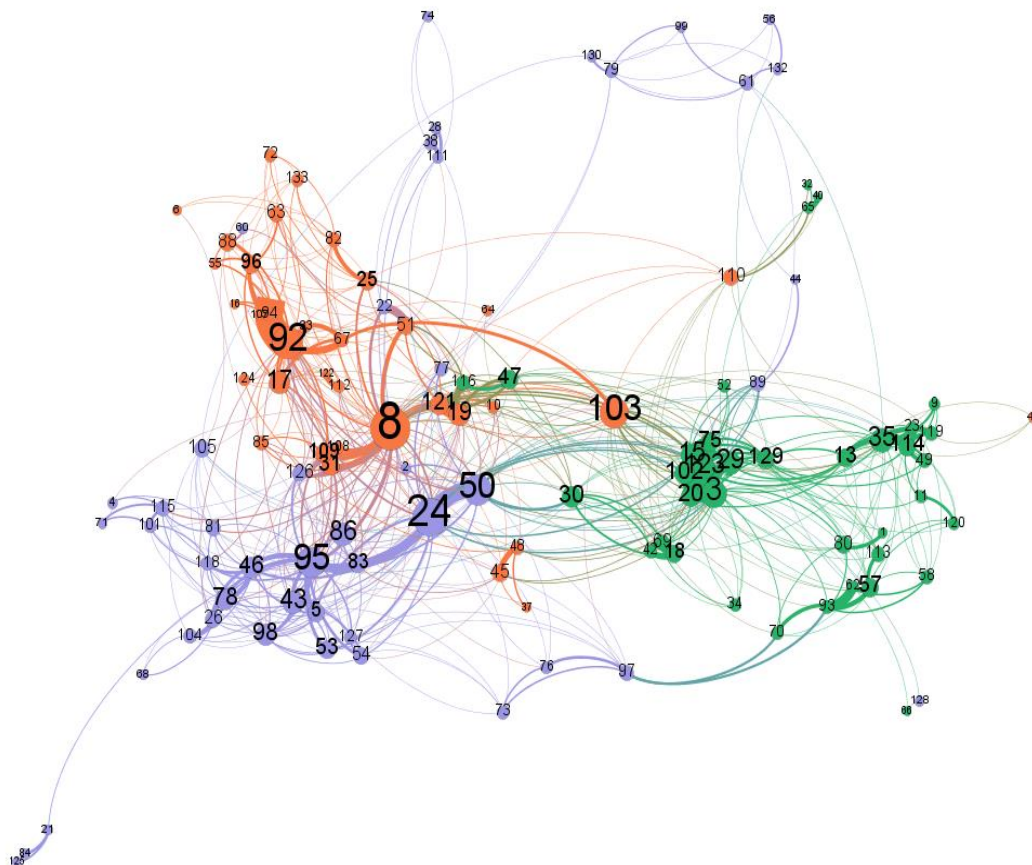
Σημεώνεται ότι η κατά την ως άνω αναπαράσταση των κόμβων, δεν συμπεριελήθησαν κόμβοι με βαθμό μικρότερο από 2 (degree range=2), ούτως ώστε να είναι πιο ευανάγνωστο.

Τα 3 πρώην ισοτιούτα αναπαρίστανται με διαφορετικό χρωματισμό, ενώ τα έναντι μεγέθη αφορούν στην αντιπροσώπευσή τους εντός του δικτύου (βλ. παρακάτω). Προφανώς ισχύει για αμφότερες τις περιόδους, βάσει υπόθεσης (σταθερός αριθμός κόμβων).



Εικόνα 20: Πρώην ισοτιούτα (αρχικές δομές)

Αντίστοιχα ως προς τη β' χρονική περίοδο.



Διάγραμμα 13: Οπτικοποίηση του υπό μελέτη δικτύου της β' περιόδου (2013-2017), βάσει της προέλευσης/αρχικής κατανομής των ερευνητών

### 5.3. Συγκριτική επισκόπηση των πρώην ινστιτούτων, κατά τις 2 περιόδους

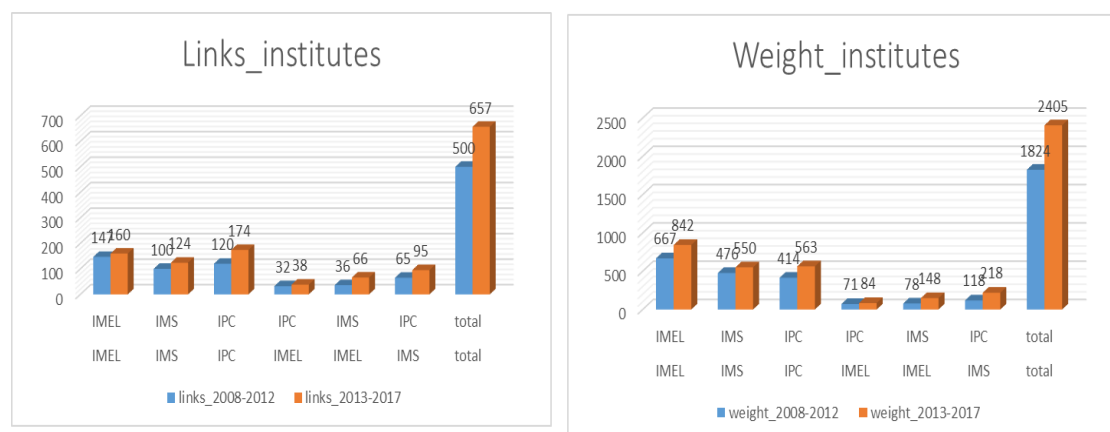
#### 5.3.1. Κύρια σύγκριση

Στην παρούσα ενότητα εξετάζονται συγκριτικά, κάποια βασικά ποσοτικά στοιχεία ως προς τα 3 πρώην ινστιτούτα.

