



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΙΩΑΝΝΙΝΩΝ

ΣΧΟΛΗ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΠΜΣ: ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΔΙΚΤΥΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**Βιβλιομετρική ανάλυση των επιστημονικών μελετών σχετικά με τα
ευφυή προγραμματιζόμενα δίκτυα (SDN)**

Μπάου Σπυριδούλα

Επιβλέπουσα: Δρ.Μαργαρίτη Σπυριδούλα

Άρτα, Μάρτιος, 2024

**A REVIEW AND BIBLIOMETRIC ANALYSIS OF SOFTWARE
DEFINED NETWORKING (SDN) RESEARCH**

Εγκρίθηκε από τριμελή εξεταστική επιτροπή

Αρτα, Μάρτιος 2024

ΕΠΙΤΡΟΠΗ ΑΞΙΟΛΟΓΗΣΗΣ

- Επιβλέπουσα καθηγήτρια
Σπυριδούλα Μαργαρίτη,
ΕΔΙΠ
- Μέλος επιτροπής
Ελευθέριος Στεργίου,
Αναπληρωτής καθηγητής
- Μέλος επιτροπής
Κωνσταντίνος Αγγέλης,
Καθηγητής

© Μπάου, Σπυριδούλα,2023.

Με επιφύλαξη παντός δικαιώματος. Allrightsreserved.

Δήλωση μη λογοκλοπής

Δηλώνω υπεύθυνα και γνωρίζοντας τις κυρώσεις του Ν. 2121/1993 περί Πνευματικής Ιδιοκτησίας, ότι η παρούσα μεταπτυχιακή εργασία είναι εκ ολοκλήρου αποτέλεσμα δικής μου ερευνητικής εργασίας, δεν αποτελεί προϊόν αντιγραφής ούτε προέρχεται από ανάθεση σε τρίτους. Όλες οι πηγές που χρησιμοποιήθηκαν (κάθε είδους, μορφής και προέλευσης) για τη συγγραφή της περιλαμβάνονται στη βιβλιογραφία.

Μπάου, Σπυριδούλα

Υπογραφή

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω στους συγγενείς και φίλους που ήταν δίπλα μου μέχρι την ολοκλήρωση της ακαδημαϊκής μου φοίτησης. Η ενθάρρυνση και η υποστήριξη όλων έπαιξαν σημαντικό ρόλο στην περάτωση της διπλωματικής μου εργασίας που ήταν ένα κομμάτι τόσο ενδιαφέρον και τόσο απαιτητικό επίσης. Ιδιαίτερες ευχαριστίες καθώς και την βαθύτατη εκτίμηση μου θα ήθελα να εκφράσω στην καθηγήτριά μου κ. Σπυριδούλα Μαργαρίτη για την καθοδήγηση, την υποστήριξη και την βοήθεια που μου πρόσφερε. Ευχαριστώ θερμά.

ΠΙΝΑΚΑΣ ΣΥΝΤΟΜΟΓΡΑΦΙΩΝ

SDN=Software Defined Network

OF=OpenFlow

ONF= Open Networking Foundation

IoT=Internet of Things

ΒΔ= Βάση Δεδομένων

H= h-index

IF= Impact Factor

VOS= Visualization Of Similarities

ISI= Institute for Scientific Information

CI= Collaboration Index

CC= Collaboration Coefficient

NCP= Number of Cited Publications

PCP= Proportion of Cited Publications

CCP= Citations per Cited Publications

G= g-index

I= i-index

TC= Total Citations

AC= Average Citations

TP= Total Publications

NCA= Number of Contributing Authors

SA= Sole-Authored publications

CA= Co-Authored publications

NAY= Number of Active Year of publication

PAY= Productivity per Active Year of publication

ΠΕΡΙΛΗΨΗ

Η βιβλιομετρική ανάλυση αφορά συνήθως έρευνες που θέλουν να εξερευνήσουν και να αναλύσουν μεγάλο όγκο δεδομένων. Μέσω αυτής γίνεται η ανάλυση, η επεξεργασία και η καταγραφή συμπερασμάτων. Οι τεχνικές αυτής της ανάλυσης χωρίζονται σε δύο κατηγορίες: την επιστημονική χαρτογράφηση και την ανάλυση απόδοσης. Οι βιβλιομετρικές μελέτες έχουν αποκτήσει μεγάλη δημοτικότητα και αυτό οφείλεται στο ελεύθερο λογισμικό και στις έγκυρες βάσεις δεδομένων. Είναι χρήσιμες για την κατανόηση της επιστημονικής γνώσης και της εξέλιξης των καθιερωμένων πεδίων και μπορούν να βοηθήσουν τον μελετητή να κατανοήσει πολλά μη δομημένα γεγονότα. Παρέχουν επίσης μια σταθερή βάση για την προώθηση ενός τομέα με νέους και ουσιαστικούς τρόπους. Τα SDN είναι μία εξελισσόμενη τεχνολογία που ασχολείται με την διαχείριση των δικτύων προσβλέποντας στην αποτελεσματικότητα και την μείωση της πολυπλοκότητας τους. Σκοπός της παρούσας διπλωματικής εργασίας είναι η βιβλιομετρική ανάλυση των επιστημονικών μελετών που αφορούν το δικτυακό περιβάλλον SDN κάνοντας χρήση δεδομένων από έγκυρες βάσεις δεδομένων. Στη συνέχεια της εργασίας θα γίνει χρήση των λογισμικών VOSViewer, Gephi και της γλώσσας R με σκοπό την ανάδειξη ερευνητικών τάσεων και την ανάδειξη της επιρροής που υπάρχει, την ανακάλυψη σχέσεων συνεργασίας και την απεικόνισή τους. Η βάση δεδομένων που θα χρησιμοποιηθεί στην παρούσα μελέτη είναι το Scopus, μια διεθνώς αναγνωρισμένη πηγή επιστημονικής βιβλιογραφίας που χρησιμοποιείται ευρέως για βιβλιομετρικές μελέτες λόγω της ποιότητας των δημοσιεύσεων και των αναλυτικών μεταδεδομένων. Ολοκληρώνοντας την έρευνα θα καταγραφούν τα αποτελέσματα της μελέτης με σκοπό την αποκάλυψη κύριων κατευθύνσεων της έρευνας στον τομέα των SDN, τη χαρτογράφηση των επιστημονικών μελετών, την ανάλυση της απόδοσής τους, την αποτύπωση των διαδρομών εξέλιξης της έρευνας τα τελευταία χρόνια.

Λέξεις-κλειδιά: Βιβλιομετρική Ανάλυση, Δικτύωση καθορισμένη από λογισμικό, Ελεγκτές, Πρωτόκολλο, Βάσεις Δεδομένων

ABSTRACT

Bibliometric analysis usually concerns researchers who want to explore and analyze a large amount of data. Through this, the analysis, processing, and recording of conclusions are done. The techniques of this analysis fall into two categories: scientific mapping and performance analysis. Bibliometric studies have gained a lot of popularity and this is due to free software and authoritative databases. They are useful for understanding scientific knowledge and the evolution of established fields and can help the scholar understand many unstructured facts. They also provide a solid foundation for advancing a field in new and meaningful ways. SDN is an advanced technology that deals with managing networks with an eye toward efficiency and reducing their complexity. The purpose of this thesis is the bibliometric analysis of the scientific studies concerning the SDN network environment using data from valid databases. In the continuation of the work, the software VOSViewer, Gephi and the R language will be used in order to highlight research trends and highlight the influence that exists, discover collaborative relationships and to depict them. The database used in this study is Scopus, recognized for bibliometric studies source of scientific literature that is widely used due to the quality of its publications and analytical metadata. Completing the survey that records the results of the study in order to reveal the main directions of research in the field of SDN, mapping scientific studies, analyzing their performance, and mapping the paths of research developments in recent years.

Keywords: Bibliometric Analysis, Software Defined Networking-SDN, Controllers, Protocol, Databases

1 Περιεχόμενα

ΕΥΧΑΡΙΣΤΙΕΣ.....	6
ΠΙΝΑΚΑΣΣΥΝΤΟΜΟΓΡΑΦΙΩΝ	7
ΠΕΡΙΛΗΨΗ.....	8
ABSTRACT	9
1. Εισαγωγή.....	12
1.1 Γενικά.....	12
1.2 Κίνητρο και στόχος	12
1.3 Δομή εργασίας	13
2. SDN	14
2.1 Επισκόπηση της δικτύωσης SDN	14
2.2 Αρχιτεκτονική SDN.....	14
2.3 Πλεονεκτήματα SDN δικτύωσης.....	15
2.4 Μειονεκτήματα SDN δικτύωσης	16
2.5 Τρέχουσα κατάσταση	16
2.6 Ιστορική αναδρομή.....	17
2.7 OpenFlow	18
3. Βιβλιομετρία ως μέθοδος ανάλυσης.....	19
3.1 Βιβλιομετρία.....	19
3.1.1 1η κατηγορία: Ανάλυση απόδοσης (Performance analysis)	19
3.1.2 2η κατηγορία: Επιστημονική χαρτογράφηση (Science mapping)	22
3.1.3 3η κατηγορία: Ανάλυση δικτύου (Network analysis)	23
3.2 Τεχνικές ανάλυσης.....	24
3.2.1 Μέθοδος Layer –Ταξινόμηση σε επίπεδα.....	25
3.2.2 Μέθοδος Clustering - Ομαδοποίηση	26
3.3 Πηγές δεδομένων και εργαλεία λογισμικού.....	26
3.3.1 VOSViewer	28
3.3.2 Gephi	29
3.3.3 R Studio.....	29
3.4 Ανάλυση SDN.....	30
4. Μεθοδολογία	32
4.1 Συλλογή δεδομένων	32
4.2 Μέθοδοι ανάλυσης στο VOSViewer	34
4.3 Μέθοδοι ανάλυσης του Gephi	35

4.4	Μέθοδοι ανάλυσης του R Studio	35
5.	Ανάλυση δεδομένων	37
5.1	Ανάλυση μέσω του λογισμικού VOSViewer.....	37
5.1.1	Χαρτογράφηση συνύπαρξης λέξεων-κλειδιών (Co-occurrence of Author keywords) 37	
5.1.2	Χαρτογράφηση κοινής συνεργασίας μεταξύ χωρών (Co-authorship countries) 42	
5.1.3	Χαρτογράφηση κοινής συγγραφής μεταξύ οργανισμών (Co-authorship organizations).....	47
5.1.4	Χαρτογράφηση συν-συγγραφέων (Co-authorship authors)	50
5.2	Ανάλυση δεδομένων μέσα από το λογισμικό Gephi.....	53
5.2.1	Ανάλυση της σχέσης μεταξύ “Authors&Citedby”	54
5.2.2	Ανάλυση της σχέσης μεταξύ “Authors&Authorkeywords”	61
5.3	Ανάλυση δεδομένων με την χρήση της Rlanguage	69
5.3.1	Επιστημονικοί χάρτες - Χάρτες συνεργασίας.....	75
6.	Συμπεράσματα	79
	Βιβλιογραφία.....	80

1. Εισαγωγή

1.1 Γενικά

Η ραγδαία ανάπτυξη των προγραμματιζόμενων δικτύων (Software Defined Networking-SDN) έχει προσελκύσει το ενδιαφέρον της ερευνητικής και ακαδημαϊκής κοινότητας και έχει οδηγήσει σε πολυάριθμες μελέτες τα τελευταία χρόνια. Οι ερευνητές μελετούν, διερευνούν, δοκιμάζουν την πρόσφατη καινοτομία από διαφορετικές πλευρές, αξιολογούν τις διαθέσιμες τεχνολογίες (π.χ., ελεγκτές SDN) και προτείνουν νέες λύσεις σε σημαντικά προβλήματα, όπως δρομολόγηση ή ανακάλυψη τοπολογίας, κ.ά. Ένα σημαντικό εργαλείο για την ταξινόμηση της επιστημονικής γνώσης, την παρακολούθηση της πορείας της εξέλιξης, και τον προσδιορισμό των πεδίων που μελετήθηκαν ή δεν μελετήθηκαν είναι η βιβλιομετρική ανάλυση. Οι βιβλιομετρικές μελέτες έχουν καθιερωθεί ήδη από τον 18ο αιώνα και έχουν αποκτήσει μεγάλη δημοτικότητα. Αυτό οφείλεται στο ελεύθερο διαθέσιμο λογισμικό και στις έγκυρες βάσεις δεδομένων επιστημονικών αναφορών.

Στα πλαίσια αυτής της μελέτης, με τη βοήθεια της βιβλιομετρικής ανάλυσης θα αξιοποιηθούν πρωτογενή δεδομένα από την ευρέως γνωστή πηγή επιστημονικών δημοσιεύσεων, το Scopus. Θα εφαρμοστούν τεχνικές ομαδοποίησης και κατάταξης με τη χρήση κατάλληλων εργαλείων λογισμικού και βιβλιομετρικών δεικτών και αναμένεται να:

- αναδειχθούν οι κύριες κατευθύνσεις και τα κύρια χαρακτηριστικά της έρευνας στον τομέα των SDN,
- αποκαλυφθούν/παρουσιαστούν οι διαδρομές εξέλιξης της έρευνας τα τελευταία χρόνια,
- χαρτογραφηθούν οι επιστημονικές μελέτες,
- αναλυθεί/αποτιμηθεί η απόδοσή τους,
- αναδειχθούν ανοιχτά ζητήματα προς μελέτη.

1.2 Κίνητρο και στόχος

Η βιβλιομετρική ανάλυση αποτελεί ένα χρήσιμο εργαλείο για τη διαχείριση του μεγάλου όγκου δεδομένων που αφορούν επιστημονικές δημοσιεύσεις και την αποτύπωση του ερευνητικού αντίκτυπου. Μέσω της βιβλιομετρικής ανάλυσης, οι ερευνητές μπορούν να αποκαλύψουν τις αναδυόμενες τάσεις σε μια ερευνητική περιοχή, τα μοτίβα συνεργασίας και τα δομικά στοιχεία μιας έρευνας.

Αυτή η ευρεία εφαρμογή αποτέλεσε το κίνητρο για την παρούσα διατριβή που αποσκοπεί να αναλύσει και να ερμηνεύσει τον τεράστιο όγκο επιστημονικών δεδομένων για τα προγραμματιζόμενα ευφυή δίκτυα. Στόχος είναι να χρησιμοποιηθεί η βιβλιομετρική ανάλυση για να αποκτηθεί ενιαία εικόνα για την υπό μελέτη επιστημονική περιοχή, να αντληθούν γνώσεις, να εντοπιστούν τυχόν κενά και να επισημανθούν νέες ιδέες για έρευνα.

1.3 Δομή εργασίας

Το πρώτο κεφάλαιο της διπλωματικής ξεκινά με το εισαγωγικό κομμάτι στο οποίο γίνεται αναφορά στο θέμα της έρευνας. Συνεχίζοντας στο δεύτερο κεφάλαιο παρουσιάζεται η τεχνολογία των δικτύων SDN, η δομή τους, τα πλεονεκτήματα και μειονεκτήματα της δικτύωσης αυτής, καθώς και την εξέλιξή τους. Επίσης γίνεται και μία αναφορά στο πρωτόκολλο Openflow. Το τρίτο κεφάλαιο εξηγεί την βιβλιομετρία και παρουσιάζει τα εργαλεία VOSViewer καιGerhi, επιπλέον ένα μέρος του αναφέρεται στις τεχνικές ανάλυσης. Στο τέταρτο κεφάλαιο συγκεντρώθηκε η μεθοδολογία για την δημιουργία της παρούσας εργασίας. Γίνεται αναφορά στην βάση δεδομένων και στην ανάλυση των δεδομένων μέσα από τα δύο λογισμικά και την γλώσσα R. Στο πέμπτο κεφάλαιο παρουσιάζονται τα αποτελέσματα από το πειραματικό κομμάτι της εργασίας. Στο έκτο και τελευταίο κεφάλαιο αναλύονται τα συμπεράσματα που προκύπτουν από την παρούσα διπλωματική.

2. SDN

2.1 Επισκόπηση της δικτύωσης SDN

Η δικτύωση καθορισμένη από λογισμικό αφορά την προσπάθεια ελέγχου και διαχείρισης των δικτύων από λογισμικό και όχι από υλικό όπως γινόταν μέχρι πριν λίγα χρόνια. Στόχος της προσέγγισης αυτής είναι η αποτελεσματική διαχείριση των δικτύων και η εφαρμογή πολιτικών ασφάλειας. Ο διαχωρισμός του επιπέδου ελέγχου με το επίπεδο δεδομένων τα κάνει πιο ευέλικτα, ξεφεύγοντας από τα πρότυπα της παραδοσιακής δικτύωσης και προσφέροντάς τους πέρα από την ευελιξία και την επεκτασιμότητα. Αυτός ο διαχωρισμός έχει σημαντικό αντίκτυπο τόσο στο επίπεδο δεδομένων όσο και στο επίπεδο ελέγχου του δικτύου[1-2].

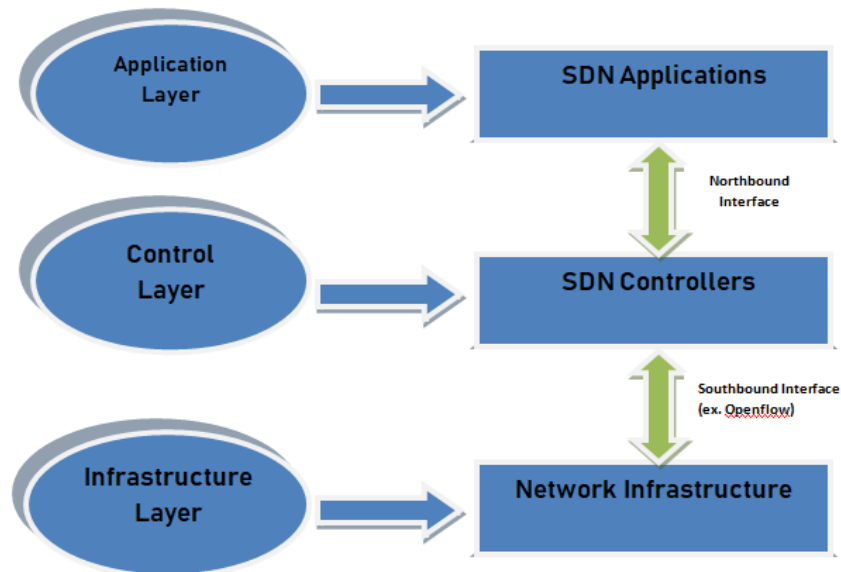
Πιο συγκεκριμένα ο διαχωρισμός αφορά:

- Επίπεδο δεδομένων (ή επίπεδο προώθησης δεδομένων): βασίζεται στην λειτουργία του επιπέδου ελέγχου και προωθεί δεδομένα αμφίδρομα μέσω των διεπαφών.
- Επίπεδο ελέγχου: είναι υπεύθυνο για την δρομολόγηση των δεδομένων μέσα σε ένα δίκτυο και για την διαχείριση σε ό,τι αφορά το δίκτυο, όπως είναι ο πίνακας προωθήσεων ή δρομολογήσεων, η σχεδίαση της τοπολογίας του δικτύου, ακόμα και για το πώς μεταδίδονται τα δεδομένα από τον δρομολογητή καθώς και για την λήψη αποφάσεων.

2.2 Αρχιτεκτονική SDN

Όπως φαίνεται και στην Εικόνα 2.1, η αρχιτεκτονική SDN αποτελείται από τρία επίπεδα. Τα επίπεδα αυτά είναι:

- Επίπεδο εφαρμογών
- Επίπεδο ελέγχου
- Επίπεδο προώθησης δεδομένων



Εικόνα 2.1Η αρχιτεκτονική SDN

Στο επίπεδο εφαρμογών περιλαμβάνονται προγράμματα τα οποία είναι υπεύθυνα για να ρυθμίζουν τον τρόπο λειτουργίας και τις απαιτήσεις που υπάρχουν σε ένα δίκτυο, καθώς και να επικοινωνούν με τον ελεγκτή SDN. Η ρύθμιση της λειτουργίας του δικτύου γίνεται αυτόματα ή και με τη βοήθεια κάποιου χειριστή.

Στο επίπεδο ελέγχου παρατηρούμε ότι υπάρχει ο κεντρικός ελεγκτής και δύο διεπαφές. Η βόρεια διεπαφή συνδέει τις εφαρμογές με τον ελεγκτή, έπειτα γίνεται η επεξεργασία των πληροφοριών που μεταφέρθηκαν από το πρώτο στο δεύτερο επίπεδο και τέλος η νότια διεπαφή επικοινωνεί με το επίπεδο δεδομένων μεταφέροντας τις απαραίτητες πληροφορίες στις συσκευές που περιλαμβάνει.

Το επίπεδο δεδομένων ασχολείται με την προώθηση των δεδομένων. Σε αυτό το επίπεδο περιλαμβάνονται συσκευές προώθησης (forwarding devices) που εκτελούν την προώθηση πακέτων σύμφωνα με τις οδηγίες του ελεγκτή και βάση τις απαιτήσεις του δικτύου και στοιχεία δικτύου (network elements) όπου παρέχουν υπηρεσίες δικτύου όπως τείχη προστασίας, εξισορροπητές φορτίου και συστήματα ανίχνευσης εισβολών [3].

2.3 Πλεονεκτήματα SDN δικτύωσης

Βασικότερο πλεονέκτημα που έφερε αυτή η τεχνολογία είναι η εικονοποίηση του δικτύου. Με το network virtualization η διαχείριση του δικτύου γίνεται ευκολότερη και αποτελεσματικότερη. Μέσα από τις κατάλληλες εικονικές μηχανές, οι διαχειριστές μπορούν να παρακολουθούν σε πραγματικό χρόνο την απόδοση και συμπεριφορά του δικτύου και να επεμβαίνουν όπου κρίνεται αναγκαίο [4].

Η εικονοποίηση με την σειρά της επιφέρει στην τεχνολογία SDN την αύξηση της ασφάλειας μέσα σε ένα δίκτυο. Στον εκπαιδευτικό τομέα εντάσσεται η χρήση SDN και η εκμάθηση γίνεται πιο προσιτή λόγω του μηδενικού κόστους διότι δεν χρειάζεται εξοπλισμό.

Για τον επιχειρηματικό κλάδο η διαχείριση των επιχειρήσεων και των οργανισμών γίνεται ευκολότερη και με μειωμένο λειτουργικό κόστος διότι δεν χρειάζονται συντηρήσεις και αντικατάσταση του εξοπλισμού. Οι χειριστές με την σειρά τους μπορούν να «δοκιμάζουν» και να παραμετροποιούν χωρίς να υπάρχει κίνδυνος στο να επιφέρει οικονομικές ζημιές σε περίπτωση που συμβεί κάποιο λάθος.

Στον ευρύτερο χώρο η αρχιτεκτονική αυτή κάνει την μεταφορά δεδομένων πιο γρήγορη και αποτελεσματική και προσφέρει σε υπηρεσίες cloud καλύτερη διαχείριση απ' ότι οι παραδοσιακές αρχιτεκτονικές δικτύου.

2.4 Μειονεκτήματα SDN δικτύωσης

Παρόλο που τα πλεονεκτήματα που συναντήσαμε παραπάνω είναι αρκετά για μία τόσο νέα τεχνολογία, τα SDN έχουν ορισμένα μειονεκτήματα που χρήζουν αντιμετώπιση σε βάθος χρόνου.

Αρχικά, όταν μία επιχείρηση ή ένας οργανισμός ξεκινήσει να χρησιμοποιήσει τα SDN θα χρειαστεί να χτίσει το δίκτυό του εξ' ολοκλήρου, αυτό συνεπάγεται πρωτίστως ότι θα υπάρχει οικονομική επιβάρυνση. Σε δεύτερο μέρος, πέραν της οικονομικής αγοράς του εξοπλισμού και της κατασκευής και λόγω του ότι η τεχνολογία είναι πολύ νέα στον χώρο, το προσωπικό χρειάζεται επιπλέον εκπαίδευση διότι θα έρθει σε επαφή με νέες γλώσσες προγραμματισμού, API's και νέα εργαλεία απ' ότι ήταν συνηθισμένο μέχρι πριν την χρήση SDN.

Επίσης, επειδή η ασφάλεια του δικτύου βασίζεται στον controller ευδαιμονεί κινδύνους στην κατάρρευση του συστήματος σε περίπτωση απροσεξίας των χειριστών.

Όπως αναφέρθηκε και πιο πάνω η δικτύωση καθορισμένη από το λογισμικό εμφανίσθηκε το 2008. Παρόλο που είναι πρόσφατη είναι ταχέως εξελισσόμενη κάτι που δεν αποδεικνύει ότι η αξιοπιστία της και η αποτελεσματικότητά της μπορεί να είναι έμπιστη.

Τέλος, είναι ευάλωτο σε επιθέσεις και αυτό πιθανότατα να οφείλεται στον ψηφιακό εξοπλισμό που δεν έχει συμβατό κάποιο τοίχος προστασίας [5].

2.5 Τρέχουσα κατάσταση

Τα δίκτυα καθορισμένα από λογισμικό είναι μία τεχνολογία ταχέως εξελισσόμενη τα τελευταία χρόνια προσφέροντας πολλαπλά οφέλη όπως αυξημένη απόδοση και ευελιξία. Έχει προσφέρει επίσης σημαντική ανάπτυξη στη διαχείριση του δικτύου από μέρους της αποτελεσματικότητας και της ευελιξίας. Τα κύρια χαρακτηριστικά αυτής της δικτύωσης είναι η ανάπτυξη νέων αρχιτεκτονικών και πρωτόκολλων δικτύου συμπεριλαμβάνοντας

ελεγκτές, μηχανισμούς εξισορρόπησης φορτίου οι οποίοι βελτιώνουν την επεκτασιμότητα και την απόδοση των SDN δικτύων.

Εξίσου σημαντικό στην τεχνολογία αυτή είναι η εμφάνιση νέων εφαρμογών στα δίκτυα 5G, στο cloud computing και στο Internet of Things (IoT). Στα δίκτυα 5G το SDN επιτρέπει τον διαχωρισμό του δικτύου και την δυναμική του διαχείριση. Στον τομέα του cloud computing επιτρέπει τη δημιουργία εικονικών δικτύων τα οποία μπορούν να διαχειριστούν και στο IoT βοηθάει στην επεκτασιμότητα και στην ασφάλεια του δικτύου.

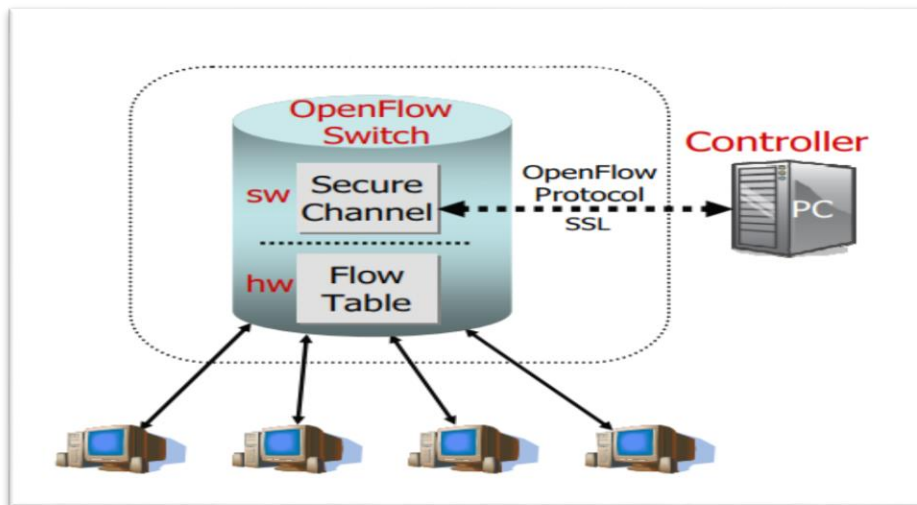
Η εξέλιξη της SDN τεχνολογίας στον κλάδο της μηχανικής μάθησης είναι σημαντική δεδομένου ότι οι αλγόριθμοι μηχανικής μάθησης μπορούν να βοηθήσουν την αυτοματοποίηση των εργασιών διαχείρισης δικτύου, στην βελτιστοποίηση της απόδοσης του δικτύου και της ασφάλειας. Καθώς το SDN είναι ακόμη εξελισσόμενο και πολλά υποσχόμενο για το μέλλον των δικτύων οι ειδικοί διευρύνουν τεχνικές μηχανικής μάθησης για την διαχείριση του δικτύου (εποπτευόμενη ή χωρίς επίβλεψη μάθηση) και καλούνται να αντιμετωπίσουν προβλήματα που αφορούν την ασφάλεια ή την ενορχήστρωση των δικτύων SDN.

2.6 Ιστορική αναδρομή

Το 2008 μια ομάδα ερευνητών από το πανεπιστήμιο του Στάνφορντ παρουσίασαν μία μελέτη με τίτλο «Software-defined Networking with Openflow» όπου έκαναν λόγο για τον διαχωρισμό των επιπέδων ελέγχου και δεδομένων στις αρχιτεκτονικές δικτύου. Οι κλασσικές δομές δικτύου βασίζονται σε πραγματικό υλικό και λογισμικό, κάτι που έκανε τα δίκτυα να μην μπορούν να διαχειριστούν και να αυξηθούν καθώς αυτά μεγαλώνουν σε μέγεθος και πολυπλοκότητα. Οι ερευνητές πρότειναν μία νέα αρχιτεκτονική που θα διαχώριζε το επίπεδο ελέγχου από το επίπεδο δεδομένων με πρωταγωνιστή έναν κεντρικό ελεγκτή SDN ο οποίος θα επικοινωνούσε με συσκευές δικτύου χρησιμοποιώντας ένα πρωτόκολλο όπως το OpenFlow. Με αυτή την προσέγγιση η διαχείριση του δικτύου θα ήταν πιο ευέλικτη και επεκτάσιμη καθώς και η εφαρμογή νέων υπηρεσιών δικτύου θα γινόταν με ευκολότερο τρόπο.

