

ΠΜΣ ΟΙΚΟΝΟΜΙΚΗ ΦΥΣΙΚΗ, ΤΜΗΜΑ
ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ, ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΙΑΣ

**Η ΑΝΑΔΕΙΞΗ ΤΗΣ ΣΗΜΑΣΙΑΣ ΤΗΣ
ΠΟΛΥΠΛΟΚΟΤΗΤΑΣ ΤΩΝ
ΤΡΟΦΟΔΟΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΣΤΗΝ
ΑΝΑΠΤΥΞΗ ΤΗΣ ΤΕΧΝΟΓΝΩΣΙΑΣ ΤΩΝ
ΠΡΟΙΟΝΤΩΝ ΤΟΥ ΗΛΕΚΤΡΟΝΙΚΟΥ
ΚΛΑΔΟΥ, ΜΕ ΤΗΝ ΧΡΗΣΗ ΤΗΣ ΘΕΩΡΙΑΣ
ΔΙΚΤΥΩΝ**

ΠΑΠΑΔΗΜΗΤΡΙΟΥ ΠΑΝΑΓΙΩΤΗΣ ΑΝΑΡΓΥΡΟΣ-
ΖΗΣΗΣ

ΕΠΙΒΛΕΠΩΝ
ΚΑΘΗΓΗΤΗΣ ΚΑΡΑΚΑΣΙΔΗΣ ΘΕΟΔΩΡΟΣ

**ΒΟΛΟΣ
2023**

Περιεχόμενα

Περίληψη	3
Abstract	4
1. Εισαγωγή	5
2. Βιβλιογραφική ανασκόπηση	5
3. Δεδομένα.....	9
4. Μεθοδολογία Έρευνας.....	10
4.1 Θεωρία δικτύων	10
4.1.1 Κοινότητες	15
4.1.2 Σημαντικότητα των κόμβων	16
4.1.3 Κατασκευή δικτύου	20
4.2 Ανάλυση χρονοσειρών.....	28
4.2.1 Περιγραφική στατιστική	28
4.2.2 Συνάρτηση αυτοσυσχέτισης.....	30
4.2.3 Έλεγχος Dickey-Fuller.....	32
4.2.4 Αυτοπαλίνδρομα υποδείγματα.....	33
5. Εμπειρικά Αποτελέσματα	35
6. Συζήτηση-Σχολιασμός	54
7. Συμπεράσματα προτάσεις	56
8. Βιβλιογραφία	58

Βεβαιώνω ότι είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στη διπλωματική εργασία. Επίσης έχω αναφέρει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επίσης βεβαιώνω ότι αυτή η πτυχιακή εργασία προετοιμάστηκε από εμένα προσωπικά ειδικά για τις απαιτήσεις του προγράμματος μεταπτυχιακών σπουδών στην Εφαρμοσμένη Οικονομική του Τμήματος Οικονομικών Επιστημών του Πανεπιστημίου Θεσσαλίας. Βόλος, Ιανουάριος 2023.

Περίληψη

Η παρούσα μελέτη ασχολείται με την πολυπλοκότητα των ηλεκτρονικών προϊόντων. Η θεωρία γράφων αποτελεί βασικό εργαλείο για των υπολογισμό του δείκτη πολυπλοκότητας των προϊόντων. Στην παρούσα εργασία όμως η θεωρία γράφων χρησιμοποιείται επιπλέον για την αναπαράσταση του συνόλου των διασυσχετίσεων της πολυπλοκότητας των ηλεκτρονικών προϊόντων. Έτσι κατασκευάζεται ένα δίκτυο γραμμικών συσχετίσεων το οποίο αναπαριστά τους σχετισμούς ανάμεσα στους δείκτες πολυπλοκότητας ηλεκτρονικών προϊόντων (δηλαδή με κωδικό HS που ξεκινάει από 85). Εν συνεχεία πραγματοποιείται αξιολόγηση των κόμβων ως προς την σημαντικότητα τους και διακρίνονται οι πιο κεντρικοί κόμβοι. Παράλληλα πραγματοποιείται και η καταγραφή κοινοτήτων που παρουσιάζονται στο δίκτυο. Τέλος εκπονείται ανάλυση χρονοσειρών στους σημαντικότερους κόμβους. Βασικό συμπέρασμα της εργασίας είναι η ύψιστη κεντρικότητα του κωδικού 8504 ο οποίος αναπαριστά τα επίπεδα πολυπλοκότητας των συστατικών τροφοδοτικών προϊόντων. Έτσι αναδεικνύεται η καθοριστική σημασία που φέρει η τεχνογνωσία κατασκευής τροφοδοτικών προϊόντων για το σύνολο των προϊόντων του ηλεκτρονικού κλάδου.

Abstract

The current study talks about the complexity of electronic products. Graph theory is a basic tool in order to calculate the product complexity index. In the current study graph theory is used to portrait cross-correlation of the complexity of electronic products. We create a linear correlation network that shows the bonds between the complexity indexes of electronic products (those who start with HS 85 code). Then, we valuate the nodes as from their statistical significance to show the most central nodes. We run time series analysis on the most important nodes. Key conclusion is the high centrality of HS code 5804 that shows the complexity of components of power supply products. The construction knowledge of power supply products has a very high effect on the whole electronic products industry.

1. Εισαγωγή

Σημαντική καινοτομία της εργασίας είναι η δισδιάστατη αξιοποίηση της θεωρίας δικτύων. Αρχικά χρησιμοποιούνται οι δείκτες πολυπλοκότητας των ηλεκτρονικών προϊόντων, δείκτες παράγωγα της θεωρίας δικτύων καθώς αυτή χρησιμοποιείται για τον υπολογισμό τους. Αυτοί οι δείκτες εισάγονται ως δεδομένα και σε δεύτερο χρόνο υπολογίζεται το δίκτυο διασυσχετίσεων τους. Το κομμάτι της βιβλιογραφικής ανασκόπησης ασχολείται με την επεξήγηση του δείκτη πολυπλοκότητας και της σημασίας του καθώς και την θεωρία γράφων ως βασικό εργαλείο για τον υπολογισμό του δείκτη πολυπλοκότητας. Το κομμάτι της μεθοδολογίας έρευνας ασχολείται με τον τρόπο κατασκευής του δικτύου διασυσχετίσεων των δεικτών πολυπλοκότητας, τις βασικές μετρικές του δικτύου που θα χρησιμοποιηθούν ως προς την κεντρικότητα των κόμβων και τις κοινότητες και τέλος με τις τεχνικές ανάλυσης χρονοσειρών που θα χρησιμοποιηθούν για τους σημαντικούς κόμβους.

2. Βιβλιογραφική ανασκόπηση

Η πολυπλοκότητα του προϊόντος έχει γίνει ένας ολοένα και πιο σημαντικός παράγοντας στις σημερινές ανταγωνιστικές αγορές. Οι εταιρείες προσπαθούν συνεχώς να αναπτύσσουν πιο σύνθετα προϊόντα με υψηλότερα επίπεδα απόδοσης και λειτουργικότητας. Ως εκ τούτου, είναι σημαντικό να κατανοήσουμε πώς η πολυπλοκότητα του προϊόντος μπορεί να μετρηθεί και να παρακολουθηθεί με την πάροδο του χρόνου. Ο Δείκτης Πολυπλοκότητας Προϊόντος (PCI) έχει προταθεί ως μέτρηση για τη μέτρηση της πολυπλοκότητας του προϊόντος σε διάφορους κλάδους. Αυτή η εργασία παρουσιάζει μια ανάλυση του PCI της βιομηχανίας των ηλεκτρονικών, χρησιμοποιώντας χρονοσειρές και σύνθετες μεθόδους ανάλυσης δικτύου.

Ο δείκτης πολυπλοκότητας προϊόντος (PCI) είναι ένα μέτρο της πολυπλοκότητας ενός προϊόντος ή συστήματος. Χρησιμοποιείται για να ποσοτικοποιηθεί την πολυπλοκότητα ενός προϊόντος ή

συστήματος όσον αφορά τα στοιχεία, τα χαρακτηριστικά και τις λειτουργίες του. Το PCI είναι ένας σημαντικός παράγοντας για τον καθορισμό του κόστους και του χρόνου που απαιτείται για την ανάπτυξη, τη συντήρηση και την υποστήριξη. Βασική του διάσταση είναι το γεγονός ότι εκφράζει το επίπεδο πολυπλοκότητας κάθε προϊόντος ως προβολή της πολυπλοκότητας των χαρακτηριστικών που φέρει (Mealy et al. ,2019)

Το PCI υπολογίζεται λαμβάνοντας υπόψη τον αριθμό των στοιχείων, των χαρακτηριστικών και των λειτουργιών που συνθέτουν ένα προϊόν ή ένα σύστημα. Όσο πιο περίπλοκο είναι το προϊόν ή το σύστημα, τόσο υψηλότερη θα είναι η βαθμολογία PCI του. Η βαθμολογία PCI μπορεί να κυμαίνεται από 0 έως 10, με το 0 να είναι το λιγότερο σύνθετο και το 10 να είναι το πιο σύνθετο.

Ο υπολογισμός του PCI περιλαμβάνει την ανάλυση διαφόρων πτυχών ενός προϊόντος ή συστήματος, όπως η αρχιτεκτονική, ο σχεδιασμός, η υλοποίηση, η δοκιμή, η ανάπτυξη και η συντήρησή του. Λαμβάνει επίσης υπόψη παράγοντες όπως η επεκτασιμότητα και η επεκτασιμότητα.

Το PCI μπορεί να χρησιμοποιηθεί για τη σύγκριση διαφορετικών προϊόντων ή συστημάτων ως προς την πολυπλοκότητά τους. Μπορεί επίσης να χρησιμοποιηθεί για να εκτιμηθεί πόση προσπάθεια απαιτείται για την ανάπτυξη και τη συντήρηση με την πάροδο του χρόνου. Αυτές οι πληροφορίες μπορούν να χρησιμοποιηθούν για τη λήψη αποφάσεων σχετικά με το ποια προϊόντα ή συστήματα πρέπει να αναπτυχθούν πρώτα ή ποια πρέπει να έχουν προτεραιότητα όταν πρόκειται για την κατανομή πόρων.

Οι κωδικοί HS (Harmonized Code, Κωδικοί Εναρμονισμένου Συστήματος) είναι ένα διεθνώς τυποποιημένο σύστημα ονομάτων και αριθμών που χρησιμοποιούνται για την ταξινόμηση προϊόντων που διακινούνται. Χρησιμοποιούνται από τις τελωνειακές αρχές σε όλο τον κόσμο για τον εντοπισμό προϊόντων κατά την εκτίμηση δασμών και φόρων και για τη συλλογή εμπορικών στατιστικών. Οι κωδικοί HS είναι οργανωμένοι σε 21 ενότητες, 96 κεφάλαια και περισσότερες από 5.000 πενταψήφιες κατηγορίες προϊόντων. Κάθε κατηγορία προϊόντος προσδιορίζεται από έναν εξαψήφιο κωδικό που χρησιμοποιείται για την ταξινόμηση των εμπορευμάτων για εισαγωγή και εξαγωγή. Τα δύο πρώτα ψηφία του κωδικού προσδιορίζουν την ενότητα, τα επόμενα δύο

ψηφία προσδιορίζουν το κεφάλαιο και τα δύο τελευταία ψηφία προσδιορίζουν το προϊόν σε αυτό το κεφάλαιο.

Η θεωρία γράφων είναι ένας κλάδος των μαθηματικών που μελετά τις ιδιότητες των γράφων και τις εφαρμογές τους σε διάφορους τομείς. Χρησιμοποιείται για την ανάλυση δικτύων, όπως κοινωνικά δίκτυα, δίκτυα μεταφορών, δίκτυα επικοινωνίας, επιχειρηματικά δίκτυα ή βιολογικά δίκτυα. Η θεωρία γράφων έχει επίσης εφαρμοστεί στην ανάλυση των Κωδικών Εναρμονισμένων Συστημάτων (HS Codes). Οι κωδικοί HS χρησιμοποιούνται για την ταξινόμηση αγαθών για διεθνές εμπόριο. Χρησιμοποιώντας τη θεωρία γράφων, είναι δυνατό να αναλυθούν οι σχέσεις μεταξύ διαφορετικών Κωδικών HS και να κατανοηθεί πώς σχετίζονται. (Hidalgo & Hausmann, 2009).

Η βασική ιδέα πίσω από τη θεωρία γράφων είναι ότι χρησιμοποιεί ένα σύνολο κόμβων (ή κορυφών) που συνδέονται με ακμές (ή τόξα) για να αναπαραστήσουν ένα δίκτυο. Κάθε κόμβος αντιπροσωπεύει μια οντότητα στο δίκτυο και κάθε άκρη αντιπροσωπεύει μια σχέση μεταξύ δύο οντοτήτων. Στην περίπτωση των κωδικών HS, κάθε κόμβος αντιπροσωπεύει ένα μεμονωμένο προϊόν ή υπηρεσία και κάθε άκρη αντιπροσωπεύει μια σχέση μεταξύ δύο προϊόντων ή υπηρεσιών. Αναλύοντας τη δομή αυτού του γραφήματος, είναι δυνατό να αποκτήσετε μια εικόνα για το πώς συνδέονται διαφορετικά προϊόντα ή υπηρεσίες και πώς αλληλοεπιδρούν μεταξύ τους.

Η θεωρία γράφων μπορεί να χρησιμοποιηθεί για την ανάλυση των σχέσεων προϊόντων του Κώδικα HS εξετάζοντας τη δομή του γραφήματος και προσδιορίζοντας ποια προϊόντα σχετίζονται μεταξύ τους. Για παράδειγμα, εάν δύο προϊόντα έχουν παρόμοια χαρακτηριστικά ή παράγονται με παρόμοιους τρόπους, τότε πιθανότατα θα συνδέονται με μια ακμή στο γράφημα. Επιπλέον, εάν δύο προϊόντα έχουν διαφορετικά χαρακτηριστικά ή παράγονται με διαφορετικούς τρόπους, τότε πιθανότατα δεν θα συνδέονται με μια άκρη στο γράφημα. Αναλύοντας αυτές τις σχέσεις μεταξύ διαφορετικών Κωδικών HS, είναι δυνατό να αποκτήσετε μια εικόνα για το πώς διαφορετικά προϊόντα αλληλοεπιδρούν μεταξύ τους και πώς σχετίζονται μεταξύ τους σε μια δεδομένη αγορά ή κλάδο.

Ενώ η θεωρία γράφων αποτελεί βασικό εργαλείο για την εκτίμηση της οικονομική πολυπλοκότητας, η παρούσα μελέτη αυξάνει την χρησιμότητα της συγκεκριμένη θεωρίας περαιτέρω. Βασική διάσταση της συγκεκριμένη μελέτη είναι η αξιοποίηση της θεωρίας γράφων προκειμένου να αναλυθεί το σύνολο των διασυσχετίσεων μεταξύ των κωδικών πολυπλοκότητας των προϊόντων. Με αυτόν τον τρόπο δίνεται η δυνατότητα αρχικά να εκφραστεί η απαιτούμενη τεχνογνωσία ανά προϊόν σε ένα ευρύτερο, διακλαδικό σύνολο.

Επιπλέον, βασική εκροή της ανάλυσης είναι η ανάδειξη των κωδικών προϊόντων που έχουν την υψηλότερη κεντρικότητα στην τοπολογία του δικτύου. Έτσι θα αναδειχθούν οι ομάδες προϊόντων με την μεγαλύτερη επιρροή, την νευραλγικότερη θέση δηλαδή στο σύστημα απαιτούμενης τεχνογνωσίας. Επιπρόσθετα, μία ακόμη εκροή είναι η καταγραφή ενδεχόμενων κοινοτήτων στο δίκτυο, κάτι που θα δείξει ποιες ομάδες προϊόντων σχετίζονται ή αλληλεξαρτώνται εντονότερα μεταξύ τους.

Τέλος, βασική καινοτομία της συγκεκριμένης εργασίας είναι η στόχευση της σε συγκεκριμένη ομάδα προϊόντων, αυτών που ξεκινάνε με κωδικό 85XX, καθιστώντας την έτσι μία εξειδικευμένη μελέτη της τεχνογνωσίας των ηλεκτρονικών προϊόντων και δίνοντας την δυνατότητα να αναδειχθούν οι σημαντικότεροι κωδικοί στου συγκεκριμένου κλάδου.

Συμπερασματικά, η θεωρία γράφων μπορεί να χρησιμοποιηθεί για την ανάλυση των σχέσεων προϊόντων του Εναρμονισμένου Κώδικα Συστήματος, εξετάζοντας τη δομή του γραφήματος και προσδιορίζοντας ποια προϊόντα σχετίζονται μεταξύ τους. Αυτό μπορεί να προσφέρει πολύτιμες πληροφορίες για το πώς αλληλοεπιδρούνε διαφορετικά προϊόντα σε μια δεδομένη αγορά ή κλάδο και μπορεί να βοηθήσει τις επιχειρήσεις να κατανοήσουν καλύτερα το ανταγωνιστικό τους τοπίο. Επιπρόσθετα, η συγκεκριμένη εργασία χρησιμοποιεί την θεωρία γράφων και σε ένα δεύτερο επίπεδο, δηλαδή στην καταγραφή και ανάλυση των αλληλεξαρτήσεων ανάμεσα στα επίπεδα πολυπλοκότητας ομάδων προϊόντων.

3. Δεδομένα

Στην συγκεκριμένη ανάλυση ως μεταβλητές χρησιμοποιούνται οι δείκτες πολυπλοκότητας προϊόντος. Οι αρχικές μεταβλητές είναι $k=48$ και όλες αναφέρονται σε ηλεκτρονικά προϊόντα. Ως ονόματα των μεταβλητών χρησιμοποιείται ο κωδικός HS 1992 της κάθε ομάδας προϊόντος. Κάθε κωδικός, δηλαδή κάθε μεταβλητή αντιστοιχεί σε μία ομάδα προϊόντων. Ο συγκεκριμένος δείκτης εκφράζει την ποικιλία, την πολυπαραγοντικότητα και το σύνολο των χαρακτηριστικών καθώς και το πόσο απαιτητικά είναι αυτά τα χαρακτηριστικά σε γνώση προκειμένου να επέλθει η παραγωγή των προϊόντων της συγκεκριμένη κλάσης.

Οι 48 μεταβλητές που εισάγονται στην ανάλυση και σχετίζονται με τον κλάδο των ηλεκτρικών προϊόντων ξεκινούν με τον κωδικό 8501 και καταλήγουν στον αριθμό 8548. Δηλαδή η ανάλυση και κατασκευή του δικτύου σε πρώτη φάση υπολογίζει όλες τις ομάδες προϊόντων που σχετίζονται με τον ηλεκτρολογικό κλάδο, μέχρις ότου εξαιρεθούν οι περιττές μεταβλητές από το δίκτυο.

Οι συγκεκριμένες μεταβλητές επιτρέπουν την ολιστική καταγραφή του επιπέδου τεχνογνωσιακής ανάπτυξης του ηλεκτρικού κλάδου. Δηλαδή έχουν ληφθεί υπ' όψιν όλες οι διαθέσιμες παράμετροι όσον αφορά το επίπεδο πολυπλοκότητας και συνεπώς τεχνογνωσίας που υπάρχει στα επί μέρους προϊόντα της ευρύτερης ηλεκτρονικής οικογένειας. Συνεπώς η συμπερασματολογία όσον αφορά την θεωρία δικτύων είναι ολιστική όσον αφορά τον συγκεκριμένο κλάδο, καθώς στο δίκτυο ως συστατικά αξιοποιούνται όλα τα διαθέσιμα στοιχεία.

