

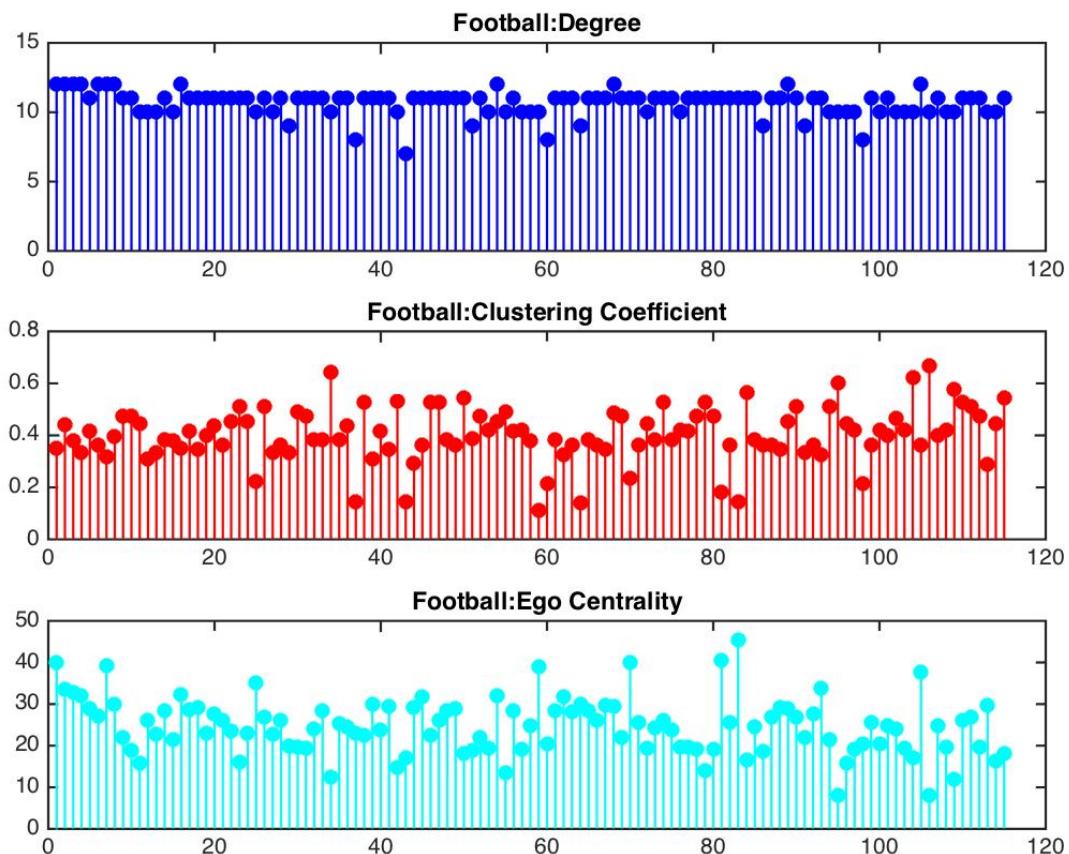
## Μέρος Α : Μελέτη Πραγματικών Τοπολογιών

Στο ερώτημα αυτό θα ασχοληθούμε με την μελέτη πραγματικών τοπολογιών και με την βοήθεια του Matlab θα εξάγουμε και θα σας παρουσιάσουμε τα εξής :

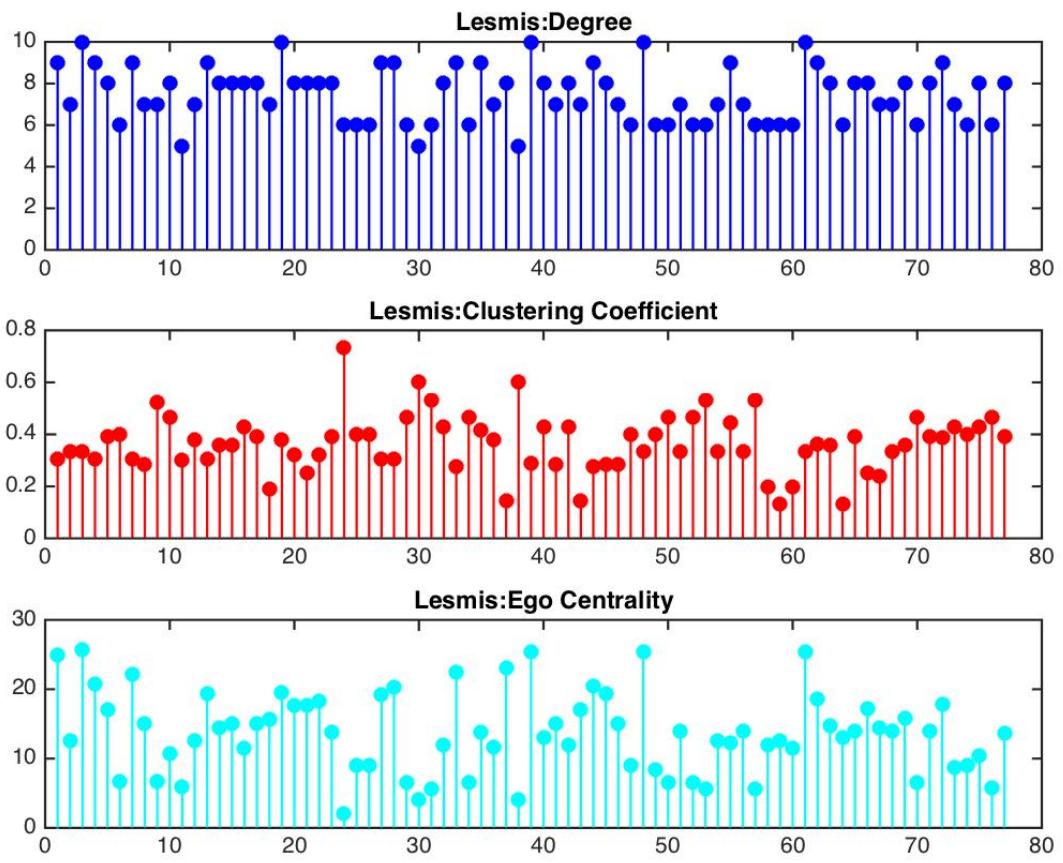
- Βαθμός κάθε κόμβου και μέσος βαθμός κάθε τοπολογίας(football, lesmis, dolphins).
- Ο συντελεστής ομαδοποίησης κάθε κόμβου και ο μέσος συντελεστής ομαδοποίησης κάθε τοπολογίας.
- Την ενδιαμεσική εγω-κεντρικότητα κάθε κόμβου και τη μέση ενδιαμεσική εγω-κεντρικότητα.

