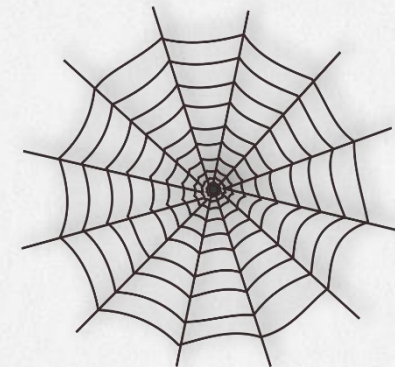




Pajek

Εφαρμογή ανάλυσης μεγάλων δικτύων

Έκδοση 4.03

<http://mrvar.fdv.uni-lj.si/pajek>





Η εφαρμογή Pajek

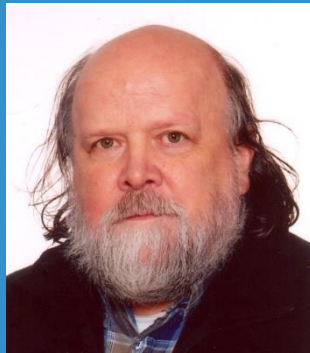
Χρησιμοποιείται για την ανάλυση και απεικόνιση μεγάλων δικτύων.



Andrej Mrvar

Καθηγητής Κοινωνικής Πληροφορικής

Πανεπιστήμιο
της Λιουμπλιάνα
Σλοβενία



Vladimir Batagelj

Μαθηματικός

Πανεπιστήμιο
της Λιουμπλιάνα
Σλοβενία

Γενικές πληροφορίες για το Pajek

- Η ανάπτυξή του ξεκίνησε το 1996
- Έχει γραφτεί σε Delphi
- Μόνο για το λειτουργικό σύστημα Windows
- Διατίθεται σε 32 και 64bit εκδόσεις
- Είναι δωρεάν, αλλά όχι ανοικτό λογισμικό
- Διαθέσιμο για μη εμπορική χρήση
- Η λέξη “Pajek” στα σλοβενικά σημαίνει “αράχνη”

Μια ειδική έκδοση του Pajek, είναι το Pajek-XXL. Έχει σημαντικά λιγότερες απαιτήσεις σε μνήμη (2 έως 3 φορές) και είναι ταχύτερο σε περιπτώσεις διεργασιών με συχνή χρήση μνήμης (π.χ. παραγωγή τυχαίων δικτύων, εξαγωγή δεδομένων δικτύων, συρρίκνωση δικτύων κ.α.).

Χρησιμοποιείται ευρέως στην επιστημονική έρευνα



1ο μέρος



Γενικά χαρακτηριστικά

Δίκτυα που υποστηρίζει το Pajek και δυνατότητες υπολογισμών στα δίκτυα αυτά.



Δίκτυα που αναλύονται ή/και απεικονίζονται με το Pajek:

- ✓ Δίκτυα οργανικών μορίων
- ✓ Αλληλεπιδράσεις δικτύων πρωτεϊνών - υποδοχέων
- ✓ Κοινωνικά δίκτυα
- ✓ Γενεαλογικά δέντρα
- ✓ Δίκτυα Internet
- ✓ Δίκτυα βιβλιογραφικών παραπομπών
- ✓ Δίκτυα διάχυσης (ιών, ειδήσεων κ.α.)
- ✓ Και πολλά άλλα

Με το Pajek είναι δυνατή:

- ✓ Η εύρεση συστάδων (υποσύνολα, γειτονικές κορυφές που παρουσιάζουν κάποιο νόημα, πυρήνες, κλπ.) σε ένα δίκτυο.
- ✓ Η εξαγωγή κορυφών που ανήκουν σε ίδιες συστάδες και η παρουσίασή τους ξεχωριστά (λεπτομερής τοπική θέαση).
- ✓ Η συμπύκνωση κορυφών σε συστάδες.
- ✓ Η παρουσίαση των σχέσεων που υπάρχουν μεταξύ συστάδων (ευρεία σφαιρική θέαση).

Τύποι δικτύων που υποστηρίζονται

- ✓ Πολυσχεσιακά δίκτυα
- ✓ Δίκτυα 2-mode (διμερείς γράφοι)
- ✓ Δυναμικά δίκτυα

Τι είναι δίκτυο

Ένα δίκτυο αποτελείται από ένα γράφο και ορισμένες συμπληρωματικές πληροφορίες σχετικά με τις κορυφές ή/και τις γραμμές του γραφήματος.



Οι τύποι δεδομένων που υποστηρίζει το Pajek

Το Pajek υποστηρίζει έξι διαφορετικούς τύπους δεδομένων (δίκτυα, κατατμήσεις, διανύσματα, μεταθέσεις, συστάδες και ιεραρχίες).

.net

Δίκτυα (Networks)

Στο Pajek τα δίκτυα είναι τα κύρια αντικείμενα. Ένα δίκτυο αποτελείται από ένα γράφο και ορισμένες συμπληρωματικές πληροφορίες σχετικά με τις κορυφές ή/και τις γραμμές του γραφήματος αυτού.

.clu

Κατατμήσεις (Partitions)

Ένα διαμέρισμα ενός δικτύου είναι μια ταξινόμηση ή ομαδοποίηση των κορυφών του δικτύου, έτσι ώστε κάθε κορυφή έχει εκχωρηθεί σε ακριβώς μία κλάση (ή ένα σύμπλεγμα).

.vec

Διανύσματα (Vectors)

Ένα διάνυσμα περιλαμβάνει και εκχωρεί μια αριθμητική τιμή για κάθε κορυφή ενός δικτύου.

.per

Μεταθέσεις (Permutations)

Μια μετάθεση ενός δικτύου είναι μια νέα αρίθμηση (δηλαδή μια αναδιάταξη) των κορυφών του.

.cls

Συστάδες (Clusters)

Μια συστάδα είναι ένα υποσύνολο των κορυφών ενός δικτύου.

.hie

Ιεραρχίες (Hierarchies)

Μια ιεραρχία περιλαμβάνει ιεραρχικά ταξινομημένες κορυφές ενός δικτύου.

Διαφορά Partition από Hierarchy

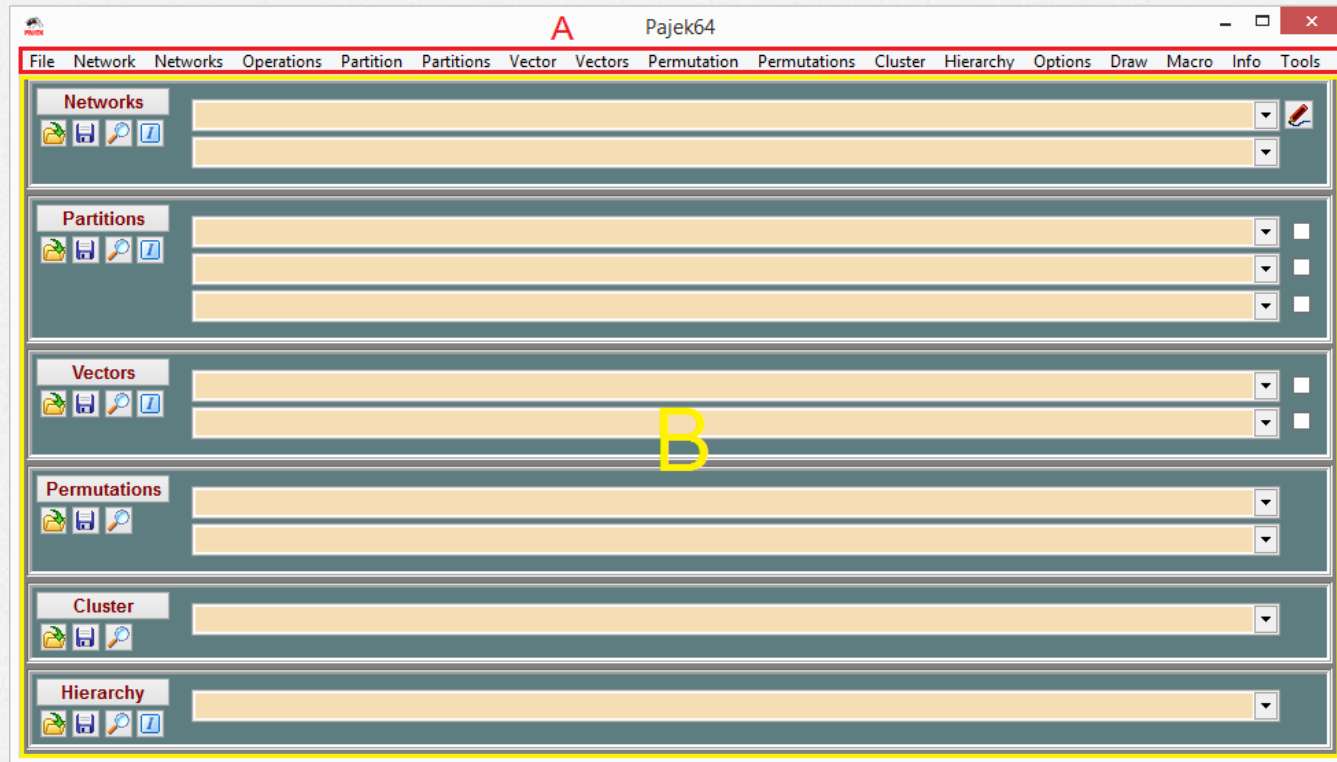
Σε ένα Partition, μια κορυφή (ή group κορυφών) δεν μπορεί να ανήκει ταυτόχρονα σε περισσότερες από μία κλάσεις όπως συμβαίνει σε μια ιεραρχία.



Το γραφικό περιβάλλον εργασίας του Pajek



Το κεντρικό παράθυρο της εφαρμογής αποτελείται από δύο περιοχές. Την περιοχή του κυρίως μενού (Κόκκινη – A) και την περιοχή των δεδομένων (Κίτρινη B).



Μοιάζει με ένα επιστημονικό κομπιουτεράκι για μαθηματικές πράξεις σε δίκτυα

Τα στοιχεία ελέγχου της περιοχής B

Είναι το Βασικό περιβάλλον εργασίας.

6 μεγάλα κουμπιά: Προσθήκη – αφαίρεση πεδίων εισαγωγής, σε όσα είναι στον πληθυντικό.

Drop-down lists:

- Πολλαπλά αρχεία δεδομένων
- Πράξεις μεταξύ αρχείων
- Απεικόνιση επιλεγμένου αρχείου

Εικονίδια:

- Άνοιγμα – εισαγωγή αρχείου (φάκελος)
- Αποθήκευση αλλαγών (δισκέτα)
- Επεξεργασία ιδιοτήτων (φακός)
- Λήψη πληροφοριών με κριτήριο (info)
- Απεικόνιση δικτύου (μολύβι)

Checkboxes: Μορφοποίηση της εμφάνισης των αντικειμένων κατά την απεικόνισή τους, όταν υπάρχουν πολλά όμοια αντικείμενα.



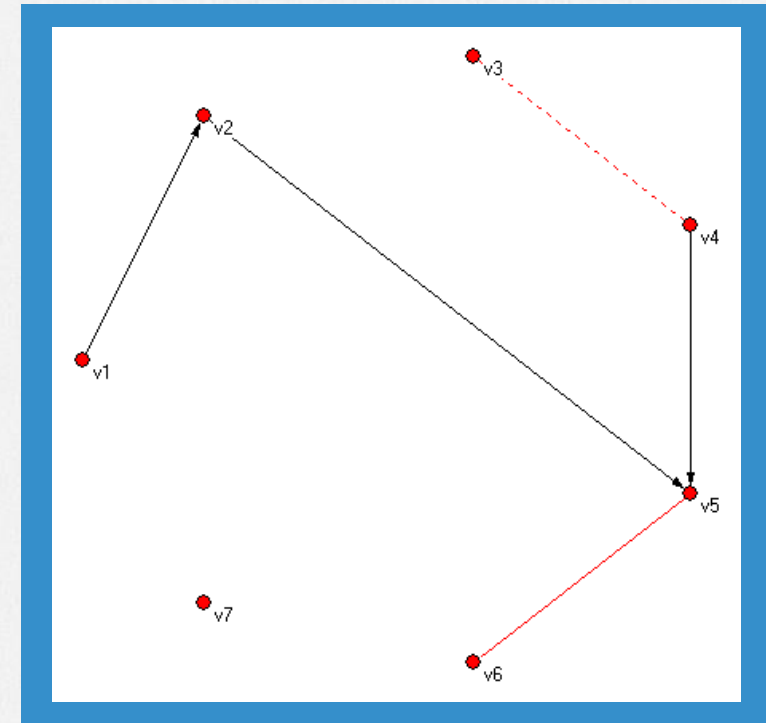
Δομή αρχείου .net



Σε ένα αρχείο .net περιλαμβάνονται όλες οι πληροφορίες σχεδίασης και απεικόνισης των πρωτογενών δεδομένων του δικτύου.

```
*Vertices 7
1 "v1"          0.1000  0.5000  0.5000
2 "v2"          0.2506  0.1873  0.5000
3 "v3"          0.5890  0.1100  0.5000
4 "v4"          0.8604  0.3264  0.5000
5 "v5"          0.8604  0.6736  0.5000
6 "v6"          0.5890  0.8900  0.5000
7 "v7"          0.2506  0.8127  0.5000
*Arcs
1 2 1
2 5 2.500000000
4 5 1
*Edges
3 4 -1
5 6 1
```

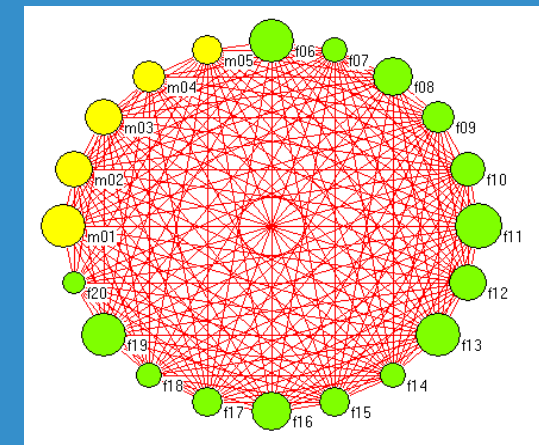
Αρχείο κατευθυνόμενου δικτύου 7 κορυφών, 3 τόξων και δύο γραμμών.



Παράδειγμα: Μαϊμού Wolfe (2-mode network)

Τα δεδομένα αφορούν τις αλληλεπιδράσεις, το φύλο, την ηλικία και το ranking των ατόμων αυτού του είδους μαϊμούς.

inter.net	inter.net	sex.clu	age.vec	rank.per
*Vertices 20		*vertices 20	*vertices 20	*vertices 20
1 "m01"	1 6 5	1	15	1
2 "m02"	1 7 9	1	10	2
3 "m03"	1 8 7	1	10	3
4 "m04"	1 9 4	1	8	4
5 "m05"	1 10 3	1	7	5
6 "f06"	1 11 3	2	15	10
7 "f07"	1 12 7	2	5	11
8 "f08"	1 13 3	2	11	6
9 "f09"	1 14 2	2	8	12
10 "f10"	1 15 5	2	9	9
11 "f11"	1 16 1	2	16	7
12 "f12"	1 17 4	2	10	8
13 "f13"	1 18 1	2	14	18
14 "f14"	2 3 5	2	5	19
15 "f15"	2 4 1	2	7	20
16 "f16"	2 5 3	2	11	13
17 "f17"	2 6 1	2	7	14
18 "f18"	2 7 4	2	5	15
19 "f19"	2 8 2	2	15	16
20 "f20"	2 9 6	2	4	17
*Edges				
1 2 2	2 10 2			
1 3 10	2 11 5			
1 4 4	2 12 4			
- - -	2 13 3			
	2 14 2			
	...			



Εντολή απεικόνισης δικτύου: Draw > Network + First Partition + First Vector



Δημιουργία αρχείων δικτύου για το Pajek



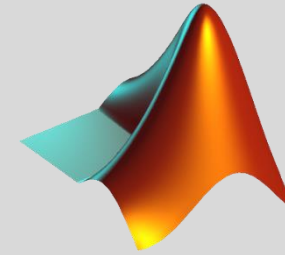
Οι έξι βασικότεροι τρόποι δημιουργίας αρχείων .net. Υπάρχουν και άλλοι τρόποι, όπως για παράδειγμα, μέσω άλλων γλωσσών προγραμματισμού.



Pajek



Επεξεργαστής Κειμένου



Matlab



Python



Excel2Pajek.exe



Σχεσιακή Βάση Δεδομένων



Δημιουργία αρχείου δικτύου με το Pajek

Έχει νόημα όταν εισάγουμε πρωτογενή δεδομένα τα οποία έχουμε συλλέξει οι ίδιοι και θα πρέπει να εισαχθούν χειροκίνητα.

