

ΟΙΚΟΝΟΜΙΚΟ  
ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΑΘΗΝΩΝ



ATHENS UNIVERSITY  
OF ECONOMICS  
AND BUSINESS

# ΤΗΚΕΥΑΣΕ ΚΣΑΝΑ ΕΚΦΟΝΗΣΙ ΓΙΑ ΤΟ ΤΙ ΖΗΤΑ

Your Subtitle

February 15, 2024

# Περιεχόμενα

1	Εισαγωγή. . . . .	3
2	Λήψη δεδομένων . . . . .	3
3	Γραφική Αναπαράσταση Δικτύου . . . . .	4
4	Βασικά Στοιχεία Δικτύου . . . . .	7
5	Component Measures . . . . .	8
6	Degree Measures .....	10
6.1	Maximum Degree.....	10
6.2	Average Node Degree .....	10
6.3	Degree Distribution .....	11
6.3.1	In-Degree.....	11
6.3.2	Out-Degree .....	13
6.3.3	Total Degree .....	16
7	Centrality measures.....	19
7.1	Betweenness Centrality.....	19
7.2	Closeness Centrality .....	21
7.3	Eigenvector Centrality .....	24
8	Clustering Effects.....	27
8.1	Average Clustering Coefficient.....	27
8.2	Number of Triangles .....	28
8.3	Clustering Coefficient Distribution .....	30
8.4	Existence of the Triadic Closure Phenomenon in the Friendship Neighborhood .....	34
9	Bridges and Local Bridges.....	36
10	Gender and Homophily .....	38
11	Graph Density .....	40
12	Community Structure(Modularity).....	41
12.1	Community Structure - Modularity Resolution 0.1 .....	42
12.2	Community Structure - Modularity Resolution 0.5 .....	44

12.3	Community Structure - Modularity Resolution 0.9 .....	46
13	PageRank.....	48
14	Συμπεράσματα.....	51

# 1 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.

## 2 Λήψη Δεδομένων

Τα δεδομένα για την ανάλυση μας τα πήραμε με τη χρήση του [Bernhard Reiner's Tool](#) χρησιμοποιώντας τα YouTube Data Tools. Αρχικά, χρησιμοποιώντας το link του καναλιού στο YouTube, βρήκαμε το id του καναλιού μέσω του [Channel Info Module](#). Έπειτα με τη χρήση του [Channel Network Module](#), 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 μπορούμε 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 μπορούσαν να

### 3 Γραφική Αναπαράσταση Δικτύου







## 4 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**. Δηλαδή 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.

- Διάμετρος δικτύου: Η **διάμετρος** ενός δικτύου είναι η μακρύτερη συντομότερη διαδρομή που μπορούμε να βρούμε. Στην περίπτωση 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**.

### Parameters:

Network Interpretation: directed

### Results:

Diameter: 3

Radius: 0

AveragePathlength: 1.9760319573901464



## 5 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.

### Connected Components Report

---

#### Parameters:

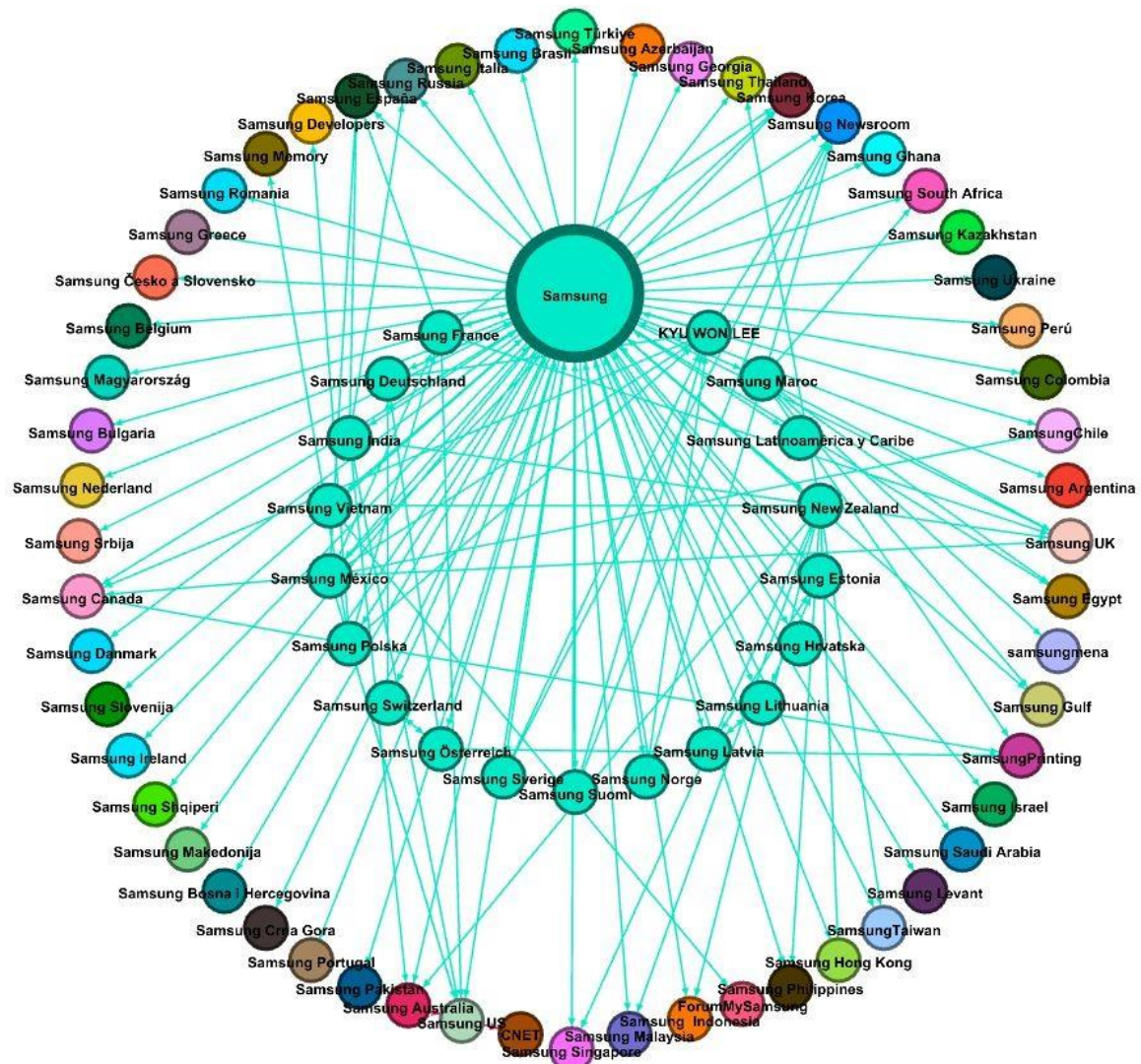
Network Interpretation: directed

#### Results:

Number of Weakly Connected Components: 1

Number of Strongly Connected Components: 57

Παρατηρώντας την πιο πάνω εικόνα λοιπόν, μπορούμε να επιβεβαιώσουμε τον αριθμό των weakly connected components. Όσον αφορά τον αριθμό των strongly connected components μέσω 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.

## 6 Degree Measures

### 6.1 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.

Label	Degree ▾
Samsung	87
Samsung New Zealand	11
Samsung México	10
Samsung Polska	9
KYU WON LEE	9
Samsung France	8
Samsung Maroc	8
Samsung India	7
Samsung US	6
Samsung Österreich	6
Samsung Latvia	6

### 6.2 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 Report

---

##### Results:

Average Degree: 1.961

## 6.3 Degree Distribution

### 6.3.1 In-Degree

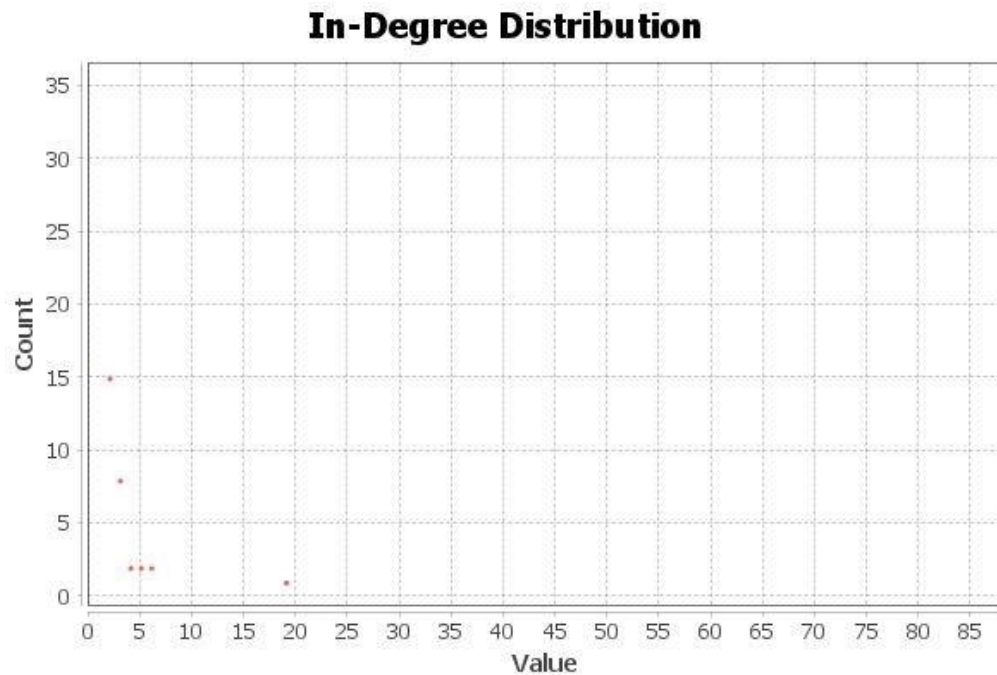
Το In-Degree είναι οι εισερχόμενες προς κάποιον κόμβο ακμές. Στην περίπτωση μας, ο αριθμός αυτός αποτελεί τον αριθμό των καναλιών που έχουν ως Featured Channel το κανάλι που εξετάζουμε. Έτσι για κάθε κανάλι με τη βοήθεια του Gephi για το δίκτυο μας έχουμε:

Label	In-Degree ▾
Samsung	19
KYU WON LEE	6
Samsung US	6
Samsung Newsroom	5
Samsung UK	5
Samsung Canada	4
Samsung España	4
Samsung Australia	3
Samsung Deutschland	3
Samsung Estonia	3
Samsung France	3
Samsung Korea	3
Samsung Latvia	3
Samsung Lithuania	3
SamsungPrinting	3
SamsungChile	2
Samsung Egypt	2
Samsung Gulf	2
Samsung Hong Kong	2
Samsung Indonesia	2
Samsung Malaysia	2
Samsung Österreich	2
Samsung Philippines	2
Samsung Russia	2
Samsung Singapore	2
Samsung South Africa	2
Samsung Switzerland	2
SamsungTaiwan	2
Samsung Thailand	2
Samsung Vietnam	2
CNET	1
ForumMySamsung	1
Samsung Argentina	1
Samsung Azerbaijan	1
Samsung Belgium	1
Samsung Bosna i Hercegovina	1
Samsung Brasil	1