Τον Μάρτιο του 2011 δημιουργήθηκε επίσημα από τον οργανισμό Open Networking Foundation (ONF) με σκοπό να προωθήσει την τεχνολογία SDN και να εμπορευματοποιήσει το Openflow. Ένα χρόνο μετά κυκλοφόρησε το Openflow v1.3 ξεκινώντας την αλλαγή στον κόσμο των δικτύων.

2.7 OpenFlow



Εικόνα 2.2 Openflow[34]

Το OpenFlow (OF) είναι ένα πρωτόκολλο επικοινωνίας που εφαρμόζεται μεταξύ συσκευών software & hardware με την προϋπόθεση ότι οι συσκευές έχουν τη δυνατότητα υποστήριξης του. Δημιουργήθηκε στο πανεπιστήμιο του Stanford και είναι το πρώτο πρωτόκολλο για επικοινωνία δικτύων SDN.

Μέσω αυτού οι χρήστες μπορούν να εκτελούν πειράματα στα δίκτυα που χρησιμοποιούν καθημερινά. Βασίζεται σε ένα Ethernet switch, με έναν εσωτερικό πίνακα ροής και μία τροποποιημένη διεπαφή για την προσθήκη και αφαίρεση καταχωρήσεων ροής. Με το OF μπορεί να γίνει προγραμματισμός στον πίνακα ροής, σε διάφορους διακόπτες και δρομολογητές. Ένας διαχειριστής δικτύου μπορεί να χωρίσει την κυκλοφορία σε ροές παραγωγής και έρευνας. Οι χρήστες μπορούν να ελέγχουν τις δικές τους ροές, επιλέγοντας τις διαδρομές που ακολουθούν τα πακέτα τους και την επεξεργασία που λαμβάνουν.

Με αυτόν τον τρόπο οι διαχειριστές δικτύου μπορούν να δοκιμάζουν νέα πρωτόκολλα δρομολόγησης, μοντέλα ασφάλειας, σχήματα διευθύνσεων ακόμη και εναλλακτικές λύσεις αντί της IP[34].

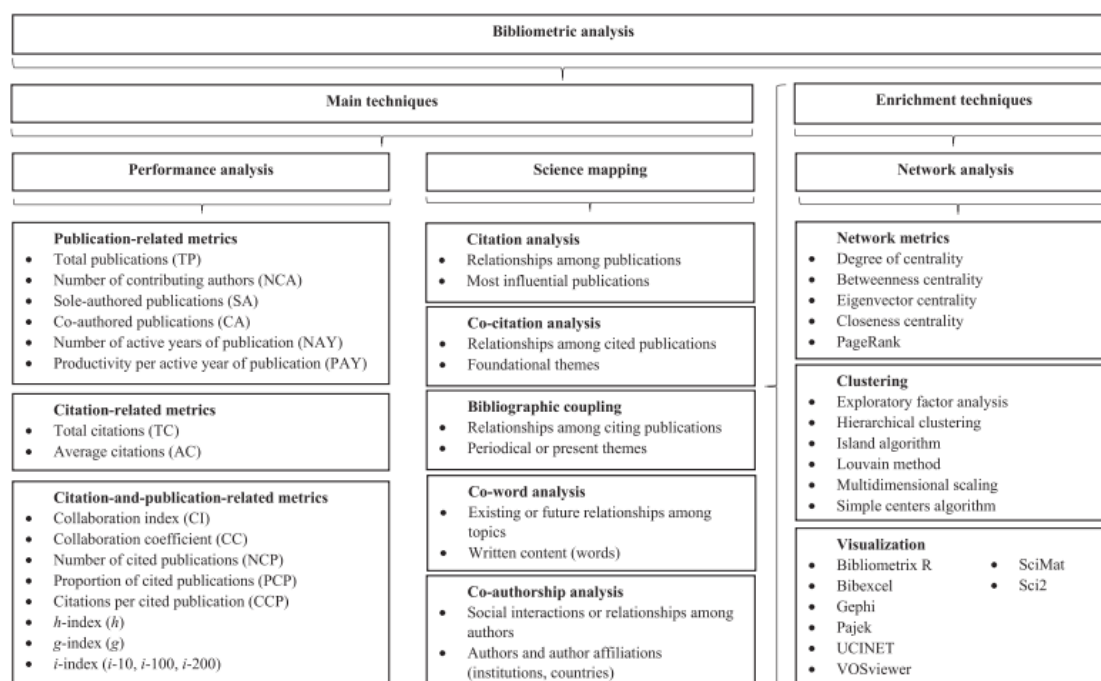
Η χρήση του Openflow στον κόσμο των δικτύων επιφέρει αρκετά πλεονεκτήματα, ορισμένα αναφέρονται στη συνέχεια:

- Οι διαδρομές προσαρμόζονται, η κυκλοφορία στο δίκτυο εξισορροπείται και εφαρμόζονται πολιτικές όταν είναι αναγκαίο. Αυτό συμβαίνει λόγω του λεπτομερούς ελέγχου που έχει ο ελεγκτής στην κυκλοφορία του δικτύου.
- Πιο εύκολη διαχείριση μεγάλων και πολύπλοκων δικτύων.
- Πιο εύκολη αναβάθμιση του δικτύου.
- Συνεχής εξέλιξη και κατά συνέπεια δυνατότητα βελτίωσης.
- Ευκολότερο στήσιμο δικτύου.
- Ευκολότερη κλωνοποίηση.

3. Βιβλιομετρία ως μέθοδος ανάλυσης

3.1 Βιβλιομετρία

Η μελέτη δημοσιεύσεων, άρθρων και η ανάλυση βιβλίων στον επιστημονικό χώρο είναι μία προσέγγιση για την περιγραφή της βιβλιομετρίας. Η βιβλιομετρία αντικατέστησε την στατική βιβλιογραφία στις αρχές του 19^{ου} αιώνα με τον Βέλγο συγγραφέα Paul Otlet να την αναφέρει ως «η μέτρηση όλων των πτυχών που σχετίζονται με την δημοσίευση και την ανάγνωση βιβλίων και εγγράφων» [8].



Εικόνα 3.1 Ταξινόμηση των τεχνικών βιβλιομετρικής ανάλυσης [23].

Σπουδαίο ρόλο στην βιβλιομετρική κοινότητα έχουν οι 3 κατηγορίες που φαίνονται στην Εικόνα 3.1. Η ανάλυση της απόδοσης, η επιστημονική χαρτογράφηση και η ανάλυση δικτύου είναι τεχνικές που βασίζονται οι βιβλιομετρικές έρευνες [24].

3.1.1 1η κατηγορία: Ανάλυση απόδοσης (Performance analysis)

Η ανάλυση απόδοσης εξετάζει τις συνεισφορές που υπάρχουν από συγγραφείς, χώρες, ιδρύματα σε ένα πεδίο μελέτης. Επικεντρώνεται σε μετρικές, όπως θα αναλυθούν στη συνέχεια, με στόχο να αξιολογηθεί η παραγωγικότητα και η επιρροή [23].

Η ανάλυση απόδοσης χωρίζεται σε τρία μέρη ανάλογα με το τι αφορούν οι μετρήσεις.

1) Μετρικές που σχετίζονται με την δημοσίευση

- **Συνολικές Δημοσιεύσεις (Total Publications):** ο δείκτης αυτός αναφέρεται στο πλήθος των δημοσιεύσεων που έχει αναφερθεί για παράδειγμα μία έρευνα μέσα από παραπομπές.

- **Αριθμός συνεισφερόντων συγγραφέων (Number of Contributing Authors):** ο NCA δείκτης αναφέρεται στο πλήθος των ατόμων που έχουν συνεργαστεί για να ολοκληρώσουν μια μελέτη.
- **Δημοσιεύσεις με μοναδικό συγγραφέα (Sole-Authored Publications):** η μετρική S αναφέρεται στον αριθμό των δημοσιεύσεων που έχει κάθε συγγραφέας, χωρίς όμως να προσμετρά τις συν-συγγραφές.
- **Δημοσιεύσεις συνεργασίας (Co-authored Publications):** παρόμοιος δείκτης με τον SA, υπολογίζει τις δημοσιεύσεις που ένας συγγραφέας έχει σε συνεργασία με άλλους.
- **Αριθμός των ενεργών ετών δημοσίευσης (Number of Active Years of publication):** αναφέρεται στα έτη όπου υπάρχουν δημοσιεύσεις από έναν συγγραφέα ή από συνεργασίες.
- **Παραγωγικότητα ανά ενεργό έτος δημοσιεύσεων (Productivity per Active Year of publications):** μετρά την παραγωγικότητα του συγγραφέα ανάλογα με τις δημοσιεύσεις που γίνονται μέσα στον χρόνο. Ο υπολογισμός του δείκτη γίνεται σύμφωνα με τον τύπο:

$$PAY = \frac{\text{Total Publications}}{\text{Number of Active Years of publications}}$$

2) Μετρικές που σχετίζονται με την αναφορά

- **Συνολικές Αναφορές (Total Citations):** ο δείκτης TC δείχνει τον αριθμό αναφορών που είχε μία δημοσίευση μέσα σε άλλες εργασίες.
- **Μέσος όρος αναφορών (Average Citations):** αφορά την μέση τιμή της αναφοράς που έχει λάβει μία δημοσίευση από άλλες. Υπολογίζεται σύμφωνα με τον τύπο:

$$AC = \frac{\text{Total Citations}}{\text{Number of Publications}}$$

3) Μετρικές που σχετίζονται με την αναφορά και την δημοσίευση

- **Δείκτης Συνεργασίας (Collaboration Index):** Αφορά την συνεργασία μεταξύ των ερευνητικών ομάδων. Ο δείκτης CI σούεται με το σύνολο των συν-συγγραφέων προς το σύνολο των δημοσιεύσεων.

$$CI = \frac{\text{Number of Collaborators}}{\text{Number of Contributing Authors}}$$

- **Συντελεστής Συνεργασίας (Collaboration Coefficient):** είναι το ποσοστό κοινής συγγραφής που έχει ένας ερευνητής (ή μία ομάδα).

$$CC = \frac{\text{Number of Collaborative Publications}}{\text{Total Number of Publications}}$$

- **Αριθμός αναφερόμενων δημοσιεύσεων (Number of Cited Publications):** αναφέρεται στο πλήθος των φορών που έχει αναφερθεί ένα έργο ή ένας συγγραφέας σε άλλες δημοσιεύσεις. Όσο μεγαλύτερη είναι η τιμή του δείκτη συνεπάγεται ότι έχει και μεγάλη επιρροή στην επιστημονική κοινότητα.

- **Αναλογία αναφερόμενων δημοσιεύσεων (Proportion of Cited Publications):** ο δείκτης αυτός μετρά το ποσοστό των δημοσιεύσεων που έχουν λάβει παραπομπές σε σχέση με τον συνολικό αριθμό των δημοσιεύσεων.

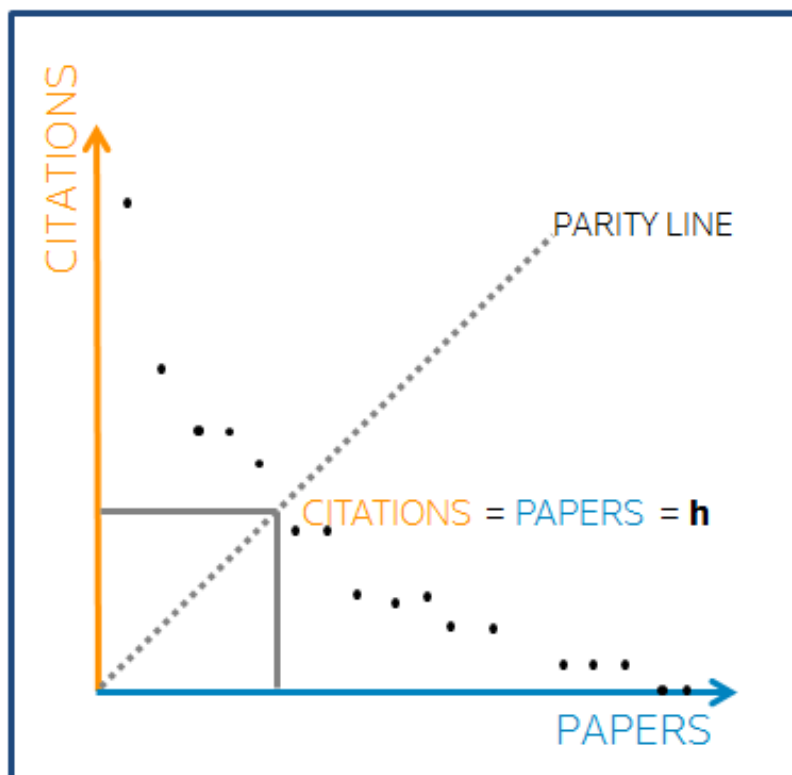
$$PCP = \left(\frac{\text{Number of Cited Publications}}{\text{Total Publications}} \right) 100$$

- **Αναφορές ανά αναφερόμενες δημοσιεύσεις (Citations per Cited Publications):** μετρά τον μέσο όρο των αναφορών που έχουν λάβει οι δημοσιεύσεις.

$$CCP = \frac{\text{Total Citations}}{\text{Total Cited Publications}}$$

- **I-index:** είναι δείκτης που προσδιορίζει τον αριθμό των δημοσιεύσεων που έχουν τουλάχιστον *i* αναφορά. Η τιμή του μπορεί να είναι 10, 100 ή 200
- **G-index:** παρόμοιος δείκτης με τον H-index. Μετρά την σημαντικότητα και την επιρροή των δημοσιεύσεων με βάση τον αριθμό των αναφορών τους.
- **H-index:** Ο δείκτης αυτός μετρά την σημαντικότητα και την επιρροή των παραπομπών ενός συγγραφέα από άλλους συγγραφείς ή ομάδες συγγραφέων. Ο τρόπος υπολογισμού του είναι εύκολος και δίνεται από την συνάρτηση

$$h - index(f) = \max\{i \in \mathbb{N} : f(i) \geq i\}$$



Εικόνα 3.2 H-index [9]

Οι δημοσιεύσεις πρέπει να ταξινομηθούν με φθίνουσα σειρά, έπειτα βρίσκουμε τον αριθμό των δημοσιεύσεων (*i*) που έχουν αναφορές τουλάχιστον ίσες με *i*. Τότε ο δείκτης έχει την τιμή με το μεγαλύτερο *i*. Έστω για παράδειγμα ένας συγγραφέας

έχει τις δημοσιεύσεις (ABCDE) και οι αναφορές σε αυτές τις δημοσιεύσεις είναι: $f(A)=10$, $f(B)=8$, $f(C)=5$, $f(D)=4$, $f(E)=3$. Ο αριθμός του δείκτη θα έχει την τιμή 4 διότι η 4^η δημοσίευση έχει 4 παραπομπές και η 5^η δημοσίευση έχει 3 παραπομπές. Με άλλα λόγια αυτό σημαίνει ότι υπάρχουν 4 δημοσιεύσεις με τουλάχιστον 4 αναφορές.

- **Παράγοντας επίδρασης (Impact Factor):** Ο παράγοντας επίδρασης μετρά την επιρροή των περιοδικών σε ένα συγκεκριμένο εύρος χρόνου που συνήθως είναι το ένα έτος. Ο υπολογισμός του γίνεται από το Ινστιτούτο Επιστημονικής Πληροφορίας (Institute for Scientific Information - ISI) και δίνεται από την ακόλουθη συνάρτηση [12]:

$$IF(y) = \frac{Citations(y)}{Publications(y-1) + Publications(y-2)}$$

Αυτό υπολογίζεται ως τον αριθμό των αναφορών που έλαβε ένα περιοδικό σε ένα συγκεκριμένο έτος, προς το άθροισμα των άρθρων που δημοσιεύθηκαν σε αυτό το περιοδικό σε διάστημα δύο ετών πριν.

3.1.2 2η κατηγορία: Επιστημονική χαρτογράφηση (Science mapping)

Η επιστημονική χαρτογράφηση βοηθά τον ερευνητή να κατανοήσει την αλληλεπίδραση και την σύνδεση που υπάρχει από διαφορετικές μελέτες ή έννοιες σε ένα συγκεκριμένο πεδίο. Περιλαμβάνει πληροφορίες, όπως η εξέλιξη που υπάρχει σε ένα ερευνητικό πεδίο, με αποτέλεσμα την ανάδειξη της απήχησης που υπάρχει στην επιστημονική κοινότητα[23].

Περιλαμβάνει τεχνικές που χρησιμοποιούνται για την ανάλυση και απεικόνιση των σχέσεων στο κομμάτι της βιβλιομετρίας. Στην συνέχεια αναφέρονται κάποιες από αυτές.

- **Ανάλυση παραπομπών (Citation Analysis):** εξετάζει πόσες φορές ένα έργο αναφέρεται σε άλλα. Ουσιαστικά αυτή η τεχνική μελετά τις σχέσεις των έργων (πχ άρθρων, βιβλίων, εργασιών, κ.α.) μέσω των παραπομπών.
- **Ανάλυση συμπαραπομπών (Co-citation Analysis):** αναζητά πως διάφορα έργα αναφέρονται μαζί σε άλλα. Εστιάζει στις σχέσεις που προκύπτουν από παραπομπές και από κοινές θεματολογίες.
- **Βιβλιογραφική σύζευξη (Bibliographic Coupling):** εξηγεί τις σχέσεις μεταξύ των παραπομπών και των ερευνητικών έργων με στόχο να αναδειχθούν θεματικές και μεθοδολογικές σχέσεις.
- **Ανάλυση συν-λέξης ή κοινής λέξης (Co-word Analysis):** αναλύει τις συχνότητες εμφάνισης λέξεων με στόχο την ανίχνευση θεμάτων με βάση τις λέξεις που χρησιμοποιούνται.
- **Ανάλυση συν-συγγραφής (Co-authorship Analysis):** αναλύει τις σχέσεις συνεργασίας των συγγραφέων με στόχο την ανακάλυψη δικτύων κοινής συγγραφής και ερευνητικής συνεισφοράς.

3.1.3 3η κατηγορία: Ανάλυση δικτύου (Network analysis)

Η ανάλυση δικτύου έχει τη δυνατότητα να αναδείξει ερευνητικά θέματα και τρέχουσες τάσεις, καθώς και επιστημονικά άρθρα που συνδέονται μεταξύ τους. Αυτά τα άρθρα μπορούν να χωριστούν σε δύο κατηγορίες:

- 1) τα άρθρα που ασκούν επίδραση (influential articles), τα οποία επηρεάζουν και προκαλούν αλλαγές στην έρευνα.
- 2) τα κορυφαία άρθρα (prestigious articles), τα οποία έχουν αναγνωρισθεί ως εξαιρετικά σημαντικά λόγω των υψηλών παραπομπών ή των βραβείων που έχουν λάβει[33].

Σε αυτή την κατηγορία περιλαμβάνονται οι δείκτες κεντρικότητας που θα αναφερθούν στο 5^ο κεφάλαιο της εργασίας και συγκεκριμένα μέσα από το λογισμικό Gephi. Οι δείκτες κεντρικότητας μετρούν την σημαντικότητα του κόμβου από διαφορετικές πτυχές. Στη συνέχεια γίνεται μία αναφορά στους δείκτες αυτούς.

- **Κεντρικότητα βαθμού (Degree Centrality):** ο δείκτης αυτός μετρά τις συνδέσεις των κόμβων. Όσο περισσότερες συνδέσεις έχει ο κόμβος τόσο πιο υψηλή θα είναι η τιμή της κεντρικότητας βαθμού. Υπολογίζεται ως :

$$C(x) = \frac{\text{Number of connection of node } x}{N - 1}$$

Όπου N: ο συνολικός αριθμός κόμβων.

- **Κεντρικότητα Εγγύτητας (Closeness Centrality):** η τιμή υπολογίζεται σύμφωνα με την απόσταση των κόμβων. Δηλαδή οι κόμβοι που είναι πιο κοντά έχουν υψηλότερη κεντρικότητα από τους άλλους. Ο υπολογισμός της γίνεται σύμφωνα με την εξίσωση:

$$C(x) = \frac{N - 1}{\sum_{x=1}^N d(x, y)}$$

Όπου

N: ο συνολικός αριθμός κόμβων

D(x,y): η απόσταση μεταξύ κόμβου x και y.

- **Κεντρικότητα μεταξύ 2 (Betweenness Centrality):** αφορά τον κόμβο που βρίσκεται στο ενδιάμεσο του συντομότερου μονοπατιού μεταξύ άλλων δύο κόμβων. Οι κόμβοι που εμφανίζουν υψηλή «κεντρικότητα ενδιάμεσων» είναι σημαντικοί για την ροή πληροφοριών στο δίκτυο. Ο υπολογισμός της γίνεται σύμφωνα της εξίσωσης που ακολουθεί:

$$C(X) = \sum \frac{\sigma_{st}(x)}{\sigma_{st}}$$

Όπου:

σ_{st} : αριθμός σύντομων μονοπατιών μεταξύ κόμβων s και t

$\sigma_{st}(x)$: αριθμός σύντομων μονοπατιών μεταξύ κόμβων s και t που διέρχονται από τον κόμβο x .

- **Κεντρικότητα ιδιοδιανύσματος (Eigenvector Centrality)**: λαμβάνει υπόψη την σημαντικότητα των γειτονικών κόμβων. Ένας κόμβος με υψηλό δείκτη «κεντρικότητας κατά ιεραρχία» είναι συνδεδεμένος με κόμβους υψηλής σημαντικότητας. Ο δείκτης υπολογίζεται ως εξής :

$$C(u) = \frac{1}{\lambda} \sum_{u \in N(u)} x(u)$$

Όπου:

$N(u)$: είναι η γειτνίαση του κόμβου u

λ : η μεγαλύτερη ιδιοτιμή της adjacency matrix A

- **PageRank**: ο αλγόριθμος PageRank μετρά την σημαντικότητα των κόμβων στο δίκτυο. Όταν η τιμή του είναι μεγάλη, σημαίνει ότι έχει πολλές και σημαντικές συνδέσεις.

3.2 Τεχνικές ανάλυσης

Στην βιβλιομετρία υπάρχουν δύο κατηγορίες τεχνικών μελέτης, οι σχεσιακές και οι αναλυτικές. Οι σχεσιακές εστιάζουν στην βαρύτητα των σχέσεων ανάμεσα σε έργα, συγγραφείς, παραπομπές και τις αλληλεπιδράσεις που υπάρχουν μεταξύ αυτών. Η επιλογή της κατάλληλης τεχνικής εξαρτάται από τους στόχους και τα ερευνητικά ερωτήματα που έχουν τεθεί. Τέτοιες σχέσεις είναι [23]:

- **Ανάλυση παραπομπών**: σε αυτή την ανάλυση αντικατοπτρίζονται δεσμοί μεταξύ των δημοσιεύσεων που δημιουργούνται όταν μία δημοσίευση παραθέτει κάποια άλλη. Ο αντίκτυπος της δημοσίευσης καθορίζεται από τον αριθμό των παραπομπών που έχει λάβει. Όσο μεγαλύτερος είναι ο αριθμός των παραπομπών, τόσο μεγαλύτερη είναι και η επιρροή της δημοσίευσης.
- **Ανάλυση συν-παραπομπών**: αναφέρεται σε ζεύγη άρθρων τα οποία αναφέρονται μαζί. Αυτό δημιουργεί ένα δίκτυο συν-παραπομπών το οποίο μπορεί να χρησιμοποιηθεί για να αναδείξει σχέσεις σημασιολογικής συνάφειας.
- **Βιβλιογραφική σύζευξη**: βασίζεται στο κοινό σημείο αναφοράς μεταξύ δύο δημοσιεύσεων που λειτουργεί προϋποθέτει ότι αυτές (δημοσιεύσεις) είναι παρόμοιες ως προς το περιεχόμενό τους. Οι δημοσιεύσεις χωρίζονται σε θεματικά συμπλέγματα βάσει των κοινών αναφορών, παρέχοντας μια αναπαράσταση του ερευνητικού πεδίου. Η βιβλιογραφική σύζευξη είναι χρήσιμη για την αποκάλυψη των τελευταίων εξελίξεων στην επιστημονική έρευνα.
- **Ανάλυση συν-συγγραφέων**: εστιάζει στα πρότυπα συνεργασίας μεταξύ των ερευνητών, εντοπίζοντας έτσι βασικούς συγγραφείς, δίκτυα και συνεργασίες.
- **Χαρτογράφηση περιοδικών**: ψάχνει πρότυπα δημοσίευσης σε συγκεκριμένα περιοδικά, αποκαλύπτοντας τις τάσεις και τα ενδιαφέροντα της έρευνας.
- **Ανάλυση λέξεων-κλειδιών**: επικεντρώνεται στην εμφάνιση και την χρήση συγκεκριμένων λέξεων-κλειδιών για τον εντοπισμό σημαντικών ερευνητικών θεμάτων.

- **Ανάλυση κοινωνικών δικτύων:** αφορά τις κοινωνικές σχέσεις και τις αλληλεπιδράσεις τους μεταξύ ερευνητών ή ιδρυμάτων.

Οι αναλυτικές τεχνικές επικεντρώνονται στις συνδέσεις που υπάρχουν μεταξύ διαφορετικών στοιχείων. Τέτοιες σχέσεις είναι:

- **Ανάλυση δικτύου (Network Analysis):** Αυτή η μέθοδος χρησιμοποιείται για να προσδιορίσει την δομή και την σημασιολογική σχέση μεταξύ συγγραφέων, ιδρυμάτων και χωρών μιας επιστημονικής έρευνας. Χρησιμοποιεί μετρικές δικτύου (κεντρικότητες, PageRank) για να αξιολογήσει την επίδραση και να παρέχει πληροφορίες σχετικές με την εξέλιξη της έρευνας.
- **Συσταδοποίηση (Cluster Analysis):** η συσταδοποίηση ομαδοποιεί σύνολα δεδομένων με κύριο κριτήριο τα κοινά χαρακτηριστικά, αυτά μπορεί να είναι η θεματική ενότητα, οι συγγραφείς, το αντικείμενο μελέτης.
- **Ανάλυση πολυπλοκότητας (Complexity Analysis):** μελετά την δομή, της ιδιότητες και την δυναμική του δικτύου. Στην βιβλιομετρία γ συγκεκριμένη πολυπλοκότητα αφορά τις σχέσεις των συγγραφέων και την κατηγοριοποίηση των άρθρων.
- **Ανάλυση κύριων στοιχείων (PCA- Principal Component Analysis):** η ανάλυση κύριου στοιχείου είναι μία μέθοδος που χρησιμοποιείται για μεγάλα σύνολα δεδομένων με στόχο να μειώσει την διάστασή τους κρατώντας τις κύριες πληροφορίες που επηρεάζουν την έρευνα.
- **Παραγοντική ανάλυση (Factor Analysis):** η παραγοντική ανάλυση μειώνει το σύνολο των μεταβλητών που αφορούν τις δημοσιεύσεις εξάγοντας τα κοινά σημεία. Τα κοινά πρότυπα που εξάγει ονομάζονται παράγοντες ή διαστάσεις.
- **Ανάλυση συμπαραπομπών (Co-Citation Analysis):** η ανάλυση συμπαραπομπών περιλαμβάνει την παρακολούθηση ζευγών άρθρων που παραθέτονται μαζί. Όταν αυτά αναφέρονται από πολλούς συγγραφείς αρχίζουν και σχηματίζουν ομάδες έρευνας.

3.2.1 Μέθοδος Layer-Ταξινόμηση σε επίπεδα

Η μέθοδος ταξινόμησης Layer επικεντρώνεται στην ανάλυση των προτύπων παραπομπών, δηλαδή την συχνότητα και τον τρόπο με τον οποίο τα επιστημονικά άρθρα παραθέτουν άλλες δημοσιεύσεις σε ένα ερευνητικό πεδίο. Με αυτόν τον τρόπο τα έγγραφα κατηγοριοποιούνται σε διαφορετικά επίπεδα με βάση τον αριθμό αναφορών τους. Σε αυτή την μέθοδο περιλαμβάνονται τα εξής βήματα:

- 1) Συλλογή δεδομένων: από τις βάσεις δεδομένων που αναφέρθηκαν σε προηγούμενο κεφάλαιο συλλέγονται οι πληροφορίες (για βιβλία, αναφορές, κλπ.) για την μελέτη ενός συγκεκριμένου πεδίου.
- 2) Υπολογισμός αριθμού αναφορών: κάθε έγγραφο συγκεντρώνει έναν αριθμό το οποίο βασίζεται στο πόσες φορές έχει παρατεθεί μέσα σε άλλα έγγραφα (π.χ. εργασίες, δημοσιεύσεις κλπ.)
- 3) Ταξινόμηση επιπέδων: σύμφωνα με τον αριθμό των αναφορών ταξινομούνται σε διαφορετικά επίπεδα. Συνήθως ταξινομούνται στην κορυφή αυτά με τον

μεγαλύτερο αριθμό αναφορών και στην βάση με τον μικρότερο αριθμό. Τα κριτήρια βέβαια μπορεί να διαφέρουν ανάλογα με την μελέτη.