1

Network > Create Random Network > Total No. of Arcs

2

Erdos-Renyi Random Network

Number of Vertices: 28

Number of Arcs: 0

OK Cancel

3

Draw > Network

4

Εμφανίζεται ασύνδετος γράφος (αφού Arcs=0)

5

File > Network > View/Edit
Για προσθήκη συνδέσεων

6

Edit lines incident to vertex

Select vertex number or vertex label: v7

OK Cancel

7

Editing Network: 1. Erdos-Renyi random network [0] (28). Vertex:7

1:	23.7	val=1.00000
1:	7.15	val=1.00000
1:	7-19	val=1.00000

Newline

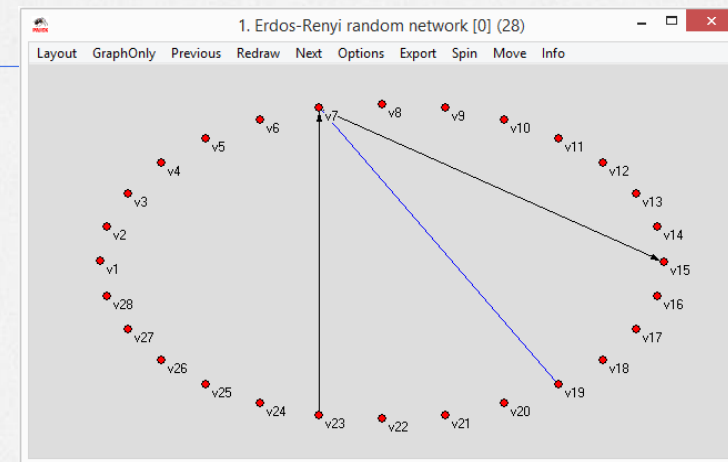
8

Add line In/Out of vertex:7

Into vertex (+), Out of vertex (-), Edge (nosign): +23

OK Cancel

9





Δημιουργία αρχείου δικτύου με κειμενογράφο



Έχει νόημα για πρωτογενή δεδομένα, τα οποία έχουμε συλλέξει οι ίδιοι και πρέπει στη συνέχεια να τα εισάγουμε χειροκίνητα.

```
*Vertices 7
1 "v1"          0.1000    0.5000    0.5000
2 "v2"          0.2506    0.1873    0.5000
3 "v3"          0.5890    0.1100    0.5000
4 "v4"          0.8604    0.3264    0.5000
5 "v5"          0.8604    0.6736    0.5000
6 "v6"          0.5890    0.8900    0.5000
7 "v7"          0.2506    0.8127    0.5000
*Arcs
1 2 1
2 5 2.5000000000
4 5 1
*Edges
3 4 -1
5 6 1
```

Αρχείο κατευθυνόμενου δικτύου 7 κορυφών, 3 τόξων και δύο γραμμών.

Μορφοποίηση του στυλ εμφάνισης

Το Pajek, επιτρέπει επιπλέον μορφοποίηση και καθορισμό των χαρακτηριστικών των στοιχείων που ανήκουν στο δίκτυο (συντεταγμένες των κορυφών στο χώρο, σχήμα και χρώμα των κορυφών και των γραμμών). Για παράδειγμα, η λέξη "box" δίπλα από μια κορυφή, την κάνει να εμφανίζεται με τετραγωνάκι αντί κύκλου.

Σύνταξη αρχείου .net σε κειμενογράφο

Η πρώτη γραμμή (*Vertices 7) ακολουθείται από επτά γραμμές οι οποίες περιλαμβάνουν τις επτά κορυφές του δικτύου. Καθεμιά από αυτές τις γραμμές περιλαμβάνει αντίστοιχα, τον αριθμό της κορυφής (1 έως 7), το σύμβολό της (v1 έως v7), και τρεις τιμές που αντιπροσωπεύουν τις συντεταγμένες της στο χώρο (x, y, z). Δεν επιτρέπονται κενές γραμμές και tabs.

Κάτω από την γραμμή με την επικεφαλίδα *Arcs, υπάρχουν τρεις γραμμές που αφορούν τα τόξα κατευθυνόμενων συνδέσεων. Ο πρώτος αριθμός αφορά την κορυφή από την οποία ξεκινά το βέλος και ο δεύτερος, την κορυφή στην οποία δείχνει. Ο τρίτος επηρεάζει το στυλ της γραμμής. Για παράδειγμα αν έχει αρνητικό πρόσημο σχεδιάζεται διακεκομμένη.

Τέλος, οι μη-κατευθυνόμενες συνδέσεις περιγράφονται κάτω από τον τίτλο *Edges.

Συναρτήσεις Matlab

- `get_random_graph()` – Επιστρέφει ένα μη-κατευθυνόμενο, τυχαίο δίκτυο 20 κόμβων, με μέσο όρο 3 συνδέσεις ανά κόμβο (πιθανότητα 3/20 ή 15%), σε μορφή .mat.
- `write_matrix_to_pajek()` – Μετατρέπει δεδομένα Matlab που βρίσκονται σε μορφή .mat, σε αρχείο δικτύου Pajek μορφής .net.

Σενάριο: Νευρωνικές συνδέσεις μεταξύ διαφορετικών εγκεφαλικών περιοχών.

Βήματα

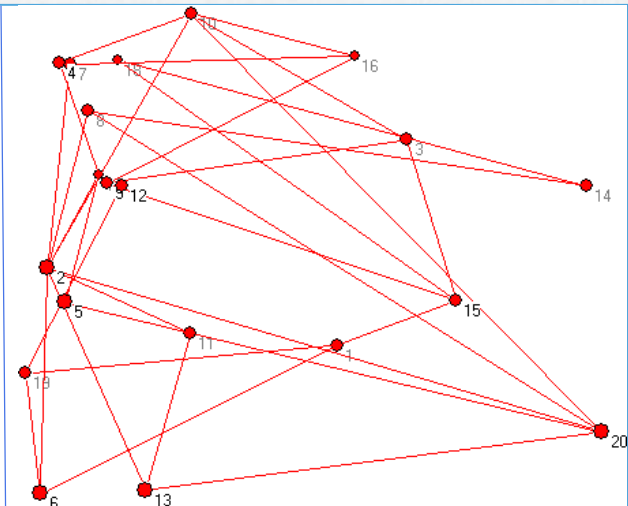
1. `AM = get_random_graph(20,3)`
2. Τα περιεχόμενά του είναι όμως μηδενικά και άσσοι. Επειδή θέλουμε το δίκτυό μας να είναι κατευθυνόμενο, εκτελούμε
➤ `AM = AM.*rand(20);`
3. `coordinates = rand(20,3)` (για μετατροπή σε 3D)
4. `write_matrix_to_pajek(AM,'random.net','weighted',true,'directed',false,'coords',coordinates);`
5. Ανοίγμα του αρχείου random.net με το Pajek
6. Απεικόνιση (περιστροφή με τα πλήκτρα x,y,z)

Ο πίνακας γειτνίασης AM

AM =

0	1	0	1	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0
0	1	0	0	0	1	0	1	0	0
1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0
0	0	1	1	0	0	0	1	0	0
0	0	0	1	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	1	0	0
0	0	0	0	0	1	0	0	0	0
1	0	0	0	0	0	1	0	0	0
0	0	0	0	0	1	0	0	1	0
0	0	0	0	0	0	0	0	0	0

Το 3d νευρωνικό δίκτυο στο Pajek



coordinates =

0.5312	0.7067	0.3508	(1 ^η γραμμή)
0.1088	0.5578	0.6855	
0.6318	0.3134	0.2941	
0.1265	0.1662	0.5306	
0.1343	0.6225	0.8324	
⋮	⋮	⋮	
0.9138	0.8711	0.6537	(20 ^η γραμμή)



Δημιουργία αρχείου δικτύου με την Python



Η γλώσσα προγραμματισμού Python είναι πλέον η δεύτερη δημοφιλέστερη, μετά την Java (και στην ακαδημαϊκή κοινότητα, η πρώτη).

Ή υποστήριξη της γλώσσας προγραμματισμού Python για χειρισμό γράφων και δικτύων

- ✓ Βιβλιοθήκη NetworkX
- ✓ Βιβλιοθήκη iGraph
- ✓ Και οι δύο διαθέτουν υποστήριξη για το Pajek

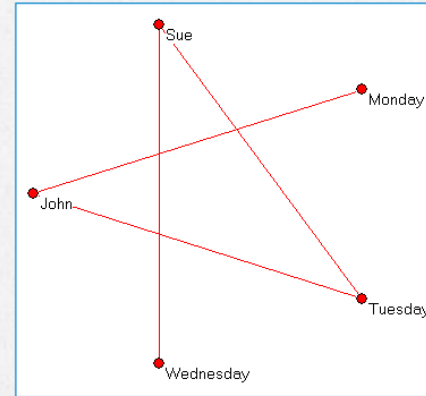
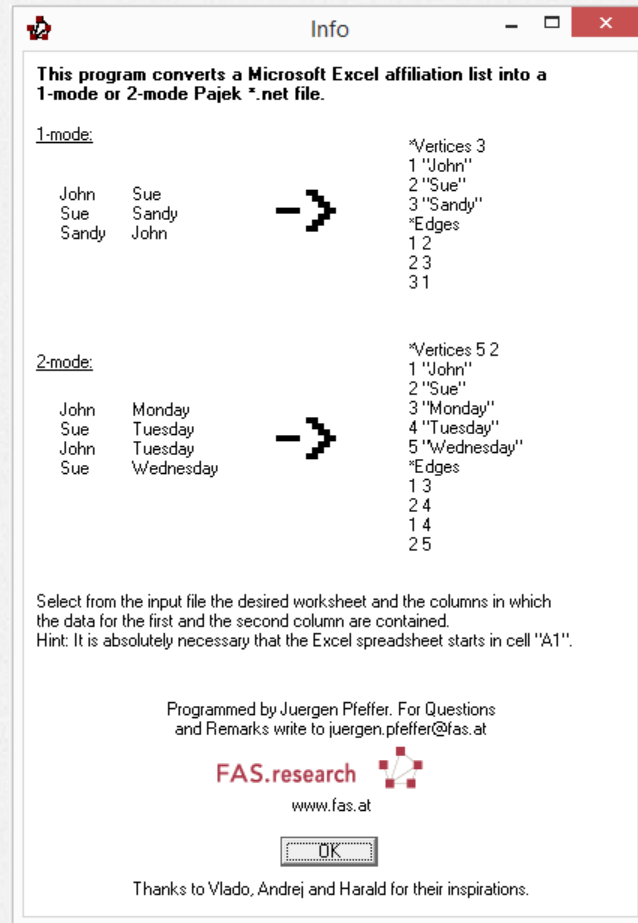
Η υλοποίηση της βιβλιοθήκης NetworkX για το Pajek, υποστηρίζει κατευθυνόμενους και μη-κατευθυνόμενους γράφους, συμπεριλαμβανομένων εκείνων με κόμβους που με βρόχους προς τον εαυτό τους, αλλά και εκείνων με παράλληλες ακμές.

<code>read_pajek</code> (path[, encoding])	Read graph in Pajek format from path.
<code>write_pajek</code> (G, path[, encoding])	Write graph in Pajek format to path.
<code>parse_pajek</code> (lines)	Parse Pajek format graph from string or iterable.



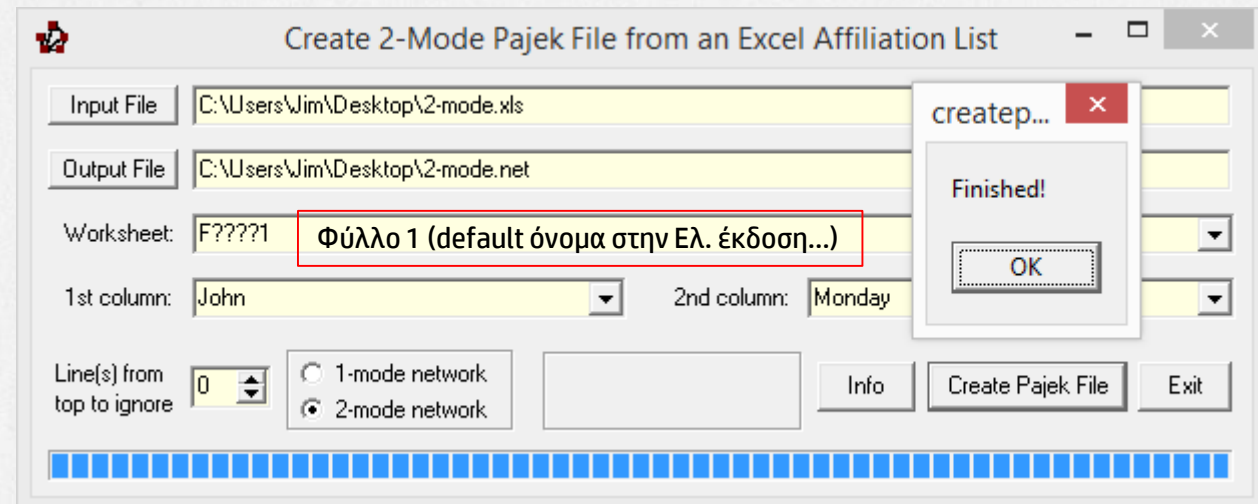
Δημιουργία αρχείου δικτύου με το Excel2Pajek

Εφαρμογή που μετατρέπει αρχείο Excel σε αρχεία δικτύου .net.



Βήματα

1. Δημιουργία ενός νέου αρχείου Excel
2. Εισαγωγή δεδομένων στις στήλες όπως στο υπόδειγμα της εικόνας (Info).
3. Εισαγωγή του xls στην εφαρμογή Excel2Pajek
4. Επιλογή για 1 ή 2-mode δίκτυο
5. Επιλογή directed ή undirected (μόνο για 1-mode)
6. Κλικ στο Create Pajek File
7. Άνοιγμα και απεικόνιση του αρχείου .net που προέκυψε, με το Pajek



Το αρχείο Excel θα πρέπει να έχει επέκταση xls (και όχι xlsx).

Δημιουργία αρχείου δικτύου μέσω σχεσιακής ΒΔ

Δύο tables μπορούν να αντιπροσωπεύσουν ένα δίκτυο one-mode. Το ένα table, θα περιλαμβάνει τις κορυφές και το άλλο, τις συνδέσεις.

Όλοι οι πίνακες

- Countries**
 - Countries : Πίνακας
- Imports: Cereals**
 - Imports: Cereals : Πίνακας
- Imports: Manufactures**
 - Imports: Manufactures : Πίνακας
- Neighbors**
 - Neighbors : Πίνακας
- Σφάλματα μετατροπής**
 - Σφάλματα μετατροπής : Πίνακας
- Αντικείμενα που δεν έχουν σχέση**
 - Imports: Cereals
 - Export a partition
 - Export a vector
 - Export network to Pajek: arcs and edges
 - Export network to Pajek: just arcs
 - Export network to Pajek: just edges
 - List of arcs
 - List of edges
 - List of vertices

Βήματα

1. Δημιουργία μιας νέας βάσης στην Access
2. Κατασκευή των πινάκων που θα παίξουν το ρόλο των κορυφών και των συνδέσεών τους (πίνακας γειτνίασης).
3. Σύνταξη ερωτημάτων για την εξαγωγή δεδομένων με τη μορφή partition, vector κ.α.
4. Σχεδίαση φόρμας εισαγωγής δεδομένων
5. Εισαγωγή δεδομένων
6. Με δεξί κλικ στο ερώτημα “Export network to Pajek: arcs and edges” μπορούμε να το εξάγουμε σε αρχείο .net

vertex1	vertex2	value
United States	Canada	1
Egypt	Israel	1
Morocco	Algeria	1
Indonesia	Malaysia	1
Malaysia	Thailand	1
Bangladesh	India	1
India	Pakistan	1
India	China	1
Pakistan	China	1
Jordan	Israel	1
Turkey	Greece	1
Croatia	Slovenia	1
Croatia	Hungary	1

Export network to Pajek: arcs and edges

- Άνοιγμα
- Προβολή διάρθρωσης
- Προβολή σχεδίασης
- Εξαγωγή**
 - Excel
 - Λίγα του SharePoint
 - Αρχείο RTF του Word
 - PDF ή XPS
 - Access
 - Αρχείο κειμένου
 - Αρχείο XML
 - Βάση δεδομένων ODBC
 - Σγγραφο HTML
 - Συγχώνευση με το Word
- Μετονομασία
- Απόκρυψη σε αυτή την ομάδα
- Διαγραφή
- Αποκοπή
- Αντιγραφή
- Επικόλληση
- Εκτύπωση...
- Προεπισκόπηση εκτύπωσης
- Ιδιότητες προβολής

Το παράδειγμα αυτό μπορεί να βρεθεί στο βιβλίο: [Exploratory Social Network Analysis with Pajek](#)



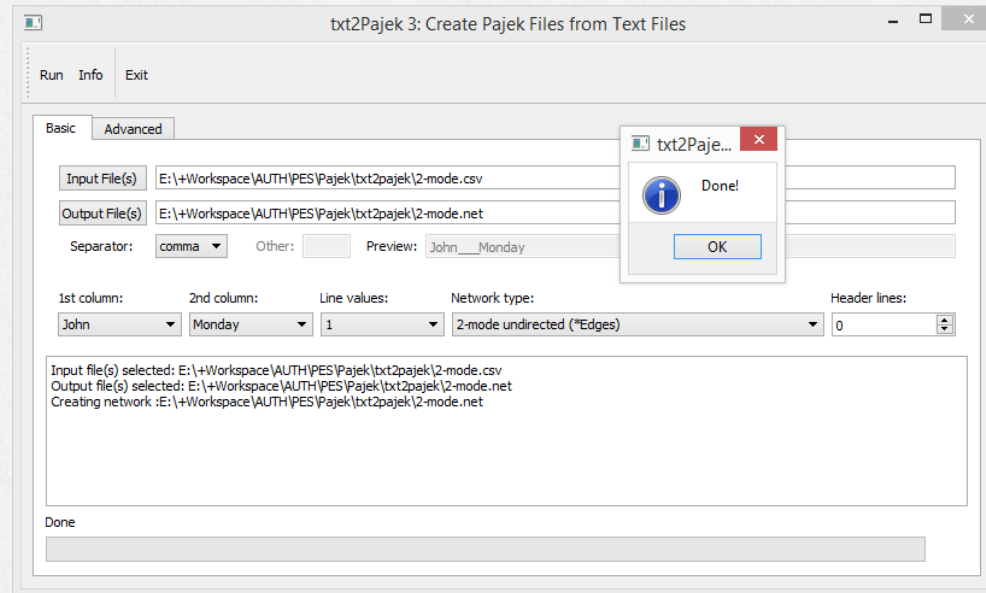
Άλλοι τρόποι δημιουργίας αρχείου δικτύου



Το Pajek μπορεί να επικοινωνήσει μέσω αρχείων κειμένου σχεδόν με οποιαδήποτε γλώσσα προγραμματισμού και συναφή εφαρμογή.