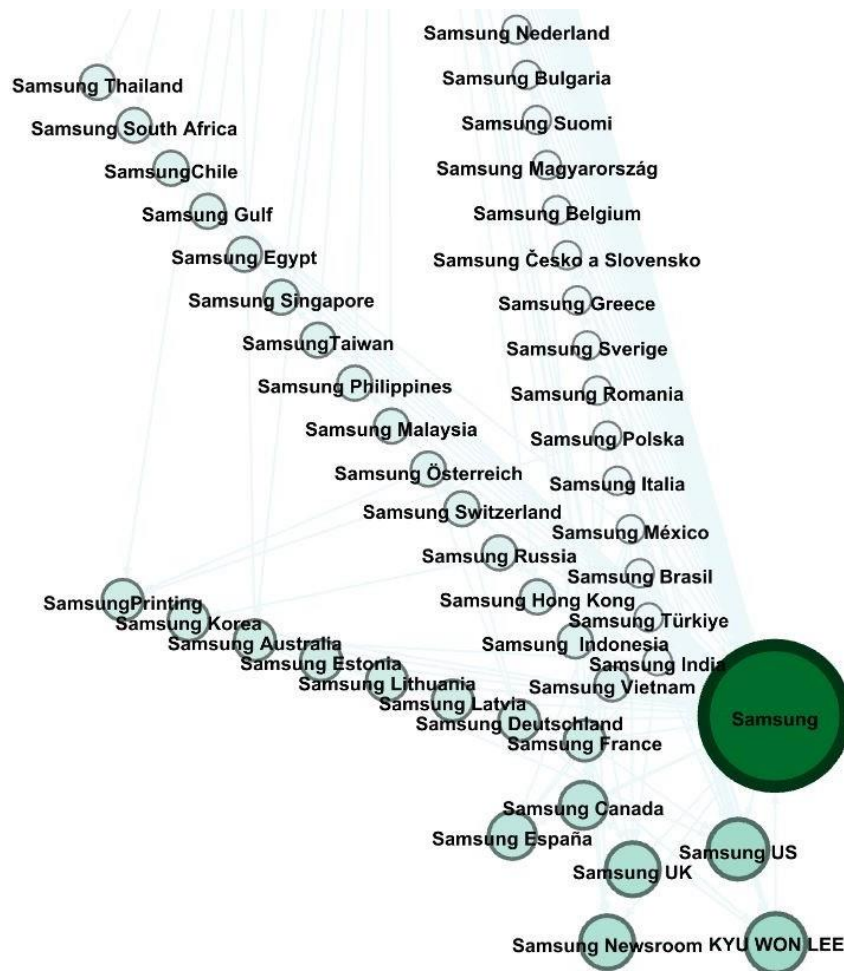
Samsung Suomi	1
Samsung Bulgaria	1
Samsung Nederland	1
Samsung Srbija	1
Samsung Norge	1
Samsung Danmark	1
Samsung Hrvatska	1
Samsung Slovenija	1
Samsung Ireland	1
Samsung Shqiperi	1
Samsung Makedonija	1
Samsung Bosna i Hercegovina	1
Samsung Crna Gora	1
Samsung Portugal	1
Samsung Pakistan	1
Samsung New Zealand	1
Samsung Levant	1
Samsung Latinoamérica y Caribe	1
Samsung Saudi Arabia	1
Samsung Israel	1
Samsung Maroc	1
Samsung Argentina	1
Samsung Colombia	1
Samsung Perú	1
Samsung Ukraine	1
Samsung Kazakhstan	1
Samsung Ghana	1
Samsung Georgia	1
Samsung Azerbaijan	1
ForumMySamsung	1
Samsung Developers	1
Samsung Memory	1
CNET	1
samsungmena	1



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 κ.ο.κ.

### 6.3.2 Out-Degree

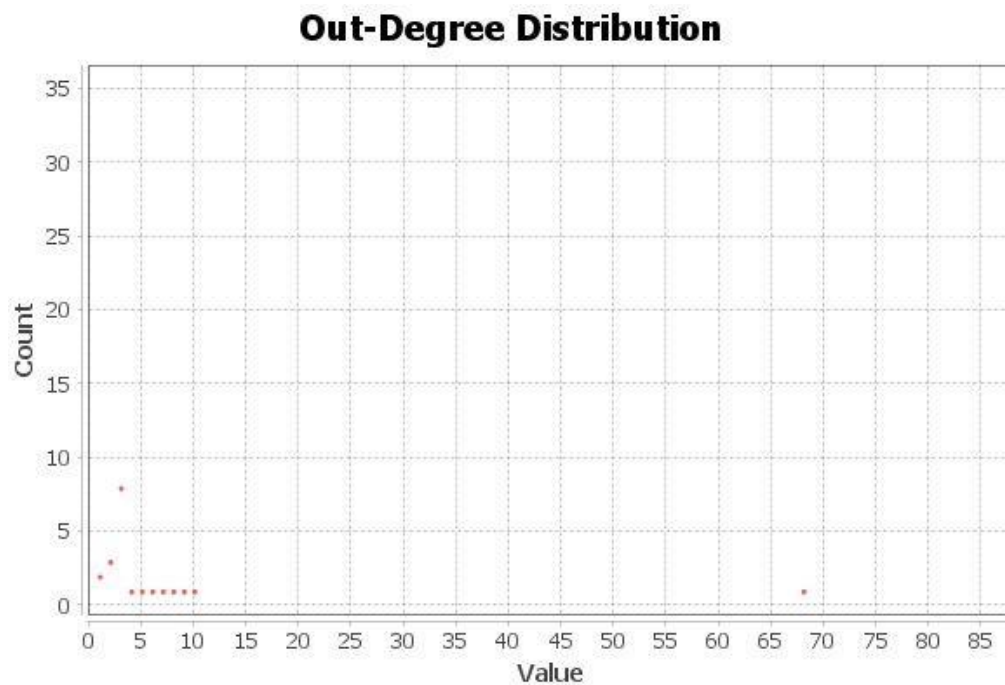
Το Out-Degree είναι οι εξερχόμενες από τον κάθε κόμβο ακμές. Με το δίκτυο το οποίο μελετάμε είναι ο αριθμός των Featured Channels που μπορεί να έχει ένα κανάλι όπως βλέπουμε παρακάτω.

Label	Out-Degree
Samsung	68
Samsung New Zealand	10
Samsung México	9
Samsung Polska	8
Samsung Maroc	7
Samsung India	6
Samsung France	5
Samsung Österreich	4
KYU WON LEE	3
Samsung Estonia	3
Samsung Latvia	3
Samsung Lithuania	3
Samsung Switzerland	3
Samsung Norge	3
Samsung Suomi	3
Samsung Sverige	3
Samsung Australia	2
Samsung Deutschland	2
Samsung Vietnam	2
Samsung Hrvatska	1
Samsung Latinoamérica y Caribe	1
Samsung US	0
Samsung Newsroom	0
Samsung UK	0
Samsung Canada	0
Samsung España	0
Samsung Korea	0
SamsungPrinting	0
SamsungChile	0
Samsung Egypt	0
Samsung Gulf	0
Samsung Hong Kong	0
Samsung Indonesia	0
Samsung Malaysia	0
Samsung Philippines	0
Samsung Russia	0
Samsung Singapore	0

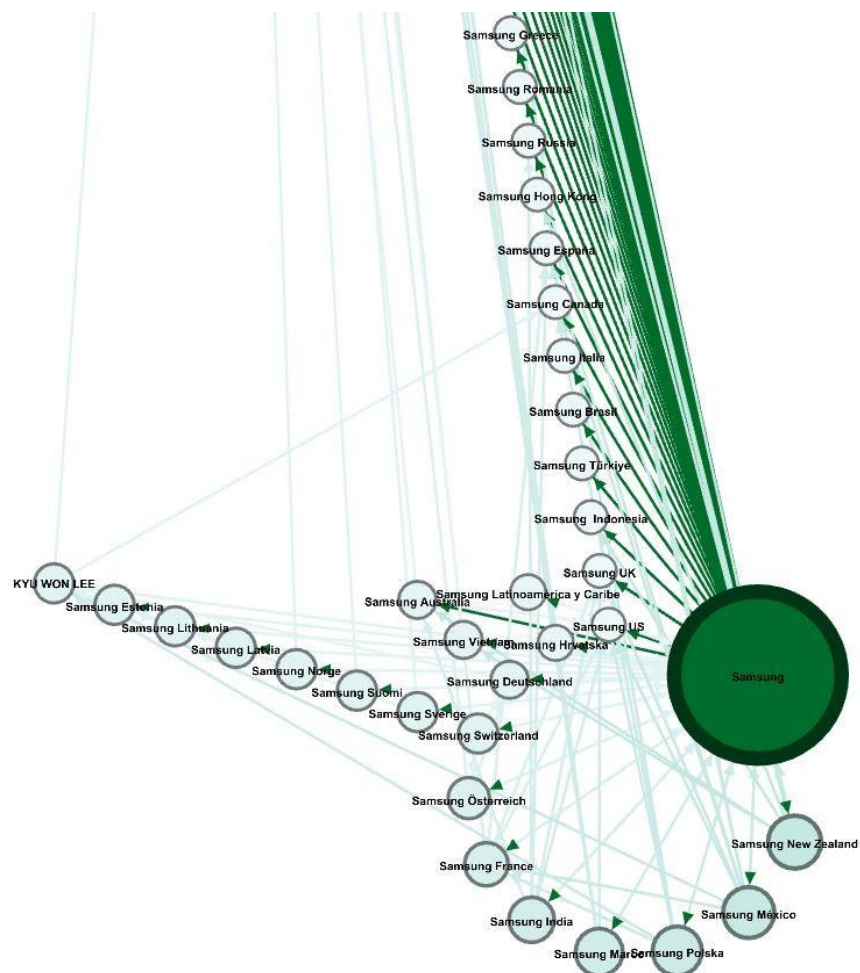
Samsung South Africa	0
SamsungTaiwan	0
Samsung Thailand	0
CNET	0
ForumMySamsung	0
Samsung Argentina	0
Samsung Azerbaijan	0
Samsung Belgium	0
Samsung Bosna i Hercegovina	0
Samsung Brasil	0
Samsung Bulgaria	0
Samsung Česko a Slovensko	0
Samsung Colombia	0
Samsung Crna Gora	0
Samsung Danmark	0
Samsung Developers	0
Samsung Georgia	0
Samsung Ghana	0
Samsung Greece	0
Samsung Ireland	0
Samsung Israel	0
Samsung Italia	0
Samsung Kazakhstan	0
Samsung Levant	0
Samsung Magyarország	0
Samsung Makedonija	0
Samsung Memory	0
samsungmena	0
Samsung Nederland	0
Samsung Pakistan	0
Samsung Perú	0
Samsung Portugal	0
Samsung Romania	0
Samsung Saudi Arabia	0
Samsung Shqiperi	0
Samsung Slovenija	0
Samsung Srbija	0
Samsung Türkiye	0
Samsung Ukraine	0