Σε πρώτη φάση παρατίθενται στοιχεία σε απόλυτα μεγέθη ως προς τον αριθμό των συνδέσμων (*links*) και των βαρών αυτών (*weights*), ανά πρώην ινστιτούτο και ανά χρονική περίοδο (τόσο εσωτερικά των ινστιτούτων όσο και μεταξύ αυτών). Όλα τα μεγέθη καταγράφουν άνοδο (βλ. σχετικό πίνακα και διαγράμματα).

institute_1	institute_2	links_2008-2012	links_2013-2017	weight_2008-2012	weight_2013-2017
IMEL	IMEL	147	160	667	842
IMS	IMS	100	124	476	550
IPC	IPC	120	174	414	563
IMEL	IPC	32	38	71	84
IMEL	IMS	36	66	78	148
IMS	IPC	65	95	118	218
total	total	500	657	1824	2405

Πίνακας 8: Απόλυτα μεγέθη συνδέσμων (*links*) και των βαρών αυτών (*weights*), ανά πρώην ινστιτούτο και ανά χρονική περίοδο



Διάγραμμα 14-15: Πλήθος συνδέσμων (*links*) και των βαρών αυτών (*weights*), ανά πρώην ινστιτούτο και ανά χρονική περίοδο

Περαιτέρω, παρατίθενται οι ποσοστιαίες μεταβολές ως προς τα παραπάνω μεγέθη.

institute_1	institute_2	links_change (%)	weight_change (%)
IMEL	IMEL	8,84%	26,24%
IMS	IMS	24,00%	15,55%
IPC	IPC	45,00%	35,99%
IMEL	IPC	18,75%	18,31%
IMEL	IMS	83,33%	89,74%
IMS	IPC	46,15%	84,75%
total	total	31,40%	31,85%

Πίνακας 9: Ποσοστιαίες μεταβολές συνδέσμων (*links*) και των βαρών αυτών (*weights*), ανά πρώην ινστιτούτο και ανά χρονική περίοδο



Διάγραμμα 16-17: Ποσοστιαίες μεταβολές συνδέσμων (links) και των βαρών αυτών (weights), ανά πρώην ινστιτούτο και ανά χρονική περίοδο

Εντός των ινστιτούτων η μεγαλύτερη αύξηση, τόσο σε συνδέσμους όσο και σε βάρη αυτών παρατηρείται στο *IPC*. Ακολουθούν κατά φθίνουσα σειρά οι μεταβολές στο *IMS* -και στο *IMEL*, ως προς τους συνδέσμους (με διαφορετική ωστόσο σειρά ως προς τα βάρη αυτών).

Περαιτέρω, μεταξύ των ινστιτούτων, παρατηρούμε ότι η μεγαλύτερη μεταβολή (αύξηση), της τάξεως πέραν του 80% αφορά στον αριθμό των συνδέσμων μεταξύ κόμβων των *IMS-IMEL* (ομοίως και ως προς τα βάρη αυτών).Επίσης σημαντική μεταβολή καταγράφεται ως προς τα αντίστοιχα μεγέθη (ιδίως ως προς τα βάρη των συνδέσμων) μεταξύ των *IPC-IMS*.

### 5.3.2. Μία άλλη προσέγγιση (σύγκριση “intra” – “inter”)

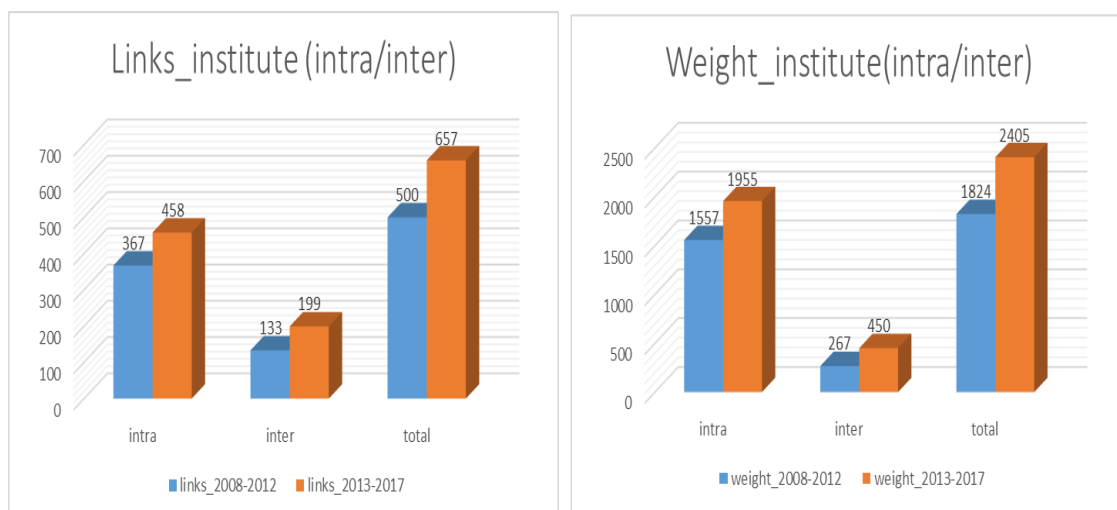
Επιπρόσθετα, κρίθηκε σκόπιμη μια έτερη κατηγοριοποίηση των συνδέσμων των κόμβων (και των βαρών αυτών αντίστοιχα). Ειδικότερα οι σύνδεσμοι ταξινομήθηκαν σε 2 κατηγορίες, ως εξής:

- *Intra*, για συνδέσμους κόμβων του ίδιου ινστιτούτου (*IMEL*, *IMS* ή *IPC*) και
- *Inter*, για συνδέσμους μεταξύ κόμβων/ερευνητών διαφορετικών ινστιτούτων

Ομοίως όπως παραπάνω, παρατίθενται σχετικοί πίνακες και διαγράμματα.

	links_2008-2012	links_2013-2017	weight_2008-2012	weight_2013-2017
intra	▼ 367	▲ 458	▼ 1557	▲ 1955
inter	▼ 133	▲ 199	▼ 267	▲ 450
total	▼ 500	▲ 657	▼ 1824	▲ 2405

Πίνακας 10: Απόλυτα μεγέθη συνδέσμων (links) και των βαρών αυτών (weights) επιμέρους ομαδοποιήσεων



Διάγραμμα 18-19: Πλήθος συνδέσμων (links) και των βαρών αυτών (weights) επιμέρους ομαδοποιήσεων

Καταρχάς πρέπει να ληφθεί ότι η πλειοψηφία των συνδέσμων (όπως και τα βάρη αυτών), σε αμφότερες τις χρονικές περιόδους, αφορά σε συνδέσμους κόμβων εντός του ίδιου ινστιτούτου (*intra*).

Ωστόσο, εξετάζοντας τη μεταβολή στο χρόνο (βλ. παρακάτω πίνακα), παρατηρούμε ότι αυτή είναι συγκριτικά μεγαλύτερη για τους συνδέσμους των κόμβων διαφορετικών ινστιτούτων (*inter*), μάλιστα της τάξεως 50% περίπου, πάνω από τη γενική μεταβολή (31,4%) την οποία φαίνεται να συμπαρασύρει.

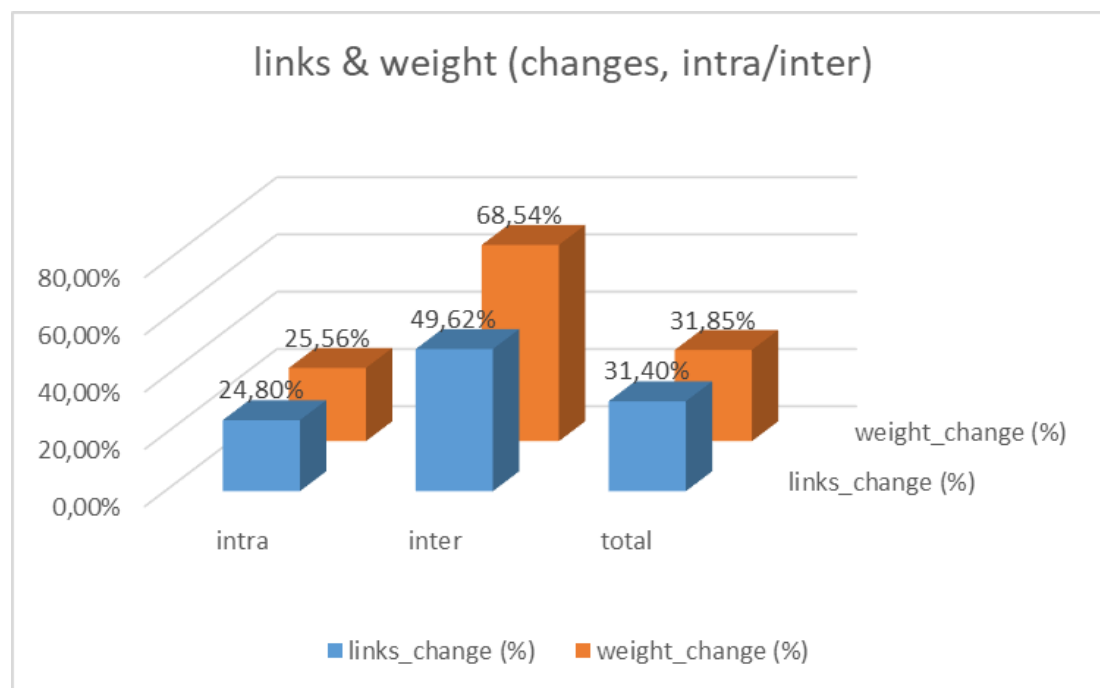
Το παραπάνω επιβεβαιώνεται από την ακόμα μεγαλύτερη μεταβολή των βαρών των συνδέσμων αυτών (*inter*), της τάξεως του 70% περίπου.

	links_change (%)	weight_change (%)
intra	24,80%	25,56%
inter	49,62%	68,54%
total	31,40%	31,85%

Πίνακας 11: Ποσοστιαίες μεταβολές συνδέσμων (links) και των βαρών αυτών (weights) επιμέρους ομαδοποιήσεων

Ακολουθεί το αντίστοιχο διάγραμμα.