Η εφαρμογή TXT2Pajek.exe

Λαμβάνει ως είσοδο ένα αρχείο κειμένου (separator: tab, comma, semicolon, κ.α.) και επιστρέφει ένα αρχείο .net .



<http://www.pfeffer.at/txt2pajek>

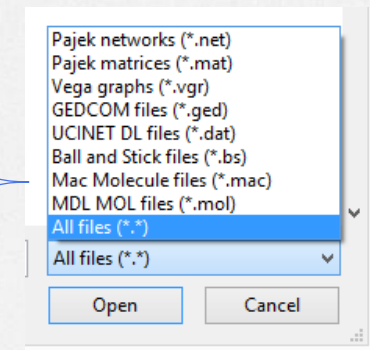
Αρχεία άλλων εφαρμογών

Ένας άλλος τρόπος δημιουργίας δικτύου με το Pajek, είναι πρώτα αυτό να εξαχθεί από κάποια άλλη εφαρμογή, σε κάποιο αποδεκτό format, και στη συνέχεια να γίνει η εισαγωγή του μέσω του Pajek.

Εντολή εισαγωγής αρχείου: **File > Network > Read**

Οι τύποι που υποστηρίζονται είναι:

- ✓ UCINET DL (gephi)
- ✓ Vega (Tomaz Pisanski project)
- ✓ GEDCOM (Genealogical Data)
- ✓ BS (Ball and Stick)
- ✓ MAC (Mac Molecule)
- ✓ MOL (MDL MOLfile)



Και η λίστα δεν τελειώνει εδώ...



Καθορισμός διάταξης δικτύου

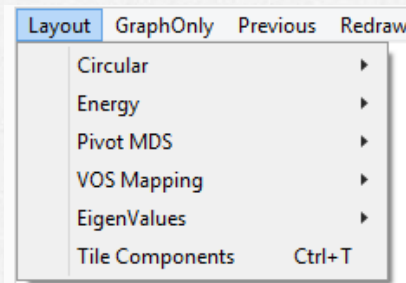


Η κεντρική ιδέα ενός δικτύου συνήθως λαμβάνεται από τον άνθρωπο μέσω της απεικόνισής της διάταξής του.

Αυτόματες ρουτίνες καθορισμού διάταξης δικτύου

Οι κυριότερες είναι:

- Ενέργειας
- Pivot MDS
- VOS Mapping



Μέσω των ρουτινών ενέργειας, το δίκτυο παρουσιάζεται ως ένα φυσικό σύστημα (πχ: μια πρωτεΐνη) του οποίου ψάχνουμε την βέλτιστη (ελάχιστη) ενεργειακή κατάσταση. Υπάρχουν δύο αλγόριθμοι:

- Kamada – Kawai. Σχεδίαση σε ένα επίπεδο.
- Fruchterman – Reingold. Σχεδίαση 2D ή 3D.

Οι ρουτίνες Pivot MDS και VOS Mapping εφαρμόζονται σε πολύ μεγάλα δίκτυα, των εκατό χιλιάδων κορυφών (Pivot MDS) και άνω (VOS Mapping).

Χειροκίνητος καθορισμός διάταξης δικτύου

- Λειτουργεί συμπληρωματικά της αυτόματης, όταν κρίνονται χρήσιμες ορισμένες διορθώσεις οι οποίες βελτιώνουν την τελική απεικόνιση του δικτύου.
- Μπορεί όμως να χρησιμοποιηθεί και μόνο αυτή, ιδιαίτερα όταν πρόκειται για πολύ απλά δίκτυα.

Χειρισμός

- Μπορούμε με το ποντίκι να μετακινούμε την κάθε κορυφή κατά βούληση.
- Περιστροφή εικόνας 3d ή 2d δικτύου με τα πλήκτρα: x, X, y, Y, z, Z (γράμμα = άξονας περιστροφής, κεφαλαία ή μικρά = θετική ή αρνητική διεύθυνση, αντίστοιχα).
- Μπορεί να επιλεγεί οποιοσδήποτε άξονας περιστροφής επιλέγοντας Spin > Normal. Στη συνέχεια, η περιστροφή γύρω από τον επιλεγμένο άξονα γίνεται με τα πλήκτρα S και s.



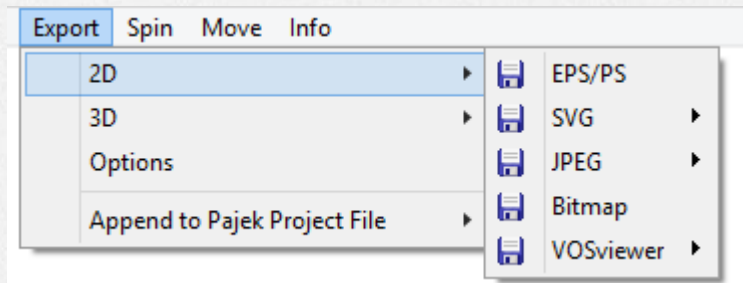
Μορφές αρχείων εξόδου του Pajek



Οι απεικονίσεις των δικτύων που δημιουργούνται μέσω του Pajek μπορούν να εξαχθούν σε διάφορες μορφές αρχείων.

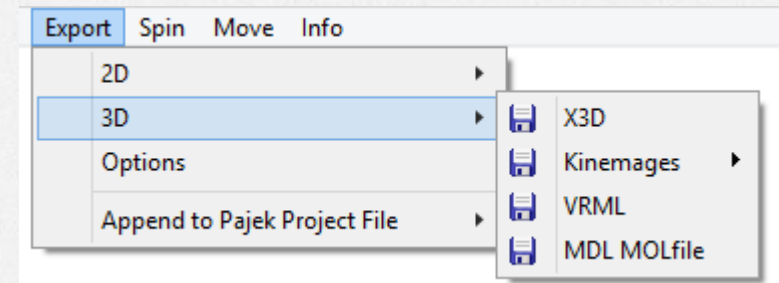
Δισδιάστατη μορφή

Μενού: Export > 2D



Τρισδιάστατη μορφή

Μενού: Export > 3D

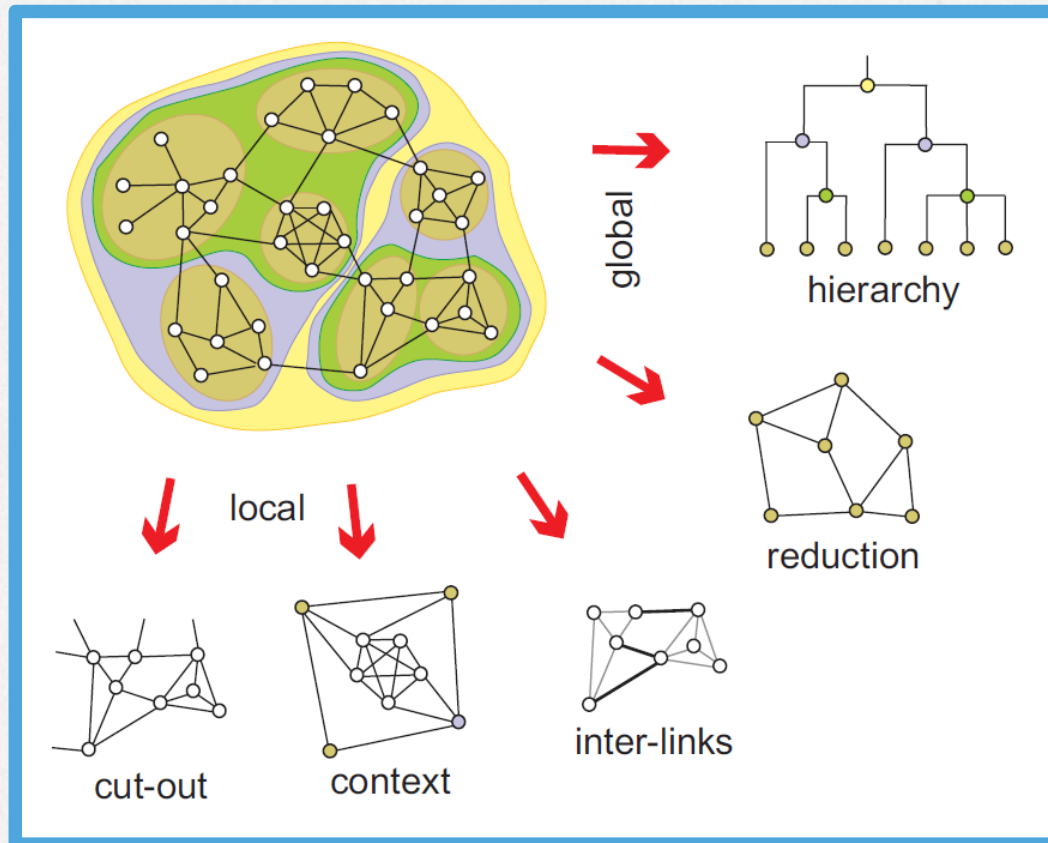




2ο μέρος

Τοπικές και ευρείες προβολές δικτύου

Οι τοπικές προβολές αφορούν συνδέσεις κορυφών μέσα στο ίδιο cluster, ενώ οι ευρείες προβολές, μεταξύ διαφορετικών.



Τοπική προβολή (local view)

Δημιουργείται επιλέγοντας ένα υποδίκτυο το οποίο εξάγεται από κάποιο σύμπλεγμα κορυφών (πχ: σε μια τάξη μαθητών, οι σχέσεις μεταξύ των αγοριών μόνο).

Ευρεία προβολή (global view)

Επιτυγχάνεται μέσω της συρρίκνωσης κορυφών που βρίσκονται στο ίδιο σύμπλεγμα (cluster) σε μια νέα κορυφή.

Με αυτό τον τρόπο αναδεικνύονται οι σχέσεις που υπάρχουν μεταξύ των clusters (πχ: σε μια τάξη μαθητών, οι σχέσεις μεταξύ αγοριών και κοριτσιών, δηλαδή ο συνολικός αριθμός των τόξων που ενώνουν τις δύο αυτές ομάδες).

Προβολή συνάφειας (contextual view)

Contextual = συνδυασμός Local & Global

Προβάλλονται οι σχέσεις που υπάρχουν μεταξύ των clusters (πχ: για κάθε κορίτσι της τάξης, ποιος είναι ο συνολικός αριθμός των τόξων που συνδέεται με αγόρια)



Βήματα εξαγωγής υποδικτύων



Προϋπόθεση: η ύπαρξη ενός αρχείου δικτύου (.net) κι ενός αρχείου partition (.clu) το οποίο έχει προκύψει από το δίκτυο αυτό.

Επιλογή υποδικτύου

Από το συνολικό δίκτυο επιλέγουμε το υποδίκτυο που μας ενδιαφέρει.

Επιλογή αριθμών και διαστημάτων των clusters

Select clusters

Example: 1-3,6,10-15

OK Cancel

1
↓
2
↓
3

Μενού

Operations > Network + Partition > Extract > SubNetwork

Παράδειγμα θα μπορούσε να αποτελεί και το υποδίκτυο που προκύπτει διαχωρίζοντας τις εισαγωγές και τις εξαγωγές μετάλλου, εντός μιας ηπείρου, από το παγκόσμιο εμπόριο μετάλλου.



Μείωση μεγέθους δικτύου



Η συρρίκνωση ενός δικτύου παρουσιάζει πολλά πλεονεκτήματα. Π.χ.: Η διευκόλυνση της οπτικής παρατήρησής και της ανάλυσής του.

1. Συρρίκνωση

Operations > Network + Partition > Shrink Network

Shrinking according to Partition

Minimum Number of Lines between Clusters

Cluster that will not be Shrunk

OK Cancel

2. Αφαίρεση των γραμμών με χαμηλή τιμή

Μπορούμε να αφαιρέσουμε τις γραμμές αυτές επιλέγοντας:

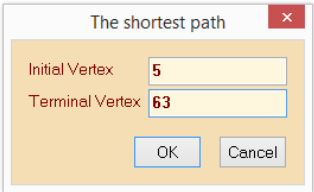
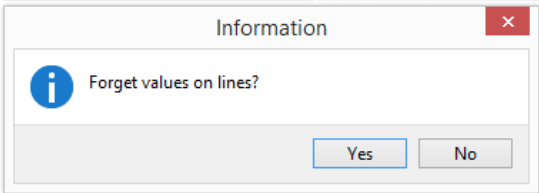
Network > Create New Network > Transform > Remove > Lines with Value > lower

Και δίνοντας εν συνέχεια στο Pajek, το αντίστοιχο αριθμητικό κατώφλι που θα λειτουργεί ως κριτήριο.

Παράδειγμα χρήσης θα μπορούσε να αποτελεί και η συρρίκνωση του δικτύου του παγκόσμιου εμπορίου μετάλλου, από το επίπεδο χωρών στο επίπεδο ηπείρων.

Μονοπάτι ή διαδρομή: Μία πεπερασμένη ή άπειρη ακολουθία διακριτών ακμών, που συνδέουν μια ακολουθία διακριτών κορυφών.

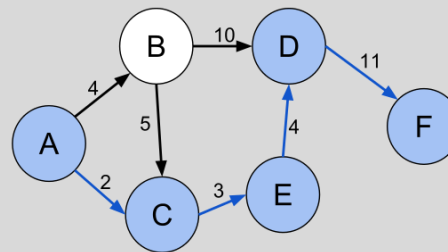
Εύρεση συντομότερου μονοπατιού

1. **Network > Create New Network > SubNetwork with Paths > One Shortest Path between Two Vertices**
2. 
3. 
4. Identify vertices in source network? **No**
5. **Layout > Energy > Kamada Kawai > Fix first and last**
6. **Options > Mark Vertices Using > Labels**
7. **Network > Create New Network > Transform > Arcs to Edges > All**
8. **Network > Create New Network > SubNetwork with Paths > All Shortest Paths between Two Vertices**

Εύρεση διαμέτρου

1. **Network > Create New Network > SubNetwork with Paths > Info on Diameter**
2. Το Pajek επιστρέφει μόνο τις δύο πιο απομακρυσμένες κορυφές
3. Αν επιθυμούμε να έχουμε και το μονοπάτι ανάμεσά τους, θα πρέπει στο αποτέλεσμα, να εκτελέσουμε επιπλέον και την εντολή εύρεσης του συντομότερου μονοπατιού ανάμεσα σε αυτές τις δύο κορυφές.

Παράδειγμα shortest path



K-neighbors

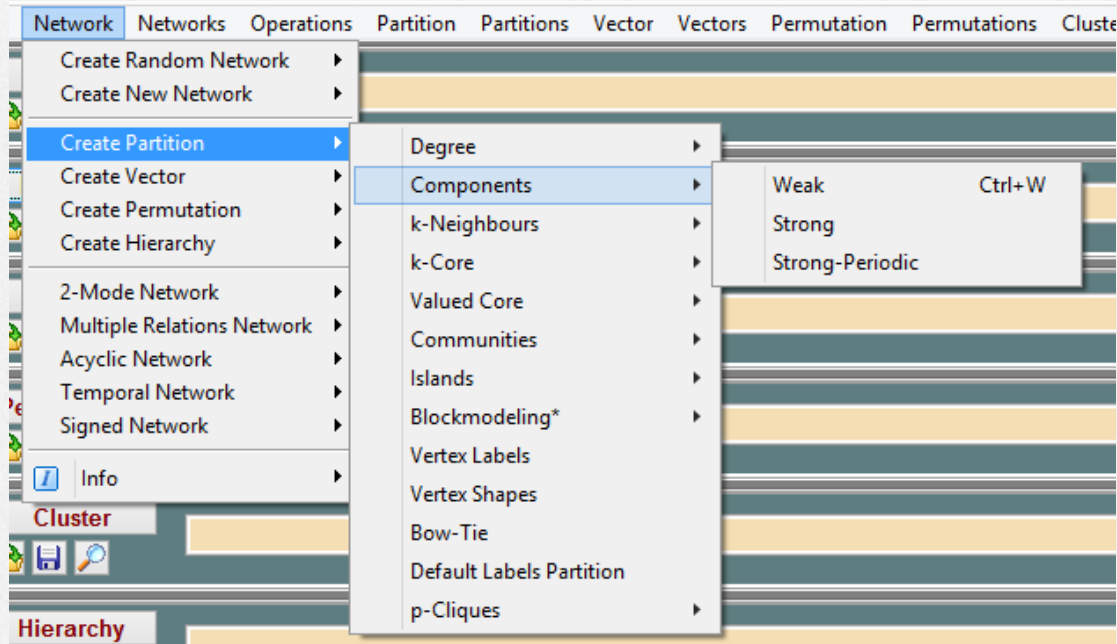
1. **Network>Create Partition > k-Neighbors > Output**, αποστάσεις όλων των κορυφών από συγκεκριμένη κορυφή
 2. Εισάγουμε την επιλεγμένη κορυφή και την απόσταση που θέλουμε να διερευνήσουμε (εισάγοντας το μηδέν υπολογίζονται όλες οι αποστάσεις).
 3. Το αποτέλεσμα είναι ένα partition – πίνακας (2xη), όπου για κάθε κορυφή παίρνουμε και την αντίστοιχη απόστασή της από την επιλεγμένη.
 4. **Network>Create Partition > k-Neighbors > Input**, υπολογίζουμε τα βήματα που απαιτούνται προκειμένου από όλες τις άλλες κορυφές, να φτάσουμε στην επιλεγμένη
 5. **Network>Create Partition > k-Neighbors > All**, υπολογίζουμε τις αποστάσεις χωρίς να λαμβάνουμε υπόψη τις κατευθύνσεις των γραμμών
- ❖ **Partition > Info**, ο αριθμός (που θα επιλέξουμε) των κοντινότερων (+) και των πιο απομακρυσμένων (-) κορυφών από την επιλεγμένη κορυφή.