- 4) Ανάλυση: αφού τα έγγραφα ταξινομηθούν, αναλύονται και υπολογίζονται διάφοροι βιβλιομετρικοί δείκτες και μετρήσεις για κάθε επίπεδο. Τέτοιοι δείκτες είναι ο μέσος όρος αναφορών και ο δείκτης h. Η ανάλυση εφαρμόζεται για να γίνει κατανοητή η επιρροή και η σημασία του κάθε εγγράφου σε κάθε ερευνητικό πεδίο.

Η τεχνική ανάλυσης Layer επιτρέπει στους ερευνητές να εντοπίσουν εργασίες με μεγάλη ή μικρή επιρροή, λαμβάνοντας σημαντικό ή μικρότερο αριθμό αναφορών. Παρέχει έναν τρόπο αξιολόγησης του αντίκτυπου και της προόδου της έρευνας σε ένα πεδίο, εξετάζοντας μοτίβα αναφορών σε διαφορετικά επίπεδα.

3.2.2 Μέθοδος Clustering - Ομαδοποίηση

Η ανάλυση με την μέθοδο της ομαδοποίησης ή διαφορετικά η ανάλυση συστάδων όπως ονομάζεται, χρησιμοποιείται για την ταξινόμηση δεδομένων με γνώμονα τις ομοιότητες ή τις ανομοιότητες τους. Σε μία βιβλιομετρική εργασία η τεχνική αυτή χρησιμοποιείται για τον εντοπισμό σχέσεων μεταξύ δημοσιεύσεων, συγγραφέων, ερευνητικών θεμάτων με κοινά χαρακτηριστικά. Η μέθοδος ανάλυσης της ομαδοποίησης χρησιμοποιεί διάφορους αλγόριθμους και τεχνικές ομαδοποίησης όπως οι ακόλουθες:

- 1) K-means: η ομαδοποίηση k μέσου διαχωρίζει τα δεδομένα σε έναν καθορισμένο αριθμό συστάδων όπου κάθε σημείο δεδομένων εκχωρείται στο σύμπλεγμα με τον πλησιέστερο μέσο όρο ή κέντρο.
- 2) Ιεραρχική ομαδοποίηση: δημιουργεί μία ιεραρχική δομή συστάδων με επαναληπτική συγχώρηση ή διαχωρισμό συστάδων με βάση των ομοιότητα ή την ανομοιότητά τους.
- 3) Ομαδοποίηση με βάση την πυκνότητα: διάφοροι αλγόριθμοι προσδιορίζουν συστάδες με βάση περιοχές υψηλής πυκνότητας δεδομένων. Ένας τέτοιος αλγόριθμος είναι ο DBSCAN.
- 4) Ομαδοποίηση βάση μοντέλου: σε αυτή την τεχνική τα δεδομένα παράγονται από ένα μείγμα κατανομών πιθανοτήτων και επιδιώκει να προσδιορίσει τον βέλτιστο αριθμό και τα χαρακτηριστικά των συστάδων.

3.3 Πηγές δεδομένων και εργαλεία λογισμικού

Στα πλαίσια μιας βιβλιομετρικής ανάλυσης ενός θέματος χρησιμοποιείται μια σειρά από προγράμματα ή εφαρμογές υπολογιστών που έχουν σχεδιαστεί ειδικά για να διευκολύνουν την μελέτη και να εξάγουν αποτελέσματα από την επεξεργασία των δεδομένων που εξάγουν από επιστημονικές βάσεις δεδομένων. Τα εργαλεία αυτά διακρίνονται στις επιστημονικές βάσεις δεδομένων και στα προγράμματα λογισμικού και μπορούν να ταξινομηθούν σε 4 κατηγορίες ανάλογα με το τι εξετάζουν. Οι κατηγορίες είναι οι εξής:

- **Πρώτη κατηγορία: Συλλογή δεδομένων ή αλλιώς Βάσεις δεδομένων**

Οι πιο γνωστές πηγές δεδομένων είναι οι ακόλουθες επιστημονικές βάσεις:

1. Web of Science: είναι μία από τις πιο δημοφιλείς βάσεις δεδομένων για βιβλιομετρική ανάλυση. Περιέχει επιστημονικές δημοσιεύσεις που μπορούν να εισαχθούν στο λογισμικό σε μορφή BibTeX ή RIS.
2. Scopus: έχει στην κατοχή του πάνω από 77 εκατομμύρια εγγραφές μέσα στις οποίες βρίσκονται ακαδημαϊκές και επιστημονικές δημοσιεύσεις, περιοδικά, πρακτικά συνεδρίων κα.
3. PubMed: είναι μία βιοιατρική βάση δεδομένων όπου τα δεδομένα που εισάγονται στο λογισμικό είναι σε μορφή MEDLINE.
4. Microsoft Academic: είναι μία ακαδημαϊκή βάση δεδομένων που παρέχει πρόσβαση σε επιστημονικές δημοσιεύσεις, συγγραφείς και ανάλυση αναφορών σε διάφορους κλάδους. Τα δεδομένα που εξάγονται είναι σε μορφή JASON.

- **Δεύτερη κατηγορία: Ανάλυση παραπομπών**

Σε αυτή την κατηγορία εντάσσονται λογισμικά κατάλληλα για την ανάλυση παραπομπών που προσδιορίζουν πρότυπα και σχέσεις μεταξύ δημοσιεύσεων, συγγραφέων και περιοδικών. Τέτοια εργαλεία είναι το BibExcel, CiteSpace, HistCite.

- **Τρίτη κατηγορία: Ανάλυση δικτύων:**

Λογισμικά κατάλληλα για την ανάλυση σχέσεων μεταξύ βιβλιογραφικών δεδομένων και την οπτικοποίηση συν-παραπομπών, συν-συγγραφέων και άλλους τύπους βιβλιομετρικών δικτύων. Τα λογισμικά αυτά περιλαμβάνουν αλγορίθμους όπου μετρούν την κεντρικότητα ή την ομαδοποίηση για να βοηθήσουν τους χρήστες να αποκτήσουν πληροφορίες σχετικά με την δομή των δεδομένων. Σε αυτή την κατηγορία ανήκουν προγράμματα όπως το Gephi και το Pajek.

- **Τέταρτη κατηγορία: Οπτικοποίηση δεδομένων:**

Είναι παρόμοια κατηγορία με την προηγούμενη. Εδώ ανήκουν λογισμικά που επιτρέπουν στους χρήστες να οπτικοποιούν τα δεδομένα και να φτιάχνουν χάρτες παραπομπών, λέξεων-κλειδιών και συν-συγγραφέων. Τέτοια προγράμματα είναι το VOSViewer. Διαφέρει από την προηγούμενη κατηγορία στο ότι τα λογισμικά οπτικοποίησης δεδομένων χρησιμοποιούνται για την παρουσίαση των αποτελεσμάτων των βιβλιομετρικών αναλύσεων, ενώ τα λογισμικά ανάλυσης δικτύων εντοπίζουν και αναλύουν πρότυπα και τάσεις μεταξύ οντοτήτων στα δεδομένα.

3.3.1 VOSViewer



Εικόνα 3.3 VOSViewer

Το ελεύθερο λογισμικό VOSViewer (Visualization Of Similarities) χρησιμοποιείται για βιβλιομετρική ανάλυση και οπτικοποίηση επιστημονικών δημοσιεύσεων. Δημιουργήθηκε το 2010 από τους Nees van Ect και Ludo Waltman στο πανεπιστήμιο του Leiden. Είναι μία μέθοδος μελέτης που με την πάροδο των χρόνων γίνεται όλο και πιο δημοφιλής διότι η χρήση του δίνει στους ερευνητές την δυνατότητα να αναλύσουν και να αναπαραστήσουν έναν μεγάλο όγκο στοιχείων ο οποίος είναι αποθηκευμένος σε βάσεις δεδομένων. Αυτός ο όγκος δεδομένων πιθανολογείτε ότι φτάνει τα 5000 έγγραφα. Το VOS είναι χαρισματικό εργαλείο κατάλληλο για την δημιουργία βιβλιογραφικών χαρτών που οπτικοποιούν τις σχέσεις μεταξύ διαφορετικών ερευνητικών περιοχών, συγγραφέων και δημοσιεύσεων με βάση τις λέξεις κλειδιά που ορίζει ο ερευνητής.

Έχει την ικανότητα να αναλύει και να οπτικοποιεί μεγάλης κλίμακας δικτύων συν-συγγραφέων, παραπομπών και συν-εμφάνισης όρων. Με την χρήση αλγορίθμων εντοπίζει συνδέσεις προσφέροντας πολύτιμες πληροφορίες.

Αναγνωρίζει και αναπαριστά οπτικά συμπλέγματα ή ομάδες συνδεδεμένων κόμβων μέσα σε ένα βιβλιομετρικό δίκτυο και το κάνει να ξεχωρίζει από άλλα λογισμικά στην κατηγορία του. Αυτή η ικανότητα επιτρέπει στους ερευνητές να διακρίνουν θεματικές περιοχές ή ερευνητικά πεδία κάνοντας πιο εύκολη την κατανόηση της δομής σε ένα συγκεκριμένο πεδίο. Το λογισμικό παρέχει διάφορους αλγορίθμους ομαδοποίησης και μέτρα ομοιότητας που μπορούν να προσαρμοστούν θέτοντας συγκεκριμένους στόχους.

Η χρήση του, δίνει στους ερευνητές την δυνατότητα από μια σειρά επιλογών οπτικοποίησης. Οι χρήστες δηλαδή μπορούν να εξατομικεύσουν την διάταξη, τα σχήματα των κόμβων και τα χρώματα για να τονίσουν τις σχετικές πληροφορίες και τα βασικά ευρήματα. Λόγω της διαδραστικής εξερεύνησης οι ερευνητές μπορούν να πραγματοποιήσουν ενέργειες όπως να αυξομειώσουν το μέγεθος σε συγκεκριμένο τμήμα του χάρτη, να εστιάσουν ή να μετακινήσουν συγκεκριμένα στοιχεία του δικτύου, να εφαρμόσουν φίλτρα. Αυτή η άμεση αλληλεπίδραση που υπάρχει με την οθόνη βοηθάει τους χρήστες του να μελετάνε πιο εύκολα τα δεδομένα τους [14-18].

3.3.2 Gephi



Εικόνα 3.4Gephi

Το Gephi είναι ένα λογισμικό ανοιχτού κώδικα και έχει σχεδιαστεί από τους Mathien Bastian, Sebastien Heymann και Mathien Jacomy. Η πρώτη του έκδοση κυκλοφόρησε τον Ιούλιο του 2008 με στόχο την εξερεύνηση και τον χειρισμό δικτύων. Είναι γραμμένο σε γλώσσα Java που αυτομάτως το κάνει φιλικό ως προς τον κάθε χρήστη καθώς τον διευκολύνει στην οπτικοποίηση, στην ανάλυση και τον χειρισμό πολύπλοκων δεδομένων δικτύου. Η εισαγωγή των δεδομένων στο λογισμικό είναι εύκολη και γίνεται μέσα από διαφορετικές μορφές αρχείων, όπως csv, excel και άλλα, επιτρέποντας τον διαδραστικό χειρισμό του δικτύου στον τομέα της οπτικοποίησης.

Ο χρήστης αφού εισάγει τα δεδομένα μέσα στο Gephi, παίρνει έναν οπτικό χάρτη με κόμβους και ακμές. Αυτό του επιτρέπει να κάνει εις βάθος αναλύσεις και να αναδειξεί πρότυπα, κοινότητες δικτύων, να υπολογίσει μετρήσεις όπως η κεντρικότητα βαθμών ή η ενδιάμεση κεντρικότητα χάρη των στατιστικών εργαλείων που είναι ενσωματωμένα σε αυτό. Τα εργαλεία αυτά εντοπίζουν την δυναμική του δικτύου, τις σημαντικές σχέσεις, τους βασικούς κόμβους και τα συμπλέγματα επιρροής.

Τέλος, το Gephi δίνει την δυνατότητα στον χρήστη να τροποποιήσει την δομή του δικτύου με την προσθήκη ή την αφαίρεση κόμβων ή ακμών, να αλλάξει τα χρώματα και γενικά την εμφάνιση του οπτικού χάρτη και να εξάγει τα δεδομένα σε διαφορετικές μορφές για περαιτέρω ανάλυση.[19-20]

3.3.3 R Studio

Το RStudio είναι ένα ολοκληρωμένο περιβάλλον ανάπτυξης (IDE) για την γλώσσα προγραμματισμού R. Σχεδιάστηκε με στόχο η ανάλυση των δεδομένων να γίνεται πιο εύκολα και πιο αποτελεσματικά από τους αναλυτές R. Studio παρέχει τη δυνατότητα στον χρήστη να αποθηκεύσει και να επαναχρησιμοποιήσει τα script που έχει φτιάξει, να αναζητά βοήθεια σε οτιδήποτε χρειαστεί, να εγκαθιστά και να διαχειρίζεται με ευκολία τις βιβλιοθήκες. Το R παρέχει επίσης στον χρήστη την προεπισκόπηση των γραφημάτων, την εύκολη εναλλαγή μεταξύ κονσόλας και τερματικού, την αυτόματη συμπλήρωση κώδικα και πολλά ακόμη. Είναι διαθέσιμο σε δύο μορφές [25-26]:

- Το RStudio Desktopόπου εκτελείται τοπικά στον υπολογιστή ως κανονική εφαρμογή και
- Το RStudio Server το οποίο επιτρέπει την πρόσβαση στο R, χρησιμοποιώντας ένα πρόγραμμα περιήγησης ιστού ενώ εκτελείται σε έναν απομακρυσμένο διακομιστή Linux.

3.4 Ανάλυση SDN

Η ανάλυση της δικτύωσης που καθορίζεται από το λογισμικό (SDN) είναι μία πολυδιάστατη εξέταση βασισμένη στις προόδους και τις επιπτώσεις της τεχνολογίας SDN. Ουσιαστικά αναγνωρίζονται τα πιθανά οφέλη που επιφέρει αυτή η τεχνολογία, παρέχοντας μία ολοκληρωμένη κατανόηση των χαρακτηριστικών, των προκλήσεων και των πλεονεκτημάτων.

Στο επίκεντρο βρίσκεται ο κεντρικός έλεγχος που προσφέρει το SDN. Σε αντίθεση με τα παραδοσιακά δίκτυα, κάνοντας χρήση ενός κεντρικού ελεγκτή, τα δίκτυα SDN διαθέτουν μία συνολική προβολή του δικτύου. Αυτό τους επιτρέπει την καλύτερη διαχείριση, διαμόρφωση και έλεγχο του δικτύου.

Ο προγραμματισμός είναι βασικό χαρακτηριστικών των SDN. Λόγω των διεπαφών προγραμματιζόμενων εφαρμογών (Application Programming Interface-API), οι διαχειριστές ή οι εφαρμογές δικτύου μπορούν να ελέγχουν και να διαμορφώνουν δυναμικά το δίκτυο. Αυτή η δυνατότητα προγραμματισμού επιτρέπει την αυτοματοποίηση εργασιών διαχείρισης δικτύου, την προσαρμογή της συμπεριφοράς του δικτύου για την κάλυψη συγκεκριμένων απαιτήσεων.

Η τεχνολογία SDN επιτρέπει την εικονοποίηση του δικτύου όπου μπορούν να δημιουργηθούν πολλαπλά δίκτυα πάνω από μία φυσική υποδομή. Αυτό επιτρέπει την καλύτερη χρήση πόρων και την απομόνωση μεταξύ διαφορετικών υπηρεσιών δικτύου.

Η βελτιστοποίηση της κυκλοφορίας είναι εξίσου βασική πτυχή των SDN. Χάρη στον κεντρικό έλεγχο δίνεται στους ελεγκτές η δυνατότητα να προσαρμόσουν το δίκτυο με βάση τις συνθήκες κυκλοφορίας, την εξισορρόπηση του φορτίου ή τις απαιτήσεις ποιότητας υπηρεσίας (QoS) με στόχο την βελτίωση της απόδοσης/αποδοτικότητας του δικτύου.

Η επεκτασιμότητα και η ευελιξία είναι πλεονεκτήματα που απορρέουν από τον διαχωρισμό των επιπέδων ελέγχου και δεδομένων. Αυτός ο διαχωρισμός επιτρέπει στους διαχειριστές δικτύου να κλιμακώσουν το επίπεδο ελέγχου διευκολύνοντας την επέκταση και τις αναβαθμίσεις του δικτύου. Ο διαχωρισμός επιτρέπει επίσης και την ανάπτυξη νέων υπηρεσιών δικτύου ή πρωτοκόλλων χωρίς τροποποίηση της υποδομής, αποφεύγοντας μεγάλες αλλαγές και επιτρέποντας γρηγορότερη υλοποίηση.

Το πλεονέκτημα της ασφάλειας δίνει την δυνατότητα να ενισχύσει την προστασία του δικτύου. Μέσα από την δημιουργία ενός κεντρικού σημείου ελέγχου και ορατότητας επιτρέπεται η εφαρμογή πολιτικών ασφαλείας σε επίπεδο ελέγχου διευκολύνοντας τον εντοπισμό πιθανών κινδύνων. Όταν εντοπιστεί ο κίνδυνος το SDN διευκολύνει την εφαρμογή της τμηματοποίησης και απομόνωσης του «μολυσμένου» δικτύου, μία διαδικασία που είναι πολύ χρήσιμη σε περιπτώσεις εισβολών σε ένα δίκτυο.

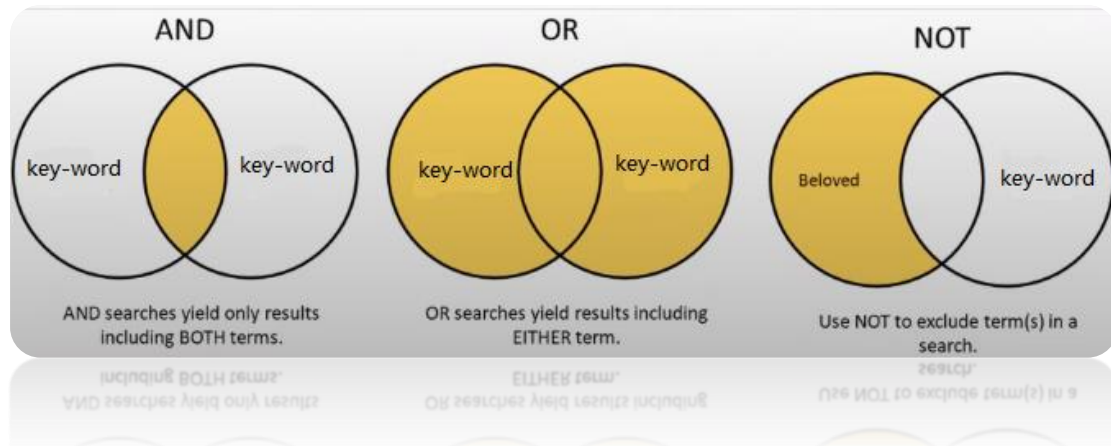
Παρά τα θετικά της τεχνολογίας SDN υπάρχουν ακόμα και κάποιες προκλήσεις. Μία τέτοια πρόκληση είναι η ανάγκη για διαλειτουργικότητα και τυποποίηση μεταξύ των υλοποιήσεων καθώς το SDN είναι νέα τεχνολογία. Ζωτικής σημασίας είναι επίσης η εγγύηση ασφαλείας

και αξιοπιστίας για την αποτροπή κακόβουλων επιθέσεων [Error! Reference source not found].

4. Μεθοδολογία

4.1 Συλλογή δεδομένων

Για να καλυφθεί ένα ευρύ φάσμα της έρευνας διαμορφώθηκε μια αναζήτηση προσαρμοσμένη στους όρους **"software-defined networks" OR "software-defined networking" OR "software defined networks" OR "software defined networking"**. Σημαντικός είναι ο ρόλος των τελεστών Boolean για την εμφάνιση του κατάλληλου εύρους.



Εικόνα 4.1 Τελεστές Boolean για την εμφάνιση κατάλληλου εύρους όρων

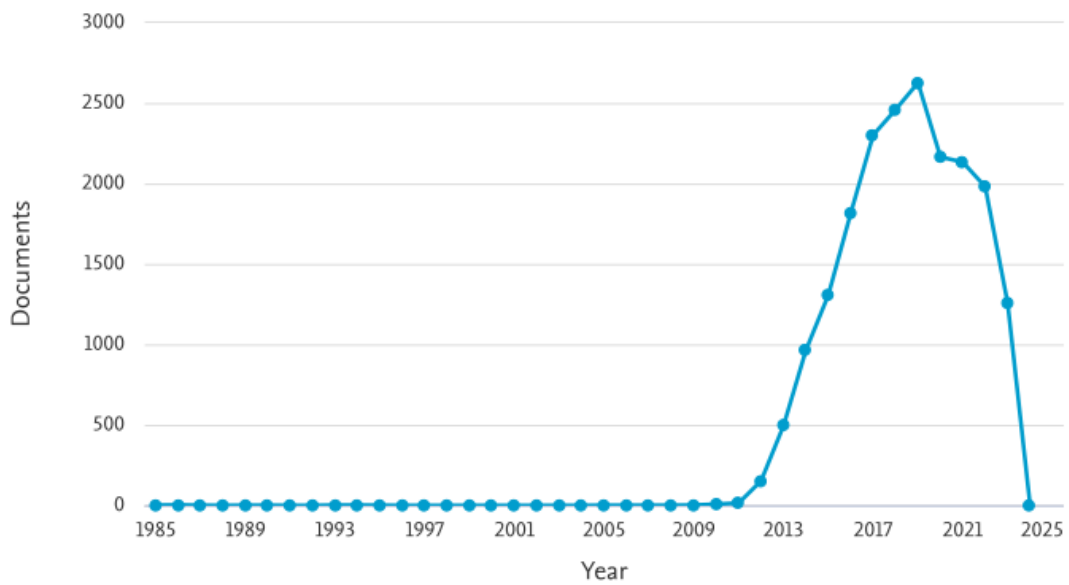
Συνολικά απαριθμήθηκαν 19.687 εγγραφές και έπειτα από την προεπεξεργασία (preprocessing) των δεδομένων οι τελικές εγγραφές που κρίθηκαν πιο αντιπροσωπευτικές ανέρχονται στις 11.200. Στην προεπεξεργασία αφαιρέθηκαν δεδομένα που θα επηρέαζαν την ορθότητα των αποτελεσμάτων, όπως είναι οι διπλότυπες καταχωρήσεις και τα δεδομένα που δεν είχαν πληρότητα και ορθότητα στις εγγραφές τους.

Το διάστημα που έχει μελετηθεί είναι από το 1985 έως το 2024. Όπως φαίνεται και στην Εικόνα 4.2 την δεκαετία του 2010 παρατηρούνται οι πρώτες εγγραφές και η εμφάνισή τους αυξάνεται ραγδαία φτάνοντας τις 2.500 εγγραφές το 2019. Έπειτα υπάρχει μια πτώση στις δημοσιεύσεις που πιθανολογείται ότι οφείλεται στους εξής λόγους:

1. στην πανδημία του COVID όπου όλοι οι οργανισμοί ήταν κλειστοί και έπειτα λειτουργούσαν με περιορισμούς
2. στην μη ενσωμάτωση των πρόσφατων δημοσιεύσεων στη βάση δεδομένων του Scopus και
3. στην πιθανή αλλαγή των λέξεων κλειδιών και της περαιτέρω εξειδίκευσης στον επιστημονικό τομέα.

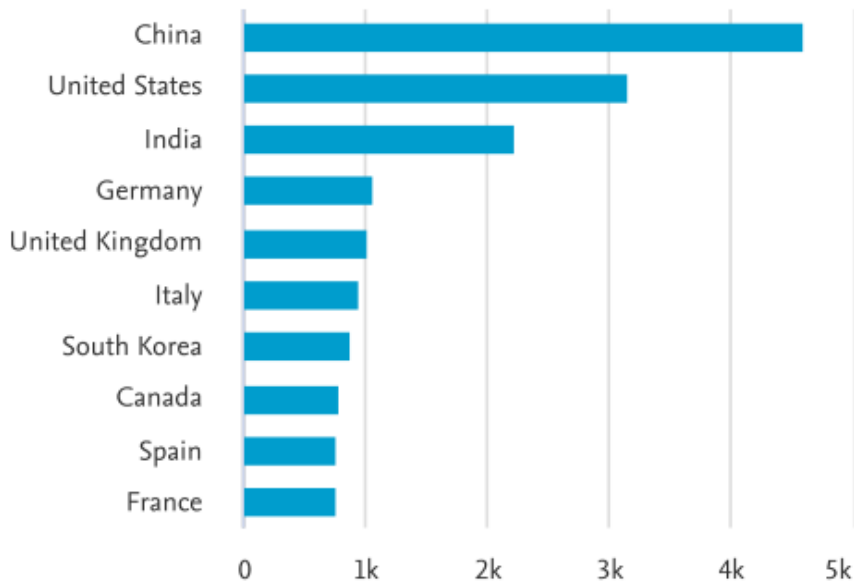
Πρώτη χώρα σε δημοσιεύσεις είναι η Κίνα με πάνω από 4.500 έγγραφα, ακολουθούν οι Ηνωμένες Πολιτείες και η Ινδία με 3.150 και 2.209 έγγραφα αντίστοιχα.

Documents by year



Εικόνα 4.2 Χρονική δημοσίευση εγγράφων

Documents by country/territory



Εικόνα 4.3 Οι 10 χώρες με τις περισσότερες μελέτες

4.2 Μέθοδοι ανάλυσης στο VOSViewer

Για την ανάγκη της διπλωματικής έγινε χρήση της έκδοσης 1.6.16 για περιβάλλον windows του ελεύθερου λογισμικού VOSViewer. Το συγκεκριμένο εργαλείο προσφέρει στον ερευνητή μία ολοκληρωμένη σουίτα από μεθόδους ανάλυσης με σκοπό την εξερεύνηση και την αποσαφήνιση περίπλοκων μοτίβων, συνδέσεων και τάσεων στην επιστημονική βιβλιογραφία.

Μία από τις βασικές μεθόδους ανάλυσης είναι η ανάλυση συνεργασίας (Co-Authorship). Σε αυτή την μέθοδο ανάλυσης εξετάζονται τα δίκτυα συνεργασίας μεταξύ των ερευνητών-συγγραφέων ή των οργανώσεων ή των χωρών. Το πρόγραμμα προσδιορίζει ομάδες συνεργασίας που συμμετέχουν σε επαναλαμβανόμενες δημοσιεύσεις.

Επιπλέον, το VOSViewer προσφέρει την ανάλυση συμ-παραπομπής (Co-citation). Εξετάζοντας τις σχέσεις συν-αναφοράς μεταξύ των εγγράφων το λογισμικό ανακαλύπτει την αλληλοσυσχέτιση μεταξύ των δημοσιεύσεων με βάση τις κοινές τους αναφορές σε ένα πεδίο μελέτης.

Η βιβλιογραφική ανάλυση σύζευξης (Bibliographic coupling) για την ανακάλυψη της σύγκλισης των εγγράφων που αναφέρονται σε παρόμοιες παραπομπές είναι ακόμη μια μέθοδος ανάλυσης που προσφέρει το εργαλείο αυτό. Αυτή η μέθοδος βοηθάει στον εντοπισμό εγγράφων που έχουν ως κοινό το θεματικό περιεχόμενο, εμπλουτίζοντας την κατανόηση των ερευνητικών σχέσεων.

Το λογισμικό αξιοποιεί την συνεμφάνιση λέξεων-κλειδιών (Co-occurrence) με σκοπό τον εντοπισμό και την καταγραφή τους σε έναν ερευνητικό τομέα.

Επιπροσθέτως, ακόμη μία ικανότητά του είναι να ομαδοποιεί όρους, συγγραφείς και έγγραφα με βάση τα μοτίβα συν-εμφάνισής τους. Αυτό βοηθάει ουσιαστικά στην διευκόλυνση των περίπλοκων θεματικών σχέσεων.

Τέλος, δε θα μπορούσε να απέχει η μέθοδος της απεικόνισης. Τρεις απεικονίσεις είναι διαθέσιμες από το λογισμικό:

1. Οπτικοποίηση δικτύου: τα δεδομένα εμφανίζονται στο χώρο με την ετικέτα τους και από έναν κύκλο. Το μέγεθος αυτών των δύο αντιπροσωπεύουν και την δύναμη του αντικειμένου. Όσο μεγαλύτερα είναι τόσο μεγαλύτερη δυναμική έχει το αντικείμενο. Τα διαφορετικά χρώματα κύκλων υποδηλώνουν τις κλάσεις μέσα στις οποίες ανήκουν οι ομάδες των αντικειμένων και οι γραμμές που ενώνουν τους κύκλους αφορούν τις συνδέσεις.
2. Απεικόνιση επικάλυψης: είναι παρόμοια με την προηγούμενη και η μόνη διαφορά είναι στον χρωματισμό των αντικειμένων. Εδώ το χρώμα υποδηλώνει την βαρύτητα του αντικειμένου που μπορεί να μετρηθεί είτε ως χρονολογία είτε ως βαθμολογία. Τα χρώματα από προεπιλογή είναι οι αποχρώσεις του μπλε-πράσινου και κίτρινου. Το μπλε δηλώνει την λιγότερη δύναμη και πλησιάζοντας το κίτρινο δηλώνεται η υψηλότερη δυναμική (είτε είναι βαθμολογία, είτε κάποια εξέλιξη σε βάθος χρόνου)

3. Απεικόνιση πυκνότητας: επίσης παρόμοια με τις προηγούμενες. Η προεπιλεγμένη βεντάλια των χρωμάτων κυμαίνεται από το μπλε-πράσινο και κίτρινο. Ο κόμβος που έχει χρώμα κίτρινο δηλώνει την μεγάλη πυκνότητα στην συγκεκριμένη περιοχή.

