

HO-450 Verão 2013: Pajek

Material baseado nos capítulos 2, 3 e 6 do livro ESNA, 2011. Nooy, Mrvar & Batagelj, Ed. Cambridge, como material auxiliar na disciplina do curso de verão HO-450, ministrada em 2013 no Instituto de Economia da Universidade Estadual de Campinas (Unicamp).

Trata-se de um resumo dos principais comandos apresentados nos capítulos indicados para utilização do Pajek 3.10. O material não visa nenhum fim lucrativo, tendo fins exclusivos de ensino e pesquisa, e respeitando integralmente os direitos dos autores do livro.

Alterações realizadas na última versão do programa encontram-se detalhadas em <http://mrvar.fdv.unicamp.br/pajek/pajekman.pdf>

Partições

Uma partição ou clustering consiste em classificar os nós de uma rede de acordo com algum critério previamente definido.

1. Carregar rede `imports_manufactures.net` via `File/Network/Read`

Analisar arquivo (abrir via txt): Rede direcionada composta por 80 países com arcos representando relações de comércio internacional (importação). O valor dos arcos representa o valor de importação em milhares de USD.



2. Carregar partição `World_system.clu` : `File/Partition/Read`

- a. Para ver a classificação dos vértices clicar em `Draw/Network+First Partition`

Obs.: para salvar todos os arquivos em um único projeto: `File/Pajek Project File/Save`

- b. Para editar a partição: `File/Partition/Edit`

- c. Para analisar a distribuição dos clusters: `Partition/Info` -> 1,+ 5 (os cinco com valores mais altos)

```
1. D:\iluna 2013\aulas-pg\esna2\data\Chapter2\World_system.clu (80)
=====
Dimension: 80
The lowest value: 1
The highest value: 3

Frequency distribution of cluster values:
-----
```

| Cluster | Freq | Freq% | CumFreq | CumFreq% | Representative |
|---------|------|----------|---------|----------|----------------|
| 1 | 12 | 15.0000 | 12 | 15.0000 | Austria |
| 2 | 51 | 63.7500 | 63 | 78.7500 | Algeria |
| 3 | 17 | 21.2500 | 80 | 100.0000 | Bangladesh |
| Sum | 80 | 100.0000 | | | |

```
-----
```

- d. Para reordenar o sociograma

- i. `Layout/Energy/Kamada-Kawai/Free`

- ii. `Move/Circles` -> 3, 80 (valor dos argumentos solicitados)

- iii. Ajuste manual

- iv. `Options>Mark Vertices Using/Partition Clusters`

- v. Para criar uma partição: `Partition/Create Constant Partition`

- vi. Para eliminar uma partição: `File/Partition/Dispose` (atenção com a escolha da partição no drop-box).

- vii. Para deslocar todos os nós de um cluster: clicar o mouse próximo a um nó do cluster e deslocar.
- viii. Para separar os clusters, alguns comandos úteis:
 1. `Layout>Circular>Using Partition`
 2. `Options/Transform/Fit Area/max(x),max(y),max(z)`
 3. `Layout/Energy/Kamada-Kawai/Optimize inside Clusters only`
 4. `Options/Transform/Resize Clusters área`

Obs. :

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

Visualização/Manipulação da rede (Shrinking)

Três tipos de manipulação:

1. Extrair uma das partes (Local View) – subrede/subgrafo
2. Transformar um cluster em um único vértice (Global View)
3. Selecionar uma parte e reduzir as classes vizinhas para se concentrar na estrutura interna e posição geral desta classe (Contextual View)

1. **Visão local:** o objetivo é analisar uma classe em particular

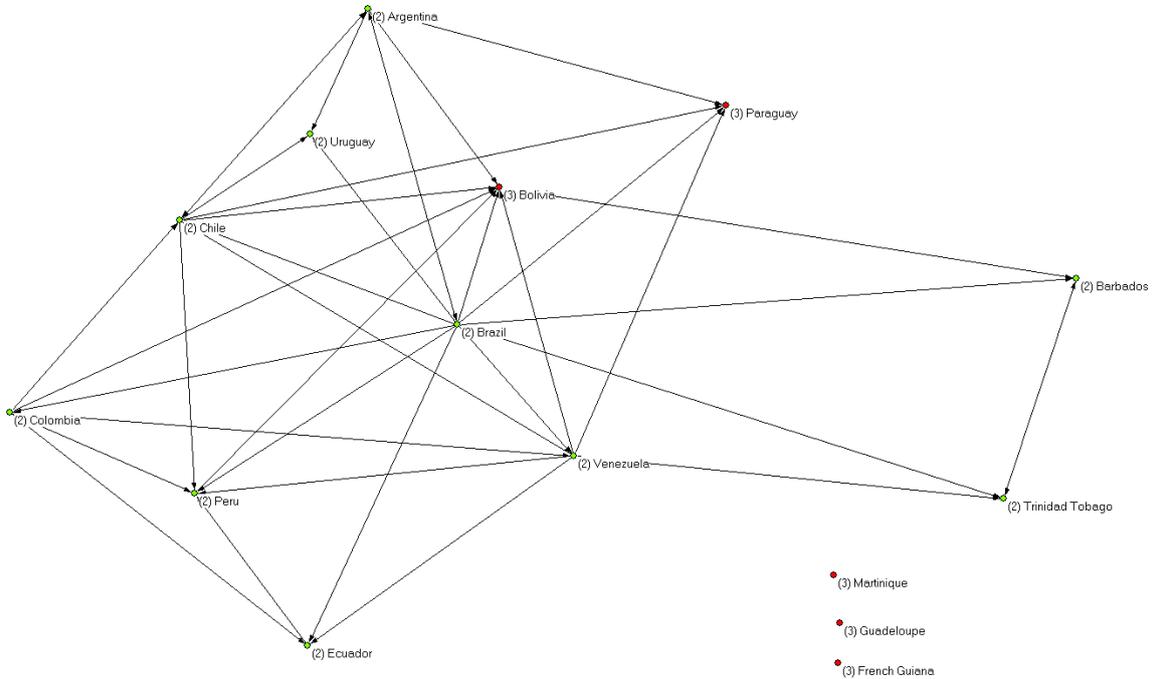
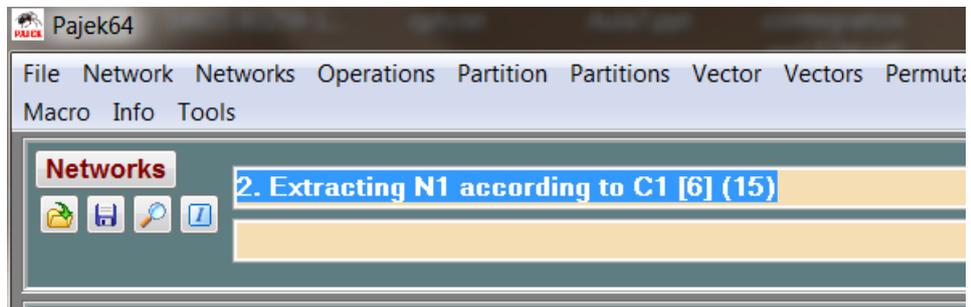
- a. Para extrair “América do Sul” (classe 6 – ver partição): precisamos de uma rede e de uma partição que indique os vértices a extrair – ver menu Operations.
 - i. `Operations/Network+Partition/Extract SubNetwork ->6`
 - ii. A subrede energizada considera apenas relações geográficas (continente como atributo em comum). Para considerar a estrutura original das relações comerciais (centro, semiperiferia, periferia que são classes de “World_system.clu”), precisamos trabalhar com duas partições (ver tela) pelo comando