Ισχυρώς και Ασθενώς συνδεδεμένα στοιχεία



Το αποτέλεσμα του υπολογισμού είναι ένα partition στο οποίο οι κορυφές που ανήκουν στο ίδιο στοιχείο (συνιστώσα) έχουν τον ίδιο αριθμό.



Υπολογισμός στο Pajek

Network > Create Partition > Components > Weak

Network > Create Partition > Components > Strong

Ορισμοί

Ένας κατευθυνόμενος γράφος ονομάζεται ασθενώς συνδεδεμένος αν η αντικατάσταση όλων των κατευθυνόμενων ακμών του με μη κατευθυνόμενες ακμές, παράγει ένα συνδεδεμένο (μη-κατευθυνόμενο) γράφημα.

Ισχυρώς συνδεδεμένος ονομάζεται αν περιέχει ένα κατευθυνόμενο μονοπάτι από τον u στον v και ένα κατευθυνόμενο μονοπάτι V και U για κάθε ζεύγος κορυφών u, v .

[http://en.wikipedia.org/wiki/Connectivity_\(graph_theory\)](http://en.wikipedia.org/wiki/Connectivity_(graph_theory))

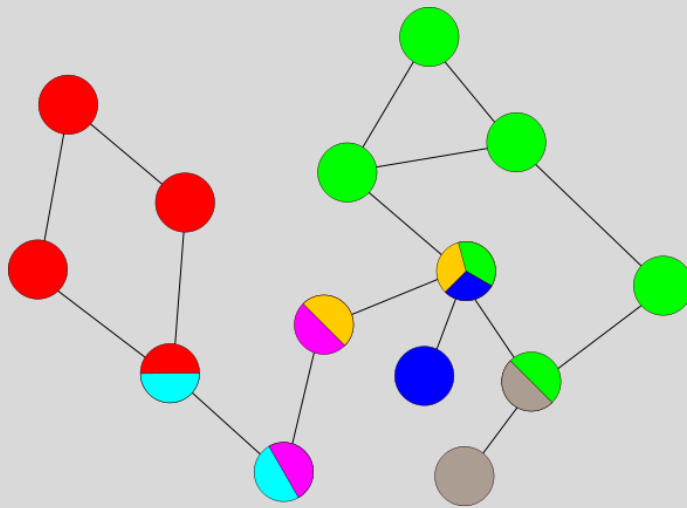


Αμφισυνδεδεμένα στοιχεία - κορυφές



Στη θεωρία γράφων, ένα αμφισυνδεδεμένο (2-connected) στοιχείο είναι ένας μέγιστος αμφισυνδεδεμένος υπό-γράφος.

Παράδειγμα



Οι κορυφές με τα 2 χρώματα ονομάζονται αμφισυνδεδεμένες

http://en.wikipedia.org/wiki/Biconnected_component

Υπολογισμός αμφιστοιχείων στο Pajek

Network > Create New Network > with Bi-Connected Components stored as Relation Numbers

Τρόπος αποθήκευσης αμφιστοιχείων

Τα αμφισυνδεδεμένα (biconnected ή 2-connected) στοιχεία, αποθηκεύονται σε ιεραρχίες. Ένας κόμβος σε μια ιεραρχία αντιπροσωπεύει ένα αμφισυνδεδεμένο στοιχείο.

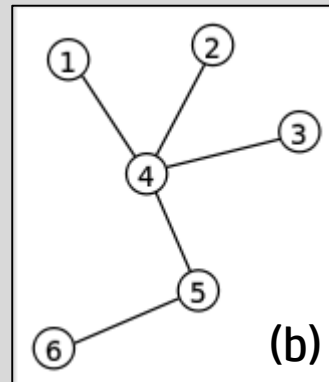
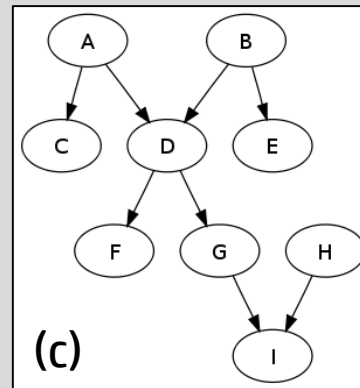
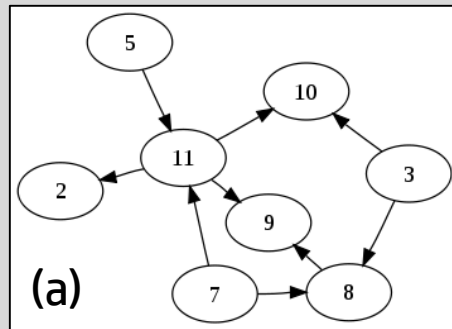
Σε μια ιεραρχία μπορούμε να πλοηγηθούμε όπως για παράδειγμα και στις δομές καταλόγων ενός παραθυρικού λειτουργικού συστήματος.

Εξαγωγή αμφιστοιχείων

Εξαγωγή ενός συγκεκριμένου αμφιστοιχείου (ή αμφικορυφής, ή αμφικόμβου): **Hierarchy > Extract Cluster**, δίνοντας στη συνέχεια τον αριθμό του κόμβου της ιεραρχίας που μας ενδιαφέρει.

Για την εξαγωγή του υποδικτύου που προκύπτει από την πιο πάνω ενέργεια: **Operations > Network + Cluster > Extract SubNetwork**.

Παραδείγματα: Κατευθυνόμενο άκυκλο δίκτυο (a), Δέντρο (b), Πολύδεντρο (c)



Δομή άκυκλων δικτύων

Σε ένα άκυκλο δίκτυο υπάρχουν οι πρώτες κορυφές (επίπεδο 1) στις οποίες δεν φθάνουν γραμμές, οι ενδιάμεσες, και οι τελικές (επίπεδο n) από τις οποίες δε ξεκινά καμία γραμμή σύνδεσης.

Επομένως κάθε άκυκλο δίκτυο έχει ένα βάθος, το μέγεθος του οποίου εξαρτάται από το σύνολο των επιπέδων των κορυφών του.

Υπολογισμός βάθους άκυκλου δικτύου στο Pajek

Network > Acyclic Network > Depth Partition > Acyclic

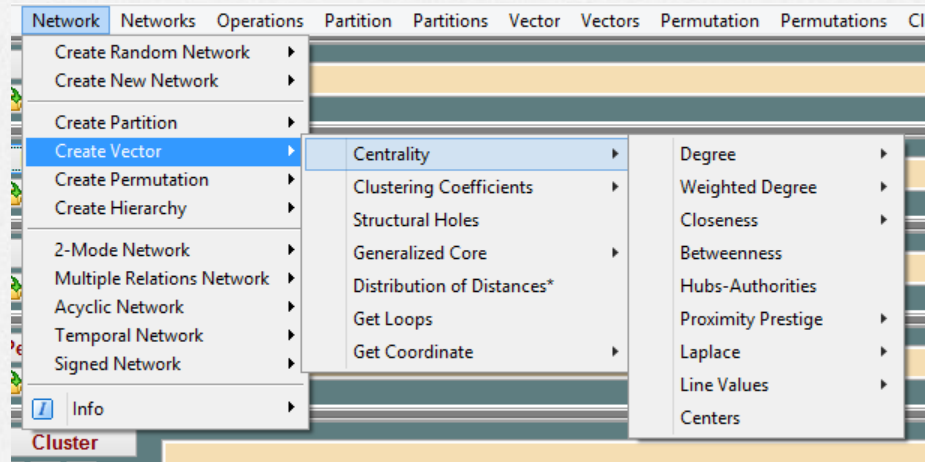
http://en.wikipedia.org/wiki/Acyclic_graph



Δείκτες κεντρικότητας (centrality indicators)



Οι δείκτες κεντρικότητας χρησιμοποιούνται για τον εντοπισμό των πιο σημαντικών κορυφών σε ένα δίκτυο.



Κυριότεροι υπολογισμοί κεντρικότητας στο Pajek

Network > Create Vector > Centrality > Betweenness

Network > Create Vector > Centrality > Closeness

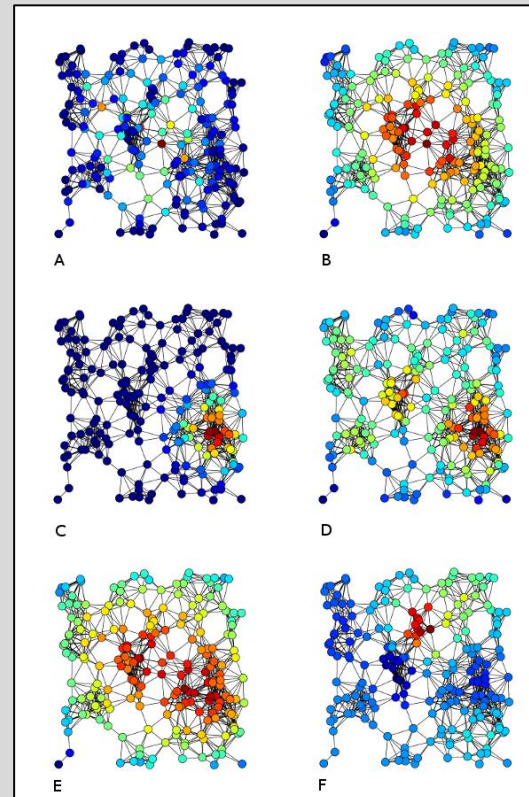
Network > Create Vector > Centrality > Degree

<http://en.wikipedia.org/wiki/Centrality>

Θεωρία

Δείκτες κεντρικότητας

- A. Ενδιαμεσότητα – Δείχνει το πόσο κεντρικός είναι ένας κόμβος του δικτύου
- B. Πυκνότητα – Απόσταση μεταξύ όλων των ζευγών των κόμβων, που ορίζεται από το μήκος των shortest paths τους
- C. Ιδιοδιάνυσμα – Μέτρο της επίδρασης ενός κόμβου σε ένα δίκτυο
- D. Βαθμός – Ο προσπίπτων αριθμός συνδέσεων σε ένα κόμβο
- E. Αρμονική – Συνδέσεις κόμβων προς κόμβο, από απόσταση 1 έως n
- F. Leo Katz



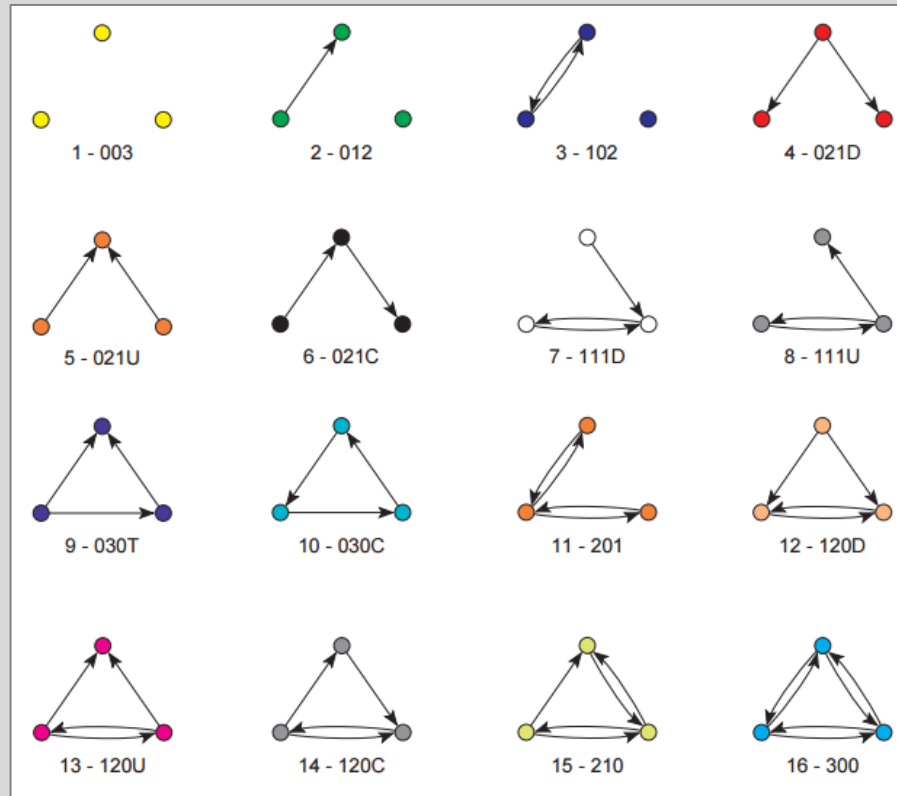


Τριαδική απογραφή (triadic census)

Ονομάζεται η εύρεση όλων των μοτίβων (patterns ή cliques) τριάδων ενός κατευθυνόμενου γράφου – δικτύου.



Όλα τα πιθανά τριαδικά μοτίβα (αριθμοί ακολουθιών Pajek)



Υπολογισμός τριαδικής απογραφής στο Pajek

Network > Info > Triadic Census

<http://vlado.fmf.uni-lj.si/pub/networks/doc/triads/triads.pdf>



Πυρήνες (cores)

Πυρήνας: Μετασχηματισμός ενός αρχικού δικτύου σε ένα άλλο, που διατηρεί τις σχέσεις μεταξύ των στοιχείων του πρώτου (ομομορφικός).



Παραδείγματα πυρήνων

- Πλήρεις γράφοι
- Κύκλοι περιττού μήκους
- Κύκλοι άρτιου μήκους (ανά δύο)

Βήματα υπολογισμού πυρήνων στο Pajek

1. **Network > Create Partition > K-Core**
2. **Input, Output ή All**

Αποτέλεσμα: Ένα partition για κάθε κορυφή του οποίου δίνεται και ο αριθμός του πυρήνα που ανήκει.

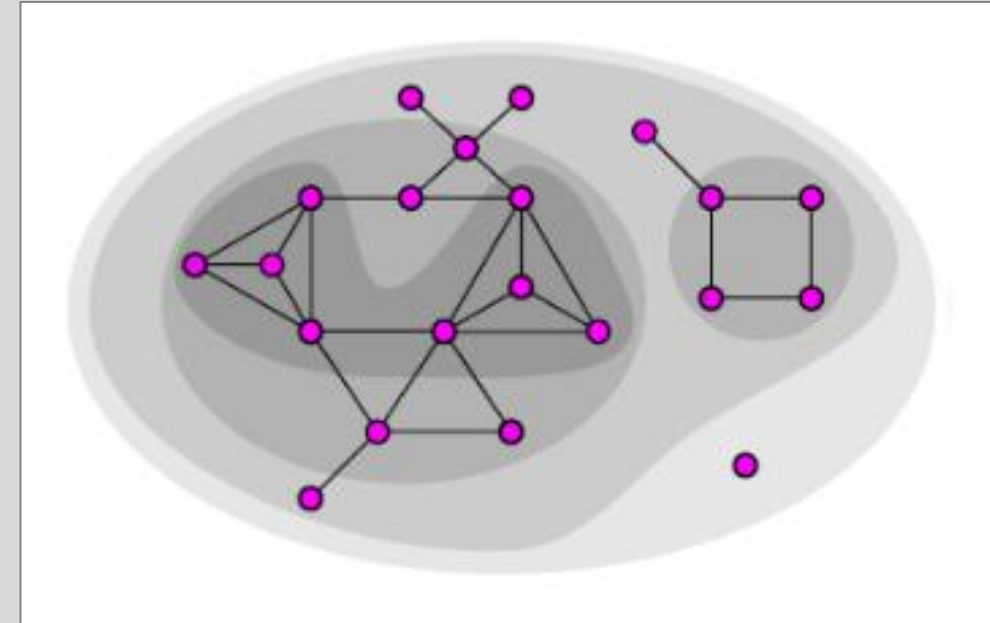
Στις περισσότερες περιπτώσεις ενδιαφερόμαστε μόνο για τον μεγαλύτερο πυρήνα (ή πυρήνες):

3. **Operations > Network + Partition > Extract SubNetwork**
4. Εισάγουμε τον ίδιο αριθμό (του μεγαλύτερου πυρήνα) δύο φορές

[http://en.wikipedia.org/wiki/Core_\(graph_theory\)](http://en.wikipedia.org/wiki/Core_(graph_theory))

Θεωρία: k-πυρήνας (k-core)

Ένα υποσύνολο κορυφών ονομάζεται k-πυρήνας αν κάθε κορυφή του υποσυνόλου είναι συνδεδεμένη με τουλάχιστον k κορυφές του υποσυνόλου αυτού. Στην εικόνα διακρίνονται πυρήνες διαφόρων μεγεθών (0, 1, 2 και 3):





Κλίκες (cliques)

Η κλίκα είναι ένας ειδικός τύπος πυρήνα



Αναζήτηση κλικών

- Είναι επεξεργαστικά πιο απαιτητική σε σχέση με την αναζήτηση πυρήνων
- Προτιμάται η αναζήτηση κλικών μεγέθους 3 ή 4 σε μικρά δίκτυα

Αναζητώντας κλίκες μεγέθους 3 (3-cliques), μπορούμε να βρούμε 16 διαφορετικούς συνδυασμούς όπως είδαμε ήδη και προηγουμένως (τριαδική απογραφή).