4.3 Μέθοδοι ανάλυσης του Gephi

Το συγκεκριμένο εργαλείο χρησιμοποιείται τόσο για την ανάλυση γράφων όσο και για την ανάλυση και οπτικοποίηση δεδομένων. Αναλόγως με τον στόχο που θέτουμε το εργαλείο παρέχει και την κατάλληλη μέθοδο ανάλυσης. Σε αυτές τις μεθόδους γίνεται αναφορά στη συνέχεια:

- Ανίχνευση κοινοτήτων (Community Detection): εντοπίζει τις ομάδες κόμβων που συνδέονται μεταξύ τους. Αυτές οι ομάδες (κοινότητες) μπορεί να είναι φυσικές ή μη φυσικές συστάδες στο δίκτυο. Αυτό βοηθά στην κατανόηση των συνδέσεων μεταξύ των ομάδων που βρίσκονται σε ένα δίκτυο.
- Ανάλυση κεντρικότητας (Centrality Analysis): αναδεικνύει τους κόμβους που επιδρούν στην δομή και την επικοινωνία του δικτύου.
- Κατάταξη κόμβων (Node Ranking): θέτοντας συγκεκριμένα χαρακτηριστικά μπορούμε να εντοπίσουμε την επιρροή μεταξύ κόμβων σε ένα δίκτυο ανάλογα με την τιμή rank που έχουν.
- Ανάλυση διαδρομής (Path Analysis): η συγκεκριμένη ανάλυση εμφανίζει τις διαδρομές ή τα μονοπάτια που υπάρχουν στους κόμβους και στον τρόπο που μεταδίδεται η πληροφορία στο δίκτυο.
- Ανάλυση κοινής γειτνίασης (Common Neighbor Analysis): αναλύει τις σχέσεις μεταξύ γειτονικών κόμβων.
- Συνεργασία συγγραφέων (Authorship Collaboration): η συγκεκριμένη μέθοδος επικεντρώνεται μεταξύ της συνεργασίας συγγραφέων για την δημιουργία νέων ερευνητικών έργων.
- Ανάλυση κλικ (Clique Analysis): εστιάζει σε ανάλυση κόμβων οι οποίες είναι πλήρως συνδεδεμένες μεταξύ τους. Όπως και η προηγούμενη ανάλυση μπορεί να αποκαλύψει σχέσεις συγγραφέων που συνδέονται στενά.
- Ανάλυση διακριτικότητας (Modularity Analysis): αναλύει το δίκτυο με σκοπό να βρει θέματα τα οποία ανήκουν σε διαφορετικές κατηγορίες.

4.4 Μέθοδοι ανάλυσης του R Studio

Οι μέθοδοι ανάλυσης που μπορούν μέσα από το RStudio να εφαρμοστούν εξαρτώνται από τις ανάγκες της ανάλυσης και της δυνατότητας της γλώσσας R. Για την παρούσα έρευνα χρησιμοποιήθηκε η R 4.2.3 και το RStudio για βιβλιομετρική ανάλυση. Για αυτό το είδος μελέτης κύριο ρόλο είχε η βιβλιοθήκη “bibliometrix” η οποία παρέχει πληθώρα συναρτήσεων και αργαλειών κατάλληλα για αυτό το σκοπό. Μέσα από το bibliometrix το

λογισμικό είναι σε θέση να «διαβάσει» τα δεδομένα, να τα αναλύσει και να δημιουργήσει τα αντίστοιχα network plots.

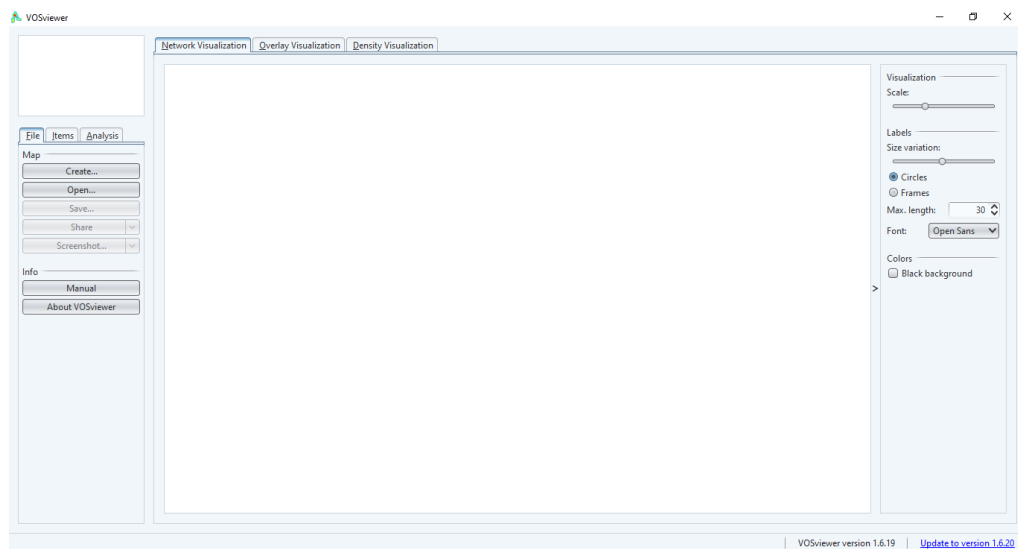
Το RStudio πέρα από την βιβλιομετρική χρησιμοποιείται και για αναλύσεις όπως:

- Στατιστική Ανάλυση
- Ανάλυση δεδομένων μηχανικής μάθησης
- Ανάλυση κειμένου
- Εξερευνητική ανάλυση δεδομένων (EDA)
- Ανάλυση χρονοσειρών
- Ανάλυση επιβίωσης
- Ανάλυση μεγάλου όγκου δεδομένων

5. Ανάλυση δεδομένων

5.1 Ανάλυση μέσω του λογισμικού VOSViewer

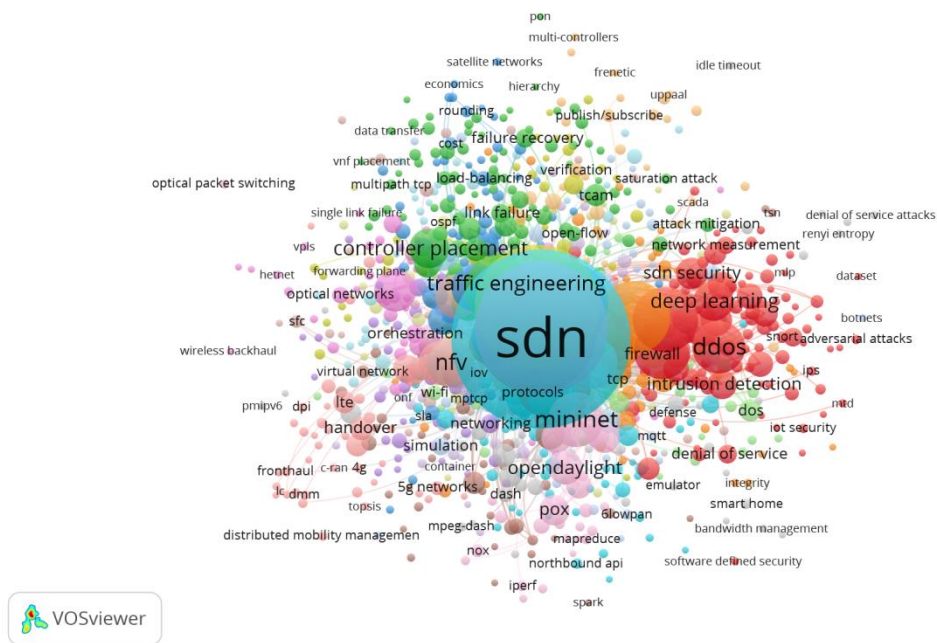
Κατά την εκκίνηση του λογισμικού ο χρήστης θα συναντήσει ένα φιλικό περιβάλλον (Εικόνα 5.1). Στο κέντρο είναι ο κύριος πίνακας που παρουσιάζει μία απεικόνιση των δεδομένων που έχουν εισαχθεί. Διαθέτει τρεις απεικονίσεις όπως θα δούμε και στην συνέχεια, την απεικόνιση δικτύου (network visualization), την απεικόνιση επικάλυψης (overlay virtualization) και την οπτικοποίηση πυκνότητας (density virtualization). Ακριβώς από κάτω είναι ο πίνακας πληροφοριών που αναφέρει τις πληροφορίες για τα στοιχεία του χάρτη. Στα δεξιά ο χρήστης συναντά έναν πίνακα επιλογών όπου του δίνεται η δυνατότητα να προσαρμόσει την οπτικοποίηση και τέλος στα αριστερά υπάρχει μία επισκόπηση του χάρτη καθώς και ο πίνακας ενεργειών



Εικόνα 5.1 Περιβάλλον VOSViewer

5.1.1 Χαρτογράφηση συνύπαρξης λέξεων-κλειδιών (Co-occurrence of Author key-words)

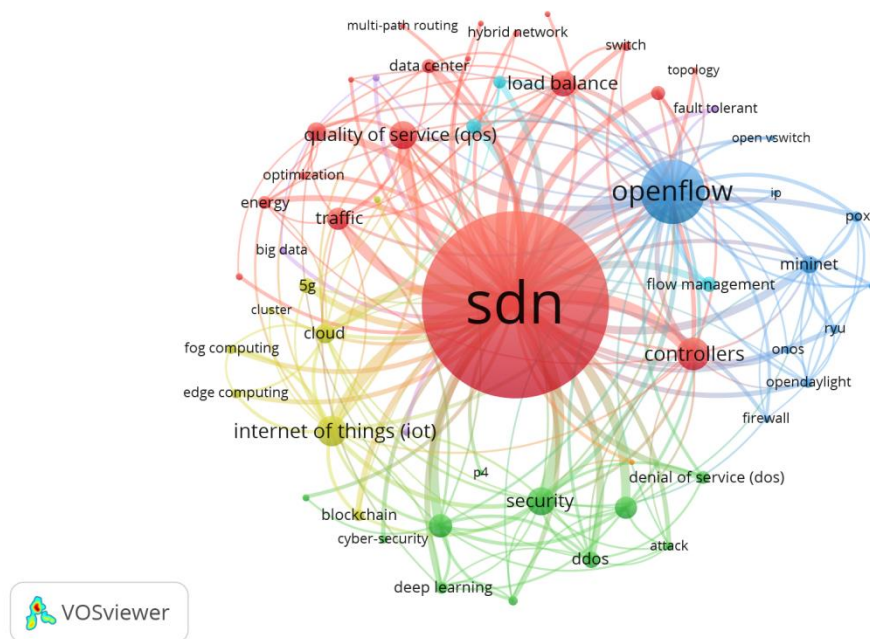
Η μελέτη της συνύπαρξης των λέξεων κλειδιών θα προσεγγίσει την τρέχουσα κατάσταση των δικτύων SDN. Μέσα από τις επιστημονικές δημοσιεύσεις θα αντληθεί ένα εύρος κλειδιών με σκοπό να υπάρξει κάποιο συμπέρασμα. Στην Εικόνα 5.2 φαίνονται όλες οι συνδέσεις που υπάρχουν χωρίς κάποιο φιλτράρισμα. Η προεπιλεγμένη τιμή που αφορά τον ελάχιστο αριθμό εμφάνισης της λέξης κλειδί είναι 5, αυτό συνεπάγεται ότι οι 13.480 λέξεις θα μειωθούν στις 1.030.



Εικόνα 5.2 Βιβλιογραφικός χάρτης λέξεων κλειδιών με τις προεπιλεγμένες τιμές

Όπως θα δούμε και στη συνέχεια τα διαφορετικά χρώματα δηλώνουν τα clusters που υπάρχουν, το μέγεθος του κύκλου αντιπροσωπεύει την δύναμη και οι γραμμές δείχνουν τις συνδέσεις.

Στην Εικόνα 5.3 φαίνεται ίδιος χάρτης με αλλαγμένες τις προεπιλεγμένες τιμές. Συγκεκριμένα, ο ελάχιστος αριθμός εμφάνισης ορίστηκε στο 50 για να περιοριστεί το αρχικό πλήθος των δεδομένων και να εστιαστεί η ανάλυση σε πιο σημαντικές αναφορές.



Εικόνα 5.3 Χάρτης λέξεων-κλειδιών έπειτα από φιλτράρισμα

Ο χάρτης που δημιουργήθηκε εμφανίζει τα 7 cluster όπως φαίνονται στους παρακάτω πίνακες.

- Cluster 1: διακρίνεται με κόκκινο χρώμα και αφορά την διαχείριση των δικτύων. Οι λέξεις που περιέχονται μέσα σε αυτό αναφέρονται στην διαχείριση και την ασφάλεια των δικτύων καθώς και στην επεξεργασία δεδομένων. Λέξεις όπως «SDN, Load Balance, QoS, Traffic» έχουν σχέση με την λειτουργία και την απόδοση αυτών. Είναι εμφανές ότι ο κύκλος που περιβάλλει την λέξη SDN είναι πολύ μεγαλύτερος λόγω της κυριαρχίας και δυναμικής που έχει.
- Cluster 2: εμφανίζεται με πράσινο χρώμα και περιέχει λέξεις που έχουν στενή σχέση με την ασφάλεια και την ανίχνευση απειλών στην επεξεργασία δεδομένων.
- Cluster 3: είναι το σύμπλεγμα συνδέσεων με χρώμα μπλε. Περιλαμβάνει λέξεις που αφορούν τον έλεγχο του δικτύου. «ONOS, OpenDaylight, POX, Ryu» είναι ονόματα ελεγκτών που ενδεχομένως υπάρχουν σε κάποιο δίκτυο. Επίσης το λογισμικό ελέγχου mininet και το πρωτόκολλο OF έχουν σχέση με τον αυτόματο έλεγχο του δικτύου.
- Cluster 4: διακρίνεται με κίτρινο χρώμα στον χάρτη και εστιάζει σε θέματα που έχουν σχέση με την συνδεσιμότητα, την επεξεργασία δεδομένων και το IoT. Γίνεται λόγος για την 5G τεχνολογία, την Blockchain που ως επί το πλείστον βοηθά τα δεδομένα να διατηρήσουν αναλλοίωτη την ταυτότητά τους. Επίσης στο συγκεκριμένο cluster συναντάμε έννοιες που αναφέρονται στην υπηρεσία cloud και στις τεχνολογίες edge και fog computing.
- Cluster 5: χρώματος μοβ. Η αναφορά του σχετίζεται γύρω από την χρήση των τεχνολογιών δικτύου στην διαχείριση μεγάλου όγκου δεδομένων (bigdata), την σωστή λειτουργία των συστημάτων ακόμα και σε συνθήκες σφάλματος (fault tolerant) και στην βελτίωση της ενέργεια μέσα από τα έξυπνα δίκτυα (smartgrid).

- Cluster 6: το χρώμα του είναι γαλάζιο και αφορά την εικονοποίηση του δικτύου (network virtualization) και την διαχείριση της ροής (Flow management).
- Cluster 7: χρώματος πορτοκαλί. Σε αυτό το cluster μονοπωλεί το ενδιαφέρον του ελέγχου πρόσβασης (Access control) το οποίο είναι σημαντικό για την ασφάλεια του δικτύου διότι καθορίζεται σε ποιόν θα επιτραπεί η προσβασιμότητα.

Πίνακας 5.1 Πληροφορίες λέξεων-κλειδιών (ταξινόμηση με βάση το TotalLinkStrength)

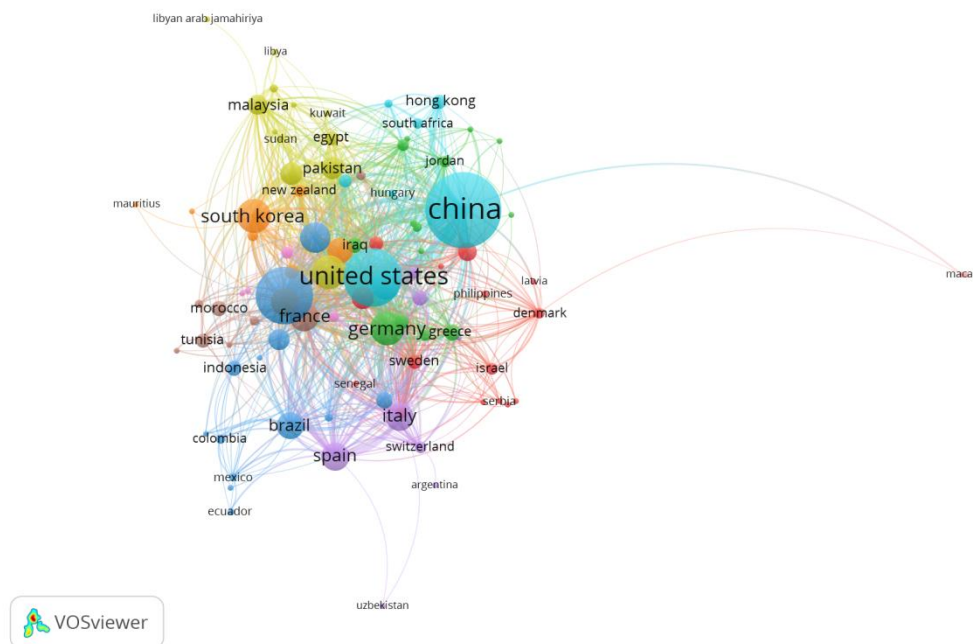
Key-words	Occurrences	TotalLinkStrength	Cluster
P4	57	87	2
Multi-PathRouting	56	96	1
GeneticAlgorithms	53	101	1
Quality of Experience (QoE)	50	103	1
Access Control	50	103	7
DeepReinforcementLearning	71	113	1
Cluster	51	116	4
Mobility	82	124	4
Smart Grid	72	125	5
Delay	50	127	1
HybridNetwork	80	134	1
OpenvSwitch	51	135	3
Bandwidth	58	140	1
Topology	63	142	1
Computer Networks	76	145	5
IP	66	156	3
FaultTolerant	88	163	5
Cyber-Security	87	164	2
ArtificialIntelligence	73	171	2
BigData	75	172	5

Optimization	90	192	1
Firewall	72	205	3
Switch	107	237	1
Floodlight	72	239	3
ONOS	85	242	3
FogComputing	102	254	4
EdgeComputing	106	255	4
Attack	101	258	2
IntrusionDetectionSystem (IDS)	106	261	2
NetworkVirtualization	167	310	6
OpenDaylight	115	328	3
Energy	179	349	1
Blockchain	156	353	4
DeepLearning	158	357	2
Ryu	119	359	3
5G	215	369	4
Denial of Service (DoS)	180	391	2
POX	125	394	3
DataCenter	200	409	1
VirtualNetwork	225	426	6
FlowManagement	228	442	6
ControlLayer	189	443	1
Routing	289	570	1
DDoS (Distributed Denial of Service)	366	685	2
Traffic	364	686	1

Cloud	324	726	4
Mininet	269	785	3
MachineLearning	408	919	2
LoadBalance	476	967	1
Quality of Service (QoS)	539	1138	1
Security	537	1217	2
Controllers	646	1249	1
Internet of Things (IoT)	599	1337	4
OpenFlow	1738	3298	3
Software-DefinedNetworking (SDN)	8380	10223	1

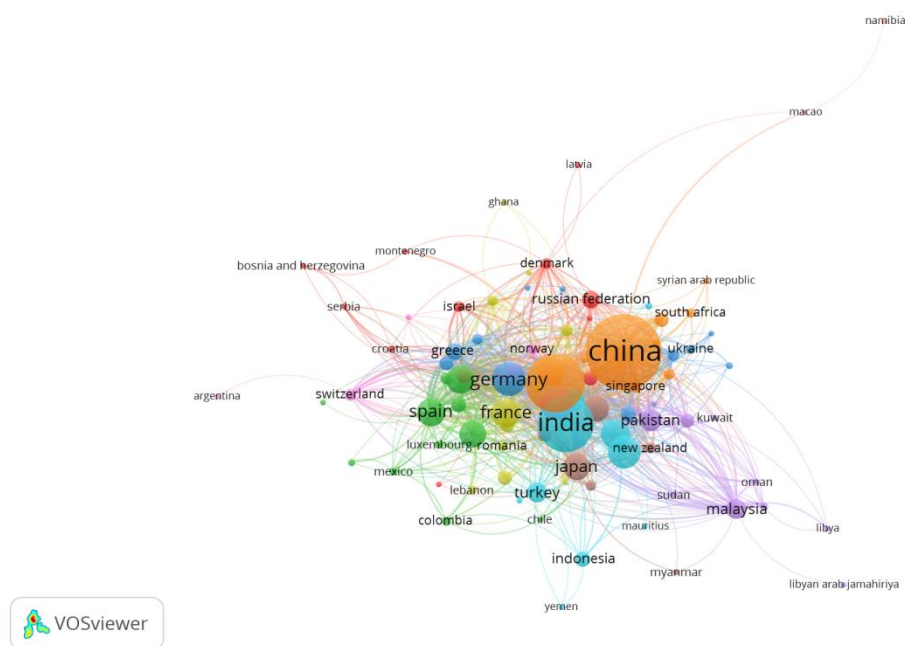
5.1.2 Χαρτογράφηση κοινής συνεργασίας μεταξύ χωρών (Co-authorship countries)

Η συγκεκριμένη μέθοδος χαρτογράφησης επικεντρώνεται στις σχέσεις μεταξύ των χωρών όπου συνεργάστηκαν για να συνεισφέρουν στο αντικείμενο της επιστημονικής μελέτης. Η αρχική μορφή του χάρτη, με τις προεπιλεγμένες τιμές, μοιάζει με την ακόλουθη εικόνα (Εικόνα 5.4).



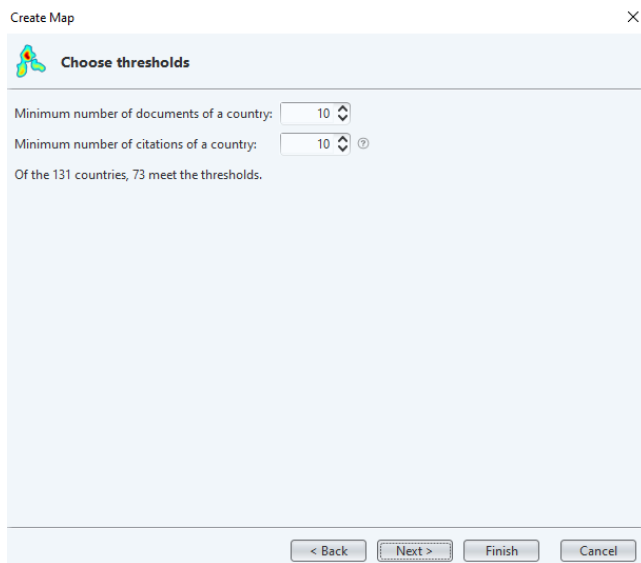
Εικόνα 5.4 Χάρτης κοινής συνεργασίας χωρών χωρίς φιλτράρισμα

Στη συνέχεια η ελάχιστη τιμή εμφανίσεων των αναφορών ορίστηκε σε 10. Αυτή η επιλογή έγινε για να έχουμε έναν πιο ευδιάκριτο χάρτη διότι επιλέγοντας την τιμή 10 ορίζουμε στο πρόγραμμα να εμφανίσει τις χώρες με την συνεργασία σε τουλάχιστον 10 επιστημονικές μελέτες. Η όψη του χάρτη φαίνεται στην συνέχεια (Εικόνα 5.5).

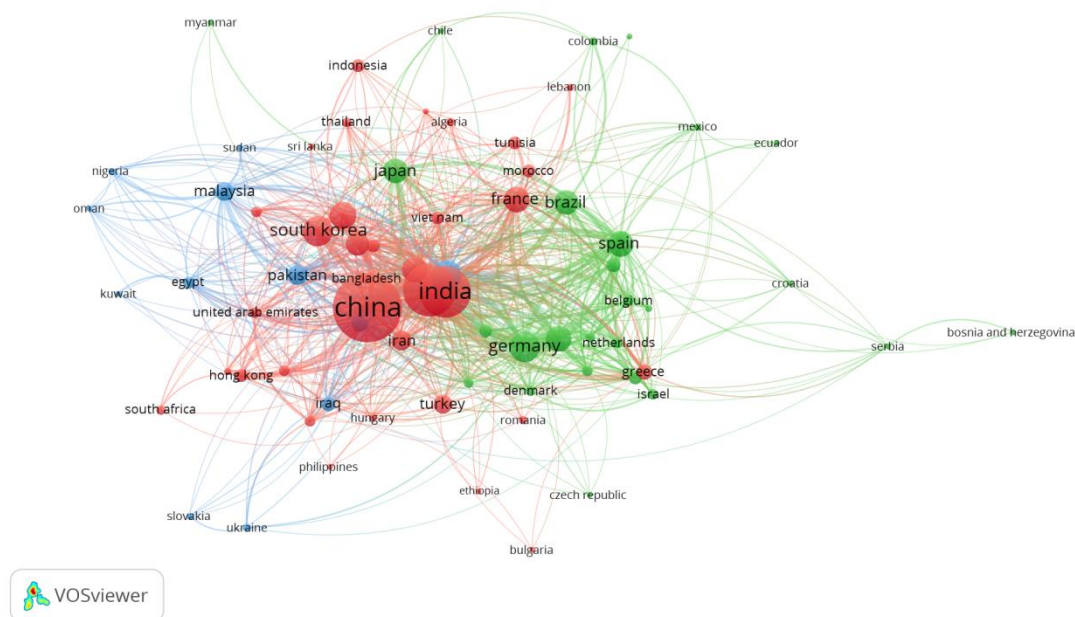


Εικόνα 5.5 Χάρτης κοινής συνεργασίας χωρών με τουλάχιστον 10 εμφανίσεις

Για να εμφανιστεί ένας πιο ξεκάθαρος χάρτης έγινε ακόμη ένα φιλτράρισμα στις προεπιλεγμένες τιμές. Συγκεκριμένα αλλάχθηκαν οι τιμές σε 10 στα πεδία που αναφέρονται στις ελάχιστες αναφορές που πρέπει να έχει μία χώρα και τα ελάχιστα έγγραφα. Όπως φαίνεται στην Εικόνα 5.6, από τις 131 χώρες οι 73 έχουν τιμές μεγαλύτερες ή ίσες από το κατώφλι που έχουμε ορίσει. Έπειτα στην Εικόνα 5.7 φαίνεται ο τελικός χάρτης.



Εικόνα 5.6 Τιμές κατωφλίου



Εικόνα 5.7 Τελικός χάρτης κοινής συνεργασίας χωρών

Για να έχουμε αυτή την οπτική μορφή επιλέχθηκε η μέθοδος ανάλυσης “Fractionalization” για τον λόγο ότι διασπά σε τμήματα τις συνεργασίες. Στην προκειμένη περίπτωση η διάσπαση γίνεται για να εμφανιστούν ποιές χώρες συνεργάζονται περισσότερο. Για να έχουμε αυτή την τελική μορφή του χάρτη οι τιμές στα πεδία “Resolution” και “Min. Clustersize” οι τιμές αλλάχθηκαν σε resolution=3 & min.cluster size=10. Η τιμή της ανάλυσης (resolution) όσο μεγαλύτερη δημιουργεί περισσότερες ομάδες, ενώ μειώνοντάς της οι ομάδες μικραίνουν. Αντίστοιχα ορίζοντας έναν ελάχιστο μέγεθος στα clusters εμφανίζει μόνο τις σημαντικές σχέσεις που έχουν την μεγαλύτερη συνεργασία.

