

ΟΙΚΟΝΟΜΙΚΟ  
ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΑΘΗΝΩΝ



ATHENS UNIVERSITY  
OF ECONOMICS  
AND BUSINESS

# ΤΗΚΕVASE KSANA ΕΚFONISI GΙΑ TO TΙ ΖΗΤΑ

**Your Subtitle**

January 27, 2024

# Περιεχόμενα

1	Εισαγωγή . . . . .	2
2	Λήψη δεδομένων . . . . .	2
3	Γραφική Αναπαράσταση Δικτύου . . . . .	3
4	Βασικά στοιχεία Δικτύου . . . . .	6
5	Component Measures . . . . .	7
6	Degree Measures . . . . .	9
6.1	Maximum Degree . . . . .	9
6.2	Average Node Degree . . . . .	9
6.3	Degree Distribution . . . . .	10
6.3.1	In-Degree . . . . .	10
6.3.2	Out-Degree . . . . .	12
6.3.3	Total Degree . . . . .	15
7	Centrality measures. . . . .	18
7.1	Degree . . . . .	18
7.2	Betweenness Centrality . . . . .	18
7.3	Closeness Centrality . . . . .	20
7.4	Eigenvector Centrality . . . . .	23
8	Clustering Effects. . . . .	26
8.1	Average Clustering Coefficient . . . . .	26
8.2	Number of Triangles . . . . .	27
8.3	Clustering Coefficient Distribution . . . . .	27
8.4	Existence of the Triadic Closure Phenomenon in the Friendship Neighborhood . . . . .	27

# 1 Εισαγωγή

Το Youtube είναι ένας ιστότοπος 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 ιστότοπο 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. Έτσι 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 του ομίλου εταιρειών Samsung που έχει ως σκοπό την ενημέρωση σχετικά με εκδηλώσεις, καινοτομες τεχνολογίες, εφαρμογές και υπηρεσίες, B2B solutions, παρουσιάσεις, και τις τελευταίες και καινοτόμες τεχνολογίες του ομίλου.

## 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 παραμέτροι που χρησιμοποιήθηκαν ήταν το seed (αρχικό κανάλι) με τη χρήση του id με crawl depth ίσον 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 υπόψην 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

επηρεάσουν την ανάλυση μας όπως για παράδειγμα ο έλεγχος δυπλοτύπων, όπου σε μια περίπτωση υπήρχε διπλότυπο όπου και εντιμετωπίστηκε μέσω του Gephii, ο έλεγχος για null τιμές κ.α. Σε μερικές περιπτώσεις υπήρχαν μη διαθέσιμες τιμές. Για παράδειγμα σε ορισμένους κόμβους, δεν υπήρχε στο αντίστοιχο κελί η χώρα ενώ ήταν γνωστή. Επομένως εισήχθησαν χοιροκίνητα οι τιμές όπου ήταν εφικτό. Σε άλλες περιπτώσεις, τυχόν σφάλματα αντιμετωπίζονται αναλόγως τη δεδομένη στιγμή όπου και αναφέρονται.

### 3 Γραφική Αναπαράσταση Δικτύου

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







Βλέποντας τα δεδομένα του δικτύου μας απο το Data Laboratory του Gephi, παρατηρήσαμε πως υπάρχουν κανάλια απο διάφορες χώρες. Επομένως θεωρήσαμε ενδιαφέρον να κάνουμε μία παραμετροποίηση με τις χώρες ως εξής. Ο χρωματισμός έγινε μέσω διαφορετικών χρωματων, τωσων, όσος και ο αριθμός των διαφορετικών χωρών, μέσω του partition tab. Στο σημείο αυτο, θεώρησαμε επίσης σημαντικό και την αναφορά του seed. Αυτό έγινε μεσω του μεγέθους των κόμβων μέσω του seedrank(αντίστοιχη μεταβλητη με την isseed εαν 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.



Απο 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 απο αυτο το κόμβο είναι επίσης με πράσινο χρώμα αφού και αυτοι οι κόμβοι είναι κανάλια απο την ίδια χώρα, την Νότιο Κορέα.