`Partitions/Extract SubPartition (Second from First) ->6`



iii. `Draw/Network+First Partition`

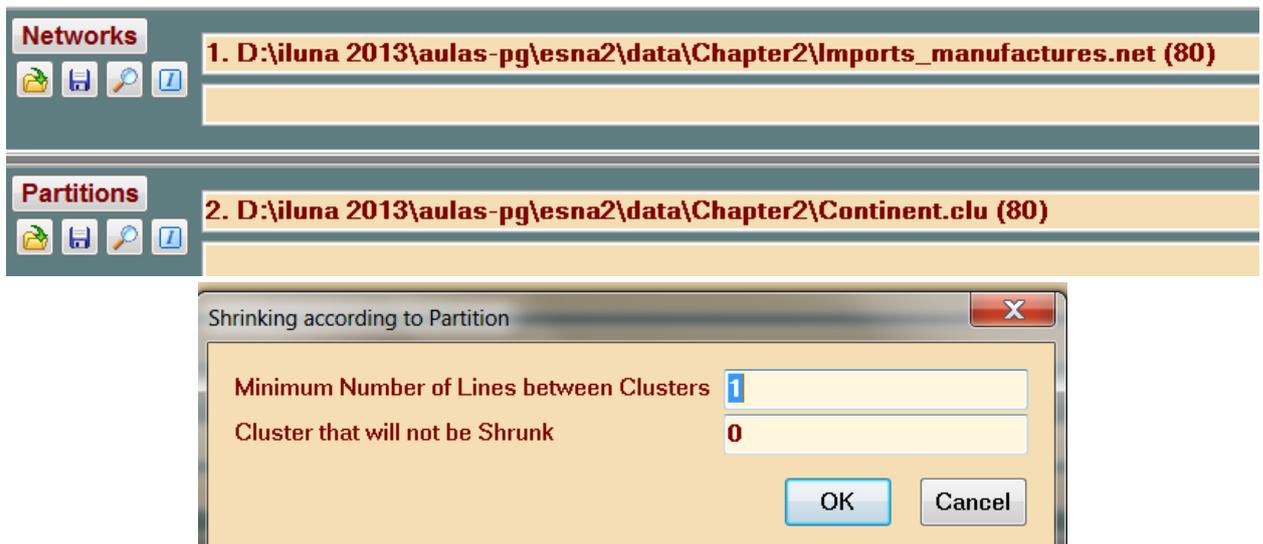
Atenção: a rede no drop-box Network deve ser a resultante do PRIMEIRO processo de redução.



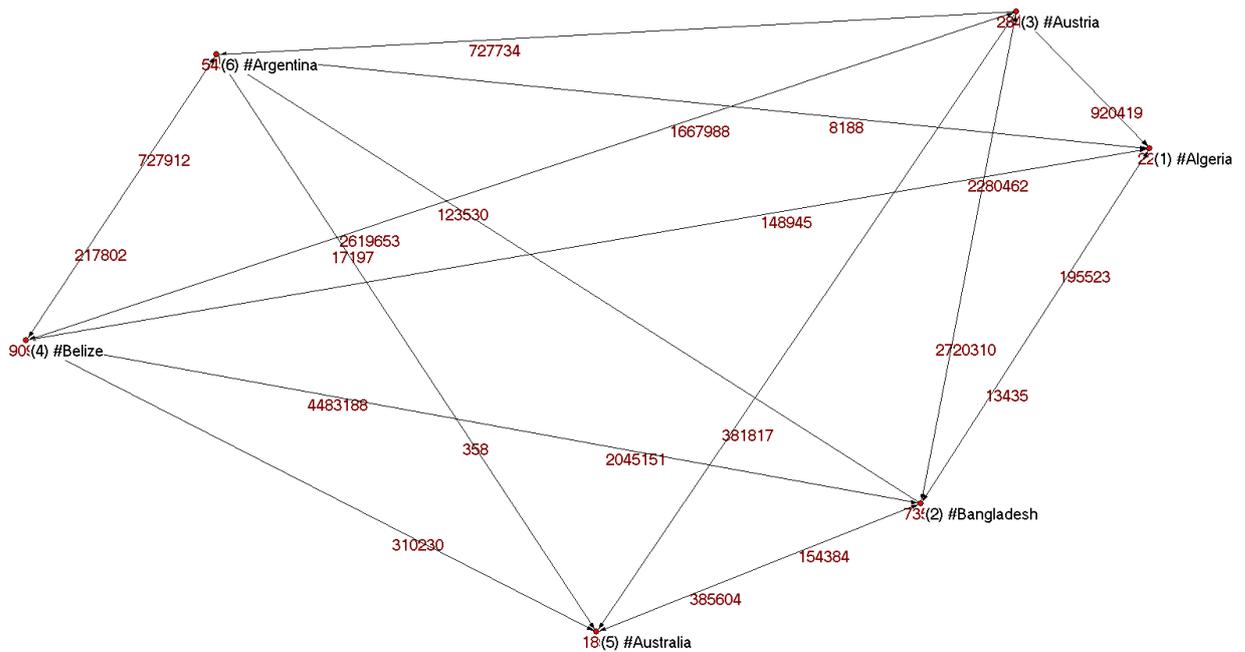
2. Visão Global:

O objetivo é analisar relações entre classes, ou seja, entre continentes. Para reduzir a rede:

- Antes, a visualização dos valores agregados de importação entre continentes será visível somente se a opção “Mark lines with values”. Assim, antes de mais nada, na tela de Draw, selecionar **Options/Lines/Mark Lines/With Values**.
- Na tela principal: selecionar a rede “Imports_manufactures.net” e a partição “Continent.clu”. Logo ir para **Operations/Network+Partition/Shrink Network**



Cada continente será representado por um nó, onde o número entre ()s representa o índice do continente.

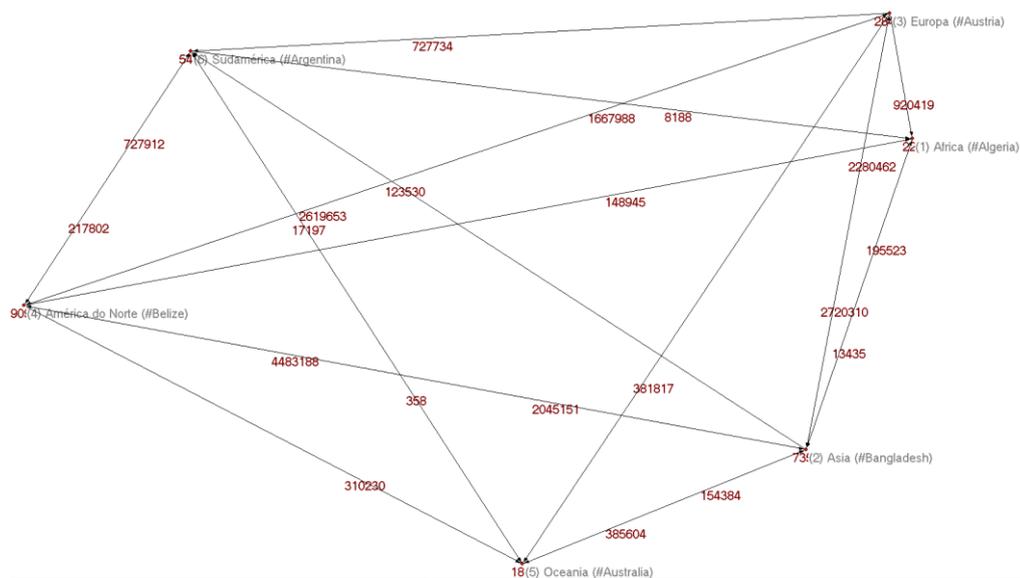


c. Para editar os nomes das classes: **File/Partition/Edit**

Editing Partition: 6. Shrinking C3 [0] (6)

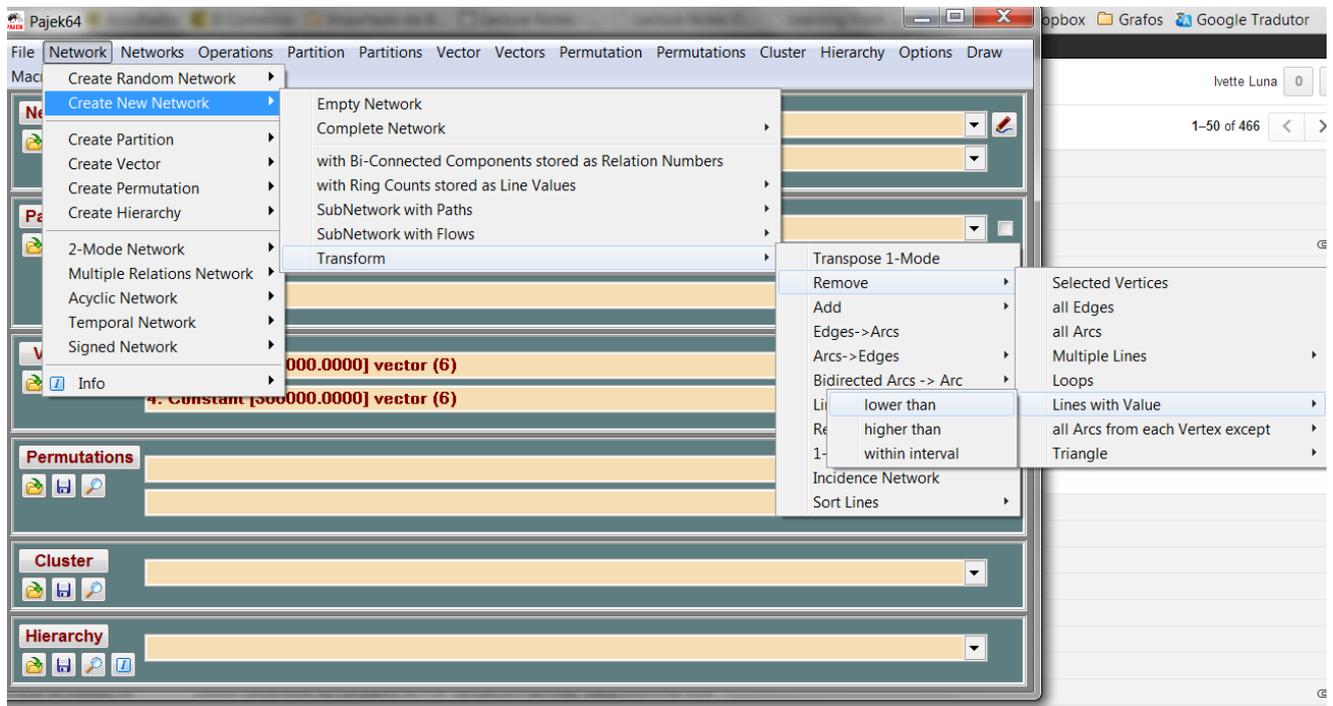
Redisplay

| Vertex | Val | Label |
|--------|-----|----------------------------|
| 1 | 1 | Africa (#Algeria) |
| 2 | 2 | Sudamérica (#Argentina) |
| 3 | 3 | Oceania (#Australia) |
| 4 | 4 | Europa (#Austria) |
| 5 | 5 | Asia (#Bangladesh) |
| 6 | 6 | América do Norte (#Belize) |