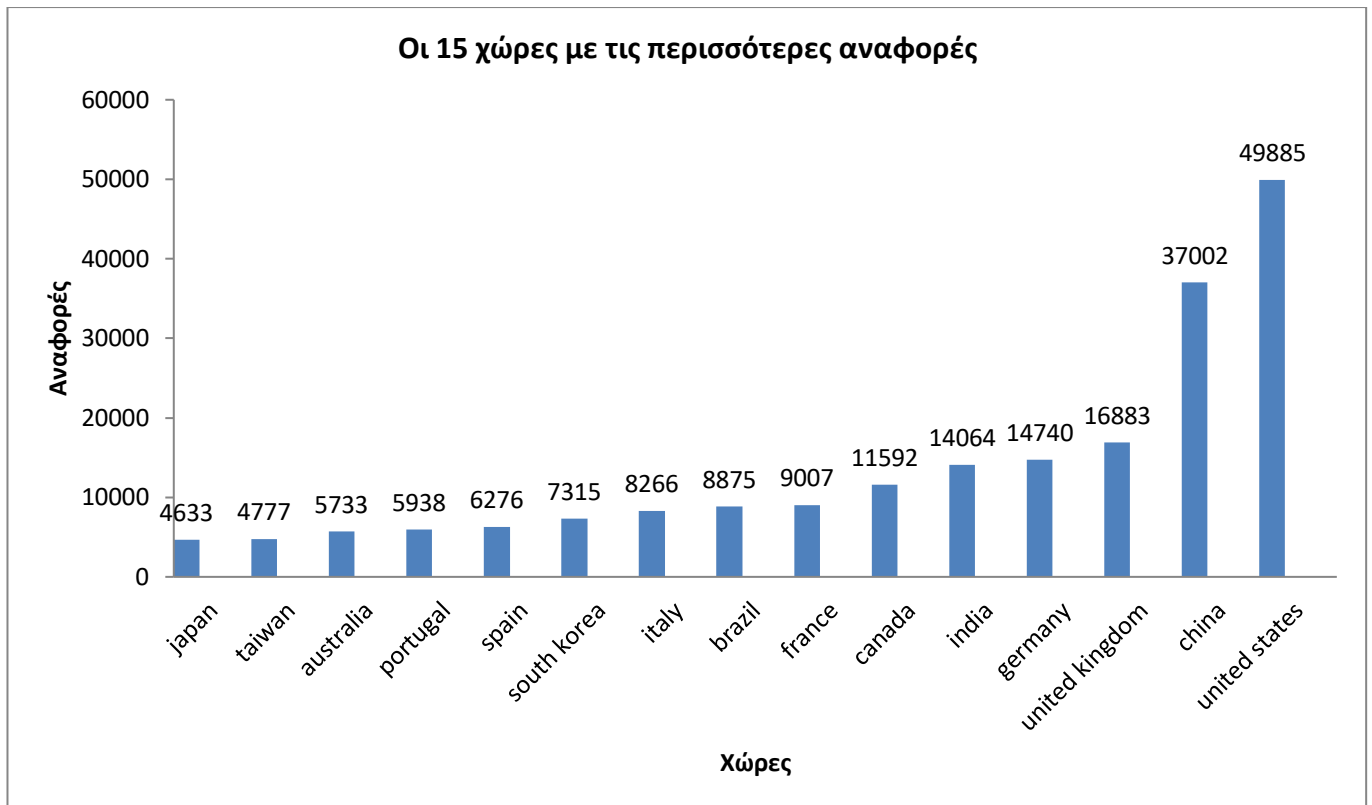
Πίνακας 5.2Χώρες ταξινομημένες με βάση τις αναφορές

Countries	Documents	Citations	TotalLinkStrength
SENEGAL	10	15	3
BULGARIA	13	33	4
ECUADOR	25	70	9
MYANMAR	19	73	5
SERBIA	20	79	19
PHILIPPINES	26	82	6
SLOVAKIA	28	84	11
CZECHREPUBLIC	19	88	7
ETHIOPIA	11	88	6
CROATIA	13	97	15
BOSNIA AND HERZEGOVINA	20	98	5
OMAN	13	98	23
CHILE	14	113	14
SRILANKA	17	114	8
ESTONIA	11	131	20
NIGERIA	32	184	39
SUDAN	20	212	27
KUWAIT	21	283	11
ROMANIA	45	300	11
UKRAINE	48	352	22
LEBANON	35	385	12
THAILAND	52	405	34
INDONESIA	109	409	22
HUNGARY	28	455	37
COLOMBIA	42	463	28
ALGERIA	42	480	28
SOUTHAFRICA	50	490	9
MEXICO	33	648	31
JORDAN	72	710	87
TUNISIA	92	737	45
BANGLADESH	66	742	67
AUSTRIA	53	770	74
IRAQ	115	776	59
NORWAY	51	832	65
EGYPT	91	839	73
UNITEDARABEMIRATES	62	855	88
NEWZEALAND	62	914	65
DENMARK	61	1050	79
QATAR	33	1116	40

VIETNAM	98	1142	79
NETHERLANDS	62	1289	72
RUSSIANFEDERATION	157	1382	71
BELGIUM	69	1405	95
ISRAEL	62	1431	53
POLAND	98	1669	86
MOROCCO	106	1691	22
SWITZERLAND	67	1756	92
SWEDEN	99	1815	117
IRELAND	107	1962	104
GREECE	134	2259	145
MALAYSIA	197	2266	167
HONGKONG	94	2283	112
FINLAND	94	2453	94
IRAN	238	2510	107
SINGAPORE	75	2532	55
TURKEY	199	2952	52
LUXEMBOURG	35	3570	35
PAKISTAN	242	3851	270
JAPAN	361	4633	211
TAIWAN	417	4777	167
AUSTRALIA	304	5733	268
PORTUGAL	121	5938	99
SPAIN	374	6276	354
SOUTH KOREA	527	7315	240
ITALY	379	8266	348
BRAZIL	338	8875	170
FRANCE	398	9007	318
CANADA	372	11592	312
INDIA	1493	14064	333
GERMANY	543	14740	465
UNITEDKINGDOM	522	16883	573
CHINA	2658	37002	985
UNITED STATES	1586	49885	1029

Στον Πίνακα 5.2 παρατίθενται όλες οι πληροφορίες που έχουν προκύψει από την επεξεργασία και οπτικοποιούνται στην Εικόνας 5.8.

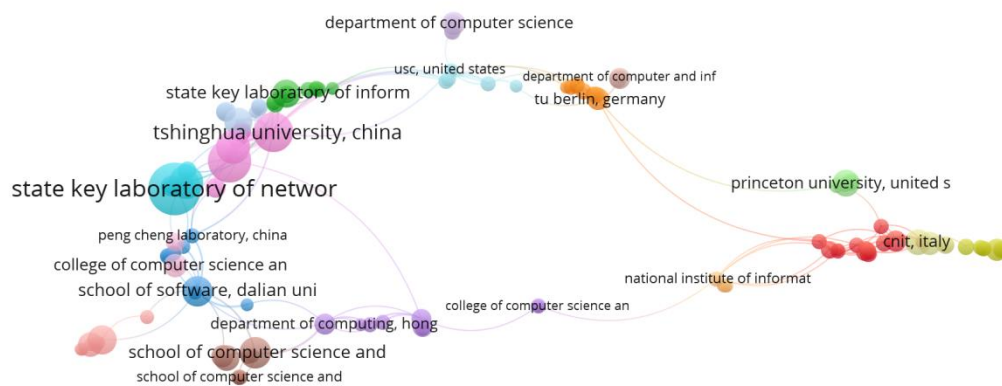
Δίνοντας στην συνέχεια μία άλλη οπτική στα παραπάνω δεδομένα, εμφανίζονται οι 15 χώρες με τις περισσότερες αναφορές.



5.8 Οπτικοποίηση 15 χωρών με περισσότερες αναφορές

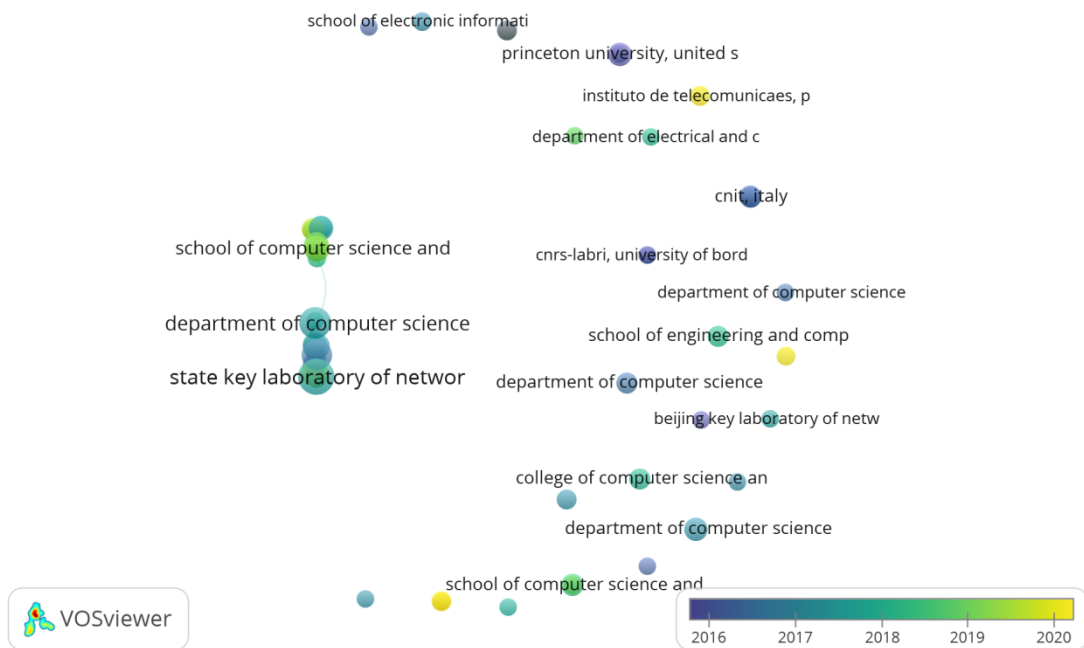
5.1.3 Χαρτογράφηση κοινής συγγραφής μεταξύ οργανισμών (Co-authorship organizations)

Ο συγκεκριμένος χάρτης παρουσιάζει οργανισμούς που έχουν συνεργαστεί στην επιστημονική κοινότητα βάσει των δημοσιεύσεων και των αναφορών που έχουν λάβει. Ο χάρτης που ακολουθεί (Εικόνα 5.9) είναι με τις προεπιλεγμένες τιμές του προγράμματος.



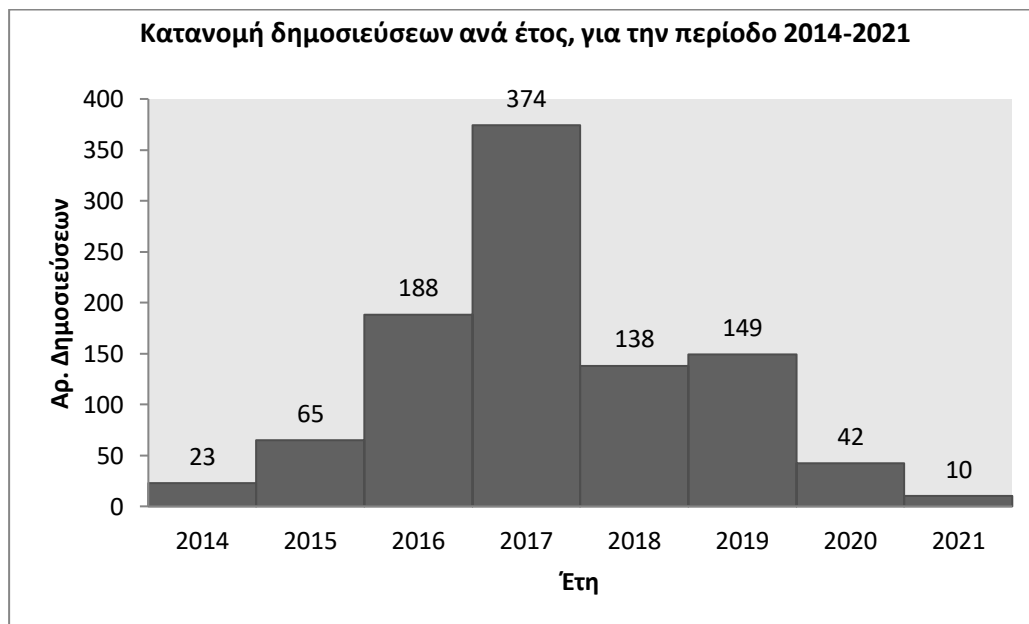
Εικόνα 5.9 Χάρτης κοινής συγγραφής οργανισμών χωρίς φιλτράρισμα

Έπειτα από την απαλοιφή των διπλότυπων εγγραφών και θέτοντας τον αριθμό 10 στους περιορισμούς με τα ελάχιστα έγγραφα και τις ελάχιστες αναφορές ανά οργανισμό, ο χάρτης συν-συγγραφής μεταξύ των οργανισμών έχει την μορφή της Εικόνας 5.10. Η απεικόνιση επικάλυψης (Overlay Virtualization) περιγράφει από τους 16589 οργανισμούς τους 54 εκ των οποίων οι 22 συνδέονται μεταξύ τους. Μέσω αυτής της απεικόνισης θέλουμε να δούμε ανά έτος τι άθροισμα δημοσιεύσεων υπάρχει.



Εικόνα 5.10 Χάρτης Overlay Virtualization

Το χρονικό εύρος είναι από το 2014 μέχρι το 2021 και οι δημοσιεύσεις ανέρχονται στις 989. Στο διάγραμμα που ακολουθεί φαίνονται ότι οι δημοσιεύσεις των οργανισμών ανά έτος με τον μέγιστο αριθμό δημοσιεύσεων να είναι το 2017.



5.11 Κατανομή δημοσιεύσεων

5.1.4 Χαρτογράφηση συν-συγγραφέων (Co-authorship authors)

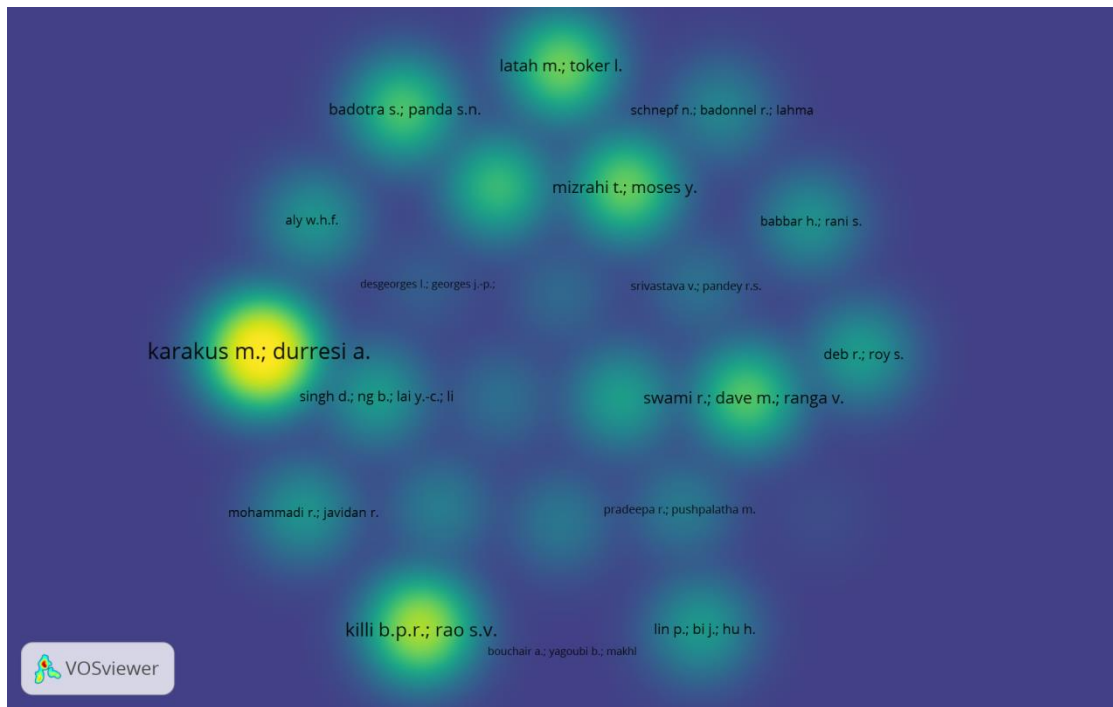
Στη συγκεκριμένη υποενότητα θα δημιουργεί ο χάρτης των συν-συγγραφέων. Ουσιαστικά θα φανούν τα αποτελέσματα του προγράμματος ως προς την αλληλεπίδραση που υπάρχει των συγγραφέων. Όπως φαίνεται στην συνέχεια (Εικόνα 5.12) η αλληλεπίδραση που έχουν οι συγγραφείς μεταξύ τους είναι μηδαμινή. Αυτό είναι πολύ πιθανό να συμβαίνει γιατί μπορεί να προέρχονται από διαφορετικούς τομείς (πληροφορική, μηχανική, analytics, κα).

Στην ακόλουθη εικόνα φαίνεται οι σχέσεις που έχουν εντοπιστεί.



Εικόνα 5.12 Χάρτης αλληλεπίδρασης συγγραφέων

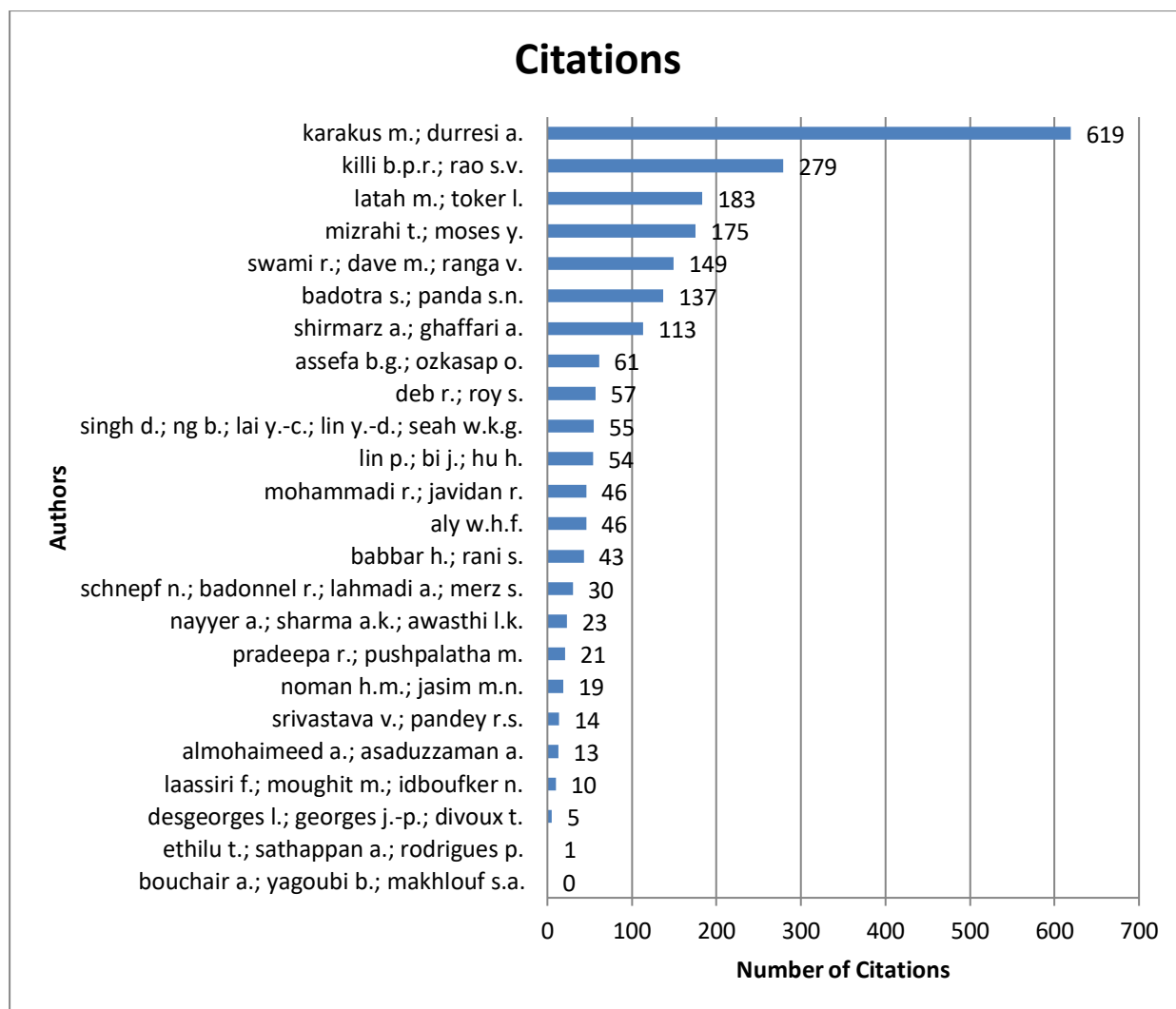
Αφήνοντας τις προεπιλεγμένες τιμές του προγράμματος και επιλέγοντας μία διαφορετική οπτική του χάρτη μας οδηγεί στο πιο κάτω αποτέλεσμα (Εικόνα 5.13). Εδώ βλέπουμε με οπτικοποίηση πυκνότητας (Density Virtualization) τους συγγραφείς που έχουν τις περισσότερες αναφορές.



Εικόνα 5.13 Επικάλυψη πυκνότητας για συγγραφείς βάση των αναφορών

Πίνακας 3 Συγγραφείς με βάση τις αναφορές

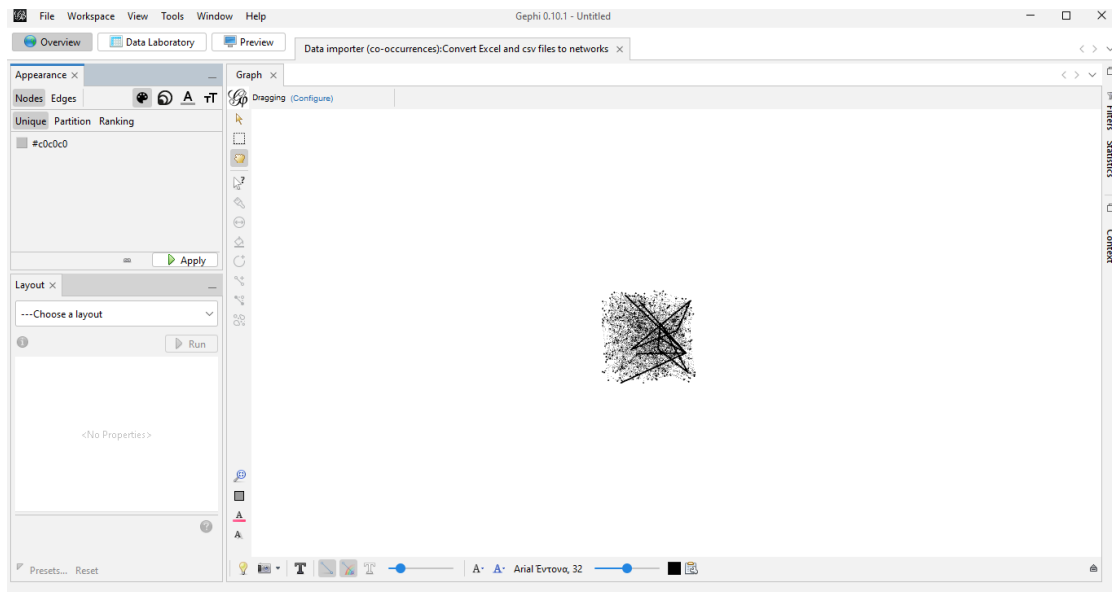
Authors	Citations
BOUCHAIR A.; YAGOUBI B.; MAKHLOUF S.A.	0
ETHILU T.; SATHAPPAN A.; RODRIGUES P.	1
DESGEORGES L.; GEORGES J.-P.; DIVOUX T.	5
LAASSIRI F.; MOUGHIT M.; IDBOUFKER N.	10
ALMOHAIMEED A.; ASADUZZAMAN A.	13
SRIVASTAVA V.; PANDEYR.S.	14
NOMANH.M.; JASIMM.N.	19
PRADEEPA R.; PUSHPALATHA M.	21
NAYYER A.; SHARMAA.K.; AWASTHIL.K.	23
SCHNEPF N.; BADONNEL R.; LAHMADI A.; MERZ S.	30
BABBAR H.; RANI S.	43
ALYW.H.F.	46
MOHAMMADI R.; JAVIDAN R.	46
LIN P.; BI J.; HU H.	54
SINGH D.; NG B.; LAI Y.-C.; LIN Y.-D.; SEAHW.K.G.	55
DEB R.; ROY S.	57
ASSEFAB.G.; OZKASAP O.	61
SHIRMARZ A.; GHAFARI A.	113
BADOTRA S.; PANDA S.N.	137
SWAMI R.; DAVE M.; RANGA V.	149
MIZRAHI T.; MOSES Y.	175
LATAH M.; TOKER L.	183
KILLIB.P.R.; RAOS.V.	279
KARAKUS M.; DURRESI A.	619



5.14 Ταξινόμηση ομάδων συγγραφέων βάση των αναφορών

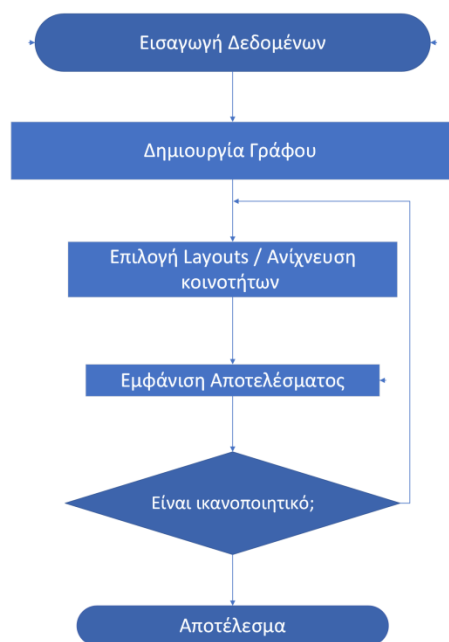
5.2 Ανάλυση δεδομένων μέσα από το λογισμικό Gephi

Το περιβάλλον εργασίας του Gephi είναι λίγο πιο πολύπλοκο σε σχέση με του VOSViewer όπως φαίνεται και στην Εικόνα 5.15. Ο αναλυτής αφού εισάγει τα δεδομένα αντικρίζει ένα γράφημα από κόμβους και ακμές (nodes&edges) στην περιοχή εργασίας. Από το toolbar που βρίσκεται στην κορυφή του δίνεται πρόσβαση στις βασικές λειτουργίες του προγράμματος. Τα τρία διαθέσιμα παράθυρα Overview, Data Laboratory και Preview επιτρέπουν στον χρήστη να έχει πρόσβαση στις επιμέρους ενέργειες που διαθέτει το καθένα όπως είναι η επιλογή κατάλληλων αλγορίθμων για την διάταξη του γράφου, αφαίρεση κόμβων ή ακμών, προεπισκόπηση του γράφου κ.α.



Εικόνα 5.15 Περιβάλλον εργασίας Gephi

Στο διάγραμμα ροής που ακολουθεί γίνεται προσπάθεια να εξηγηθεί η διαδικασία που γίνεται στο Gephi για να προκύψουν τα τελικά αποτελέσματα.



Εικόνα 5.16 Διάγραμμα ροής Gephi

5.2.1 Ανάλυση της σχέσης μεταξύ “Authors& Citedby”

Το Gephi θα δημιουργήσει έναν γράφο ή αλλιώς ένα δίκτυο συνδέσεων που οι κόμβοι (nodes) θα εκπροσωπούν τους συγγραφείς και οι ακμές (edges) θα δείχνουν τις συνδέσεις.

Οι συνδέσεις αντικατοπτρίζουν τις αναφορές (ή την επιρροή) που λαμβάνει ο κόμβος (συγγραφέας) από άλλους κόμβους (συγγραφείς).