## 4 Βασικά στοιχεία Δικτύου

Το δίκτυο που μελετάμε έχει τα εξής βασικά στοιχεία:

- Αριθμός κόμβων: **76** διαφορετικά **κανάλια-κόμβοι**
- Αριθμός ακμών: **149** **συνδέσμοι** μέσω των οποίων συνδέονται τα κανάλια-κόμβοι
- Ο γράφος μας είναι **κατευθυνόμενος**. Δηλαδή κάθε σύνδεσμος απο 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.
- **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 απο 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.





## 6 Degree Measures

(mikri eiagogi AN DEN FKENNEI EN OK) Στο σημείο αυτό της analysis mas tha aafterthoume sta degree measures. ta degree measures einai...

### 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

isos na valw mia mikri isagogi?

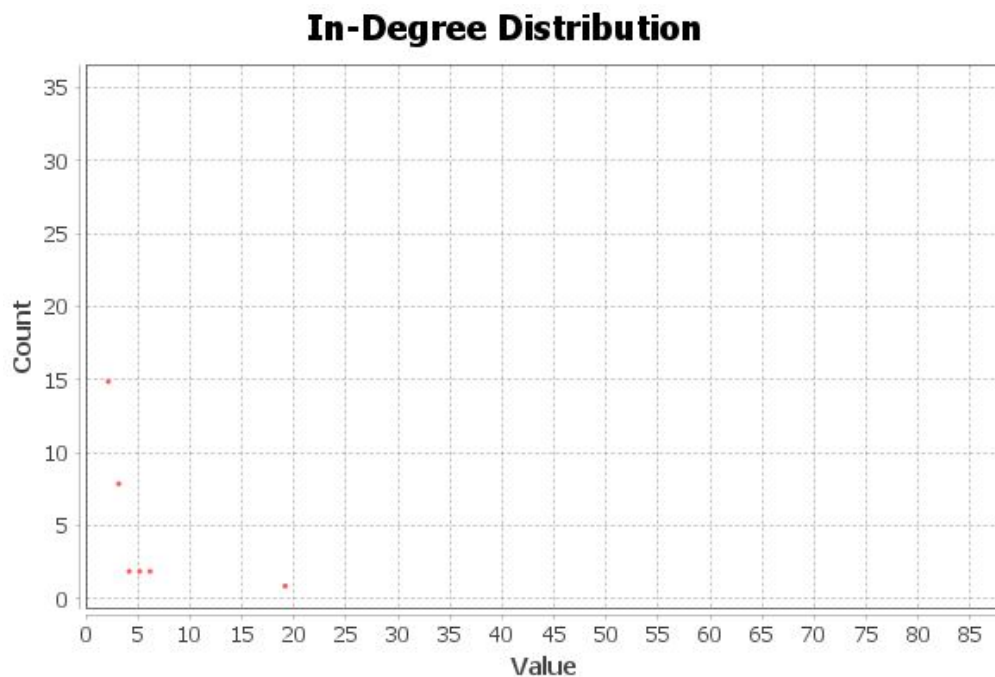
### 6.3.1 In-Degree

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:

