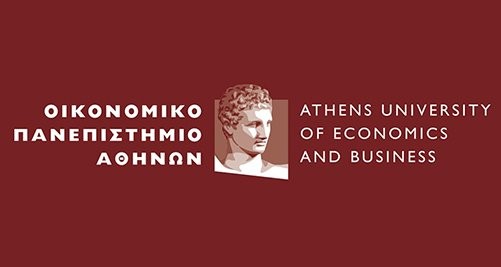
A close-up of a sign

Description automatically generated

Οικονομικό Πανεπιστήμιο Αθηνών

Τμήμα Διοικητικής Επιστήμης και Τεχνολογίας

Εργασία στο μάθημα Ανάλυση Κοινωνικών Δικτύων

Κωνσταντίνος Κατσάμης 3190237

A computer screen with red circles

Description automatically generated

February 15, 2024

**Περιεχόµενα**

1. [Εισαγωγή](#_bookmark0) …………………………………………………………………………………………………. 3
2. [Λήψη δεδοµένων](#_bookmark1) ………………………………………………………………………………………… 3
3. [Γραφική Αναπαράσταση ∆ικτύου](#_bookmark2) ……………………………………………………………………….. 4
4. [Βασικά Στοιχεία ∆ικτύου](#_bookmark3) ………………………………………………………………………………. 7
5. [Component Measures](#_bookmark4) ……………………………………………………………………………………. 8
6. [Degree Measures](#_bookmark5) 10
   1. [Maximum Degree](#_bookmark6) 10
   2. [Average Node Degree](#_bookmark7) 10
   3. [Degree Distribution](#_bookmark8) 11
      1. [In-Degree](#_bookmark9) 11
      2. [Out-Degree](#_bookmark10) 13
      3. [Total Degree](#_bookmark11) 16
7. [Centrality measures](#_bookmark12) 19
   1. [Betweenness Centrality](#_bookmark13) 19
   2. [Closeness Centrality](#_bookmark14) 21
   3. [Eigenvector Centrality](#_bookmark15) 24
8. [Clustering Effects](#_bookmark16) 27
   1. [Average Clustering Coefficient](#_bookmark17) 27
   2. [Number of Triangles](#_bookmark18) 28
   3. [Clustering Coefficient Distribution](#_bookmark19) 30
   4. [Existence of the Triadic Closure Phenomenon in the Friendship Neighborhood](#_bookmark20) 34
9. [Bridges and Local Bridges](#_bookmark21) 36
10. [Gender and Homophily](#_bookmark22) 38
11. [Graph Density](#_bookmark23) 40
12. [Community Structure(Modularity)](#_bookmark24) 41
    1. [Community Structure - Modularity Resolution 0.1](#_bookmark25) 42
    2. [Community Structure - Modularity Resolution 0.5](#_bookmark26) 44
    3. [Community Structure - Modularity Resolution 0.9](#_bookmark27) 46
13. [PageRank](#_bookmark28) 48
14. [Συµπεράσµατα](#_bookmark29) 51
15. **Εισαγωγή**

Το YouTube είναι ένας ισότοπος κοινοποίησης, αποθήκευσης, αναζήτησης και αναπαραγωγής βίντεο. Κάθε χρήστης µπορεί να δημιουργήσει λογαριασμό και να ανεβάζει τα δικά του βίντεο ή ακόμα και να αναπαράγει σε πραγματικό χρόνο. Εκτός από τους χρήστες, πρόσβαση έχει ο οποιοσδήποτε στον ισότοπο αυτό όπου µπορεί µόνο να παρακολουθεί τα βίντεο άλλων χρηστών. Το προφίλ του χρήστη παρουσιάζεται ως κανάλι όπου άλλοι χρήστες µπορούν να εγγραφούν ώστε να παρακολουθούν και να ενημερώνονται για βίντεο ή για πραγματικού χρόνου αναπαραγωγές που τους ενδιαφέρουν. Τα βίντεο που ανεβάζει ο κάθε χρήστης είναι συνήθως αποθηκευμένα σε playlists αναλόγως µε την μορφή και το θέμα που έχουν. Επίσης στο κανάλι του ο κάθε χρήστης µπορεί να έχει κανάλια άλλων χρηστών που όπως αναφέρονται στην αγγλική ορολογία ”Featured channels”. Τα επιλεγμένα αυτά κανάλια αποτελούν κανάλια όπου ένας χρήστης επιλέγει να τα συμπεριλάβει στο δικό του κανάλι(δεν φαίνονται στο κοινό). Ο λόγος που γίνεται αυτό είναι για να προωθούν οι χρήστες και να εμφανίζουν άλλα κανάλια που τους αρέσουν, µε τα οποία µπορεί να συνεργάζονται ή να θέλουν να τα προτείνουν στους θεατές τους. Με αυτό τον τρόπο, οι χρήστες µπορούν να προσεγγίσουν πολλά είδη κοινού και να αυξήσουν έτσι τις εγγραφές και τις προβολές τους. Στην ανάλυση αυτή θα εξετάσουμε το κανάλι Samsung. Το κανάλι αυτό είναι το κανάλι της εταιρείας Samsung που έχει ως σκοπό την ενημέρωση σχετικά µε εκδηλώσεις, εφαρμογές και υπηρεσίες, B2B solutions, παρουσιάσεις, και τις τελευταίες καινοτόμες τεχνολογίες της εταιρείας.

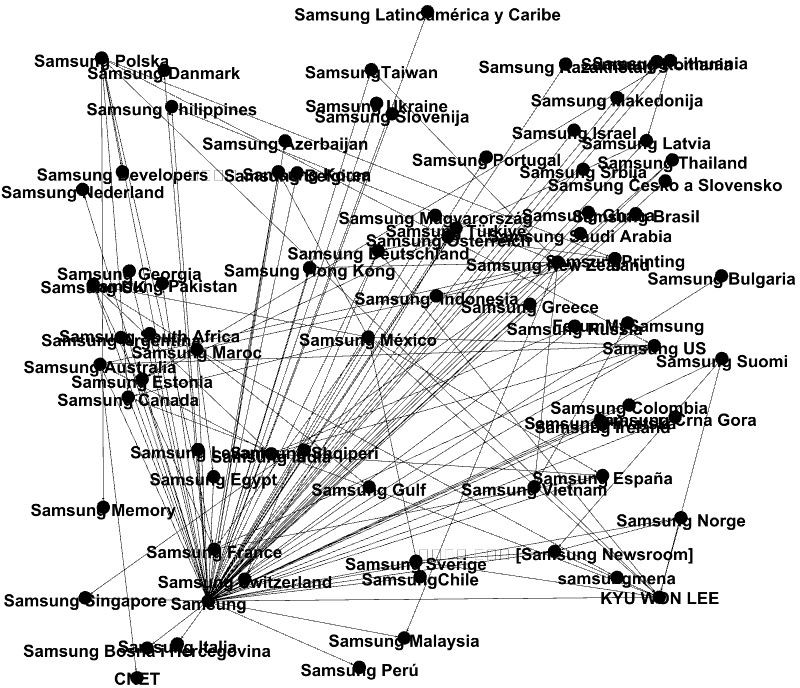
1. **Λήψη Δεδοµένων**

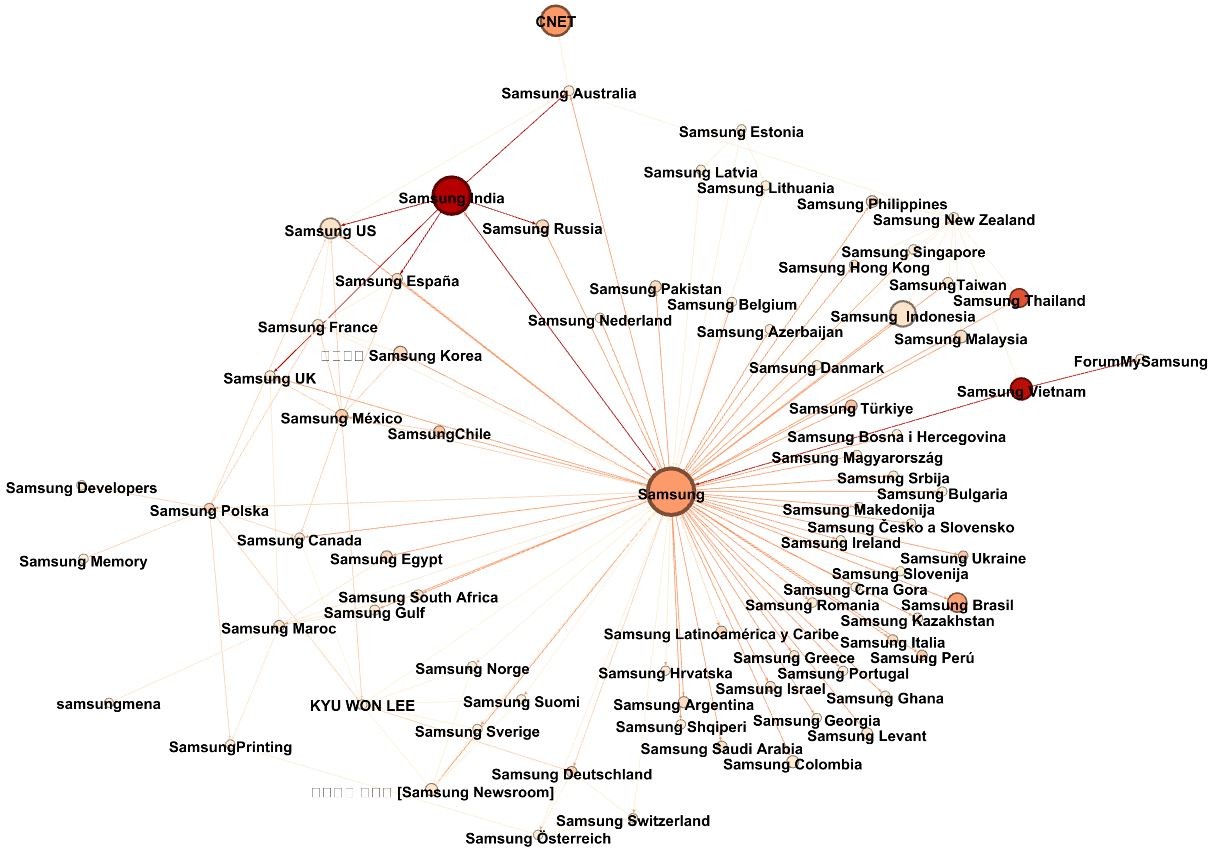
Τα δεδομένα για την ανάλυση µας τα πήραμε µε τη χρήση του [Bernhard Reiner’s Tool](https://labs.polsys.net/) χρησιμοποιώντας τα YouTube Data Tools. Αρχικά, χρησιμοποιώντας το link του καναλιού στο YouTube, βρήκαμε το id του καναλιού µέσω του [Channel Info Module](https://ytdt.digitalmethods.net/mod_channel_info.php). ΄Έπειτα µε τη χρήση του [Channel Network Module](https://ytdt.digitalmethods.net/mod_channel_info.php), πήραμε δεδομένα για το δίκτυο του καναλιού. Οι παράμετροι που χρησιμοποιήθηκαν ήταν το seed(αρχικό κανάλι – κανάλι πηγή) µε τη χρήση του id µε crawl depth ίσο µε 2(το crawl depth καθορίζει πόσο βαθιά στο δίκτυο µπορουµε να φτάσουμε. Για παράδειγμα µε depth=0 το εργαλείο αυτό επιστρέφει το δίκτυο µε τις συσχετίσεις ανάμεσα στα seeds που δίνονται, µε dept=1 επιστρέφει τα featured channels που έχει ο χρήστης στο κανάλι του και µε depth=2 επιστρέφει τα featured channels που υπάρχουν στα κανάλια που βρήκαμε στο depth=1). Η επιλογή για της εγγραφές δεν λήφθηκε υπόψιν διότι θέλαμε τα δεδομένα να είναι µόνο µε τα featured channels. Μετά από αυτά τα βήματα το εργαλείο δημιούργησε ένα gdf αρχείο το οποίο φορτώσαμε στο πρόγραµµα Gephi για ανάλυση. Εδώ να σημειωθεί ότι µέσω του Gephi έγινε έλεγχος των δεδομένων για τυχόν σφάλματα που θα µπορθούσαν να

επηρεάσουν την ανάλυση µας όπως για παράδειγμα ο έλεγχος διπλοτύπων, όπου σε µια περίπτωση υπήρξε διπλότυπο όπου και αντιμετωπίστηκε µέσω του Gephi, ο έλεγχος για null τιμές κ.α. Σε ορισµένες περιπτώσεις υπήρχαν µη διαθέσιμες τιμές. Για παράδειγμα σε ορισμένους κόμβους, δεν υπήρχε στο αντίστοιχο κελί η χώρα ενώ ήταν γνωστή. ΄Έτσι συμβουλευμένοι τα κανάλια, εντοπίστηκαν και συμπληρώθηκαν.