Μέσα από τον αλγόριθμο “Modularity” και το Layout “ForceAtlas 2” δημιουργήθηκαν 1643 ομάδες στα δεδομένα μας. Η τιμή της τμηματικότητας (modularity) μπορεί να πάρει τιμές από [-1,1]. Ο υπολογισμός της γίνεται μέσα από τον τύπο [21]:

$$Q = \frac{1}{2m} \sum_{ij} (A_{ij} - \frac{k^i k^j}{2m}) \delta(C^i, C^j)$$

όπου:

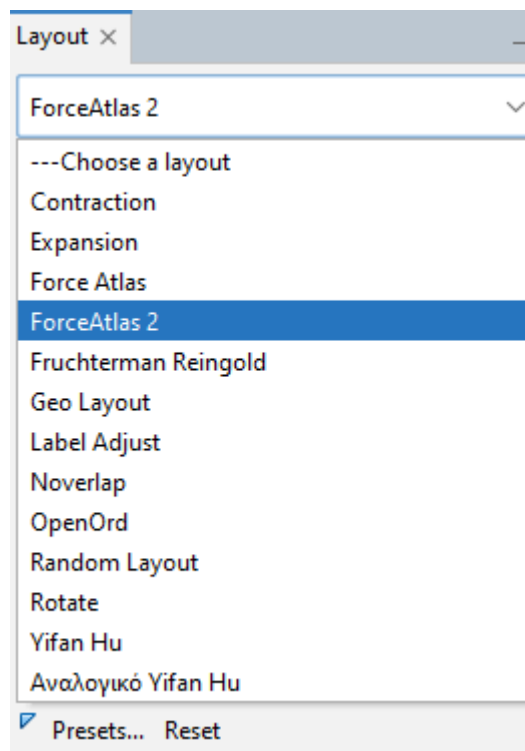
A: ο πίνακας γειτνίασης (adjacency matrix)

m: ο αριθμός συνδέσεων μέσα στο δίκτυο

k_i : ο βαθμός του κόμβου i (πόσες συνδέσεις έχει ο κόμβος i προς άλλους κόμβους)

C_i : ο τύπος του κόμβου i (η κοινότητα στην οποία ανήκει ο κόμβος i)

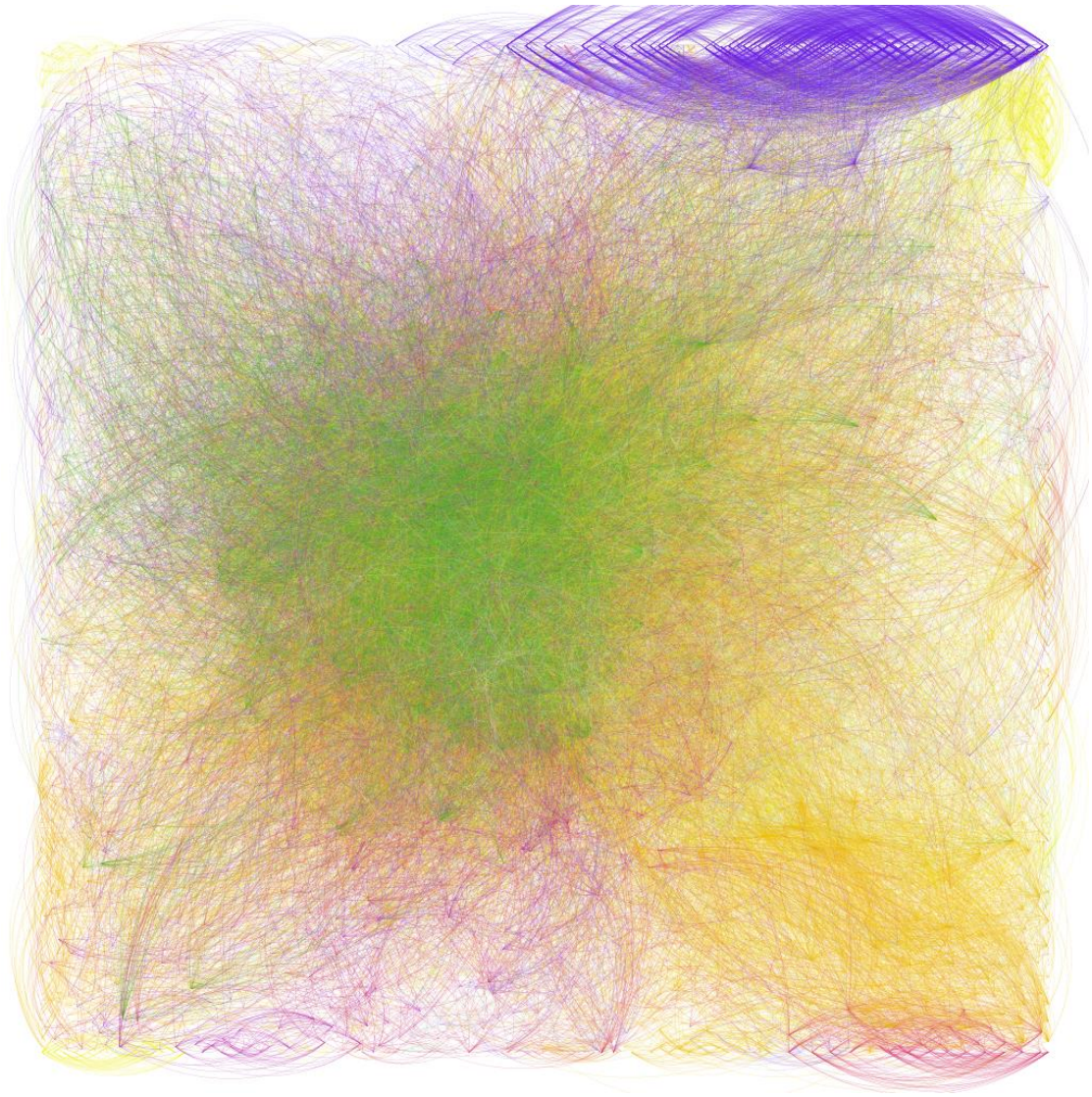
$\delta(C^i, C^j)$: η συνάρτηση δέλτα του Κρόνεκερ στην οποία ισχύει ότι $\delta(x,y) = 1$ αν $x=y$ ή 0 για διαφορετική περίπτωση.



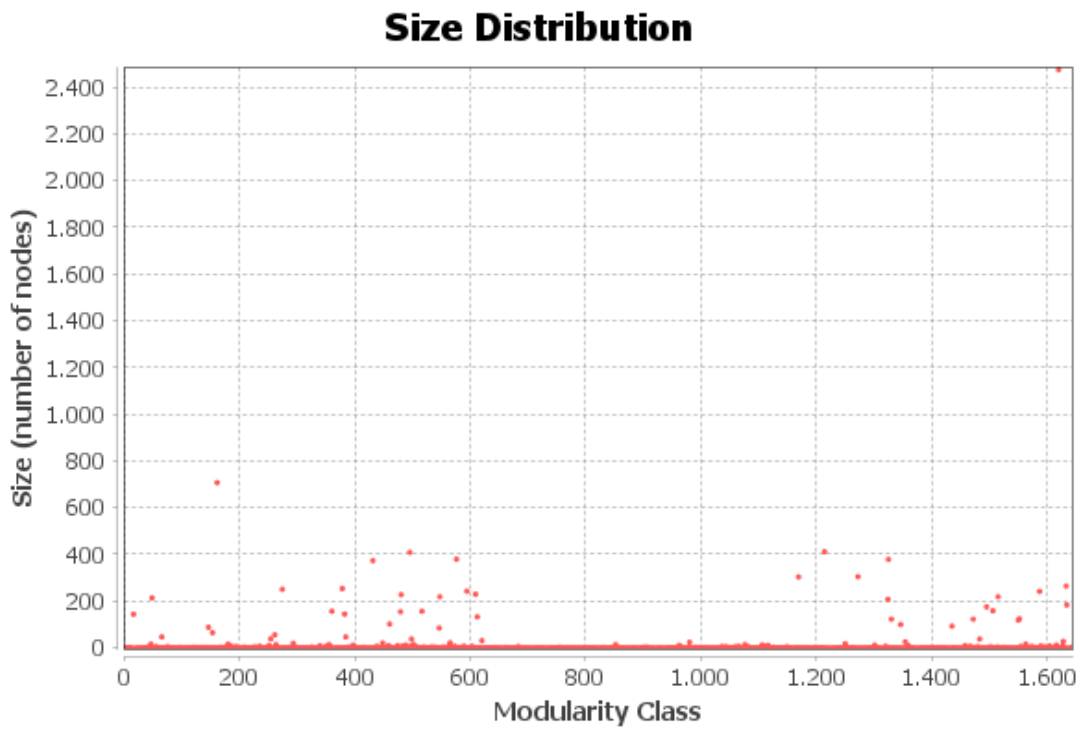
Εικόνα 5.17 Επιλογή κατάλληλου Layout

Για να είναι πιο ευδιάκριτος ο γράφος χρησιμοποιήθηκε το Layout ForceAtlas 2 όπου χρησιμοποιώντας τον Force-Directed αλγόριθμο κρατά τους κόμβους που συνδέονται κοντά και απωθεί αυτούς που δεν συνδέονται δημιουργώντας με αυτόν τον τρόπο μία απόσταση

μεταξύ τους. Η οπτικοποίηση του δικτύου φαίνεται στην Εικόνα 5.18 όπου τα διαφορετικά χρώματα διευκολύνουν τον αναλυτή να ξεχωρίζει τις κοινότητες. Επειδή ο γράφος διαθέτει πολλές ομάδες δικτύων, δεν είναι εφικτό να ξεχωρίζονται απευθείας για αυτό στο διάγραμμα της Εικόνας 5.19 γίνεται μια αναπαράσταση της κατανομής του πλήθους των κόμβων σε σχέση με τις κλάσεις (modularityclass).

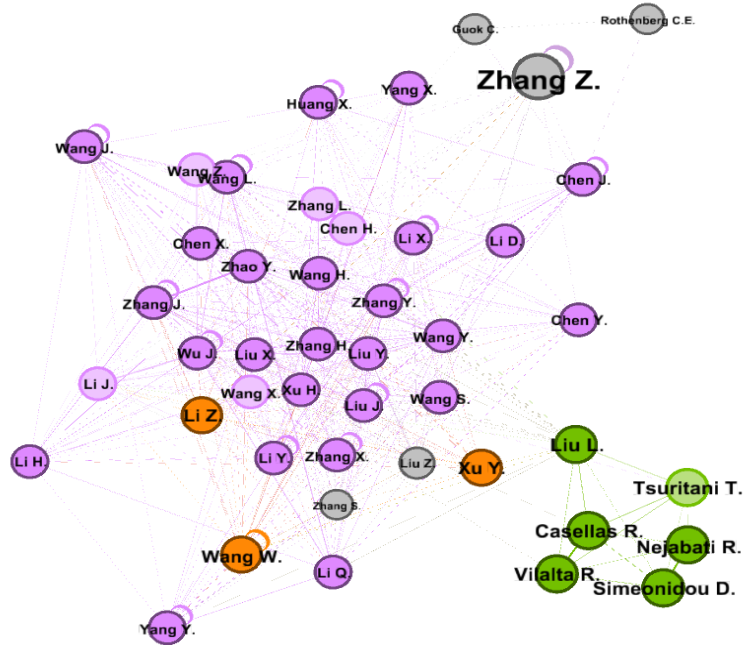


Εικόνα 5.18Χρωματισμός κοινότητων



Εικόνα 5.19 Διανομή κοινοτήτων

Χρησιμοποιώντας το φίλτρο Degree Range επιλέξαμε να ασχοληθούμε με κόμβους που έχουν ένα σχετικά υψηλό επίπεδο συνδεσιμότητας. Πιο συγκεκριμένα, το φίλτρο αυτό εμφάνισε ένα εύρος από 0 συνδέσεις έως 292 συνδέσεις όπου καλύπτουν όλο το φάσμα του δικτύου. Πειραματικά ορίστηκε το range να πάρει τις τιμές 100 έως 292 με σκοπό ο χάρτης να μικρύνει και να εμφανιστούν οι συνδέσεις που έχουν υψηλή επίδραση στο δίκτυο όπως φαίνεται στην Εικόνα 5.20.



Εικόνα 5.20 Αποτέλεσμα φίλτρου "degree range 100-292"

Στον πίνακα 4 δίνονται ορισμένες πληροφορίες που αφορούν τους γράφους των εικόνων 5.18 και 5.20.

Πίνακας 4 Πληροφορίες γράφων

Range	0-292	100-292
Nodes	17481	45
Edges	60829	451
Modularity	0.813	0.284
Num. of communities	1643	5
Average Degree	6.971	20.044
Avg. Weighted Degree	9.516	57.778
Network Diameter	19	4
Graph Density	0	0.456
Connected Components	1582	1

Nodes: Είναι οι κόμβοι που αντιπροσωπεύουν τους συγγραφείς.

Edges: Οι συνδέσεις που υπάρχουν μεταξύ των κόμβων.

Modularity: Προσδιορίζει την ποιότητα της δομής του γράφου, δηλαδή είναι μέτρο που αναφέρεται στο κατά πόσο το δίκτυο (γράφος) είναι σπονδυλωτός. Μία υψηλή τιμή (κοντά στο 1) δείχνει ότι το δίκτυο έχει πολλές και πυκνές κοινότητες με στενές συνδέσεις μεταξύ τους.

Number of Communities: Είναι ο αριθμός των κοινοτήτων του γράφου.

Average Degree: Ο μέσος βαθμός συνδέσεων κάθε κόμβου.

Network Diameter: Αναφέρεται στην μεγαλύτερη απόσταση μεταξύ δύο κόμβων.

Network Density: Αναφέρεται στην πυκνότητα του δικτύου, δηλαδή στο ποσοστό των πιθανών συνδέσεων στο δίκτυο.

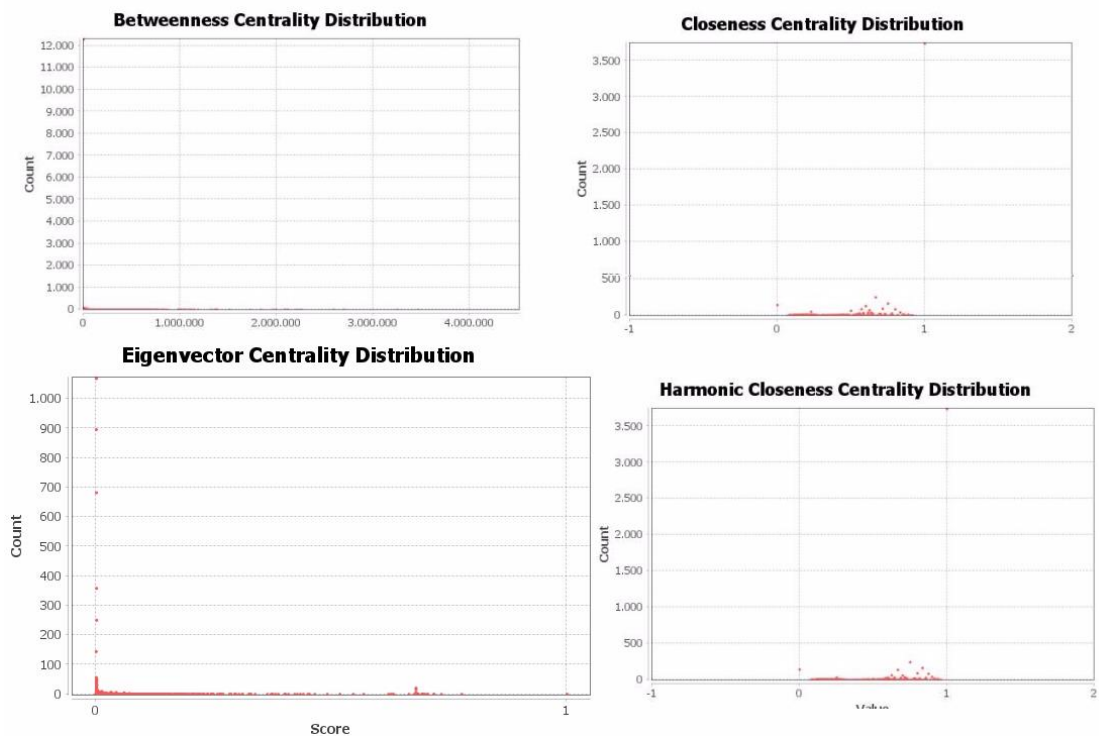
PageRank: Η πιθανότητα μετάβασης από έναν συγγραφέα σε έναν άλλο, επηρεάζοντας την σημασία των συνδέσεων στο δίκτυο.

Connected Components: Είναι τα συνδεδεμένα τμήματα που δημιουργούν το δίκτυο. Τα τμήματα αυτά είναι μικρά υποδίκτυα τα οποία δεν είναι απαραίτητο να συνδέονται μεταξύ τους.

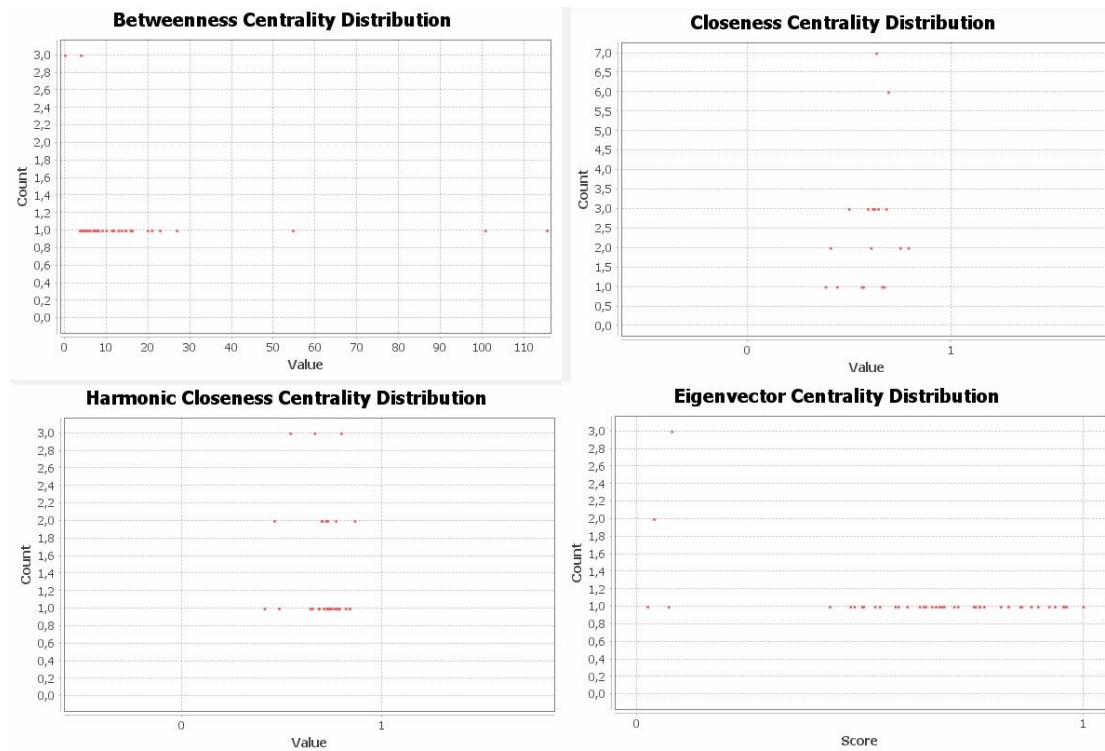
Το δίκτυο που έχει δημιουργηθεί με την χρήση του φίλτρου DegreeRange είναι πολύ μικρότερο υποδίκτυο του αρχικού. Με μόλις 45 κόμβους και 451 ακμές που αντιστοιχεί στο 0,26% των κόμβων και 0,74% των ακμών του αρχικού γράφου δημιουργούν 5 κοινότητες. Η συνδεσιμότητα φαίνεται να είναι πιο ισχυρή και πιθανόν αυτό συμβαίνει γιατί υπάρχει ένα υποσύνολο συγγραφέων με μεγαλύτερη επιρροή στο δίκτυο. Αυτό συμπεραίνεται από την αυξημένη τιμή στην πυκνότητα του δικτύου και του μέσου βαθμού.

Από την καρτέλα statistics εξήχθησαν οι δύο παρακάτω εικόνες (Εικόνα 5.21 και Εικόνα 5.22) όπου φαίνονται διαγραμματικά οι κεντρικότητες και για τους δύο γράφους που αναφέρθηκαν.

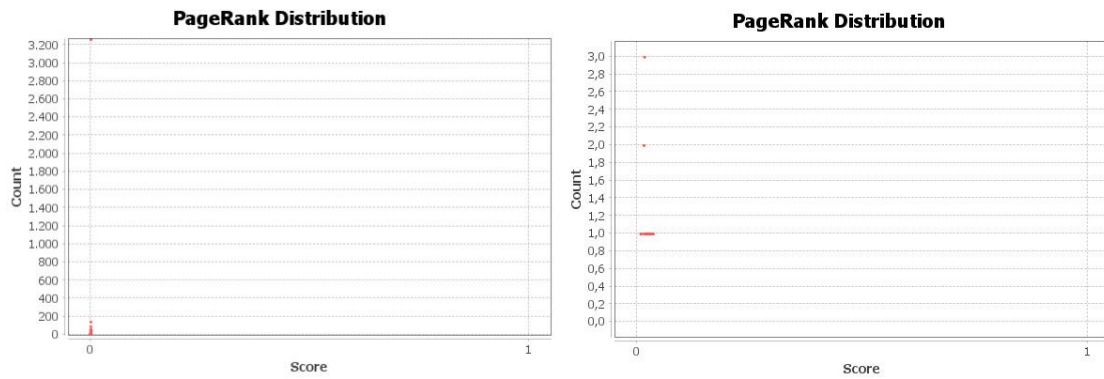
Επίσης απεικονίζεται το PageRank των γράφων στην Εικόνα 5.23 όπου το πρώτο διάγραμμα αντιπροσωπεύει τον πρώτο γράφο και το δεύτερο διάγραμμα τον γράφο που έχει επιλεγεί το εύρος. Έπειτα (Εικόνα 5.24) φαίνεται το δίκτυο των συγγραφέων που έχουν την υψηλότερη PageRank τιμή.



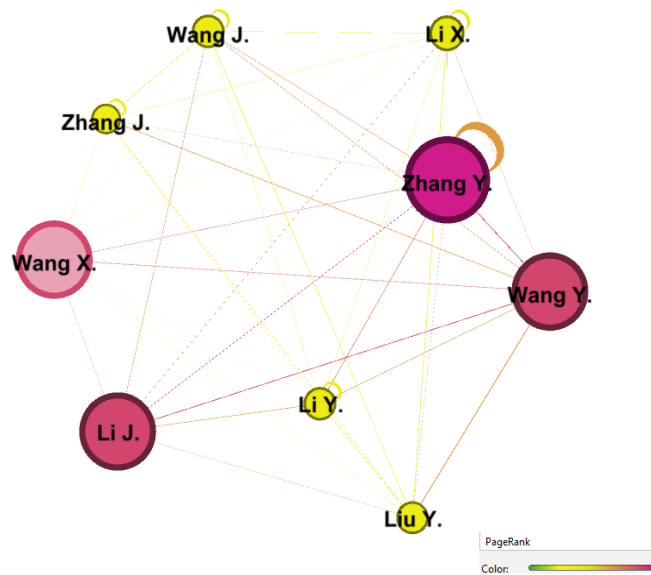
Εικόνα 5.21 Κεντρικότητες αρχικού γράφου



Εικόνα 5.22 Κεντρικότητες γράφου με Range 100-292



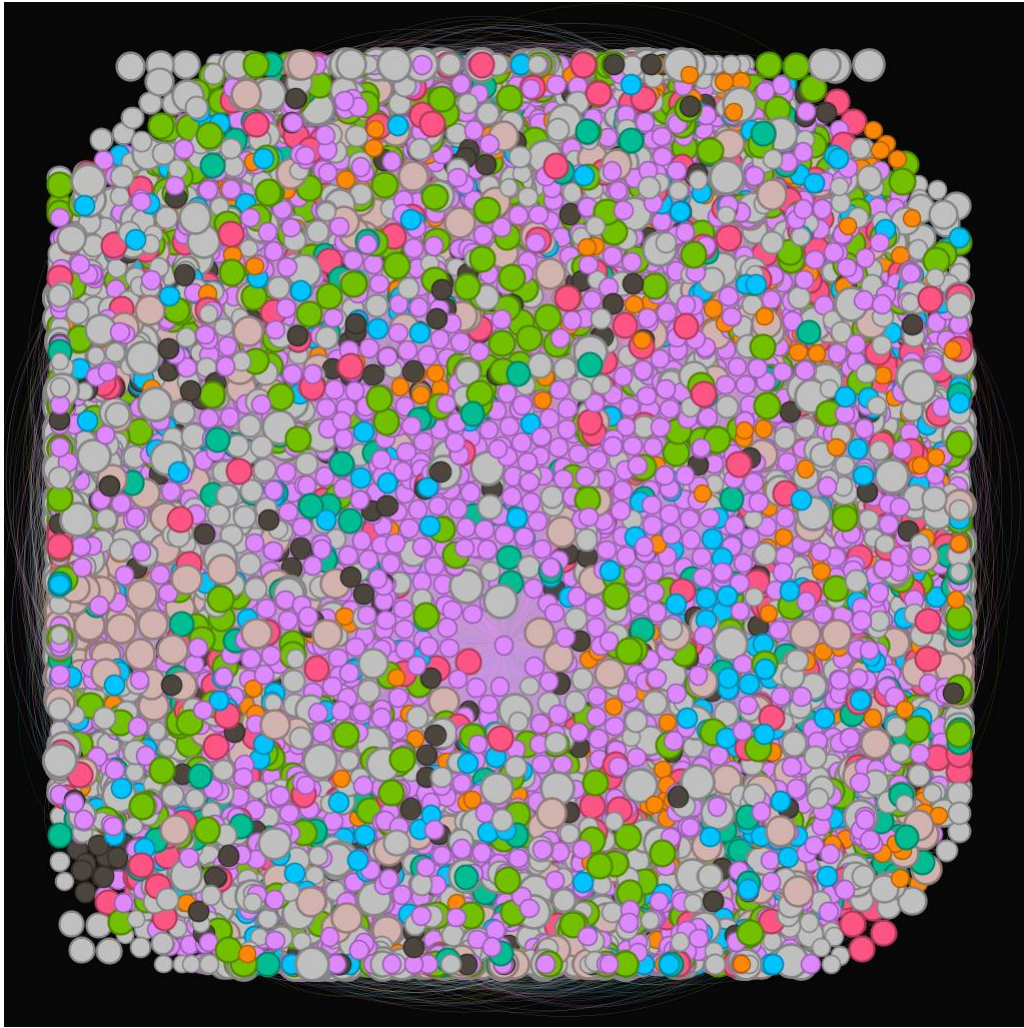
Εικόνα 5.23 PageRank γράφων



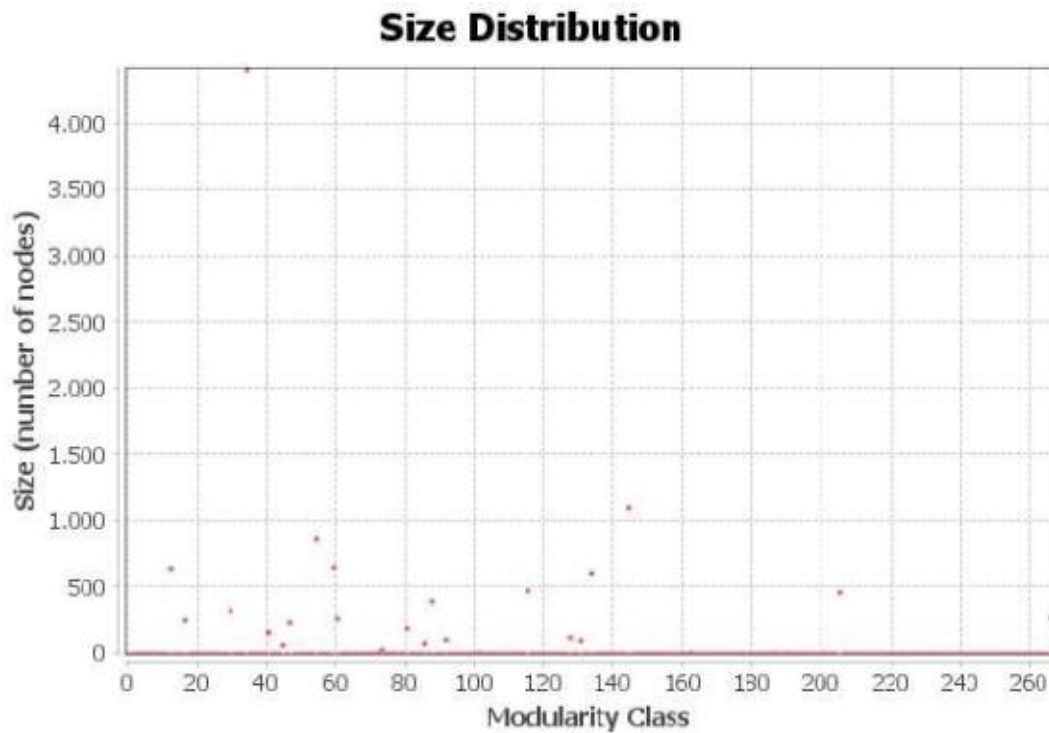
Εικόνα 5.24 Δίκτυο συγγραφέων με υψηλή τιμή PageRank

5.2.2 Ανάλυση της σχέσης μεταξύ “Authors&Author keywords”

Μελετώντας αυτή την σχέση πιθανολογούμε ότι ως προσδοκώμενο αποτέλεσμα θα είναι η ανακάλυψη ομάδων όπου λέξεις-κλειδιά θα αποκαλύπτουν τα ενδιαφέροντα και την μελέτη που έχει κάνει κάθε συγγραφέας στο επιστημονικό πεδίο των δικτύων. Στην Εικόνα 5.25 φαίνεται ο γράφος που έχει δημιουργηθεί, με τα χρώματα να αποτυπώνουν τα συμπλέγματα που έχουν δημιουργηθεί και στη συνέχεια (Εικόνα 5.26) απεικονίζεται διαγραμματικά η κατανομή τους.



Εικόνα 5.25 Απεικόνιση κοινοτήτων με διαφορετικά χρώματα



Εικόνα 5.26 Κατανομή κοινοτήτων

Οι πληροφορίες που πήραμε από τον γράφο είναι οι εξής:

Nodes:12875

Edges:63536

Modularity: 0.373

NumberofCommunities:267

Average Degree: 9.87

Network Diameter: 6

Network Density: 0.001

Connected Components: 43

Οι κεντρικότητες διαμορφώνονται όπως φαίνεται στην Εικόνα 5.27.