### 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.

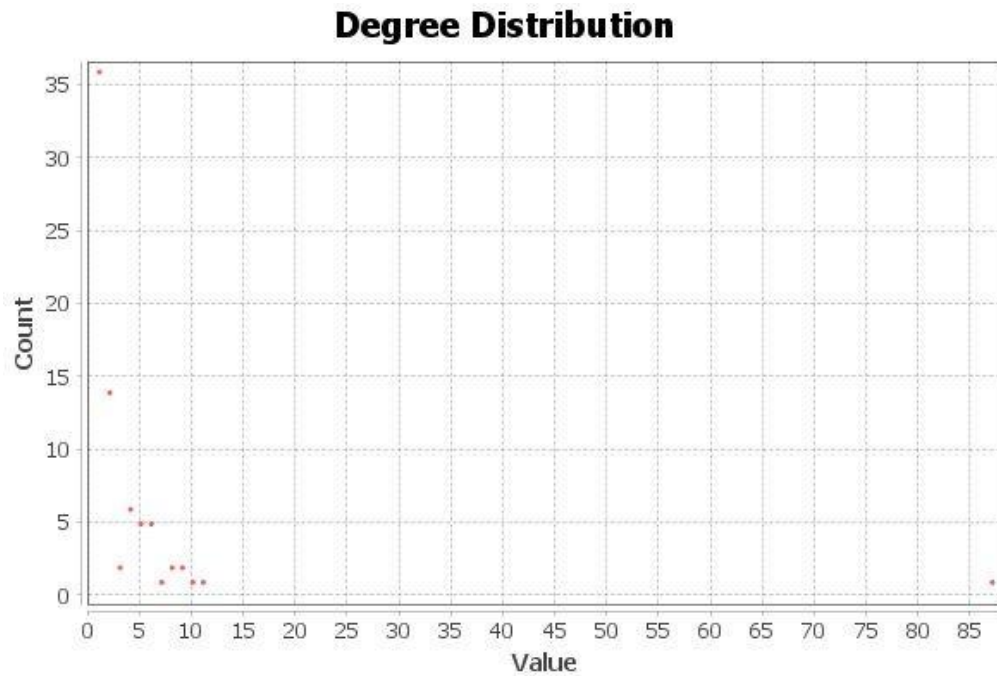
### 6.3.3 Total Degree

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

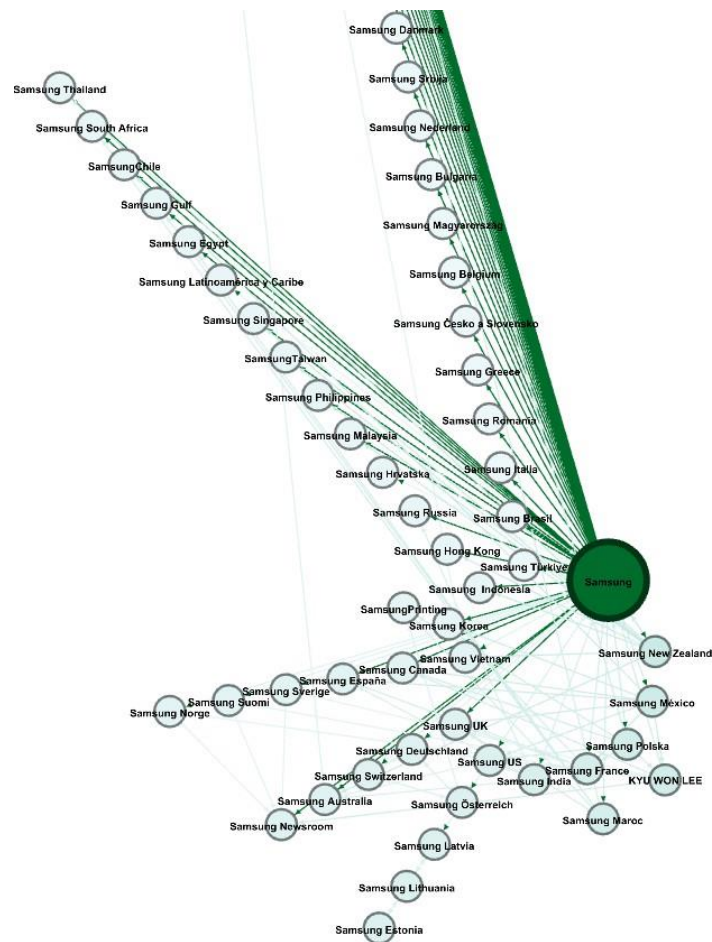
Label	Degree ▼
Samsung	87
Samsung New Zealand	11
Samsung México	10
Samsung Polska	9
KYU WON LEE	9
Samsung France	8
Samsung Maroc	8
Samsung India	7
Samsung US	6
Samsung Österreich	6
Samsung Latvia	6
Samsung Lithuania	6
Samsung Estonia	6
Samsung UK	5
Samsung Deutschland	5
Samsung Switzerland	5
Samsung Australia	5
Samsung Newsroom	5
Samsung Vietnam	4
Samsung Canada	4
Samsung España	4
Samsung Sverige	4
Samsung Suomi	4
Samsung Norge	4
Samsung Korea	3
SamsungPrinting	3
Samsung Indonesia	2
Samsung Hong Kong	2
Samsung Russia	2
Samsung Hrvatska	2
Samsung Malaysia	2
Samsung Philippines	2
SamsungTaiwan	2
Samsung Singapore	2
Samsung Latinoamérica y Caribe	2
Samsung Egypt	2
Samsung Gulf	2

SamsungChile	2
Samsung South Africa	2
Samsung Thailand	2
Samsung Türkiye	1
Samsung Brasil	1
Samsung Italia	1
Samsung Romania	1
Samsung Greece	1
Samsung Česko a Slovensko	1
Samsung Belgium	1
Samsung Magyarország	1
Samsung Bulgaria	1
Samsung Nederland	1
Samsung Srbija	1
Samsung Danmark	1
Samsung Slovenija	1
Samsung Ireland	1
Samsung Shqiperi	1
Samsung Makedonija	1
Samsung Bosna i Hercegovina	1
Samsung Crna Gora	1
Samsung Portugal	1
Samsung Pakistan	1
Samsung Levant	1
Samsung Saudi Arabia	1
Samsung Israel	1
Samsung Argentina	1
Samsung Colombia	1
Samsung Perú	1
Samsung Ukraine	1
Samsung Kazakhstan	1
Samsung Ghana	1
Samsung Georgia	1
Samsung Azerbaijan	1
ForumMySamsung	1
Samsung Developers	1
Samsung Memory	1
CNET	1
samsungmena	1

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.

## 7 Centrality measures

### 7.1 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  $n_i$  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  $n_i$ :  $\Sigma_{uw}(n_i)$
2. Με τον αριθμό των συντομότερων διαδρομών που περνούν από τον κόμβο  $x$ (τα μονοπάτια των  $u$  προς  $w$ ):  $\Sigma_{uw}$

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  $C_B(n_i) = \frac{\sum (\Sigma_{uw}(n_i) / \Sigma_{uw})}{\sum (\Sigma_{uw}(n_i) / \Sigma_{uw})}$ .

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



Label	Betweenness Centrality
Samsung	1334.5
Samsung Polska	45.166667
Samsung Maroc	24.333333
Samsung Australia	20.5
Samsung Vietnam	19.0
Samsung Österreich	6.333333
Samsung Deutschland	3.833333
KYU WON LEE	3.5
Samsung Sverige	1.833333
Samsung Suomi	1.833333
Samsung Norge	1.833333
Samsung México	1.833333
Samsung France	1.5
Samsung US	0.0
Samsung Ukraine	0.0
Samsung UK	0.0
Samsung Türkiye	0.0
Samsung Thailand	0.0
SamsungTaiwan	0.0
Samsung Switzerland	0.0
Samsung Srbija	0.0
Samsung South Africa	0.0
Samsung Slovenija	0.0
Samsung Singapore	0.0
Samsung Shqiperi	0.0
Samsung Saudi Arabia	0.0
Samsung Russia	0.0
Samsung Romania	0.0
SamsungPrinting	0.0
Samsung Portugal	0.0
Samsung Philippines	0.0
Samsung Perú	0.0
Samsung Pakistan	0.0
Samsung New Zealand	0.0
Samsung Newsroom	0.0
Samsung Nederland	0.0
samsungmena	0.0

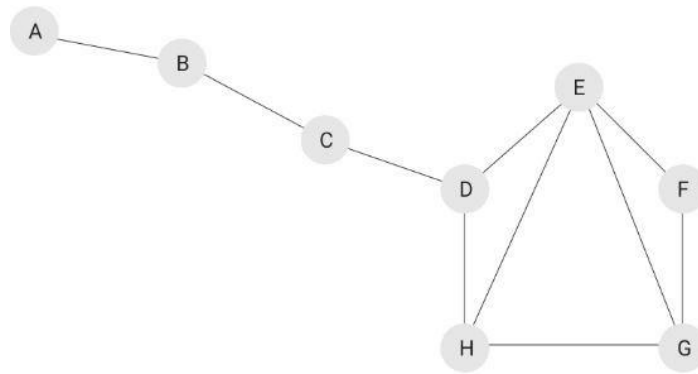
Samsung Memory	0.0
Samsung Malaysia	0.0
Samsung Makedonija	0.0
Samsung Magyarország	0.0
Samsung Lithuania	0.0
Samsung Levant	0.0
Samsung Latvia	0.0
Samsung Latinoamérica y Caribe	0.0
Samsung Korea	0.0
Samsung Kazakhstan	0.0
Samsung Italia	0.0
Samsung Israel	0.0
Samsung Ireland	0.0
Samsung Indonesia	0.0
Samsung India	0.0
Samsung Hrvatska	0.0
Samsung Hong Kong	0.0
Samsung Gulf	0.0
Samsung Greece	0.0
Samsung Ghana	0.0
Samsung Georgia	0.0
Samsung Estonia	0.0
Samsung España	0.0
Samsung Egypt	0.0
Samsung Developers	0.0
Samsung Danmark	0.0
Samsung Crna Gora	0.0
Samsung Colombia	0.0
SamsungChile	0.0
Samsung Česko a Slovensko	0.0
Samsung Canada	0.0
Samsung Bulgaria	0.0
Samsung Brasil	0.0
Samsung Bosna i Hercegovina	0.0
Samsung Belgium	0.0
Samsung Azerbaijan	0.0
Samsung Argentina	0.0
ForumMySamsung	0.0
CNET	0.0

Όπως φαίνεται και από τους πιο πάνω πίνακες λοιπόν, είναι λίγες οι χώρες που έχουν μη μηδενικό Betweenness 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.