Τα αποτελέσματα τα οποία πήραμε φαίνονται παρακάτω :

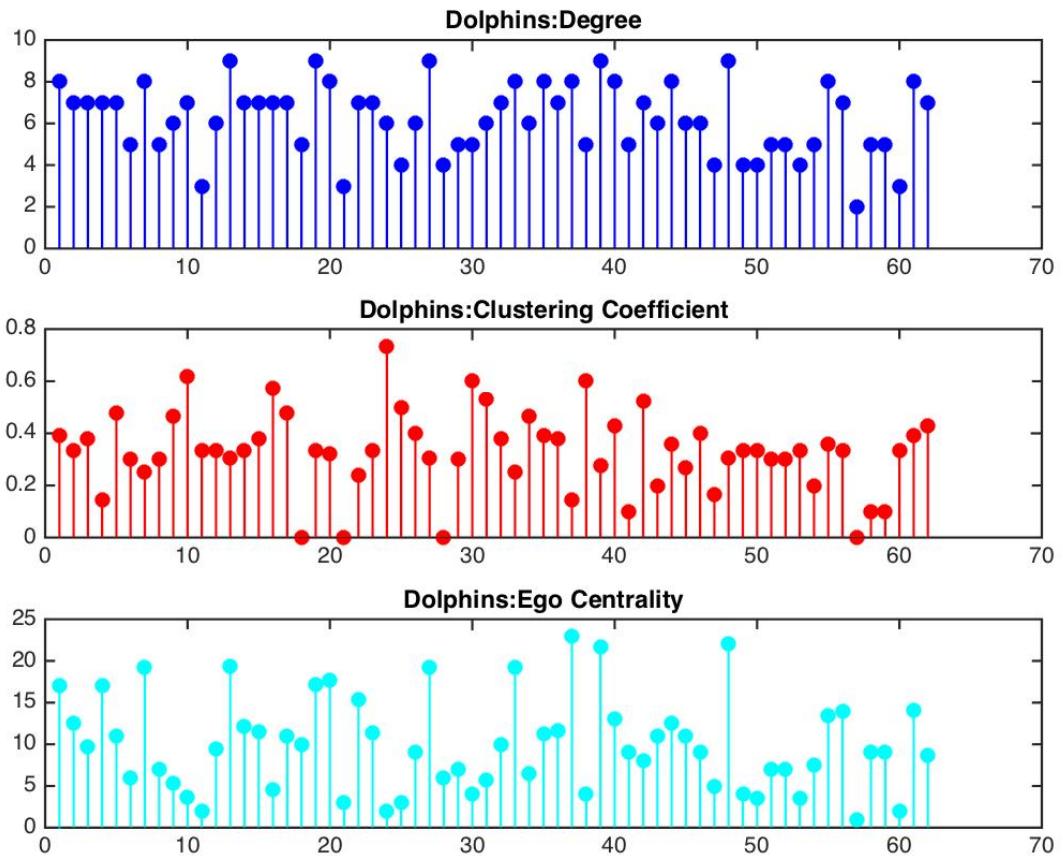
### American College Football :



## Les Miserables :



## Dolphin Social Network :



Οι μέσες τιμές που ζητούνται παρουσιάζονται συγκεντρωμένες στον παρακάτω πίνακα:

Τοπολογία Ι Μετρική	mean degree	mean CC	mean Ego
<b>Football</b>	10.6609	0.40322	24,7221
<b>Dolphins</b>	6.2258	0.33033	13,6898
<b>Lesmis</b>	7.4545	0.36546	10,0048

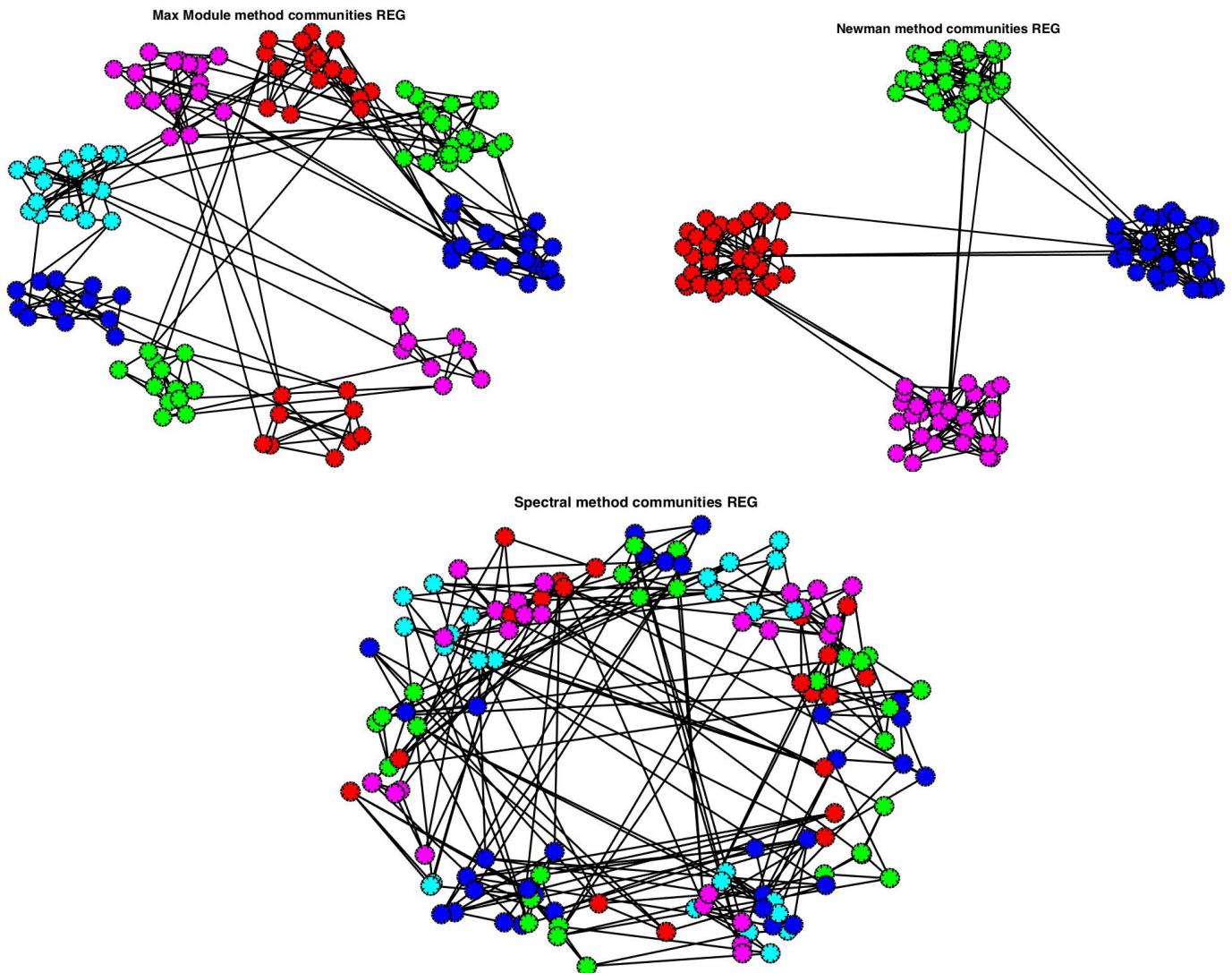
### Σχολιασμός Αποτελεσμάτων :