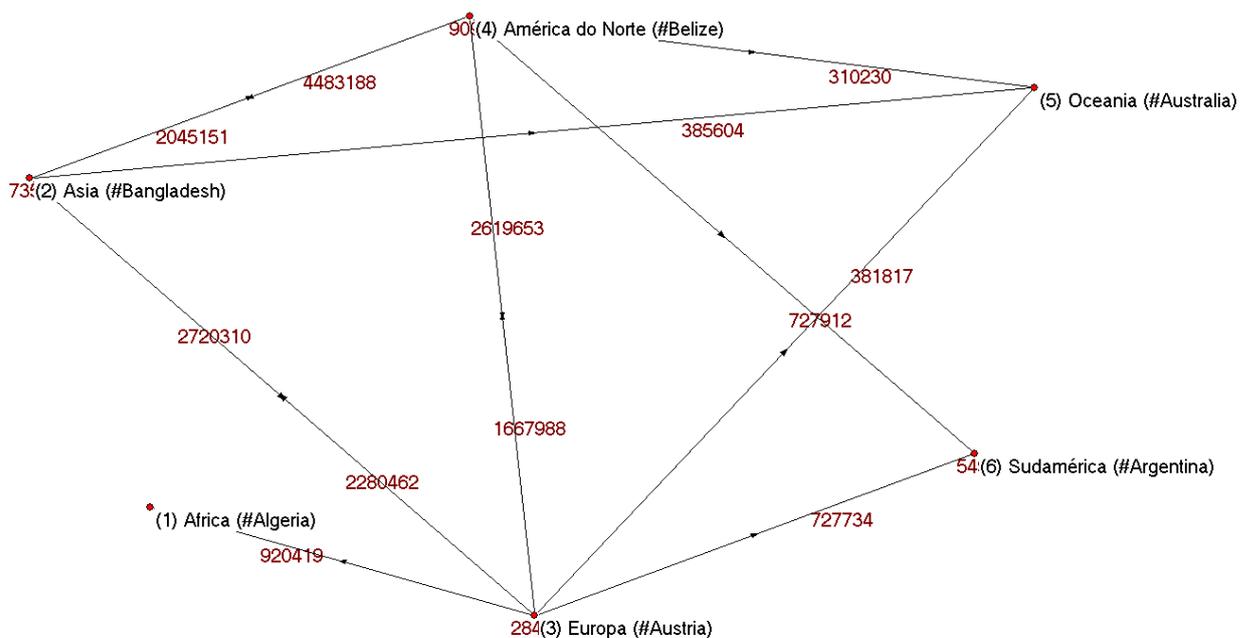


Os loops não são visualizados, mas os seus valores encontram-se do lado do respectivo vértice.

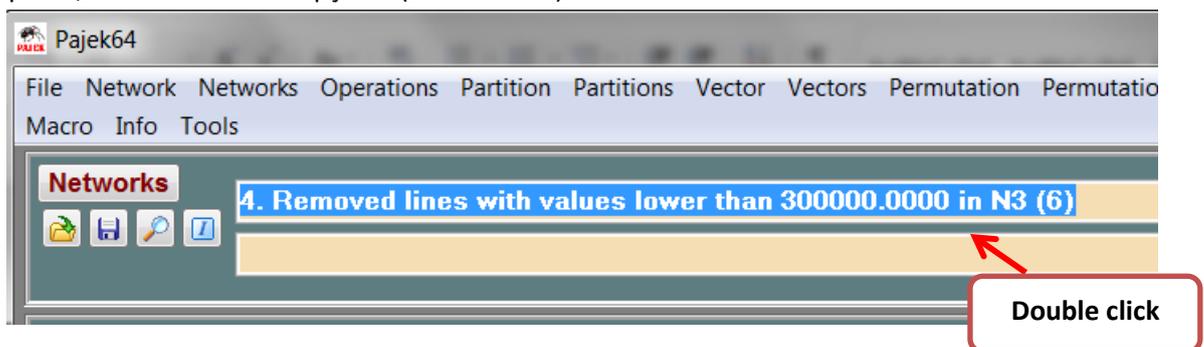
d. Para podar as arestas com valor inferior a um determinado limiar, ir para **Network/Create New Network/Transform/Remove/Lines With Value/lower than**



Obtendo assim, uma nova rede com arestas de valores superiores ao valor especificado (300 000):



Uma opção para analisar as relações resultantes é via a matriz de adjacência, dando dois clicks sobre a rede no drop-box, e selecionando a opção 2 (Lines Values).