## 7.2 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 άθροισμα των συντομότερων μονοπατιών από τον κόμβο που εξετάζουμε προς όλους τους υπόλοιπους. Επομένως για τον κόμβο C έχουμε:

Αρχικός κόμβος	Τελικός Κόμβος	Συντομότερη Διαδρομή	Κόστος Διαδρομής
C	A	C→B→A	2
C	B	C→B	1
C	D	C→D	1
C	E	C→D→E	2
C	F	C→D→E→F	3
C	G	C→D→H→G	3
C	H	C→D→H	2
Σύνολο:			14

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

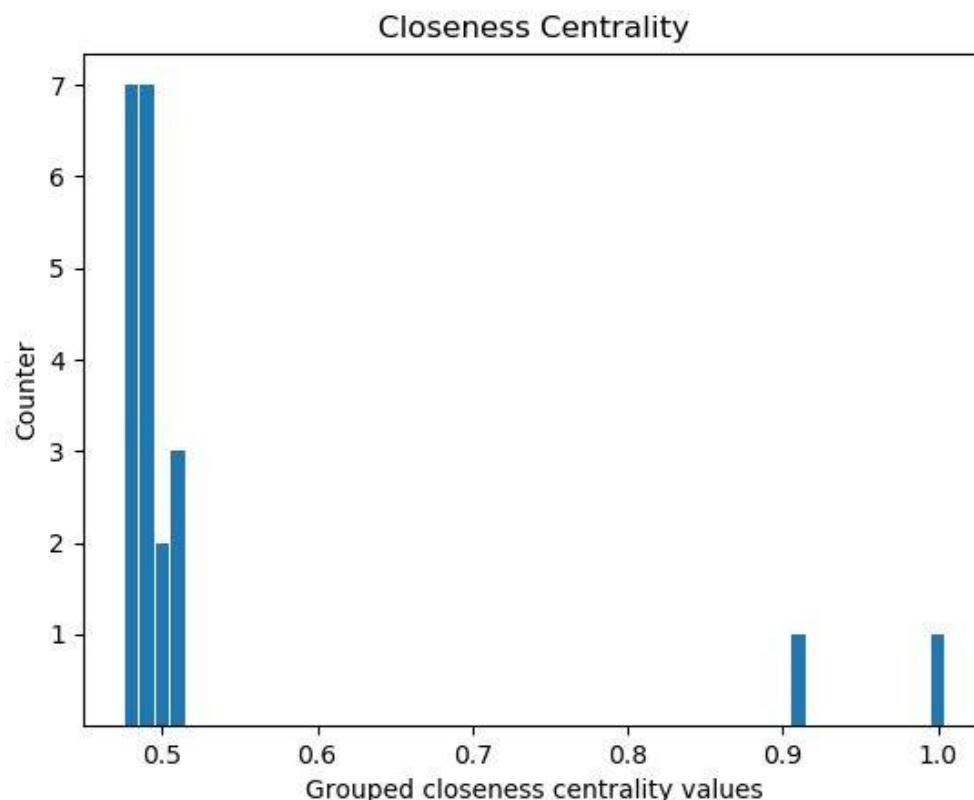
Να πούμε επίσης πως σε κατευθυνόμενους γράφους, παίζει ρόλο η φορά των ακμών. Για παράδειγμα σε περιπτώσεις όπου δεν υπάρχει διαδρομή λόγω φοράς ακμών, θεωρούμε μηδενικό το μονοπάτι. Έτσι γενικεύοντας το πιο πάνω υπάρχει περίπτωση κάποιος κόμβος να έχει μηδενικό Betweenness Centrality.

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

Label	Closeness Centrality ▾
Samsung Australia	1.0
Samsung	0.914634
Samsung Polska	0.517241
Samsung New Zealand	0.517241
Samsung México	0.510204
Samsung Maroc	0.506757
Samsung India	0.5
Samsung Österreich	0.496689
Samsung France	0.493421
Samsung Switzerland	0.493421
KYU WON LEE	0.490196
Samsung Sverige	0.490196
Samsung Suomi	0.490196
Samsung Norge	0.490196
Samsung Vietnam	0.487013
Samsung Deutschland	0.487013
Samsung Lithuania	0.487013
Samsung Latvia	0.487013
Samsung Estonia	0.487013
Samsung Latinoamérica y Caribe	0.480769
Samsung Hrvatska	0.480769
Samsung US	0.0
Samsung Ukraine	0.0
Samsung UK	0.0
Samsung Türkiye	0.0
Samsung Thailand	0.0
SamsungTaiwan	0.0
Samsung Srbija	0.0
Samsung South Africa	0.0
Samsung Slovenija	0.0
Samsung Singapore	0.0
Samsung Shqiperi	0.0
Samsung Saudi Arabia	0.0
Samsung Russia	0.0
Samsung Romania	0.0
SamsungPrinting	0.0
Samsung Portugal	0.0

Samsung Philippines	0.0
Samsung Perú	0.0
Samsung Pakistan	0.0
Samsung Newsroom	0.0
Samsung Nederland	0.0
samsungmena	0.0
Samsung Memory	0.0
Samsung Malaysia	0.0
Samsung Makedonija	0.0
Samsung Magyarország	0.0
Samsung Levant	0.0
Samsung Korea	0.0
Samsung Kazakhstan	0.0
Samsung Italia	0.0
Samsung Israel	0.0
Samsung Ireland	0.0
Samsung Indonesia	0.0
Samsung Hong Kong	0.0
Samsung Gulf	0.0
Samsung Greece	0.0
Samsung Ghana	0.0
Samsung Georgia	0.0
Samsung España	0.0
Samsung Egypt	0.0
Samsung Developers	0.0
Samsung Danmark	0.0
Samsung Crna Gora	0.0
Samsung Colombia	0.0
SamsungChile	0.0
Samsung Česko a Slovensko	0.0
Samsung Canada	0.0
Samsung Bulgaria	0.0
Samsung Brasil	0.0
Samsung Bosna i Hercegovina	0.0
Samsung Belgium	0.0
Samsung Azerbaijan	0.0
Samsung Argentina	0.0
ForumMySamsung	0.0

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.

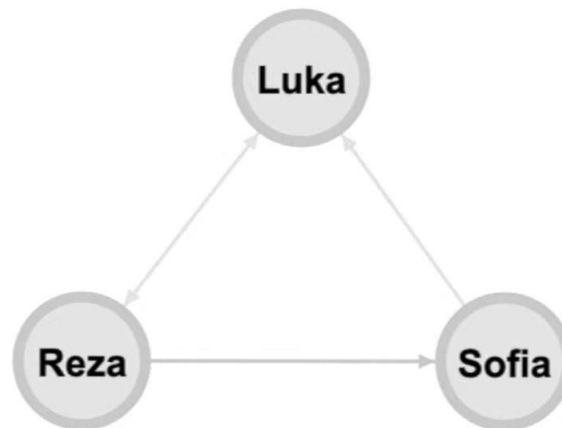


Βλέπουμε λοιπόν πως υπάρχει μια ομαδοποίηση των κόμβων μεταξύ των τιμών 0.48 και 0.51. Αντίθετα όμως, μόνο δύο κόμβοι έχουν υψηλές τιμές της τάξης των 0.91 και 1.0. Άρα αμέσως καταλαβαίνουμε ότι τα κανάλια που έχουν πιο σημαντικό ρόλο στο δίκτυο μας είναι αυτά με 0.91 και 1.0 με το όνομα αυτών να είναι 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.

### 7.3 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. Δείχνει 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.

		From		
		Reza	Sofia	Luka
To	Reza	0	0	1
	Sofia	1	0	0
	Luka	1	1	0



Στη συνέχεια χρησιμοποιώντας αλγόριθμους γραμμικής άλγεβρας βρίσκουμε το ιδιοδιάνυσμα που αντιστοιχεί στη μεγαλύτερη ιδιότημή 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.

Label	Eigenvector Centrality		
Samsung	1.0	Samsung Greece	0.207843
Samsung US	0.461073	Samsung Česko a Slovensko	0.207843
Samsung UK	0.399569	Samsung Belgium	0.207843
Samsung Newsroom	0.398726	Samsung Magyarország	0.207843
Samsung España	0.356162	Samsung Suomi	0.207843
Samsung Canada	0.355318	Samsung Bulgaria	0.207843
Samsung Latvia	0.352047	Samsung Nederland	0.207843
Samsung Lithuania	0.352047	Samsung Srbija	0.207843
Samsung Estonia	0.352047	Samsung Norge	0.207843
Samsung Deutschland	0.317048	Samsung Danmark	0.207843
Samsung Korea	0.312754	Samsung Hrvatska	0.207843
Samsung France	0.294658	Samsung Slovenija	0.207843
Samsung Australia	0.294658	Samsung Ireland	0.207843
KYU WON LEE	0.282834	Samsung Shqiperi	0.207843
Samsung Switzerland	0.262445	Samsung Makedonija	0.207843
Samsung Österreich	0.262445	Samsung Bosna i Hercegovina	0.207843
Samsung Vietnam	0.251251	Samsung Crna Gora	0.207843
Samsung Indonesia	0.251251	Samsung Portugal	0.207843
Samsung Hong Kong	0.251251	Samsung Pakistan	0.207843
Samsung Russia	0.251251	Samsung New Zealand	0.207843
Samsung Malaysia	0.251251	Samsung Levant	0.207843
Samsung Philippines	0.251251	Samsung Latinoamérica y Caribe	0.207843
SamsungTaiwan	0.251251	Samsung Saudi Arabia	0.207843
Samsung Singapore	0.251251	Samsung Israel	0.207843
Samsung Egypt	0.251251	Samsung Maroc	0.207843
Samsung Gulf	0.251251	Samsung Argentina	0.207843
SamsungChile	0.251251	Samsung Colombia	0.207843
Samsung South Africa	0.251251	Samsung Perú	0.207843
Samsung Thailand	0.251251	Samsung Ukraine	0.207843
Samsung India	0.207843	Samsung Kazakhstan	0.207843
Samsung Türkiye	0.207843	Samsung Ghana	0.207843
Samsung Brasil	0.207843	Samsung Georgia	0.207843
Samsung México	0.207843	Samsung Azerbaijan	0.207843
Samsung Italia	0.207843	SamsungPrinting	0.141417
Samsung Polska	0.207843	CNET	0.061504
Samsung Romania	0.207843	ForumMySamsung	0.052455
Samsung Sverige	0.207843	Samsung Developers	0.043407
		Samsung Memory	0.043407
		samsungmena	0.043407



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

Πέραν από το πιο πάνω, το πλέον σημαντικό που μπορούμε να πούμε για τα αποτελέσματα αυτά είναι πως υπάρχει μια μεγάλη αρθρική 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.