1. **Γραφική Αναπαράσταση Δικτύου**

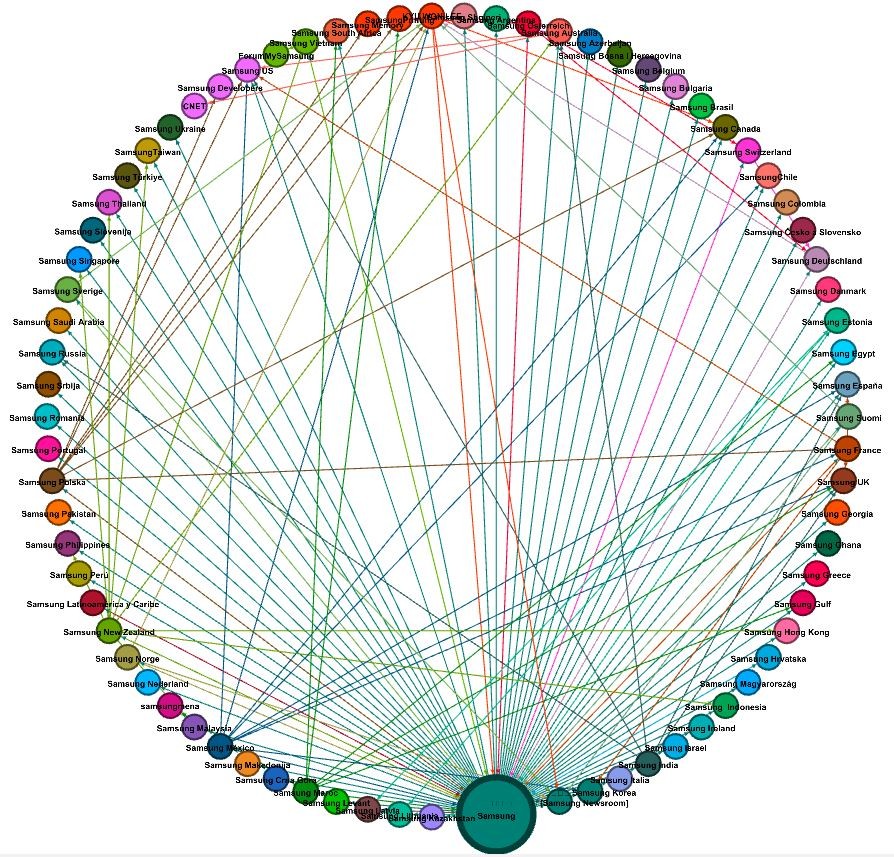
Το δίκτυο µόνο µε τα ονόματα των κόμβων(καναλιών) χωρίς κάποια παραμετροποίηση.



Επίσης µέσω του Gephi µπορούµε να θέσουμε διάφορες παραμέτρους όσον αφορά τoν χρωματισμό και την διάταξη ανάλογα µε ορισµένες ιδιότητες που έχει το δίκτυο µας. Για παράδειγμα, για τη μορφή των κόμβων θέσαμε το μέγεθος του κάθε κόμβου ανάλογα µε το πλήθος των εγγραφών(subscribercount) που έχει το κανάλι που αντιπροσωπεύει και για τον χρωματισμό θέσαμε άσπρο-πορτοκαλί-κόκκινο στο χαρακτηριστικό των προβολών(viewcount(100s)). Για την διάταξη, τρέξαμε τον Atlas Force 2 για να αραιώσουμε τον γράφο µας και τον Label Adjust για να διαχωριστούν οι ετικέτες ονομάτων των κόμβων. ΄Έτσι προέκυψε η παρακάτω εικόνα:

Από την εικόνα αυτή, τα δεδομένα που λαμβάνουμε είναι ο αριθμός των εγγραφών σε ένα κανάλι παίζει αρκετό ρόλο µε τις προβολές που µπορεί να έχει, πράγμα αναμενόμενο για τον ισότοπο που συζητάμε.

Βλέποντας τα δεδομένα του δικτύου µας από το Data Laboratory του Gephi, παρατηρήσαμε πως υπάρχουν κανάλια από διάφορες χώρες. Επομένως θεωρήσαμε ενδιαφέρον να κάνουμε µία παραμετροποίηση µε τις χώρες ως εξής. Ο χρωματισμός έγινε µέσω διαφορετικών χρωμάτων, τόσων, όσος και ο αριθμός των διαφορετικών χωρών, µέσω του partition tab. Στο σημείο αυτό, θεωρήσαμε επίσης σημαντικό και την αναφορά του seed. Αυτό έγινε µεσω του µεγέθους των κόμβων µέσω του seedrank(αντίστοιχη μεταβλητή µε την isseed εάν χρησιμοποιούσαμε τον χρωματισμό). Στη συνέχεια µέσω του Plugin Circular Layout που κατεβάσαμε µέσω των Tools του Gephi, δημιουργήσαμε την πιο κάτω διάταξη θέτοντας στην ιδιότητα ”Order Nodes By” την χώρα. Για άλλη µια φορά, χρησιμοποιήσαμε τον Label Adjust για διαχωρισμό των ετικετών.

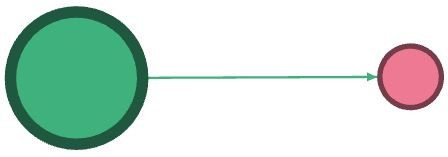


Από την πιο πάνω εικόνα µπορούµε εύκολα να παρατηρήσουμε πως ο κεντρικός και ίσως ο πιο σημαντικός κόμβος να είναι ο ”Samsung” ο οποίος είναι µε πράσινο χρώμα. Οι δύο δεξιές θέσεις από αυτό το κόμβο είναι επίσης µε πράσινο χρώμα αφού και αυτοί οι κόμβοι είναι κανάλια από την ίδια χώρα, την Νότιο Κορέα.

1. **Βασικά στοιχεία Δικτύου**

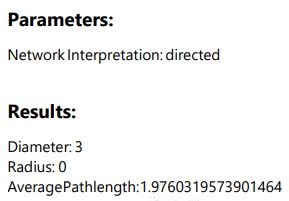
Το δίκτυο που μελετάμε έχει τα εξής βασικά στοιχεία:

* Αριθμός κόμβων: **76** διαφορετικά **κανάλια-κόμβοι**
* Αριθμός ακμών: **149 σύνδεσμοι** µέσω των οποίων συνδέονται τα κανάλια-κόμβοι
* Ο γράφος µας είναι **κατευθυνόμενος**. ∆ηλαδή κάθε σύνδεσμος από ένα κανάλι προς ένα άλλο έχει κατεύθυνση όπως φαίνεται στην πιο κάτω εικόνα:



Ο πράσινος κόμβος-κανάλι έχει ως featured channel τον κόμβο-κανάλι µε ροζ χρώμα.

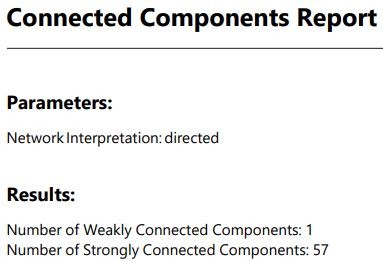
* ∆ιάµετρος δικτύου: Η **διάμετρος** ενός δικτύου είναι η µακρύτερη συντομότερη διαδρομή που µπορούµε να βρούμε. Στην περίπτωσή µας είναι **3**. Τιμή αναμενόμενη λόγω του depth µε τιμή 2 που επιλέξαμε(Η αρίθμηση του depth ξεκινάει από μηδέν).
* **Average path length**: Είναι ο **µέσος όρος των συντομότερων µονοπατιών** για όλα τα ζεύγη κόμβων. Στο δίκτυο µας είναι **1.9760**.



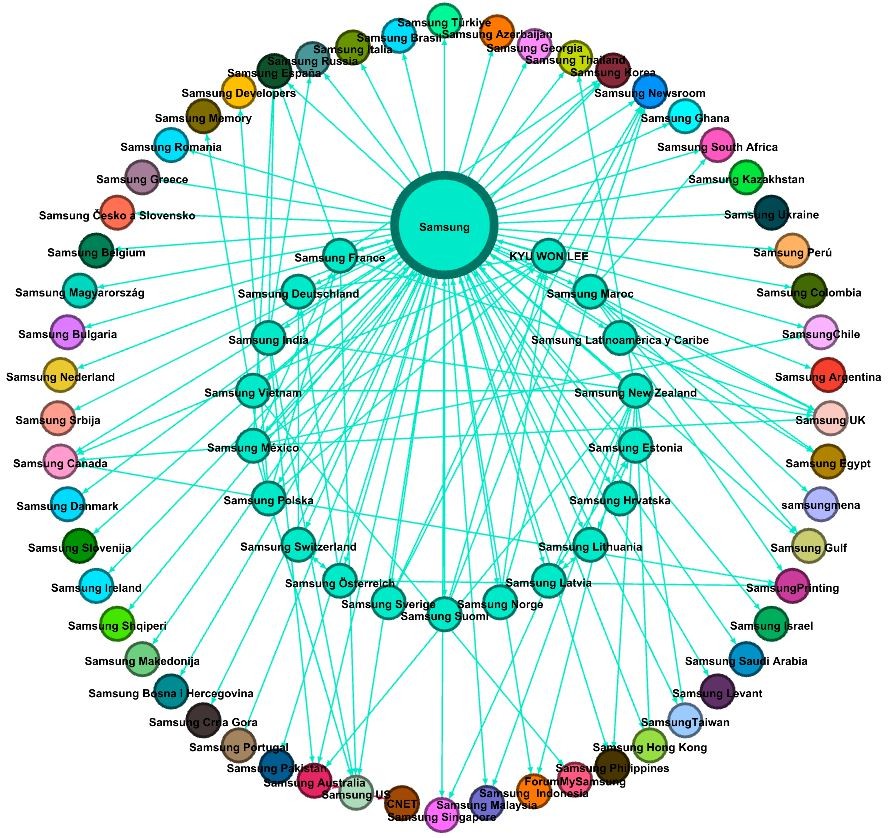
1. **Component Measures**

Στο δίκτυο µας, όλοι οι κόμβοι είναι συνδεδεμένοι µεταξύ τους(έµµεσα είτε άμεσα). Άρα µπορούµε να πούμε πως υπάρχει **ένα giant component**. Επομένως ο αριθμός των **weakly connected components** είναι ίσος µε **1**.

Αναφορικά µε τον αριθμό των **strongly connected components**, αυτό που πρέπει να δούμε στην περίπτωση µας είναι αν υπάρχουν κανάλια-κόμβοι τα οποία δεν έχουν Featured Channels, δηλαδή δεν έχουν εξερχόμενους συνδέσμους. Έτσι µέσω του Connected Components tool από το πεδίο Statistics του Gephi έχουμε την ακόλουθη αναφορά.



Παρατηρώντας την πιο πάνω εικόνα λοιπόν, µπορούµε να επιβεβαιώσουμε τον αριθμό των weakly con- nected components. Όσον αφορά τον αριθμό των strongly connected components µεσω του Gephi βλέπουμε πως είναι **57**. Στο σημείο αυτό µπορούµε να εφαρμόσουμε µια διάταξη για να δούμε σχηματικά αυτούς τους κόμβους ώστε να καταλάβουμε καλύτερα τι συμβαίνει. Χρησιμοποιώντας λοιπόν τον αλγόριθμο Dual Circle Layout, µε Upper Order Count ίσο µε 20(Πλήθος κόμβων - strong connected components + weakly connected components) µε σκοπό να πάρουμε στον εξωτερικό κύκλο τα κανάλια που δεν έχουν Featured Channels(20 κανάλια, 20 διαφορετικά χρώματα). Έτσι όπως φαίνεται και πιο κάτω, στον εξωτερικό κύκλο, τα κανάλια αυτά έχουν ακμές που φτάνουν σε αυτά και κανένα δεν έχει ακμή που να ξεκινάει από αυτά.

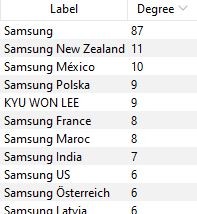


Να σημειωθεί ότι κρατήσαμε τη διαμόρφωση των κόμβων σχετικά µε το µέγεθος στην σχέση seedrank χωρίς αυτό να παίζει κάποιο ρόλο, γι’ αυτό και ο κόμβος Samsung έχει µεγαλύτερο µέγεθος.