Διάγραμμα 20: Ποσοστιαίες μεταβολές συνδέσμων (links) και των βαρών αυτών (weights) επιμέρους ομαδοποιήσεων

Θα πρέπει βέβαια να σημειωθεί ότι οι εν λόγω μεταβολές σχετίζονται και με τις αρχικές τιμές των παραπάνω μεγεθών (π.χ. οι “intra” σύνδεσμοι είναι σχεδόν τριπλάσιοι των “inter” συνδέσμων, κατά την α’ περίοδο).

#### 5.4. Συγκριτική επισκόπηση των πρώην ινστιτούτων ως αυτοτελή δίκτυα, κατά τη β’ περίοδο

##### 5.4.1. Κατασκευή νέων δικτύων & παραδοχές

Το τελευταίο σκέλος της ανάλυσης, εστιάζει στα επιμέρους ινστιτούτα, σύμφωνα με την αρχική διάρθρωση του INN και λαμβάνοντας υπόψη τα παραπάνω συγκριτικά στατιστικά μεγέθη, επανέρχεται στη θεωρία γράφων/δικτύων, συγκρίνοντας πλέον μέτρα κεντρικότητας κ.λπ. μεταξύ των πρώην ινστιτούτων, για την β’ χρονική περίοδο.

Ακολουθεί σχετικός πίνακας, για τον οποίο πρέπει να λεχθούν τα εξής:

- Τα στοιχεία που καταγράφονται για κάθε πρώην ινστιτούτο αφορούν στους συνδέσμους κόμβων αποκλειστικά εντός των ίδιων (αναφέρθηκαν προηγουμένως ως “intra”). Ως εκ τούτου σύνδεσμοι μεταξύ κόμβων διαφορετικών ινστιτούτων έχουν αφαιρεθεί (κάθε ινστιτούτο μελετάται ως ξεχωριστό-αποκομμένο-απομονωμένο δίκτυο, έστω  $N_1, N_2, N_3$ ).
- Λόγω του παραπάνω, το άθροισμα των επιμέρους συνδέσμων κόμβων είναι μικρότερο του συνόλου των συνδέσμων, καθώς αφορά “intra” συνδέσμους, δηλ.  $N_1 + N_2 + N_3 < N_{total}$  ( $458 < 657$ ).

Κατά αυτόν τον τρόπο επιχειρούνται να αναδείχθουν ενδεχόμενες διαφοροποιήσεις στην δομή/αρχιτεκτονική των επιμέρους δικτύων, εν συγκρίσει με όσα αναφέρθηκαν στις προηγούμενες ενότητες.

Ποσοτικά στοιχεία-μετρικές/πρώην ισοτιμότητα	IMS (N1)	IPC (N2)	IMEL (N3)	TOTAL (N)
Γενικές πληροφορίες				
▪ Κόμβοι ( <i>nodes</i> )	42	52	39	133
▪ Σύνδεσμοι ( <i>edges</i> )	124	174	160	657
Επισκόπηση του δικτύου ( <i>Network Overview</i> )				
▪ Μέση βάρθρωση ( <i>average degree</i> )	5,905	6,692	8,205	9,88
▪ Μέση σταθμισμένη βάρθρωση ( <i>average weighted degree</i> )	26,19	21,654	43,179	36,165
▪ Διάμετρος δικτύου ( <i>network diameter</i> )	4	6	5	6
▪ Πυκνότητα δικτύου ( <i>graph density</i> )	0,144	0,131	0,216	0,075
▪ Διαμέριση δικτύου ( <i>δείκτης modularity</i> )	0,511	0,48	0,524	0,628
✓ Πλήθος κοινοτήτων ( <i>Number of communities</i> )	11	11	6	13
Επίσκόπηση των κόμβων ( <i>Node Overview</i> )				
▪ Μέσος συντελεστής συσσωμάτωσης ( <i>average clustering coefficient</i> )	0,687	0,7	0,64	0,632
✓ Πλήθος τριαδικών σχέσεων ( <i>total triangles</i> )	177	316	313	1533
Επισκόπηση των συνδέσμων ( <i>Edge Overview</i> )				
▪ Μέση απόσταση κόμβων/μήκος μονοπατιού ( <i>average path length</i> )	2,137	2,67	2,249	2,693

Πίνακας 12: Βασικά ποσοτικά χαρακτηριστικά 3 νέων δικτύων και αρχικού (β' περίοδος)