Οι μεταβλητές είναι καταγεγραμμένες σε ετήσιες παρατηρήσεις και αναφέρονται στην χρονική περίοδο 1995-2019. Η κάθε χρονοσειρά δηλαδή αποτελείται από $n=25$ παρατηρήσεις. Ο πίνακας δεδομένων είναι διαστάσεων (25×38) .

Τα δεδομένα προέρχονται από τον Άτλα του Χάρβαρντ για την οικονομική πολυπλοκότητα.

<https://atlas.cid.harvard.edu/rankings/product>

4. Μεθοδολογία Έρευνας

4.1 Θεωρία δικτύων

Η θεωρία δικτύων (ή αλλιώς θεωρία γράφων) αποτελεί αντικείμενο θεμελίωσης και μελέτης της μαθηματικής επιστήμης με δυνατότητα εφαρμογής σε όλα τα επιστημονικά πεδία. Η διεπιστημονικότητα του αντικειμένου των δικτύων έγκειται στο γεγονός της πολυπαραγοντικότητας ή πολυμεταβλητότητας η οποία συναντάται σε όλους τους επιστημονικούς κλάδους. Σε αντίστοιχο πνεύμα με τα συστήματα εξισώσεων ένα δίκτυο αναπαριστά το σύνολο των αλληλεπιδράσεων ή συσχετίσεων των οντοτήτων, ξεφεύγοντας από διμεταβλητά πλαίσια και εξετάζοντας το περιβάλλον στο οποίο ανήκουν οι μεταβλητές. Βασικές πληροφορίες που μπορεί να παρέχει ένα δίκτυο είναι το πιο κόμβοι έχουν σημαντικότερο ρόλο και ποιες κοινότητες δημιουργούνται στο δίκτυο.

Ένα πλεονέκτημα της θεωρίας δικτύων έναντι των συστημάτων εξισώσεων είναι ότι δεν αντιμετωπίζει περιορισμούς ως προς την εμπειρική εκτίμηση. Έναντι των γράφων τα συστήματα εξισώσεων είναι ανώτερα καθώς προσδιορίζουν την ακριβή συναρτησιακή σχέση μεταξύ των μεταβλητών με αποτέλεσμα να είναι ακριβέστερα και να έχουν την δυνατότητα διεξαγωγής προβλέψεων. Κατά την εμπειρική εκτίμηση συστημάτων εξισώσεων όμως δημιουργούνται προβλήματα όπως αυτό της αλληλεξάρτησης που έχει ως αποτέλεσμα την μη εξωγένεια των διαταρακτικών όρων και συνεπώς μεροληπτικούς εκτιμητές, καθώς επίσης και το πρόβλημα της ταυτοποίησης.

Ένα δίκτυο μπορεί να παρακάμψει τα παραπάνω καθώς δεν αποτελείται από ακριβείς συναρτησιακές σχέσεις, αλλά από ενδείξεις διασυσχετίσεων με την χρήση δεικτών, προσπερνώντας έτσι μαθηματικούς και οικονομετρικούς περιορισμούς. Τα δίκτυα αντίστοιχα μπορούν να καταγράψουν την ένταση της σχέσης, την φορά της επίδρασης, καθώς και το πρόσημο της σχέσης.

Ως δίκτυο νοείται ένα σύνολο κόμβων (vertices) και ακμών (edges), με τις ακμές να αναπαριστούν τις σχέσεις μεταξύ κόμβων. Ουσιαστικά αποτελεί ένα σύνολο διασυσχετίσεων και δύναται να περιέχει και αυτοσυσχετίσεις-αυτοεπιδράσεις των μεταβλητών (self loops). Από εδώ και πέρα η έννοια του κόμβου ταυτίζεται με την έννοια της μεταβλητής καθώς ο κάθε κόμβος αναπαριστά μία μεταβλητή. Υπάρχουν δύο τρόποι κατά τους οποίους μπορεί να εκφραστεί, επεξεργασθεί και αποθηκευτεί ένα δίκτυο.

Η μαθηματική οντότητα που αναπαριστά το δίκτυο ονομάζεται Γράφος (graph) και ορίζεται ως $G(V,E)$ όπου V το σύνολο των κόμβων του δικτύου και E το σύνολο των ακμών του. Ο πρώτος τρόπος για την ακριβή έκφραση του δικτύου είναι μέσω του συνόλου των κόμβων $V\{v_1, v_2, \dots, v_n\}$ και του συνόλου των ακμών $E\{e_1e_2, e_2e_4, \dots, e_ne_1\}$ όπου e_1e_2 εκφράζει είτε την ύπαρξη ακμής ανάμεσα στον κόμβο v_1 και v_2 . (West, 2001).

Ο δεύτερος τρόπος ο οποίος χρησιμοποιείται και στην συγκεκριμένη εργασία είναι μέσω του πίνακα γειτνίασης. Ο πίνακας γειτνίασης (adjacency matrix) είναι διαστάσεων $\overline{adj}(N*N)$ όπου N το σύνολο των κόμβων του δικτύου. Κάθε στοιχείο του πίνακα αναπαριστά την ύπαρξη ή όχι συσχέτισης ανάμεσα στον κόμβο της αντίστοιχης γραμμής και στον κόμβο της αντίστοιχης στήλης. Αν το στοιχείο i,j του πίνακα είναι διάφορο του μηδενός, σημαίνει ότι υπάρχει συσχέτιση ανάμεσα στους δύο κόμβους, ή αλλιώς ότι οι κόμβοι γειτονεύουν. Η σύσταση του πίνακα εξαρτάται από το είδος του γραφήματος και συνεπώς πρώτου παρουσιασθεί ο πίνακας γειτνίασης προηγείται η κατηγοριοποίηση των γράφων σε δύο βασικές διακρίσεις. Σε βεβαρημένα και μη βεβαρημένα δίκτυα και σε κατευθυνόμενα και μη κατευθυνόμενα δίκτυα. (West, 2001).

Δίκτυα Χωρίς Βάρη, Δίκτυα Με Βάρη

Όταν ένα δίκτυο είναι βεβαρημένο, δεν καταγράφεται απλώς η ύπαρξη της σχέσης ανάμεσα στις δύο μεταβλητές αλλά και η βαρύτητα της σχέσης αυτής. Συνεπώς πέραν από την ύπαρξη της σχέσης, ενδεχομένως και το πρόσημο της, σε κάθε συσχέτιση αντιστοιχεί και ένα βάρος το οποίο εκφράζει το πόσο ισχυρή η είναι η σχέση ανάμεσα στους δύο κόμβους. Η συγκεκριμένη ιδιότητα είναι αρκετά σημαντική καθώς επιτρέπει στον διαχωρισμό των σχέσεων ως προς την ισχύ τους. Σε ένα μη βεβαρημένο δίκτυο όλες οι σχέσεις είναι εξίσου ισχυρές καθώς καταγράφεται απλώς η ύπαρξη τους και όχι η βαρύτητα τους, συμπληρώνοντας απλώς μία μονάδα στον πίνακα γειτνίασης. Ως αποτέλεσμα πολύ ισχυρές σχέσεις ανάμεσα σε κόμβους εξισώνονται με όχι και

τόσο ισχυρές ή οριακά υπαρκτές σχέσεις. Χάνεται συνεπώς σημαντική πληροφορία. Σε έναν βεβαρημένο δίκτυο κατά την κατασκευή του πραγματοποιείται όχι μόνο υπολογισμός της σχέσης, αλλά και υπολογισμός της βαρύτητας της.

Γενικά ενδείκνυται τα βάρη ενός δικτύου να είναι κανονικοποιημένα. Εφόσον ο γράφος εκφράζει τις σχέσεις σε ένα ενιαίο πλαίσιο-δίκτυο, τα βάρη πρέπει να είναι συγκρίσιμα και να αναφέρονται σε αντίστοιχες τάξεις μεγέθους. Συνεπώς η βέλτιστη λύση είναι όλα τα βάρη να εκφράζονται στο διάστημα 0 έως 1, με την ερμηνεία της μονάδας να εξαρτάται από τον τρόπο κατασκευής του δικτύου. Εάν για παράδειγμα το δίκτυο έχει κατασκευαστεί με τον συντελεστή pearson, η μονάδα ως βάρος μίας ακμής αντιστοιχεί στην πλήρη ευθυγράμμιση-συσχέτιση των μεταβλητών, ενώ εάν τα βάρη αποτελούν αποτέλεσμα του συντελεστή παλινδρόμησης και διαιρεθούν όλα τα βάρη με τον μέγιστο συντελεστή για να είναι κανονικοποιημένα, τότε η μονάδα αντιστοιχεί στην ακμή με την μεγαλύτερη ισχύ έναντι των άλλων. Σε κάθε περίπτωση η κανονικοποίηση είναι θεμιτή, και η χρήση βεβαρημένου δικτύου προσδίδει περισσότερη πληροφορία από την χρήση ενός μη βεβαρημένου δικτύου.

Προκειμένου να εκφραστούν τα βάρη το διάνυσμα όπου φιλοξενεί τις ακμές του πίνακα συνοδεύεται από ένα διάνυσμα το οποίο φιλοξενεί το βάρος της κάθε ακμής. Συμπληρωματικά και με τον τρόπο που ακολουθείται στην υπό μελέτη εργασία, τα βάρη φιλοξενούνται στον πίνακα γειτνίασης. Έτσι ο πίνακας γειτνίασης αντί να παίρνει τιμές 0 ή 1, παίρνει τιμές από 0 έως 1, όπου εάν ένα στοιχείο του πίνακα γειτνίασης είναι διάφορο του μηδενός δεν δίνεται απλώς πληροφορία για την ύπαρξη συσχέτισης, αλλά και για την ένταση της

Μη Κατευθυνόμενα και Κατευθυνόμενα δίκτυα ~ Συσχετίσεις και Επιδράσεις

Οι σχέσεις που εκφράζουν οι ακμές του δικτύου που μπορεί γενικά να αναπαριστούν την συσχέτιση των μεταβλητών, μπορεί όμως η αναπαράσταση της σχέσης να γίνει πιο εξειδικευμένη και έτσι ο γράφος να φιλοξενεί τις επιδράσεις που ασκεί η μία μεταβλητή από την άλλη. Στην περίπτωση της απλής συσχέτισης για κάθε ζεύγος κόμβων γίνεται να διατυπωθεί το πολύ μία ακμή, ενώ στην περίπτωση της καταγραφής επίδρασης γίνεται να υπάρχουν το πολύ δύο ακμές, δηλαδή η επίδραση της μεταβλητής A στην μεταβλητή B και η επίδραση της μεταβλητής B στην μεταβλητή A.

Από την παραπάνω διάκριση ανάμεσα στην καταγραφή αμφίδρομης συσχέτισης ή την καταγραφή της αλληλεπίδρασης των μεταβλητών προκύπτει και η διάκριση του δικτύου σε κατευθυνόμενο και μη κατευθυνόμενο δίκτυο. Στην περίπτωση της καταγραφής συσχέτισης, δηλαδή σε ένα μη κατευθυνόμενο δίκτυο, αναμένεται ο πίνακας γειτνίασης να είναι συμμετρικός και συνεπώς η ανάλυση μπορεί να αρκестεί στο τρίγωνο κάτω αριστερά ή πάνω δεξιά της διαγώνιου.

Στην περίπτωση καταγραφής επιδράσεων ανάμεσα στις μεταβλητές χρειάζεται το σύνολο του πίνακα γειτνίασης για την έκφραση των αλληλεπιδράσεων. Σε αυτήν την περίπτωση η ανάγνωση του πίνακα γειτνίασης γίνεται είτε ανά γραμμή, είτε ανά στήλη. Σε περίπτωση ανάγνωσης ανά γραμμή η κάθε γραμμή αναπαριστά τις επιδράσεις του κόμβου που αντιστοιχεί στην εκάστοτε γραμμή προς τους κόμβους που βρίσκονται στην κάθε στήλη. Στην περίπτωση ανάγνωσης ανά στήλη ο κάθε στήλη αναπαριστά τις επιδράσεις που δέχεται ο κόμβος της στήλης από τον κόμβο της εκάστοτε γραμμής.

Ανάλογα με τον τρόπο κατασκευής ή μέτρησης του δικτύου παράγεται και η αντίστοιχη κατηγορία δικτύου. Πιο συγκεκριμένα εάν ο υπολογισμός των διασυσχετίσεων γίνει με δείκτη ο οποίος υπολογίζει απλή αμφίδρομη συσχέτιση μεταξύ των μεταβλητών, το δίκτυο το οποίο θα προκύψει θα είναι μη κατευθυνόμενο (συμμετρικό), εάν ο υπολογισμός των διασυσχετίσεων γίνει πιο εξειδικευμένα, δηλαδή με δείκτη που εκφράζει την επίδραση που ασκεί η μία μεταβλητή στην άλλη, το δίκτυο που θα προκύψει θα είναι κατευθυνόμενο (μη συμμετρικό).

Γενικά ένα κατευθυνόμενο δίκτυο παρέχει περισσότερη πληροφορία από ένα συμμετρικό. Αυτό συμβαίνει διότι στην περίπτωση του κατευθυνόμενου δικτύου μπορεί μία διασυσχέτιση να εκφραστεί με δύο σχέσεις, ενώ στην περίπτωση του μη κατευθυνόμενου με μία, γενικεύοντας έτσι την έκφραση της σχέσης μεταξύ των μεταβλητών. Έτσι εάν φεροιπίν σε ένα μη κατευθυνόμενο δίκτυο έχουμε μία ακμή ανάμεσα σε δύο κόμβους, σε ένα κατευθυνόμενο δίκτυο αυτή η ακμή θα μπορούσε να αντιστοιχεί σε επίδραση της A μεταβλητής στην B ή σε επίδραση της B μεταβλητής στην A ή σε αλληλεπίδραση των μεταβλητών. Προκύπτει δηλαδή ότι από ένα κατευθυνόμενο δίκτυο θα μπορούσαμε να κατασκευάσουμε ένα δίκτυο αμφίδρομων συσχετίσεων, καθώς για την έκφραση συσχέτισης θα αρκούσε είτε επίδραση της A στην B, είτε επίδραση της B στην A. Από ένα μη κατευθυνόμενο δίκτυο όμως δεν θα μπορούσαμε να κατασκευάσουμε ένα κατευθυνόμενο, καθώς δεν γνωρίζουμε εάν η ύπαρξη ακμής (δηλαδή αμφίδρομης συσχέτισης) αντιστοιχεί σε επίδραση της A στην B ή σε επίδραση της B στην A.

Επιπρόσθετα, η καταγραφή αμφίδρομων συσχετίσεων και συνεπώς η κατασκευή ενός μη κατευθυνόμενου δικτύου, ενδέχεται να περιέχει θόρυβο ή μερολήψια ως προς την ύπαρξη συσχέτισης. Πιο συγκεκριμένα ενδέχεται η καταγραφή της συσχέτισης να μην εκφράζει άμεση συσχέτιση μεταξύ των μεταβλητών, αλλά μία αμοιβαία επίδραση ή αμοιβαία τάση που ‘βιώνουν’ οι δύο μεταβλητές. Εάν υπάρχει μία μεταβλητή Γ που ασκεί όμοιας τάξης επίδραση στην μεταβλητή Α και Β, μπορεί οι δύο μεταβλητές να φαίνεται πως συσχετίζονται, ενώ στην πραγματικότητα δεν σχετίζονται άμεσα και απλώς βιώνουν μία αμοιβαία επίδραση. Επιπλέον, εάν και οι δύο μεταβλητές βρίσκονται σε μία σταθερή ανοδική πορεία ή σταθερή καθοδική πορεία, ακολουθούν δηλαδή κάποια τάση, ο απλός υπολογισμός συσχέτισης ενδέχεται να εκφράσει συσχέτιση ενώ στην πραγματικότητα δεν υπάρχει.

Στην εικόνα 1 βλέπουμε αριστερά ένα μη κατευθυνόμενο δίκτυο συμμετρικό. Τα βεβαρημένα δίκτυα έχουν βάρη στο διάστημα 0 έως 1, με την ερμηνεία της μονάδας να εκφράζεται στο διάστημα 0 έως 1

Εικόνα 1: Κατηγορίες δικτύων

Μη Κατευθυνόμενο Δίκτυο					Κατευθυνόμενο Δίκτυο				
<u>Χωρίς Βάρη</u>					<u>Χωρίς Βάρη</u>				
	χ	ψ	ω	ζ		χ	ψ	ω	ζ
χ	0				χ	0	0	1	0
ψ	1	0			ψ	1	0	0	1
ω	1	0	0		ω	0	1	0	1
ζ	0	1	1	0	ζ	0	0	1	0
<u>Βεβαρημένο</u>					<u>Βεβαρημένο</u>				
	χ	ψ	ω	ζ		χ	ψ	ω	ζ
χ	0				χ	0	0	0.23	0
ψ	0.63	0			ψ	0.33	0	0	0.1
ω	0.26	0	0		ω	0	0.64	0	0.92
ζ	0	0.83	0.92	0	ζ	0	0	0.7	0

4.1.1 Κοινότητες

Οι κοινότητες ενός δικτύου αποτελούν υποδίκτυο ή ομάδες κόμβων οι οποίες έχουν εντονότερες σχέσεις μεταξύ τους από ότι έχουν με το υπόλοιπο δίκτυο. Αποτελούν δηλαδή μέρη του γραφήματος των οποίων οι κόμβοι τείνουν να σχετίζονται εντονότερα μεταξύ τους(σε όρους βαρύτητας ακμών) από ότι σχετίζονται με κόμβους εκτός της κοινότητας. Δεν συναντάται σε όλους τα δίκτυα κοινότητας, όταν όμως τις συναντάμε μπορούν να δώσουν σημαντική πληροφορία αναφορικά με την τοπολογία και την υπόσταση του δικτύου. Επιπλέον αποτελούν εργαλείο στον εντοπισμό παραγόντων (factors-clusters) στην ανάλυση χρονοσειρών, γεγονός αρκετά σημαντικά ειδικότερα στην ανάλυση μεγάλων δεδομένων. (Ferreira & Zhao, 2016).

Ένα δίκτυο δεν αποτελείται απαραίτητα από κοινότητες, ή ακόμα και να αποτελείται αυτές μπορεί να είναι ασθενείς, δηλαδή η διαφορά της βαρύτητας των ακμών μεταξύ των μελών της να μην είναι σημαντικά μεγαλύτερη από τις ακμές των μελών με κόμβους εκτός κοινότητας. Προκειμένου να εξεταστεί κατά πόσο η δομή ενός δικτύου βασίζεται σε κοινότητες χρησιμοποιείται ο δείκτης modularity. Ο συγκεκριμένος δείκτης μετράει πόσο δυνατή είναι η διαίρεση ενός δικτύου σε κοινότητες.

Ο συγκεκριμένος δείκτης παίρνει τιμές από -1 έως 1. Το 0 υποδηλώνει ότι οι ενδεχόμενες κοινότητες που εντοπίστηκαν θα μπορούσαν κάλλιστα να είχαν δημιουργηθεί τυχαία και συνεπώς να μην ανταποκρίνονται σε ιδιαίτερη μορφολογία του δικτύου, άρα υποδηλώνει ότι το δίκτυο δεν απαρτίζεται από κοινότητες. Αρνητικές τιμές υποδηλώνουν πως οι διαίρεση του γράφου σε κοινότητες οδηγεί σε αρκετά τυχαίες κοινότητες γεγονός που αποκλείει την ύπαρξη κοινοτήτων. Αν η τιμή είναι θετική και όσο πιο κοντά είναι στην μονάδα, φαίνεται πως η τοπολογία του γράφου είναι όντως δομημένη σε κοινότητες και οι κόμβοι έχουν την τάση να δημιουργούν ομάδες όπου οι μεταξύ τους συνδέσεις είναι σημαντικά ισχυρότερες από τις συνδέσεις που φέρουν με κόμβους εκτός κοινότητας.