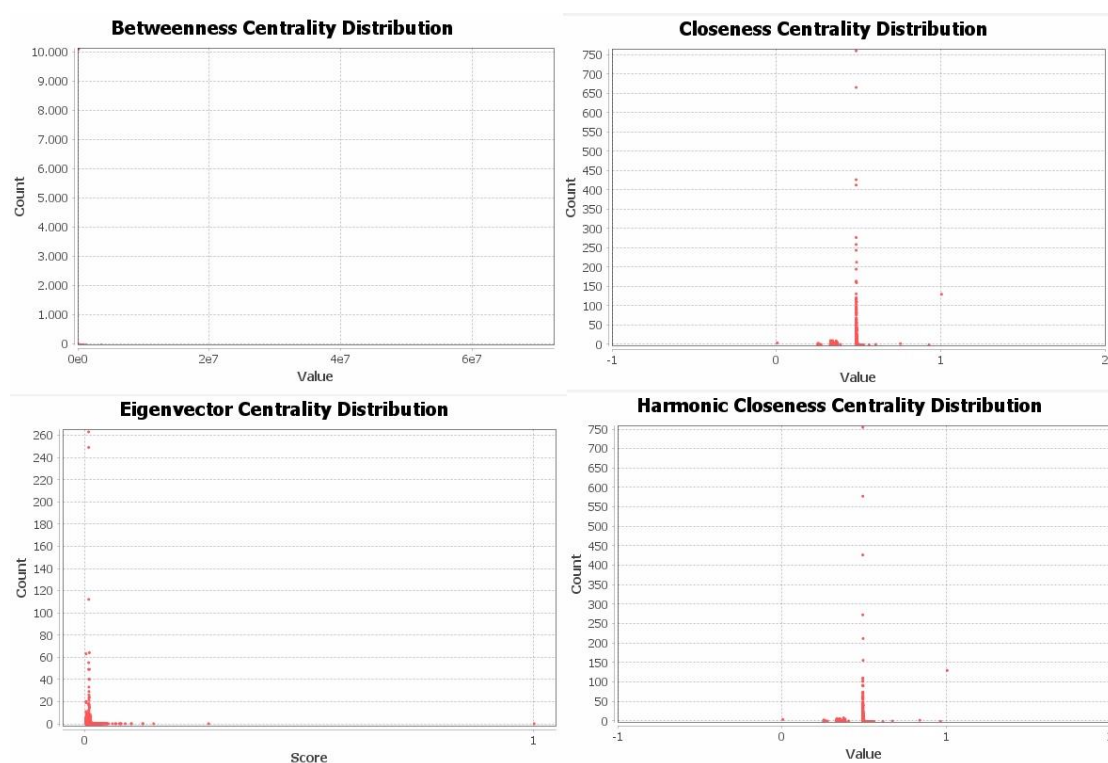
Το πρώτο διάγραμμα αναφέρεται στην κεντρικότητα μεσολάβησης όπου παρατηρείται η συγκέντρωση των κόμβων στην αρχή. Αυτό σημαίνει ότι οι περισσότεροι κόμβοι δε συμμετέχουν στα μονοπάτια επικοινωνίας που έχει το δίκτυο γι αυτό το λόγω και η συγκεκριμένη κεντρικότητα είναι χαμηλή.

Στο δεύτερο διάγραμμα η κεντρικότητα εγγύτητας μας δείχνει ότι οι κόμβοι είναι σχετικά κοντά κάνοντας το δίκτυο ομοιόμορφο. Αυτό το συμπέρασμα βγήκε διότι οι περισσότεροι

κόμβοι είναι συγκεντρωμένοι σε μία τιμή του άξονα x που συνεπάγεται ότι η τιμή της κεντρικότητας closeness είναι σχεδόν ίδια για τους κόμβους.

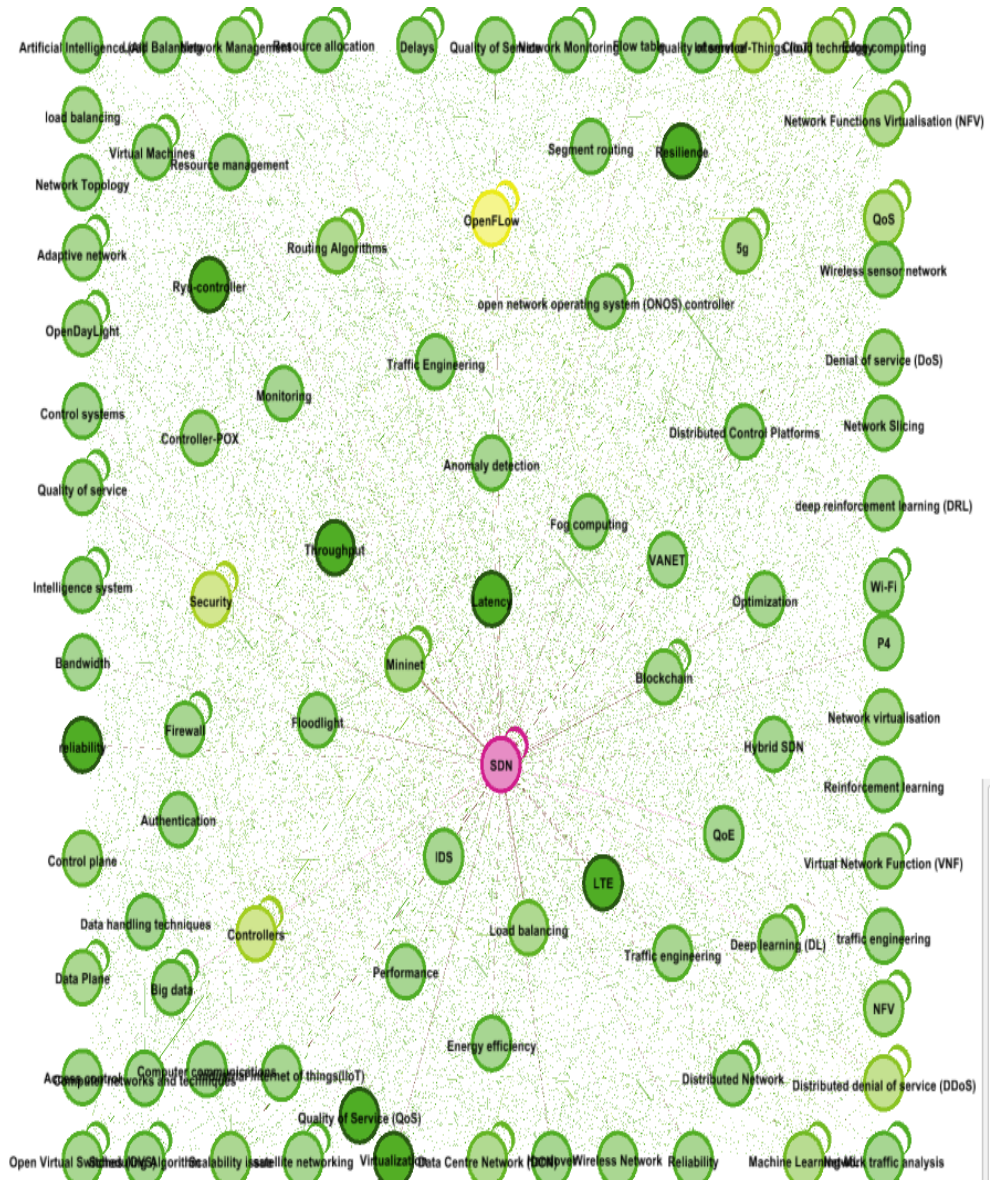
Το τρίτο διάγραμμα μοιάζει πολύ με το πρώτο. Εδώ οι περισσότεροι κόμβοι βρίσκονται στην αρχή, με την κεντρικότητα ιδιοδιανύσματος να είναι χαμηλή. Αυτό γίνεται γιατί οι κόμβοι δεν είναι καλά συνδεδεμένοι με τους κόμβους που έχουν σημαντική επιρροή.

Τέλος, στο διάγραμμα την κεντρικότητας αρμονικής εγγύτητας παρατηρούμε ότι οι κόμβοι συγκεντρώνονται μεταξύ από τις τιμές 0 και 1. Αυτό θα μπορούσαμε να το μεταφράσουμε πως οι πλειοψηφία των κόμβων είναι καλά συνδεδεμένοι διότι έχουν σύντομες αποστάσεις μεταξύ τους μέσα στο δίκτυο.

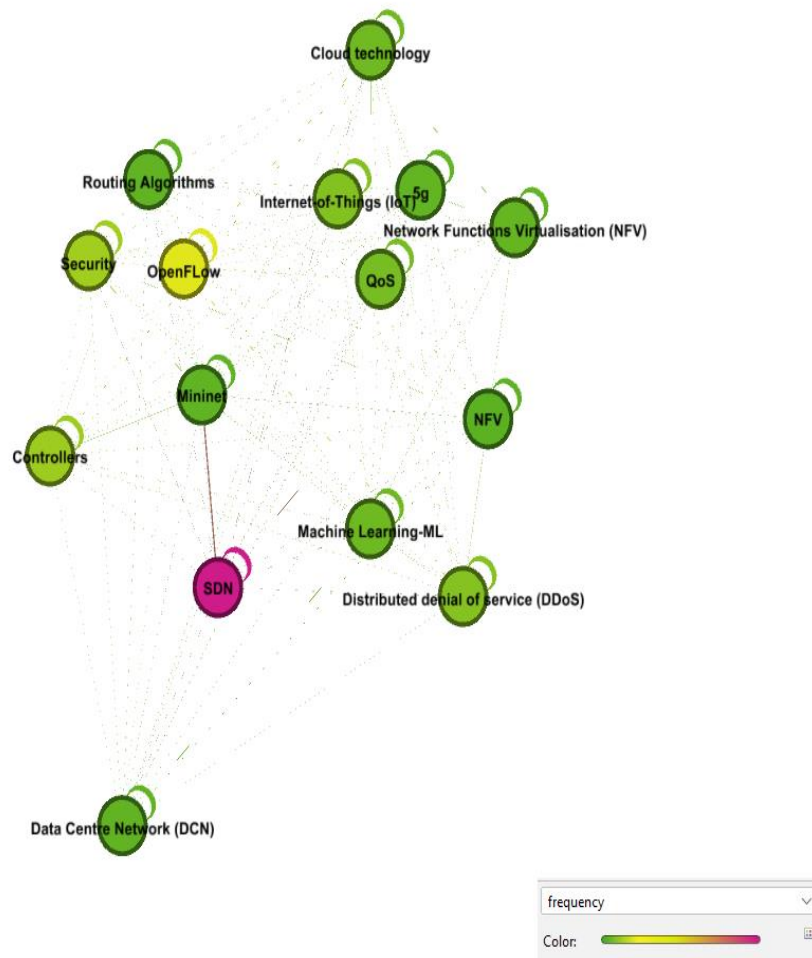


Εικόνα 5.27 Κεντρικότητες αρχικού γράφου

Όπως μπορούμε να διακρίνουμε στην Εικόνα 5.25 η οπτικοποίηση του γράφου είναι ιδιαίτερα πυκνή που δυσκολεύει την μελέτη της. Χρησιμοποιώντας πιο αραιή αναπαράσταση μέσα από τα Layout και επιλέγοντας από την καρτέλα “Filters” το “DegreeRange” με αρχικό εύρος «100-11693» και έπειτα με εύρος «400-11693» θα προσπαθήσουμε να εστιάσουμε σε πιο σημαντικούς κόμβους που λόγο του εύρους που έχουμε επιλέξει θα έχουν και περισσότερες συνδέσεις.

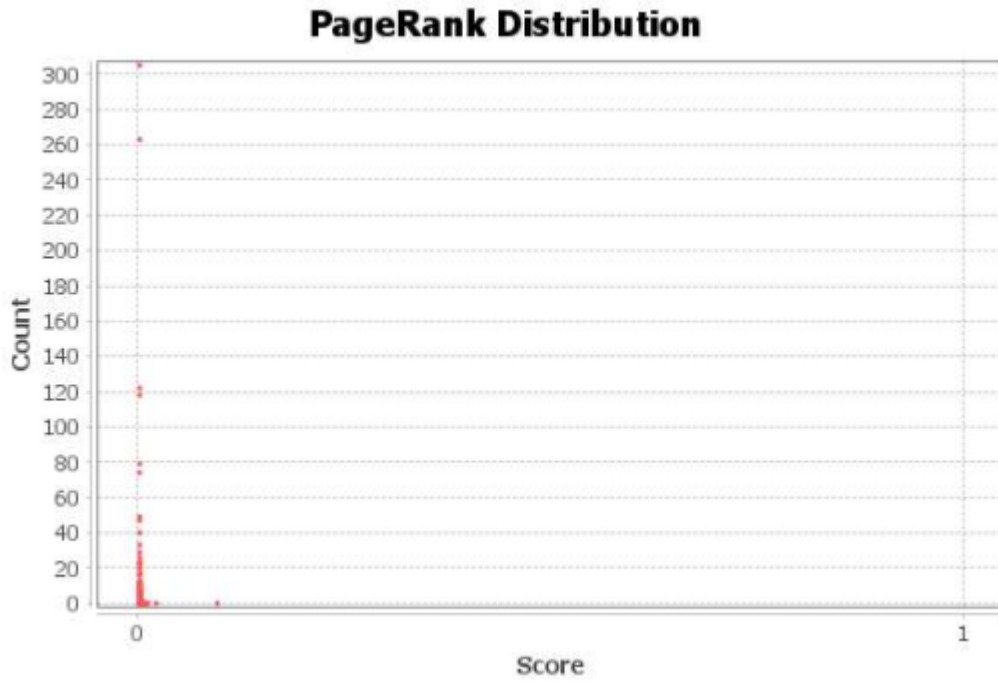


Εικόνα 5.28 Αποτελέσματα "DegreeRange 100-11693"

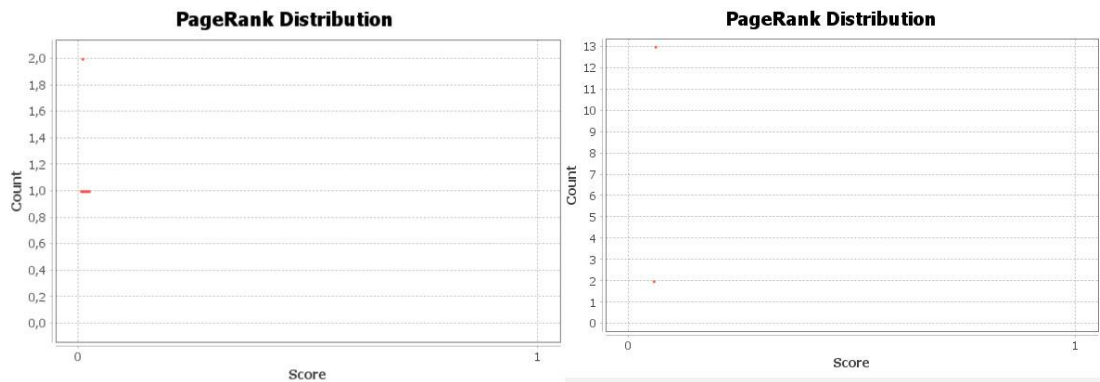


Εικόνα 5.29 Αποτελέσματα " Degree Range 400-11693"

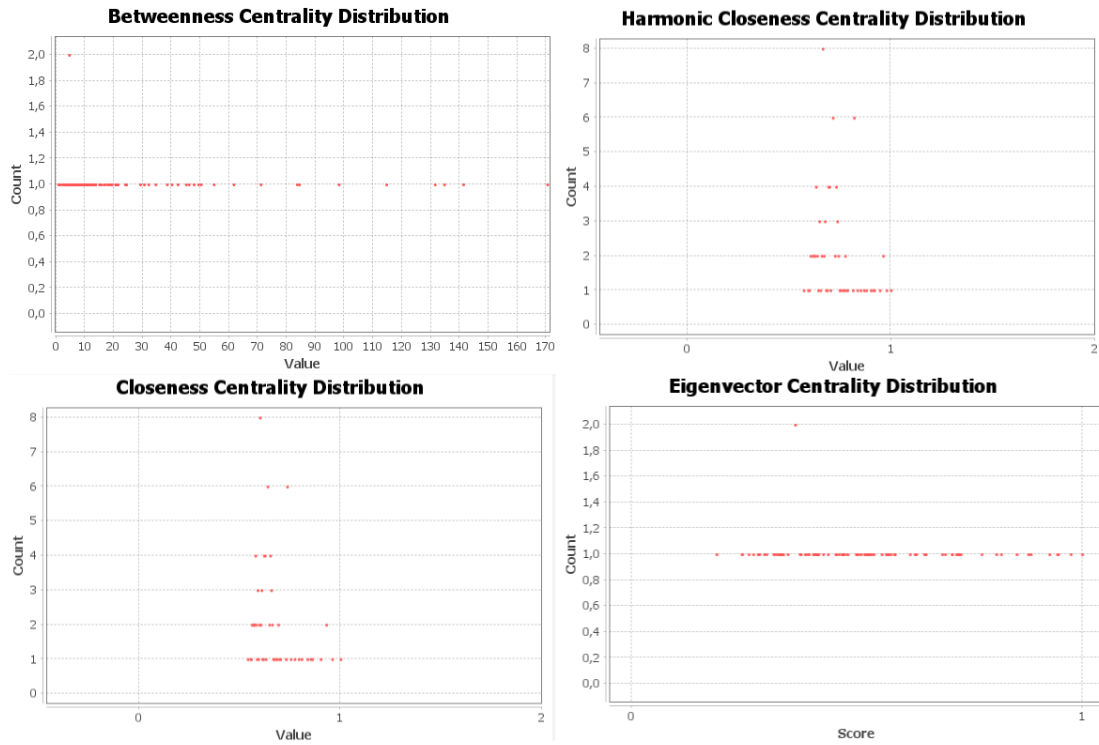
Και στις δύο εικόνες (Εικόνα 5.28 & Εικόνα 5.29) φαίνεται ο γράφος που δημιουργήθηκε με βάση το εύρος. Στην πρώτη το εύρος είναι από 100 έως 11693 και στην δεύτερη από 400 έως 11693 και τα χρώματα του γράφου να χαρακτηρίζουν την συχνότητα της κάθε λέξης-κλειδί. Στην συνέχεια (Εικόνα 5.30 - Εικόνα 5.33) απεικονίζονται τα διαγράμματα των κεντρικότητων και το PageRank.



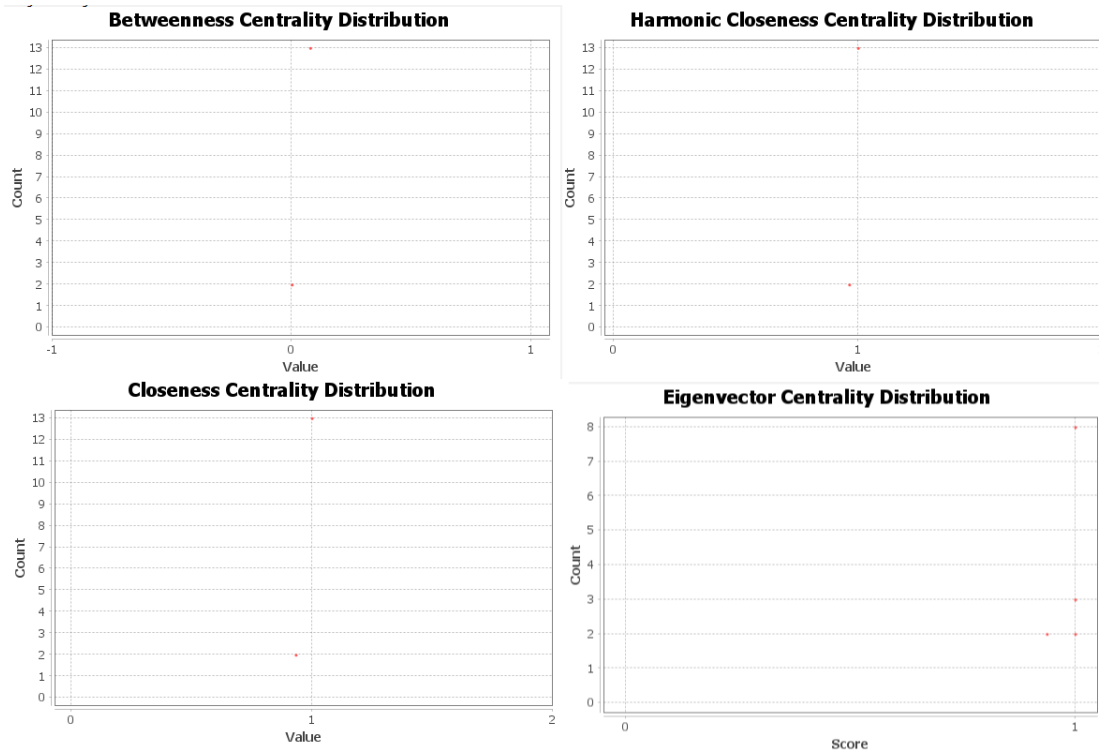
Εικόνα 5.30 PageRank του αρχικού δικτύου



Εικόνα 5.31 PageRank με range 100-11693 & 400-11693



Εικόνα5.32 Centralities Range=100-11693



Εικόνα5.33 Centralities Range=400-11693

Οι πληροφορίες που μας δίνουν οι δύο γράφοι φαίνονται στον παρακάτω πίνακα (Πίνακας 5.5).

Πίνακας 5.5 Πληροφορίες γράφων

Range	0-11693	100-11693	400-11693
Nodes	12875	92	15
Edges	63536	1914	119
Modularity	0.373	0.161	0.323
Num. of communities	267	4	2
Average Degree	9.87	41.609	15.867
Avg. Weighted Degree	11.062	80.37	35.333
Network Diameter	6	2	2
Graph Density	0.001	0.457	1.133
Connected Components	43	1	1

5.3 Ανάλυση δεδομένων με την χρήση της R language

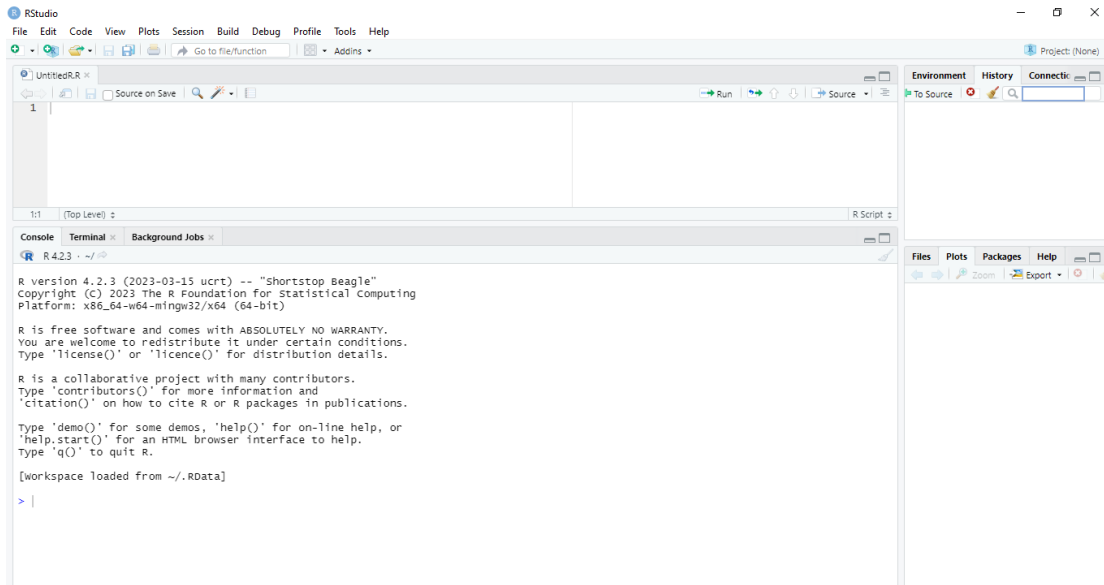
Μέσα από το RStudio και την R 4.2.3 θα γίνει μία ανάλυση των βιβλιομετρικών δεδομένων που χρησιμοποιούμε στην παρούσα εργασία. Στην Εικόνα 5.34 απεικονίζεται το περιβάλλον εργασίας του RStudio. Όπως φαίνεται αποτελείται από τέσσερα παράθυρα.

Το πρώτο παράθυρο είναι η κύρια περιοχή για την συγγραφή και την εκτέλεση κώδικα.

Το δεύτερο παράθυρο διαθέτει την κονσόλα όπου ο χρήστης μπορεί να εκτελέσει ή να δοκιμάσει άμεσα εντολές και να δει τα αποτελέσματά τους.

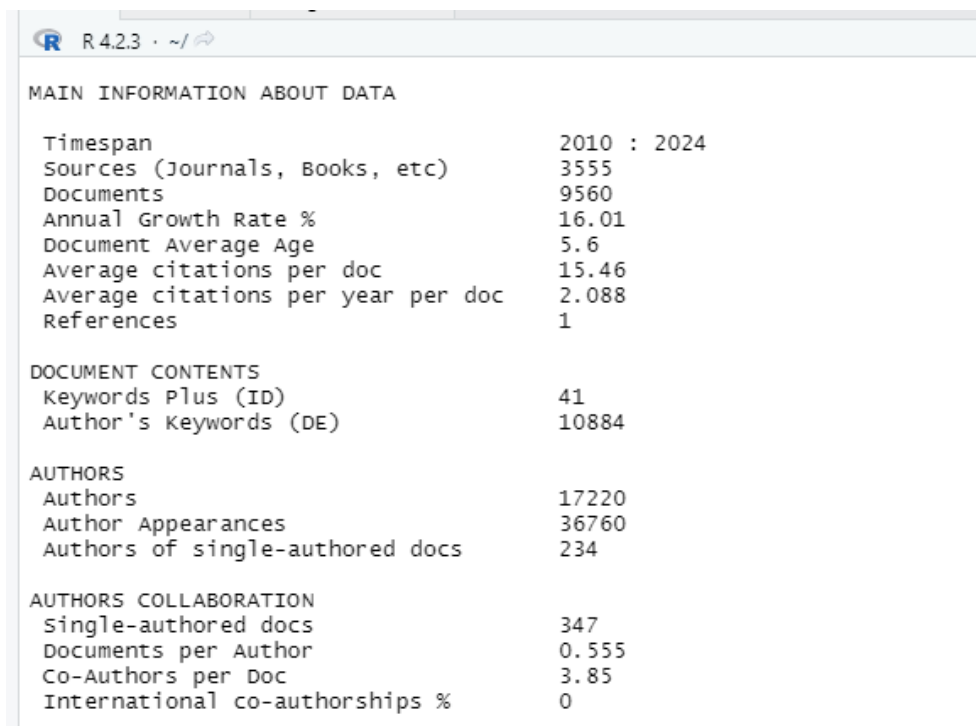
Το τρίτο παράθυρο εμφανίζει το ιστορικό των εντολών που έχουν εκτελεστεί στην κονσόλα.

Στο τέταρτο και τελευταίο παράθυρο εμφανίζονται τα διαγράμματα που δημιουργεί ο χρήστης.



Εικόνα 5.34 Περιβάλλον RStudio

Στην Εικόνα 5.35 εμφανίζονται περιγραφικές πληροφορίες για τα δεδομένα που εξήχθησαν με την βοήθεια της R. Όπως φαίνεται αυτές οι πληροφορίες είναι γενικές και προσδιορίζουν το χρονικό διάστημα των δεδομένων μας, το είδος των πηγών, και στατιστικά στοιχεία όπως μέσους όρους παραπομπών, εμφάνιση συγγραφέων, συν-συγγραφή, κ.ά.

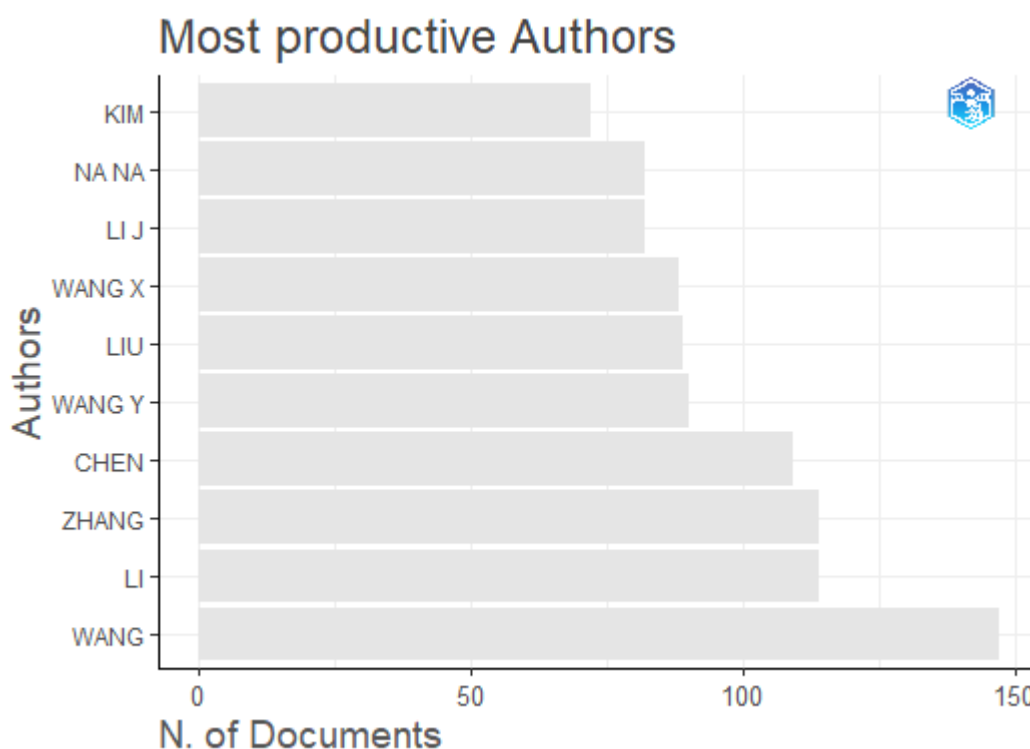


Εικόνα 5.35 Γενικές πληροφορίες από τα δεδομένα

Στο κομμάτι της προεπεξεργασίας των βιβλιομετρικών δεδομένων που έγινε μέσα από την R, σημαντικός ήταν ο ρόλος της λημματοποίησης (Lemmatization) που περιορίζει της διαφορετικές μορφές μιας λέξης που θα μελετηθεί (λ.χ. SDN/Software Defined Networks/ Software Defined Networking). Επιπλέον ενέργειες που έγιναν ήταν για να αφαιρεθούν διπλότυποι ορισμοί, πιθανές ασάφειες, απομάκρυνση κειμένου που δεν έχει σημασία. Η εναρμόνιση πεζής και κεφαλαίας γραμματοσειράς, η επιλογή να μείνουν οι πλήρης λέξεις και η αφαίρεση των παυλών είναι μερικές από τις ενέργειες αυτές.