#### 5.4.2. Σύγκριση των δικτύων $N_1$ , $N_2$ , $N_3$ - $N_{total}$

Συγκρίνοντας τα παραπάνω μεγέθη, δηλ. των νέων δικτύων ( $N_1$ ,  $N_2$ ,  $N_3$ ) και του αρχικού δικτύου που αφορά στο σύνολο του INN, όπως έχει περιγραφεί στις παραπάνω ενότητες, παρατηρούμε ότι η πλειοψηφία των μεγεθών είναι αρκετά διαφοροποιημένη. Ειδικότερα:

- Ο μέσος βαθμός των κόμβων των νέων δικτύων ( $N_1$ ,  $N_2$ ,  $N_3$ ) είναι μικρότερος σε σχέση με το αντίστοιχο μέγεθος του  $N_{total}$  (9,88). Στην ίδια συνάφεια κυμαίνεται και ο σταθμισμένος μέσος βαθμός κόμβων, με εξαίρεση το  $N_3$  (IMEL).
- Η διάμετρος των νέων δικτύων ( $N_1$ ,  $N_2$ ,  $N_3$ ) είναι ίδια ή μικρότερη, σε σχέση με το αντίστοιχο μέγεθος του  $N_{total}$ . Ομοίως και ως προς τη μέση απόσταση κόμβων.
- Η πυκνότητα των νέων δικτύων ( $N_1$ ,  $N_2$ ,  $N_3$ ) είναι μεγαλύτερη (διπλάσια, ακόμα και τριπλάσια, για το  $N_3$ -IMEL), σε σχέση με το αντίστοιχο μέγεθος του  $N_{total}$ .
- Ο μέσος συντελεστής συσσωμάτωσης των νέων δικτύων ( $N_1$ ,  $N_2$ ,  $N_3$ ) είναι μεγαλύτερος, σε σχέση με το αντίστοιχο μέγεθος του  $N_{total}$ .
- Η τιμή της modularity των νέων δικτύων ( $N_1$ ,  $N_2$ ,  $N_3$ ) είναι μικρότερη, σε σχέση με το αντίστοιχο μέγεθος του  $N_{total}$ . Παράλληλα, το μέγεθος των κοινοτήτων είναι μικρότερο (μάλιστα στο  $N_3$ -IMEL, είναι περίπου οι μισές).

Βάσει των προαναφερομένων, γίνεται κατανοητό ότι τα νέα δίκτυα, που είναι βεβαίως πιο μικρά σε μέγεθος είναι πιο συνεκτικά εν συγκρίσει με το  $N_{total}$ , όπως ήταν αναμενόμενο.

#### 5.4.3. Επιμέρους σύγκριση των δικτύων $N_1$ , $N_2$ , $N_3$

Συγκρίνοντας τα νέα δίκτυα μεταξύ τους  $N_1$ ,  $N_2$ ,  $N_3$  προκύπτουν τα εξής:

- ο μικρότερος μέσος βαθμός κόμβων παρατηρείται στο  $N_1$  (*IMS*), ενώ ο μεγαλύτερος στο  $N_3$  (*IMEL*). Ομοίως, ως προς τον υψηλότερο σταθμισμένο μέσο βαθμό κόμβων ( $N_3$  - *IMEL*), ενώ παρατηρείται μη αντιστοίχιση ως προς τα υπόλοιπα δίκτυα σε σχέση με τον μέσο βαθμό κόμβων.
- η μικρότερη διάμετρος παρατηρείται  $N_1$  (*IMS*). Ομοίως και ως προς την μέση απόσταση κόμβων.

Όπως έχει αναφερθεί στη θεωρία, στα στα δίκτυα του «μικρού κόσμου» (small world) παρατηρείται η μικρή σχετικά τιμή της μέσης («ελάχιστης») απόστασης κόμβων καθώς επίσης η υψηλή συσταδοποίησης (ενδεχόμενη εφαρμογή στο νέο δίκτυο  $N_1$ ).

- η μεγαλύτερη πυκνότητα παρατηρείται στο  $N_3$  (*IMEL*).
- η μεγαλύτερη τιμή του συντελεστή συσσωμάτωσης παρατηρείται στο  $N_2$  (*IPC*). ομοίως και ως προς τον υψηλότερο αριθμό τριαδικών σχέσεων.
- η μεγαλύτερη τιμή της modularity παρατηρείται στο  $N_3$  (*IMEL*), όπου επίσης καταγράφεται ο μικρότερος αριθμός κοινοτήτων.

#### **Χρήσιμη επισήμανση**

- το νέο δίκτυο  $N_3$  (*IMEL*), το οποίο φαίνεται πιο συνεκτικό -συγκριτικά με τα υπόλοιπα- έχει επίσης το μικρότερο αριθμό κόμβων, γεγονός που ενδεχομένως σχετίζεται με τα ως άνω ευρήματα, και
- τα παραπάνω ευρήματα θα πρέπει να εξετασθούν συνδυαστικά με τα ευρήματα της ενότητας (5.3.).