# Degree Measures

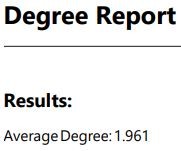
## Maximum Degree

Το Maximum Degree είναι ο µέγιστος αριθμός ακμών που έχει ένας κόμβος μέσα στο δίκτυο. Στην περίπτωση που εξετάζουμε, αφορά τον κόμβο ”Samsung” µε τιμή 87. Αποτέλεσμα αναμενόμενο, αφού ο συγκεκριμένος κόμβος παίζει τον πιο σημαντικό ρόλο στο δίκτυο µας όπως έχουμε δει και σε άλλες περιπτώσεις. Αυτό φαίνεται µέσω του πιο κάτω στιγμιότυπου που πήραμε από το Gephi αφού βρήκαμε πρώτα το degree του κάθε κόμβου.



## Average Node Degree

Το Average Node Degree είναι ο µέσος αριθμός ακμών που υπάρχουν στο δίκτυο. Στο δίκτυο µας είναι ίσο µε 1.961 σύμφωνα µε το Degree Report που φτιάξαμε µέσω του Gephi από το µενού Statistics.



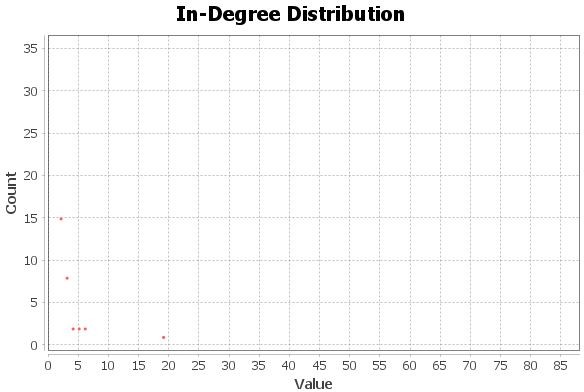
## Degree Distribution

### In-Degree

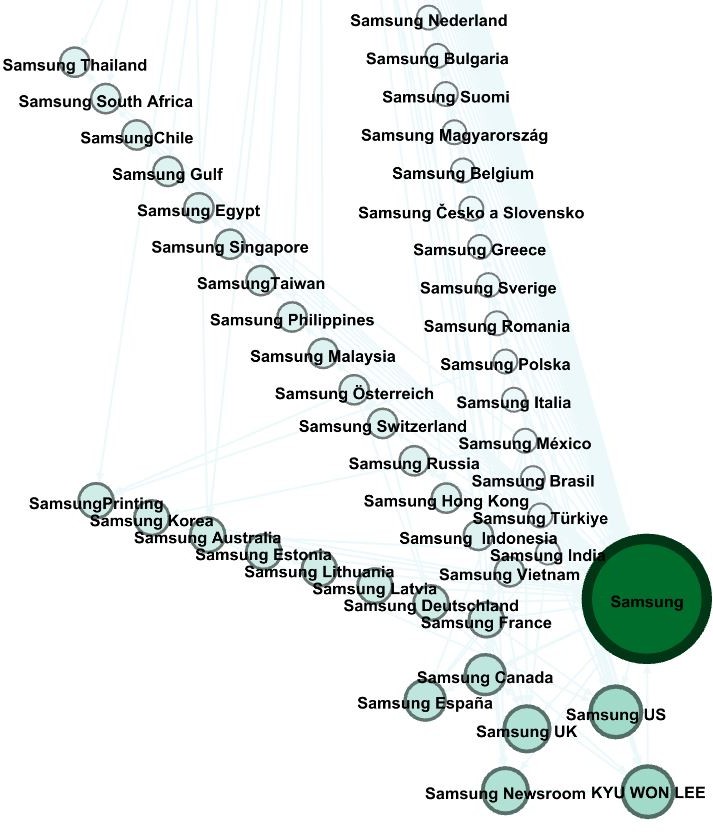
Το In-Degree είναι οι εισερχόμενες προς κάποιον κόμβο ακμές. Στην περίπτωση µας, ο αριθμός αυτός αποτελεί τον αριθμό των καναλιών που έχουν ως Featured Channel το κανάλι που εξετάζουμε. Έτσι για κάθε κανάλι µε τη βοήθεια του Gephi για το δίκτυο µας έχουμε:



Κατανομή του In-Degree µέσω γραφικής παράστασης:



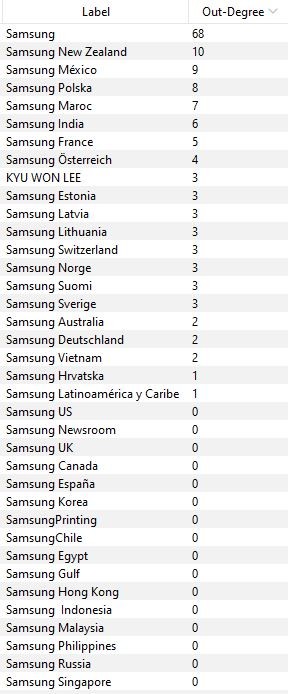
Μετά από τα πιο πάνω, θα ήταν αρκετά ενδιαφέρον να δούμε πως αλλάζει το δίκτυο όσον αφορά µέγεθος και χρώμα κόμβων σε συνάρτηση µε το In-Degree.



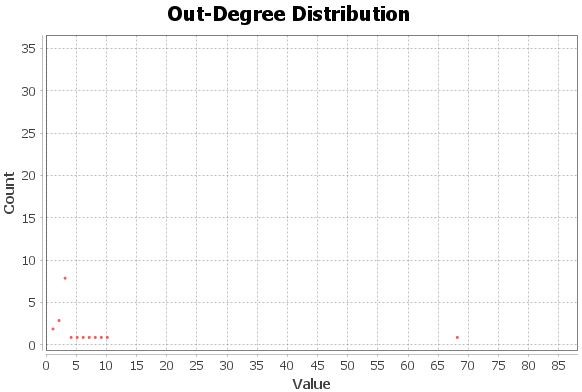
Έτσι χρησιμοποιώντας τον αλγόριθμο Radial Axis Layout µπορούµε να δούμε τον διαχωρισμό που γίνεται ανάμεσα στους κόμβους σε σχέση µε το in-degree του κάθε καναλιού. Τα κανάλια λοιπόν χωρίστηκαν σε 7 διαφορετικές ομάδες σε οριζόντιους άξονες αφού οι διαφορετικές τιμές που παρατηρούνται είναι 7 όπως είδαμε και στους πιο πάνω πίνακες. Έτσι στο σημείο αυτό µπορούµε εύκολα να δούμε τα κανάλια τα όποια υπάρχουν κατά πολύ περισσότερες φορές ως Featured channels σε άλλα. Πρωταγωνιστικό ρόλο έχει το κανάλι της Samsung για ακόμα µια φορά ενώ ακολουθούν στη συνέχεια τα κανάλια SamsungUS, KYO WON LEE κοκ.

### Out-Degree

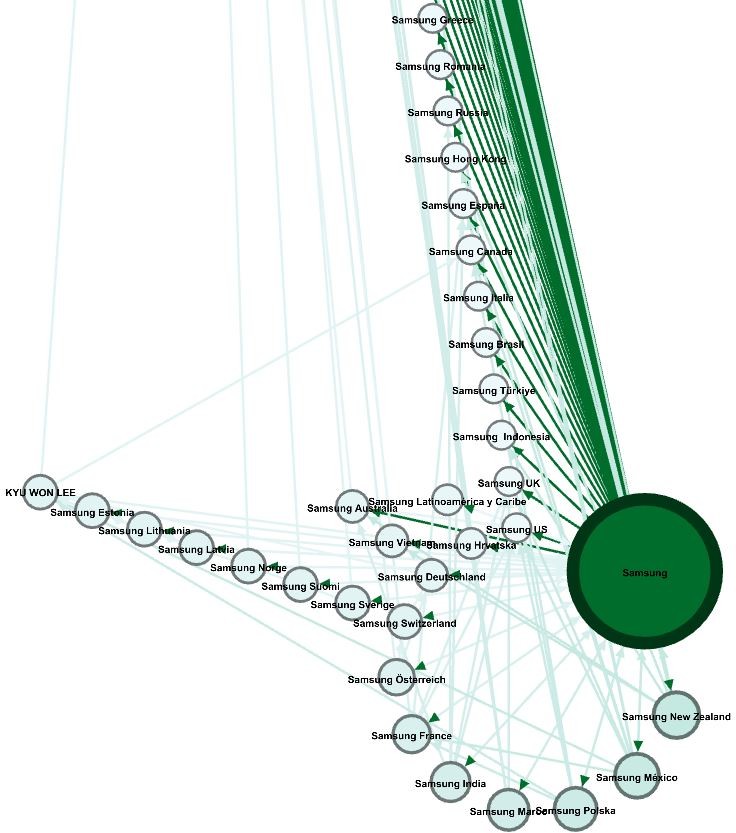
Το Out-Degree είναι οι εξερχόμενες από τον κάθε κόμβο ακμές. Με το δίκτυο το οποίο µελετάµε είναι ο αριθμός των Featured Channels που µπορεί να έχει ένα κανάλι όπως βλέπουμε παρακάτω.



Κατανομή του Out-Degree µέσω γραφικής παράστασης:



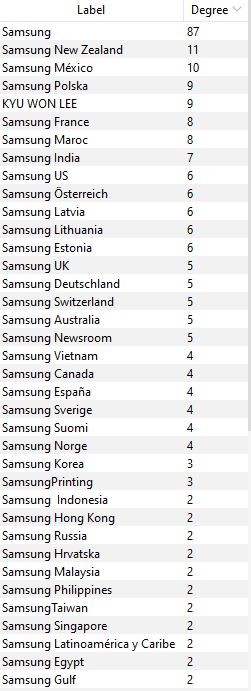
Αντίστοιχα µε το In-Degree θα δούμε πως αλλάζει το δίκτυο όσον αφορά µέγεθος και χρώμα κόμβων σε συνάρτηση µε το Out-Degree αυτή τη φορά.



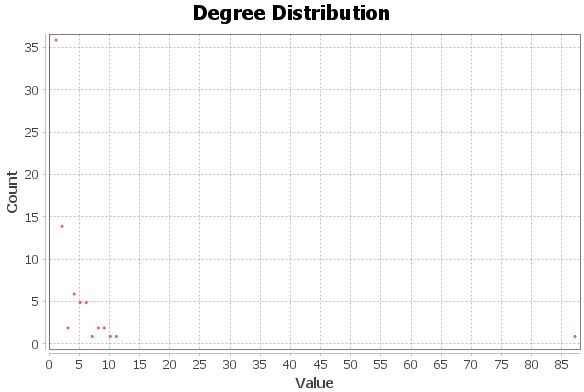
Με τον αντίστοιχο τρόπο που δουλέψαμε για το In-Degree προηγουμένως, δουλέψαμε και τώρα. Όπως παρατηρούμε, στην πρώτη θέση εξακολουθεί να είναι το κανάλι Samsung ενώ στο προσκήνιο έχουν προστεθεί αρκετά κανάλια σε σχέση µε πριν. Λογικό, αφού όσο πιο πολλά Featured Channels έχει ένα κανάλι τόσο πιο εύκολα µπορεί να προσεγγίσει κοινό και να αυξήσει τις προβολές και τις εγγραφές του. Επίσης ένα παράδειγμα που πολλές φορές συμβαίνει είναι ότι µε αυτόν τον τρόπο ο κόσμος µπορεί να ενημερωθεί πολύ πιο γρήγορα για ένα καινούργιο προϊόν που έχει παρουσιαστεί σε µια άλλη χώρα βλέποντας ένα προτεινόμενο κανάλι που θα προτείνει η ίδια η πλατφόρμα του YouTube µέσω των Featured Channels που έχει το κανάλι το οποίο ακολουθεί ένας χρήστης.

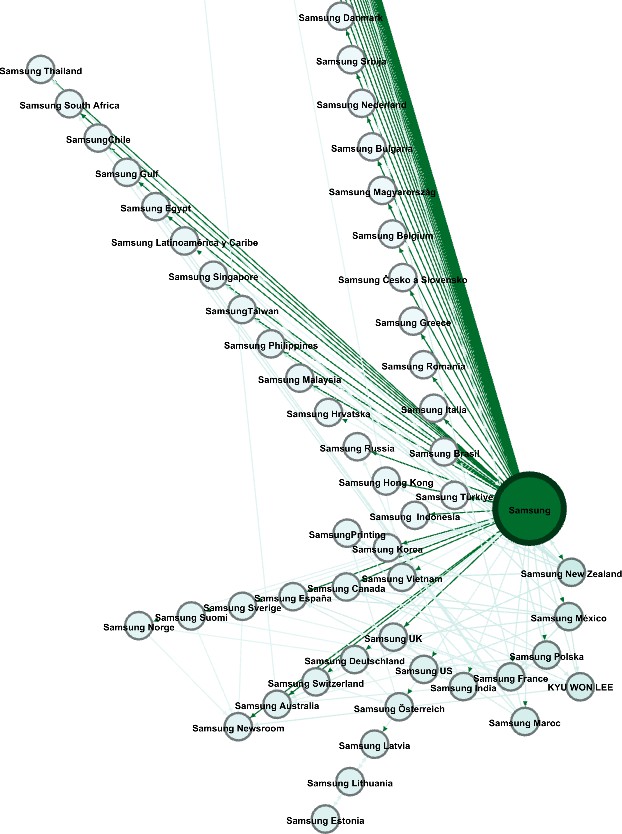
### Total Degree

Το Total Degree είναι το σύνολο των ακμών που ξεκινούν ή που καταλήγουν σε ένα κόμβο. Με άλλα λόγια, είναι ουσιαστικά το άθροισα του In-Degree και του Out-Degree.