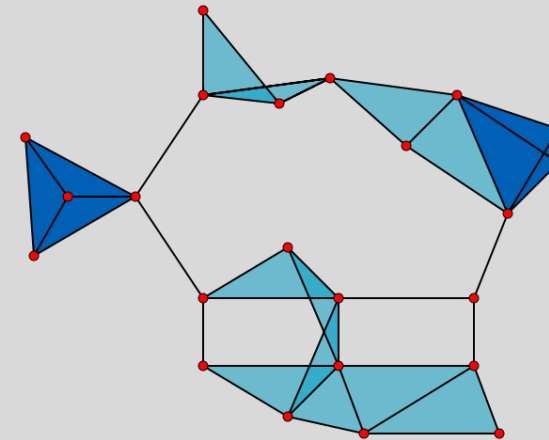
Αν όμως θέλουμε να βρούμε όλες τις εμφανίσεις των κλικών μεγέθους 3, δηλαδή και τύπους τριάδων που εμφανίζονται περισσότερες από μία φορές ?

Τότε κάνουμε χρήση μιας γενικότερης διαδικασίας, κατά την οποία γίνεται αναζήτηση τμημάτων ή αλλιώς προτύπων (Fragments), στο δίκτυο.

[http://wikipedia.org/wiki/Clique_\(graph_theory\)](http://wikipedia.org/wiki/Clique_(graph_theory))

Θεωρία

Ένα υποσύνολο ενός δικτύου ονομάζεται κλίκα, εάν καθεμιά κορυφή του υποσυνόλου αυτού συνδέεται με την καθεμιά από τις άλλες κορυφές του.



23 × 1-vertex cliques (οι κορυφές)
42 × 2-vertex cliques (οι γραμμές σύνδεσης)
19 × 3-vertex cliques (τα ανοικτά και σκούρα μπλε τρίγωνα)
2 × 4-vertex cliques (μόνο οι σκούρες μπλε περιοχές)



Αναζήτηση προτύπων (fragments)



Είναι η διαδικασία κατά την οποία γίνεται μια προσπάθεια εντοπισμού μικρότερων δικτύων σε ένα δίκτυο (πχ: αναζήτηση των 3-cliques)

Αναζήτηση προτύπων στο Pajek

1. Σχεδιάζουμε το πρότυπο οι ίδιοι, ως ένα νέο, ανεξάρτητο και πλήρες κατευθυνόμενο δίκτυο:

Network > Create New Network > Complete Network > Directed

2. Ορίζουμε το νέο μικρό δίκτυο ως fragment, επιλέγοντάς το ώστε να είναι το πρώτο δίκτυο στην αντίστοιχη dropdown λίστα. Φροντίζουμε το αρχικό δίκτυο να είναι το δεύτερο

3. Αν υπάρχουν γραμμές θα πρέπει να μετατραπούν σε αμφίδρομα τόξα, επειδή ο αλγόριθμος αναζήτησης κάνει διάκριση μεταξύ edges και bidirected arcs:

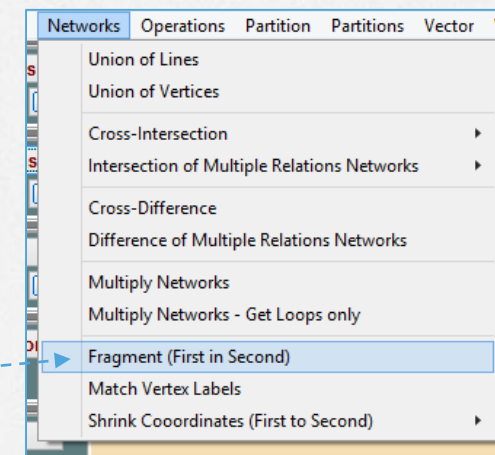
Network > Create New Network > Transform > Edges > Arcs

4. Πραγματοποιούμε την αναζήτηση:

Networks > Fragment (First in Second) (> Find)

Αποτελέσματα αναζήτησης

- ✓ Ένα υποδίκτυο το οποίο περιέχει μόνο όλα τα fragments που είναι ίδια με το δοσμένο κριτήριο αναζήτησης.
- ✓ Ένα partition με τιμές από μηδέν και άνω. Η τιμή μηδέν σημαίνει ότι η αντίστοιχη κορυφή δεν ανήκει σε κάποιο fragment. Μία τιμή a , σημαίνει ότι η αντίστοιχη κορυφή ανήκει σε a fragments.
- ✓ Μια ιεραρχία με όλα τα fragments





Μέθοδοι εντοπισμού κοινοτήτων (communities)



Οι δομές κοινοτήτων είναι πυκνά clusters όπου στο εσωτερικό τους υπάρχουν περισσότερες συνδέσεις απ' ότι στο εξωτερικό τους

Μέθοδοι εντοπισμού δομών κοινοτήτων στο Pajek

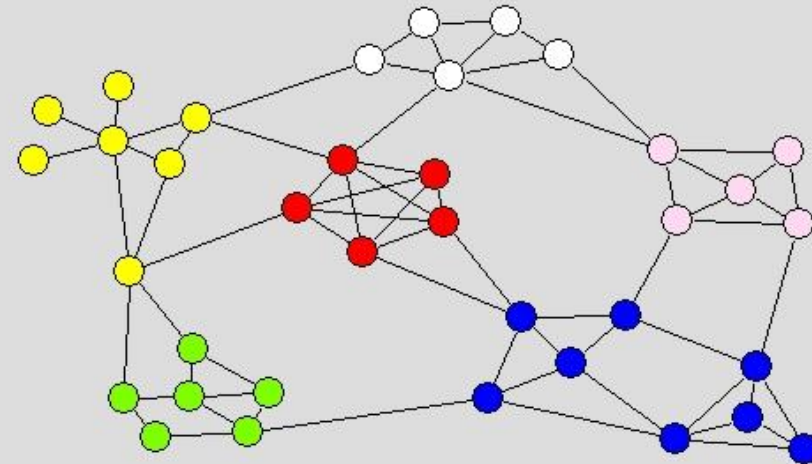
Το Pajek διαθέτει δύο μεθόδους εντοπισμού δομών κοινοτήτων:

- ✓ Μέθοδος Louvain
- ✓ Μέθοδος VOS clustering

<http://mrvar.fdv.uni-lj.si/pajek/community/CommunityDrawExample.htm>

<http://mrvar.fdv.uni-lj.si/pajek/community/LouvainVOS.htm>

Παράδειγμα



http://en.wikipedia.org/wiki/Community_structure



Δίκτυα δύο τύπων (2-mode)



Ένα 2-mode δίκτυο, αποτελείται από δύο είδη μονάδων - υποσυνόλων (πχ: υπάλληλοι και εταιρείες), τα οποία συσχετίζονται με συνδέσεις.

Δημιουργία 2-mode δικτύου στο Pajek

Network > 2-Mode Network > Partition into 2 Modes

Όπου στις κορυφές του πρώτου υποσυνόλου (υπάλληλοι) δίνεται ο αριθμός 1, και σε εκείνες του δεύτερου (εταιρείες), δίνεται ο αριθμός 2.

Βήματα μετασχηματισμού δικτύου 2-mode σε 1-mode

1. Εξάγουμε το πρώτο υποσύνολο των κορυφών:

Network > 2-Mode Network > 2-Mode to 1-Mode > Rows

2. Εξάγουμε το δεύτερο υποσύνολο:

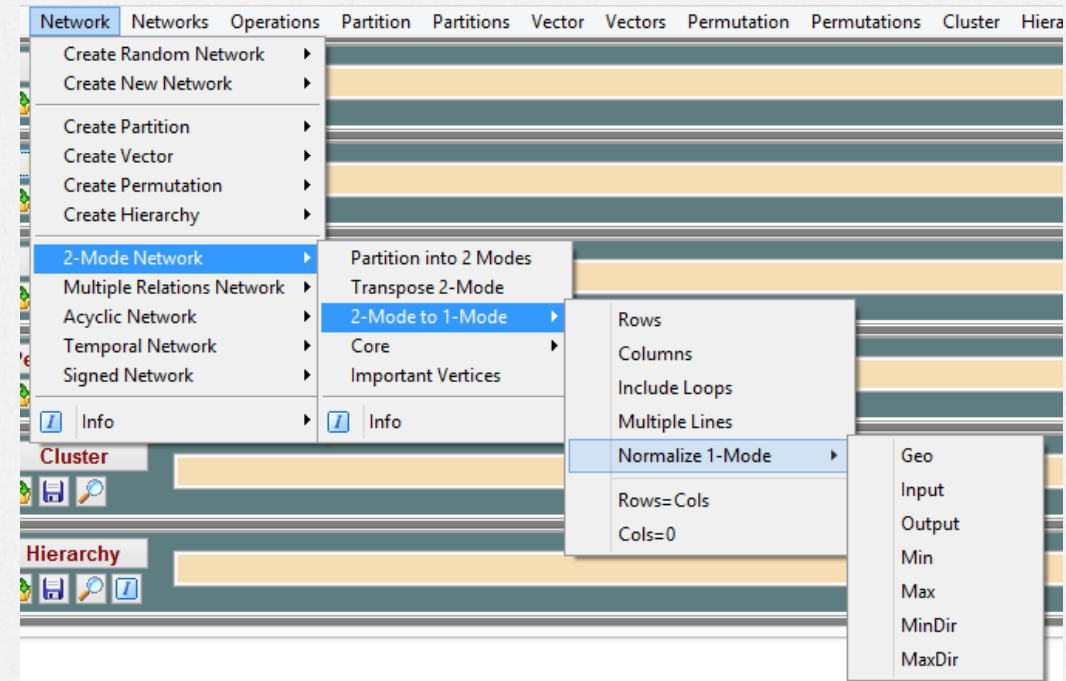
Network > 2-Mode Network > 2-Mode to 1-Mode > Columns

Επίσης μπορούμε να δημιουργήσουμε δίκτυο με βρόχους ή χωρίς:

Network > 2-Mode Network > 2-Mode to 1-Mode > Include Loops

Για τη δημιουργία δικτύων με βάρη (πχ: αριθμός υπαλλήλων της ίδια εταιρείας), ή δικτύων με πολλαπλές γραμμές:

Network > 2-Mode Network > 2-Mode to 1-Mode > Multiple Lines



<http://mrvar.fdv.uni-lj.si/sola/info4/uvod/twomode.pdf>



Απεικόνιση δικτύου με βάρη (αξίες, κόστη κλπ.)



Ένα σταθμισμένο δίκτυο αντιστοιχίζει μια ετικέτα (βάρος) με κάθε κόμβο στο δίκτυο αυτό.

Βήματα απεικόνισης στο Pajek

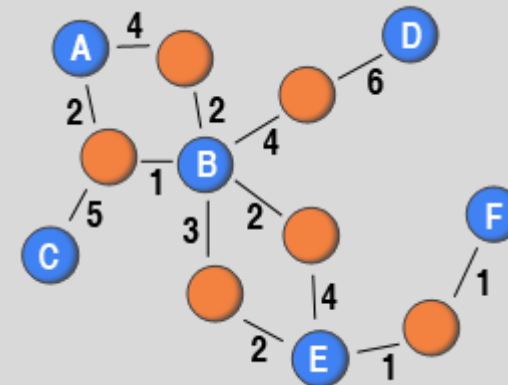
1. Αποθηκεύουμε πρώτα τις αξίες αυτές σε ένα αρχείο vector:
Network > Create Vector > Get Loops
2. Χρησιμοποιούμε το vector αυτό για να καθορίσουμε τα μεγέθη των κορυφών:
Draw > Network + Vector

Αφού δημιουργηθεί δίκτυο με τις αξιολογημένες γραμμές, μπορούμε να το απεικονίσουμε με πολλούς τρόπους:

- Απεικόνιση πλήρους δικτύου:
 - a. απεικονίζουμε το δίκτυο χρησιμοποιώντας την μέθοδο Energy, έχοντας επιλέξει **Options > Values of Lines > Similarities** (για να εμφανιστούν πιο κοντά οι γραμμές με τις υψηλότερες τιμές)
 - b. Οι αξίες των μπορούν να αναπαρασταθούν με διαφορετικά πάχη γραμμής (**Options > Lines > Different Widths**) ή/και με χρήση διαβαθμίσεων του γκρι (**Options > Lines > Grey-Scale**).
 - c. Αφού εξάγουμε την εικόνα σε μορφή SVG, μπορούμε να προσθέσουμε γραμμές ανάλογα με την αξία τους **Export > SVG > Line-Values > Nested Classes**
- ii. Απεικόνιση του πιο σημαντικού μέρους του δικτύου: Εμφανίζονται μόνο οι γραμμές με αξία πάνω από ένα κατώφλι.

Θεωρία: Βάρη, ή αξίες, ή κόστη, ή σταθμίσεις

Τα βάρη είναι συνήθως πραγματικοί αριθμοί. Μπορεί όμως να περιορίζονται σε ρητούς αριθμούς ή σε ακραίους. Ορισμένοι αλγόριθμοι απαιτούν περαιτέρω περιορισμούς σχετικά με τα βάρη. Για παράδειγμα, ο αλγόριθμος του Dijkstra δουλεύει σωστά μόνο για θετικά βάρη. Το βάρος ενός μονοπατιού ή το βάρος ενός δέντρου σε ένα σταθμισμένο δίκτυο, είναι το άθροισμα των βαρών των γραμμών (συνδέσεων) που το απαρτίζουν.



http://en.wikipedia.org/wiki/Glossary_of_graph_theory#Weighted_graphs_and_networks



Νησίδες (islands)

Σε ένα δίκτυο όπου δίνονται ορισμένες ιδιότητες (τιμές) των κόμβων ή των γραμμών, μπορούμε να βρούμε νησίδες



Υπολογισμός νησίδων στο Pajek

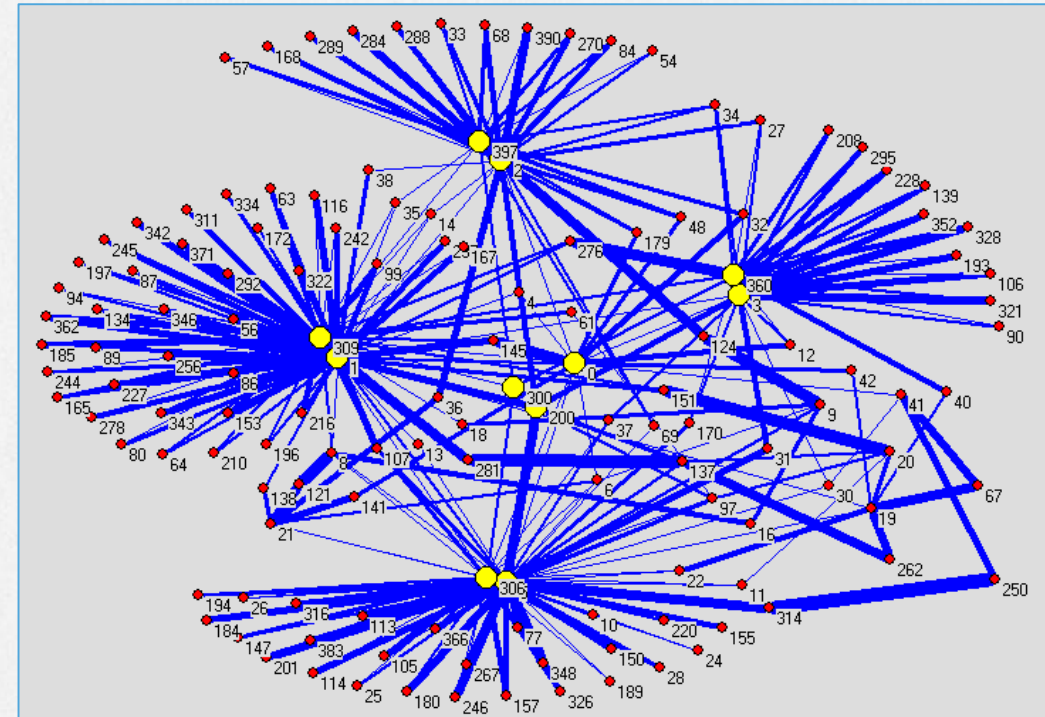
Οι νησίδες, ονομάζονται νησίδες κόμβων αν δίνονται οι αξίες των κόμβων, και νησίδες γραμμών, αν δίνονται οι αξίες των γραμμών.

Οι νησίδες μοιάζουν με τις κλίκες, δηλαδή, μπορούμε να τις εντοπίσουμε σε clusters κόμβων με αξίες σύνδεσης μεγαλύτερες εντός του cluster και μικρότερες εκτός.

Για τον εντοπισμό τους θα πρέπει επίσης να επιλέξουμε το μικρότερο και το μεγαλύτερο επιτρεπόμενο μέγεθος της νησίδας.