Viewing Network as a Matrix --- 4. Removed lines with values lower than 300000.0000 in N3 (6)

| File | 1 | 2 | 3 | 4 | 5 | 6 | Label |
|------|--------|--------|--------|----------|---------|---------|-------------|
| 1. | 0 | 0 | 0 | 0 | 0 | 0 | #Algeria |
| 2. | 0 | 549445 | 0 | 0 | 0 | 0 | #Argentina |
| 3. | 0 | 0 | 0 | 0 | 0 | 0 | #Australia |
| 4. | 920419 | 727734 | 381817 | 28402992 | 2720310 | 2619653 | #Austria |
| 5. | 0 | 0 | 385604 | 2280462 | 7355707 | 4483188 | #Bangladesh |
| 6. | 0 | 727912 | 310230 | 1667988 | 2045151 | 9095279 | #Belize |

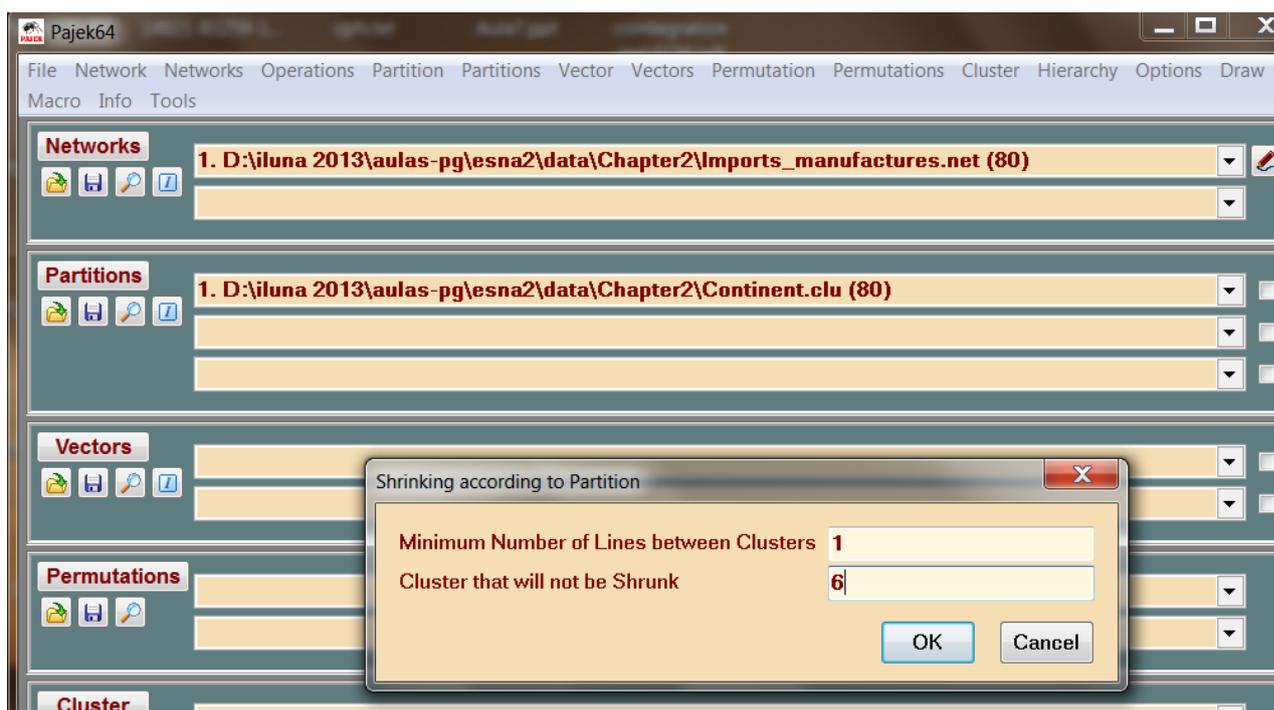
3. Visão Contextual

A visão local nos permitiu isolar um continente e detectar os vértices centrais na região (Brasil). A visão global nos permitiu analisar as relações entre continentes, destacando a posição econômica de cada continente a nível mundial. A visão contextual nos permitirá analisar uma classe, porém considerando também as relações (agregadas) com o mundo afora: trata-se de uma visão mista.

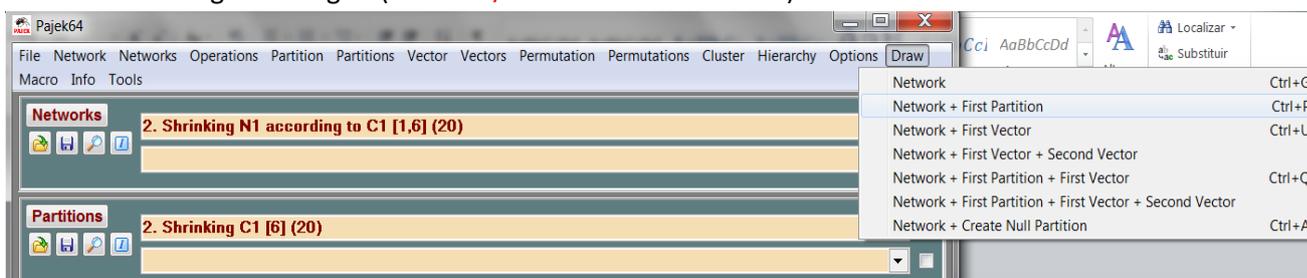
Nesta visão, todas as classes são “compactadas”, exceto aquela que se deseja analisar. Usando novamente o comando