Τεχνικά, από τον εντοπισμό και ανάλυση κοινοτήτων στα πλαίσια της θεωρίας δικτύων μπορούμε να πάρουμε τρεις βασικές πληροφορίες. Πόσες κοινότητες υπάρχουν, από πόσους κόμβους απαρτίζεται η κάθε κοινότητα και τέλος ποια είναι τα μέλη της κάθε κοινότητας.

Η σημασία και ερμηνεία των κοινοτήτων ανάγεται στο πεδίο μελέτης. Φεροιπίν σε ένα μικρό χωριό, σχεδόν όλοι οι κάτοικοι έχουν ακμές-σχέσεις μεταξύ τους, αλλά δημιουργούνται κοινότητες, δηλαδή ομάδες όπου τα μέλη έχουν ισχυρότερες σχέσεις μεταξύ τους ισχυρότερες από ότι με τους λοιπούς κατοίκους του χωριού. Τέτοιες ομάδες μπορεί να είναι τα νοικοκυριά, το καφενείο, ο συνεταιρισμός ή το σχολείο.

Σε επίπεδο ανάλυσης χρονοσειρών οικονομικού ενδιαφέροντος, οι κοινότητες έχουν ενδιαφέρον, καθώς εάν φεροιπίν αναζητούνται ερμηνευτικές μεταβλητές για μία χρονοσειρά, η εστίαση του ερευνητή θα έπρεπε να στραφεί στα μέλη της κοινότητας που αυτή η χρονοσειρά ανήκει.

Στην περίπτωση του δείκτη πολυπλοκότητας προϊόντος, η κοινότητα θα δείξει ποιες ομάδες αγαθών τείνουν να αναπτύσσουν την τεχνογνωσία τους ταυτόχρονα. Έτσι μπορούν να αναδειχθούν είτε σχετισμοί που βασίζονται σε κλάδους, είτε σχετισμοί που βασίζονται σε συμπληρωματικά ως προς την κατασκευή προϊόντος αγαθά. Τα αποτελέσματα κοινοτήτων μπορούν να δώσουν αρκετή πληροφορία για την μελέτη των σχετισμών της εφοδιαστικής αλυσίδας.

4.1.2 Σημαντικότητα των κόμβων

Βασική ιδιότητα της θεωρίας δικτύων είναι ότι δίνει την δυνατότητα αξιολόγησης των κόμβων. Οι κεντρικότητες που θα αναπτυχθούν παρακάτω δύναται να χρησιμοποιηθούν ως δείκτες προκειμένου να εκφράσουν την καθοριστικότητα (definiteness) και την σημασία ενός κόμβου στο δίκτυο. Όσο υψηλότερες τιμές ένας κόμβος συγκεντρώνει στις αντίστοιχες κεντρικότητες του, τόσο πιο νευραλγική κρίνεται η θέση του στο δίκτυο. Έτσι μπορεί κανείς να δει μέσω της καθοριστικότητας των κόμβων που μπορεί να εστιάσει στην βελτίωση ενός δικτύου, η να αναζητήσει την πηγή των φαινομένων που εξελίσσονται στο δίκτυο (Friedkin, 1991).

Η βαθμική κεντρικότητα δείχνει το σύνολο των συνδέσεων του κόμβου, ενώ η ιδιοκεντρικότητα συνυπολογίζει και το πόσο σημαντικές είναι αυτές οι συνδέσεις. Η εγγύτητα δείχνει πόσο απέχει ο κόμβος από τους υπόλοιπους, ενώ η διαμεσότητα πόσα μονοπάτια περνούν από τον κόμβο. Η ερμηνεία του κάθε δείκτη καθώς και η σημασία του καθορίζεται από το πεδίο εφαρμογής και το εννοιολογικό πλαίσιο του δικτύου.

Βαθμική κεντρικότητα (Degree Centrality)

Ως βαθμός $\deg(v)$ ορίζεται το σύνολο των γειτόνων ή το σύνολο των ακμών ενός κόμβου.

$$\text{Βαθμική κεντρικότητα} = \frac{\deg(v)}{n-1} \quad (1)$$

Προκειμένου να πάρουμε την βαθμική κεντρικότητα, διαιρείται ο βαθμός ενός κόμβου με το σύνολο των ακμών που δύναται ο κόμβος να έχει. Αποτελεί δηλαδή μία κανονικοποίηση του βαθμού. Σε δίκτυα με βάρη στον υπολογισμό της βαθμικής κεντρικότητας συνυπολογίζεται και το βάρος της ακμής, δηλαδή η ισχύς της. Σε κατευθυνόμενα δίκτυα μπορούμε να έχουμε έσω-βαθμό και έξω-βαθμό δηλαδή ακμές που εξέρχονται και ακμές που εισέρχονται από τον κόμβο, ή και τις συνολικές ακμές.

Ο συγκεκριμένος δείκτης στο επεξεργαστικό περιβάλλον της R με την χρήση της βιβλιοθήκης `igraph` υπολογίζεται με την χρήση της παρακάτω εντολής:

`Degree(G,normalized=TRUE)`

Ιδιοκεντρικότητα (Eigencentality)

Η ιδιοκεντρικότητα βασίζεται στον βαθμική κεντρικότητα και αποτελεί μία εξελιγμένη μορφή της. Δεν λαμβάνει υπ' όψιν μόνο τον αριθμό των γειτόνων ενός κόμβου αλλά και πόσους γείτονες έχει ο γείτονας. Το συγκεκριμένο μέτρο ουσιαστικά σταθμίζει τον βαθμική κεντρικότητα με την βαθμική κεντρικότητα των γειτόνων. Όσο υψηλότερο βαθμό έχουν οι γείτονες ενός κόμβου, τόσο υψηλότερη τιμή παίρνει. Αποτελεί δηλαδή ένδειξη όχι μόνο της άμεσης διασύνδεσης του κόμβου αλλά και το κατά πόσο συνορεύει με σημαντικούς γείτονες. Έτσι εάν έχουμε δύο κόμβους με μόλις μία ακμή ο καθένας (δηλαδή η βαθμική κεντρικότητα τους ισοψηφίζει) η ιδιοκεντρικότητα

θα έρθει να δείξει το πόσο σημαντικός είναι ο γείτονας του κάθε κόμβου. Βασίζεται στις ιδιοτιμές του πίνακα γειτνίασης.

Ο συγκεκριμένος δείκτης στο επεξεργαστικό περιβάλλον της R με την χρήση της βιβλιοθήκης `igraph` υπολογίζεται με την χρήση της παρακάτω εντολής:

```
eigen_centrality(G,directed=FALSE,scale=TRUE,weights=NULL)
```

Κεντρικότητα Εγγύτητας (Closeness centrality)

Αρχικά υπολογίζεται η μέση απόσταση:

$$l_i = \frac{1}{n-1} \sum_{i \neq j} d_{ij} \quad (2)$$

Με d_{ij} συμβολίζουμε την απόσταση του κόμβου i από τον κόμβο j , δηλαδή το μικρότερο μονοπάτι που ενώνει τους δύο κόμβους. Ως μονοπάτι νοείται ένα σύνολο ακμών που ξεκινάει από έναν κόμβο και καταλήγει σε έναν άλλον. Ο συγκεκριμένος δείκτης εκφράζει την μέση απόσταση ενός κόμβου από όλους τους υπολοίπους. Δηλαδή όσο υψηλότερη είναι η τιμή του, τόσο πιο απομακρυσμένος είναι ο δείκτης (έχει δηλαδή αρνητική σημασία ως προς την καθοριστικότητα του κόμβου).

Η κεντρικότητα εγγύτητας είναι το αντίστροφο μέγεθος της μέσης απόστασης ενός κόμβου από τους υπολοίπους:

$$C_i = \frac{1}{l_i} \quad (3)$$

Όσο υψηλότερη κεντρικότερη εγγύτητα, τόσο πιο κοντά έγγυται στους υπόλοιπους κόμβους και τόσο μεγαλύτερη είναι η ένδειξη ότι ο συγκεκριμένος κόμβος είναι κεντρικός ως προς την τοπολογία του δικτύου.

Ο συγκεκριμένος δείκτης στο επεξεργαστικό περιβάλλον της R με την χρήση της βιβλιοθήκης `igraph` υπολογίζεται με την χρήση της παρακάτω εντολής:

```
Closeness(G, normalized=TRUE)
```

Διαμεσότητα- Ενδιάμεση Κεντρικότητα (Betweenes Centrality)

Η διαμεσότητα μας δείχνει τι ποσοστό από το σύνολο των μονοπατιών που συνδέουν τους υπόλοιπους κόμβους μεταξύ τους, περνάει από τον κόμβο αναφοράς. Μελετώντας τον κόμβο i , έχουμε $\mu_{κλ}$ το σύνολο των μονοπατιών που συνδέουν τον κόμβο $κ$ με τον κόμβο $λ$ και $\nu_{κλ}$ το σύνολο των μονοπατιών που συνδέουν τον κόμβο $κ$ με τον κόμβο $λ$ και περνάνε από τον κόμβο i . Δηλαδή το ν είναι υποσύνολο του μ .

Εξετάζουμε λοιπόν τι ποσοστό των συνολικών μονοπατιών του δικτύου περνάει από τον συγκεκριμένο κόμβο:

$$\delta_i = \sum_{κ.λ} \frac{\nu_{κλ}}{\mu_{κλ}} \quad (4)$$

Ο συγκεκριμένος δείκτης στο επεξεργαστικό περιβάλλον της R με την χρήση της βιβλιοθήκης `igraph` υπολογίζεται με την χρήση της παρακάτω εντολής:

```
Betweenness(G, directed=FALSE, normalized=TRUE)
```

Η αξιολόγηση κόμβων παρέχει σημαντικές πληροφορίες κατά την μελέτη ενός πεδίου. Αρχικά εντοπίζοντας τους κόμβους με την χαμηλότερη κέντρωση ο ερευνητής μπορεί να γλυτώσει χρόνο από την ενασχόληση με όχι τόσο σημαντικές παραμέτρους του προβλήματος που μελετάει. Από εκεί και πέρα ο εντοπισμός των πιο κεντρικών κόμβων μπορεί να κάνει την διαδικασία έρευνας πιο στοχευμένη. Εάν φεροιπίν κάποιος ερευνητής αναζητά ερμηνευτικές μεταβλητές, μπορεί να ξεκινήσει από τις μεταβλητές με την μεγαλύτερη κέντρωση, καθώς είναι πιο πιθανό να έχουν επίδραση ως οι σημαντικότερες του δικτύου.

Η χρησιμότητα της επίγνωσης των σημαντικών κόμβων όμως ξεπερνούν την διμεταβλητότητα. Σε περίπτωση που θέλει κανείς να επηρεάσει το δίκτυο, δηλαδή το σύνολο των μεταβλητών προς μία κατεύθυνση, μπορεί να εστιάσει στο να επηρεάσει άμεσα τις μεταβλητές με την μεγαλύτερη κέντρωση, οι οποίες έμμεσα θα επιδράσουν στις υπόλοιπες. Αν σκοπός είναι η επίθεση σε δίκτυο, αυτή πρέπει να γίνει στις μεταβλητές με την μεγαλύτερη κέντρωση. Αν πάλι η επίθεση (ή η αρνητική εξωτερικότητα) γίνει αναγκαστικά, συνειδητά, αλλά αθέμιτα, προκειμένου να μετριαστεί η ζημεία στο δίκτυο μπορούν να επιλεγθούν οι μεταβλητές με την χαμηλότερη κεντρικότητα (Friedkin, 1991).

Σε περίπτωση που κάποιος επιθυμεί να προστατέψει ένα δίκτυο, οφείλει αρχικά να προστατέψει τους κόμβους με την μεγαλύτερη κέντρωση, διότι αν για παράδειγμα αρχίσουν αυτοί να έχουν αρνητικές μεταβολές, τότε αυτές θα μεταφερθούν και στο υπόλοιπο δίκτυο. Αν μία παράμετρος ενός συστήματος αναμένεται να αποκτήσει κάποιο πρόβλημα, πρέπει να είναι γνωστή η κεντρικότητα της έτσι ώστε να υπάρχει προϋδεασμός για την ένταση των επιπτώσεων του προβλήματος στο υπόλοιπο σύστημα.

Επιπλέον σημαντική πληροφορία σχετικά με την κεντρικότητα είναι ο εντοπισμός μηχανισμών, σχετισμών και τάσεων. Εάν για παράδειγμα υπάρχει μία τάση σε ένα σύνολο χρονοσειρών, είναι πολύ πιθανό να δημιουργείται ή να διαχέεται από την χρονοσειρά με την υψηλότερη κεντρικότητα.

Όσον αφορά τον δείκτη πολυπλοκότητας προϊόντων, η σημασία της κεντρικότητας ενός δείκτη, δηλαδή το επίπεδο τεχνογνωσίας που απαιτεί μία ομάδα προϊόντων για να παραχθεί, είναι μεγάλη. Μέσω της κεντρικότητας μπορούν να ληφθούν ενδείξεις όσον αφορά την γνωσιακή εξάρτηση μεταξύ διαφορετικών πεδίων, δηλαδή από ποιο πεδίο εξαρτάται η ανάπτυξη της κατάρτισης των άλλων πεδίων. Επιπλέον μέσα στην καταγραφή της συνολικής αλληλεξάρτησης τεχνογνωσίας ανάμεσα σε ομάδες προϊόντων που παρέχει το δίκτυο, οι πιο κεντρικοί κόμβοι περιμένουμε να είναι αυτοί που είναι οι πιο χρήσιμοι στους υπόλοιπους, καθώς επίσης οι πιο κεντρικές ομάδες προϊόντος να είναι αυτές που οι βασικές τους γνώσεις ή το γνωστικό τους πεδίο είναι πιο κοντά στο γνωστικό πεδίο των υπολοίπων πεδίων.

4.1.3 Κατασκευή δικτύου

Προκειμένου να κατασκευαστεί ένα δίκτυο αρκεί η καταγραφή όλων των στατιστικά σημαντικών συσχετίσεων μεταξύ των εν δυνάμει κόμβων-μεταβλητών, οι οποίες έπειτα θα εισαχθούν στον πίνακα γειτνίασης. Ο τρόπος που θα πραγματοποιηθεί η παραπάνω διαδικασία εξαρτάται κυρίως από το είδος του δικτύου που θα κατασκευαστεί και τον δείκτη ο οποίος θα χρησιμοποιηθεί. Για την ακρίβεια δεν είναι συμβατοί όλοι οι δείκτες με όλα τα είδη δικτύου και ο δείκτης αποτελεί πολλές φορές ερευνητικό περιορισμό ως προς το βέλτιστο δυνατό αποτελέσματα. Δεδομένου ότι για την κατασκευή του δικτύου θα χρησιμοποιηθεί ο συντελεστής γραμμικής συσχέτισης Pearson, το καλύτερο αποτέλεσμα που μπορεί να παραχθεί είναι ένα μη κατευθυνόμενο δίκτυο (συμμετρικό) με βάρη.

Ο τρόπος κατασκευής του δικτύου είναι ο εξής: όποιο ζεύγος μεταβλητών έχει στατιστικά σημαντική διασυσχέτιση, η τιμή του συντελεστή συσχέτισης εισάγεται ως βάρος στον πίνακα γειτνίασης.

Ο συντελεστής συσχέτισης του Pearson

Βασικό εργαλείο εντοπισμού συσχέτισης ανάμεσα σε δύο μεταβλητές είναι η συνδιακύμανση ($\sigma_{\chi\psi}$).

Η συνδιακύμανση εντοπίζει και εκφράζει αμοιβαίες συμπεριφορές των μεταβλητών ως προς την θέση των παρατηρήσεων σε σχέση με τον μέσο όρο. Ο τύπος της συνδιακύμανσης δίνεται από την παρακάτω σχέση (Taylor, 1990):

$$\sigma_{\chi\psi} = \frac{\sum_{i=1}^N (\chi_i - \bar{\chi})(\psi_i - \bar{\psi})}{N-1} \quad (5)$$

Η λειτουργία της συνδιακύμανσης μπορεί να εκφραστεί ως η μέση αμοιβαία απόκλιση των μεταβλητών από τον μέσο όρο τους. Παρατηρούμε στον αριθμητή ότι βασικά στοιχεία της είναι η διαφορά της παρατήρησης από τον μέσο όρο. Για κάθε παρατήρηση υπολογίζεται το γινόμενο της απόκλισης των δύο μεταβλητών και υπολογίζεται το άθροισμα των αμοιβαίων αποκλίσεων όλων των παρατηρήσεων. Όσο μεγαλύτερη είναι η απόκλιση της κάθε μεταβλητής, τόσο μεγαλύτερη η τιμή του αριθμητή. Δεδομένου ότι έχουμε το γινόμενο των αποκλίσεων, εάν οι δύο μεταβλητές παρουσιάζουν ταυτόχρονα μεγάλη απόκλιση από τον μέσο όρο, αυτό το φαινόμενο καταγράφεται έντονα καθώς ο αριθμητής καταγράφει το γινόμενο των αποκλίσεων και λειτουργεί πολλαπλασιαστικά και όχι αθροιστικά.

Η συνδιακύμανση μπορεί να χαρακτηριστεί ως η αμοιβαία τυπική απόκλιση, εκφράζει δηλαδή κατά πόσο οι δύο μεταβλητές έχουν την τάση να απέχουν ταυτόχρονα από τον μέσο όρο. Όσο μεγαλύτερο είναι το γινόμενο της απόκλισης, τόσο μεγαλύτερη είναι η αμοιβαία απόκλιση από τον μέσο όρο και συνεπώς τόσο μεγαλύτερη είναι η τάση των μεταβλητών να απομακρύνονται από τα μέσα επίπεδα τους.

Αφού υπολογιστεί το άθροισμα του γινομένου των αποκλίσεων των μεταβλητών από τον μέσο, το μέγεθος διαιρείται με N-1 προκειμένου να εκφραστεί η μέση αμοιβαία διακύμανση των μεταβλητών και το μέγεθος να κανονικοποιηθεί ως προς το μέγεθος του δείγματος και το αποτέλεσμα του αριθμητή να αναχθεί στο σύνολο των παρατηρήσεων.

Βασική προϋπόθεση είναι οι παρατηρήσεις να είναι ευθυγραμμισμένες ως προς την συχνότητα τους. Επιπλέον κρίνεται δόκιμο να υπολογίζεται σε μεταβλητές οι οποίες είναι όμοιας τάξης μεγεθών. Εάν εφαρμοστεί σε μεταβλητές οι οποίες έχουν διαφορετική τάξη μεγεθών, για παράδειγμα μία μεταβλητή που ανήκει στο διάστημα $[0,1]$ και μία μεταβλητή που έχει τιμές εκατομμυρίου, οι μεταβολές της πρώτης μεταβλητής θα θεωρηθούν αμελητέες από τον δείκτη.