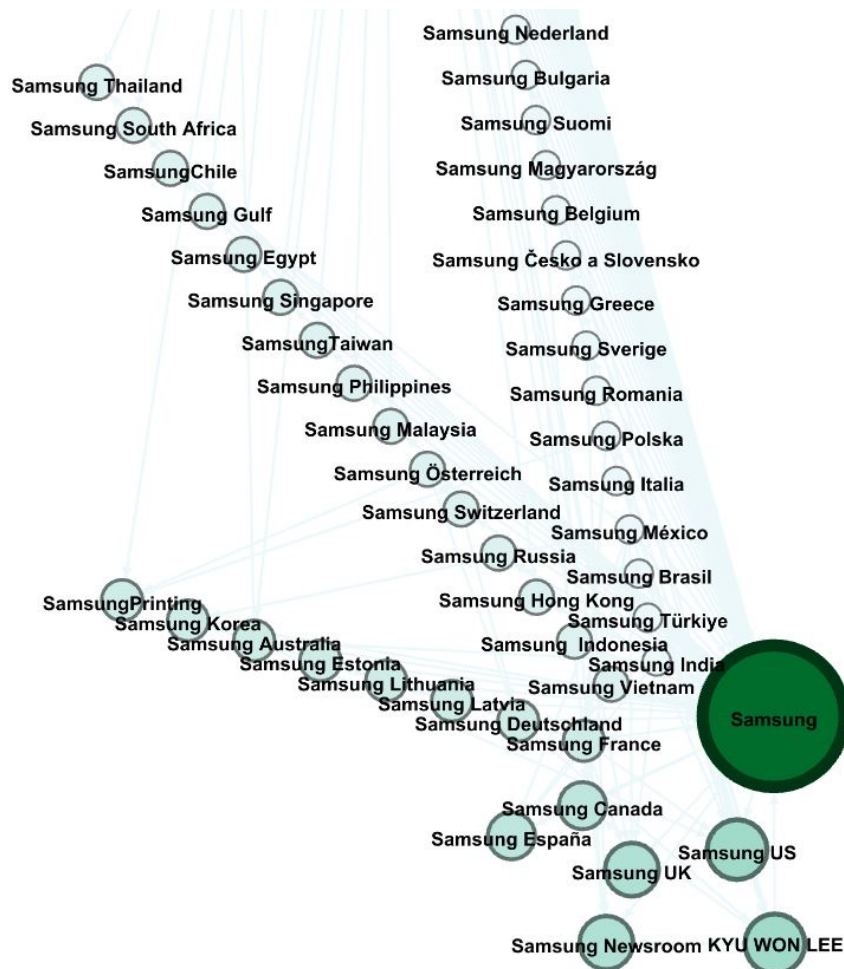
Label	In-Degree ▾
Samsung	19
KYU WON LEE	6
Samsung US	6
Samsung Newsroom	5
Samsung UK	5
Samsung Canada	4
Samsung España	4
Samsung Australia	3
Samsung Deutschland	3
Samsung Estonia	3
Samsung France	3
Samsung Korea	3
Samsung Latvia	3
Samsung Lithuania	3
SamsungPrinting	3
SamsungChile	2
Samsung Egypt	2
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 απο τα πιο πάνω, θα ήταν αρκετα ενδιαφέρον να δούμε πως αλλάζει το δίκτυο όσον αφορά μέγεθος και χρώμα κόμβων σε σε συνάρτηση με το In-Degree. dipla pou thn pio katw na mpei h ipolipi ths





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 χωρήστηκαν 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. Προταγωνιστικό 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 South Africa	0
Samsung New Zealand	10	SamsungTaiwan	0
Samsung México	9	Samsung Thailand	0
Samsung Polska	8	CNET	0
Samsung Maroc	7	ForumMySamsung	0
Samsung India	6	Samsung Argentina	0
Samsung France	5	Samsung Azerbaijan	0
Samsung Österreich	4	Samsung Belgium	0
KYO WON LEE	3	Samsung Bosna i Hercegovina	0
Samsung Estonia	3	Samsung Brasil	0
Samsung Latvia	3	Samsung Bulgaria	0
Samsung Lithuania	3	Samsung Česko a Slovensko	0
Samsung Switzerland	3	Samsung Colombia	0
Samsung Norge	3	Samsung Crna Gora	0
Samsung Suomi	3	Samsung Danmark	0
Samsung Sverige	3	Samsung Developers	0
Samsung Australia	2	Samsung Georgia	0
Samsung Deutschland	2	Samsung Ghana	0
Samsung Vietnam	2	Samsung Greece	0
Samsung Hrvatska	1	Samsung Ireland	0
Samsung Latinoamérica y Caribe	1	Samsung Israel	0
Samsung US	0	Samsung Italia	0
Samsung Newsroom	0	Samsung Kazakhstan	0
Samsung UK	0	Samsung Levant	0
Samsung Canada	0	Samsung Magyarország	0
Samsung España	0	Samsung Makedonija	0
Samsung Korea	0	Samsung Memory	0
SamsungPrinting	0	samsungmena	0
SamsungChile	0	Samsung Nederland	0
Samsung Egypt	0	Samsung Pakistan	0
Samsung Gulf	0	Samsung Perú	0
Samsung Hong Kong	0	Samsung Portugal	0
Samsung Indonesia	0	Samsung Romania	0
Samsung Malaysia	0	Samsung Saudi Arabia	0
Samsung Philippines	0	Samsung Shqiperi	0
Samsung Russia	0	Samsung Slovenija	0
Samsung Singapore	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. dipla pou thn pio katw na mpei h ipolipi ths



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.