- Για την τοπολογία football από τα παραπάνω αποτελέσματα και γραφικές παρατηρούμε ότι μοιάζουν αρκετά με την SW τοπολογία διότι τα degree δεν έχουν μεγάλες διακυμάνσεις, το mean CC είναι πολύ κοντά με αυτό που είχαμε υπολογίσει στην 1η εργαστηριακή άσκηση για την SW και τα αποτελέσματα του ego επίσης ενισχύουν αυτό τον ισχυρισμό. Μετά από δοκιμές κατέληξα ότι αν επιλέξουμε την πιθανότητα ανασύνδεσης ίση με 0.34 τα αποτελέσματα της football είναι πολύ κοντινά με την SW. Το clustering coefficient ήταν αυτό το οποίο με έκανε να καταλήξω στη SW τοπολογία αφού είναι το σημαντικότερο για τη δομή του δικτύου.
- Για την τοπολογία lesmis παρατηρούμε ότι τα degree μοιάζουν και αυτά με την τοπολογία SW. Όσο αφορά τα αποτελέσματα των CC έχουμε ότι μοιάζουν με αυτά του SW αλλά λόγο του μικρού αριθμού κόμβων η ένδειξη αυτή δεν είναι και πολύ ασφαλής για συμπεράσματα. Τέλος τα αποτελέσματα του Ego μοιάζουν περισσότερα με αυτά του RGG. Αν επιθυμούσαμε να πάρουμε μια προσέγγιση του lesmis θα μπορούσαμε να πάρουμε την SW τοπολογία με πιθανότητα 0.48 και όχι την τοπολογία RGG διότι δίνει πολύ άσχημα αποτελέσματα στα CC κι επίσης το δίκτυο αυτό είναι ενα δίκτυο ανθρώπων που δεν είναι άγνωστοι μεταξύ τους, άρα θα μπορούσαμε να πούμε ότι δεν ταιριάζει με μοντέλα τυχαίων γράφων.
- Για την τοπολογία dolphins παρατηρούμε ότι τα degree μοιάζουν και αυτά με την τοπολογία SW. Όσο αφορά τα αποτελέσματα των CC έχουμε ότι μοιάζουν με αυτά του SW αλλά λόγο του μικρού αριθμού κόμβων η ένδειξη αυτή δεν είναι και πολύ ασφαλής για συμπεράσματα, όπως ακριβώς και με τοπολογία lesmis. Τέλος τα αποτελέσματα του Ego μοιάζουν περισσότερα με αυτά του RGG. Αν επιθυμούσαμε να πάρουμε μια προσέγγιση του lesmis θα μπορούσαμε να πάρουμε την SW τοπολογία με πιθανότητα 0.46-0.50. Επίσης θα μπορούσαμε να πούμε ότι η τοπολογία dolphin θα μπορούσε να προσεγγιστεί από SF τοπολογία με κατάλληλες παραμέτρους, κάτι το οποίο φαίνεται και από τις μετρικές.