Κατανομή του Total Degree µέσω γραφικής παράστασης:



Όπως και στις δύο προηγούμενες περιπτώσεις, θα δούμε πως διαμορφώνεται το δίκτυο µας λαμβάνοντας υπόψιν το Total Degree αυτή τη φορά.

Το πρώτο πράγμα που µπορεί να προσέξει κανείς για το σχήμα που προέκυψε µε µετρική το Total Degree είναι πως υπάρχει ο ίδιος αριθμός ομάδων κατά πλήθος κόμβων σε σχέση µε πριν. Η διαφορά όμως εγγυάται στο γεγονός πως όλοι σχεδόν οι κόμβοι που υπήρχαν και πριν στο Out-Degree, πέραν από τον προφανές της Samsung, υπάρχουν και τώρα. Άρα φαίνεται πως το Out-Degree είναι αυτό που παίζει τον πιο σημαντικό ρόλο αφού όπως είπαμε και προηγουμένως είναι αυτό που καθορίζει ποια κανάλια θα προωθηθούν περισσότερο από τον τρόπο που δουλεύει το YouTube µέσω των Featured Channels.

# Centrality measures

## Betweenness Centrality

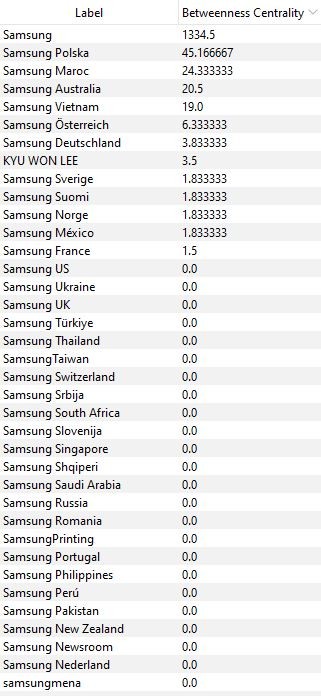
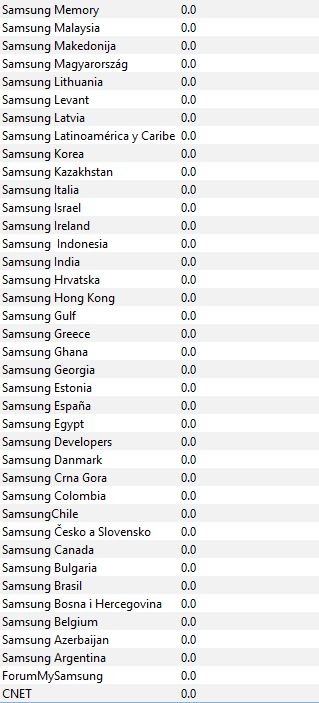
Το Betweenness Centrality δείχνει πόσο σημαντικός είναι ένας κόμβος(ως ενδιάμεσος) όταν θέλουμε να συνδέσουμε όλους τους κόμβους µεταξύ τους µέσω αυτού. Για παράδειγμα, για τον κόμβο *ni* βρίσκουμε για κάθε ζεύγος κόμβων(u, w) του δικτύου τις εξής τιμές όπου και τις διαιρούμε:

1. Το σύνολο των συντομότερων µονοπατιών από τον κόμβο *ni*: Σ*uω*(*ni*)
2. Με τον αριθμό των συντομότερων διαδρομών που περνούν από τον κόμβο x(τα μονοπάτια των u προς w):Σ*uω*

Αθροίζοντας το πηλίκο των διαιρέσεων των σημείων 1 και 2 βρίσκουμε το Betweenness Centrality του κόμβου

x. Ο τύπος για την πιο πάνω διαδικασία δίνεται από την σχέση *CB*(*ni*) = Σ(Σ*uω*(*ni*) / Σ*uω*).

Αφού καταλάβαμε πως προκύπτει το Betweenness Centrality, µπορούµε µε την χρήση του Gephi να το βρούμε αυτόματα για όλους τους κόμβους µέσω των Statistics.

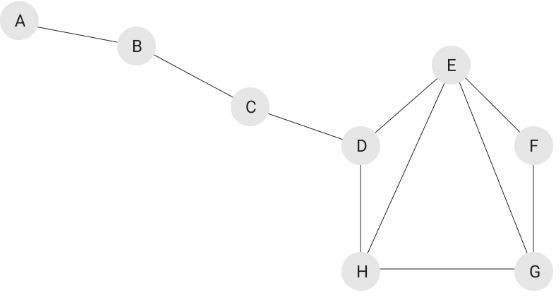
Όπως φαίνεται και από τους πιο πάνω πίνακες λοιπόν, είναι λίγες οι χώρες που έχουν µη µηδενικό Between- ness Centrality. Στην κορυφή των µετρήσεων µας είναι για ακόμη µια φορά το κανάλι της Samsung ενώ έχουν ανέβει στην κορυφή τώρα ορισμένα κανάλια όπου σε προηγούμενες μετρήσεις δεν ήταν σε τόσο υψηλή θέση.

Όπως βλέπουμε, υπάρχουν µία ή περισσότερες χώρες από κάθε ήπειρο εκτός από την Ευρώπη που συγκεντρώνει 7 χώρες.

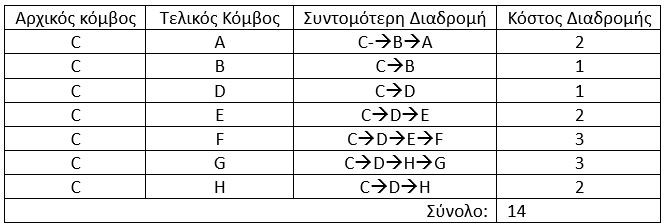
## Closeness Centrality

Το Closeness Centrality είναι µια µετρική που αποσκοπεί στο πόσο κοντά είναι ένας κόμβος σε όλους τους άλλους. Να σημειωθεί επίσης ότι µικρότεροι αριθμοί δείχνουν πως ένας κόμβος έχει υψηλό Closeness Centrality µε τις τιμές να κυμαίνονται από 0 έως 1.

Τον τρόπο µε τον οποίο µπορούµε να υπολογίσουμε τη μετρική αυτή σε ένα κατευθυνόμενο δίκτυο όπως το δικό µας µπορούµε να τον δούμε µέσω του ακόλουθου παραδείγματος.



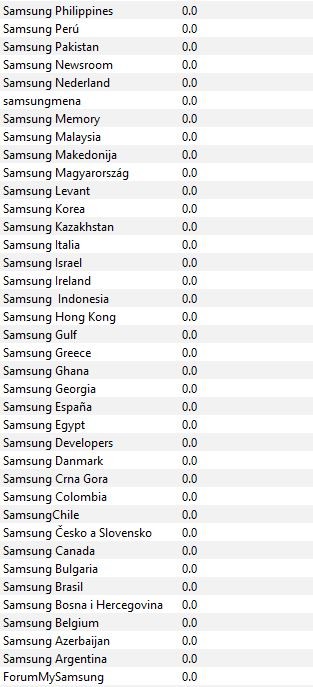
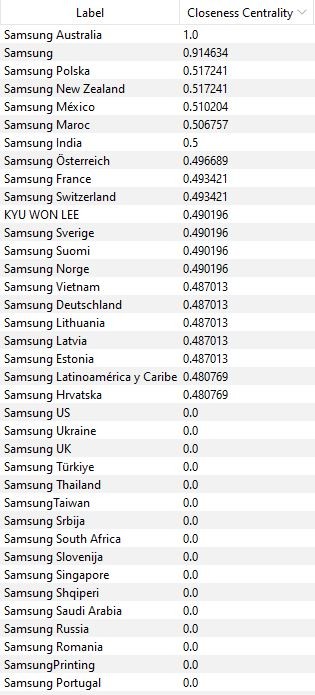
Έστω πως θέλουμε να βρούμε το Closeness Centrality για τον κόμβο C. Βρίσκουμε τον συνολικό αριθμό κόμβων του δικτύου µας και αφαιρούμε ένα, και τον διαιρούμε µε το άθροισα των συντομότερων µονοπατιών από τον κόμβο που εξετάζουμε προς όλους τους υπόλοιπους. Επομένως για τον κόμβο C έχουμε:



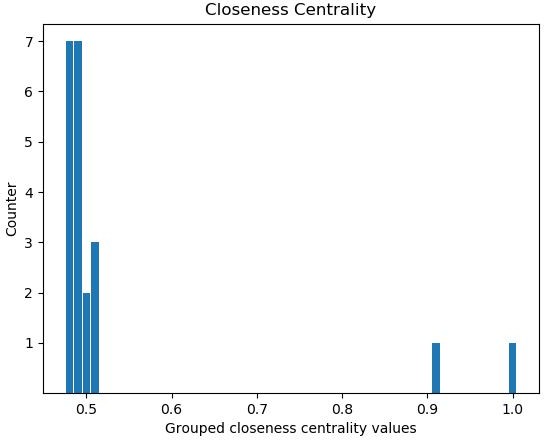
Άρα το Closeness Centrality του κόμβου C είναι 7/14 = 0.5

Να πούμε επίσης πως σε κατευθυνόμενους γράφους, παίζει ρόλο η φορά των ακμών. Για παράδειγμα σε περιπτώσεις όπου δεν υπάρχει διαδρομή λόγω φοράς ακμών, θεωρούμε µηδενικό το µονοπάτι. Έτσι γενικεύοντας το πιο πάνω υπάρχει περίπτωση κάποιος κόμβος να έχει µηδενικό Betweenness Centrality.

Τώρα µε την βοήθεια του Gephi µέσω των Statistics έχουμε τις εξής τιμές για τη µετρική αυτή.



Παρατηρώντας τις πιο πάνω τιμές βλέπουμε πως υπάρχει µια κατανομή για το Closeness Centrality των καναλιών του δικτύου µας. Έτσι για να καταλάβουμε καλύτερα τι συμβαίνει μπορούμε να δούμε τις τιμές αυτές µέσω του ακόλουθου διαγράµµατος με την χρήση των δεδομένων από το Gephi και με την μετατροπή τους σε διάγραμμα μέσω της βιβλιοθήκης matplotlib.pyplot της Python.

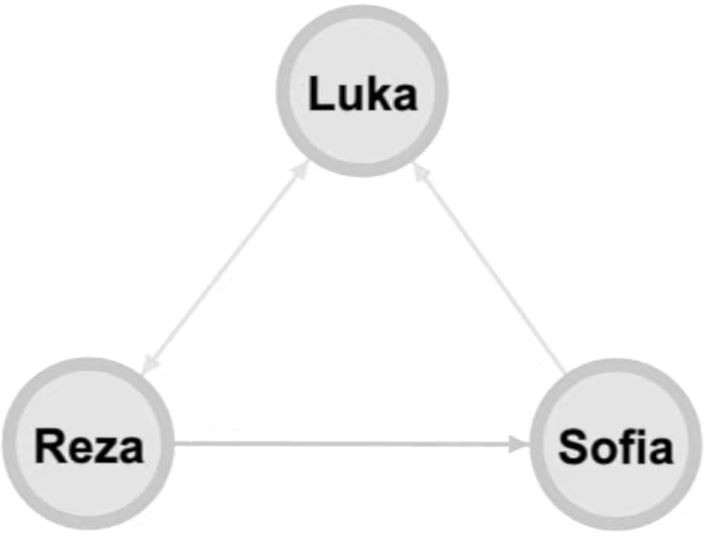


Βλέπουμε λοιπόν πως υπάρχει µια ομαδοποίηση των κόμβων µεταξύ των τιμών 0.48 και 0.51. Αντίθετα όμως, µόνο δύο κόμβοι έχουν υψηλές τιμές της τάξης των 0.91 και 1.0. Άρα αμέσως καταλαβαίνουμε ότι τα κανάλια που έχουν πιο σημαντικό ρόλο στο δίκτυο µας είναι αυτά µε 0.91 και 1.0 µε το όνομα αυτών να είναι Samsung και Samsung Australia αντίστοιχα. Στο σημείο αυτό, να σημειωθεί ότι υπάρχουν αρκετά κανάλια τα οποία έχουν µηδενικό Closeness Centrality αφού είναι κανάλια τα οποία είναι featured channels άλλων καναλιών ενώ ταυτόχρονα τα κανάλια αυτά δεν έχουν δικά τους featured channels.