## 8 Clustering Effects

### 8.1 Average Clustering Coefficient

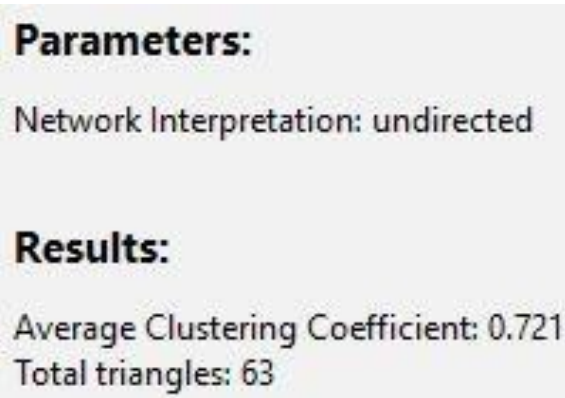
Το Average Clustering Coefficient μας λέει την πιθανότητα με την οποία 2 γειτονικοί κόμβοι τυχαία επιλεγμένοι ενός τυχαίου κόμβου να είναι συνδεδεμένοι. Ο τρόπος με τον οποίο υπολογίζουμε τη μετρική αυτή γίνεται ως εξής. Επιλέγουμε έναν τυχαίο κόμβο όπου στη συνέχεια επιλέγουμε 2 γείτονες του τυχαία 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.

#### **Results:**

Average Clustering Coefficient: 0.311

## 8.2 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, μέσω 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.

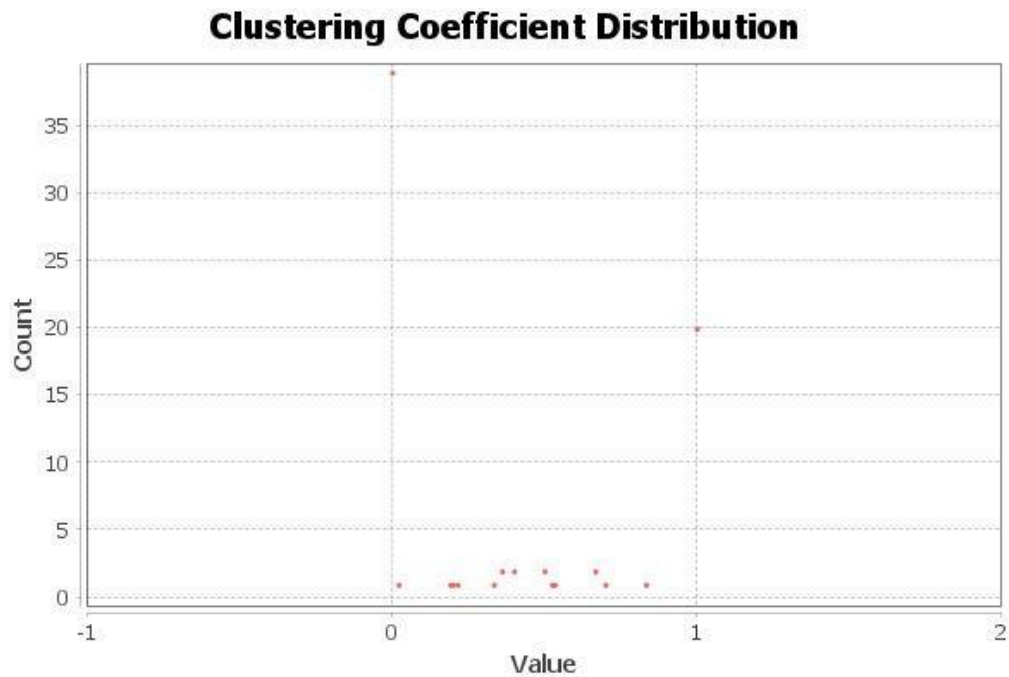
Πιο κάτω φαίνονται αναλυτικότερα οι τιμές των Triangles ανά κανάλι.

Label	Number of triangles ▼
Samsung	50
Samsung México	13
KYU WON LEE	13
Samsung France	11
Samsung New Zealand	9
Samsung US	8
Samsung Newsroom	7
Samsung India	6
Samsung Polska	6
Samsung UK	5
Samsung Canada	5
Samsung Deutschland	4
Samsung España	4
Samsung Australia	4
Samsung Maroc	4
Samsung Switzerland	3
Samsung Österreich	3
Samsung Sverige	3
Samsung Suomi	3
Samsung Norge	3
Samsung Latvia	3
Samsung Lithuania	3
Samsung Estonia	3
Samsung Korea	3
Samsung Vietnam	1
Samsung Indonesia	1
Samsung Hong Kong	1
Samsung Russia	1
Samsung Malaysia	1
Samsung Philippines	1
SamsungTaiwan	1
Samsung Singapore	1
Samsung Egypt	1
Samsung Gulf	1
SamsungChile	1
Samsung South Africa	1
Samsung Thailand	1
Samsung Türkiye	0

Samsung Brasil	0
Samsung Italia	0
Samsung Romania	0
Samsung Greece	0
Samsung Česko a Slovensko	0
Samsung Belgium	0
Samsung Magyarország	0
Samsung Bulgaria	0
Samsung Nederland	0
Samsung Srbija	0
Samsung Danmark	0
Samsung Hrvatska	0
Samsung Slovenija	0
Samsung Ireland	0
Samsung Shqiperi	0
Samsung Makedonija	0
Samsung Bosna i Hercegovina	0
Samsung Crna Gora	0
Samsung Portugal	0
Samsung Pakistan	0
Samsung Levant	0
Samsung Latinoamérica y Caribe	0
Samsung Saudi Arabia	0
Samsung Israel	0
Samsung Argentina	0
Samsung Colombia	0
Samsung Perú	0
Samsung Ukraine	0
Samsung Kazakhstan	0
Samsung Ghana	0
Samsung Georgia	0
Samsung Azerbaijan	0
ForumMySamsung	0
Samsung Developers	0
SamsungPrinting	0
Samsung Memory	0
CNET	0
samsungmena	0

### 8.3 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.

Label	Clustering Coefficient <span>▼</span>
Samsung Indonesia	1.0
Samsung Hong Kong	1.0
Samsung Russia	1.0
Samsung Switzerland	1.0
Samsung Sverige	1.0
Samsung Suomi	1.0
Samsung Norge	1.0
Samsung Latvia	1.0
Samsung Lithuania	1.0
Samsung Estonia	1.0
Samsung Malaysia	1.0
Samsung Philippines	1.0
SamsungTaiwan	1.0
Samsung Singapore	1.0
Samsung Egypt	1.0
Samsung Gulf	1.0
SamsungChile	1.0
Samsung South Africa	1.0
Samsung Korea	1.0
Samsung Thailand	1.0
Samsung Canada	0.833333
Samsung Newsroom	0.7
Samsung Deutschland	0.666667
Samsung España	0.666667
Samsung US	0.533333
Samsung France	0.52381
Samsung UK	0.5
Samsung Österreich	0.5
Samsung India	0.4
Samsung Australia	0.4
Samsung México	0.361111
KYU WON LEE	0.361111
Samsung Vietnam	0.333333
Samsung Polska	0.214286
Samsung New Zealand	0.2
Samsung Maroc	0.190476
Samsung	0.021313

Samsung Türkiye	0.0
Samsung Brasil	0.0
Samsung Italia	0.0
Samsung Romania	0.0
Samsung Greece	0.0
Samsung Česko a Slovensko	0.0
Samsung Belgium	0.0
Samsung Magyarország	0.0
Samsung Bulgaria	0.0
Samsung Nederland	0.0
Samsung Srbija	0.0
Samsung Danmark	0.0
Samsung Hrvatska	0.0
Samsung Slovenija	0.0
Samsung Ireland	0.0
Samsung Shqiperi	0.0
Samsung Makedonija	0.0
Samsung Bosna i Hercegovina	0.0
Samsung Crna Gora	0.0
Samsung Portugal	0.0
Samsung Pakistan	0.0
Samsung Levant	0.0
Samsung Latinoamérica y Caribe	0.0
Samsung Saudi Arabia	0.0
Samsung Israel	0.0
Samsung Argentina	0.0
Samsung Colombia	0.0
Samsung Perú	0.0
Samsung Ukraine	0.0
Samsung Kazakhstan	0.0
Samsung Ghana	0.0
Samsung Georgia	0.0
Samsung Azerbaijan	0.0
ForumMySamsung	0.0
Samsung Developers	0.0
SamsungPrinting	0.0
Samsung Memory	0.0
CNET	0.0
samsungmena	0.0

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:



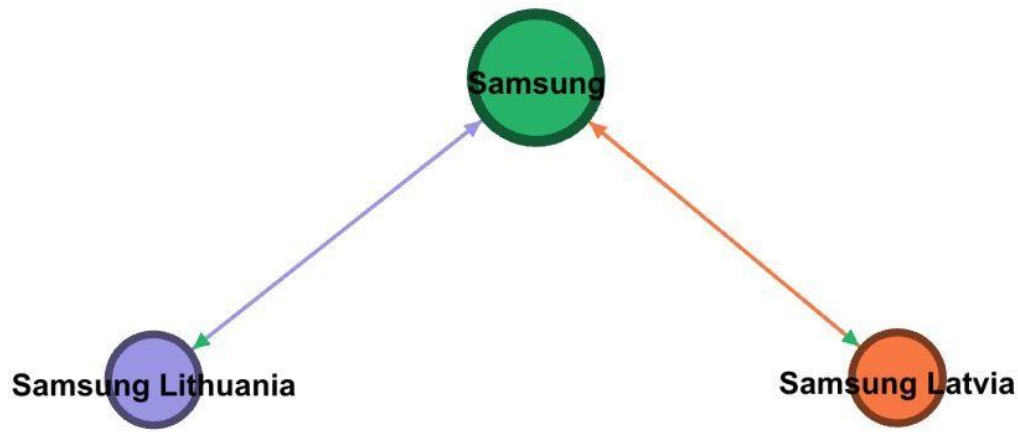
Κανάλι με υψηλό Clustering Coefficient:



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.

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

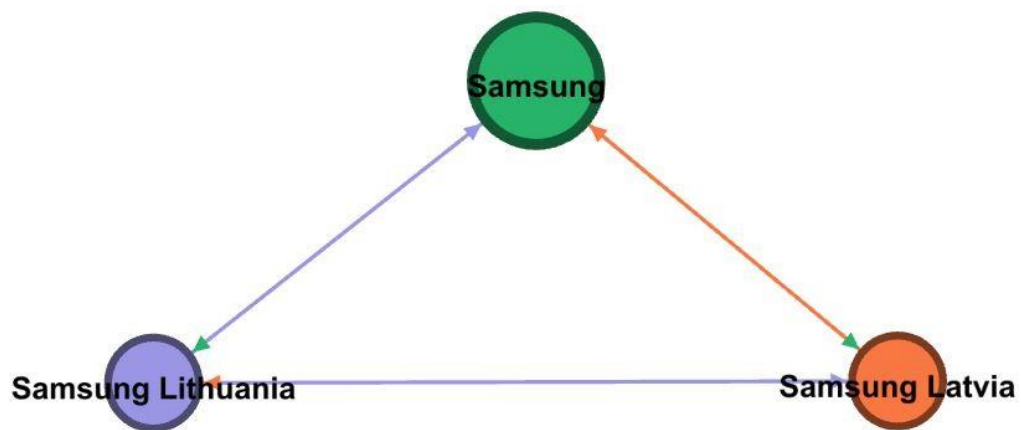
Η ύπαρξη τριαδικής κλειστότητας σε Friendship Neighborhood είναι ένα φαινόμενο το οποίο παρακολουθείται σε βάθος χρόνου. Για παράδειγμα στη πραγματική ζωή, όταν δύο άτομα(που δεν γνωρίζονται αρχικά) κάνουν παρέα με ένα κοινό άτομο, είναι πολύ πιθανό πως μελλοντικά, το κοινό αυτό άτομο που γνωρίζουν θα 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.