## Μέρος Β : Εξοικοίωση με τα εργαλεία εντοπισμού κοινοτήτων

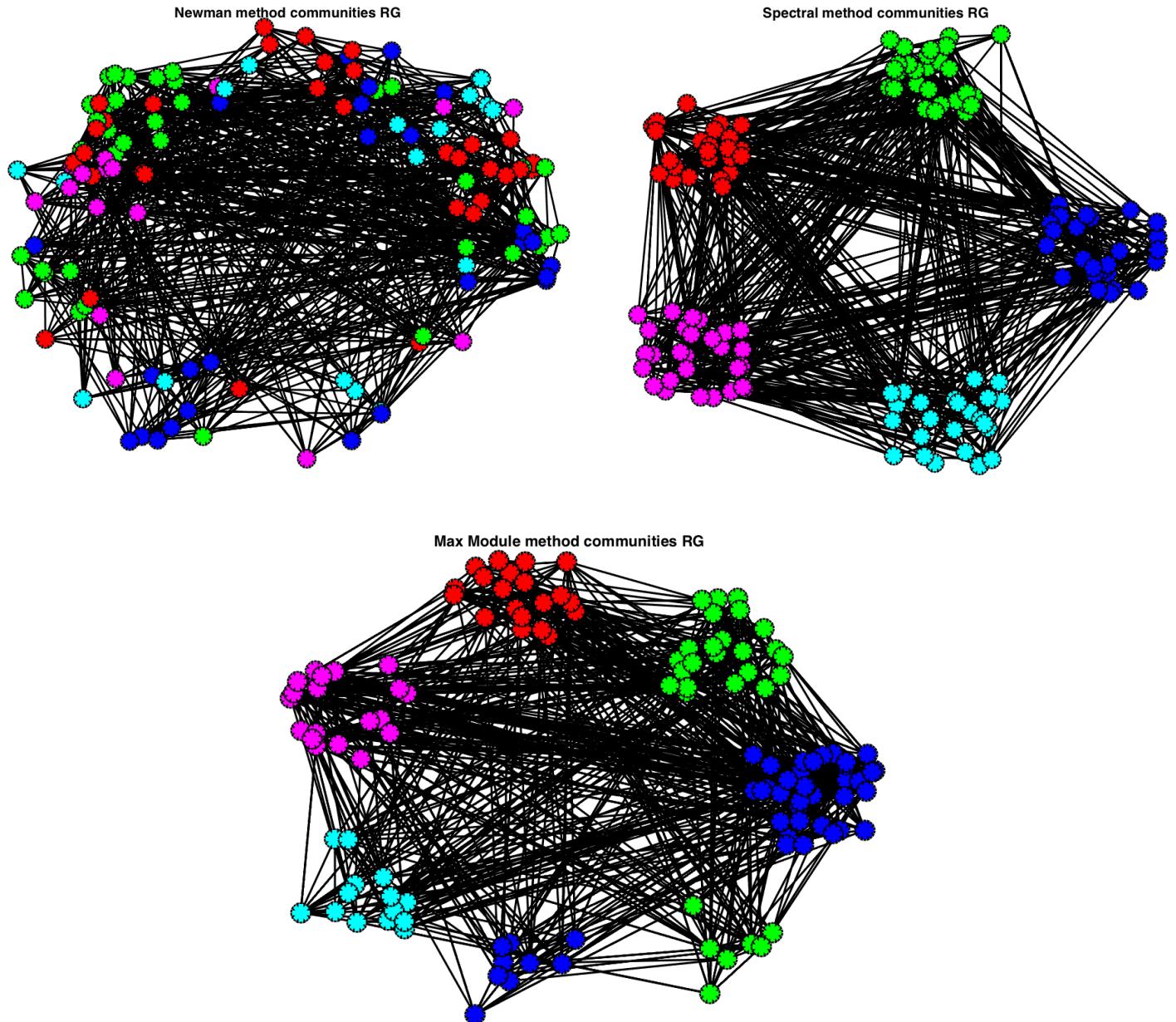
Στο μέρος αυτό της άσκησης αυτό που έχουμε να κάνουμε είναι για κάθε αλγόριθμο εντοπισμού κοινοτήτων και για κάθε τοπολογία(σύνθετη και πραγματική) να υπολογίσουμε το modularity και συγκρίνουμε τους αλγορίθμους με βάση την τιμή αυτή. Αυτό υλοποιήθηκε με τη βοήθεια του matlab και τα αποτελέσματα τα οποία πήρα φαίνονται παρακάτω :

**REG :**



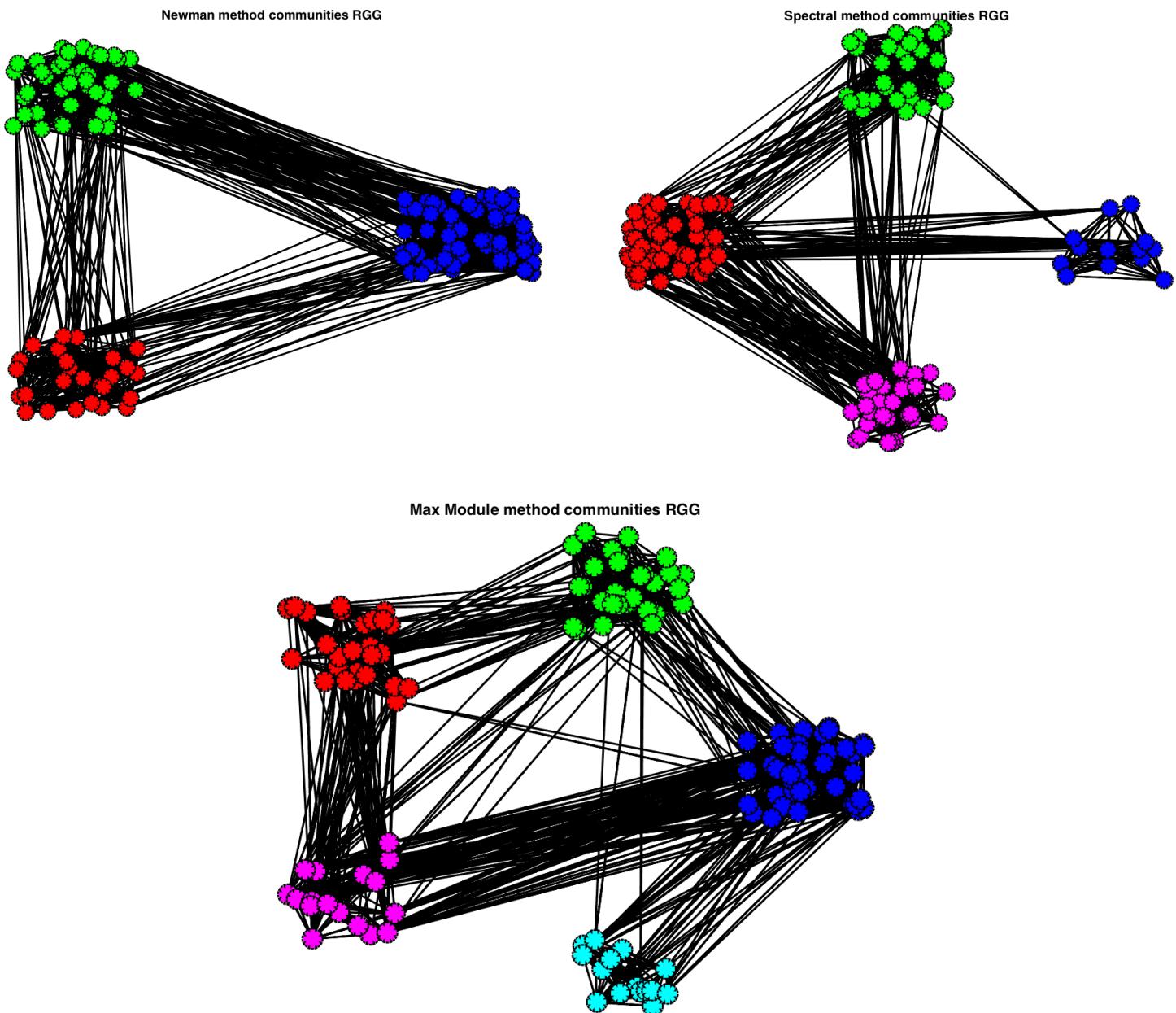
Method	Maximum Module
Newman(newman_comm_fast.m)	0.6786
Spectral(GCSpectralClust2.m)	0.5304
MaxModularity(GCModulMax1.m)	0.7187