Στη συνέχεια φαίνονται οι δέκα πιο παραγωγικοί συγγραφείς από την οπτική του Rplot (Εικόνα 5.36) και από την οπτική του RStudio(Εικόνα 5.37).

Στο γράφημα της εικόνας 5.38 όπου είναι η ετήσια επιστημονική παραγωγή και στην Εικόνα 5.39 απεικονίζονται οι πληροφορίες από το περιβάλλον του λογισμικού. Είναι φανερό ότι υπάρχει μία αυξητική τάση μέχρι το 2018 όπου εκείνο το έτος υπάρχει το peak πάνω από 1200 δημοσιευμένα άρθρα και έπειτα ξεκινά να φθίνει.

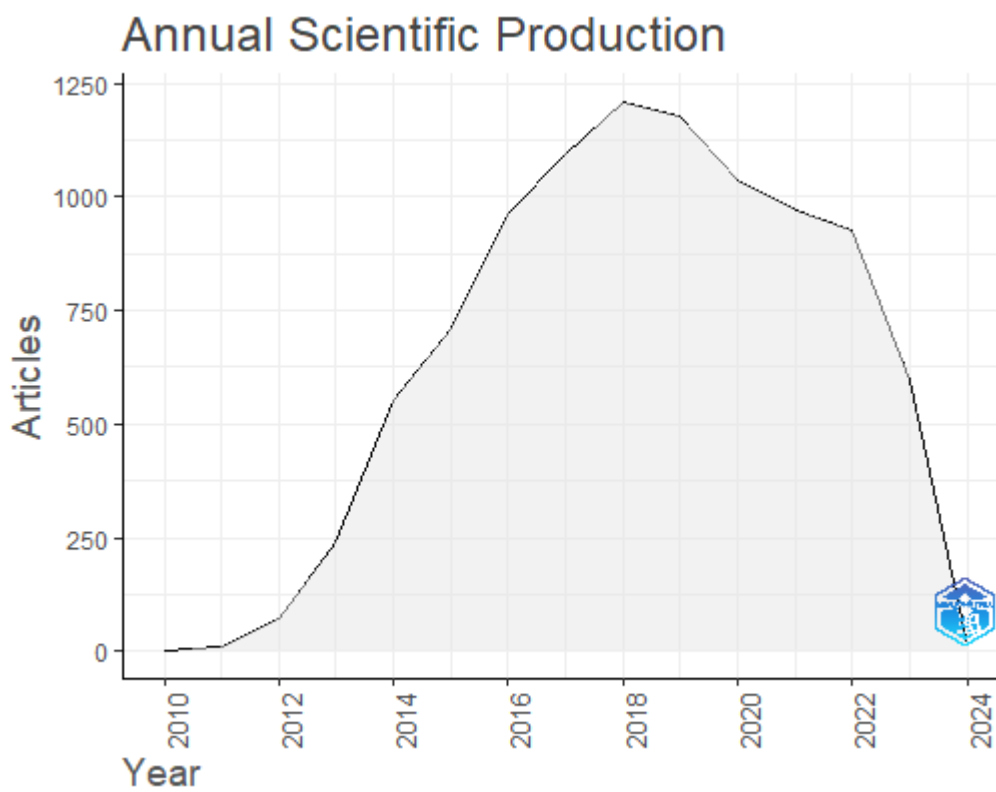


Εικόνα 5.36 Οι 10 πιο παραγωγικοί συγγραφείς

Most Productive Authors

	Authors	Articles	Authors	Articles Fractionalized
1	WANG	147	NA NA	82.0
2	LI	114	WANG	44.1
3	ZHANG	114	CHEN	30.7
4	CHEN	109	LI	29.2
5	WANG Y	90	ZHANG	29.1
6	LIU	89	LIU	24.1
7	WANG X	88	KIM	23.4
8	LI J	82	WANG Y	20.9
9	NA NA	82	WANG X	20.0
10	KIM	72	XU	19.9

Εικόνα 5.37 Οι 10 πιο παραγωγικοί συγγραφείς στο RStudio



Εικόνα 5.38 Ετήσια επιστημονική παραγωγή

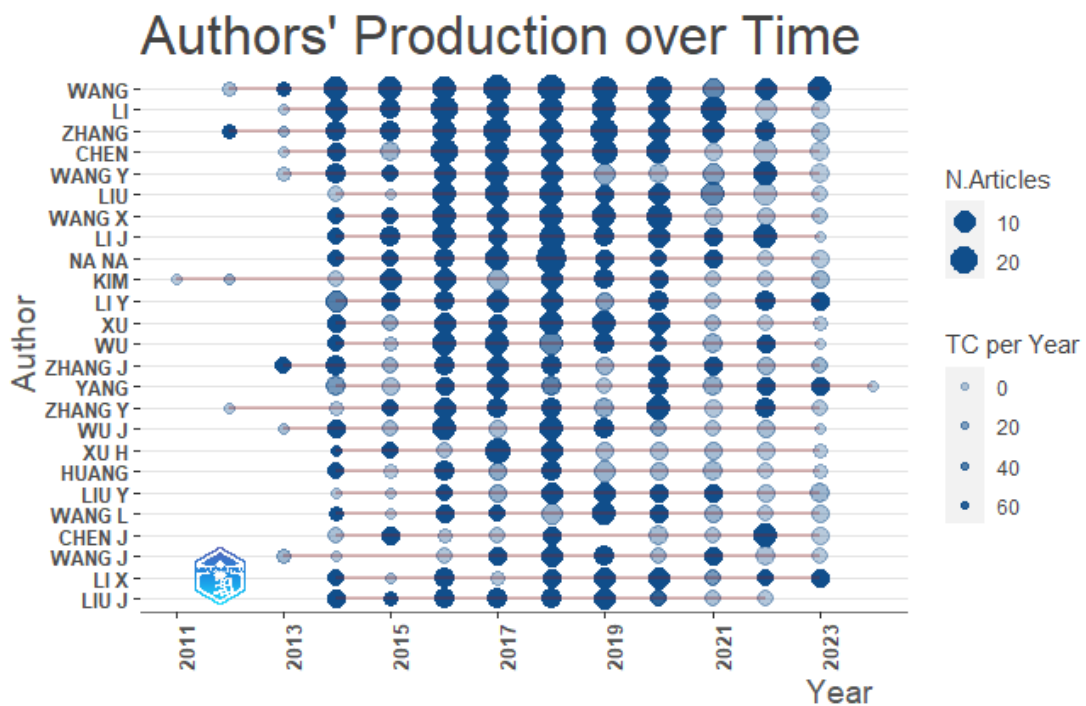
Annual Scientific Production

Year	Articles
2010	1
2011	9
2012	72
2013	237
2014	553
2015	709
2016	960
2017	1092
2018	1209
2019	1177
2020	1037
2021	972
2022	926
2023	596
2024	8

Annual Percentage Growth Rate 16.01

Εικόνα 5.39 Ετήσια επιστημονική παραγωγή RStudio

Ορίζοντας το πρόγραμμά μας να εμφανίσει τους 25 πιο παραγωγικούς συγγραφείς από το 2011 μέχρι το 2023 πήραμε τα αποτελέσματα που φαίνονται παρακάτω (Εικόνα 5.40 & Εικόνα 5.41).

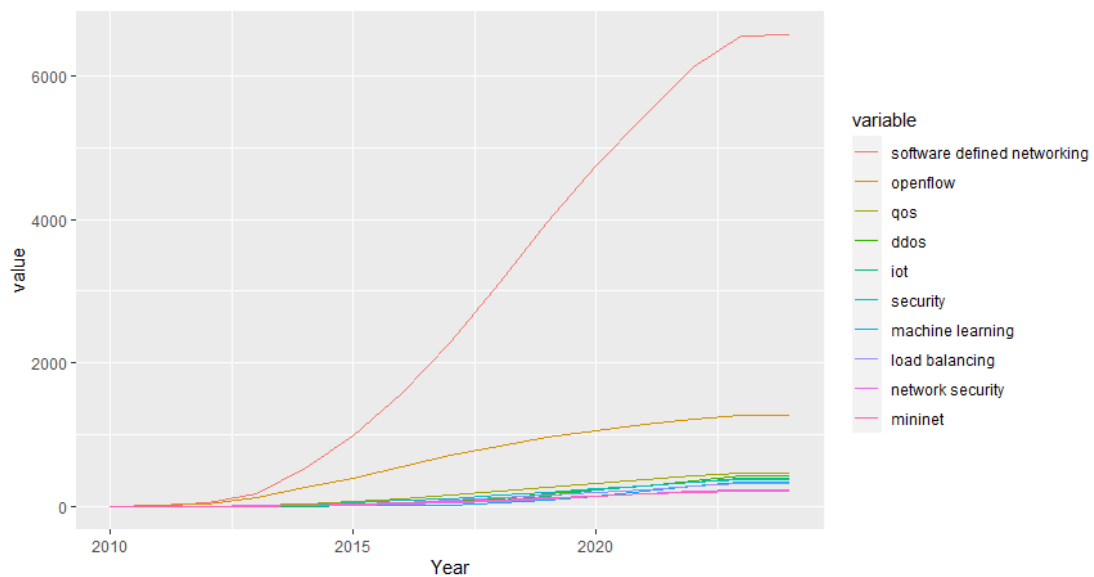


Εικόνα 5.40 Οι 25 πιο παραγωγικοί συγγραφείς σε σχέση με τον χρόνο

	Author	Dominance Factor	Tot Articles	Single-Authored	Multi-Authored	First-Authored	Rank by Articles	Rank by DF
1	ZHANG X	0.4473684	38	0	0	38	17	23
2	LI H	0.4042553	47	0	0	47	19	15
3	XU H	0.3928571	57	1	0	56	22	8
4	ZHAO Y	0.3750000	40	0	0	40	15	21
5	CHEN J	0.3673469	49	0	0	49	18	11
6	WANG J	0.3673469	49	0	0	49	18	11
7	ZHANG Y	0.3636364	66	0	0	66	24	6
8	WANG H	0.3409091	44	0	0	44	15	17
9	WANG S	0.3191489	47	0	0	47	15	15
10	WANG W	0.3170732	41	0	0	41	13	19
11	WANG L	0.2592593	54	0	0	54	14	9
12	WANG Y	0.2555556	90	0	0	90	23	1
13	LI Z	0.2500000	44	0	0	44	11	17
14	LI Q	0.2500000	40	0	0	40	10	21
15	LIU Y	0.2407407	54	0	0	54	13	9
16	LI Y	0.2394366	71	0	0	71	17	4
17	LI J	0.2317073	82	0	0	82	19	3
18	LI X	0.2291667	48	0	0	48	11	13
19	LIU J	0.2291667	48	0	0	48	11	13
20	HU Y	0.1951220	41	0	0	41	8	19
21	CASELLAS R	0.1891892	37	0	0	37	7	24
22	WANG Z	0.1891892	37	0	0	37	7	24
23	WANG X	0.1704545	88	0	0	88	15	2
24	ZHANG J	0.1343284	68	1	0	67	5	9
25	WU J	0.1311475	61	0	0	61	8	7

Εικόνα 5.41 Κατάταξη των 25 πιο παραγωγικών συγγραφέων στο RStudio

Στη συνέχεια, στο χρονικό εύρος μιας δεκαετίας (από το 2011 μέχρι το 2024) απεικονίζονται οι λέξεις κλειδιά και η μεταβλητότητά τους σε σχέση με τον χρόνο (Εικόνα 5.42 & Εικόνα 5.43).



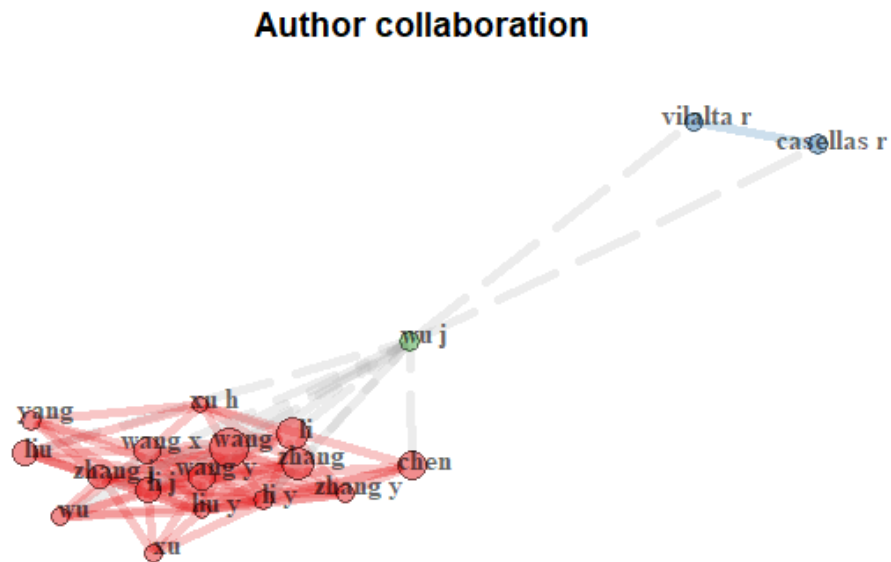
Εικόνα 5.42 Χρονική εξέλιξη των λέξεων-κλειδιών

Year	software	defined	networking	openflow	qos	ddos	iot	security	machine	learning	load balancing	network	security	mininet
1 2010		1	1	0	0	0	0	0		0	0	0	0	0
2 2011		5	6	0	0	0	0	0		0	0	1	1	0
3 2012		47	38	2	0	0	2	2		0	1	1	1	0
4 2013		174	114	9	0	0	7	7		0	4	3	3	0
5 2014		536	258	30	4	1	20	20		3	18	7	7	6
6 2015		972	392	72	15	11	40	40		3	32	15	15	21
7 2016		1579	552	102	38	32	75	75		9	55	31	31	46
8 2017		2290	702	157	68	57	102	102		17	84	52	52	61
9 2018		3115	836	214	109	91	148	148		47	122	70	70	88
10 2019		3964	960	255	173	145	195	195		83	163	102	102	122
11 2020		4738	1049	317	223	221	244	244		140	194	139	139	140
12 2021		5452	1135	377	287	274	289	289		207	234	181	181	165
13 2022		6128	1220	420	359	339	335	335		275	280	204	204	194
14 2023		6566	1260	462	421	387	378	378		334	308	223	223	212
15 2024		6573	1260	463	422	387	378	378		335	308	224	224	212

Εικόνα 5.43 Εξέλιξη των λέξεων-κλειδιών στο RStudio

5.3.1 Επιστημονικοί χάρτες - Χάρτες συνεργασίας

Το δίκτυο συνεργασίας συγγραφέων που θα δούμε στην συνέχεια (Εικόνα 5.44) αποκαλύπτει ομάδες μελέτης και σημαντικούς συγγραφείς. Έπειτα εμφανίζονται τα κύρια χαρακτηριστικά του δικτύου συν-συγγραφέων (Εικόνα 5.45).



Εικόνα 5.44 Δίκτυο συνεργασίας συγγραφέων

Main statistics about the network

Size	17220
Density	0
Transitivity	0.246
Diameter	22
Degree Centralization	0.018
Average path length	5.981

Εικόνα 5.45 Κύρια χαρακτηριστικά δικτύου συν-συγγραφέων

Στην Εικόνα 5.46 γίνεται υπολογισμός μέσα από την R στις μετρικές:

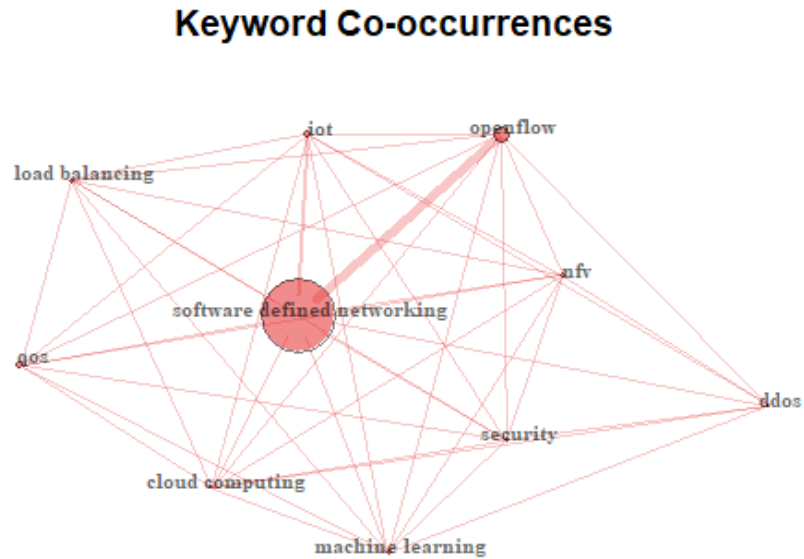
- H-index: μετρά την επιρροή των παραπομπών ενός συγγραφέα.
- G-index: μετρά την επιρροή των δημοσιεύσεων βάση των αναφορών.
- M-index: υπολογίζει την συχνότητα που ένας συγγραφέας δημοσιεύει έργα
- TC: μετρά τις συνολικές αναφορές που έχει λάβει ο συγγραφέας για όλες του τις εργασίες.
- NP: μετρά τον συνολικό αριθμό των εργασιών που έχει δημοσιεύσει ένας συγγραφέας.
- PY_start: αναφέρεται το έτος που ξεκίνησαν να δημοσιεύουν εργασίες.

Οι πληροφορίες που παίρνουμε από τον γράφο της Εικόνας 5.44 φαίνονται στη συνέχεια.

```
Element h_index g_index m_index TC NP PY_start
1 CHEN 16 25 1.3333333 803 91 2013
2 CHEN J 12 18 1.0909091 399 44 2014
3 HUANG 14 21 1.2727273 535 49 2014
4 KIM 15 27 1.0714286 827 65 2011
5 LI 17 37 1.4166667 1497 93 2013
6 LI J 17 38 1.5454545 1525 70 2014
7 LI X 13 28 1.1818182 826 38 2014
8 LI Y 17 32 1.5454545 1119 56 2014
9 LIU 14 27 1.2727273 844 78 2014
10 LIU J 16 36 1.4545455 1332 41 2014
11 LIU Y 14 31 1.2727273 967 45 2014
12 NA NA 8 13 0.7272727 219 57 2014
13 WANG 22 36 1.6923077 1565 115 2012
14 WANG J 12 25 1.0000000 684 44 2013
15 WANG L 12 20 1.0909091 467 46 2014
16 WANG X 16 26 1.4545455 841 66 2014
17 WANG Y 17 27 1.4166667 932 79 2013
18 WU 17 28 1.5454545 899 61 2014
19 WU J 15 28 1.2500000 852 53 2013
20 XU 16 28 1.4545455 889 58 2014
21 XU H 16 24 1.6000000 632 51 2015
22 YANG 15 30 1.3636364 958 60 2014
23 ZHANG 18 34 1.5000000 1355 90 2013
24 ZHANG J 18 30 1.5000000 989 56 2013
25 ZHANG Y 12 22 0.9230769 544 56 2012
```

Εικόνα 5.46 Υπολογισμός διάφορων μετρικών

Τέλος υπάρχει ο χάρτης συνεμφάνισης λέξεων-κλειδιών (Εικόνα 5.47), οι κεντρικότητες (Εικόνα 5.48), και ο βαθμός των κόμβων (Εικόνα 5.49).



Εικόνα 5.47 Χάρτης συνεμφάνισης λέξεων-κλειδιών

\$cluster_res	vertex	cluster	btw_centrality	clos_centrality	pagerank_centrality
1	software defined networking	1	0.0	0.0006706908	0.37367981
2	openflow	1	0.0	0.0025641026	0.15398072
3	iot	1	0.0	0.0029585799	0.08144177
4	qos	1	0.0	0.0032362460	0.06903946
5	ddos	1	0.0	0.0035971223	0.06029903
6	security	1	0.0	0.0038910506	0.06855029
7	machine learning	1	5.0	0.0042194093	0.06096422
8	load balancing	1	11.0	0.0042016807	0.05023578
9	nfv	1	4.5	0.0040160643	0.04088780
10	cloud computing	1	11.0	0.0039840637	0.04092112

Εικόνα 5.48 Υπολογισμός Κεντρικότητων

SnodeDegree		node	degree
software defined networking		software defined networking	1.000000000
openflow		openflow	0.217749211
iot		iot	0.108930988
qos		qos	0.104645918
security		security	0.074199369
machine learning		machine learning	0.064727109
load balancing		load balancing	0.062246279
nfv		nfv	0.062246279
ddos		ddos	0.059201624
cloud computing		cloud computing	0.050631484
routing		routing	0.049391069
mininet		mininet	0.047586829
controller		controller	0.047135769
network security		network security	0.045782589
5g		5g	0.041271989
network function virtualization		network function virtualization	0.038001804
of		of	0.036535859
reliability		reliability	0.031686964
software defined networking controller		software defined networking controller	0.031348669
scalability		scalability	0.030784844
network virtualization		network virtualization	0.030672079
control plane		control plane	0.029431664
network management		network management	0.028867839
blockchain		blockchain	0.028867839
traffic engineering		traffic engineering	0.027852954
big data		big data	0.027289129
deep learning		deep learning	0.026950834
controller placement		controller placement	0.024695535
data plane		data plane	0.024244475
optimization		optimization	0.024244475
edge computing		edge computing	0.024131710
latency		latency	0.023680650
ddos attack		ddos attack	0.022891295
energy efficiency		energy efficiency	0.022553000
resource allocation		resource allocation	0.022101940

Εικόνα 5.49 Βαθμοί των κόμβων

6. Συμπεράσματα

Τα τελευταία χρόνια παρατηρείται μια αύξηση στις μελέτες με την μέθοδο της βιβλιομετρικής ανάλυσης. Αυτό συμβαίνει πιθανότατα επειδή η μελέτη του μεγάλου όγκου των δεδομένων που απαιτείται υπάρχει πλέον ηλεκτρονικά. Με αυτή την μέθοδο ο αναλυτής μελετά και οπτικοποιεί την πληροφορία με σκοπό την εύκολη ανίχνευση σχέσεων μέσω του οπτικού συστήματος.

Στην παρούσα εργασία μελετήθηκε η επιστημονική περιοχή των δικτύων SDN και αποκομίσθηκε μία σφαιρική εικόνα. Για την κατανόηση της τεχνολογίας των SDN χρησιμοποιήθηκε η βιβλιομετρική ανάλυση. Τα δίκτυα καθορισμένα από το λογισμικό βασίζονται στο πρωτόκολλο OpenFlow και η επικοινωνία γίνεται μέσω ενός κεντρικού ελεγκτή. Διαφέρει από τα παραδοσιακά δίκτυα διότι έχει επιτευχθεί ο διαχωρισμός του επιπέδου ελέγχου (control plane) με το επίπεδο δεδομένων (data plane).

Με τη βοήθεια της βιβλιομετρικής ανάλυσης και των δεδομένα που εξήχθησαν από την βάση Scopus κατέστη δυνατή η διερεύνηση, κατανόηση και συζήτηση για την παρούσα κατάσταση της επιστημονικής περιοχής των δικτύων SDN. Χρησιμοποιώντας συνδυασμούς επιλεγμένων λέξεων-κλειδιών και εφαρμόζοντας διάφορες τεχνικές ανάλυσης(π.χ. ομαδοποίηση), εκτιμήθηκαν βασικοί βιβλιομετρικοί δείκτες. Με την χρήση του λογισμικού VOSViewer διαμορφώθηκαν χάρτες συνεργασίας και ανιχνεύθηκαν θεματικές περιοχές μέσα από τις λέξεις κλειδιά. Μέσα από το Gephi εξετάζεται η επιρροή των συγγραφέων και τα ενδιαφέροντα που έχουν για μελέτη. Τέλος, με την χρήση της γλώσσας R επεξεργάσθηκαν τα δεδομένα με αποτέλεσμα να δημιουργηθούν οι επιστημονικοί χάρτες.

Βιβλιογραφία

1. Waseem, Q., Din, W. I. S. W., Aminuddin, A., Mohammed, M. H., & Aziza, R. F. A. (2022, August). Software-Defined Networking (SDN): A Review. In *2022 5th International Conference on Information and Communications Technology (ICOIACT)* (pp. 30-35). IEEE.
2. Kreutz, D., Ramos, F. M., Verissimo, P. E., Rothenberg, C. E., Azodolmolky, S., & Uhlig, S. (2014). Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, *103*(1), 14-76.
3. "SDN Architecture Overview" Opennetworking.org. Retrieved 22 November 2014.
4. Benzekki, Kamal; El Fergougui, Abdeslam; Elbelrhiti Elalaoui, Abdelbaki (2016). "Software-defined networking (SDN): A survey". *Security and Communication Networks*. 9 (18): 5803–5833.
5. Pros And Cons Of Software-Defined Networking (SDN), Technology Blog (2018)
6. SDN Technical Specifications. Open Networking Foundation. (2021, September 22). <https://opennetworking.org/software-defined-standards/specifications/>
7. McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., ... & Turner, J. (2008). OpenFlow: enabling innovation in campus networks. *ACM SIGCOMM computer communication review*, *38*(2), 69-74.
8. Rousseau, R. (2014). Forgotten founder of bibliometrics. *Nature*, *510*(7504), 218-218.
9. «[Δείκτες αξιολόγησης περιοδικών & ερευνητικής δραστηριότητας](#)». Βιβλιοθήκη και Κέντρο Πληροφόρησης. Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης.». Βιβλιοθήκη και Κέντρο Πληροφόρησης. Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης.
10. The University of British Columbia (2023). Citation Metrics Workshop. <https://guides.library.ubc.ca/citationmetricsworkshop/researchers>
11. Bornmann, L., & Daniel, H. D. (2007). What do we know about the h index?. *Journal of the American Society for Information Science and technology*, *58*(9), 1381-1385.
12. "Ομάδα Web of Science" . Web of Science Group . 5 Αυγούστου 2019 . Ανακτήθηκε στις 28 Ιουλίου 2020
13. Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., & Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. *Profesional de la Información*, *29*(1)
14. Ding, Y., Chowdhury, G. G., & Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information processing & management*, *37*(6), 817-842.
15. Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for information Science and Technology*, *62*(7), 1382-1402.
16. Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, *60*(2), 348-362.
17. Van Eck, N., & Waltman, L. (2010). Software survey: VOSViewer, a computer program for bibliometric mapping. *scientometrics*, *84*(2), 523-538.
18. <https://www.VOSViewer.com/>
19. <https://gephi.org/>
20. Bastian, M., Heymann, S., & Jacomy, M. (2009, March). Gephi: an open sourcesoftware for exploring and manipulating networks. In *Proceedings of the international AAAI conference on web and social media* (Vol. 3, No. 1, pp. 361-362).
21. Grandjean, Martin (2015). «GEPHI – Εισαγωγή στην Ανάλυση και Οπτικοποίηση Δικτύων», <https://www.martingrandjean.ch/gephi-introduction>

22. Newman, MEJ (2006). "[Διαμόρφωση και κοινοτική δομή στα δίκτυα](#)". Πρακτικά της Εθνικής Ακαδημίας Επιστημών των Ηνωμένων Πολιτειών της Αμερικής, 103(23): 85778696. [arXiv: physics/0602124](#) . [Bibcode:2006PNAS..103.8577N](#). [doi: 10.1073/pnas.0601602103](#) . [PMC 1482622](#) . [PMID 16723398](#).
23. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of businessresearch*, 133, 285-296.
24. Levallois, C. (2017). Workingwithtext in Gephi.<https://seinecle.github.io/gephi-tutorials/generated-html/working-with-text-en.html>
25. Camille, R. (2015). Create a topic map of (some of) your institution's publications in Gephi.<https://emerging.commons.gc.cuny.edu/2015/02/create-topic-map-institutions-publications-gephi/>
26. Petrovich, E. (2020). Science mapping.https://www.isko.org/cyclo/science_mapping#4.1.2
27. <https://en.wikipedia.org/wiki/RStudio>
28. Johnson, D. (2024). R Tutorial για αρχάριους: Μάθετε τη γλώσσα προγραμματισμού R. [RStudioTutorial για αρχάριους: Ένας πλήρης οδηγός | DataCamp](#)
29. The R Project for Statistical Computing. <https://www.r-project.org/>
30. Nath, S.(2023). Community Dentistry & Oral Epidemiology Bibliometric Analysis. <https://rpubs.com/sonianath/1022496>
31. Aria, M., Cuccurullo, C. (2017). A brief introduction to bibliometrix. https://www.bibliometrix.org/vignettes/Introduction_to_bibliometrix.html
32. Aria, M. (2023). Comprehensive Science Mapping Analysis.<https://cran.r-project.org/web/packages/bibliometrix/bibliometrix.pdf>
33. Hossain, N.U.I., Dayarathna, V.L., Nagahi, M. and Jaradat, R., 2020. Systems thinking: A review and bibliometricanalysis. *Systems*, 8(3), p.23.
34. MacKeown N., Anderson T., Balakrishnan H., Parulkar G., Peterson L., Rexford J., Shenker S., Turner J. 2018. OpenFlow: Enabling Innovation in Campus Networks