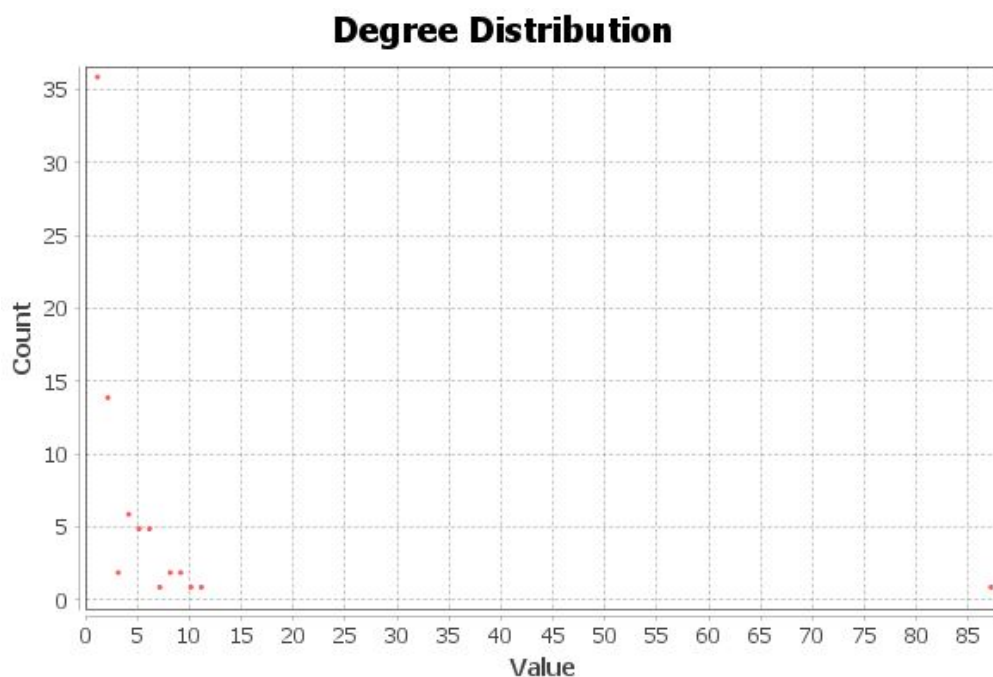
### 6.3.3 Total Degree

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.

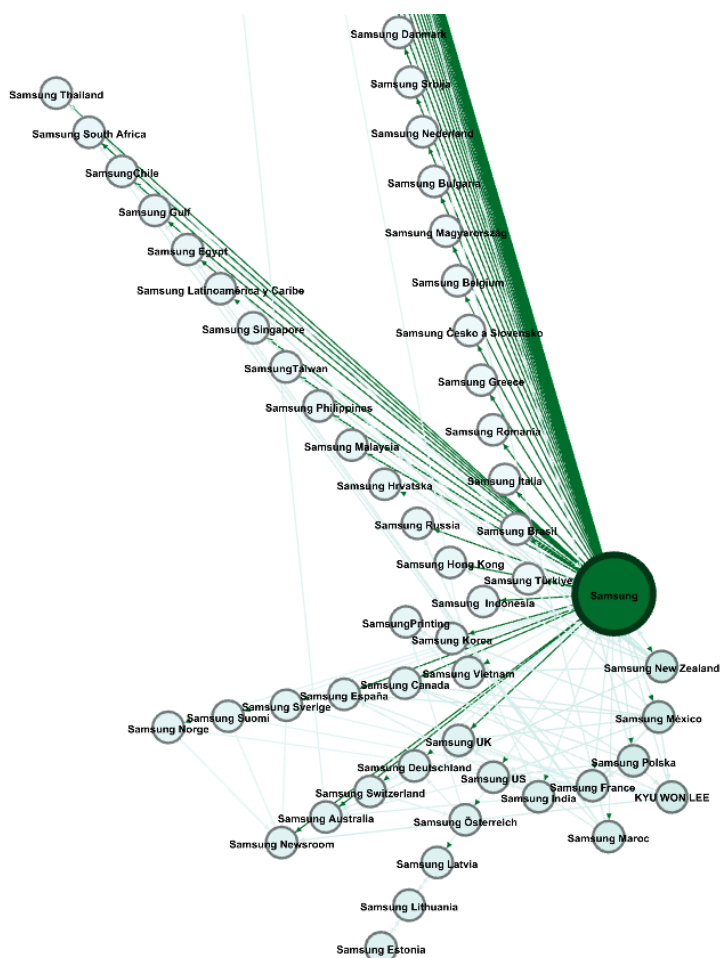
Label	Degree ▾
Samsung	87
Samsung New Zealand	11
Samsung México	10
Samsung Polska	9
KYU WON LEE	9
Samsung France	8
Samsung Maroc	8
Samsung India	7
Samsung US	6
Samsung Österreich	6
Samsung Latvia	6
Samsung Lithuania	6
Samsung Estonia	6
Samsung UK	5
Samsung Deutschland	5
Samsung Switzerland	5
Samsung 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 υπόψη το Total Degree αυτή τη φορά. **dipla pou thn pio katw na mpei h ipolipi ths**