Βασικό χαρακτηριστικό του συντελεστή είναι ο υπολογισμός του πρόσημού. Όπως θα εξηγηθεί με την χρήση του παρακάτω πίνακα, η συνδιακύμανση μπορεί να έχει είτε θετικό είτε αρνητικό πρόσημο ανάλογα με το αν οι μεταβλητές κινούνται προς την ίδια κατεύθυνση. Εάν δηλαδή υπάρχει η τάση όταν η μεταβλητή χ είναι πάνω από τον μέσο όρο της, η μεταβλητή ψ πηγαίνει κάτω από τον μέσο όρο της, η συνδιακύμανση θα είναι αρνητική. Αυτό πραγματοποιείται λόγω του γινομένου του οποίου υπάρχει και για κάθε παρατήρηση του ζεύγους του μεταβλητών μπορεί να παραχθεί το παρακάτω πρόσημο.

Πρόσημο ($\chi_i - \bar{\chi}$)	Πρόσημο ($\psi_i - \bar{\psi}$)	Πρόσημο συνδιακύμανσης παρατηρήσεων
+	+	+
-	-	+
+	-	-
-	+	-

Ο συντελεστής συσχέτισης του Pearson δίνεται από την παρακάτω σχέση:

$$\rho(\chi, \psi) = \frac{\sigma_{\chi\psi}}{\sigma_{\chi}\sigma_{\psi}} \quad (6)$$

Ουσιαστικά εκφράζει τι ποσοστό της συνολικής διακύμανσης των μεταβλητών είναι αμοιβαία. Στον αριθμητή του φιλοξενείται η συνδιακύμανση η οποία εκφράζει την αμοιβαία διακύμανση, ενώ στον αριθμητή στέκεται το γινόμενο των τυπικών αποκλίσεων το οποίο εκφράζει την συνολική μεταβλητότητα των δύο μεταβλητών. Όσο μεγαλύτερη είναι η συνδιακύμανση, τόσο μεγαλύτερο είναι το ποσοστό της συνολικής μεταβλητότητας που είναι αμοιβαία, συνεπώς τόσο περισσότερο σχετίζονται η μεταβλητές. (Taylor, 1990)

Η συσχέτιση δηλαδή υπολογίζεται ως η τάση των μεταβλητών να κυμαίνονται ταυτόχρονα μακριά από τον μέσο σε σχέση με το πόσο κυμαίνονται μακριά από τον μέσο η κάθε μία ξεχωριστά. Χρησιμοποιείται η αμοιβαία μεταβλητότητα ως ποσοστό της συνολικής μεταβλητότητας για να εντοπίσει την αμοιβαία συμπεριφορά των μεταβλητών και συνεπώς το αν σχετίζονται η όχι.

Ο αναλυτικός τύπος του συντελεστή δίνεται από την παρακάτω σχέση:

$$\rho(\chi, \psi) = \frac{\sum_{i=1}^N (\chi_i - \bar{\chi})(\psi_i - \bar{\psi})}{\sqrt{(\chi_i - \bar{\chi})^2} \sqrt{(\psi_i - \bar{\psi})^2}} \quad (7)$$

Η βασικές πληροφορίες που παίρνουμε από τον συντελεστή έχουν να κάνουν με το απόλυτο μέγεθος του και το πρόσημο του. Ο συντελεστής όντας ποσοστό ανήκει στο διάστημα $[0,1]$. Όσο πιο έντονη είναι η συσχέτιση τόσο μεγαλύτερη η τιμή του. Εάν η απόλυτη τιμή του είναι φεροιπίν 0.5, αυτό σημαίνει πως το 50% της συνολικής μεταβλητότητας των χρονοσειρών είναι αμοιβαία. Σε περίπτωση που ο συντελεστής είναι 0, η συνδιακύμανση είναι 0 και συνεπώς δεν εντοπίζεται καθόλου η τάση των μεταβλητών να απέχουν ταυτόχρονα από τον μέσο τους όρο, αρά υπό το συγκεκριμένο πρίσμα, δεν φαίνεται να σχετίζονται. Αν η τιμή του συντελεστή είναι ίση με 1, αυτό σημαίνει πως η συμπεριφορά των μεταβλητών ευθυγραμμίζεται πλήρως, δηλαδή κάθε η συνολική μεταβλητότητα προκαλείται ταυτόχρονα και για τις δύο μεταβλητές. Συνεπώς η τιμή του συντελεστή μας δίνει την ένταση της συσχέτισης.

Επιπλέον, το πρόσημο του συντελεστή μας δείχνει το είδος της συσχέτισης, αν δηλαδή οι μεταβλητές σχετίζονται αρνητικά ή θετικά. Το πρόσημο του συντελεστή εξαρτάται καθαρά από τον αριθμητή καθώς ο παρονομαστής που αποτελείται από τις δύο τυπικές αποκλίσεις, οι οποίες είναι πάντα θετικές, είναι πάντα θετικός. Συνεπώς το πρόσημο δίνεται από τον αριθμητή, την συνδιακύμανση. Εάν το πρόσημο είναι θετικό σημαίνει πως η αμοιβαία μεταβλητότητα είναι προς την ίδια κατεύθυνση, δηλαδή όταν η μία μεταβλητή ανεβαίνει πάνω από τον μέσο όρο της η δεύτερη μεταβλητή ανεβαίνει και αυτήν πάνω από τον δικό της, ενώ αν η μία μεταβλητή κατεβαίνει κάτω από τον μέσο όρο της, η δεύτερη μεταβλητή έχει την τάση να κατεβαίνει και αυτή. Άρα το πρόσημο του συντελεστή δείχνει κατά πόσο οι μεταβλητές τείνουν να κινούνται προς την ίδια ή προς την αντίθετη κατεύθυνση.

Ο συντελεστής Pearson εντοπίζει μόνο γραμμικές σχέσεις. Αυτό σημαίνει πως εάν μία ευκλείδια απόκλιση της μίας μεταβλητής από τον μέσο συνοδεύεται από μία εκθετική ή λογαριθμική

απόκλιση της δεύτερης μεταβλητής από τον μέσο, η συνδιακύμανση και συνεπώς ο συντελεστής δεν θα μπορέσουν να την αποτυπώσουν. Αυτό αποτελεί σαφή περιορισμό του συντελεστή, ειδικότερα για μεγέθη με πιο χαοτική συμπεριφορά τα οποία μπορεί ενίοτε να έχουν την τάση να αποκλίνουν ταυτόχρονα από τον μέσο, αλλά με τάξεις μεγέθους που η γραμμικότητα δεν μπορεί να ακολουθήσει. Συνεπώς αναφερόμαστε σε ένα συντελεστή ο οποίος εντοπίζει γραμμικές συσχετίσεις. (Asuero et al. , 2006)

Στατιστικός έλεγχος του συντελεστή συσχέτισης

Σημαντικό κομμάτι της ανάλυσης είναι ο έλεγχος στατιστικής σημαντικότητας του συντελεστή Pearson. Η εκτίμηση της γραμμικής συσχέτισης μπορεί να εμπεριέχει θόρυβο και υπερεκτίμηση της σχέσης η οποία αποδίδεται σε τυχαία παρόμοια συμπεριφορά των μεταβλητών και όχι σε πραγματικό μηχανισμό συσχέτισης. Επιπλέον η αξιοπιστία της συμπερασματολογίας για την συσχέτιση εξαρτάται και από την δυναμική του δείγματος, δηλαδή όσο υψηλότερο είναι το δείγμα (οι πληροφορίες που δίνονται στην σχέση) τόσο πιο αξιόπιστο κρίνεται το αποτέλεσμα της.

Προκειμένου να γίνει στατιστική αξιολόγηση, κυρίως του παράγοντα του δείγματος και να εξαχθεί εάν ο συντελεστής συσχέτισης είναι στατιστικά σημαντικός, χρησιμοποιείται ο έλεγχος t για τον συντελεστή Pearson (Assuero et al., 2006). Πιο συγκεκριμένα υποθέτουμε ότι ο συντελεστής pearson ακολουθεί την κατανομή t-student. Παρακάτω παρουσιάζεται η δομή του στατιστικού ελέγχου :

Μηδενική υπόθεση (H_0): $\rho=0$

Δηλαδή η μηδενική υπόθεση υποθέτει απουσία συσχέτισης ανάμεσα στις μεταβλητές, στο τέλος του ελέγχου παράγεται πληροφορία σχετικά με το εάν θα απορριφθεί ή όχι η συσχέτιση, ή πιο συγκεκριμένα καταλήγουμε στην πιθανότητα αυτή η υπόθεση να ισχύει. Ανάλογα και με το επίπεδο σημαντικότητας που επιλέγουμε, αν η τιμή p-value είναι χαμηλότερη του επιπέδου σημαντικότητας μπορούμε να συμπεράνουμε με α βεβαιότητα της ύπαρξης συσχέτισης.

Στατιστική ελέγχου t

Υποθέτουμε ότι ο συντελεστής συσχέτισης ακολουθεί κατανομή τύπου t-student, με το κέντρο (την μέση τιμή) της κατανομής να είναι. Δηλαδή προσδοκάτε ότι η υψηλότερη πιθανότητα είναι ο συντελεστής να είναι μηδέν. Συνεπώς υπολογίζουμε την στατιστική ελέγχου, προκειμένου να

δούμε την θέση του συντελεστή στην συγκεκριμένη κατανομή, αν δηλαδή ο συντελεστής συσχέτισης φεύγει αρκετά από το κέντρο της μηδενικής υπόθεσης. Σημειώνεται εδώ ότι η κατανομή έχει και θετικές και αρνητικές τιμές στον άξονα x , καθώς ο συντελεστής παίρνει και αυτός είτε αρνητικές είτε θετικές τιμές.

Για την πραγματοποίηση ελέγχου χρησιμοποιείται η στατιστική t η οποία ουσιαστικά τοποθετεί τον συντελεστή Pearson στον χώρο της κατανομής t -student προκειμένου να εξετασθεί εάν απέχει σημαντικά από την αναμενόμενη μηδενική τιμή. Η στατιστική ελέγχου έχει στον αριθμητή της το μέγεθος του δείγματος. Αυτό σημαίνει πως όσο μεγαλύτερο είναι το δείγμα, τόσο πιο εύκολα θα απομακρυνθεί η στατιστική ελέγχου από την τιμή της μηδενικής υπόθεσης αν ο συντελεστής συσχέτισης είναι υψηλός. Ο τύπος της στατιστικής ελέγχου είναι ο παρακάτω:

$$t = \rho * \sqrt{\frac{n-2}{1-\rho^2}} \quad (8)$$

Όπου n είναι το μέγεθος του δείγματος και η πρώτη παράγωγος της στατιστικής ελέγχου είναι θετική, γεγονός που σημαίνει ότι μία αύξηση του μεγέθους του δείγματος καθιστά το αποτέλεσμα του συντελεστή πιο αξιόπιστο.

Κρίσιμες τιμές

Οι κρίσιμες τιμές έχουν βασικό ρόλο στην απόφαση απόρριψης της μηδενικής υπόθεσης ή στην εκτίμηση της πιθανότητας ισχύς της μηδενικής υπόθεσης. Ουσιαστικά λαμβάνοντας υπ' όψιν το μέγεθος του δείγματος εξετάζουν εάν η στατιστική ελέγχου είναι αρκετά μακριά από την μηδενική υπόθεση έτσι ώστε αυτή να απορριφθεί. Συνεπώς οι βαθμοί ελευθερίας της κρίσιμης τιμής εξαρτώνται από το μέγεθος του δείγματος και δίνονται από την σχέση $\beta.ε.=n-2$. Η κρίσιμη τιμή παίρνεται από τον πίνακα της κατανομής t -student και εκφράζεται ως t_{n-2} .

P-value

Η βασική εκροή που εστιάζει ο ερευνητής κατά την εκπόνηση ενός στατιστικού ελέγχου με την χρήση κάποιου στατιστικού προγράμματος είναι η τιμή p -value. Η τιμή p -value μας δείχνει την πιθανότητα να ισχύει η μηδενική υπόθεση ή στην συγκεκριμένη περίπτωση, ποια είναι η πιθανότητα να μην υπάρχει συσχέτιση ανάμεσα στις δύο μεταβλητές. Αναφορικά με τον στατιστικό έλεγχο:

$$P - value = P(|t| > |t_{v-2}| = \text{Αληθές})$$

Το επίπεδο σημαντικότητας (α)

Το επίπεδο σημαντικότητας εκφράζει το πόσο αβέβαιοι είμαστε για την διατύπωση μας. Πιο συγκεκριμένο σε όρους στατιστικού ελέγχου δίνει την πιθανότητα να απορριφθεί η μηδενική υπόθεση, ενώ η μηδενική υπόθεση είναι αληθής:

$$\alpha = P(\text{Απόρριψη } H_0 | H_0 \text{ Αληθής})$$

Συνεπώς εάν εξαχθεί συμπερασματολογία με $\alpha=0.05$ ή 5%, αυτό σημαίνει πως υπάρχει 95% βεβαιότητα για το συμπέρασμα. Εάν το p-value είναι χαμηλότερο του α , τότε απορρίπτεται η μηδενική υπόθεση:

$$P - value < \alpha \rightarrow \text{Απόρριψη μηδενικής υπόθεσης}$$

Αλγοριθμική εκτέλεση

Συνεπώς, προκειμένου να κατασκευαστεί το δίκτυο, όποιο ζεύγος έχει στατιστικά σημαντική διασυσχέτιση και συνεπώς ισχύει η συνθήκη $P - value < \alpha$ δημιουργείται μία ακμή στο δίκτυο η οποία φέρει ως βάρος την τιμή του συντελεστή Pearson.

Σκοπός είναι η εισαγωγή ενός πίνακα δεδομένων (εν προκειμένου ένα σύνολο χρονοσειρών) διαστάσεων ($v \times k$) όπου v το σύνολο του δείγματος, δηλαδή των χρονικών παρατηρήσεων και k το σύνολο των μεταβλητών και η μετατροπή του πίνακα δεδομένων σε έναν πίνακα διασυσχετίσεων διαστάσεων ($k \times k$). Δεδομένου ότι το δίκτυο είναι συμμετρικό και οι συσχετίσεις αμφίδρομες αρκεί το κάτω αριστερά μέρος της διαγωνίου του πίνακα. Ο πίνακας διασυσχετίσεων θα φιλοξενεί ως τιμές τα βάρη των συσχετίσεων μόνο σε περίπτωση που ο συντελεστής γραμμικής συσχέτισης είναι στατιστικά σημαντικός. Εν συνεχεία ο συγκεκριμένος πίνακας διασυσχετίσεων θα χρησιμοποιηθεί ως πίνακας γειτνίασης προκειμένου να διεξαχθεί η τοπολογική ανάλυση της θεωρίας δικτύων ως προς τους σημαντικούς κόμβους και την ύπαρξη κοινοτήτων.

Προκειμένου να γίνει ο υπολογισμός των διασυσχετίσεων και των αντίστοιχων p-values, χρησιμοποιείται η βιβλιοθήκη “psych” (με την χρήση της γλώσσας R). Η συγκεκριμένη βιβλιοθήκη διαθέτει την συνάρτηση corr.test() στην οποία εισάγεται ο πίνακας δεδομένων ($v \times k$)

και εξάγονται δύο στοιχεία, ο πίνακας r (ο πίνακας διασυσχετίσεων διαστάσεων $(κ*κ)$) που φιλοξενεί τις τιμές των συντελεστών γραμμικής συσχέτισης και ο πίνακας p διαστάσεων $(κ*κ)$ που φιλοξενεί τα αντίστοιχα p -values τους.

Παρακάτω στην εικόνα 2 παρουσιάζεται ο αλγόριθμος ο οποίος χρησιμοποιώντας την παραπάνω συνάρτηση περνάει από όλα τα στοιχεία κάτω τις διαγωνίου του πίνακα p (με τα p -values), και όπου υπάρχει $p\text{-value} < \alpha$ (εν προκειμένου το α είναι 0.01 ή 1%) , τότε η τιμή του πίνακα r (με τις διασυσχετίσεις) εισάγεται στον πίνακα διασυσχετίσεων ο οποίος θα χρησιμοποιηθεί ως πίνακας γειτνίασης.

Εικόνα 2: Αλγόριθμος κατασκευής του δικτύου

```
library("psych")
xt=corr.test(sk)
pvalues=xt$p
lower_pv=lower.tri(pvalues)*pvalues
a=dim(lower_pv)
test=lower_pv
pcor=xt$r
dim(pcor)

for (i in 2:NROW(lower_pv))
{for (j in 1:NCOL(lower_pv)) {

  if (i>j){
    if (lower_pv[i,j]<= 0.01){
      test[i,j]=pcor[i,j]
    }else{
      test[i,j]=0
    }
  }
}
}
```

Στον παραπάνω αλγόριθμο εισάγεται ο πίνακας δεδομένων sk και παράγεται ο πίνακας διασυσχετίσεων $test$.

Τέλος για την κατασκευή του δικτύου στο πρόγραμμα R αρκεί η εισαγωγή του πίνακα διασυσχετίσεων $test$ ως πίνακα γειτνίασης με την χρήση της βιβλιοθήκης `igraph` και την εντολή `graph_from_adjacency_matrix()` , όπως φαίνεται στην εικόνα 3.

Εικόνα 3: Εντολή κατασκευής δικτύου

```
library(igraph)

G=graph_from_adjacency_matrix(
  test,
  mode = c("undirected"),
  weighted=TRUE
)
```

4.2 Ανάλυση χρονοσειρών

Οι χρονοσειρές είναι μεταβλητές των οποίων οι παρατηρήσεις αντιστοιχούν σε καταγραφή της τιμής μεταβλητής σε διαφορετικές χρονικές περιόδους, με την συχνότητα καταγραφής ή αλλιώς την συχνότητα των παρατηρήσεων να είναι σταθερή. Αποτελεί δηλαδή μία δυναμική έκφραση της μεταβλητής. Αν αναλογιστεί κανείς ότι για την καταγραφή των σχέσεων μεταξύ οντοτήτων είναι απαραίτητη η καταγραφή της επίδρασης της μίας στην άλλη ή καταγραφή των αμοιβαίων μεταβολών, ο χρόνος αποτελεί βασική διάσταση για την έκφραση των αλλαγών στην κατάσταση των μεταβλητών. Η χρονική έκφραση των μεταβλητών μπορεί να συμπληρώνεται και από διαστρωματική έκφραση (panel data).

4.2.1 Περιγραφική στατιστική

Μέτρα κεντρικής τάσης

Ο μέσος όρος αποτελεί θεμελιώδες μέτρο κεντρικής τάσης. Εκφράζει το μέσο επίπεδο των μεταβλητών, δηλαδή γύρω από ποια τιμή τείνουν να συγκεντρώνονται οι μεταβλητές, ως ένδειξη της κεντρικής τους τάσης. Δίνεται από τον παρακάτω τύπο:

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N} \quad (9)$$

Επιπλέον σημαντικό μέτρο κεντρικής τάσης αποτελεί η διάμεσος (Δ), η οποία εκφράζει την μεσαία παρατήρηση εάν διατάξουμε τις μεταβλητές σε αύξουσα σειρά. Τα δύο μέτρα κεντρικής τάσης στην κανονική κατανομή ταυτίζονται και στις υπόλοιπες κατανομές δρουν συμπληρωματικά. Γενικότερα σε μία κατανομή με έντονη λοξότητα επιλέγεται η διάμεσος ως ένδειξη της κεντρικής τάσης, καθώς επηρεαζόμενη λιγότερο από τις ακραίες τιμές, εντοπίζει καλύτερα τα επίπεδα υψηλής συγκέντρωσης της μεταβλητής. (Ταμπάκης & Χαψά, 2013)

Μέτρα διασποράς

Βασικό μέτρο διασποράς είναι η τυπική απόκλιση η οποία δίνεται από τον παρακάτω τύπο:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}} \quad (10)$$

Ο συγκεκριμένος δείκτης εκφράζει την μέση απόσταση των παρατηρήσεων από τον μέσο όρο, δηλαδή κατά μέσο όρο πόσο μακριά διασπείρονται οι παρατηρήσεις από το μέσο επίπεδο τους. Είναι σημαντική ένδειξη της μεταβλητότητας. Προκειμένου να εκφραστεί το απόλυτο μέτρο της απόστασης χωρίς πρόσημο, αρχικά αθροίζεται το τετράγωνο των αποστάσεων και έπειτα η αθροισμένη ποσότητα εισάγεται σε ρίζα έτσι ώστε να γυρίσει στα κανονικά της επίπεδα. Με αυτό τον τρόπο αποφεύγεται η απαλοιφή αποκλίσεων που έχουν ίδιο μέτρο αλλά διαφορετικό πρόσημο, κάτι που θα προκαλούσε υποτίμηση του μέσου επιπέδου των αποκλίσεων.