RG :



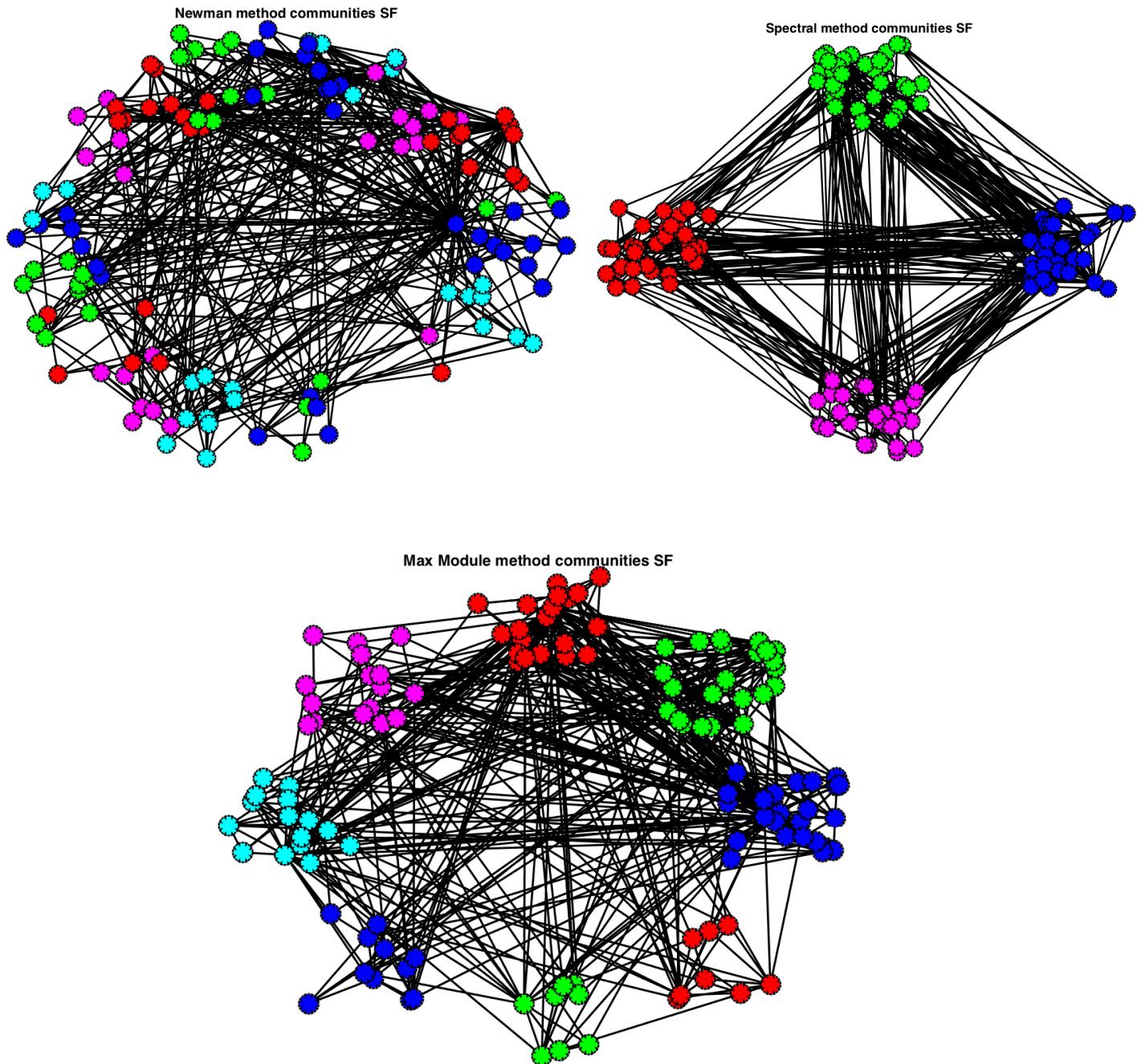
Method	Maximum Module
Newman(newman_comm_fast.m)	0.0745
Spectral(GCSpectralClust2.m)	0.1823
MaxModularity(GCModulMax1.m)	0.2410

RGG :