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 κατα πλήθος κόμβων σε σχέση με πριν. Η διαφορά όμως έγκυται στο γεγονός πως όλοι σχεδόν οι κόμβοι που υπήρχαν και πριν στο Out-Degree, πέραν απο τον προφανές της Samsung, υπάρχουν και τώρα. Άρα φαίνεται πως το Out-Degree είναι αυτο που παίζει τον πιο σημαντικό ρόλο αφού όπως είπαμε και προηγουμένως είναι αυτό που καθορίζει ποια κανάλια θα προωθηθούν περισσότερο απο τον τρόπο που δουλεύει το Youtube μέσω των Featured Channels.

Τέλος να πούμε πως δεν έγινε κάποια αναφορά για το Weight Degree αφού στο δίκτυο που μελετάμε όλες οι ακμές έχουν ίσο βάρος πραγμα που δεν επηρεάζει τα δεδομένα μας. Επομένως δεν είχε νόημα η οποιαδήποτε αναφορά σε αυτό. **sioueftou full gia touto an j nmz en k dioti j sto fire etsi elalen**



## 7 Centrality measures

### 7.1 Degree

dame enikserw ti na grapsw, sthn ekfonisi lalei gia Degree enw sto Section6 pou en ta Degree Measures lalei gia Total Degree. enen idia touta ta 2? na mpw stes dialeksis gia to section7 na dw ti lalei

### 7.2 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 παίρνουν 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  $C_B(n_i) = \sum (\Sigma_{uw}(n_i) / \Sigma_{uw})$ .

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

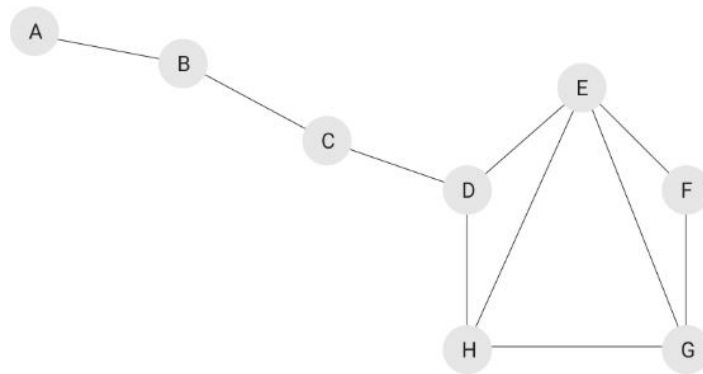
Samsung Memory	0.0
Samsung Malaysia	0.0
Samsung Makedonija	0.0
Samsung Magyarország	0.0
Samsung Lithuania	0.0
Samsung Levant	0.0
Samsung Latvia	0.0
Samsung Latinoamérica y Caribe	0.0
Samsung Korea	0.0
Samsung Kazakhstan	0.0
Samsung Italia	0.0
Samsung Israel	0.0
Samsung Ireland	0.0
Samsung Indonesia	0.0
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 απο 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 απο κάθε ήπειρο εκτός απο την Ευρώπη που συγκεντρώνει 7 χώρες.

### 7.3 Closeness Centrality

Το Closeness Centrality είναι μια μετρηκή 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 κειμένονται απο 0 εως 1.