Ένας ακόμα δείκτης διασποράς ο οποίος χρησιμοποιεί την τυπική απόκλιση ως συστατικό, είναι ο συντελεστής μεταβλητότητας ο οποίος δίνεται από τον παρακάτω τύπο:

$$CV_x = \frac{\sigma_x}{\bar{x}} \quad (11)$$

Η προστιθέμενη αξία του συγκεκριμένου δείκτη προκύπτει από το γεγονός ότι εκφράζει την τυπική απόκλιση ως ποσοστό του μέσου όρου. Έτσι ουσιαστικά κανονικοποιεί την τυπική απόκλιση, δίνοντας ένα ποσοστιαίο μέγεθος και κάνοντας πιο σαφή την τάξη της μεταβλητότητας της χρονοσειράς. Επιπλέον επιτρέπει την σύγκριση της διασποράς-μεταβλητότητας ανάμεσα σε μεταβλητές οι οποίες έχουν διαφορετική τάξη μεγεθών. Η διαφορετική τάξη μεγεθών ανάμεσα σε χρονοσειρές ενδέχεται να παράξει τυπική απόκλιση μεγαλύτερη σε μία χρονοσειρά με χαμηλότερη μεταβλητότητα. Η κανονικοποίηση ως προς τον μέσο όρο ουσιαστικά απαλείφει τα

επίπεδα της μεταβλητής και επιστρέφει ένα ποσοστό το οποίο καθιστά την μεταβλητότητα των μεταβλητών συγκρίσιμη μεταξύ τους. (Ταμπάκης & Χαψά, 2013)

Ένα ακόμα σημαντικό εργαλείο διάγνωσης του επιπέδου διασποράς μίας μεταβλητής είναι το εύρος. Το εύρος είναι η διαφορά της παρατήρησης με την μεγαλύτερη τιμή από την παρατήρηση με την χαμηλότερη τιμή και ορίζει ακριβώς το πλαίσιο τιμών στο οποίο κυμαίνονται τα δεδομένα.

4.2.2 Συνάρτηση αυτοσυσχέτισης

Μελετώντας το πως μία χρονοσειρά κυμαίνεται στο χρόνο μπορεί να αντληθεί σημαντική πληροφορία εξετάζοντας εάν η μεταβλητή ακολουθεί κάποια εσωτερικά μοτίβα. Με την χρήση του συντελεστή αυτοσυσχέτισης λαμβάνεται ένδειξη για το κατά πόσο η τιμή μίας χρονοσειράς εξαρτάται από τις τιμές που είχε στις προηγούμενες χρονικές περιόδους. Ο συντελεστής αυτοσυσχέτισης έχει ως βασικό συστατικό την αυτοσυνδιακύμανση. Ουσιαστικά αποτελεί μία έκφραση του συντελεστή γραμμική συσχέτισης Pearson στο εσωτερικό της μεταβλητής, εξετάζει δηλαδή κατά πόσο η τιμές μίας μεταβλητής την χρονική στιγμή t σχετίζεται με την τιμή της μεταβλητής την χρονική στιγμή $t+k$. (Κάτος, 2004)

Ο τύπος της αυτοσυνδιακύμανσης δίνεται από την παρακάτω σχέση:

$$\gamma_k = \sigma_{i,i-k} = \frac{\sum_{i=1}^N (x_i - \bar{x})(x_{i-k} - \bar{x})}{N-1} \quad (12)$$

Ο τύπος του συντελεστή αυτοσυσχέτισης δίνεται από τον παρακάτω τύπο:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

$$\text{Όπου } \gamma_0 = \sigma^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}$$

$$\rho_k = \frac{\frac{\sum_{i=1}^N (x_i - \bar{x})(x_{i-k} - \bar{x})}{N-1}}{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}} \quad (13)$$

Ο συγκεκριμένος δείκτης χρησιμεύει σε τρία πράγματα. Αρχικά δίνει πληροφορία για την συμπεριφορά της μεταβλητής όσον αφορά την συσχέτιση της με καταστάσεις που βρισκόταν στο παρελθόν. Βάση αυτής της πληροφορίας μπορεί να εξαχθεί συμπερασματολογία για την στασιμότητα της χρονοσειράς, έννοια που παρουσιάζεται παρακάτω. Επιπλέον εάν υπάρχει πρόθεση για κατασκευή αυτοπαλίνδρομου υποδείγματος, ο συντελεστής αυτοσυσχέτισης υποδεικνύοντας με ποιες χρονικές υστερήσεις σχετίζεται η μεταβλητή, σημαίνει ποιες χρονικές υστερήσεις είναι δόκιμο να εισαχθούν στο υπόδειγμα.

Σύμφωνα με τον ασθενή ορισμό της στασιμότητας, μια χρονοσειρά χαρακτηρίζεται στάσιμη όταν έχει σταθερό μέσο όρο, σταθερή διακύμανση και η αυτοδιακύμανση εξαρτάται μόνο από την τάξη της υστέρησης και όχι από την χρονική περίοδο που μετράται (Κάτος, 2004). Μία χρονοσειρά δεν είναι στάσιμη όταν αυτή έχει μοναδιαία ρίζα. Ως μοναδιαία ρίζα νοείται η ύπαρξη κάποιας στοχαστικής τάσης η οποία αίρει την στασιμότητα της χρονοσειράς, ομοιάζοντας την με τυχαία περίπατο.

Το πρώτο πρόβλημα που δημιουργεί η απουσία στασιμότητας ή η ύπαρξη μοναδιαίας ρίζας είναι ότι καθιστά την μεταβλητή απρόβλεπτη. Το δεύτερο πρόβλημα που προκαλεί είναι σε περίπτωση εκτίμησης γραμμικής παλινδρόμησης, είτε δυμεταβλητής είτε κατά την εκτίμηση αυτοπαλίνδρομων υποδειγμάτων, ενδέχεται τα αποτελέσματα να είναι πλασματικά, δημιουργεί δηλαδή τον κίνδυνο της πλασματικής παλινδρόμησης (Mushtaq, 2011).

Συνεπώς προκειμένου να πραγματοποιηθεί εκτίμηση αξιόπιστων υποδειγμάτων, αφενός πρέπει να πραγματοποιηθεί διάγνωση στην χρονοσειρά αναφορικά με την ύπαρξη της μοναδιαίας ρίζας και σε περίπτωση που η χρονοσειρά χαρακτηριστεί ως μη στάσιμη πρέπει να μετασχηματιστεί έτσι ώστε να απαλλαγεί από την ύπαρξη μοναδιαίας ρίζας. Οι μετασχηματισμοί που ενδείκνυνται και είναι προϋπόθεση για τον υπολογισμό αξιόπιστων αυτοπαλίνδρομων υποδειγμάτων είναι δύο. Αρχικά υπάρχει η δυνατότητα η χρονοσειρά να λογαριθμιστεί. Επιπλέον υπάρχει η δυνατότητα οι εκτιμήσεις του υποδείγματος να μην γίνουν στα επίπεδα, αλλά στις πρώτες διαφορές των μεταβλητών (Κάτος,2004).

Αυτές είναι η δύο κλασσικές τεχνικές οι οποίες μετατρέπουν μία χρονοσειρά από μη στάσιμη σε στάσιμη. Ένα πρόβλημα που προκύπτει από τον υπολογισμό των πρώτων διαφορών είναι ότι θα χαθεί μία παραταρήση από την μεταβλητή και συνεπώς θυσιάζεται πληροφορία για την αποφυγή της πλασματικής παλινδρόμησης.

Προκειμένου όμως να γίνουν οι μετασχηματισμοί πρέπει πρώτα να γίνει η διάγνωση για το εάν υπάρχει η όχι μοναδιαία ρίζα. Μία σημαντική ένδειξη για την ύπαρξη στασιμότητας της χρονοσειράς δίνεται από το διάγραμμα των αυτοσυσχετίσεων. Πιο συγκεκριμένα, εάν όσο προχωράμε στο διάγραμμα προς τα δεξιά στις μακρινές χρονικές υστερήσεις η αυτοσυσχέτιση υποχωρεί αργά, αυτό είναι ένδειξη μη στασιμότητα. Αντίθετα, εάν η αυτοσυσχέτιση μειώνεται απότομα είναι ένδειξη απουσίας μοναδιαίας ρίζας (Κάτος,2004).

4.2.3 Έλεγχος Dickey-Fuller

Ένας πιο αξιόπιστος τρόπος εντοπισμού της μοναδιαίας ρίζας είναι ο έλεγχος που προτάθηκε από τους Dickey και Fuller (1979). Ο συγκεκριμένος έλεγχος βασίζεται σε ένα αυτοπαλίνδρομο υπόδειγμα και έχει ως μηδενική υπόθεση την ύπαρξη μοναδιαίας ρίζας. Ο έλεγχος βασίζεται στο παρακάτω υπόδειγμα:

$$\Delta\chi_i = \alpha + \beta_0\chi_{i-1} + \varepsilon_i \quad (14)$$

Η μηδενική υπόθεση είναι $H_0: \beta_0=0$, εάν ισχύει αυτή τότε η εξίσωση μετασχηματίζεται σε :

$$\chi_i = \alpha + \chi_{i-1} + \varepsilon_i \quad (15)$$

Έτσι η χρονοσειρά εκφράζεται μέσω ενός τυχαίου περιπάτου, δηλαδή εξαρτάται από έναν σταθερό όρο, το επίπεδο της στην προηγούμενη χρονική περίοδο και έναν λευκό θόρυβο. Αυτό καθιστά την χρονοσειρά μία απρόβλεπτη στοχαστική διαδικασία. Σε δεύτερο χρόνο ο έλεγχος έγινε επαυξημένος προσθέτοντας κάποιες χρονικές υστερήσεις των διαφορών της μεταβλητής στο υπόδειγμα, έτσι ώστε να γίνεται καλύτερη προσαρμογή και η ακρίβεια του υποδείγματος (και συνεπώς η αξιοπιστία του στατιστικού ελέγχου) να αυξηθεί (Mushtaq, 2011).

Πιο συγκεκριμένα το υπόδειγμα του επαυξημένου ελέγχου είναι της μορφής:

$$\Delta\chi_i = \alpha + \beta_0\chi_{i-1} + \sum_{k=1}^K \beta_{i-k}\Delta\chi_{i-k} + \varepsilon_i \quad (16)$$

Η μηδενική υπόθεση του ελέγχου είναι $H_0: \beta_0=0$, η οποία αντιστοιχεί στην ύπαρξη μοναδιαίας ρίζας. Κριτήριο απόρριψης της μηδενικής υπόθεσης είναι η στατιστική ελέγχου t να είναι μικρότερη της κρίσιμης τιμής, η οποία κρίσιμη τιμή δίνεται από την κατανομή των Dickey και Fuller.

Στο επεξεργαστικό περιβάλλον της R ο συγκεκριμένος έλεγχος μπορεί να πραγματοποιηθεί με την χρήση της βιβλιοθήκης 'urca' , όπου παρέχεται η συνάρτηση `ur.df()` η οποία διεξάγει τον συγκεκριμένο έλεγχο. Για την ερμηνεία του ελέγχου εξετάζεται το p -value του συντελεστή χρονικής υστέρησης της μεταβλητής (όχι των διαφορών της μεταβλητής). Ουσιαστικά το συγκεκριμένο p -value δείχνει την πιθανότητα να υπάρχει μοναδιαία ρίζα, δηλαδή την πιθανότητα να ισχύει $\beta_0=0$. Εάν είναι χαμηλότερο του επιπέδου σημαντικότητας που επιλέγει ο ερευνητής, η ύπαρξη στασιμότητας μπορεί να απορριφθεί.

4.2.4 Αυτοπαλίνδρομα υποδείγματα

Προκειμένου αφενός να ερμηνευθεί μία χρονοσειρά ως προς τις προηγούμενες παρατηρήσεις τις με ακριβή συναρτησιακή σχέση χρησιμοποιούνται τα αυτοπαλίνδρομα υποδείγματα $AR(Q)$.

Τα υποδείγματα είναι της μορφής:

$$x_i = \alpha_0 + \sum_{q=1}^Q \alpha_q x_{i-q} + \varepsilon_i \quad (17)$$

Το Q υποδεικνύει την τάξη του υποδείγματος, δηλαδή ποια είναι η μεγαλύτερη χρονική υστέρηση που συγκαταλέγεται στο υπόδειγμα. Όσο μεγαλύτερο είναι το Q , από τόσο μακρύτερα στο παρελθόν της χρονοσειράς υπάρχουν επιδράσεις, τόσο μεγαλύτερη είναι η μνήμη της χρονοσειράς.

Προκειμένου να εντοπιστεί η τάξη του υποδείγματος χρησιμοποιείται το διάγραμμα αυτοσυσχετίσεων. Παρατηρώντας το διάγραμμα βλέπουμε σε ποιες χρονικές υστερήσεις παρουσιάζεται σχετικά υψηλή αυτοσυσχέτιση και η τελευταία τάξη που έχει υψηλή αυτοσυσχέτιση συμπεριλαμβάνεται στο υπόδειγμα ως ύστατος όρος.

Τέλος, αναγκαίο είναι το διατυπωθεί πως ο όρος ε_i ή αλλιώς διαταρακτικός όρος αποτελεί το σφάλμα του υποδείγματος, δηλαδή την απόκλιση πραγματικής και εκτιμημένης τιμής ($\varepsilon_i = \chi_i - \hat{\chi}_i$) και είναι προϋπόθεση να ακολουθεί την κανονική κατανομή, να έχει μηδενικό μέσο (δηλαδή το υπόδειγμα να έχει καλή προσαρμογή) και σταθερή διακύμανση (ομοσκεδαστικά κατάλοιπα) ($\varepsilon_i \sim N(0, \sigma^2)$) (Κάτος, 2004) .

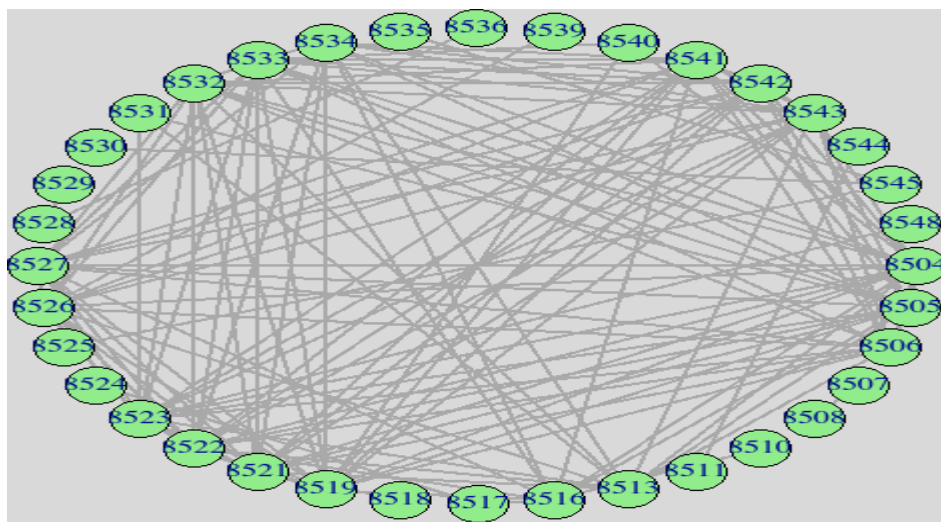
Η ανάλυση χρονοσειρών έχει πολλές συνιστώσες, ο υπολογισμός των οποίων καθώς και οι εξατομίκευση των εργαλείων στα χαρακτηριστικά και τις ανάγκες τις κάθε χρονοσειράς φέρει χρονικό και επεξεργαστικό κόστος. Κρίνεται δόκιμο ότι η εφαρμογή των παραπάνω εργαλείων να πραγματοποιείται στοχευμένα, προκειμένου να επιτευχθεί αφενός εξοικονόμηση χρόνου και πόρων, αφετέρου τα σχετικά συμπεράσματα από την ανάλυση να αποκτούν σημασία σε ένα ευρύτερο πλαίσιο του πεδίου ή της θεματικής που μελετάται.

Η θεωρία δικτύων δίνει την δυνατότητα αναγνώρισης των σημαντικότερων κόμβων-μεταβλητών στο πλαίσιο του περιβάλλοντος που συγκαταλέγονται. Μετά την αναγνώριση των σημαντικότερων κόμβων οι τεχνικές της ανάλυσης χρονοσειρών μπορούν να παρουσιαστούν εξατομικευμένα στις μεταβλητές που χαρακτηρίζονται ως καθοριστικότερες. Με αυτόν τον τρόπο η ενασχόληση του ερευνητή γίνεται πιο στοχευμένη.

5. Εμπειρικά Αποτελέσματα

Με την παραπάνω μεθοδολογία και για επίπεδο σημαντικότητας 1%, εκτιμήθηκε αρχικά ένα δίκτυο $g(48,137)$, δηλαδή ένα δίκτυο με 48 κόμβους και 137 ακμές. (Εικόνα 4).

Εικόνα 4: Γραφική αναπαράσταση δικτύου



Εντοπίστηκαν 12 χρονοσειρές οι οποίες δεν παρουσίαζαν καμία ακμή, δηλαδή ήταν γραμμικά άσχετες με τις λοιπές μεταβλητές και αυτές αφαιρέθηκαν με την χρήση των παρακάτω εντολών:

```
Isolated = which(degree(G)==0)
```

```
G2 = delete.vertices(G, Isolated)
```

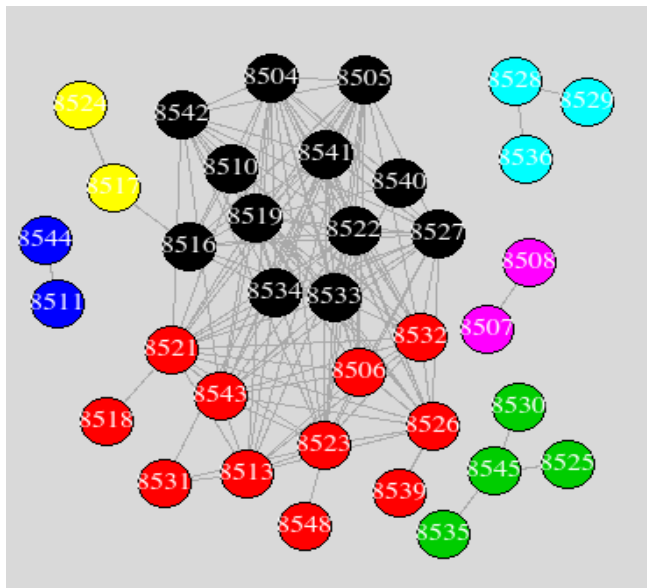
Οι αφαίρεση των ακμών έγινε έτσι ώστε να μην δημιουργείται θόρυβος στην ανάλυση από την συμπερίληψη κόμβων ασυσχέτιστων με το δίκτυο. Έτσι οι νέες διαστάσεις του δικτύου είναι $g(36,137)$.