## Eigenvector Centrality

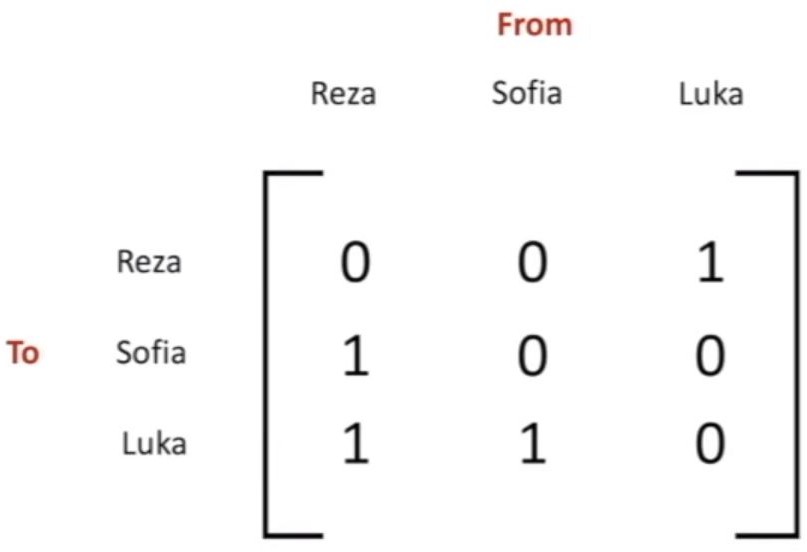
Το Eigenvector Centrality είναι ένα µετρο µε το οποίο µπορούµε να καταλάβουμε την επιρροή που µπορεί να έχει ένας κόμβος µέσα στο δίκτυο µας. ∆είχνει δηλαδή πόσο σημαντικός είναι ένας κόμβος ανάλογα µε το πόσο σημαντικοί είναι και οι κόμβοι-γείτονες που έχει.

Ο τρόπος µε τον οποίο µπορούµε να υπολογίσουμε τη µετρική αυτή σε ένα κατευθυνόμενο δίκτυο όπως το δικό µας µπορούµε να τον δούμε µέσω του εξής παραδείγματος. Έστω το ακόλουθο δίκτυο:



Αρχικά κατασκευάζουμε το γραµµικό σύστημα(πίνακας δύο διαστάσεων) για το δίκτυο µας όπου βάζουμε τους αριθμούς 1 ή 0 εφόσον υπάρχει μονοπάτι που συνδέει τους κόμβους του δικτύου µας ή όχι. Για παράδειγμα για τον κόμβο Reza η πρώτη στήλη στο γραµµικό µας σύστημα θα είναι 0(θεωρούμε πως δεν υπάρχει µονοπατι από κάποιο κόμβο προς τον εαυτό του), 1(για το µονοπάτι Reza προς Sofia) και 1(για το µονοπάτι Reza προς Luke).

Έτσι, µε αυτό το τρόπο προκύπτει ο ακόλουθος πίνακας για το δίκτυο που εξετάζουμε στο παράδειγμα μας.



Στη συνέχεια χρησιμοποιώντας αλγόριθμους γραµµικής άλγεβρας βρίσκουμε το ιδιοδιάνυσµα που αντισ- τοιχεί στη μεγαλύτερη ιδιότιµή του πίνακα που βρήκαμε στο προηγούμενο βήμα, κάνουμε κανονικοποίηση και έτσι προκύπτουν οι τιμές για το Eigenvector Centrality του κάθε κόμβου.

Τώρα µε την βοήθεια του Gephi για το Eigenvector Centrality έχουμε τις εξής τιμές.



Αρχικά, το πρώτο πράγμα που βλέπουμε είναι πως δεν υπάρχουν µηδενικές τιμές σε αντίθεση µε το Close- ness Centrality αφού λαμβάνοντας υπόψιν τον αλγόριθμο µε τον οποίο υπολογίζεται το Eigenvector Centrality όλοι οι κόμβοι είναι ενωμένοι µε όλους έστω και µε µία µονο ακμή.

Πέραν από το πιο πάνω, το πλέον σημαντικό που µπορουµε να πούμε για τα αποτελέσματα αυτά είναι πως υπάρχει µια µεγάλη αρθριτική διαφορά μεταξύ του πρώτου σε σκορ καναλιού και όλων των υπολοίπων. Στη πρώτη θέση λοιπόν βρίσκεται ξανά το κανάλι Samsung µε σκορ 1 ενώ τα υπόλοιπα κανάλια ξεκινούν από σκορ 0.46 και κάτω. Αυτό µας δείχνει ότι οι γείτονες των κόμβων µας είναι το ίδιο περίπου ισχυροί µε µικρή διαφορά κάθε φορά που ολοένα και μικραίνει. Άρα φαίνεται πως επηρεάζονται περίπου το ίδιο κανάλια που βρίσκονται ως featured channels σε άλλα αφού ο σκοπός είναι να φαίνονται όλα τα κανάλια χωρίς να υπάρχει κάποιος ιδιαίτερος λόγος διαχωρισμού προτιμήσεων αφού στο δίκτυο που µελετάµε όλες οι συνδέσεις είναι µε βάση το κριτήριο των featured channels.

# Clustering Effects

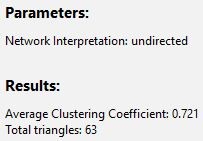
## Average Clustering Coefficient

Το Average Clustering Coefficient µας λέει την πιθανότητα µε την οποία 2 γειτονικοί κόμβοι τυχαία επιλεγμένοι ενός τυχαίου κόμβου να είναι συνδεδεμένοι. Ο τρόπος µε τον οποίο υπολογίζουμε τη µετρική αυτή γίνεται ως εξής. Επιλέγουμε έναν τυχαίο κόμβο όπου στη συνέχεια επιλέγουμε 2 γείτονες του τυχαία και στη συνέχεια ελέγχουμε εάν είναι συνδεδεμένοι. Αθροίζουμε τα αποτελέσματα της παραπάνω διαδικασίας για όλους τους κόμβους, διαιρούμε µε το πλήθος των κόμβων και έτσι βρίσκουμε το επιθυμητό αποτέλεσμα. Στην περίπτωση µας µε τη χρήση του Gephi βρίσκουμε πως το Average Clustering Coefficient για το δίκτυο µας είναι 0.311.



## Number of Triangles

Ο όρος ”Triangle” αναφέρεται σε ένα σύνολο από 3 κόμβους όπου κάθε κόμβος συνδέεται µε τους άλλους δύο. Για να υπολογίσουμε των συνολικό αριθμό τέτοιων συνόλων αρχικά θα πρέπει να θεωρήσουμε πως το δίκτυο µας είναι µη κατευθυνόμενο. Επομένως, µεσω του Gephi θα πρέπει να επιλέξουμε την αντίστοιχη επιλογή για µη κατευθυνόμενο γράφο. Τα αποτελέσματα λοιπόν φαίνονται πιο κάτω.



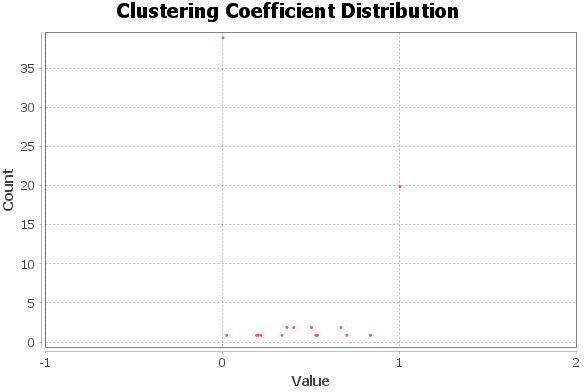
Από το στιγμιότυπο που πήραμε από το Gephi βλέπουμε αρχικά την ύπαρξη **63 Triangles**. Επίσης, παρατηρούμε πως έχει αλλάξει το Average Clustering Coefficient σε σχέση µε πριν. Αυτό οφείλεται στο γεγονός ότι ο γράφος µας είναι µη κατευθυνόμενος µε αποτέλεσμα να προσµετρούνται όλες οι συνδέσεις ανεξάρτητος κατεύθυνσης. Επομένως 2 τυχαίοι γείτονες ενός τυχαίου κόμβου θα συνδέονται µεταξύ τους µε πιθανότητα 0.721 έναντι 0.311 που ήταν προηγουμένως.

Πιο κάτω φαίνονται αναλυτικότερα οι τιμές των Triangles ανά κανάλι.

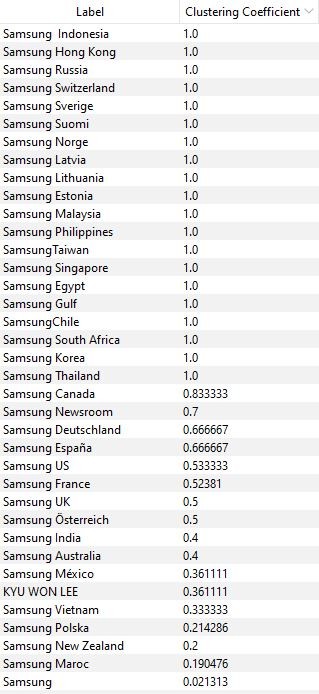
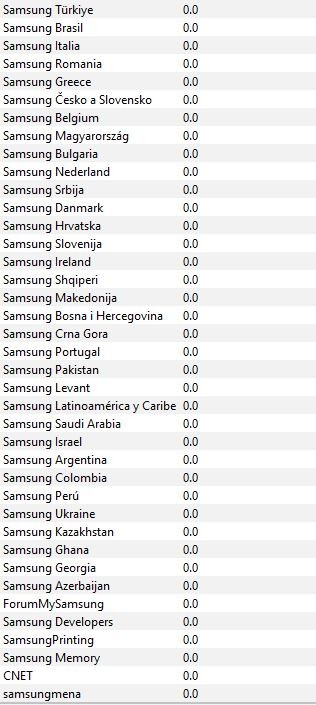


## Clustering Coefficient Distribution

Στο σημείο αυτό, µπορούµε να δούμε ξεχωριστά τις τιμές κάθε κόμβου για τη µετρική του Clustering Coefficient ξεκινώντας από τη γραφική του Gephi.



Όπως βλέπουμε οι τιμές είναι διάσκορπες στο διάστημα 0 έως 1 συµπεριλαµβανοµένων, πράγμα που φαίνεται στην επόμενη σελίδα µέσω των δύο πινάκων.

Αντίθετα µε τις υπόλοιπες µετρικές, τα αποτελέσματα δεν είναι συνηθισμένα. Κανάλια όπου σε προηγούμενες µετρικές είχαν υψηλό σκορ τα βλέπουμε τώρα να έχουν χαμηλό, ενώ κανάλια µε χαμηλό σκορ τώρα τα βλέπουμε να έχουν ψηλό Clustering Coefficient. Αυτό παρατηρείται ιδιαίτερα στα κοινωνικά δίκτυα αφού οι κόμβοι τείνουν να δημιουργούν στενά δεμένες ομάδες που χαρακτηρίζονται από σχετικά υψηλό Clustering Coefficient. Για παράδειγμα µπορούµε να δούμε κάποιες περιπτώσεις καναλιών για να καταλάβουμε πως σχηματίζονται οι ομάδες που αναφέραμε σε σχέση µε το σκορ του Clustering Coefficient.

Κανάλι µε χαµηλό Clustering Coefficient:



Κανάλι µε µεσαίου σκορ Clustering Coefficient:



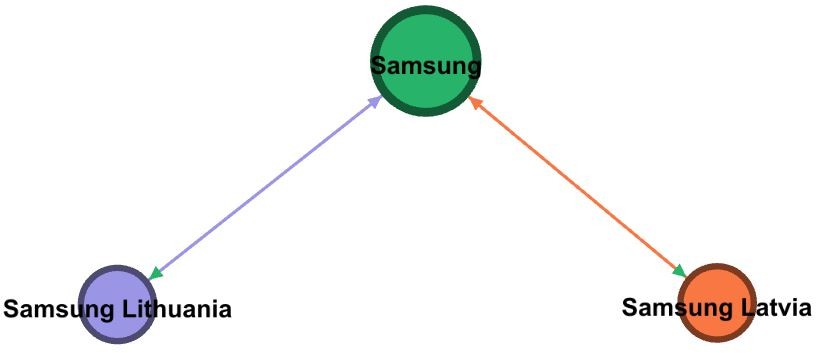
Κανάλι µε υψηλό Clustering Coefficient:



Τα σχήματα αυτά δημιουργήθηκαν ανά ομάδες µε κριτήριο το Clustering Coefficient και γι’ αυτό φαίνονται 3 ομαδοποιήσεις. Σε κάθε διαφορετική εικόνα, έχουμε µε πιο έντονο χρώμα τους κόμβους που συνδέονται µεταξύ τους ανάλογα µε το ύψος του Clustering Coefficient.