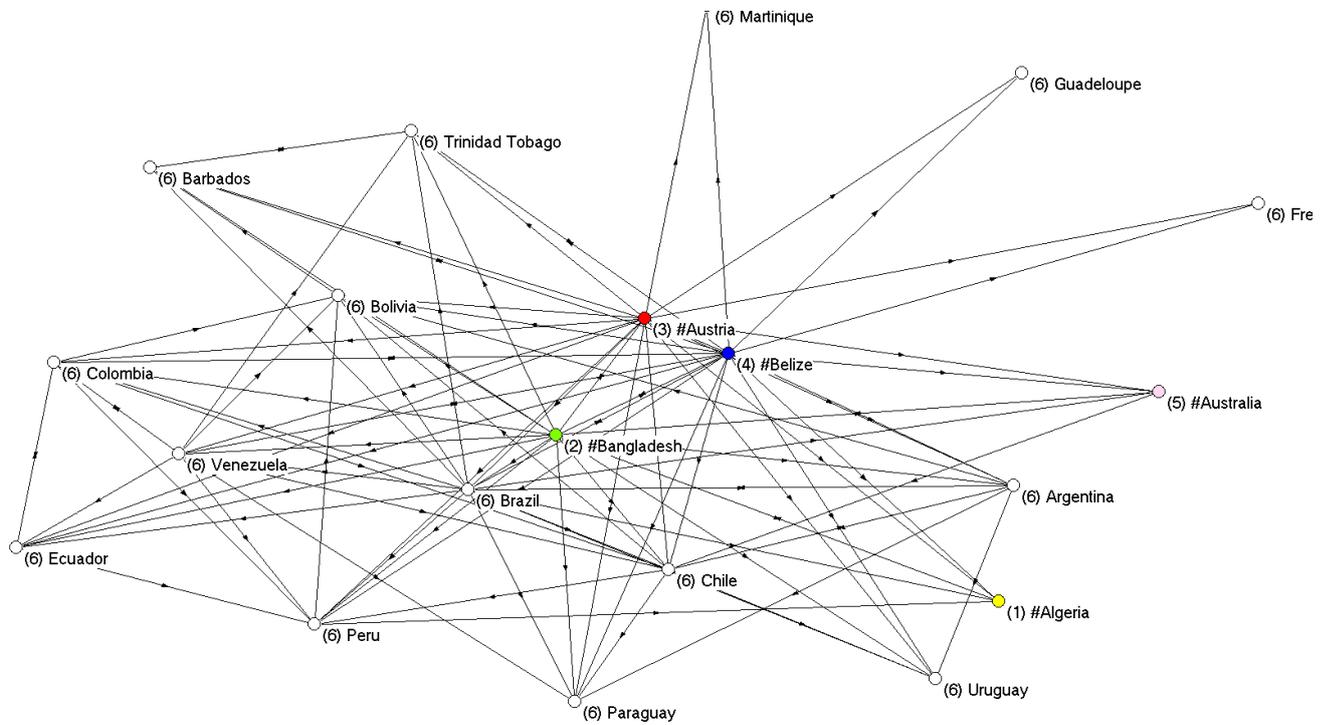
Operations/Network+Partition/Shrink Network

Optamos desta vez por não compactar o cluster 6 (América do Sul), como mostra a figura:



obtendo assim o grafo a seguir (via **Draw/Network+First Partition**)



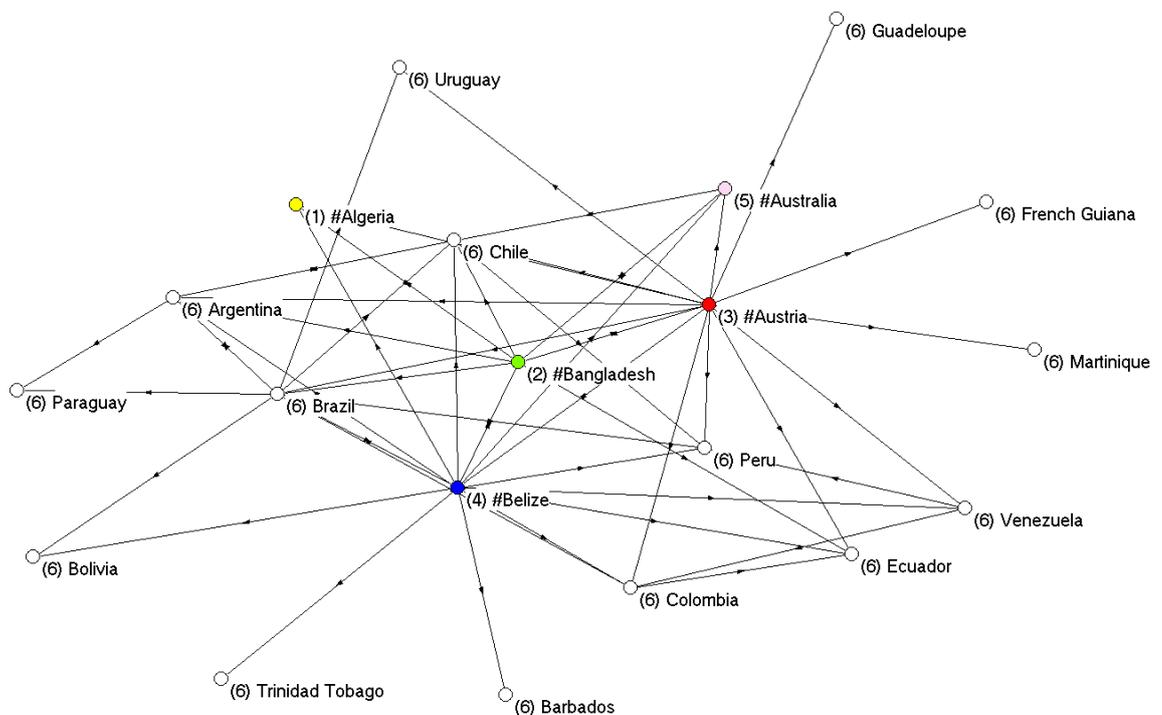


Onde os continentes vizinhos estão plotados em cores diferentes, enquanto os nós pertencentes ao cluster 6 (América do Sul) aparecem sem cor.

Para remover as arestas com valores inferiores a 10 000, clicar novamente em

Network/Create New Network/Transform/Remove/Lines With Value/lower than

Obtendo uma nova rede a partir da qual pode-se inferir a respeito das relações dos países em América do Sul e as relações com os outros continentes.



Vetores

Vetores são arranjos contendo informação numérica referente a uma propriedade e associada a cada vértice da rede selecionada. No drop-box “Vector”:

1. Carregar **vetor GDP_1995.vec** (PIB/cápita) via **File/Vector**
2. Clicar em **Vector/Info** e indicar o número de vértices com o maior (+) ou menor(-) valor e o número de partições entre o valor máximo e mínimo observado (Exemplo: #5). O relatório mostrará o seguinte resultado:

```

=====
1. D:\iluna 2013\aulas-pg\esna2\data\Chapter2\GDP_1995.vec (80)
=====
Dimension: 80
The lowest value:          115.0000
The highest value:         43034.0000

Highest values:
-----
Rank  Vertex      Value      Id
-----
1     72      43034.0000  Switzerland
2     39      40920.0000  Japan
3     53      33708.0000  Norway
4     18      33121.0000  Denmark
5     26      29419.0000  Germany
-----
Sum (all values):          819995.0000

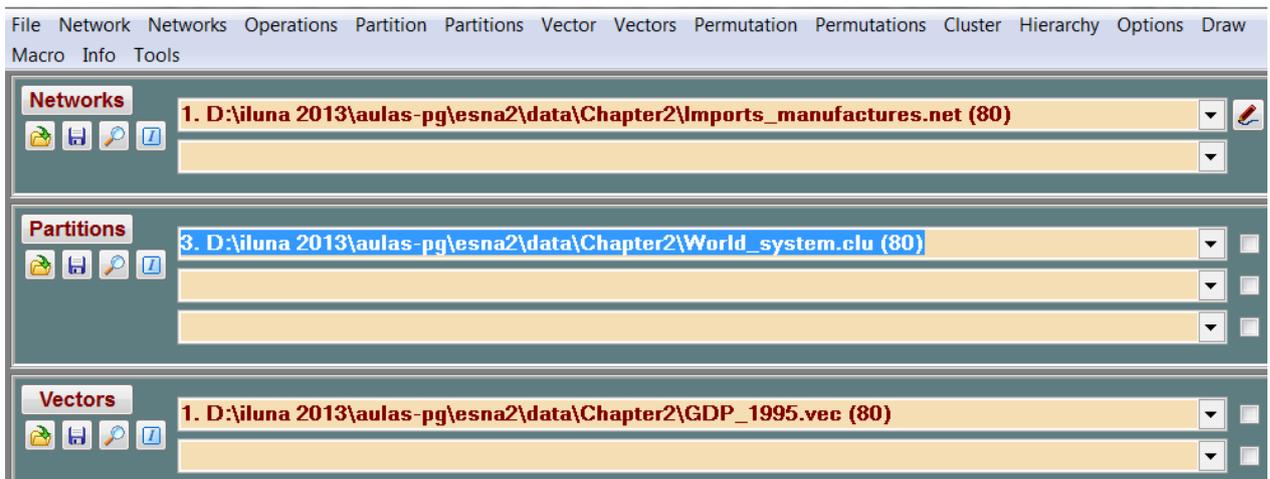
Arithmetic mean:          10249.9375
Median:                    4545.0000
Standard deviation:        10834.7491
2.5% Quantile:             327.5750
5.0% Quantile:             424.1500
95.0% Quantile:           29604.1000
97.5% Quantile:           33888.3000

-----
Vector Values              Frequency      Freq%      CumFreq      CumFreq%
-----
( 115.000 ... 115.000]      1      1.2500      1      1.2500
( 115.000 ... 10844.750]    50     62.5000     51     63.7500
( 10844.750 ... 21574.500]   14     17.5000     65     81.2500
( 21574.500 ... 32304.250]   11     13.7500     76     95.0000
( 32304.250 ... 43034.000]    4      5.0000     80    100.0000
-----
Total                      80     100.0000

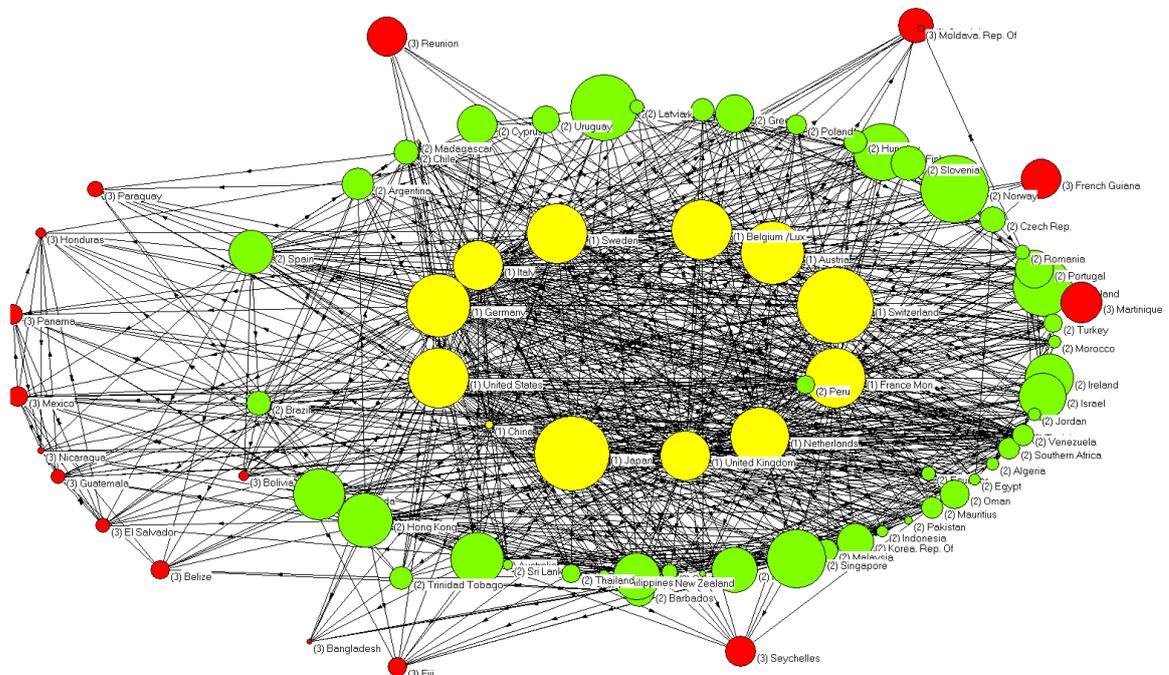
```