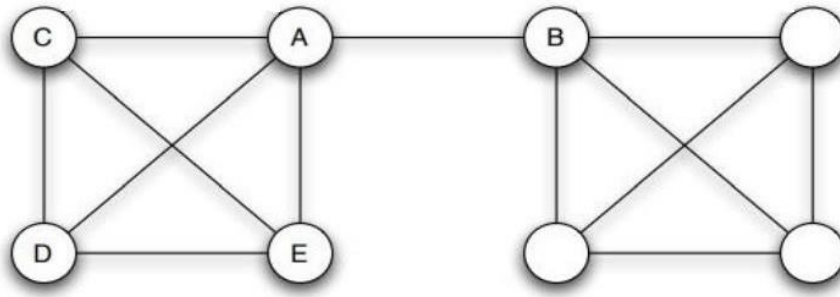
Άρα μπορεί να γίνει η σύνδεση με τα featured channels που περιγράψαμε όπως φαίνεται στην πιο κάτω εικόνα.



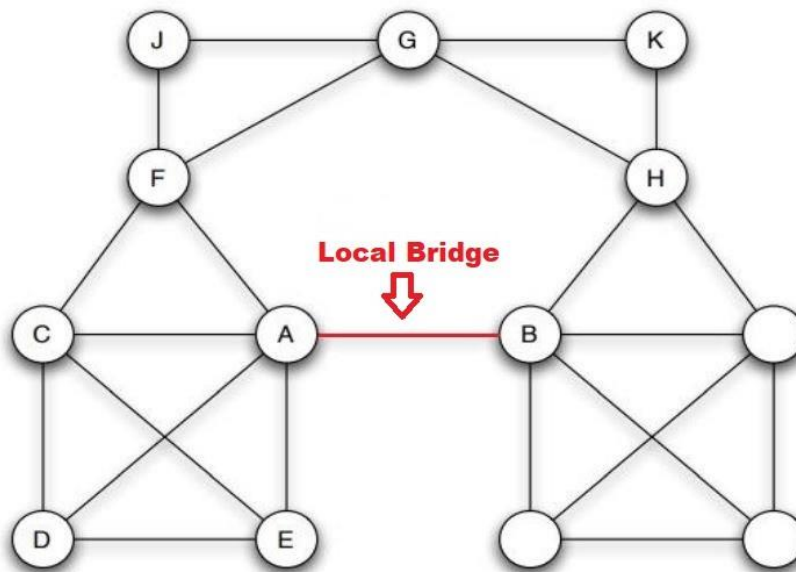
Το αποτέλεσμα μπορεί να είναι όπως φαίνεται στην τελευταία εικόνα χωρίς να είναι βέβαια απαραίτητη η αμφίδρομη σύνδεση.

## 9 Bridges and Local Bridges

Γέφυρα ή αλλιώς Bridge στην αγγλική ορολογία θεωρούμε την μοναδική 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.

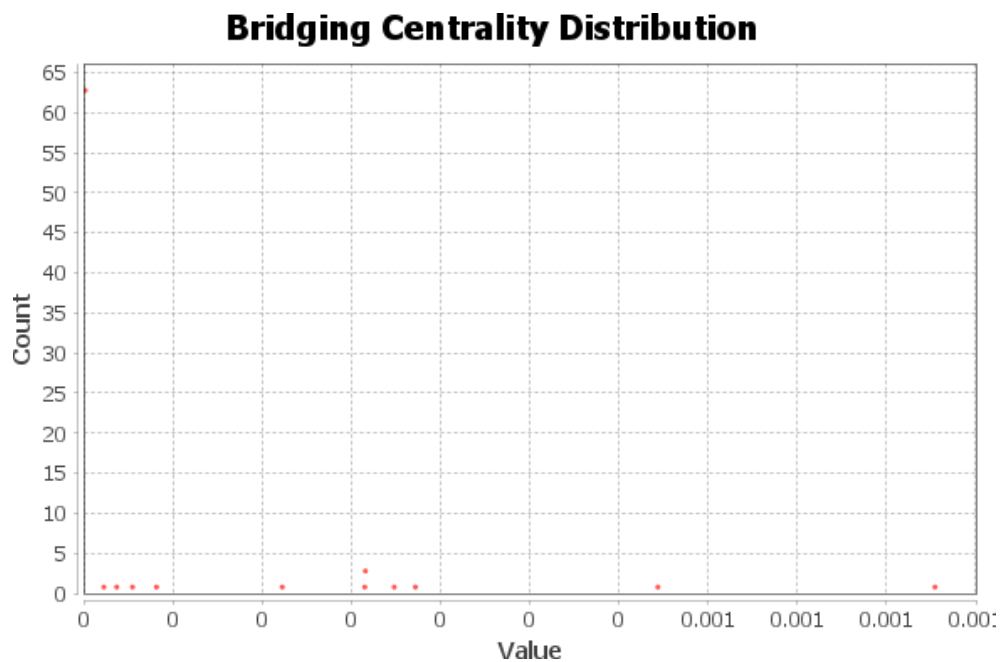


Έστω πως έχουμε παρόμοιο δίκτυο με το προηγούμενο, μόνο που η σύνδεση των δύο κόμβων που αποτελεί την ένωση των δύο κοινοτήτων δεν είναι η μοναδική ως εξής:



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 κοντά 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.

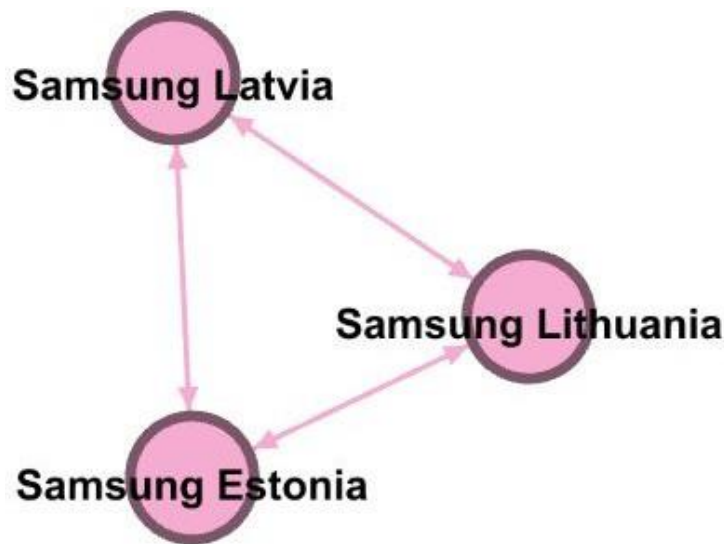




## 10 Gender and Homophily

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 άτομα του κύκλου μας τον οποίο εμείς επιλέγουμε με βάση ομοιότητες που συναντούμε στα άτομα αυτά.

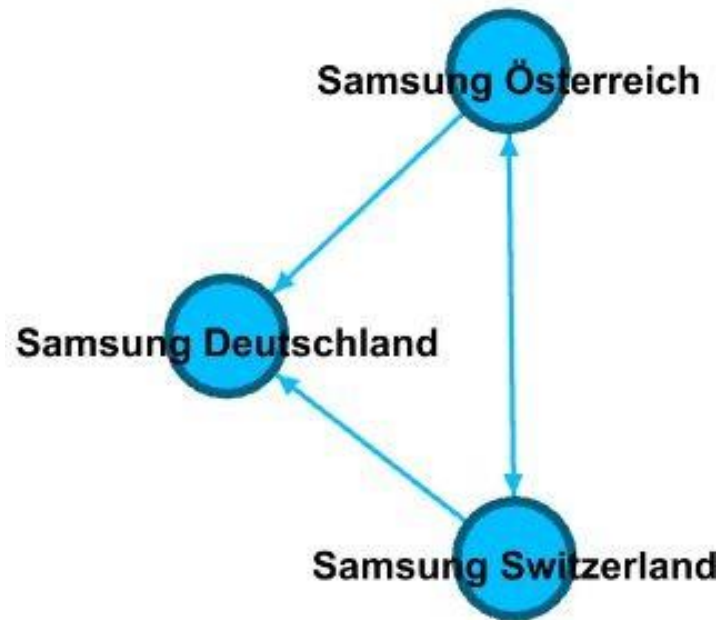
Στο σημείο αυτό λοιπόν θα δούμε πως το φαινόμενο αυτό παρατηρείται στο δικό μας δίκτυο. Θα εξετάσουμε περιπτώσεις καναλιών όπου η συναναστροφή τους οφείλεται σε ομοιότητες που υπάρχουν στα κανάλια αυτά. Τέτοιες περιπτώσεις συναντήσαμε στην ενότητα Community Structure όπου βρήκαμε σχηματισμένα communities στο δίκτυο μας. Για παράδειγμα, ας θυμηθούμε την 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.

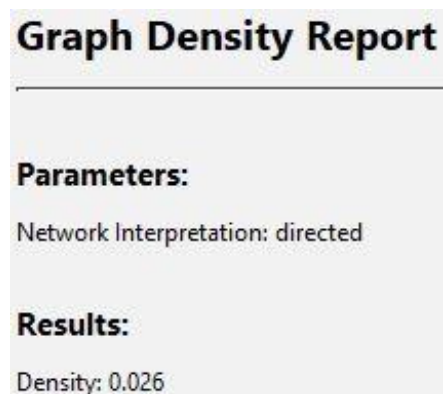
Άλλο ένα παράδειγμα τέτοιου είδους, είναι με τα κανάλια Samsung Switzerland, Samsung Deutschland(Γερμανία) και Samsung Österreich(Αυστρία).