Στο Pajek οι γραμμές των νησίδων υπολογίζονται:

Network > Create Partition > Islands > Lines Weights



<http://vlado.fmf.uni-lj.si/pub/networks/doc/sunbelt/islands.pdf>



Ιεραρχικό clustering

Μέθοδος ανάλυσης συστάδων που επιδιώκει να οικοδομήσει μια ιεραρχία των συστάδων



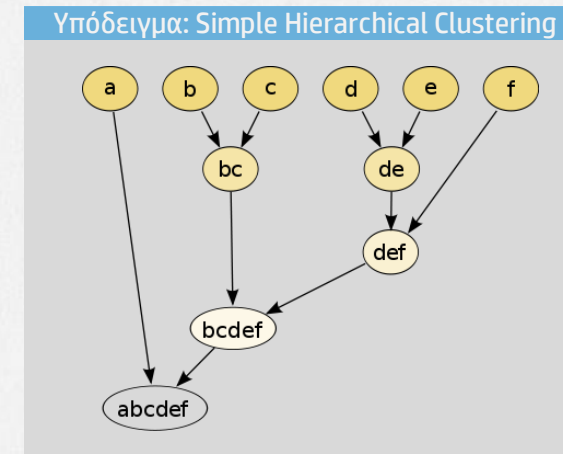
Βήματα ιεραρχικού clustering στο Pajek

1. Δημιουργία πλήρως συνδεδεμένου cluster προκειμένου οι ανομοιοότητες να υπολογιστούν για όλες τις μονάδες
Cluster > Create Complete Cluster
2. Υπολογισμός του dissimilarity matrix (πίνακας ανομοιοτήτων)
Operations > Network + Cluster > Dissimilarity > Network Based

➤ Επιλέγουμε:
 - i. **d1/All** αν θέλουμε να θεωρήσουμε το δίκτυο ως ένα δυαδικό πίνακα
 - ii. **Corrected Euclidian** ή **Corrected Manhattan**, απόσταση για δίκτυα με βάρη - αξίες.
3. Μας ζητείται να ονομάσουμε το αρχείο EPS που θα δημιουργηθεί στο οποίο θα αποθηκευτεί το δενδρόγραμμα που θα προκύψει
4. Επιλέγουμε OK προκειμένου να πραγματοποιηθεί το Ιεραρχικό clustering με βάση τον dissimilarity matrix που υπολογίστηκε.

Τα αποτελέσματα του ιεραρχικού clustering

- ✓ Ο dissimilarity matrix
- ✓ Η εικόνα EPS το δενδρογράμματος (άνοιγμα με GSview)
- ✓ Ένα αρχείο ανακατέματος .per (permutation) σύμφωνα με το δενδρόγραμμα το οποίο μπορεί να χρησιμοποιηθεί για την απεικόνιση του πίνακα αναδιάταξης (reordered matrix) στο EPS (**File > Network > Export Matrix to EPS > Using Permutation**).
- ✓ Ιεραρχία που αναπαριστά το ιεραρχικό clustering



http://en.wikipedia.org/wiki/Hierarchical_clustering



Μοντελοποίηση σε μπλοκ (Blockmodeling)



Μείωση ενός μεγάλου, και ενδεχομένως ασυνάρτητου δικτύου, σε μια μικρότερη πιο κατανοητή δομή που μπορεί να ερμηνευθεί ευκολότερα.

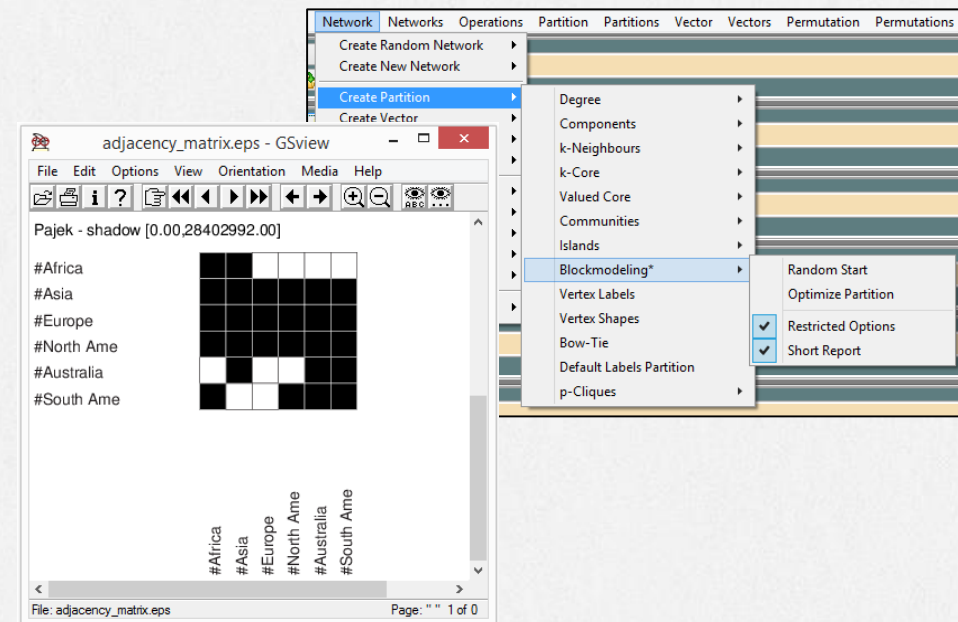
Βήματα μοντελοποίησης σε μπλοκ, στο Pajek

1. Επιλογή τρόπου μοντελοποίησης:
 - a. Ξεκινώντας με τυχαίο partition σε δοσμένο αριθμό clusters **Network > Create Partition > Blockmodeling* > Random Start**
 - b. Βελτιστοποιώντας το δοσμένο partition (πχ: partition που προέκυψε από ιεραρχικό clustering) **Network > Create Partition > Blockmodeling* > Optimize Partition**
2. Αφού επιλέξουμε τον τρόπο που θα πραγματοποιηθεί το blockmodeling, επιλέγουμε:
 - i. Τον τύπο ισοδυναμίας (typeof equivalence – **Structural** ή **Regular**). Μπορούμε επίσης να καθορίσουμε τους δικούς μας τύπους blocks επιλέγοντας **3-Define**.
 - ii. Τον αριθμό των επαναλήψεων
 - iii. Τον αριθμό των clusters

3. Run

Βήματα μοντελοποίησης σε μπλοκ, στο Pajek

Όλα τα διαφορετικά partitions με την χαμηλότερη τιμή της συνάρτησης του κριτηρίου (ένα ή περισσότερα partitions).



<http://vlado.fmf.uni-lj.si/pub/networks/doc/KN/Course4.pdf>



Χρονικώς μεταβαλλόμενα δίκτυα (temporal)



Στα χρονικώς μεταβαλλόμενα (ή δυναμικά) δίκτυα, οι κόμβοι και οι γραμμές δεν είναι απαραίτητως παρούσες σε κάθε στιγμή.

Δημιουργία δυναμικού δικτύου με το Pajek

Στο αρχείο εισόδου που δημιουργούμε για το Pajek μπορούμε για ένα δυναμικό δίκτυο να χρησιμοποιούμε τις παρακάτω μορφές χρονικής υπόστασης:

- ❖ t_i – μια μόνο χρονική στιγμή
- ❖ t_i-t_j – ένα χρονικό διάστημα
- ❖ t_i-* – από μία χρονική στιγμή και μετά

Έστω δυναμικό δίκτυο με κορυφές a, b, c οι οποίες εμφανίζονται:

- a – Σε δύο χρονικά διαστήματα. Από το 5 δευτερόλεπτο έως το 10 και από το 12 έως το 14.
- b – Από το διάστημα 1 έως 3 και έπειτα μόνο στο 7
- c – Από το 4 και μετά

Και γραμμές μεταξύ των $a-b$, και των $a-c$ οι οποίες εμφανίζονται:

- $a-b$ – Μόνο τη χρονική στιγμή 7
- $a-c$ – Στο χρονικό διάστημα 6 έως 8

Αφού ανοίξουμε το αρχείο .net μέσω του Pajek, επιλέγουμε **Network > Temporal Network > Generate in Time**

<http://vlado.fmf.uni-lj.si/pub/networks/doc/seminar/SFI-SAS04tn.pdf>

Το Pajek υποστηρίζει και ένα σύνολο χρονικών συμβάντων για μεταβαλλόμενα (ή δυναμικά) δίκτυα πραγματικού χρόνου:

Event	Explanation
TI t	initial events – following events happen when time point t starts
TE t	end events – following events happen when time point t is finished
AV uns	add vertex v with label n and properties s
HV v	hide vertex v
SV v	show vertex v
DV v	delete vertex v
AA uvs	add arc (u,v) with properties s
HA uv	hide arc (u,v)
SA uv	show arc (u,v)
DA uv	delete arc (u,v)
AE uvs	add edge $(u:v)$ with properties s
HE uv	hide edge $(u:v)$
SE uv	show edge $(u:v)$
DE uv	delete edge $(u:v)$
CV vs	change vertex property – change property of vertex v to s
CA uvs	change arc property – change property of arc (u,v) to s
CE uvs	change edge property – change property of edge $(u:v)$ to s
CT uv	change type – change (un)directedness of line (u,v)
CD uv	change direction of arc (u,v)
PE uvs	replace pair of arcs (u,v) and (v,u) by single edge $(u:v)$ with properties s
AP uvs	add pair of arcs (u,v) and (v,u) with properties s
DP uv	delete pair of arcs (u,v) and (v,u)
EP uvs	replace edge $(u:v)$ by pair of arcs (u,v) and (v,u) with properties s

Το αρχείο δικτύου .net θα έπρεπε να συνταχθεί ως εξής:

```
*Vertices 3
1 "a" [5-10,12-14]
2 "b" [1-3,7]
3 "c" [4-*]
*Edges
1 2 1 [7]
1 3 1 [6-8]
```

Δύο κορυφές πρέπει να συνδέονται από μικρούς κύκλους συνδέσεων, όπου οι δύο γειτονικοί θα πρέπει να έχουν τουλάχιστον μια κοινή κορυφή.

Δημιουργία κύκλων στο Pajek

Δίκτυο μικρών κύκλων που οι δακτύλιοί τους να είναι τριγωνικοί (3-Rings):

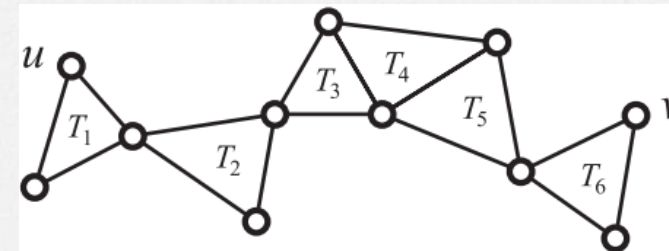
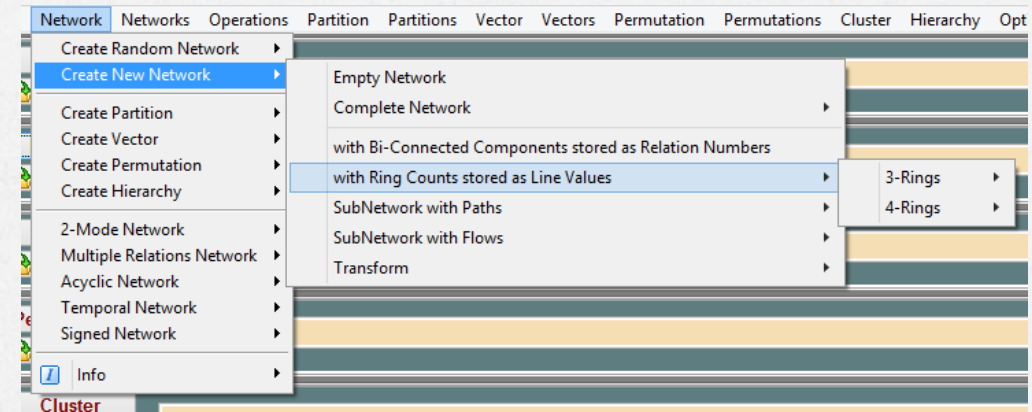
Network > Create New Network > with Ring Counts stored as Line Values > 3-Rings

Τετραγωνικοί (4-Rings):

Network > Create New Network > with Ring Counts stored as Line Values > 4-Rings

Οι δακτύλιοι μπορεί να είναι κατευθυνόμενοι ή μη κατευθυνόμενοι.

<http://arxiv.org/pdf/cs/0308011.pdf>





Εξαγωγή δεδομένων



Υπάρχει δυνατότητα εξαγωγής δεδομένων σε διάφορους τύπους αρχείων για την περαιτέρω ανάλυσή τους από άλλα λογισμικά.

Εξαγωγή στην R για στατιστική ανάλυση

Για να συνεργαστεί το Pajek με την R θα πρέπει να εντοπιστεί το εκτελέσιμο της R που δημιουργεί τα γραφικά:

Tools > R > Locate R

(εντοπισμός αρχείου C:\Program Files\...\bin\Rgui.exe ή Rterm.exe)

Αφού γίνει αυτό, τα δεδομένα μπορούν να σταλούν

Tools > R > Send to R.

Εξαγωγή στο SPSS για στατιστική ανάλυση

Εντοπίζουμε το αρχείο runsyntax.exe που βρίσκεται συνήθως στον κατάλογο C:\Program Files\SPSS:

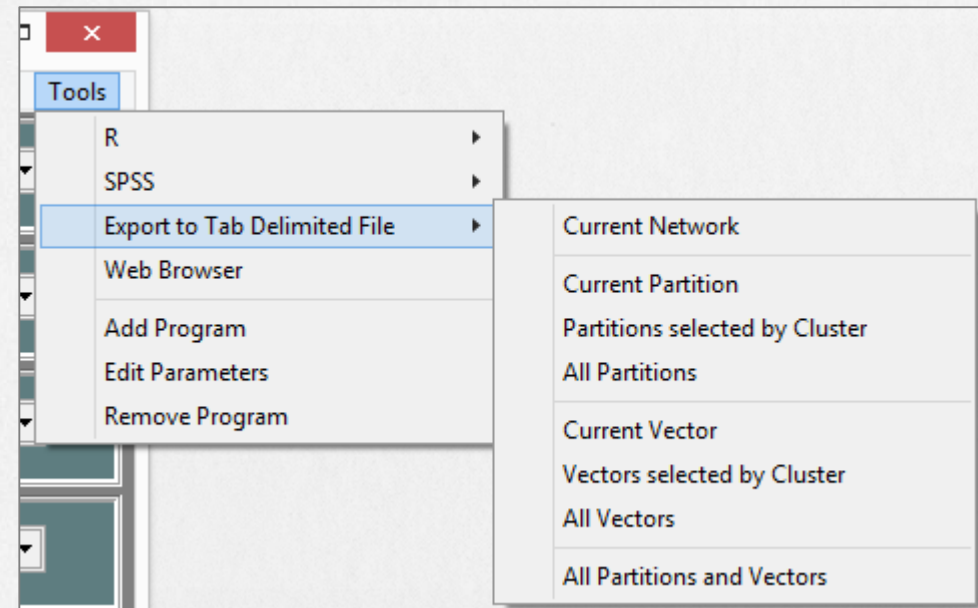
Tools > SPSS > Locate SPSS

Στη συνέχεια μπορούμε να στείλουμε δίκτυα, partitions και vectors, επιλέγοντας: **Tools > SPSS > Send to SPSS**

Εξαγωγή σε αρχείο χωρισμένο με tabs

Tools > Export to Tab Delimited File

Για άνοιγμα από στατιστικό πακέτο (πχ: Stata), ή από το Excel.





3ο μέρος



Σενάριο χρήσης με πραγματικά δεδομένα (1/7)



Το παγκόσμιο εμπόριο μετάλλου. Τα τόξα του δικτύου αντιπροσωπεύουν τις εισαγωγές μετάλλου που κάνει μία χώρα από μία άλλη.

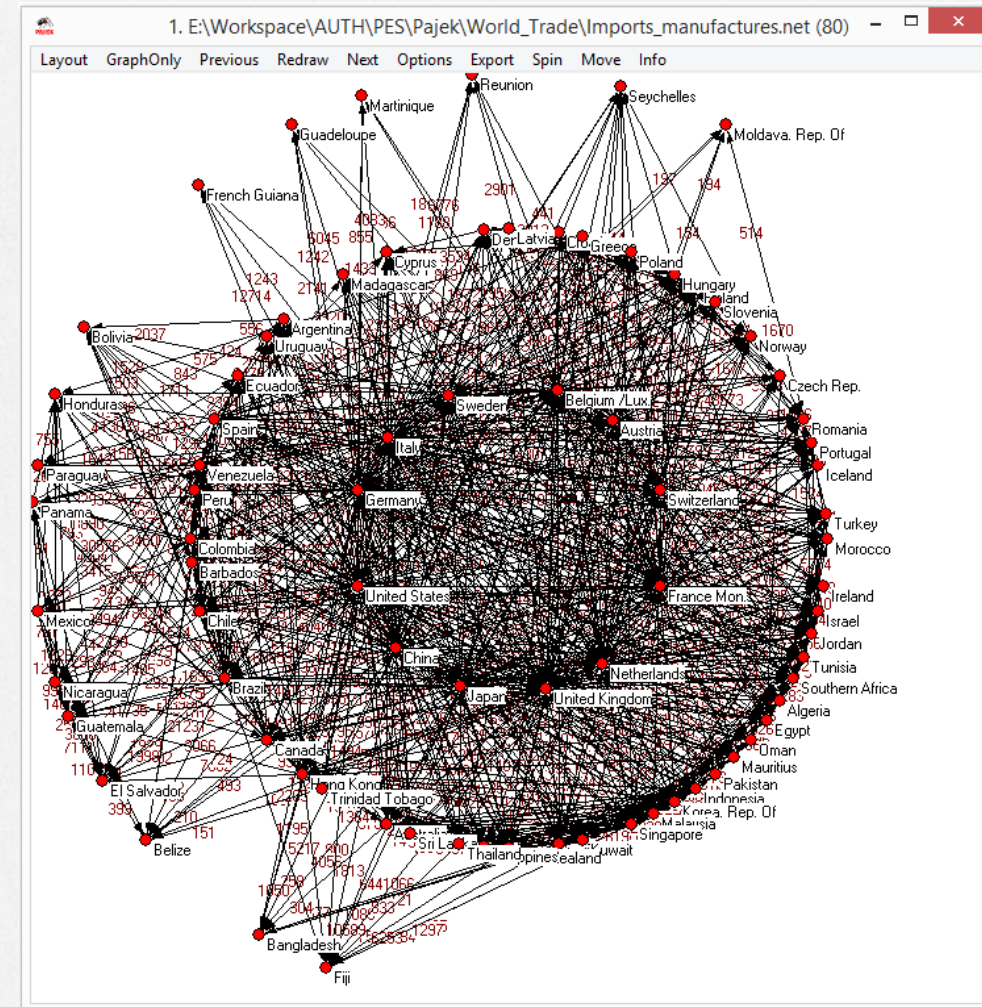
Αναλυτικά βήματα υπολογισμών - απεικόνιση

Αφού εισάγουμε το αρχείο του δικτύου στο Pajek (**File>Network>Read>Imports_manufactures.net**), επιλέγουμε **Draw>Network** και προκύπτει ο γράφος της διπλανής εικόνας:

Το μέγεθος των κόκκινων κουκίδων εισάγεται από εμάς μέσω του μενού, επιλέγοντας **Options>Size>Of Vertices**. Τα βάρη των συνδέσεων ενεργοποιούνται και απενεργοποιούνται μέσω του **Options>Lines>Mark Lines**.