Τον τρόπο με τον οποίο μπορούμε να υπολογίσουμε τη μετρηκή 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 αριθμό κόμβων του δικτύου μας και αφαιρούμε ένα, και τον διαιρούμε με το άθροισμα των συντομότερων μονοπατιών απο 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  $7/14 = 0.5$

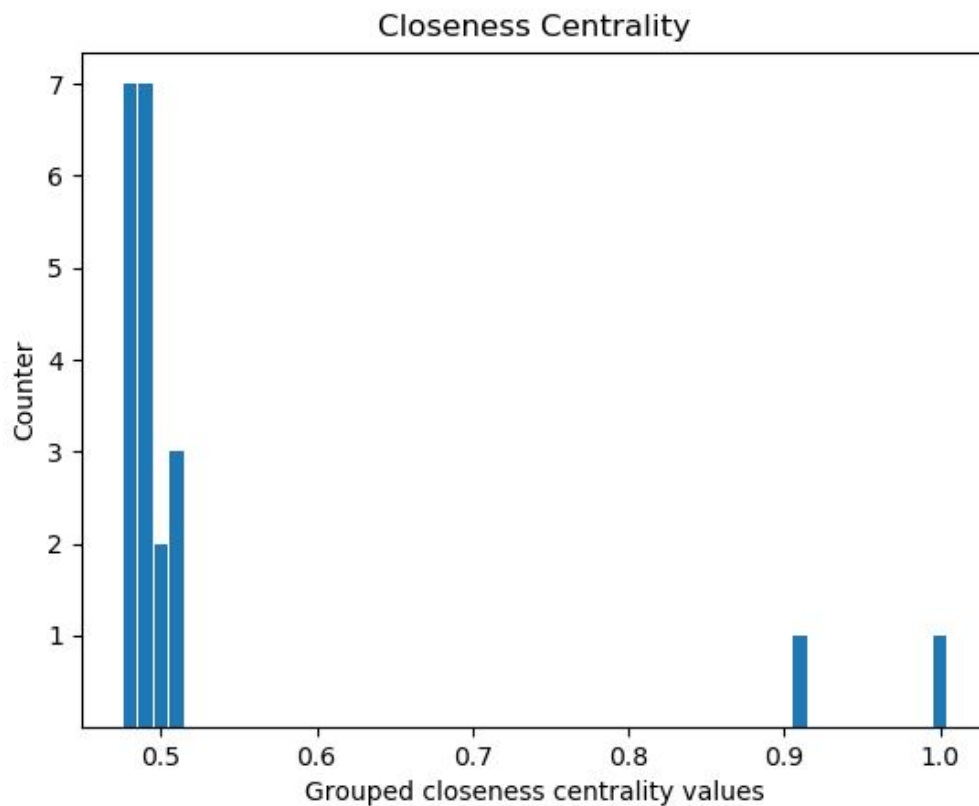
Να πούμε επίσης πως σε κατευθυνόμενους γράφους, παίζει ρόλο η φορά των ακμών. Για παράδειγμα σε περιπτώσεις όπου δεν υπάρχει διαδρομή λόγω φοράς ακμών, θεωρούμε μηδενικό το μονοπάτι. Έτσι γενικεύοντας το πιο πάνω υπάρχει περίπτωση κάποιος κόμβος να έχει μηδενικό Betweenness Centrality.

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.



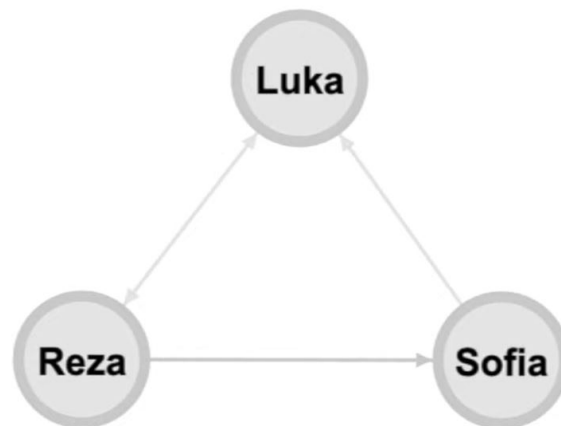
Βλέπουμε λοιπόν πως υπάρχει μια ομαδοποίηση των κόμβων μεταξύ των τιμών 0.48 και 0.51. Αντίθετα όμως, μόνο δύο κόμβοι έχουν υψηλές τιμές της τάξης των 0.91 και 1.0. Άρα αμέσως καταλαβαίνουμε ότι τα κανάλια που έχουν πιο σημαντικό ρόλο στο δίκτυο μας είναι αυτά με 0.91 και 1.0 με το όνομα αυτών να είναι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