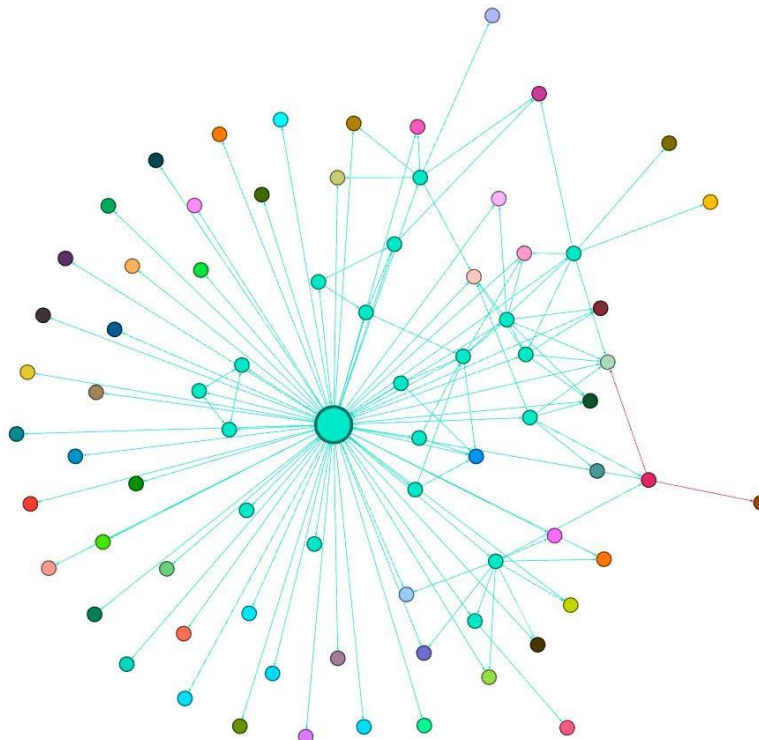
Όπως και πριν, παρατηρούμε κανάλια από χώρες της Ευρώπης τα οποία έχουν ισχυρή σύνδεση. Οι χώρες από τις οποίες προέρχονται είναι γειτονικές ενώ αυτή τη φορά έχουν ακόμα ένα κοινό, τη γλώσσα. Λογικό, αφού αυτό βοηθάει τα προτεινόμενα κανάλια να ελκύουν έτσι περισσότερο κοινό.

## 11 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  $76 \times 75 = 5700$ . Επομένως το Graph Density για το δίκτυο μας είναι  $149 / 5700 = 0.026$ . Πράγματι αυτό επιβεβαιώνεται μέσω του Gephi.

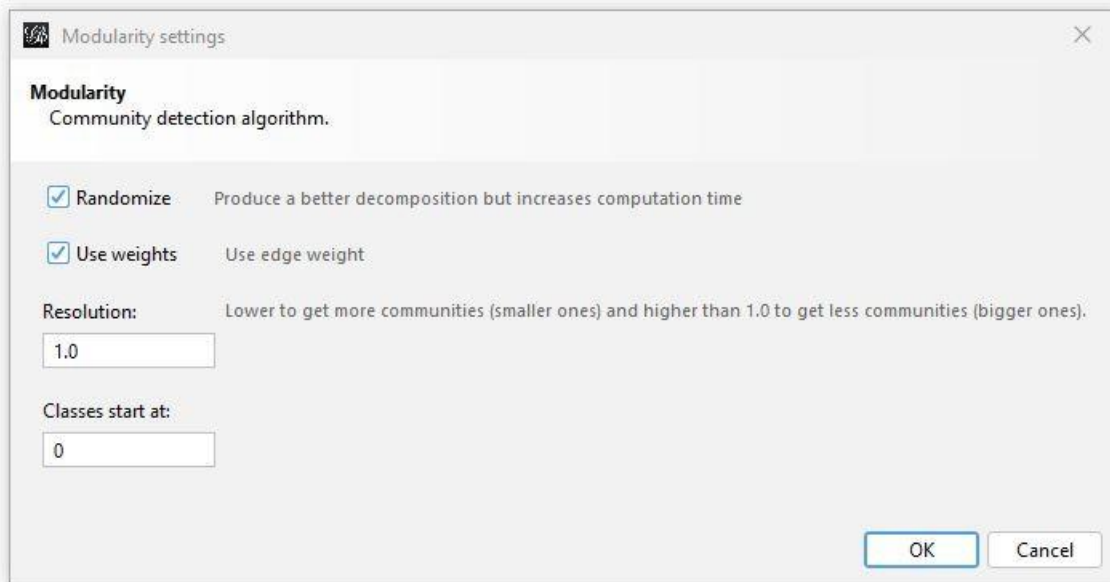


Το αποτέλεσμα που βρήκαμε είναι αρκετά μικρός αριθμός. Λογικό, αφού αν ανατρέξουμε στο δίκτυο μας θα δούμε πως 36 στους 76 κόμβους είναι ενωμένοι μόνο με έναν άλλο, ποσοστό που αγγίζει το 47.3% των κόμβων του δικτύου μας. Αυτό φαίνεται μέσω του Total Degree που είναι ίσο με 1 για τους κόμβους αυτούς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



## 12 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.

## 12.1 Community Structure - Modularity Resolution 0.1

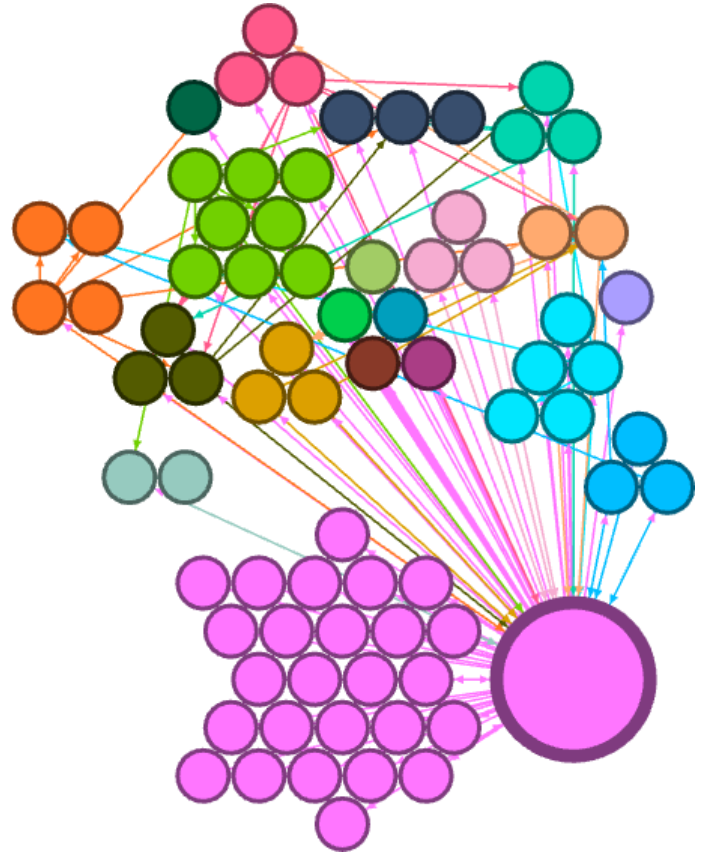
Με Resolution 0.1 έχουμε Modularity 0.283 ενώ δημιουργούνται 20 communities διαφόρων μεγεθών καθώς τα μικρότερα αποτελούνται από 1 μόνο κόμβο(7 communities) ενώ το μεγαλύτερο από 26. Όπως βλέπουμε υπάρχει μεγάλη διαφορά στο πλήθος των communities καθώς επίσης και μεγάλη κατανομή.

### Parameters:

Randomize: On  
Use edge weights: On  
Resolution: 0.1

### Results:

Modularity: 0.283  
Modularity with resolution: -0.128  
Number of Communities: 20

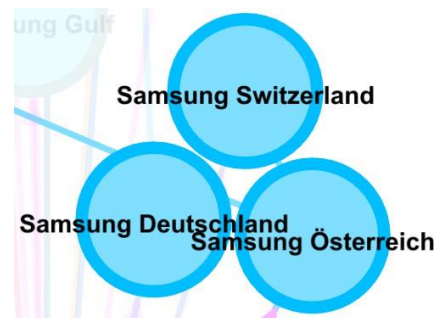


Να πούμε επίσης πως σε μικρές ομαδοποιήσεις τα κανάλια από τα οποία αποτελούνται είναι από χώρες που βρίσκονται γεωγραφικά κοντά ή στην ίδια ήπειρο όπως για παράδειγμα:

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



Το community με μπλε χρώμα αποτελείται από τα κανάλια Samsung Switzerland, Samsung Deutschland(Γερμανία), Samsung Österreich(Αυστρία).



Το community με πράσινο χρώμα αποτελείται από τα κανάλια Samsung New Zealand, Samsung Singapore, Samsung Malaysia, Samsung Hong Kong, Samsung Taiwan, Samsung Philippines, Samsung 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, samsungmena, Samsung Gulf. Σε αυτό το community βλέπουμε κανάλια από χώρες της Βορείου Αφρικής(Ηπείρου) και της Μέσης Ανατολής.



Το community με έντονο ροζ χρώμα αποτελείται από τα κανάλια Samsung South Canada, Samsung Chile, Samsung Mexico. 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.



## 12.2 Community Structure - Modularity Resolution 0.5

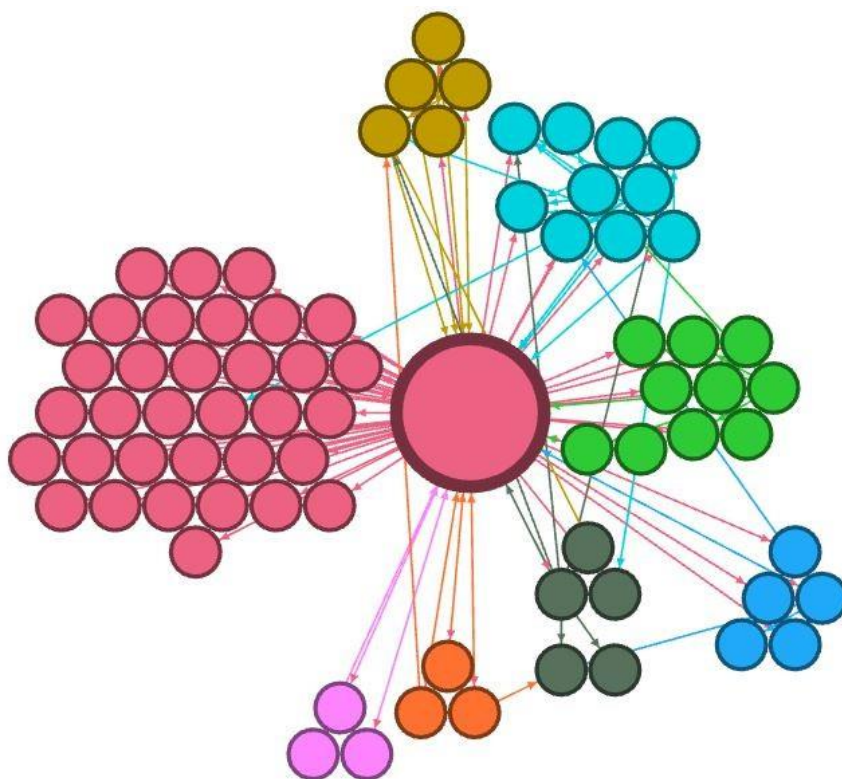
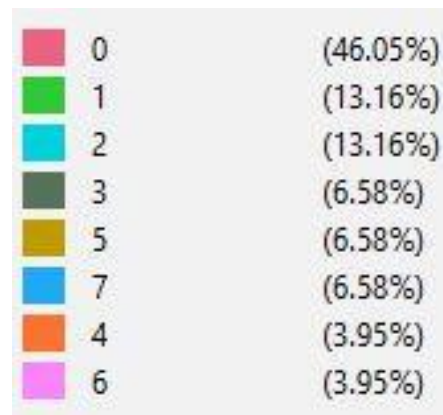
Με Resolution 0.5 έχουμε Modularity 0.396 ενώ δημιουργούνται 8 communities μεγέθους 3, 3, 5, 5, 5, 10, 10 και 35 κόμβων. Ακολουθούν τα ποσοστά σε σχέση με το μέγεθος του δικτύου καθώς επίσης οι μετρήσεις και ο τρόπος διάταξης του συγκεκριμένου community structure.