Εύκολα παρατηρούμε από την εικόνα ότι σχηματίζονται τρεις νοητοί «ομόκεντροι» κύκλοι. Τρεις ομάδες χωρών. Η κεντρική ομάδα αποτελεί τον πυρήνα του εμπορίου του μετάλλου, και καθώς κινούμαστε προς τα έξω, συναντάμε τις χώρες που παίζουν δευτερεύων και τριτεύων ρόλο, αντιστοίχως. Αυτό είναι το πρώτο και το ευκολότερο συμπέρασμα που μπορούμε να βγάλουμε.

Είναι προφανές ότι ο αριθμός των γραμμών είναι αρκετά μεγάλος και δυσχεραίνει την περαιτέρω οπτική ανάλυση του δικτύου μας. Για το λόγο αυτό θα πρέπει να επιλεγεί κάποιο κριτήριο - εργαλείο απλοποίησης ή σήμανσης που θα μας επιτρέψει να βγάλουμε επιπλέον συμπεράσματα.

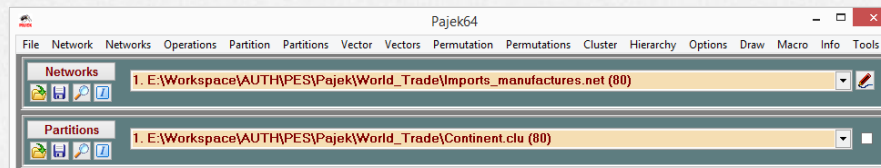


Σενάριο χρήσης με πραγματικά δεδομένα (2/7)

Το παγκόσμιο εμπόριο μετάλλου. Τα τόξα του δικτύου αντιπροσωπεύουν τις εισαγωγές μετάλλου που κάνει μία χώρα από μία άλλη.

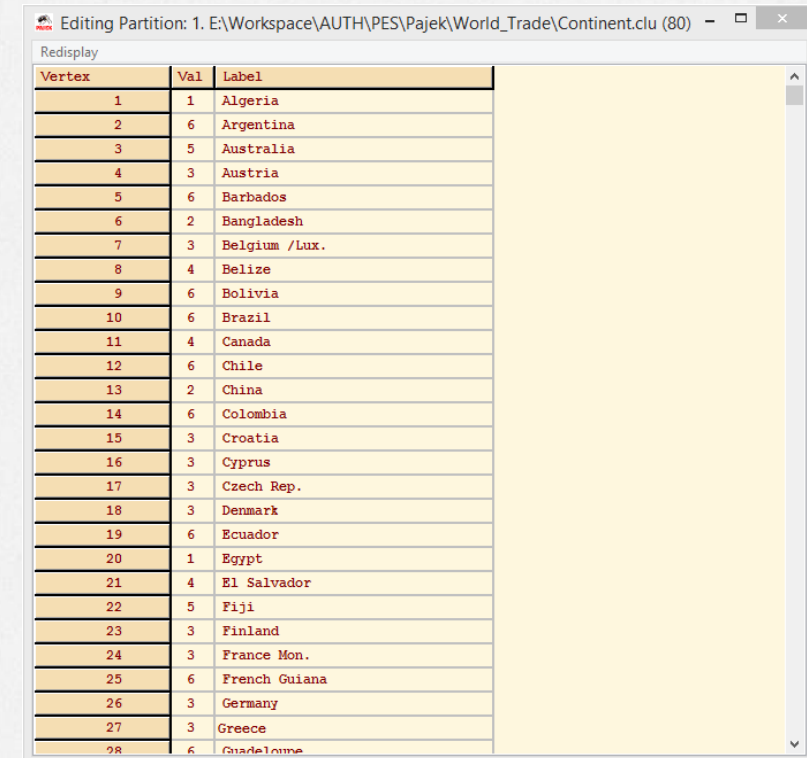
Ένας τρόπος για να επιτευχθεί κάτι τέτοιο, είναι η διαμέριση (partitioning) του δικτύου των χωρών σε ηπείρους. Αυτό επιτυγχάνεται μέσω της εισαγωγής του αντίστοιχου αρχείου (**File > Partition > Read > Continent.clu**).

Μετά την ολοκλήρωση της εισαγωγής, θα πρέπει να επιβεβαιώσουμε οπτικά ότι το αρχείο διαμέρισης .clu έχει τον ίδιο αριθμό κόμβων με το αρχείο του δικτύου .net (που στην περίπτωση μας είναι 80):



Κάνοντας κλικ στο εικονίδιο του μεγεθυντικού φακού (**View/Edit Partition**) κάτω από το κουμπί Partition, μπορούμε να δούμε την διαμέριση βάσει των ηπείρων.

Η στήλη Vertex συμβολίζει τις κορυφές του δικτύου και παίρνει τιμές 1 – 80 (αριθμός συμμετεχουσών χωρών). Η στήλη με το όνομα Val παίρνει τιμές 1 – 6 και συμβολίζει τις 6 από τις 7 ηπείρους, μιας και δεν υπάρχουν χώρες που να αντιπροσωπεύουν την έβδομη (πιθανότατα επειδή δεν συμμετέχουν στο δίκτυο εισαγωγών μετάλλου). Η Ελλάδα ανήκει στην Ευρώπη, επομένως εύκολα βγαίνει το συμπέρασμα ότι όποια χώρα (Label) έχει την τιμή (Val) 3, ανήκει στην Ευρώπη.



Vertex	Val	Label
1	1	Algeria
2	6	Argentina
3	5	Australia
4	3	Austria
5	6	Barbados
6	2	Bangladesh
7	3	Belgium /Lux.
8	4	Belize
9	6	Bolivia
10	6	Brazil
11	4	Canada
12	6	Chile
13	2	China
14	6	Colombia
15	3	Croatia
16	3	Cyprus
17	3	Czech Rep.
18	3	Denmark
19	6	Ecuador
20	1	Egypt
21	4	El Salvador
22	5	Fiji
23	3	Finland
24	3	France Mon.
25	6	French Guiana
26	3	Germany
27	3	Greece
28	6	Guadeloupe



Σενάριο χρήσης με πραγματικά δεδομένα (3/7)



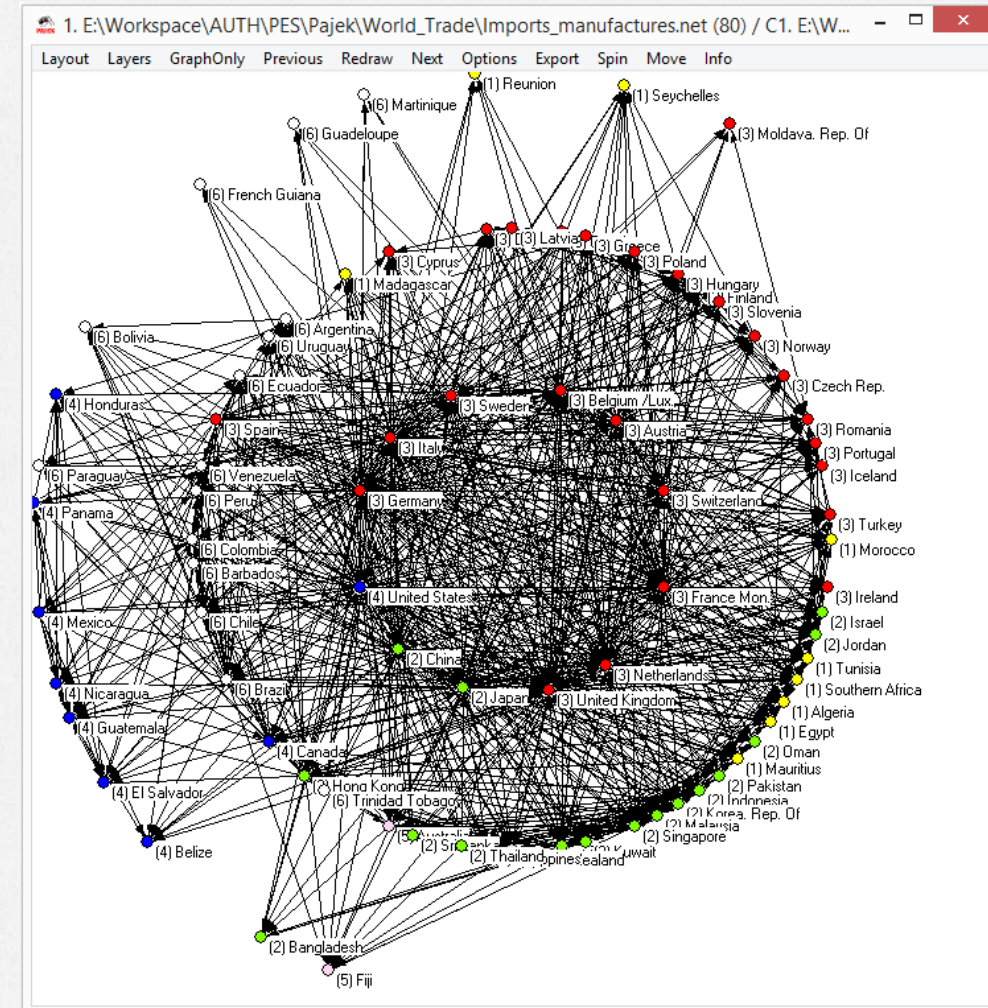
Το παγκόσμιο εμπόριο μετάλλου. Τα τόξα του δικτύου αντιπροσωπεύουν τις εισαγωγές μετάλλου που κάνει μία χώρα από μία άλλη.

Επιλέγοντας **Draw > Network + First Partition**, παίρνουμε μία λίγο πιο κατατοπιστική απεικόνιση του δικτύου, μιας και το Pajek εφαρμόζει αυτόματα διαφορετικό χρώμα για καθεμιά από τις 6 ηπείρους:

Η εμφάνιση του αριθμού της ηπείρου εφαρμόζεται επιλέγοντας **Options > Mark Vertices Using > Partition Clusters**.

Με μια σύντομη παρατήρηση συμπεραίνουμε ότι κάποιες χώρες της Ευρώπης (κόκκινο χρώμα) ανήκουν στον πυρήνα, ενώ κάποιες άλλες όχι. Οι χώρες εκτός πυρήνα είναι σίγουρα κατώτερες οικονομικά από τις χώρες εντός.

Παρ' όλα αυτά, το δίκτυό μας συνεχίζει να περιέχει αρκετή ακόμα πληροφορία που δυσχεραίνει την παρατήρησή του. Μία από τις πρώτες επεμβάσεις που μπορούν να γίνουν, είναι η συρρίκνωση του δικτύου. Αυτό μπορεί να γίνει αφού πρώτα θέσουμε ένα λογικό στόχο, όπως είναι για παράδειγμα η αναπαράσταση της συσχέτισης στο εμπόριο μετάλλου μεταξύ των 6 ηπείρων που συμμετέχουν.



Σενάριο χρήσης με πραγματικά δεδομένα (4/7)

Το παγκόσμιο εμπόριο μετάλλου. Τα τόξα του δικτύου αντιπροσωπεύουν τις εισαγωγές μετάλλου που κάνει μία χώρα από μία άλλη.

Επιλέγοντας από το κεντρικό μενού του Pajek, **Operations > Network + Partition > Shrink Network**, εμφανίζεται το παρακάτω παράθυρο διαλόγου, στο οποίο ζητείται να εισαχθεί ο ελάχιστος αριθμός συνδέσεων μεταξύ των clusters του partition και ο αριθμός του cluster που δεν επιθυμούμε να συρρικνωθεί:

Shrinking according to Partition

Minimum Number of Lines between Clusters

1

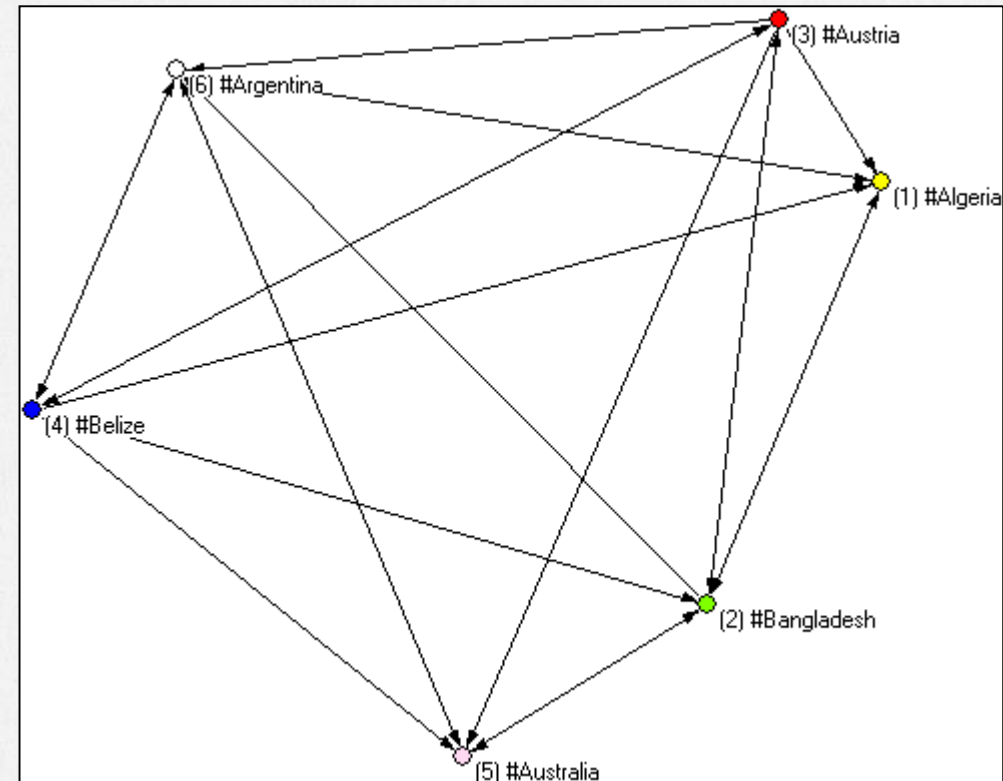
Cluster that will not be Shrunk

0

OK

Cancel

Ως αποτέλεσμα παίρνουμε κάτι πολύ πιο ευανάγνωστο αλλά με ορισμένα σημεία στα οποία απαιτείται από εμάς χειροκίνητη διόρθωση:



Σενάριο χρήσης με πραγματικά δεδομένα (5/7)

Το παγκόσμιο εμπόριο μετάλλου. Τα τόξα του δικτύου αντιπροσωπεύουν τις εισαγωγές μετάλλου που κάνει μία χώρα από μία άλλη.