Η πυκνότητα του δικτύου είναι ένας δείκτης ο οποίος παίρνει τιμές από 0 έως 1 και μας δείχνει πόσο πυκνό είναι το δίκτυο. Πιο συγκεκριμένα είναι το ποσοστό των ακμών που υπάρχουν σε σχέση με τον μέγιστο αριθμό ακμών που θα μπορούσαν να υπάρξουν αν όλοι κόμβοι συνδέονταν με όλους. Στο συγκεκριμένο δίκτυο η πυκνότητα είναι 27,1% κάτι που το καθιστά ένα σχετικά αραιό δίκτυο.

Μετά την κατασκευή του δικτύου η πρώτη ανάλυση που πραγματοποιείται είναι η ανάλυση κοινοτήτων.

Αλγόριθμος επιλογής κοινότητας	Modularity (προσαρμογή)
cluster_fast_greedy	0.1359
cluster_leading_eigen	0.1338
cluster_edge_betweenness	0.0955

Εικόνα 5: Γραφική αναπαράσταση δικτύου βάσει κοινοτήτων



Παρακάτω παρουσιάζονται οι τρεις βασικές κοινότητες που εντοπίστηκαν από τον αλγόριθμό, το χρώμα με το οποίο εμφανίζονται στο γράφημα, το μέγεθος τους και τα μέλη τους. Τα μέλη των κοινοτήτων δεν σχηματίζουν κάποια ομοιογενής κατηγορία προϊόντων.

<i>Αριθμός κοινότητας</i>	1
<i>Χρώμα</i>	Μαύρο
<i>Αριθμός μελών</i>	12
<i>Μέλη</i>	8504,8505,8510,8516,8519,8522,8527,8531,8534,8540,8541,8542

<i>Αριθμός κοινότητας</i>	2
<i>Χρώμα</i>	Κόκκινο
<i>Αριθμός μελών</i>	11
<i>Μέλη</i>	8506,8513,8518,8521,8523,8526,8531,8532,8539,8543,8548

<i>Αριθμός κοινότητας</i>	3
<i>Χρώμα</i>	Πράσινο
<i>Αριθμός μελών</i>	4
<i>Μέλη</i>	8525,8530,8535,8545

Όπως βλέπουμε στην εικόνα 6, η πρώτη κοινότητα περιέχει κωδικούς που αντιστοιχούν σε εξαρτήματα όπως ηλεκτρικοί μετασχηματιστές, ηλεκτρομαγνήτες, επαγωγείς, μηχανές για κούρεμα, συσκευές για ζέσταμα. Η δεύτερη μπαταρίες, φορητές λάμπες, μικροφωνικός εξοπλισμός κτλ. Η Τρίτη ανθρακόνημα, ηλεκτρονικό εξοπλισμό για προστασία ηλεκτρονικών συστημάτων.

Εικόνα 6. Μέλη των τριών κοινοτήτων (<https://atlas.cid.harvard.edu/rankings/product>)

ΜΕΛΗ ΚΟΙΝΟΤΗΤΑΣ 1	
8504	Electrical transformers, static converters (for example, rectifiers) and inductors; parts thereof
8505	Electromagnets; permanent magnets and articles intended to become permanent magnets after magnetization; electromagnetic or permanent magnet chucks, clamps and similar holding devices; electromagnetic couplings, clutches and brakes; electromagnetic lifting
8510	Shavers, hair clippers and hair-removing appliances, with self-contained electric motor; parts thereof
8516	Electric instantaneous or storage water heaters and immersion heaters; electric space heating apparatus and soil heating apparatus; electrothermic hairdressing apparatus (for example, hair dryers, hair curlers, curling tong heaters) and hand dryers; electric
8519	Sound recording or reproducing apparatus
8522	Parts and accessories suitable for use solely or principally with the apparatus of headings 8519 to 8521
8527	Reception apparatus for radio broadcasting, whether or not combined, in the same housing, with sound recording or reproducing apparatus or a clock
8531	Electric sound or visual signaling apparatus (for example, bells, sirens, indicator panels, burglar or fire alarms), other than those of heading 8512 or 8530; parts thereof
8534	Electronic printed circuits
8540	Thermionic, cold cathode or photocathode tubes (for example, vacuum or vapour or gas filled tubes, mercury arc rectifying tubes, cathode-ray tubes, television camera tubes); parts thereof
8541	Diodes, transistors and similar semiconductor devices; photosensitive semiconductor devices, including photovoltaic cells whether or not assembled in modules or made up into panels; light-emitting diodes; mounted piezoelectric crystals; parts thereof
8542	Electronic integrated circuits; parts thereof
ΜΕΛΗ ΚΟΙΝΟΤΗΤΑΣ 2	
8506	Primary cells and primary batteries; parts thereof
8513	Portable electric lamps designed to function by their own source of energy (for example, dry batteries, storage batteries, magnetos), other than lighting equipment of heading 8512; parts thereof
8518	Microphones and stands therefor; loudspeakers, whether or not mounted in their enclosures; headphones and earphones, whether or not combined with a microphone, and sets consisting of a microphone and one or more loudspeakers; audio frequency electric amplifiers
8521	Video recording or reproducing apparatus, whether or not incorporating a video tuner
8523	Discs, tapes, solid-state non-volatile storage devices, smart cards and other media for the recording of sound or of other phenomena, whether or not recorded, including matrices and masters for the production of discs, but excluding products of Chapter 37
8526	Radar apparatus, radio navigational aid apparatus and radio remote control apparatus
8531	Electric sound or visual signaling apparatus (for example, bells, sirens, indicator panels, burglar or fire alarms), other than those of heading 8512 or 8530; parts thereof
8532	Electrical capacitors, fixed, variable or adjustable (pre-set); parts thereof
8539	Electrical filament or discharge lamps, including sealed beam lamp units and ultraviolet or infrared lamps; arc lamps; parts thereof
8543	Electrical machines and apparatus, having individual functions, not specified or included elsewhere in this chapter; parts thereof
8548	Waste and scrap of primary cells, primary batteries and electric storage batteries; spent primary cells, spent primary batteries and spent electric storage batteries; electrical parts of machinery or apparatus, not specified or included elsewhere in this chapter
ΜΕΛΗ ΚΟΙΝΟΤΗΤΑΣ 3	
8525	Transmission apparatus for radio-broadcasting or television, whether or not incorporating reception apparatus or sound recording or reproducing apparatus; television cameras, digital cameras and video camera recorders
8530	Electrical signaling, safety or traffic control equipment for railways, streetcar lines, subways, roads, inland waterways, parking facilities, port installations or airfields (other than those of heading 8608); parts thereof
8535	Electrical apparatus for switching or protecting electrical circuits, or for making connections to or in electrical circuits (for example, switches, fuses, lightning arresters, voltage limiters, surge suppressors, plugs and other connectors, junction boxes)
8545	Carbon electrodes, carbon brushes, lamp carbons, battery carbons and other articles of graphite or other carbon, with or without metal, of a kind used for electrical purposes

Κρίνεται αναγκαίο να σχολιαστεί ότι το χαμηλό modularity (0.13/1) υποδεικνύει και χαμηλή αξιοπιστία όσον όχι την ύπαρξη της κοινότητας, αλλά της σημασίας της.

Επιπλέον οι 4 από τις 7 κοινότητες που εντοπίστηκαν αποδίδονται στο γεγονός ότι υπάρχουν μερικοί κόμβοι που συνδέονται μεταξύ τους αλλά όχι με την κύρια συνιστώσα του δικτύου. Αυτές οι κοινότητες λοιπόν μπορούν να εξηγηθούν και να μελετηθούν με δυμεταβλητούς όρους, χωρίς να αναχθούν στο δικτυακό πλαίσιο του ολισμού, στο οποίο κυμαίνεται η σημασία της κοινότητας.

Η πληροφορία που μπορεί να αντληθεί από την συγκεκριμένη ανάλυση έχει να κάνει με τις 3 κοινότητες και με την αναλυτικότερη χαρτογράφηση του δικτύου. Πιο συγκεκριμένα φαίνεται οι κόμβοι του δικτύου να έχουν ήπια τάση προς την δημιουργία κοινότητας (0.13/1) και οι κοινότητες που παρουσιάζονται (κυρίως η κοινότητες 1,2 και 3) φαίνεται να αποτελούν περιοχές των οποίων τα προϊόντα παρουσιάζουν έντονη συσχέτιση ως προς την τεχνογνωσία τους.

Αξιολόγηση κόμβων βάση κεντρικότητας

Η σημαντικότερη εκροή της παρούσας ανάλυσης είναι η διάκριση των κόμβων ως προς την κεντρικότητα τους. Με αυτόν τον τρόπο θα εντοπιστούν οι πιο σημαντικοί κόμβοι του δικτύου, η σημαντικότητα των οποίων θα αξιολογηθεί βάση τεσσάρων δεικτών κεντρικότητα, δηλαδή βάση της βαθμικής κεντρικότητας, της ιδιοκεντρικότητας, της διαμεσότητας και της εγγύτητας.

Στους παρακάτω πίνακες (Εικόνα 6 έως Εικόνα 9) παρουσιάζονται τα αποτελέσματα της κάθε κεντρικότητας διατεταγμένη σε φθίνουσα σειρά και στις αριστερές στήλες (Rank και Vertex) φαίνεται σε κάθε περίπτωση ποιοι κόμβοι εντοπίζονται ως περισσότερο κεντρικοί. Αναφορικά με την ερμηνεία, από την παρούσα ανάλυση θα αναδειχθούν οι ομάδες προϊόντων που είναι εντονότερα σχετιζόμενες και πιο κεντρικά διατυπωμένες στον χάρτη αλληλεξάρτησης τεχνογνωσίας.

Εικόνα 7: Διάταξη των κόμβων ως προς την βαθμική κεντρικότητα

Rank	Vertex	Degree
1	8504	0.485714
2	8505	0.457143
3	8519	0.457143
4	8521	0.457143
5	8522	0.457143
6	8523	0.457143
7	8533	0.457143
8	8527	0.428571
9	8541	0.428571
10	8543	0.428571
11	8506	0.4
12	8526	0.4
13	8532	0.4
14	8542	0.4
15	8534	0.342857
16	8513	0.314286
17	8516	0.285714
18	8531	0.085714
19	8540	0.085714
20	8545	0.085714
21	8517	0.057143
22	8528	0.057143
23	8507	0.028571
24	8508	0.028571
25	8510	0.028571
26	8511	0.028571
27	8518	0.028571
28	8524	0.028571
29	8525	0.028571
30	8529	0.028571
31	8530	0.028571
32	8535	0.028571
33	8536	0.028571
34	8539	0.028571
35	8544	0.028571
36	8548	0.028571

Εικόνα 8 : Διάταξη των κόμβων ως προς την ιδιοκεντρικότητα

Rank	Vertex	Eigencentrality
1	8504	1
2	8522	0.986604457
3	8519	0.957091314
4	8533	0.957091314
5	8521	0.956257046
6	8543	0.951861445
7	8527	0.943695772
8	8541	0.943695772
9	8505	0.942813216
10	8523	0.918311548
11	8542	0.90895276
12	8506	0.902525945
13	8532	0.902525945
14	8526	0.844616986
15	8534	0.75276126
16	8513	0.654700893
17	8516	0.530109014
18	8540	0.204389243
19	8531	0.176450061
20	8518	0.067068087
21	8548	0.06440674
22	8539	0.059238095
23	8517	0.037363543
24	8510	0.037179749
25	8524	0.002620531
26	8545	5.13141E-17
27	8530	3.84856E-17
28	8511	2.56571E-17
29	8528	2.56571E-17
30	8536	2.56571E-17
31	8544	2.56571E-17
32	8507	1.28285E-17
33	8525	1.28285E-17
34	8529	1.28285E-17
35	8535	1.28285E-17
36	8508	0

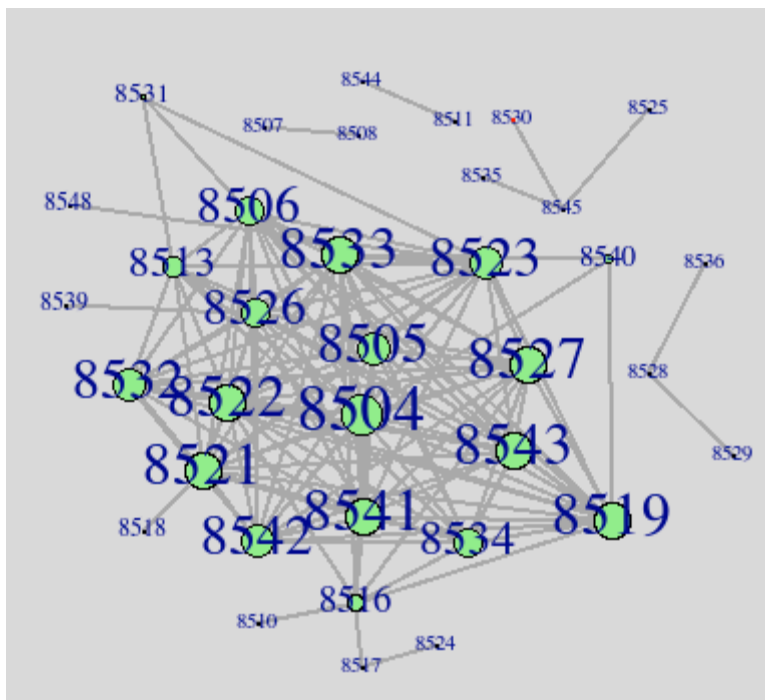
Εικόνα 9: Διάταξη των κόμβων ως προς την διαμεσότητα

Rank	Vertex	Betweness
1	8516	0.1092437
2	8523	0.0502786
3	8521	0.0407017
4	8526	0.0394958
5	8517	0.0386555
6	8505	0.0267416
7	8504	0.0262159
8	8519	0.0217741
9	8533	0.0217741
10	8522	0.0131706
11	8527	0.0106496
12	8541	0.0106496
13	8513	0.0078885
14	8534	0.0058824
15	8545	0.005042
16	8543	0.0020462
17	8528	0.0016807
18	8542	0.0012059
19	8506	0.0010989
20	8532	0.0010989
21	8507	0
22	8508	0
23	8510	0
24	8511	0
25	8518	0
26	8524	0
27	8525	0
28	8529	0
29	8530	0
30	8531	0
31	8535	0
32	8536	0
33	8539	0
34	8540	0
35	8544	0
36	8548	0

Εικόνα 10: Διάταξη των κόμβων ως προς την εγγύτητα

Rank	Vertex	Closeness
1	8505	0.083999
2	8522	0.083842
3	8541	0.083826
4	8519	0.083822
5	8533	0.083785
6	8504	0.083766
7	8526	0.083731
8	8523	0.083705
9	8506	0.083645
10	8527	0.083584
11	8542	0.083546
12	8513	0.083456
13	8516	0.083435
14	8534	0.083427
15	8521	0.083347
16	8532	0.083316
17	8543	0.083315
18	8531	0.082247
19	8540	0.081751
20	8548	0.081388
21	8517	0.081186
22	8510	0.081108
23	8518	0.081078
24	8539	0.080783
25	8524	0.078287
26	8545	0.030338
27	8535	0.03031
28	8525	0.030309
29	8530	0.030306
30	8528	0.029434
31	8536	0.029421
32	8529	0.02942
33	8511	0.028582
34	8544	0.028582
35	8507	0.028582
36	8508	0.028582

Εικόνα 11: Γραφική αναπαράσταση των κόμβων ως προς την ιδιοκεντρικότητα τους.



Αρχικά παρατηρούμε ότι όσον αφορά την βαθμική κεντρικότητα υπάρχει ισοψηφία επτά κόμβων οι οποίοι φαίνεται να έχουν το ίδιο σύνολο ακμών. Η συγκεκριμένη ισοψηφία αναδεικνύει την ανάγκη εισαγωγής περισσότερων κριτηρίων- δεικτών κεντρικότητας.

Παρατηρούμε ότι από την ανάλυση ξεχωρίζουν ως ισχυρότεροι οι κόμβοι 8504 και 8505. Πιο συγκεκριμένα ο κόμβος 8504(Ηλεκτρικοί μετασχηματιστές, στατικοί μετατροπείς, επαγωγείς) λαμβάνει πρώτη θέση στην βαθμική κεντρικότητα και στην ιδιοκεντρικότητα, ενώ ο κόμβος 8505 (ηλεκτρομαγνήτες και διάφορων ειδών μαγνήτες) λαμβάνει πρώτη θέση στην βαθμική κεντρικότητα και την εγγύτητα.

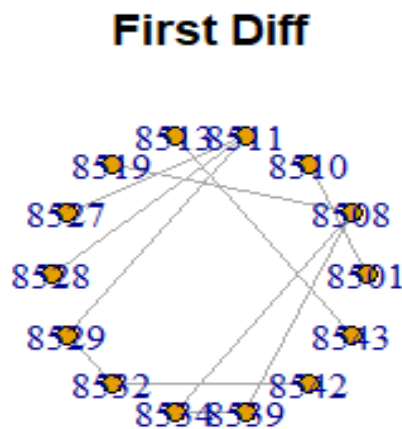
Όσον αφορά την σημασία της ιδιοκεντρικότητας φαίνεται να ξεχωρίζει από τις άλλες κεντρικότητες. Ο λόγος είναι διότι αναδεικνύει τους κόμβους που έχουν πολλούς αλλά και ισχυρούς γείτονες.

Συνεπώς, εξάγεται το συμπέρασμα ότι η τεχνογνωσία και συνεπώς η πολυπλοκότητα που απαιτείται για την παραγωγή προϊόντων της ομάδας 8504 είναι πιο καθοριστική, η πιο κεντρική και συνεπώς η σημαντικότερη στην σύνθεση και την εξέλιξη της τεχνογνωσίας για τα υπόλοιπα προϊόντα του κλάδου.

Επίσης, ενδιαφέρον παρουσιάζει και η μελέτη του δικτύου που δημιουργείται, αφαιρώντας τις γραμμικά άσχετες με τις λοιπές μεταβλητές, άμα πάρουμε τις πρώτες διαφορές των χρονοσειρών, διαστάσεων $G(14,11)$ (Εικόνα 12)

Παρατηρούμε ότι δημιουργείτε δηλαδή ένα δίκτυο με 14 κόμβους και 11 ακμές. Όπως παρατηρούμε είναι ένα μικρότερο δίκτυο. Η δομή των κοινοτήτων θα αλλάξει, συνεπώς και η ανάλυση κοινοτήτων και η αξιολόγηση κόμβων βάση κεντρικότητας.