### Parameters:

Randomize: On  
Use edge weights: On  
Resolution: 0.5

### Results:

Modularity: 0.396  
Modularity with resolution: 0.081  
Number of Communities: 8



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, Samsung 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 Βιετνάμ αλλά και για το κανάλι ForumMySamsung αφού όπως είδαμε το περιεχόμενό του είναι στα Βιετναμέζικα. Δίπλα φαίνεται η καινούργια μορφή του συγκεκριμένου community.



## 12.3 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.

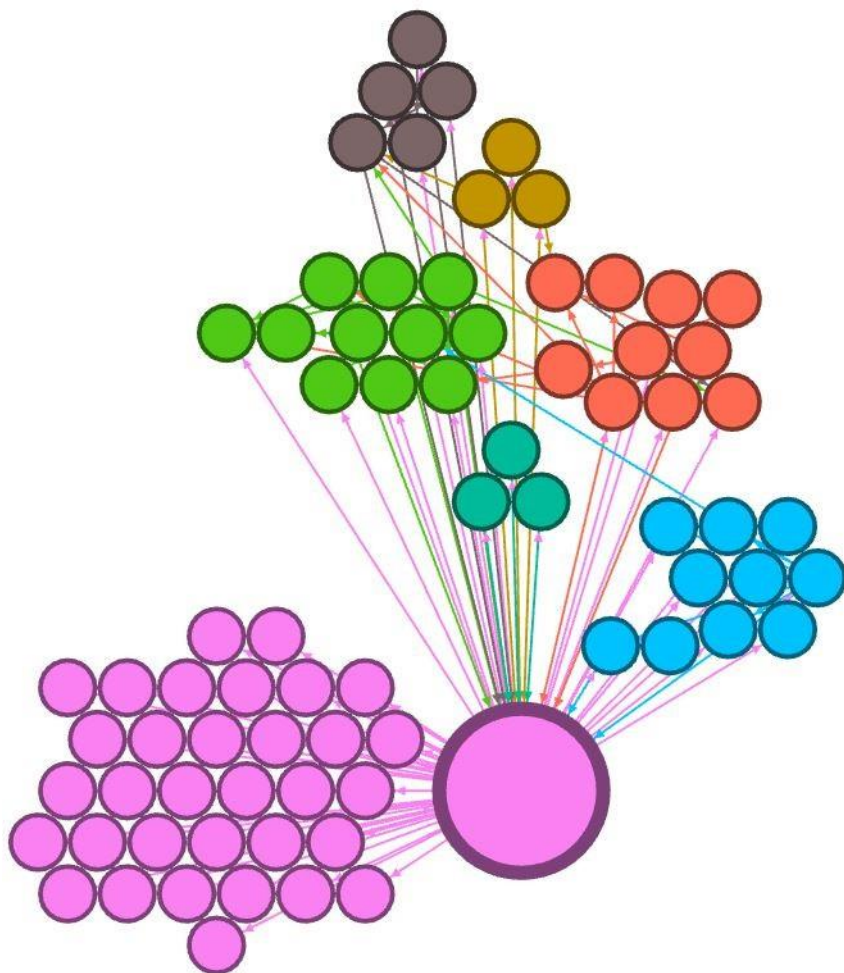
### Parameters:

Randomize: On  
Use edge weights: On  
Resolution: 0.96

### Results:

Modularity: 0.399  
Modularity with resolution: 0.374  
Number of Communities: 7

4	(44.74%)
0	(14.47%)
1	(13.16%)
6	(13.16%)
5	(6.58%)
2	(3.95%)
3	(3.95%)



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.

## 13 PageRank

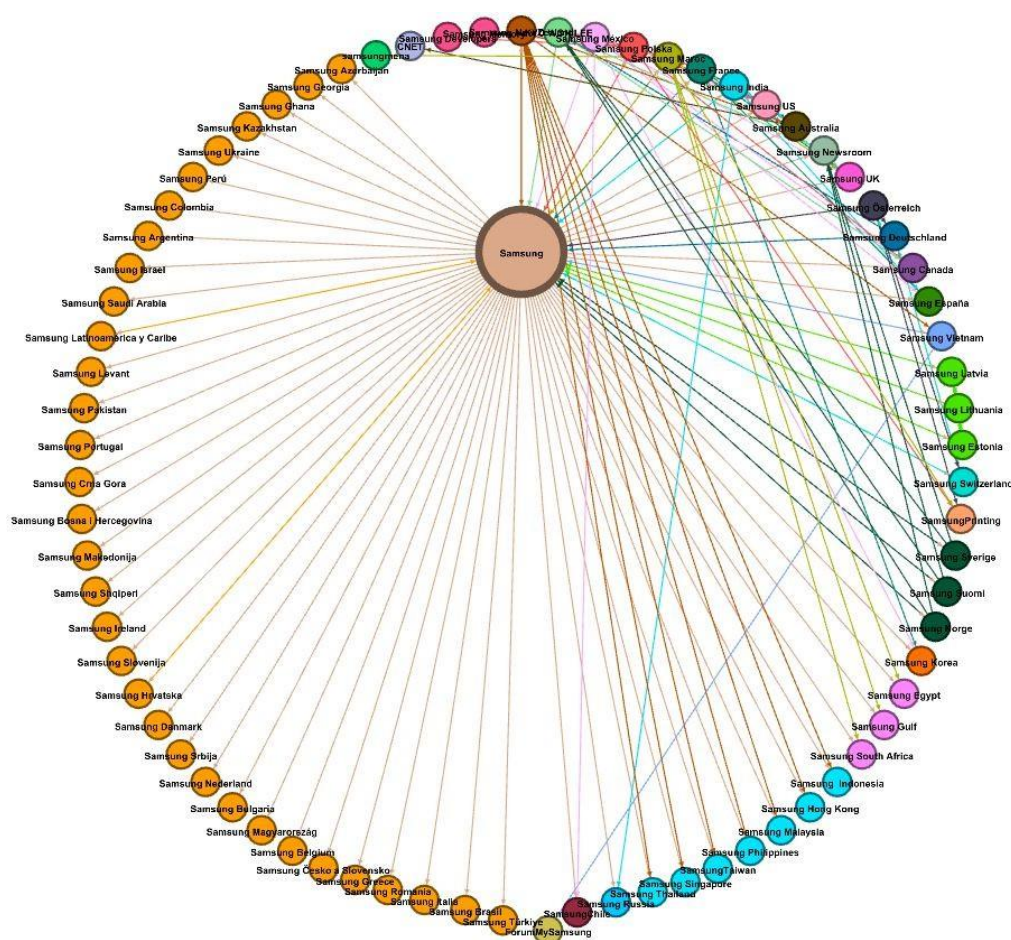
Το PageRank αποτελεί έναν αλγόριθμο ο οποίος αρχικά αναπτύχθηκε από την Google για την κατάταξη των ιστοσελίδων με σκοπό την καλύτερη και πιο αποτελεσματική εμφάνιση τους στις αναζητήσεις των χρηστών. Ο αλγόριθμος αυτός εφαρμόζεται και στην ανάλυση δικτύων αφού χρησιμοποιείται για την μέτρηση της σημασίας 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 μέσω του Gephi.

Label	PageRank
Samsung	0.265049
Samsung New Zealand	0.037077
KYU WON LEE	0.029077
Samsung México	0.028505
Samsung Polska	0.028008
Samsung Maroc	0.026474
Samsung France	0.022432
Samsung India	0.020319
Samsung US	0.019637
Samsung Australia	0.018392
Samsung Newsroom	0.017269
Samsung UK	0.016752
Samsung Österreich	0.014724
Samsung Deutschland	0.014351
Samsung Canada	0.013655
Samsung España	0.013537
Samsung Vietnam	0.01327
Samsung Latvia	0.012092
Samsung Lithuania	0.012092
Samsung Estonia	0.012092
Samsung Switzerland	0.011419
SamsungPrinting	0.011294
Samsung Sverige	0.010923
Samsung Suomi	0.010923
Samsung Norge	0.010923
Samsung Korea	0.010657
Samsung Egypt	0.008457
Samsung Gulf	0.008457
Samsung South Africa	0.008457
Samsung Indonesia	0.008395
Samsung Hong Kong	0.008395
Samsung Malaysia	0.008395
Samsung Philippines	0.008395
SamsungTaiwan	0.008395
Samsung Singapore	0.008395
Samsung Thailand	0.008395
Samsung Russia	0.00812

SamsungChile	0.007933
ForumMySamsung	0.00573
Samsung Türkiye	0.005241
Samsung Brasil	0.005241
Samsung Italia	0.005241
Samsung Romania	0.005241
Samsung Greece	0.005241
Samsung Česko a Slovensko	0.005241
Samsung Belgium	0.005241
Samsung Magyarország	0.005241
Samsung Bulgaria	0.005241
Samsung Nederland	0.005241
Samsung Srbija	0.005241
Samsung Danmark	0.005241
Samsung Hrvatska	0.005241
Samsung Slovenija	0.005241
Samsung Ireland	0.005241
Samsung Shqiperi	0.005241
Samsung Makedonija	0.005241
Samsung Bosna i Hercegovina	0.005241
Samsung Crna Gora	0.005241
Samsung Portugal	0.005241
Samsung Pakistan	0.005241
Samsung Levant	0.005241
Samsung Latinoamérica y Caribe	0.005241
Samsung Saudi Arabia	0.005241
Samsung Israel	0.005241
Samsung Argentina	0.005241
Samsung Colombia	0.005241
Samsung Perú	0.005241
Samsung Ukraine	0.005241
Samsung Kazakhstan	0.005241
Samsung Ghana	0.005241
Samsung Georgia	0.005241
Samsung Azerbaijan	0.005241
samsungmena	0.005189
CNET	0.005099
Samsung Developers	0.004949
Samsung Memory	0.004949



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.

## 14 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.