## 7.4 Eigenvector Centrality

Το Eigenvector Centrality είναι ένα μετρο με το οποίο μπορούμε να καταλάβουμε την επηροή που μπορεί να έχει ένας κόμβος μέσα στο δίκτυο μας. Δείχνει δηλαδή πόσο σημαντικός είναι ένας κόμβος ανάλογα με το ποσο σημαντικοι είναι και οι κόμβοι-γειτονες 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 0 ή 1 εφόσον υπάρχει μονοπάτι που συνδέει τους κόμβους του δικτύου μας. Για παράδειγμα για τον κόμβο Reza η πρώτη στήλη στο γραμμικό μας σύστημα θα είναι 0(θεωρούμε πως δεν υπάρχει μονοπατι απο 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.

		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 ιδιοτιμή 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 US	0.461073
Samsung UK	0.399569
Samsung Newsroom	0.398726
Samsung España	0.356162
Samsung Canada	0.355318
Samsung Latvia	0.352047
Samsung Lithuania	0.352047
Samsung Estonia	0.352047
Samsung Deutschland	0.317048
Samsung Korea	0.312754
Samsung France	0.294658
Samsung Australia	0.294658
KYU WON LEE	0.282834
Samsung Switzerland	0.262445
Samsung Österreich	0.262445
Samsung Vietnam	0.251251
Samsung Indonesia	0.251251
Samsung Hong Kong	0.251251
Samsung Russia	0.251251
Samsung Malaysia	0.251251
Samsung Philippines	0.251251
SamsungTaiwan	0.251251
Samsung Singapore	0.251251
Samsung Egypt	0.251251
Samsung Gulf	0.251251
SamsungChile	0.251251
Samsung South Africa	0.251251
Samsung Thailand	0.251251
Samsung India	0.207843
Samsung Türkiye	0.207843
Samsung Brasil	0.207843
Samsung México	0.207843
Samsung Italia	0.207843
Samsung Polska	0.207843
Samsung Romania	0.207843
Samsung Sverige	0.207843

Samsung Greece	0.207843
Samsung Česko a Slovensko	0.207843
Samsung Belgium	0.207843
Samsung Magyarország	0.207843
Samsung Suomi	0.207843
Samsung Bulgaria	0.207843
Samsung Nederland	0.207843
Samsung Srbija	0.207843
Samsung Norge	0.207843
Samsung Danmark	0.207843
Samsung Hrvatska	0.207843
Samsung Slovenija	0.207843
Samsung Ireland	0.207843
Samsung Shqiperi	0.207843
Samsung Makedonija	0.207843
Samsung Bosna i Hercegovina	0.207843
Samsung Crna Gora	0.207843
Samsung Portugal	0.207843
Samsung Pakistan	0.207843
Samsung New Zealand	0.207843
Samsung Levant	0.207843
Samsung Latinoamérica y Caribe	0.207843
Samsung Saudi Arabia	0.207843
Samsung Israel	0.207843
Samsung Maroc	0.207843
Samsung Argentina	0.207843
Samsung Colombia	0.207843
Samsung Perú	0.207843
Samsung Ukraine	0.207843
Samsung Kazakhstan	0.207843
Samsung Ghana	0.207843
Samsung Georgia	0.207843
Samsung Azerbaijan	0.207843
SamsungPrinting	0.141417
CNET	0.061504
ForumMySamsung	0.052455
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 λαμβανοντας υπόψη 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 αριθμητική διαφορά μεταξύ του πρώτου σε σκόρ καναλιού και όλων των υπολοίπων. Στη πρώτη θέση λοιπόν βρίσκεται ξανά το κανάλι Samsung με σκόρ 1 ενώ τα υπόλοιπα κανάλια ξεκινούν από σκόρ 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 βρισκονται 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 υποδεικνύει πως είναι ενσωματωμένοι οι κόμβοι στο δίκτυο σε σχέση με τους γειτονικούς τους κόμβους και υπολογίζεται ως εξής. Επιλέγουμε έναν τυχαίο κόμβο όπου στη συνέχεια επιλέγουμε 2 γείτονες του τυχαία και στη συνέχεια ελέγχουμε εάν είναι συνδεδεμένοι. Αθροίζοντας τα αποτελέσματα της πάραπανω 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**

## **8.3 Clustering Coefficient Distribution**

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