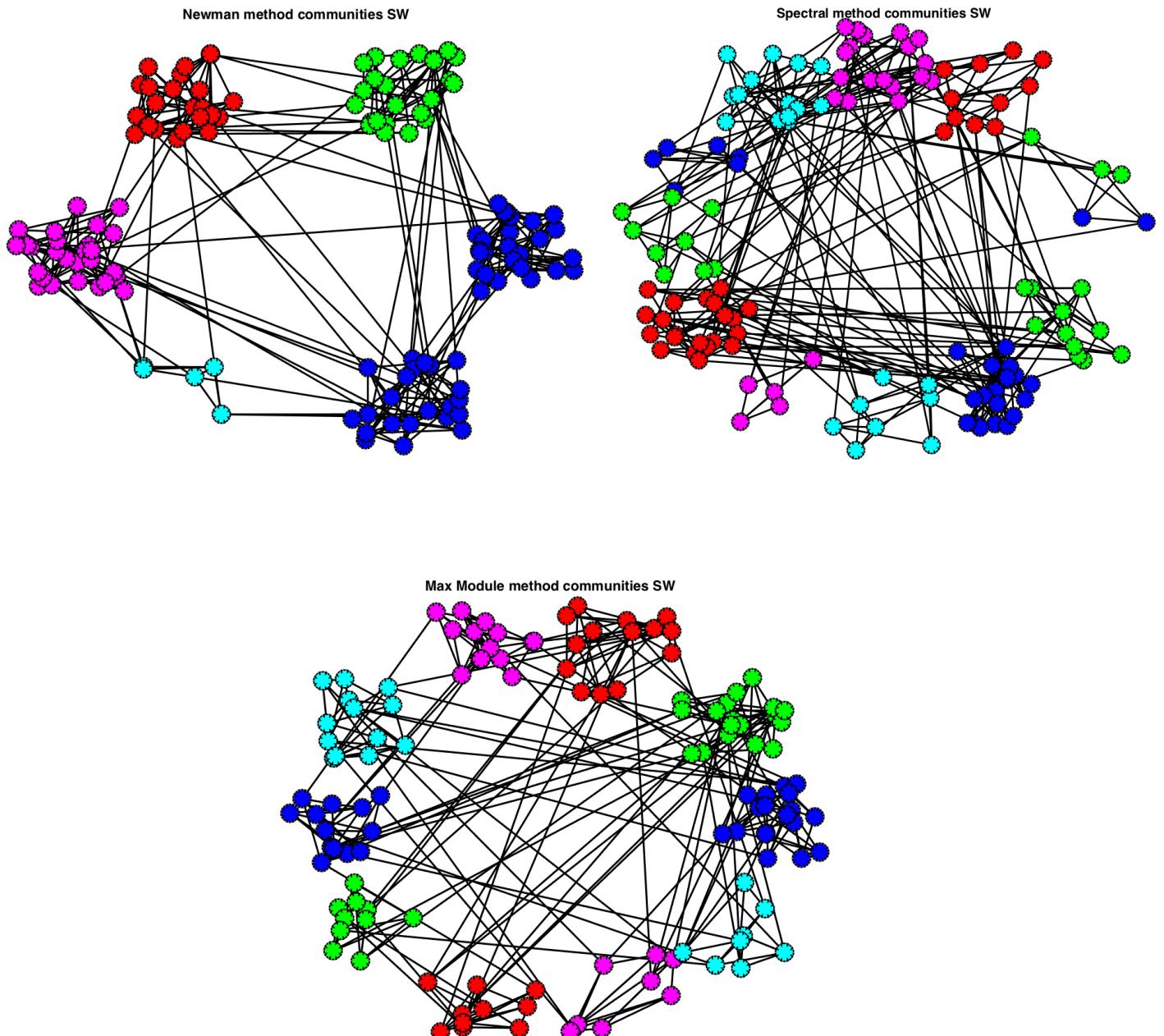
Method	Maximum Module
<b>Newman(newman_comm_fast.m)</b>	0.4698
<b>Spectral(GCSpectralClust2.m)</b>	0.5281
<b>MaxModularity(GCModulMax1.m)</b>	0.5467

SF :



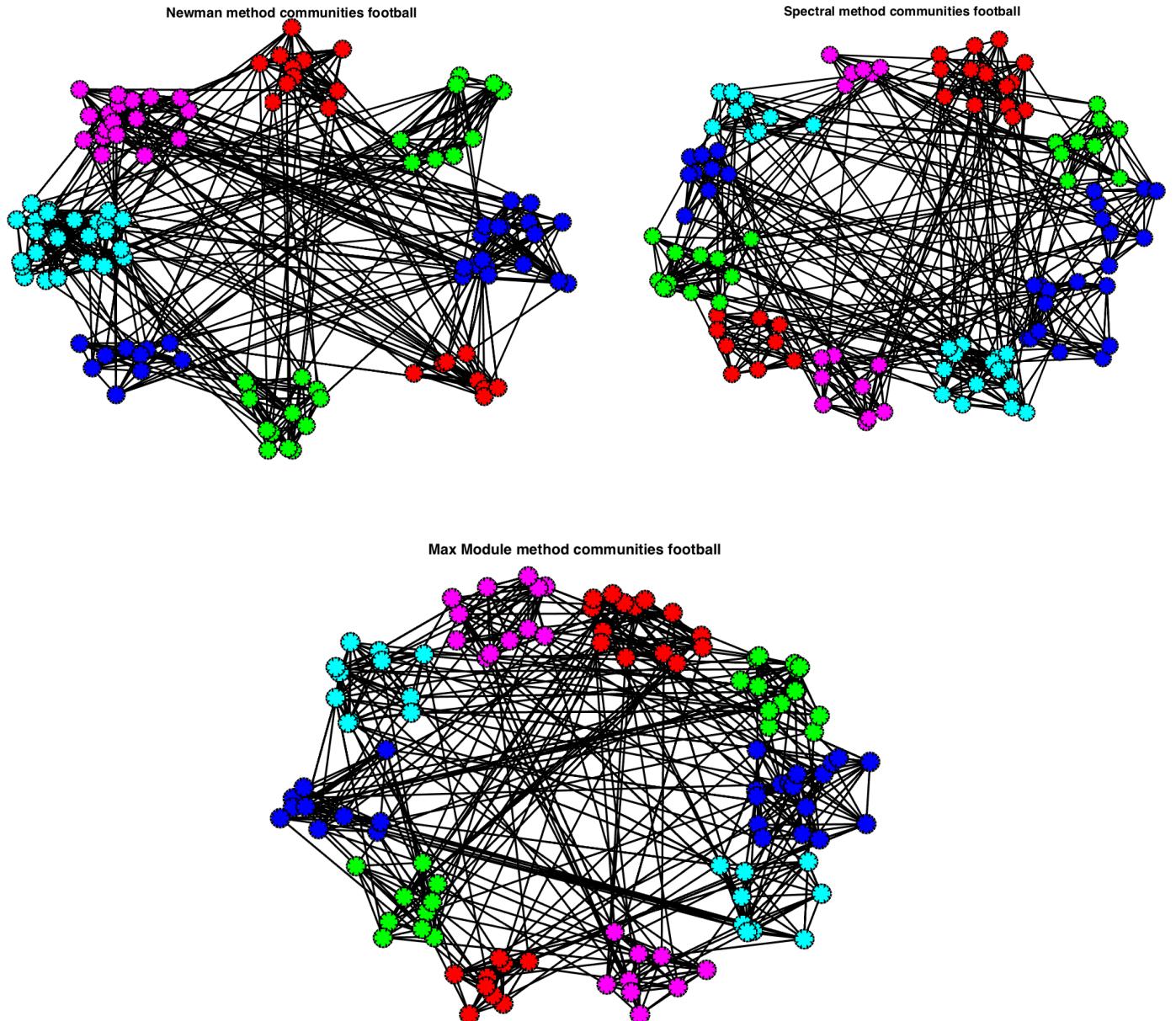
Method	Maximum Module
Newman(newman_comm_fast.m)	0.1542
Spectral(GCSpectralClust2.m)	0.1544
MaxModularity(GCModulMax1.m)	0.2868

SW :