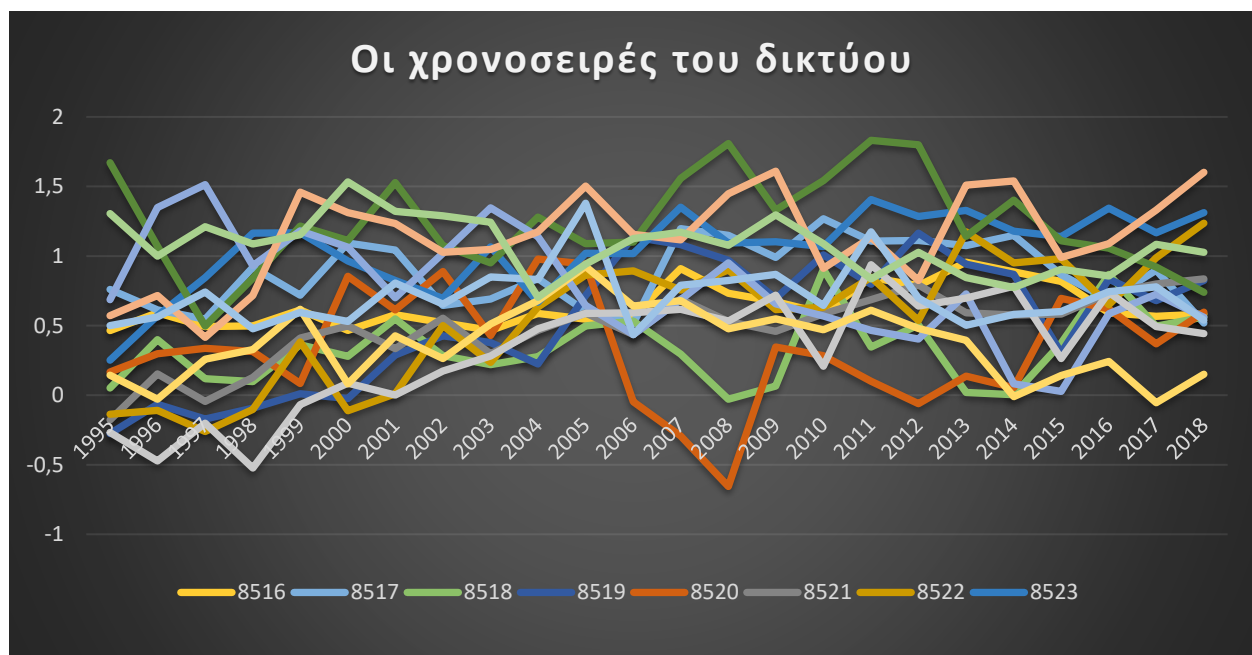
Εικόνα 12: Γραφική αναπαράσταση δικτύου (Πρώτες διαφορές)



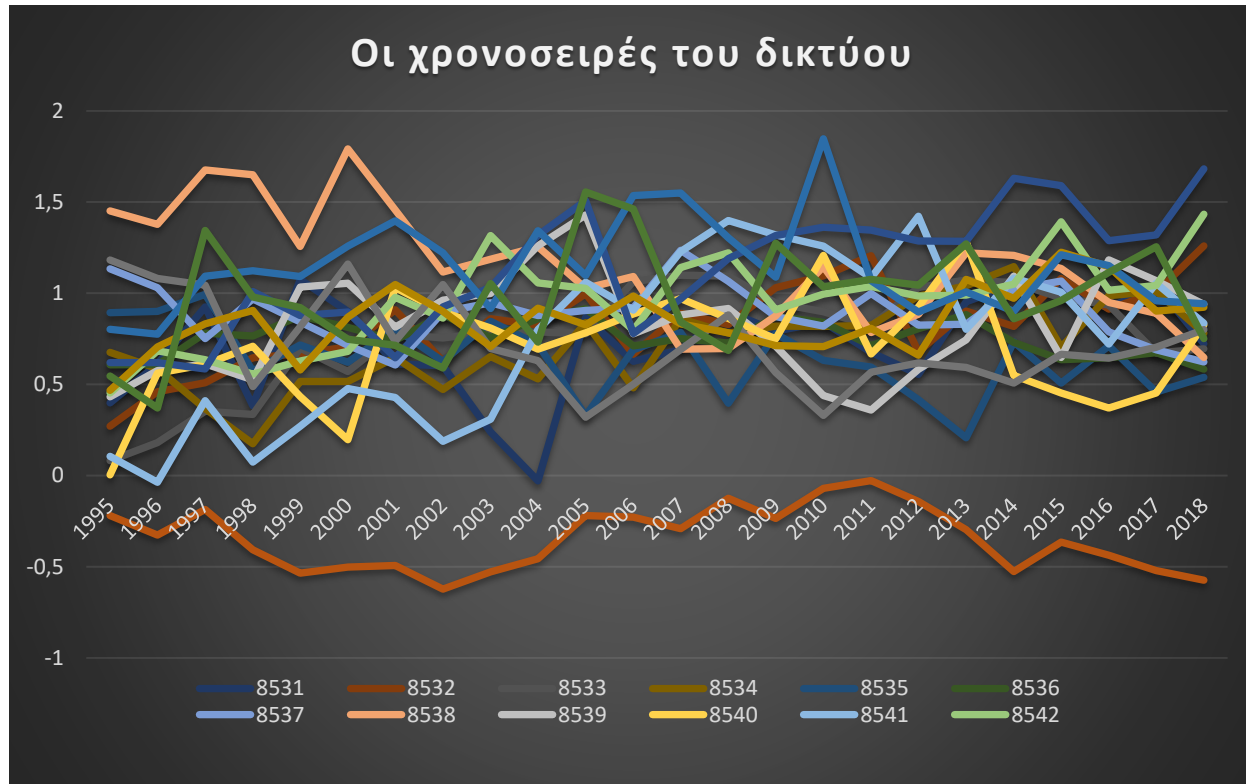
Εικόνα 13: Γραφική αναπαράσταση των χρονοσειρών του δικτύου HS8501-HS8515



Εικόνα 14: Γραφική αναπαράσταση των χρονοσειρών του δικτύου HS8516-HS8530

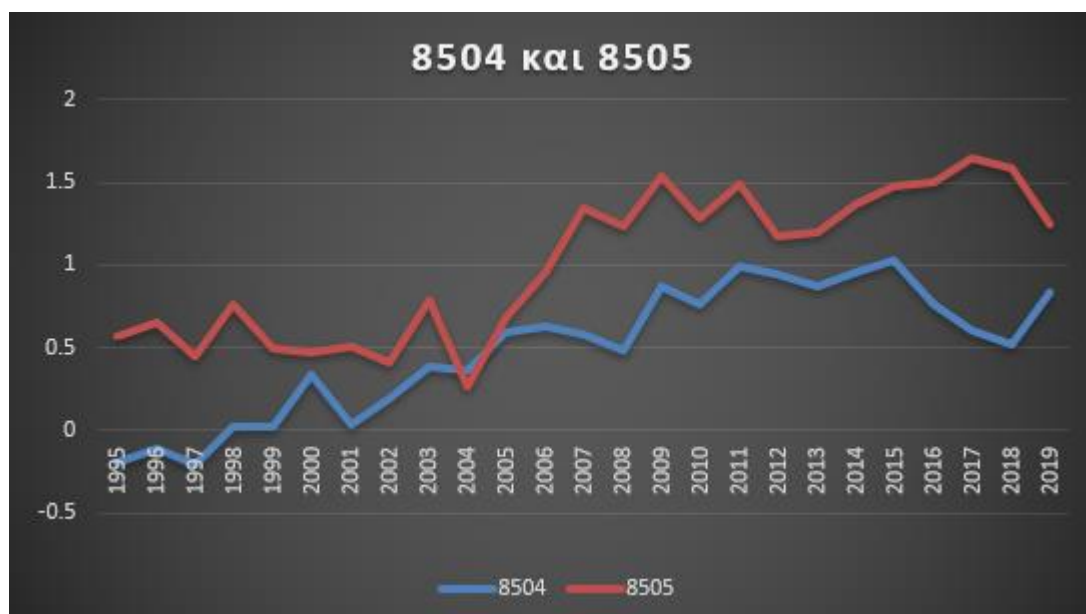


Εικόνα 15: Γραφική αναπαράσταση των χρονοσειρών του δικτύου HS8531-HS8548



Έπειτα από την ανάλυση του δικτύου και την ανάδειξη των δύο σημαντικότερων κόμβων πραγματοποιείται ανάλυση χρονοσειρών στους κωδικούς οι οποίοι κρίθηκαν σημαντικότεροι. Παρακάτω παρουσιάζονται τα αποτελέσματα περιγραφικής στατιστικής, ανάλυσης στασιμότητας και εκπόνηση υποδείγματος για της χρονοσειρές 8504 και 8505. Η μεσολάβηση της θεωρίας δικτύων επιτρέπει την στοχευμένη εφαρμογή ανάλυσης χρονοσειρών παρακάμπτοντας την εφαρμογή της ανάλυσης στο σύνολο των κόμβων του δικτύου.

Εικόνα 16: Η οπτική παρουσίαση των χρονοσειρών 8504 και 8505 για την χρονική περίοδο 1995-2019.



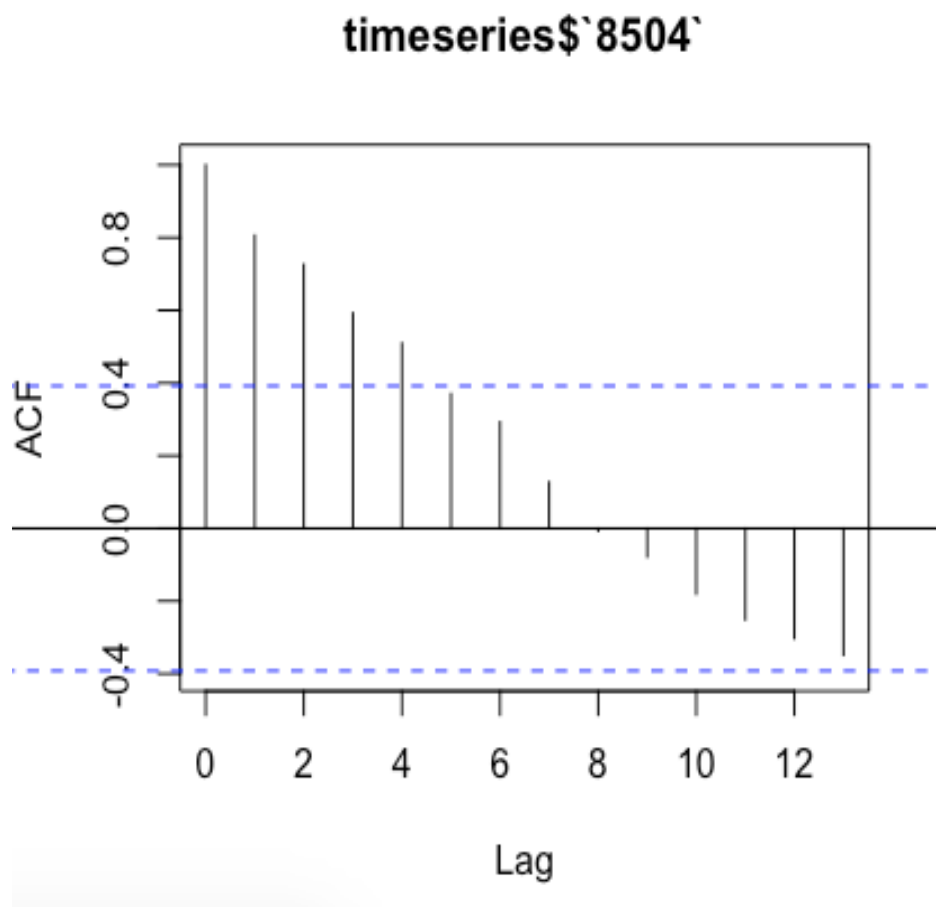
Αρχικά παρουσιάζονται τα μέτρα κεντρικής τάσης, ο μέσος όρος και η διάμεσος. Παρατηρούμε ότι τα επίπεδα πολυπλοκότητας είναι αυξημένα στην μεταβλητή 8405 σε σχέση με την 8404. Επίσης, η απόκλιση μέσου όρου και διαμέσου είναι ένδειξη ότι οι χρονοσειρές δεν ακολουθούν την κανονική κατανομή και φαίνεται πως η κεντρική τάση τους είναι δεξιά από τα επίπεδα του μέσου όρου, καθώς και στις δύο περιπτώσεις η διάμεσος είναι υψηλότερη.

Αναφορικά με την διασπορά των μεταβλητών, η μεταβλητή 8405 φαίνεται πως εκτείνεται σε μεγαλύτερο εύρος τιμών. Μολονότι η μεταβλητή 8405 παρουσιάζει μεγαλύτερη τυπική απόκλιση, κατόπιν κανονικοποίησης με τον μέσο όρο και υπολογισμού του συντελεστή μεταβλητότητας φαίνεται πως η μεταβλητότητα-διασπορά στην μεταβλητής 8404 είναι υψηλότερη.

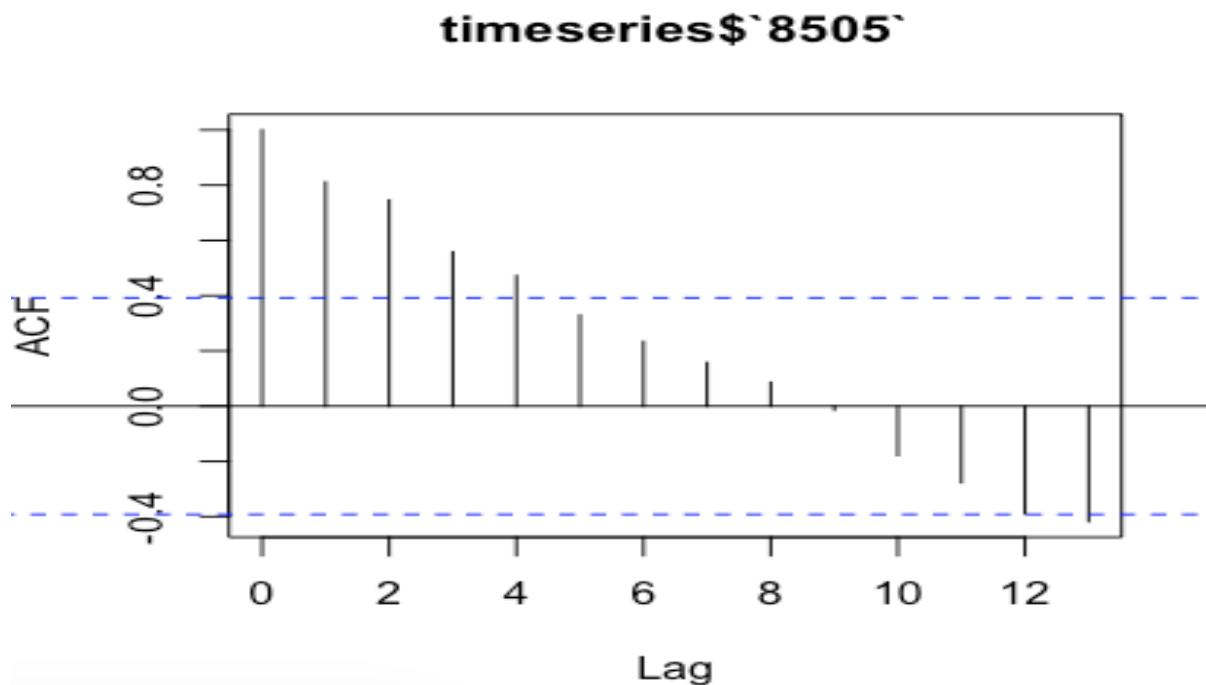
	Εύρος	Τυπική Απόκλιση	Συντελεστής Μεταβλητότητας	Μέσος Όρος	Διάμεσος
8404	1.23	0.39	0.79	0.49	0.58
8405	1.38	0.45	0.45	1	1.17

Εν συνεχεία υπολογίστηκε ο συντελεστής αυτοσυσχέτισης για τις δύο μεταβλητές με την χρήση της συνάρτησης `acf()` στο προγραμματιστικό περιβάλλον R. Αναφορικά με την στασιμότητα λαμβάνεται θετική πρώτη ένδειξη καθώς φαίνεται η αυτοσυσχέτιση να σβήνει σχετικά αργά με αποτέλεσμα να αναμένεται η ύπαρξη μοναδιαίας ρίζας.

Εικόνα 17: Το διάγραμμα αυτοσυσχετίσεων της μεταβλητής 8504



Εικόνα 18: Το διάγραμμα αυτοσυσχετίσεων της μεταβλητής 8505



Συμπληρωματικά πραγματοποιήθηκε ο αυξημένο έλεγχος στασιμότητας Dickey-Fuller με την χρήση της βιβλιοθήκης 'urca' και την εντολή `ur.df()`.

Στην περίπτωση της πρώτης χρονοσειράς 8504 η τιμή p-value του όρους χρονικής υστέρησης της μεταβλητής στα επίπεδα ήταν 0.1139. Αυτό ισοδυναμεί με 11% πιθανότητα η χρονοσειρά να εμπεριέχει μοναδιαία ρίζα. Συνεπώς ούτε για $\alpha=10\%$ δεν μπορεί να απορριφθεί η ύπαρξη μοναδιαίας ρίζας στην χρονοσειρά και η χρονοσειρά χαρακτηρίζεται ως μη στάσιμη.

Αναφορικά με την δεύτερη χρονοσειρά 8505 ο αντίστοιχος έλεγχος για την στατιστική σημαντικότητα του συντελεστή του όρου χρονικής υστέρησης της μεταβλητής στο υπόδειγμα Dickey Fuller δίνει $P\text{-value}=0.3817$. Συνεπώς δεν μπορεί να απορριφθεί η μηδενική υπόθεση $\beta=0$, δηλαδή απουσία μοναδιαίας ρίζας και έτσι χαρακτηρίζεται και αυτή η χρονοσειρά ως μη στάσιμη.

Εν συνεχεία οι δύο χρονοσειρές ανάχθηκαν στις πρώτες διαφορές του προκειμένου να εξαλείψουμε το φαινόμενο της στασιμότητας και να αποφευχθεί το πρόβλημα της πλασματικής παλινδρόμησης στο αυτοπαλίνδρομο υπόδειγμα.

Εικόνα 19: Διαγραμματική αναπαράσταση των μεταβλητών σε πρώτες διαφορές

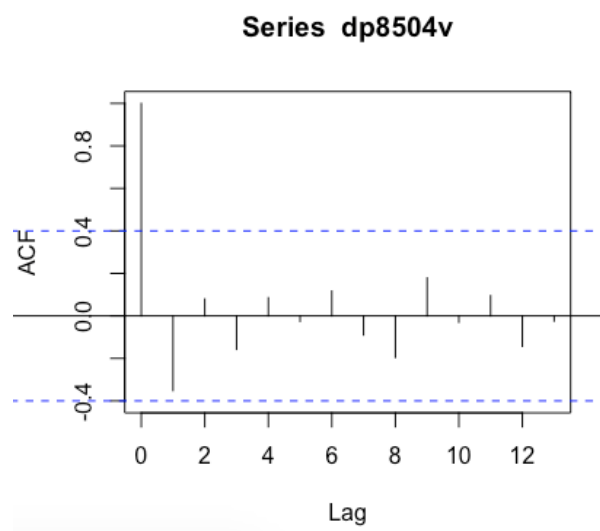


Η μετατροπή των πρώτων διαφορών έχει την παρακάτω μαθηματική έκφραση :

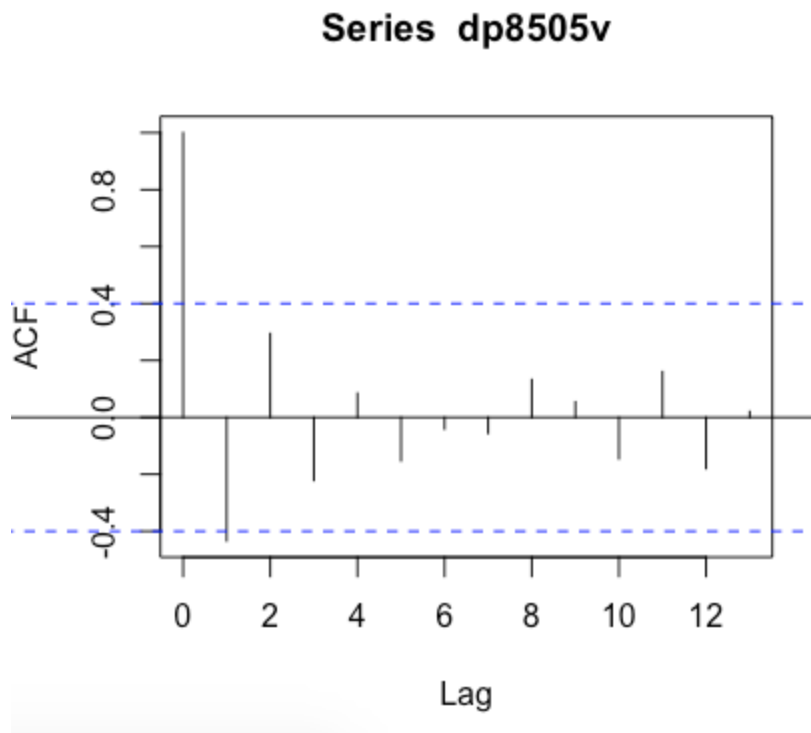
$$\chi_i \rightarrow \Delta\chi_i, \text{ όπου } \Delta\chi_i = \chi_i - \chi_{i-1}$$

Πραγματοποιήθηκε με την χρήση της εντολής `diff()` που παρέχει το υπολογιστικό περιβάλλον της R. Εν συνεχεία υπολογίζονται εκ νέου οι συναρτήσεις αυτοσυσχετίσεων των δύο μεταβλητών.