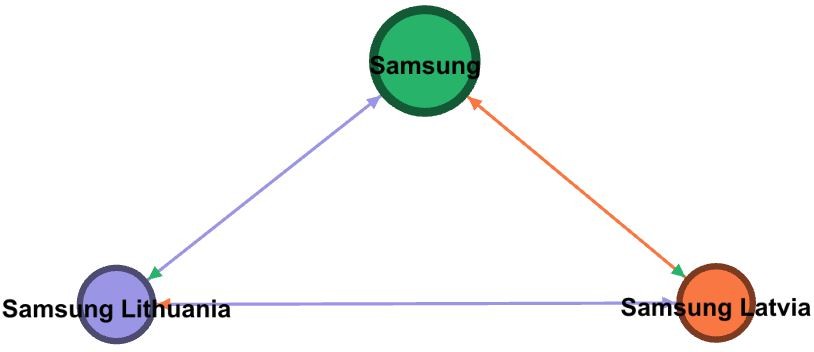
## Existence of the Triadic Closure Phenomenon in the Friendship Neighborhood

Η ύπαρξη τριαδικής κλειστότητας σε Friendship Neighborhood είναι ένα φαινόμενο το οποίο παρακολουθείται σε βάθος χρόνου. Για παράδειγμα στη πραγματική ζωή, όταν δύο άτομα(που δεν γνωρίζονται αρχικά) κάνουν παρέα µε ένα κοινό άτομο, είναι πολύ πιθανό πως μελλοντικά, το κοινό αυτό άτομο που γνωρίζουν θα τους φέρει σε επαφή ώστε να γνωριστούν. Έτσι διευρύνεται ο κύκλος ενός ατόμου µέσω ”Friendship Neighborhoods” από τον κύκλο ενός φίλου. Όσον αφορά το δίκτυο το οποίο εξετάζουμε, αυτό µπορεί να γίνει τόσο εύκολα όπως το σενάριο που περιγράψαμε προηγουμένως. Για παράδειγμα βλέπουμε πιο κάτω τα κανάλια Samsung Lithuania και Samsung Latvia τα οποία συνδέονται µε το κοινό κανάλι Samsung.



Αυτό σημαίνει πως το κανάλι Samsung έχει ως featured channels τα Samsung Lithuania και Samsung Latvia και ταυτόχρονα αυτά τα δύο έχουν ως featured channel το κανάλι Samsung. Επομένως αφού οι χρήστες του καναλιού Samsung Lithuania βλέπουν προτεινόμενα βίντεο για το κανάλι Samsung, και οι χρήστες του καναλιού Samsung βλέπουν προτεινόμενα βίντεο για το κανάλι Samsung Latvia, τότε θα είναι πιθανό οι χρήστες του Samsung Lithuania να ενδιαφέρονται για τα βίντεο του Samsung Latvia.

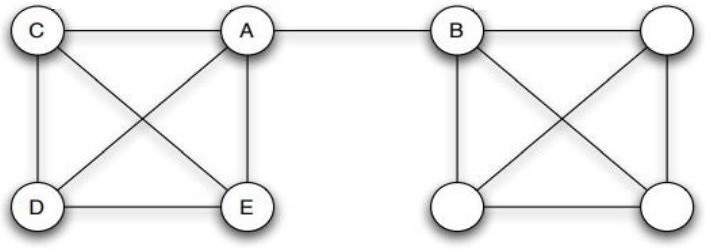
Αυτό ισχύει και για την αντίθετη περίπτωση ξεκινώντας από το κανάλι Samsung Latvia και καταλήγοντας στο κανάλι Samsung Lithuania.

Άρα µπορεί να γίνει η σύνδεση µε τα featured channels που περιγράψαμε όπως φαίνεται στην πιο κάτω εικόνα.

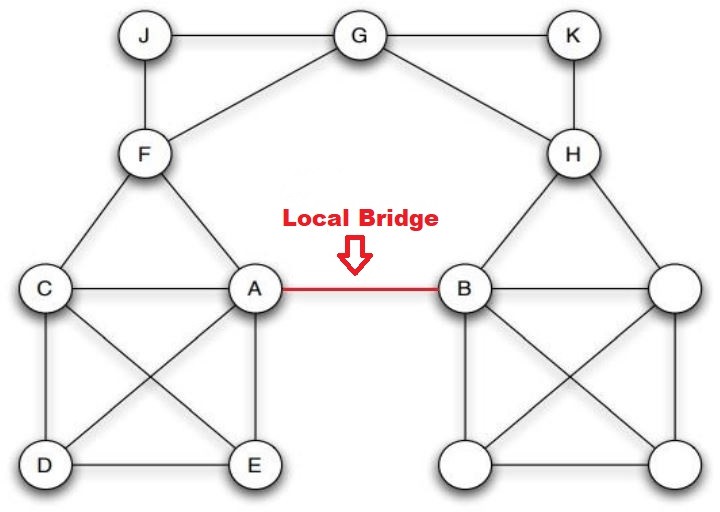
Το αποτέλεσμα µπορεί να είναι όπως φαίνεται στην τελευταία εικόνα χωρίς να είναι βέβαια απαραίτητη η αμφίδρομή σύνδεση.

# Bridges and Local Bridges

Γέφυρα ή αλλιώς Bridge στην αγγλική ορολογία θεωρούμε την µοναδίκη ακμή η οποία συνδέει 2 κόμβους όπου κάθε ένας από αυτούς ανήκει σε ένα µία διαφορετική γειτονιά(ειδικότερα όταν η κάθε γειτονιά έχει πιο ισχυρούς δεσμούς µεταξύ της). Με άλλα λόγια, αν η ακμή αυτή διαγραφτεί, θα χωρίσει το δίκτυο στο οποίο ανήκει στα δύο. Ένα παράδειγμα τέτοιου δικτύου φαίνεται πιο κάτω.

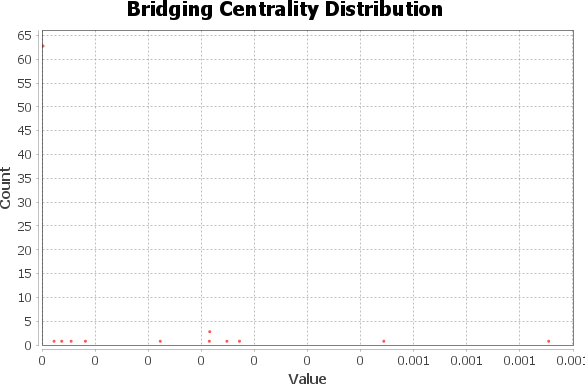


Έστω πως έχουμε παρόμοιο δίκτυο µε το προηγούμενο, µόνο που η σύνδεση των δύο κόμβων που αποτελεί την ένωση των δύο κοινοτήτων δεν είναι η µοναδική ως εξής:



Στη περίπτωση δηλαδή όπου υπάρχουν περισσότερες από µία ακμές θεωρούμε ως Local Bridge την ακμή όπου η διαγραφή της θα αυξήσει την απόσταση µεταξύ των 2 κόμβων που συνέδεε.

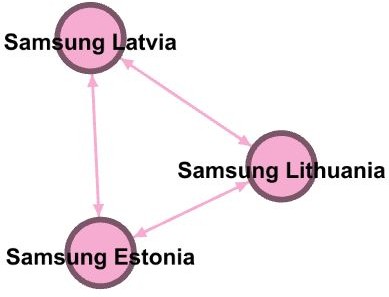
Στο σημείο αυτό θα στραφούμε ξανά στο Gephi για να εξετάσουμε το δίκτυο µας µε βάση το Bridging Centrality. Οι τιμές που έχουν εμφανιστεί είναι πολύ κoντά στο μηδέν όπως βλέπουμε. Μέσα από αυτές τις τιμές όμως και µέσα από διάφορους ελέγχους σε layouts δεν μπορέσαμε να βρούμε σχηματισμούς που να δείχνουν την αυστηρή ύπαρξη Bridges αφού πολλές κοινότητες συνδέονται µέσω αρκετών ακμών. Αυτό που μένει δηλαδή µέσα από αυτή την ενότητα είναι το πως τα κανάλια τα οποία έχουν αρκετές συνδέσεις φαίνεται να ανήκουν σε µια γειτονία µε ισχυρή αλληλεπίδραση και επηρεασμό σε σχέση µε τα άλλα κανάλια. Έτσι ερχόμαστε για άλλη µια φορά να φέρουμε στο μυαλό µας την λογική µε την οποία έχουν στηθεί οι συνδέσεις αυτές, που δεν είναι άλλη πέραν από τη λογική των Featured Channels.



# Gender and Homophily

Homophily είναι µια αρχή κατά την οποία οι άνθρωποι τείνουν να μοιάζουν µε τους φίλους τους. Η βάση για αυτή την ιδέα φαίνεται να πηγάζει µέσα από τις γραφές του Πλάτωνα(”Η ομοιότητα γεννά φιλία”) και του Αριστοτέλη(”Οι άνθρωποι ‘αγαπούν αυτούς που είναι σαν τον εαυτό τους’ ”). Η τάση αυτή φαίνεται από τα αρχαία χρόνια ενώ στον σύγχρονο κόσμο και στην καθημερινότητα µας παρατηρούμε επίσης το φαινόμενο αυτό. Από τα αρχικά στάδια της ζωής µας µέσα από το σχολείο μέχρι την ενηλικίωση και την μετέπειτα ζωή µας(δουλειά κλπ.) τείνουμε να κάνουμε παρέα, να επηρεάζουμε και να επηρεαζόμαστε από άτομά του κύκλου µας τον οποίο εμείς επιλέγουμε µε βάση ομοιότητες που συναντούμε στα άτομα αυτά.

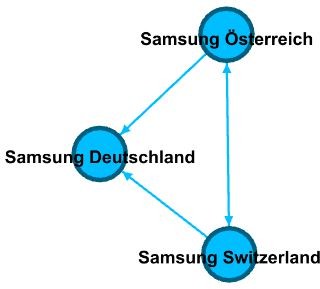
Στο σημείο αυτό λοιπόν θα δούμε πως το φαινόμενο αυτό παρατηρείται στο δικό µας δίκτυο. Θα εξετάσουμε περιπτώσεις καναλιών όπου η συναναστροφή τους οφείλεται σε ομοιότητες που υπάρχουν στα κανάλια αυτά. Για παράδειγμα, η περίπτωση των τριών καναλιών Samsung Latvia, Samsung Lithuania και Samsung Estonia.



Παρατηρούμε πως οι συνδέσεις είναι αμφίδρομες για όλα τα κανάλια σε αυτή τη κοινότητα. Αυτό δείχνει µια ιδιαίτερα ισχυρή σύνδεση στη κοινότητα αυτή. Επίσης οι χώρες από τις οποίες αποτελούνται τα κανάλια αυτά είναι πολύ κοντά γεωγραφικά, έχουν παρόμοια ιστορία και πολιτισμό αφού ήταν όλες πρώην χώρες της Σοβιετικής Ένωσης. Επομένως φαίνεται να επηρεάζει η µία την άλλη ή ακόμα να υπάρχει πιο έντονη ροή δεδομένων και πληροφοριών σε σχέση µε χώρες που βρίσκονται πιο μακριά αφού μιλάμε για ένα δίκτυο που αποτελείται από κανάλια στο YouTube µε σκοπό την παρουσίαση καινούργιων τεχνολογιών και επομένως την

ροή της πληροφορίας. Επομένως επαληθεύεται αυτό που είπαμε στην αρχή.

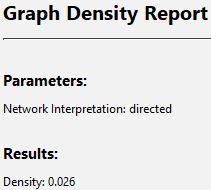
Άλλο ένα παράδειγμα τέτοιου είδους, είναι µε τα κανάλια Samsung Switzerland, Samsung Deutschland(Γερμανία) και Samsung Osterreich(Αυστρία).