Πρέπει οι κορυφές του νέου δικτύου να αποτελούνται μόνο από ονόματα ηπείρων. Επομένως θα πρέπει να γίνει ανάλογη επεξεργασία στη νέα αυτή διαμέριση που έχει προκύψει μετά τη συρρίκνωση:

Editing Partition: 2. Shrinking C1 [0] (6)

Redisplay

Vertex	Val	Label
1	1	#Algeria
2	6	#Argentina
3	5	#Australia
4	3	#Austria
5	2	#Bangladesh
6	4	#Belize

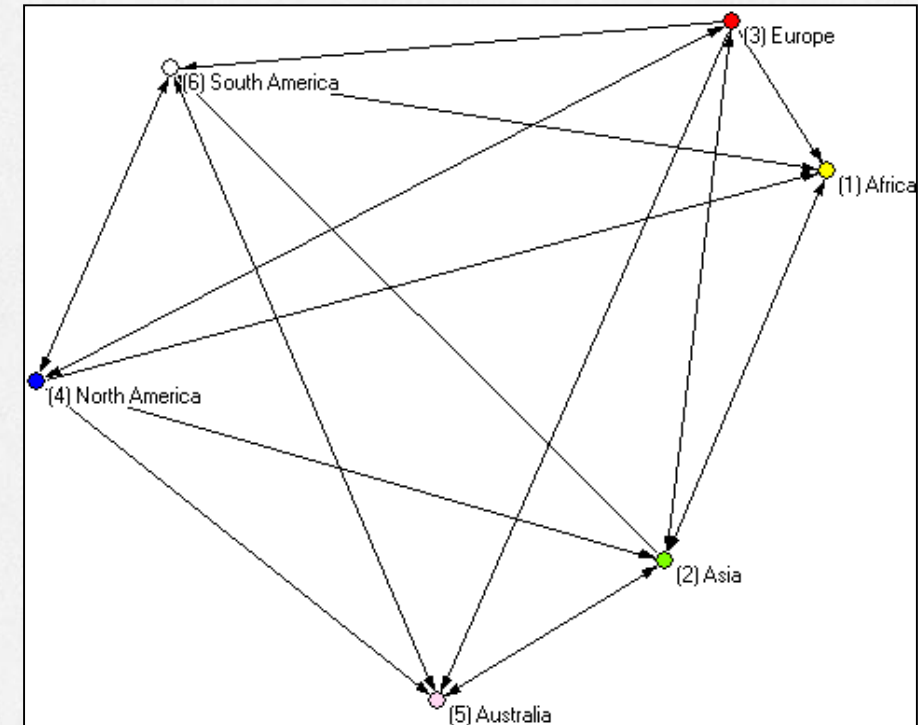
Η αλλαγή των ονομάτων γίνεται στα πεδία της στήλης Label:

Editing Partition: 2. Shrinking C1 [0] (6)

Redisplay

Vertex	Val	Label
1	1	Africa
2	6	South America
3	5	Australia
4	3	Europe
5	2	Asia
6	4	North America

Και το νέο δίκτυο αποτελείται πλέον, μόνο από τα ονόματα των 6 ηπείρων:



Αν και ο παραπάνω γράφος είναι αρκετά απλός, είναι λίγο δύσκολο κάποιος να μπορέσει να καταλάβει αν είναι πλήρης. Η χρήση ενός πίνακα γειτνίασης (adjacency matrix) μπορεί να προσφέρει μια καλύτερη εικόνα.



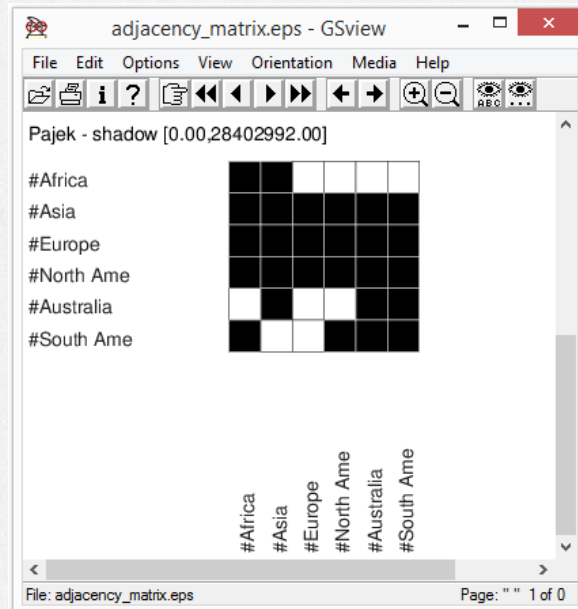
Σενάριο χρήσης με πραγματικά δεδομένα (6/7)



Το παγκόσμιο εμπόριο μετάλλου. Τα τόξα του δικτύου αντιπροσωπεύουν τις εισαγωγές μετάλλου που κάνει μία χώρα από μία άλλη.

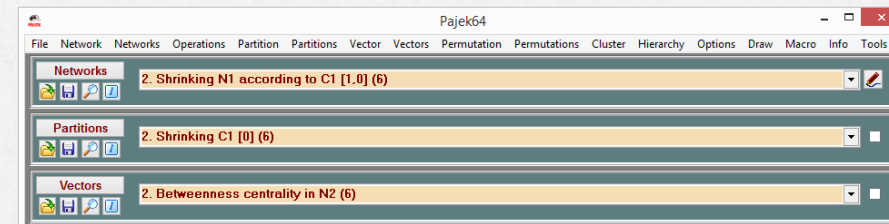
Μέσα από το κεντρικό μενού του Pajek επιλέγουμε **File > Network > Export as Matrix to EPS > Using Partition > Structural**. Στα παράθυρα διαλόγου που θα προκύψουν αφήνουμε τις προεπιλεγμένες τιμές και σώζουμε το αρχείο που περιλαμβάνει τον πίνακα γειτνίασης στο αρχείο **adjacency_matrix.eps**.

Στη συνέχεια, αν στο σύστημά μας υπάρχει εγκατεστημένο το απαιτούμενο λογισμικό, μπορούμε να προβάλουμε τον πίνακα γειτνίασης κάνοντας διπλό κλικ στο αντίστοιχο αρχείο:



Είναι προφανές με την πρώτη ματιά, ότι ο γράφος μας δεν είναι πλήρως συνδεδεμένος και ότι ο παραπάνω πίνακας δεν είναι συμμετρικός. Δηλαδή, το εμπόριο μετάλλου δεν πραγματοποιείται από, και προς, όλες τις ηπείρους. Για παράδειγμα, η Αυστραλία δε συναλλάσσεται με την Αφρική.

Σε αυτό το σημείο έχει επίσης νόημα, να υπολογίσουμε το μέγεθος της ενδιαμεσότητας (betweenness) δημιουργώντας ένα διάνυσμα μέσω του Pajek. Από το κεντρικό μενού επιλέγουμε **Network > Create Vector > Centrality > Betweenness**. Το νέο διάνυσμα εμφανίζεται στο αντίστοιχο πεδίο:



Κάνοντας κλικ στο εικονίδιο με τον μεγεθυντικό φακό που βρίσκεται κάτω από το κουμπί Vectors, εμφανίζεται ο πίνακας με τις τιμές που αφορούν την ενδιαμεσότητα της κάθε Ηπείρου:

Σενάριο χρήσης με πραγματικά δεδομένα (7/7)

Το παγκόσμιο εμπόριο μετάλλου. Τα τόξα του δικτύου αντιπροσωπεύουν τις εισαγωγές μετάλλου που κάνει μία χώρα από μία άλλη.

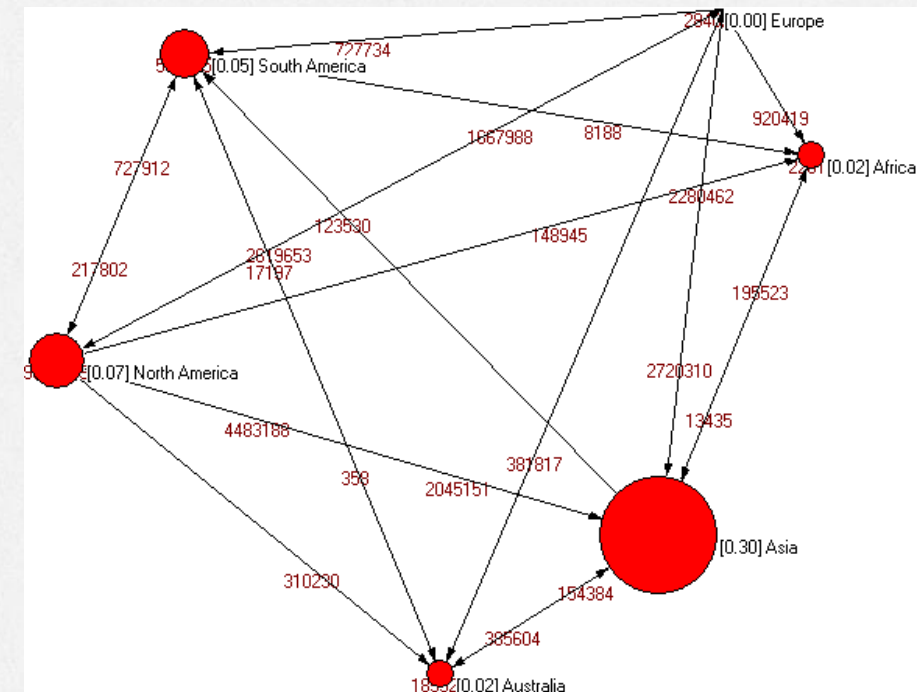
Editing Vector: 2. Betweenness centrality in N2 (6)

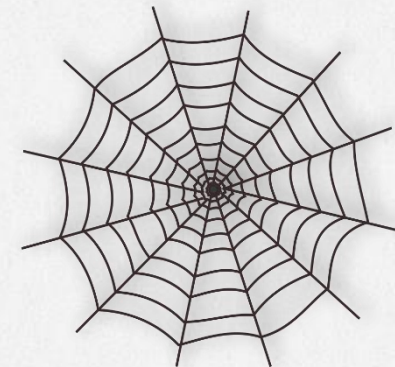
Redisplay

Vertex	Val	Label
1	0.016667	Africa
2	0.050000	South America
3	0.016667	Australia
4	0.000000	Europe
5	0.300000	Asia
6	0.066667	North America

Όπως ήταν ίσως αναμενόμενο, η ήπειρος της Ασίας παρουσιάζει τη μεγαλύτερη ενδιαμεσότητα. Αυτό σημαίνει ότι στο παγκόσμιο εμπόριο του μετάλλου (το έτος 1994) οι ασιατικές χώρες παίζουν κομβικό ρόλο και έχουν το μεγαλύτερο μερίδιο των συναλλαγών.

Αυτό είναι ακόμα πιο προφανές αν το διάνυσμα παρουσιαστεί γραφικά, επιλέγοντας **Draw > Network + First Vector**:





4ο μέρος



Βιοϋπολογισμός του συντομότερου μονοπατιού



Ο ευκαρυωτικός οργανισμός *Physarum polycephalum*, ανήκει στα πρώτιστα και μπορεί να «υπολογίσει» το συντομότερο μονοπάτι σε ένα γράφο.

Βήματα βιοϋπολογισμού

Ο συγκεκριμένος οργανισμός δε διαθέτει εγκέφαλο, ούτε κάποιου είδους νευρικό σύστημα. Όμως, παρ' όλα αυτά, είναι σε θέση να λαμβάνει αρκετά σύνθετες «αποφάσεις» προκειμένου να προσεγγίσει την τροφή του.

Παράδειγμα με λαβύρινθο:

- Αρχικά επεκτείνει τους δενδρίτες, του σε όλους τους διαδρόμους χαρτογραφώντας έτσι ολόκληρο το λαβύρινθο **(a)**.
- Αποσύρει τον κάθε βλαστό που δεν «εντόπισε» τροφή **(b)** αφήνοντας πίσω ένα είδος λάσπης που λειτουργεί σαν εξωτερική μνήμη η οποία υπενθυμίζει στο *Physarum* τα αδιέξοδα μονοπάτια.
- Αποφεύγοντας τα μονοπάτια αυτά, φύεται αποκλειστικά κατά μήκος της συντομότερης διαδρομής **(c)** από την αρχή του λαβυρίνθου ως την τροφή.

<http://arxiv.org/pdf/1106.0423v3.pdf>

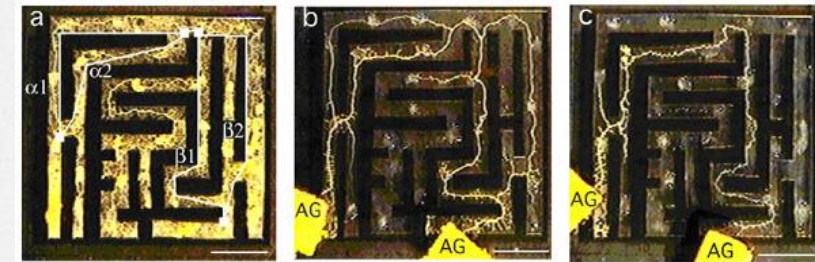
<https://vimeo.com/67686724>

<https://youtu.be/mvBSkt6LhJE>

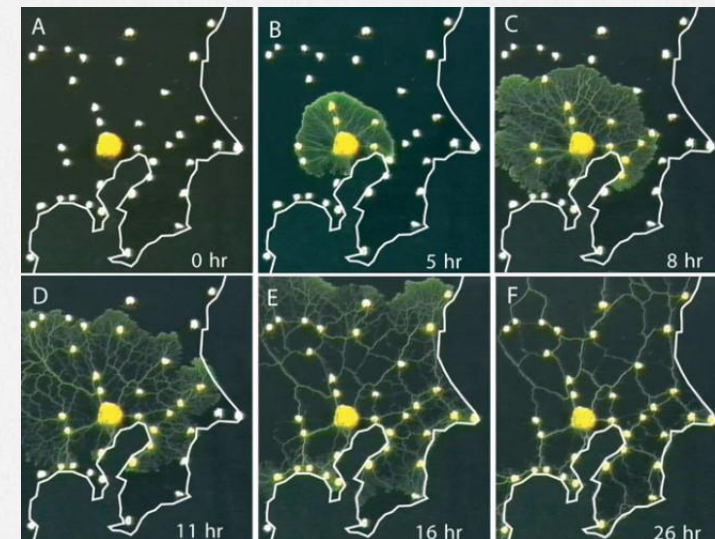
<https://youtu.be/BZUQQmcR5-g>

<https://youtu.be/11s27hu03yw>

Ο λαβύρινθος του παραδείγματος



Το σιδηροδρομικό δίκτυο του Τόκιο (χρόνος: 26ώρες)





Βιβλιογραφία



- [1] A. Mrvar και V. Batagelj, «Εγχειρίδιο χρήσης του Pajek,» 2 Απρίλιος 2015. [Ηλεκτρονικό]. Available: <http://mrvar.fdv.uni-lj.si/pajek/pajekman.pdf>.
- [2] T. Kamada και S. Kawai, «An algorithm for drawing general undirected graphs,» *Information Processing Letters*, 16 April 2003.
- [3] T. M. J. FRUCHTERMAN και E. M. REINGOLD, «Graph Drawing by Force-directed Placement,» *SOFTWARE-PRACTICE AND EXPERIENCE*, p. 36, November 1991.
- [4] S. Wasserman και K. Faust, «Part V. Dyadic and Triadic Methods,» σε *Social Network Analysis - Methods and Applications*, 1994.
- [5] V. Batagelj, A. Mrvar, A. Ferligoj και P. Doreian, «Generalized Blockmodeling with Pajek,» *Metodoloski zvezki*, pp. 455-467, 2004.
- [6] V. Batagelj και M. Zaversnik, «Short Cycles Connectivity,» 23 July 2013.
- [7] W. d. Nooy, A. Mrvar και V. Batagelj, «Exploratory Social Network Analysis with Pajek,» November 2010. [Ηλεκτρονικό]. Available: <http://vlado.fmf.uni-lj.si/pub/networks/book/esna2.htm>.

1. [Η εφαρμογή Pajek](#)
2. [Γενικά χαρακτηριστικά](#)
3. [Οι τύποι δεδομένων που υποστηρίζει το Pajek](#)
4. [Το γραφικό περιβάλλον εργασίας του Pajek](#)
5. [Δομή αρχείου .net](#)
6. [Παράδειγμα: Μαϊμού Wolfe \(2-mode network\)](#)
7. [Δημιουργία αρχείων δικτύου για το Pajek](#)
8. [Δημιουργία αρχείου δικτύου με το Pajek](#)
9. [Δημιουργία αρχείου δικτύου με κειμενογράφο](#)
10. [Δημιουργία αρχείου δικτύου με το Matlab](#)
11. [Δημιουργία αρχείου δικτύου με την Python](#)
12. [Δημιουργία αρχείου δικτύου με το Excel2Pajek](#)
13. [Δημιουργία αρχείου δικτύου μέσω σχεσιακής ΒΔ](#)
14. [Άλλοι τρόποι δημιουργίας αρχείου δικτύου](#)
15. [Καθορισμός διάταξης δικτύου](#)
16. [Μορφές αρχείων εξόδου του Pajek](#)
17. [Τοπικές και ευρείες προβολές δικτύου](#)
18. [Βήματα εξαγωγής υποδικτύων](#)
19. [Μείωση μεγέθους δικτύου](#)
20. [Εύρεση μονοπατιών \(paths\)](#)
21. [Ισχυρώς και Ασθενώς συνδεδεμένα στοιχεία](#)
22. [Αμφισυνδεδεμένα στοιχεία - κορυφές](#)
23. [Άκυκλα δίκτυα](#)
24. [Δείκτες κεντρικότητας \(centrality indicators\)](#)
25. [Τριαδική απογραφή \(triadic census\)](#)
26. [Πυρήνες \(cores\)](#)
27. [Κλίκες \(cliques\)](#)
28. [Αναζήτηση προτύπων \(fragments\)](#)
29. [Μέθοδοι εντοπισμού κοινοτήτων \(communities\)](#)
30. [Δίκτυα δύο τύπων \(2-mode\)](#)
31. [Απεικόνιση δικτύου με βάρη \(αξίες, κόστη κλπ.\)](#)
32. [Νησίδες \(islands\)](#)
33. [Ιεραρχικό clustering](#)
34. [Μοντελοποίηση σε μπλοκ \(Blockmodeling\)](#)
35. [Χρονικώς μεταβαλλόμενα δίκτυα \(temporal\)](#)
36. [Συνδεσιμότητα μικρών κύκλων](#)
37. [Εξαγωγή δεδομένων](#)
38. [Σενάριο χρήσης με πραγματικά δεδομένα \(1/7\)](#)
39. [Βιοϋπολογισμός του συντομότερου μονοπατιού](#)
40. [Βιβλιογραφία](#)