## 6. ΕΠΙΛΟΓΟΣ

### 6.1. Σύνοψη & Συμπεράσματα

Στην παρούσα?? εργασία παρουσιάσθηκε το θεωρητικό πλαίσιο της θεωρίας γράφων και συνακόλουθα της ανάλυσης δικτύων -στο βαθμό που κρίθηκε αναγκαίο- με περιγραφή των δομικών στοιχείων των δικτύων, των ποσοτικών τους χαρακτηριστικών, των κυριότερων κατηγοριών αυτών, χρήσιμων δεικτών (π.χ. μέτρα κεντρικότητας και εύρεσης κοινοτήτων), καθώς επίσης των κυριότερων μοντέλων, ως έχουν μελετηθεί από γνωστούς ερευνητές και προκύπτουν βάσει της διεθνούς βιβλιογραφίας.

Περαιτέρω, δεδομένου ότι αντικείμενο της εργασίας αποτέλεσε η κατασκευή και μελέτη -υπό τη πρίσμα της ανάλυσης δικτύων (*network analysis*)- ενός δικτύου ερευνητών βάσει των κοινών τους δημοσιεύσεων<sup>18</sup> στο μεγαλύτερο ινστιτούτο της Χώρας μας (INN), παρουσιάσθηκαν συνοπτικά σημαντικές μελέτες επί δικτύων ερευνητών και ακαδημαϊκών σε παγκόσμιο επίπεδο, στις οποίες αξιοποιήθηκαν διάφορες βιβλιογραφικές βάσεις δεδομένων.

Εξάλλου, όπως αναφέρθηκε η μελέτη και κατανόηση κάποιων ιδιαίτερων χαρακτηριστικών γύρω από την επιστήμη, εν προκειμένω των δικτύων ερευνητών, δύναται να έχει θετικό αντίκτυπο στην ίδια την επιστήμη, γεγονός που γίνεται ήδη αντιληπτό, δημιουργώντας ενδεχομένως ένα νέο επιστημονικό τομέα, επονομαζόμενο ως «*Επιστήμη της Επιστήμης*» - “*Science of Science*” (SOS).

Στο πλαίσιο αυτό συγκεντρώθηκαν τα απαιτούμενα στοιχεία του δικτύου που μελετήθηκε, μέσω της βιβλιογραφικής βάσης δεδομένων “Scopus”, τα οποία εν συνεχεία ταξινομήθηκαν και εν συνεχεία αναλύθηκαν μέσω του προγράμματος ανοικτού κώδικα *gerhi*.

Δεδομένου ότι κάθε δίκτυο, δύναται να μελετηθεί σε 3 επίπεδα, δηλ. α) κόμβων, β) υποομάδων/κοινοτήτων και γ) δικτύου, επιλέχθηκε η ακόλουθη προσέγγιση για το υπό μελέτη δίκτυο:

- Εξήχθησαν τα κυριότερα ποσοτικά χαρακτηριστικά του δικτύου, στα οποία έγινε αντιπαραβολή με αντίστοιχα στοιχεία άλλων μελετών/δημοσιεύσεων, σε παγκόσμιο επίπεδο, προκειμένου να εντοπισθούν κοινά σημεία, συγκλίσεις και διαφορές μεταξύ των δικτύων.

*Παρατηρείται η ομοιότητα των μετρικών με το δίκτυο φυσικών επιστημόνων του Newton, γεγονός που βρίσκεται σε αντιστοιχία με το επιστημονικό πεδίο μελέτης της πλειοψηφίας των ερευνητών του δικτύου του INN, παρ' όλο το σημαντικώς μικρότερο δείγμα (133 κόμβοι/ερευνητές).*

- Περαιτέρω, μέσω της κατανομής βάθμωσης των κόμβων και συνδυαστικά με ως άνω ποσοτικά στοιχεία, επιχειρήθηκε η ταξινόμηση του δικτύου, βάσει των κυρίαρχων μοντέλων.

<sup>18</sup> Ως ένας από τους δείκτες συνεργασίας εντός της ακαδημαϊκής/ερευνητικής κοινότητας.

Παρατηρείται η ομοιότητα του δικτύου της εργασίας με το μοντέλο των Barabasi-Albert (BA), δηλ. των δικτύων ελεύθερης κλίμακας (scale-free networks), καθώς παρατηρήθηκε η ύπαρξη ενός μικρού αριθμού κόμβων με μεγάλη συνδεσιμότητα, γνωστά ως “hubs”, σε αντιδιαστολή με την ύπαρξη ενός μεγάλου αριθμού κόμβων με μικρό αριθμό συνδέσεων. Σημειώνεται ότι το εν λόγω μοντέλο απαντάται κατά κόρον στα δίκτυα του πραγματικού κόσμου, όπως και το δίκτυο που μελετήθηκε.

Περαιτέρω, δεδομένου ότι εξετάστηκε η εξέλιξη του δικτύου στο χρόνο, μέσω 2 περιόδων, παρατηρήθηκε η εμφάνιση του φαινομένου γνωστού ως “rich get richer”, όπου ένας κόμβος μεγάλου βαθμού αυξάνει περισσότερο το βαθμό του σε σχέση με άλλους κόμβους που έχουν μικρότερο βαθμό, βάσει και της «προτιμητέας διασύνδεσης» (preferential attachment).

- Αναδείχθηκαν οι κρίσιμοι για το δίκτυο κόμβοι, με χρήση των μέτρων κεντρικότητας, ενώ εξετάστηκαν τυχόν μεταβολές στην πάροδο του χρόνου.