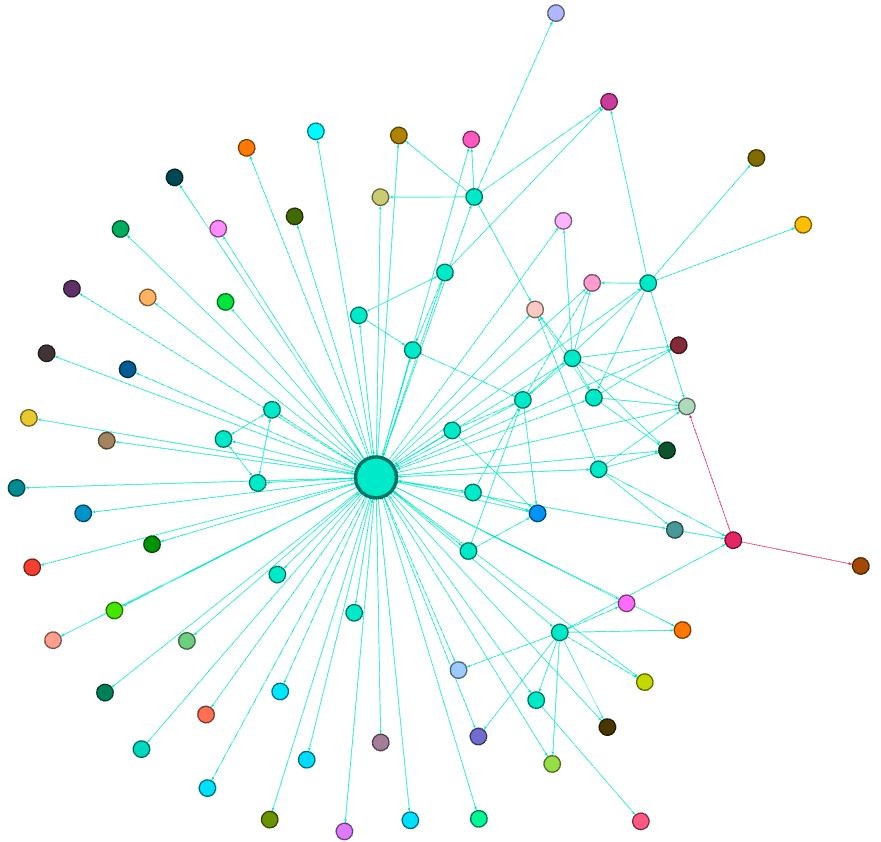


΄Όπως και πριν, παρατηρούμε κανάλια από χώρες της Ευρώπης τα οποία έχουν ισχυρή σύνδεση. Οι χώρες από τις οποίες προέρχονται είναι γειτονικές ενώ αυτή τη φορά έχουν ακόμα ένα κοινό, τη γλώσσα. Λογικό, αφού αυτό βοηθάει τα προτεινόμενα κανάλια να ελκύουν έτσι περισσότερο κοινό.

# Graph Density

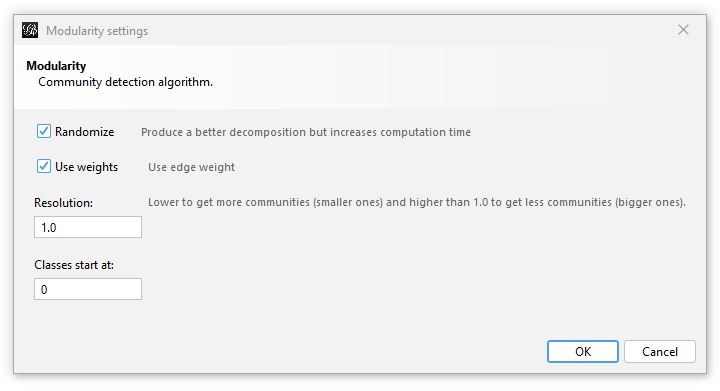
Το Graph density είναι µια µετρική για την συνδεσιμότητα των κόμβων σε ένα γράφημα. Υπολογίζεται µε το γινόμενο των ακμών που έχει το δίκτυο προς τον συνολικό αριθμό ακμών που θα μπορούσε να έχει εάν όλοι οι κόμβοι ήταν συνδεδεμένοι µε όλους. Στο δίκτυο µας όπως είδαμε και στην ενότητα 4 για τα Βασικά στοιχεία

Δικτύου, ο αριθμός κόμβων είναι 76 και ο αριθμός ακμών είναι 149. Άρα αφού πρόκειται για κατευθυνόμενο γράφο, ο συνολικός αριθμός ακμών που θα μπορούσε να έχει το δίκτυο µας είναι 76\*75 = 5700. Επομένως το Graph Density για το δίκτυο µας είναι 149 / 5700 = 0.026. Πράγματι αυτό επιβεβαιώνεται µέσω του Gephi.

Το αποτέλεσμα που βρήκαμε είναι αρκετά μικρός αριθμός. Λογικό, αφού αν ανατρέξουμε στο δίκτυο µας θα δούμε πως 36 στους 76 κόμβους είναι ενωμένοι µόνο µε έναν άλλο, ποσοστό που αγγίζει το 47.3% των κόμβων του δικτύου µας. Αυτό φαίνεται µέσω του Total Degree που είναι ίσο µε 1 για τους κόμβους αυτούς το οποίο είχαμε αναλύσει στην ενότητα 6.

# Community Structure(Modularity)

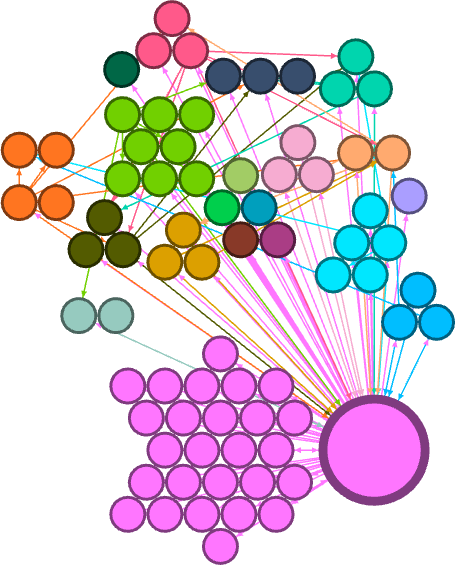
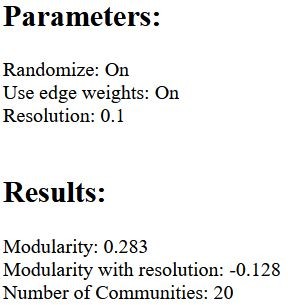
Στο σημείο αυτό, µπορούµε να δούμε πως το δίκτυο µας διαχωρίζεται σε communities ούτως ώστε να εξάγουμε κάποια συμπεράσματα. Αυτό µπορεί να γίνει µε τη χρήση του Modularity µέσω των Statistics του Gephi ώστε να δούμε τα διαφορετικά Communities που σχηματίζονται. Τρέχοντας λοιπόν το στατιστικό Modularity, εμφανίζεται το πιο κάτω παράθυρο.



Όπως βλέπουμε, υπάρχουν ορισµένες επιλογές οι οποίες µπορούν να ληφθούν υπόψιν όπως το randomize, το βάρος των ακμών, το Resolution και η αρχική τιμή από την οποία θα ξεκινάει η αρίθμηση των κόμβων για τη μετρική που θα παραχθεί. Στα πλαίσια της ανάλυσης µας ενδιαφέρει το Resolution αφού αυτό είναι που διαμορφώνεται το πλήθος των communities ανάλογα µε την τιμή που θα πάρει(μικρότερες τιμές έχουν αποτέλεσμα πιο πολλά communities και μεγαλύτερες τιμές λιγότερα communities). Να σημειωθεί επίσης πως το βάρος των ακμών είναι ίδιο σε όλο µας το δίκτυο ενώ ο βαθμός στον οποίο επηρεάζει τα αποτελέσματα µας είναι ελάχιστος έως μηδαμινός σύμφωνα µε μετρήσεις που έχουν πραγματοποιηθεί. Επομένως αφού το Resolution καθορίζει τα αποτελέσματα µας, μετά από αρκετές δοκιμασίες που κάναμε για να καταλάβουμε πως διαμορφώνονται τα communities καταλήξαμε στα πιο κάτω αποτελέσματα.

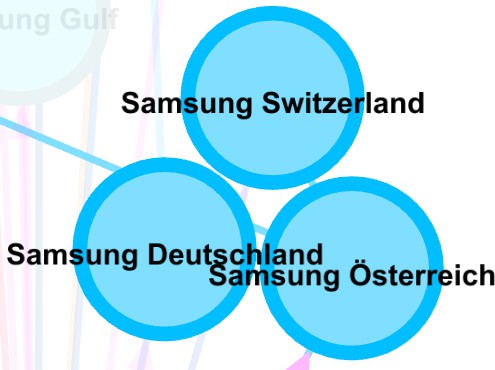
## Community Structure - Modularity Resolution 0.1

Με Resolution 0.1 έχουμε Modularity 0.283 ενώ δημιουργούνται 20 communities διαφόρων μεγεθών καθώς τα μικρότερα αποτελούνται από 1 µονο κόμβο(7 communities) ενώ το µεγαλύτερο από 26. Όπως βλέπουμε υπάρχει µεγάλη διαφορά στο πλήθος των communities καθώς επίσης και µεγάλη κατανομή.



Να πούμε επίσης πως σε μικρές ομαδοποιήσεις τα κανάλια από τα οποία αποτελούνται είναι από χώρες που βρίσκονται γεωγραφικά κοντά ή στην ίδια ήπειρο όπως για παράδειγμα:

Το community µε ροζ χρώμα αποτελείται από τα κανάλια Sam- sung Latvia, Samsung Lithuania, Samsung Estonia.

Το community µε μπλε χρώμα αποτελείται από τα κανάλια Sam- sung Switzerland, Samsung Deutschland(Γερμανία), Samsung Os- terreich(Αυστρία).

Το community µε πράσινο χρώμα αποτελείται από τα κανάλια Samsung New Zealand, Samsung Singapore, Samsung Malaysia, Samsung Hong Kong, Samsung Taiwan, Samsung Philippines, Sam- sung Thailand και Samsung Indonesia. Τα κανάλια αυτά είναι από χώρες οι οποίες μπορεί να µην είναι τόσο κοντά όπως στις προηγούμενες περιπτώσεις που είδαμε αλλά ανήκουν σχεδόν όλες στην ίδια ήπειρο εκτός από την Νέα Ζηλανδία.

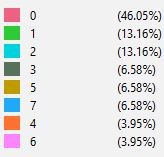
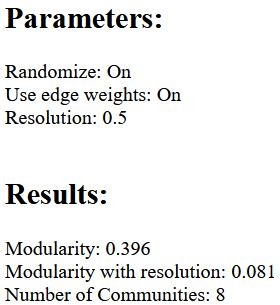
Το community µε γαλάζιο χρώμα αποτελείται από τα κανάλια Samsung South Africa, Samsung Egypt, Samsung Maroc, samsung- mena, Samsung Gulf. Σε αυτό το community βλέπουμε κανάλια από χώρες της Βορείου Αφρικής(Ηπείρου) και της Μέσης Ανα- τολής.

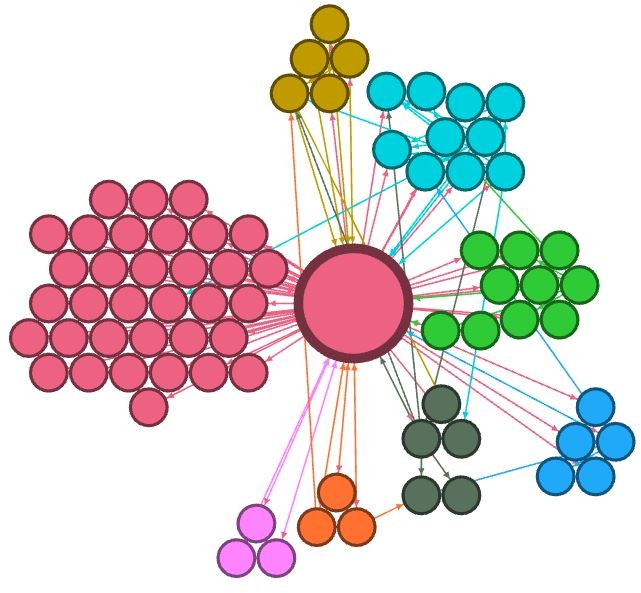
Το community µε έντονο ροζ χρώμα αποτελείται από τα κανάλια Samsung South Canada, Samsung Chile, Samsung Mex- ico. Σε αυτό το community βλέπουμε κανάλια από χώρες της Αμερικής.

Τα υπόλοιπα communities αποτελούνται από κανάλια στα οποία οι χώρες που προέρχονται δεν έχουν κάποια σχέση ή είναι communities µε ένα και μοναδικό κανάλι.

## Community Structure - Modularity Resolution 0.5

Με Resolution 0.5 έχουμε Modularity 0.396 ενώ δημιουργούνται 8 communities μεγέθους 3, 3, 5, 5, 5, 10, 10 και 35 κόμβων. Ακολουθούν τα ποσοστά σε σχέση µε το µέγεθος του δικτύου καθώς επίσης οι μετρήσεις και ο τρόπος διάταξης του συγκεκριμένου community structure.