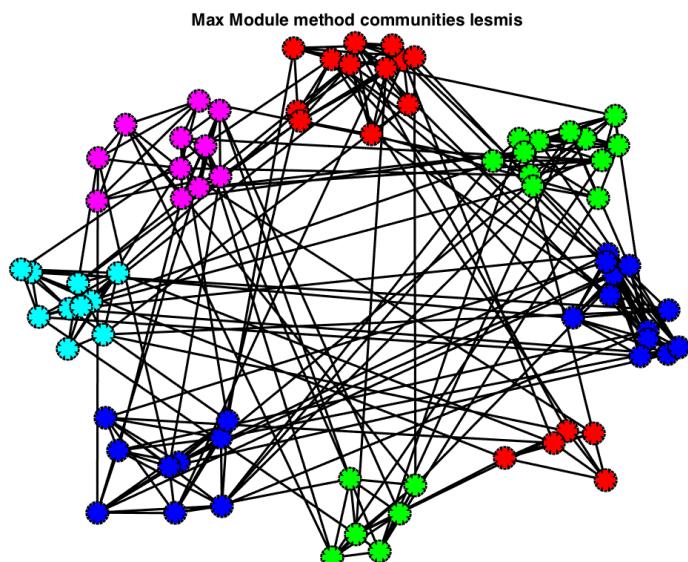
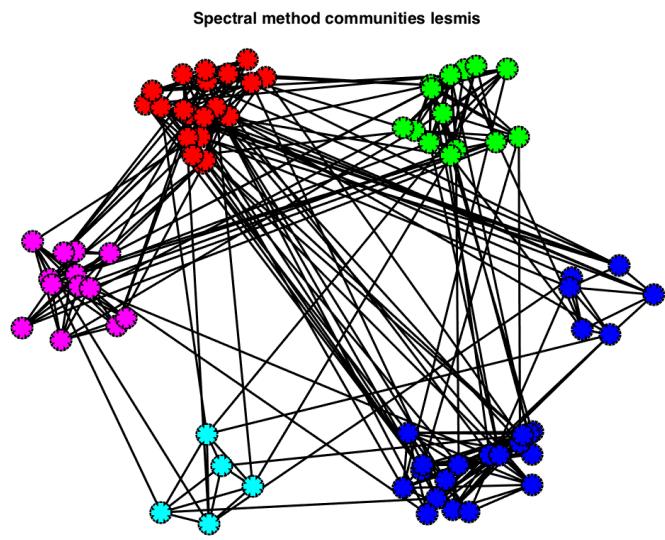
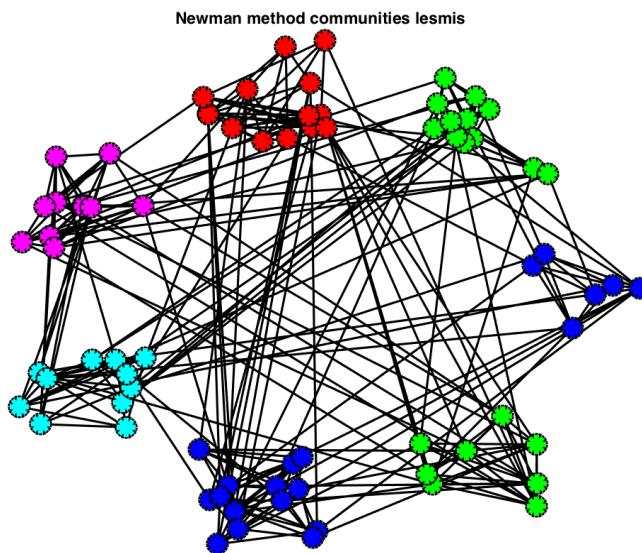
Method	Maximum Module
<b>Newman(newman_comm_fast.m)</b>	0.5916
<b>Spectral(GCSpectralClust2.m)</b>	0.4921
<b>MaxModularity(GCModulMax1.m)</b>	0.6659

Football :



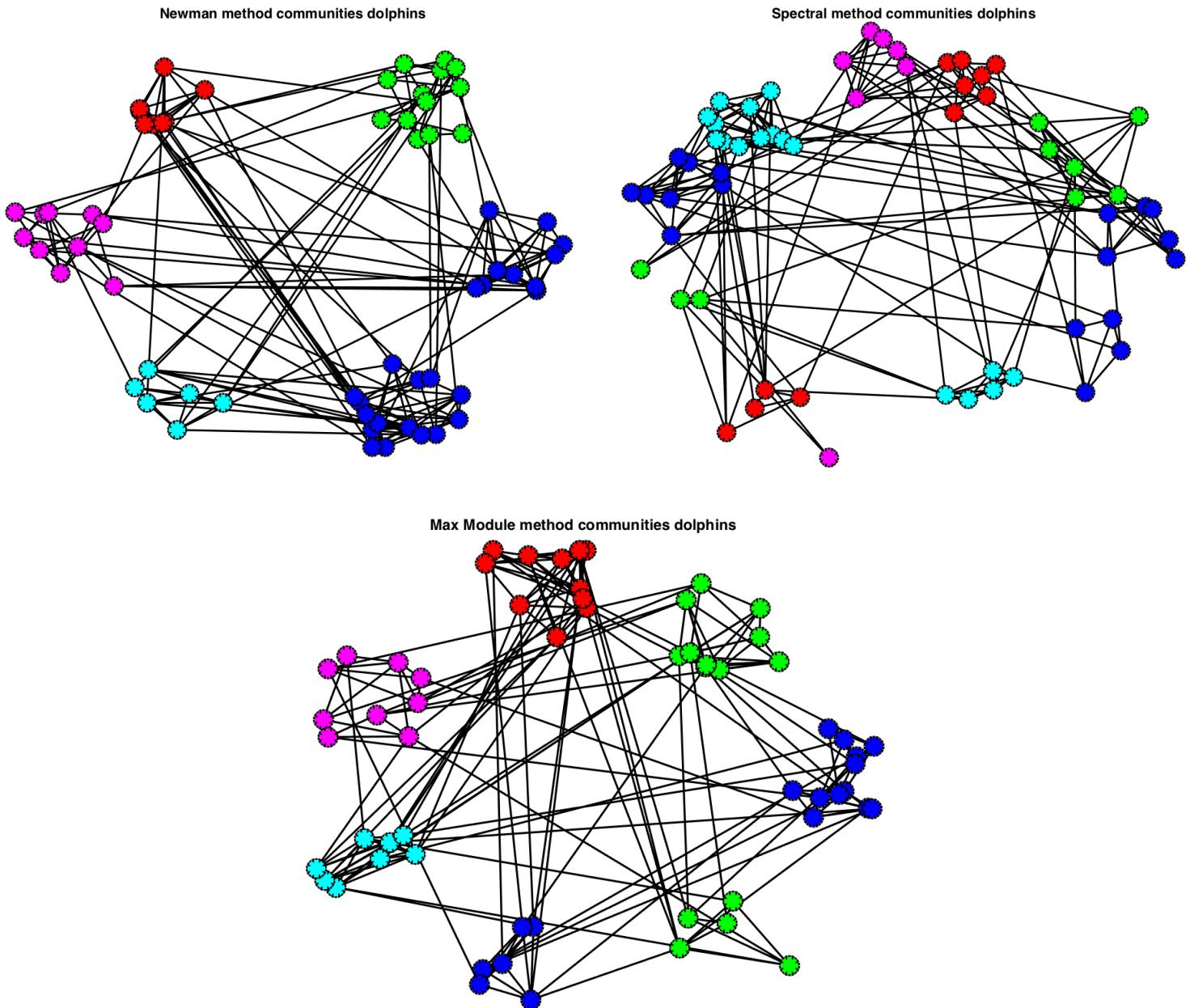
Method	Maximum Module
Newman(newman_comm_fast.m)	0.4914
Spectral(GCSpectralClust2.m)	0.5917
MaxModularity(GCModulMax1.m)	0.6046