Διαπιστώνεται η ύπαρξη διαφορετικών ρόλων εντός του δικτύου, όπως ήταν αναμενόμενο ένεκα και του μεγέθους αυτού (π.χ. κόμβοι με μεγάλο αριθμό συνδέσεων, σε συνάφεια με τα παραπάνω, ενδιάμεσοι κόμβοι κρίσιμοι για τη συνεκτικότητα συγκεκριμένων υποομάδων/κοινοτήτων κ.λπ.).

Επίσης, δεν παρατηρήθηκαν ιδιαίτερες μεταβολές ως προς τους κόμβους με σημαντικούς ρόλους, κατά τη σύγκριση μεταξύ των 2 περιόδων (πέραν ασφαλώς της γενικής αύξησης των συνδέσεων).

- Αναδείχθηκαν οι επιμέρους υποομάδες/κοιότητες που ενυπήρχαν στο δίκτυο, μέσω της κατάλληλης διαμέρισής του. Περαιτέρω, ελήφθη υπόψη η διοικητική μεταβολή που επήλθε κατά το παρελθόν και οδήγησε στη δημιουργία του εν λόγω Ινστιτούτου, μέσω δηλ. της συγχώνευσης 3 αυτοτελών Ινστιτούτων, και εξήχθησαν σχετικά συμπεράσματα ως προς τη συνεργασία ερευνητών προερχομένων από διαφορετικές δομές και επιστημονικά πεδία, με παρατηρούμενη τάση ομογενοποίησης (βλ. και παρακάτω).
- Πραγματοποιήθηκε συγκριτική επισκόπηση των αρχικών δομών/πρώην ινστιτούτων, ως προς τα σημαντικότερα μεγέθη (σύνδεσμοι και βάρη αυτών), καθώς επίσης των σχέσεων που αναπτύχθηκαν τόσο εντός της κάθε δομής όσο και μεταξύ αυτών.

Παρατηρείται μία διαφαινόμενη βελτίωση της συνεργασίας – αύξηση κοινών δημοσιεύσεων μεταξύ ερευνητών προερχομένων από διαφορετικές δομές, εν συγκρίσει -ποσοστιαία- με τις εντός των δομών κοινές δημοσιεύσεις.

- Τέλος, εξετάστηκαν τα κυριότερα ποσοτικά μεγέθη των κόμβων (ερευνητών) που προέρχονταν από τα 3 προηγούμενα αυτοτελή ινστιτούτα κατά τη β' περίοδο, με επιμέρους συγκρίσεις αυτών αλλά και με το αρχικό-κύριο δίκτυο της μελέτης (τα ευρήματα φέρεται να συσχετίζονται, μεταξύ άλλων, με το μέγεθος των δικτύων που αποτέλεσαν αντικείμενο σύγκρισης).

## 6.2. Ενδεχόμενες επεκτάσεις

*Εκτιμάται ότι επιτεύχθη η ανάδειξη της χρησιμότητας της ανάλυσης δικτύων στην μελέτη διαφόρων τύπων δεδομένων-στοιχείων που βρίσκονται σε αλληλεξάρτηση, χρησιμοποιώντας ένα καλώς ορισμένο υπόβαθρο ορισμού των μετρικών δικτύου.*

Ωστόσο, και πέραν των όσων αναλυτικά περιγράφησαν, τα οποία αποτυπώθηκαν και με πίνακες, διαγράμματα και κατάλληλες οπτικοποιήσεις κατά περίπτωση, η εργασία θα μπορούσε ενδεχομένως να επεκταθεί προς τις ακόλουθες κατευθύνσεις (ενδεικτικά, όχι περιοριστικά):

- Στη περαιτέρω ανάλυση του εν λόγω δικτύου, με πρόσθετες μετρικές, υφιστάμενες ή/και νέες.
- Στη μελέτη ενός ευρύτερου δικτύου με κοινές δημοσιεύσεις ερευνητών, π.χ. σε επίπεδο Ερευνητικού Κέντρου, χώρας, ευρωπαϊκό κ.λπ.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