Εικόνα 20: Το διάγραμμα αυτοσυσχετίσεων των πρώτων διαφορών της μεταβλητής 8504



Εικόνα 21: Το διάγραμμα αυτοσυσχετίσεων των πρώτων διαφορών της μεταβλητής 8505



Οι συναρτήσεις αυτοσυσχετίσεων φαίνεται να μειώνονται απότομα, σημαντική ένδειξη πως οι χρονοσειρές είναι στάσιμες στις πρώτες διαφορές τους.

Εν συνεχεία πραγματοποιείται επαυξημένος έλεγχος Dickey-Fuller αρχικά στην μεταβλητή d8504. Το p-value του συντελεστή του όρου χρονικής υστέρησης είναι 0.0015 , συνεπώς με επίπεδο σημαντικότητας 1% μπορούμε να απορρίψουμε την μηδενική υπόθεση περί ύπαρξης μοναδιαίας ρίζας και να συμπεράνουμε ότι η χρονοσειρά είναι στάσιμη στις πρώτες διαφορές της.

Ο αντίστοιχος έλεγχος στην χρονοσειρά d8505 δίνει P-value=0.0061 , συνεπώς και εδώ για επίπεδο σημαντικότητας 1% μπορεί να απορριφθεί η μηδενική υπόθεση και να συμπεράνουμε απουσία μοναδιαίας ρίζας και άρα ύπαρξη στασιμότητας.

Εφόσον οι χρονοσειρές είναι στάσιμες δύναται να πραγματοποιηθεί η θεμελίωση του υποδείγματος. Το υπόδειγμα που εκτιμάται είναι υπόδειγμα τύπου AR. Προκειμένου να εντοπιστεί η τάξη του υποδείγματος χρησιμοποιήθηκαν τα διάγραμμα αυτοσυσχετίσεων των πρώτων διαφορών των μεταβλητών. Αυτά υπέδειξαν την πρώτη τάξη καθώς και για τις δύο χρονοσειρές

φαίνεται πως εκεί παρουσιάζεται η υψηλότερη αυτοσυσχέτιση. Συνεπώς ακολουθεί η εκτίμηση ενός υποδείγματος τύπου AR(1) στις πρώτες διαφορές των μεταβλητών.

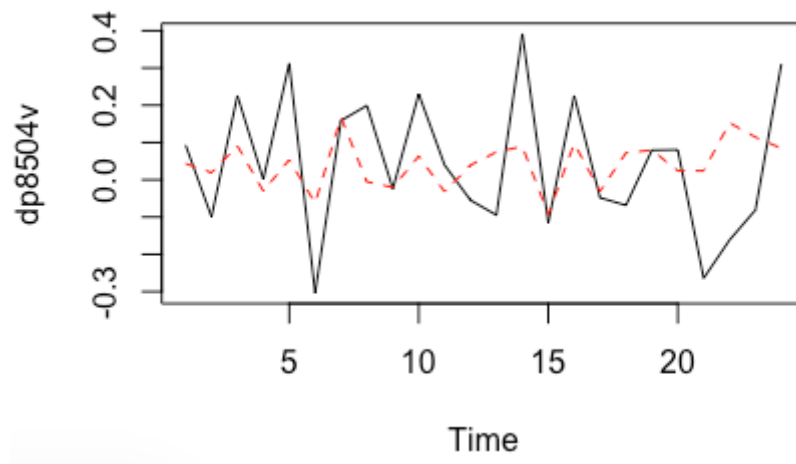
Το υπόδειγμα που εκτιμήθηκε με την χρήση της συνάρτησης `arima()`, για την εκτίμηση χρησιμοποιήθηκαν οι πρώτες διαφορές των μεταβλητών και ως τάξη του υποδείγματος ορίστηκε η $c(1,0,0)$.

Για κάθε υπόδειγμα παρουσιάζεται αρχικά η συναρτησιακή έκφραση του υποδείγματος και ένα διάγραμμα της μεταβλητής μαζί με τα κατάλοιπα του υποδείγματος. Από τα διαγράμματα και των δύο χρονοσειρών φαίνεται πως ο μέσος όρος έχει σταθερή διακύμανση και κυμαίνεται γύρω από το μηδέν.

Τα αποτελέσματα της εκτίμησης του υποδείγματος για την χρονοσειρά 8504 είναι τα παρακάτω.

$$\widehat{\Delta\chi}_i = 0.039 - 0.3736 * \Delta\chi_{i-1} + \varepsilon_i \quad (18)$$

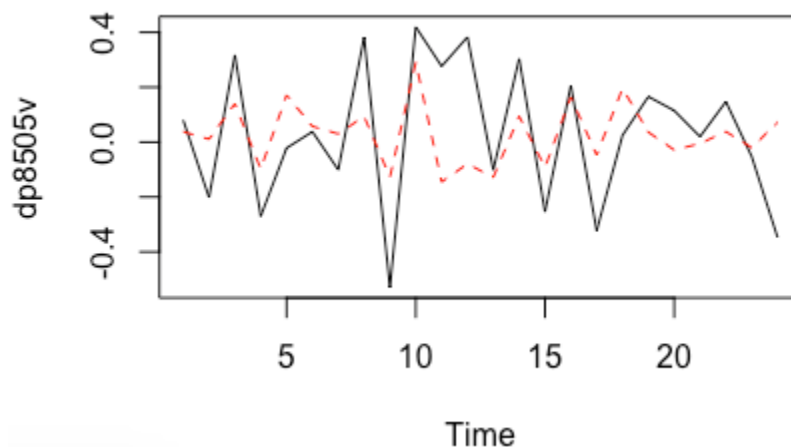
Εικόνα 22: Η προσαρμογή του υποδείγματος για την χρονοσειρά 8504



Τα αποτελέσματα της εκτίμησης του υποδείγματος για την χρονοσειρά 8504 είναι τα παρακάτω.

$$\widehat{\Delta\chi}_i = 0.0325 - 0.4588 * \Delta\chi_{i-1} + \varepsilon_i \quad (19)$$

Εικόνα 23: Η προσαρμογή του υποδείγματος για την χρονοσειρά 8505



6. Συζήτηση-Σχολιασμός

Εικόνα 24: Η περιγραφή των προϊόντων που συγκαταλέγονται στον εκάστοτε δείκτη σύμφωνα με τον Άτλα του Χάρβαρντ

8504	Electrical transformers, static converters (for example, rectifiers) and inductors; parts thereof
8505	Electromagnets; permanent magnets and articles intended to become permanent magnets after magnetization;electromagnetic or permanent magnet chucks, clamps and similar holding devices;electromagnetic couplings, clutches and brakes; electromagnetic lifting

Η ανάλυση των Κωδικών NS και των σχέσεων τους με τον Δείκτη Πολυπλοκότητας Προϊόντος (PCI) αποκάλυψε ότι ο Κωδικός Ns8504 έχει την υψηλότερη κεντρικότητα από όλους, κάτι που προκύπτει κυρίως από την ερμηνεία του δείκτη της ιδιοκεντρικότητας. Αυτός ο κωδικός περιλαμβάνει μετασχηματιστές, ανορθωτές και επαγωγείς που χρησιμοποιούνται σε μονάδες τροφοδοσίας (PSU) όλων των ηλεκτρονικών συσκευών. Το γεγονός αυτό αντιστοιχεί στον υψηλό βαθμό κεντρικότητας και δείχνει την εξάρτηση της πλειοψηφίας των άλλων κατηγοριών ηλεκτρονικών προϊόντων από αυτή την κατηγορία.

Η κεντρικότητα είναι ένα μέτρο του πόσο σημαντικός είναι ένας κόμβος σε ένα γράφημα, όπου οι κόμβοι αντιπροσωπεύουν οντότητες και οι ακμές αντιπροσωπεύουν σχέσεις μεταξύ τους. Σε αυτήν την περίπτωση, η ανάλυσή μας υποδεικνύει πόσο σημαντικός είναι ο κωδικός Ns8504 σε σχέση με άλλους κωδικούς όσον αφορά τη βαθμολογία PCI. Όσο υψηλότερη είναι η κεντρική θέση, τόσο πιο σημαντικό είναι να συνδέονται άλλοι κωδικοί σε αυτήν.

Οι μετασχηματιστές χρησιμοποιούνται ως εξαρτήματα σε PSU για τη μετατροπή της τάσης AC από το ένα επίπεδο στο άλλο, είτε αυξάνοντάς το είτε μειώνοντάς το. Οι ανορθωτές χρησιμοποιούνται επίσης στα PSU καθώς μετατρέπουν την τάση εναλλασσόμενου ρεύματος σε τάση συνεχούς ρεύματος επιτρέποντας μόνο μία κατεύθυνση ροής ρεύματος μέσω αυτών. Οι επαγωγείς χρησιμοποιούνται επίσης στα σύγχρονα παλμοτροφοδοτικά και φορτιστές καθώς βοηθούν στη ρύθμιση της ροής του ρεύματος αποθηκεύοντας ενέργεια όταν το ρεύμα περνά μέσα από αυτά και βοηθώντας στην μετατροπή του σε υψηλές συχνότητες κατά την χρήση των οποίων χρησιμοποιούνται μικρότεροι μετασχηματιστές. Αυτή η ιδιότητα τους κάνει απαραίτητους για την σμίκρυνση των συσκευών όπως Πχ. στα κινητά τηλέφωνα, οι φορτιστές, οι λάμπες led κλπ.

Ο συσχετισμός μεταξύ της χρήσης αυτής της κατηγορίας σε όλα τα προϊόντα με τον υψηλό βαθμό κεντρικότητάς της δείχνει ότι είναι απαραίτητο για τις περισσότερες ηλεκτρονικές συσκευές να περιλαμβάνονται αυτά τα εξαρτήματα στις μονάδες τροφοδοσίας τους για σωστή λειτουργία. Αυτό υποδηλώνει ότι ο κωδικός 8504 θα πρέπει να έχει προτεραιότητα κατά το σχεδιασμό μονάδων τροφοδοσίας για οποιαδήποτε ηλεκτρονική συσκευή λόγω της σημασίας και της συνάφειάς του.

Συμπερασματικά, η ανάλυσή μας αποκάλυψε ότι ο κωδικός 8504 έχει υψηλό βαθμό κεντρικότητας σε σύγκριση με άλλους κωδικούς και τα προϊόντα του αποτελούν βασικά στοιχεία για τις μονάδες τροφοδοσίας όλων των ηλεκτρονικών συσκευών λόγω της σημασίας και της συνάφειάς τους. Επομένως, οι σχεδιαστές θα πρέπει να δίνουν προτεραιότητα σε αυτόν τον κώδικα όταν σχεδιάζουν μονάδες τροφοδοσίας για οποιαδήποτε ηλεκτρονική συσκευή λόγω της σημασίας και της συνάφειάς του. Παράλληλα καταδεικνύεται η εξάρτηση του συνόλου τις αγορές των ηλεκτρονικών συσκευών από την συγκεκριμένη κατηγορία και η σημαντικότητα της.

7. Συμπεράσματα προτάσεις

Η ανάλυσή μας για τους Δείκτες Πολυπλοκότητας Προϊόντος (PCI) της βιομηχανίας ηλεκτρονικών είχε ως αποτέλεσμα την υψηλή κεντρικότητα της κατηγορίας με κωδικό HS8504. Αυτή η κατηγορία αποτελείται από βασικά υλικά των τροφοδοτικών, τα οποία τροφοδοτικά είναι απαραίτητα εξαρτήματα για τα περισσότερα ηλεκτρονικά προϊόντα. Τα αποτελέσματα δείχνουν ότι αυτή η κατηγορία δεν είναι μόνο πολύ πολύπλοκη αλλά και σε μεγάλο βαθμό οι άλλες κατηγορίες του κλάδου εξαρτώνται από αυτήν.

Οι δείκτες PCI μπορούν να χρησιμοποιηθούν για τον εντοπισμό κρίσιμων διαδρομών στο σχεδιασμό και την ανάπτυξη προϊόντων, καθώς σε συνδυασμό με την ανάλυση δικτύων μπορούν να παρέχουν επιπρόσθετη και καλύτερη κατανόηση όχι μόνο της πολυπλοκότητας αλλά και των εξαρτήσεων του προϊόντος. Αναλύοντας τους δείκτες PCI, μπορούμε να προσδιορίσουμε ποιες κατηγορίες είναι πιο σημαντικές για την εύρυθμη λειτουργία ενός κλάδου και ποιες κατηγορίες πρέπει να τύχουν μεγαλύτερης προσοχής κατά τον σχεδιασμό και την ανάπτυξη προϊόντων.

Επιπλέον, αναλύοντας τους δείκτες PCI, μπορούμε να αποκτήσουμε πληροφορίες για το πώς αλληλοεπιδρούν διαφορετικές κατηγορίες μεταξύ τους και πώς οι αλλαγές σε μια κατηγορία μπορεί να επηρεάσουν άλλες κατηγορίες. Αυτό μπορεί να μας βοηθήσει να κατανοήσουμε καλύτερα πώς οι αλλαγές σε έναν τομέα μπορεί να επηρεάσουν άλλους τομείς κατά τον σχεδιασμό και την ανάπτυξη προϊόντων.

Συνολικά, η ανάλυση μας για τους Δείκτες Πολυπλοκότητας Προϊόντων αποκάλυψε ότι τα τροφοδοτικά είναι απαραίτητα εξαρτήματα για πολλά ηλεκτρονικά προϊόντα και ότι είναι πολύ περίπλοκα και εξαρτώνται από αυτά οι άλλες κατηγορίες του κλάδου. Αναλύοντας αυτούς τους δείκτες και το δίκτυο τους, με την χρήση θεωρίας γραφών μπορούμε να επεκτείνουμε την πληροφόρηση που δίνουν οι δείκτες πολυπλοκότητας και να αποκτήσουμε γνώσεις και για τις εξαρτήσεις του προϊόντος εκτός από την πολυπλοκότητα, βοηθώντας μας να εντοπίσουμε κρίσιμες διαδρομές στο σχεδιασμό και την ανάπτυξη προϊόντων.

Βασική αδυναμία της συγκεκριμένης προσέγγισης όσον αφορά την κατασκευή του δικτύου διασυσχετίσεων των δεικτών πολυπλοκότητας είναι η αδυναμία του συντελεστή συσχέτισης να εντοπίσει και να εκφράσει μη γραμμικές σχέσεις. Αυτό ενδέχεται να προκαλεί σημαντική απώλεια πληροφορίας καθώς το είδος της συσχέτισης μεταξύ των μεταβλητών δύναται να είναι εκθετικό, λογαριθμικό ή ημιτονοειδές. Σε περίπτωση ύπαρξης σχέσης αυτής της μορφής ανάμεσα στις μεταβλητές ο υπό μελέτη συντελεστής συσχέτισης δεν θα εντοπίσει την αμοιβαία συμπεριφορά των μεταβλητών με αποτέλεσμα να υποτιμήσει την ένταση της σχέσης ή χειρότερα να δώσει απουσία συσχέτισης ενώ στην πραγματικότητα οι μεταβλητές σχετίζονται. Προκειμένου να γίνει περαιτέρω εμβάθυνση ως προς την σύσταση του δικτύου, εύλογο θα ήταν για την κατασκευή του να χρησιμοποιηθεί κάποιος πιο εξελιγμένος δείκτης με δυνατότητα εντοπισμού και μη γραμμικών σχέσεων.

Η υπόσταση του δικτύου θα μπορούσε επίσης να επεκταθεί μετατρέποντας το από μη κατευθυνόμενο σε κατευθυνόμενο. Με αυτόν τον τρόπο η πληροφορία που παρέχεται ως προς τις διασυσχετίσεις των μεταβλητών θα επεκταθεί. Ένα κατευθυνόμενο δίκτυο αποτελεί δίκτυο αιτιότητας και όχι απλώς συσχέτισης. Συνεπώς όλη η τοπολογία του δικτύου μετασχηματίζεται. Τα μονοπάτια φεροιπίν παύουν να είναι απλώς μονοπάτια συσχετίσεων και μετατρέπονται σε μονοπάτια επιδράσεων, καταγράφοντας αλυσιδωτές επιδράσεις. Επίσης η μελέτη της βαθμικής κεντρικότητας ως προς την σημαντικότητα του κάθε κόμβου μπορεί να γίνει πιο εξειδικευμένη και να διακριθεί σε “degree in” και “degree out” με το έσω βαθμό να εκφράζει τον αριθμό των κόμβων που ασκούν άμεσα επίδραση τον κόμβο αναφοράς και με τον βαθμό έξω να εκφράζει τον αριθμό των κόμβων που ασκεί άμεσα επίδραση ο κόμβος αναφοράς. Έτσι εάν φεροιπίν αναζητείται ο κόμβος που ασκεί την σημαντικότερη επίδραση μπορούν οι σχετικοί δείκτες να γίνουν πιο εξειδικευμένοι επιλέγοντας τον έξω βαθμό του κάθε κόμβου.

8. Βιβλιογραφία

- Asuero, A. G., Sayago, A., & González, A. G. (2006). The correlation coefficient: An overview. *Critical reviews in analytical chemistry*, 36(1), 41-59.
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427-431.
- Ferreira, L. N., & Zhao, L. (2016). Time series clustering via community detection in networks. *Information Sciences*, 326, 227-242.
- Friedkin, N. E. (1991). Theoretical foundations for centrality measures. *American journal of Sociology*, 96(6), 1478-1504.
- Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the national academy of sciences*, 106(26), 10570-10575.
- Mealy, P., Farmer, J. D., & Teytelboym, A. (2019). Interpreting economic complexity. *Science advances*, 5(1), eaau1705.
- Mushtaq, R. (2011). Augmented dickey fuller test.
- Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, 6(1), 35-39.
- West, D. B. (2001). *Introduction to graph theory* (Vol. 2). Upper Saddle River: Prentice hall.
- Ταμπάκης Α. Ν. , Χαψά, Α.Ν. (2013). Εφαρμοσμένη Στατιστική, εργαστηριακές ασκήσεις. Θεσσαλονίκη. Ζυγός
- Κάτος, Α. (2004). Οικονομετρία, θεωρία και εφαρμογές. Θεσσαλονίκη, Ζυγός.