Lesmis :



Method	Maximum Module
Newman(newman_comm_fast.m)	0.4364
Spectral(GCSpectralClust2.m)	0.4818
MaxModularity(GCModulMax1.m)	0.5670

Dolphins :



Method	Maximum Module
<b>Newman(newman_comm_fast.m)</b>	0.4163
<b>Spectral(GCSpectralClust2.m)</b>	0.5114
<b>MaxModularity(GCModulMax1.m)</b>	0.5540

### **Συγκεντρωτικός Σχολιασμός Αποτελεσμάτων :**

- Για τη μέθοδο ετοπισμού κοινοτήτων newman χρησιμοποίησα τη συνάρτηση newman\_comm\_fast.m η οποία μου έδινε το modularity metric for each cluster breakdown και από αυτό έπαιρνα το max για να βρω το maximum module για κάθε τοπολογία. Στις μέθοδους Newman κι Spectral διάλεξα ως αριθμό κοινότητων αυτές που έκαναν max το modularity.
- Για την REG παρατηρούμε οτι οι μέθοδοι Newman και MaxModularity δίνουν αρκετά κοντινά αποτελέσματα τα οποία είναι πολύ καλά αφού δίνουν και πολύ καλό διαχωρισμό, ενώ η μέθοδος spectral δίνει τα χειρότερα αποτελέσματα συγκριτικά με τις άλλες και ο λόγος είναι οτι βάζει τυχαία ακμές και δημιουργησε παραπάνω κοινότητες. Το καλύτερο αποτέλεσμα το δίνει η MaxModularity.
- Για την RG τοπολογία έχουμε οτι η μέθοδος Newmann δίνει το χειρότερο αποτέλεσμα και στη συνέχεια ακολουθούν οι Spectral και η MaxModularity. Τα αποτελέσματα εδω είναι αρκετά χαμηλά και ο διαχωρισμός είναι κακός και με τις τρεις μεθόδους.
- Στην τοπολογία RGG και οι 3 μέθοδοι δίνουν κοντινά αποτελέσματα αλλά και πάλι τα χειρότερα τα δίνει η Newman και τα καλύτερα η MaxModularity. Μια ακόμη παρατηρηση είναι οτι έχουμε και με τις τρεις μεθόδους έχουμε μικρό αριθμόν κοινοτήτων γεγονός το οποίο οφείλεται στις συντεταγμένες κάθε κόμβου αφού παίζουν ρόλο στην δημιουργία μιας κοινότητας.
- Στην τοπολογία SF παρατηρούμε πολύ άσχημα αποτελέσματα και με τις 3 μεθόδους, ιδιαίτερα με την Spectral και την Newman. Αυτό οφείλεται στο γεγονός οτι σε τέτοιου είδους γράφους λαμβάνουμε υπόψην τη κατανομή βαθμού, πράγμα το οποίο σημαίνει οτι κόμβοι με μεγάλο βαθμό συνδέονται μεταξύ τους και σχηματίζουν κοινότητα.
- Στην τοπολογία SW λόγω της μικρής πιθανότητας ανασύνδεσης ακμών τα αποτελέσματα μοιάζουν με εκείνα της REG, δηλαδή η Maxmodularity δίνει τα καλύτερα αποτελέσματα και η Spectral τα χειρότερα όσο αφορά επίδοση και διαχωρισμό.
- Για την τοπολογία Football έχουμε οτι τη χειρότερη επίδοση την έχει η μέθοδος Newman, ενώ οι άλλες έχουν σχεδόν την ίδια(η MaxModularity υπερισχύει ελάχιστα) και ο αριθμός των κοινοτήτων που εντοπίζονται είναι σχετικά κοντικός και για τις τρεις μεθόδους.
- Για τη τοπολογία Lesmis ισχύουν τα ίδια με την τοπολογία Football μόνο που εδω η μέθοδος MaxModularity δίνει με διαφορά καλύτερη επίδοση.
- Για τη τοπολογία Dolphins ισχύουν τα ίδια με την τοπολογία Football τόσο για την επίδοση όσο και για τον αριθμό των κοινοτήτων.
- Μια γενική παρατήρηση είναι οτι η μέθοδος Spectral παρότι στις σύνθετες τοπολογίες έδινε σχετικά άσχημα αποτελέσματα, στις πραγματικές έδωσε αρκετά καλά. Επίσης η μέθοδος Newman στις σύνθετες τοπολογίες έδινε είτε αρκετά καλά είτε πολύ άσχημα αποτελέσματα ενώ στις πραγματικές έδωσε μέτρια αποτελέσματα και στις 3 τοπολογίες, τα οποία όμως ήταν χειρότερα από τις άλλες μεθόδους. Η μέθοδος MaxModularity έδωσε και σε όλες τις τοπολογίες(εκτός την RGG) πολύ καλά αποτελέσματα(τα καλύτερα από τις 3 μεθόδους).