### A. Ελληνόγλωσση

- A. Γκαζή, Ι. Σκαρπέλος, *Social Media & Πολιτιστική Επικοινωνία*, Πολιτιστικό Εργαστήριο VI, ΠΑΝΤΕΙΟ
- A. Σέττας-Καφφετζής, *Ανάλυση κοινωνικών δικτύων, υπολογιστικές μέθοδοι και μέθοδοι απεικόνισης*, Μεταπτυχιακή Διπλωματική Εργασία, ΕΜΠ (2012)
- A. Προβατά, *Πολύπλοκα Δίκτυα και Εφαρμογές*, σημειώσεις (2017)
- B. Γκογκόγλου, *Ανάλυση online κοινωνικών δικτύων. Μια εφαρμογή σε δίκτυο από το Youtube*, Μεταπτυχιακή Διπλωματική Εργασία, ΑΠΘ (2016)
- K. Βαζούρα, *Ανάλυση του εργαλείου NodeXL*, Μεταπτυχιακή Διπλωματική Εργασία, ΠΑΠΕΙ (2015)
- K. Μαρμαρέλλη, *Μοντελοποίηση και ανάλυση κοινωνικών δικτύων*, Μεταπτυχιακή Διπλωματική Εργασία, ΕΜΠ (2014)
- M. Νίκου, *Εξελίξεις και μετασχηματισμοί στην Ευρωπαϊκή αγορά ηλεκτρικής ενέργειας: Εφαρμογή της Ανάλυσης Κοινωνικών Δικτύων*, Μεταπτυχιακή Διπλωματική Εργασία, ΕΜΠ (2017)
- Π. Κανακάκης, *Εύρεση κοινοτήτων σε γράφους Μεγάλων Δεδομένων με τη χρήση της πλατφόρμας Exagone*, Πτυχιακή Εργασία, ΕΚΠΑ (2016)
- Σ. Παππάς, *Ανάλυση Αποθετηρίων Πακέτων Λογισμικού με Τεχνικές Ανάλυσης Κοινωνικών Δικτύων*, Μεταπτυχιακή Διπλωματική Εργασία, ΟΠΑ (2016)
- X. Αληφραγκής, *Μελέτη και Ανάλυση Κοινωνικών Δικτύων Χρηστών Σε Ψηφιακές Βιβλιοθήκες Μαθησιακών Αντικειμένων*, Μεταπτυχιακή Διπλωματική Εργασία, ΠΑΠΕΙ (2015)

### B. Ξενόγλωσση

- A. Zeng, Z. Shen, J. Zhou, J. Wu, Y. Fan, Y. Wang, H. E. Stanley, *The science of science: From the perspective of complex systems*, *Physics Reports* 714–715 (2017) 1–73
- D. J. Watts and S. H. Strogatz, *Collective dynamics of small world networks* *Nature*, (London) vol. 393, p.440 (1998).
- E. Sarigöl, R. Pfitzner, I. Scholtes, A. Garas, F. Schweitzer, *Predicting Scientific Success Based on Coauthorship Networks*, arXiv:1402.7268v1 [physics.soc-ph] (2014)
- F. Mali, L. Kronegger, N. Ferligoj, *Co-Authorship Trends And Collaboration Patterns In The Slovenian Sociological Community*, *Corvinus Journal Of Sociology And Social Policy* Vol.1 (2010) 2, 29–50
- G. Melin, O. Persson, *Studying Research Collaboration Using Co-authorship*, *Scientometrics*, Vol. 36, No. 3 (1996) 363-377

- G. Nomikos, *Point centrality indices and ISP network vulnerability*, Master Thesis, NKUA (2013)
- L. da F. Costa, F. A. Rodrigues, G. Travieso, P. R. Villas Boas, *Characterization of Complex Networks: A Survey of measurements*, *Advances in Physics*, Volume 56, pages 167 - 242, Issue 1 (2007)
- L. Wardil, C. Hauert, *Cooperation and coauthorship in scientific publishing*, *Physical Review E* 91, 012825 (2015)
- M.E.J. Newman, *The structure of scientific collaboration networks*, *Proc. Natl. Acad. Sci. USA* 98 (2) (2001) 404–409.
- M.E.J. Newman, *Scientific collaboration networks. I. Network construction and fundamental results*, *Phys. Rev. E* 64 (2001) 016131.
- M.E.J. Newman, *Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality*, *Phys. Rev. E* 64 (2001) 016132.
- M. E. Newman, A. L. Barabasi and D. J. Watts *The Structure and Dynamics of Networks* Princeton University Press, 2006
- M.E.J. Newman, *Assortative mixing in networks*, *Phys. Rev. Lett.* 89 (2002) 208701.
- M. Girvan, M.E.J. Newman, *Community structure in social and biological networks*, *Proc. Natl. Acad. Sci. USA* 99 (12) (2002) 7821–7826.
- S.Uddin, L. Hossain, K. Rasmussen (2013) *Network Effects on Scientific Collaborations*. *PLoS ONE* 8(2): e57546. doi:10.1371/journal.pone.0057546
- T. H. Summers, F. L. Cortesi and J. Lygeros *On submodularity and controllability of Complex Dynamical Networks* *IEEE Transactions and Control* vol. 3, pp. 91-101 (2016)
- T. Papageorgiou, *Crawling Facebook: A social Network Analysis*, Master Thesis, AUEB (2011)
- T. Opsahl, V. Colizza, P. Panzarasa, J.J. Ramasco, *Prominence and control: The weighted rich-club effect*, *Phys. Rev. Lett.* 101 (2008) 168702.
- S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, *Complex networks: Structure and dynamics*, *Physics Reports* 424 (2006) 175 - 308
- U. Brandes, *A Faster Algorithm for Betweenness Centrality*, *Journal of Mathematical Sociology* 25(2):163-177, (2001).
- V. Constantoudis, N. Karasmani, *Networks and surfaces: First investigations*, presentation
- V. Colizza, A. Flammini, M.A. Serrano, A. Vespignani, *Detecting rich-club ordering in complex networks*, *Nat. Phys.* 2 (2) (2006) 110–115.
- V. D Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, *Fast unfolding of communities in large networks*, in *Journal of Statistical Mechanics: Theory and Experiment* 2008 (10), P1000