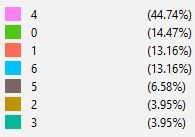
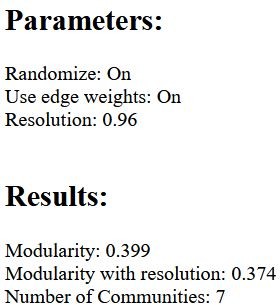


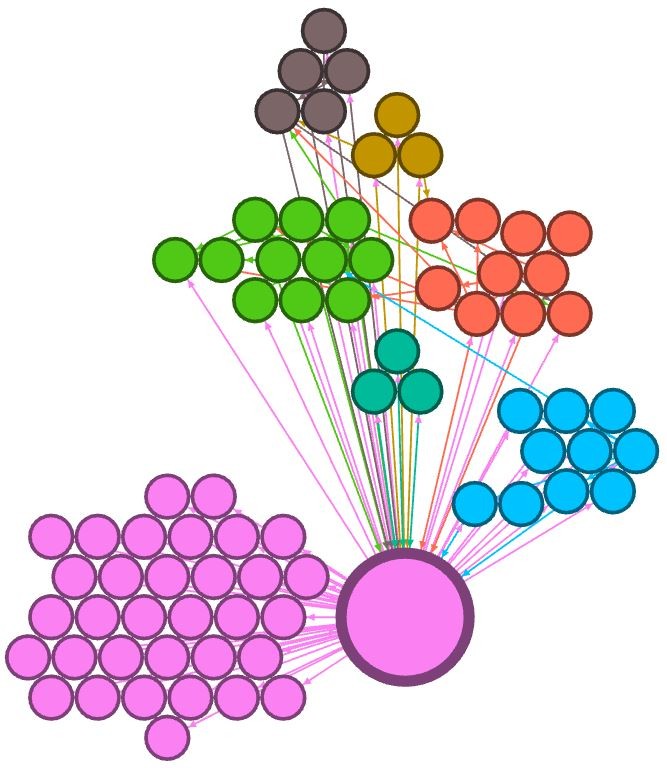
Στο σημείο αυτό και µε τα στατιστικά που έχουμε πάρει, καταλαβαίνουμε προφανώς πως ο αριθμός των communities έχει μειωθεί, δεν υπάρχουν πλέον ατομικά communities αφού έχουν συγχωνευθεί µε άλλα, ενώ ταυτόχρονα ορισμένα άλλα κανάλια έχουν ενταχθεί σε προηγούμενα communities που είχαμε στη προηγούμενη υπό-ενότητα µε Resolution 0.1. Πιο συγκεκριμένα έχουμε:

Στο community όπου προηγουμένως αποτελείτο από τα κανάλια Samsung New Zealand, Samsung Singapore, Samsung Malaysia, Samsung Hong Kong, Samsung Taiwan, Samsung Philippines, Sam- sung Thailand και Samsung Indonesia πλέον έχουν προστεθεί τα κανάλια Samsung Vietnam και ForumMySamsung. Το γεγονός αυτό δεν µας εκπλήσσει αφού πάλι πρόκειται για κανάλια που απευθύνονται σε κοινό από ασιατικές χώρες, προφανώς για το Βιετνάμµ αλλά και για το κανάλι ForumMySamsung αφού όπως είδαμε το περιεχόμενο του είναι στα Βιετναμέζικα. Δίπλα φαίνεται η καινούργια μορφή του συγκεκριμένου community.

## Community Structure - Modularity Resolution 0.9

Με Resolution 0.9 έχουμε Modularity 0.399 και δημιουργούνται 7 communities µεγέθους 3, 3, 5, 10, 10, 11 και 34 κόμβων. Ακολουθούν τα ποσοστά σε σχέση µε το µέγεθος του δικτύου καθώς επίσης οι μετρήσεις και ο τρόπος διάταξης του συγκεκριμένου community structure.

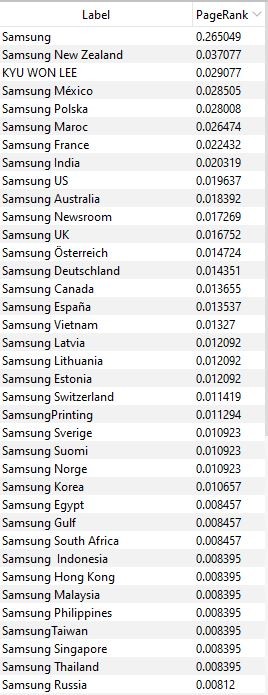




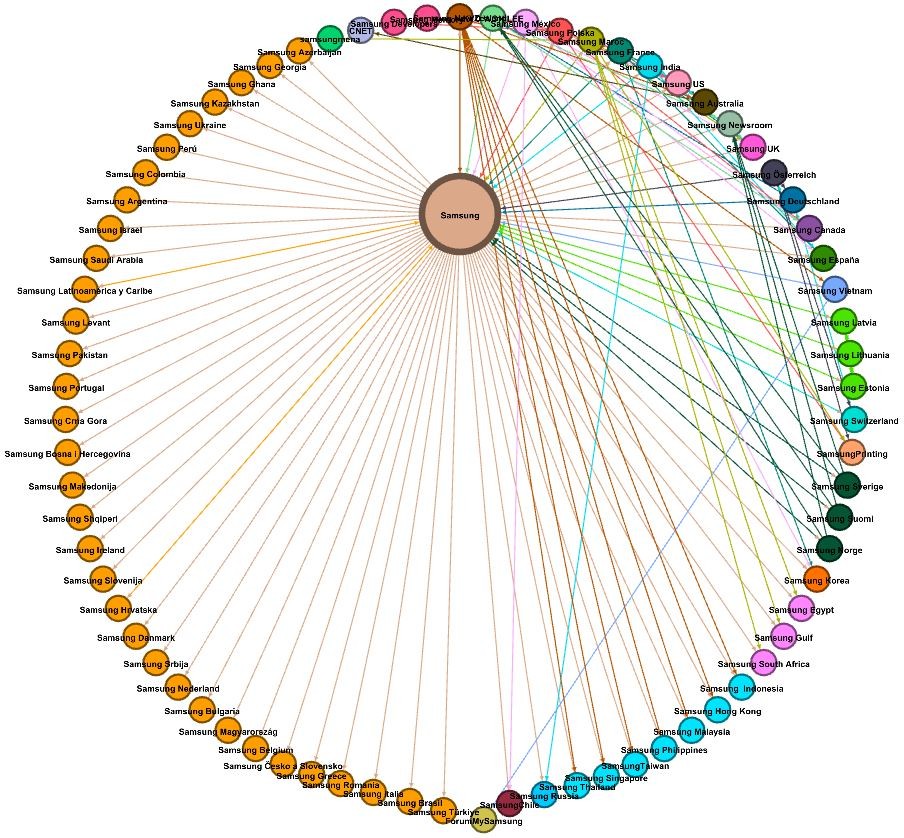
Στη παρούσα φάση όπως και πριν, βλέπουμε να συγκεντρώνονται ακόμη περισσότερο οι κόμβοι σε πιο μεγάλα communities. Πλέον φαίνεται να ξεπερνάμε το μοντέλο της ομαδοποίησης κατά χώρες. Αυτό δεν µπορούµε να πούμε πως είναι κακό αφού το κριτήριο µε το οποίο υπάρχουν οι ακμές στο δίκτυο µας δεν καθορίζεται από τις χώρες στις οποίες βρίσκονται τα κανάλια αλλά παίζει σίγουρα σημαντικό ρόλο αφού όπως είδαμε γίνονται ομαδοποιήσεις ανάλογα µε την περιοχή στην οποία βρίσκονται γεωγραφικά τα κανάλια και τη γλώσσα στην οποία μιλούν. Επομένως, μετά από όσα προηγήθηκαν, καταλαβαίνουμε ότι κάθε Resolution έπαιξε το δικό του ρόλο ώστε να εξάγουμε τα συμπεράσματα αυτά.

# PageRank

Το PageRank αποτελεί έναν αλγόριθμο ο οποίος αρχικά αναπτύχθηκε από την Google για την κατάταξη των ιστοσελίδων µε σκοπό την καλύτερη και πιο αποτελεσματική εμφάνιση τους στις αναζητήσεις των χρηστών. Ο αλγόριθμος αυτός εφαρμόζεται και στην ανάλυση δικτύων αφού χρησιμοποιείται για την μέτρηση της σημασίας ή της επιρροής των κόμβων σε ένα δίκτυο µε την ανάθεση αριθμητικής τιμής η οποία αντιπροσωπεύει την σημαντικότητα ενός κόμβου στο δίκτυο. Ο υπολογισμός του PageRank σε ένα δίκτυο γίνεται µε βάση τον αριθμό και την ποιότητα των συνδέσεων σε ένα κόμβο ενώ επίσης κόμβοι µε υψηλότερη τιμή θεωρούνται σημαντικότεροι. Πιο κάτω, βλέπουμε τις μετρήσεις των κόμβων για το δίκτυο µας µεσω του Gephi.



Όπως ήταν αναμενόμενο για άλλη µια φορά στην κορυφή βρίσκεται το κανάλι Samsung. Τα υπόλοιπα κανάλια, τα βλέπουμε να ακολουθούν µε αρκετά µεγάλη διαφορά σε σχέση µε το σκορ που σημειώνει το πρώτο κανάλι. Επομένως σύμφωνα µε το PageRank το πιο σημαντικό κανάλι είναι της Samsung. Εκτός από αυτό, παρατηρούμε πως οι τιμές των υπόλοιπων κόμβων είναι αρκετά κοντινές ενώ επίσης συναντώνται ομαδοποιήσεις. Για να δούμε σχηματικά τη διαφορά, µπορούµε να διαχωρίσουμε και να ταξινομήσουμε τους κόμβους ως προς τη µετρική που εξετάζουμε. Τα αποτελέσματα ακολουθούν πιο κάτω.



Ο σχηματισμός έγινε µε βάση το Dual Circle Layout ενώ ο χρωματισμός και η σειρά των κόμβων στον εξωτερικό κύκλο µε βάση το PageRank. Με δεδομένο ότι υπάρχει ένας µονο κόμβος µε σημαντικά μεγάλο σκορ βάλαμε λοιπόν ένα µόνο κόμβο στο εσωτερικό του σχήματος οπότε δεν µπορεί να σχηματιστεί κάποιος εσωτερικός κύκλος. Με αυτό το τρόπο όμως µπορούµε εύκολα να διακρίνουμε τους κόμβους που έχουν τις λιγότερες συνδέσεις και επομένως το μικρότερο PageRank αφού ενώνονται µόνο µε έναν κόμβο. Μιλάμε βέβαια για τους κόμβους µε τον πορτοκαλί χρωματισμό. Έτσι συνεχίζοντας στον κύκλο µπορούµε να δούμε πως αυξάνεται το PageRank µε βάσει τον αριθμό των συνδέσεων των κόμβων.

Τα πιο πάνω αποτελέσματα οφείλονται στο ότι οι κόμβοι µεταξύ τους δεν έχουν την ιδιαίτερη σχέση όπως έχει ο κόμβος που αντιπροσωπεύει το κανάλι της Samsung µε όλους τους υπόλοιπους αφού όπως είπαμε και

στην αρχή, η σχέση µε την οποία συνδέονται τα κανάλια είναι µέσω των featured channels ξεκινώντας από έναν µόνο seeder.

# Συµπεράσµατα

Στην παρούσα ανάλυση, µε αφετηρία το κανάλι της Samsung, είδαμε πως η πληροφορία διαχέεται σε όλο τον κόσμο µε βάση τα κανάλια που αντιπροσωπεύουν την κάθε χώρα. Ξεκινώντας από το κανάλι ”πηγή” είδαμε ποια κανάλια συμβάλουν στην γνωστοποίηση περιεχομένου του καναλιού αυτού. Σε αρκετές περιπτώσεις, σημαντικό ρόλο παίζει το ίδιο το κανάλι το οποίο αναλύσαμε όπως για παράδειγμα στις περιπτώσεις των Betweenness Centrality, Eigenvector Centrality και PageRank. Είδαμε να δημιουργούνται communities µε βάση την κουλτούρα, τη γλώσσα και την τοποθεσία της χώρας από την οποία προέρχεται το κάθε κανάλι. Παρατηρήσαμε στο φαινόμενο της ομοφυλίας, κανάλια να έχουν ως Featured Channels κανάλια γειτονικών χωρών και κανάλια που να απευθύνονται σε κοινό µε ίδια γλώσσα και πολιτισμό. Παρατηρούμε δηλαδή πως πράγματα τα οποία συναντούμε στον πραγματικό κόσμο, ταυτίζονται και στους εικονικούς αυτούς κόστους όπου εμείς οι ίδιοι δημιουργούμε. Επομένως μιλάμε για ένα δίκτυο το οποίο ζει και αναπτύσσεται συνεχώς µέσα από τις διάφορες ενέργειες και τάσεις των χρηστών.