3. Considera a partição “World_system.net” + vetor “GDP_1995.vec” no desenho da rede e tamanho dos vértices.

Draw/Network+First Partition+First Vector

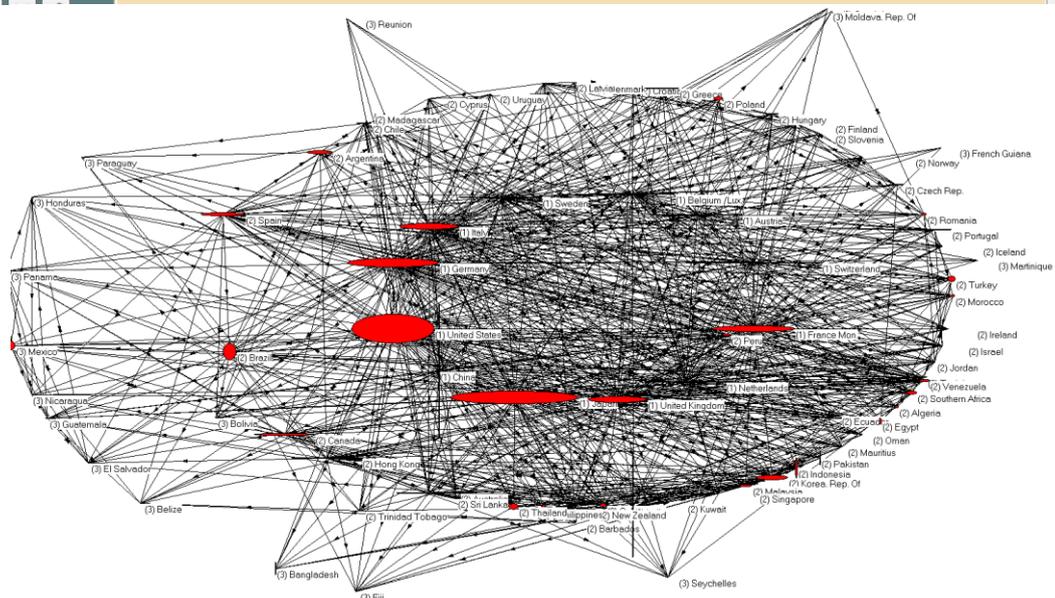
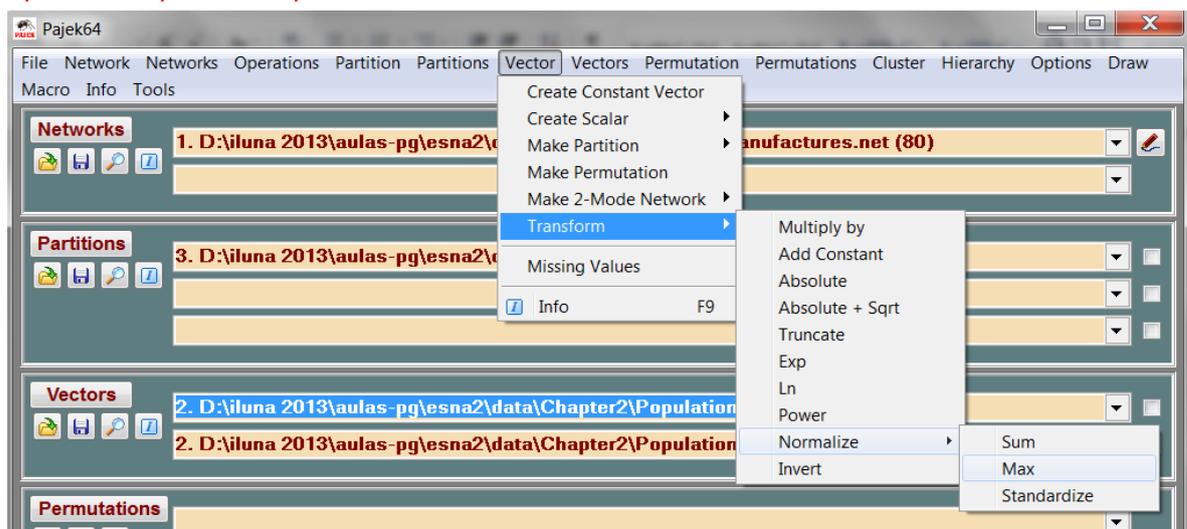


Para alterar o tamanho dos vértices clicar em **Options/Size/Of vértices**.



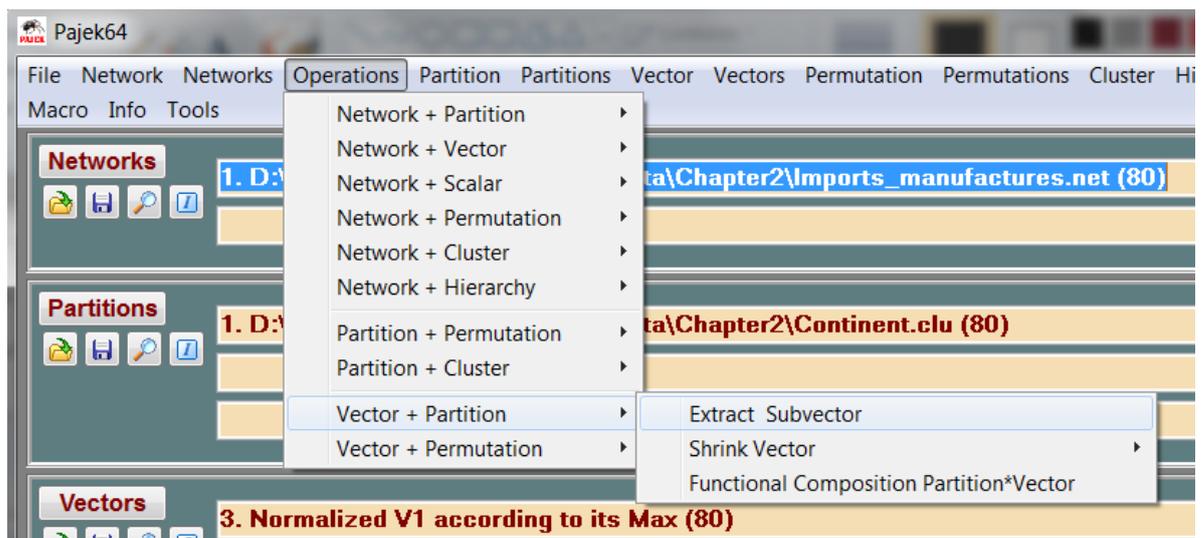
Ao considerar dois vetores, os vértices serão representados por elipses, onde o diâmetro horizontal será associado ao primeiro vetor, e o diâmetro vertical será associado ao segundo vetor. Antes, **recomenda-se normalizar os vetores ao intervalo unitário via (um de cada vez).**

Vector/Transform/Normalize/Max

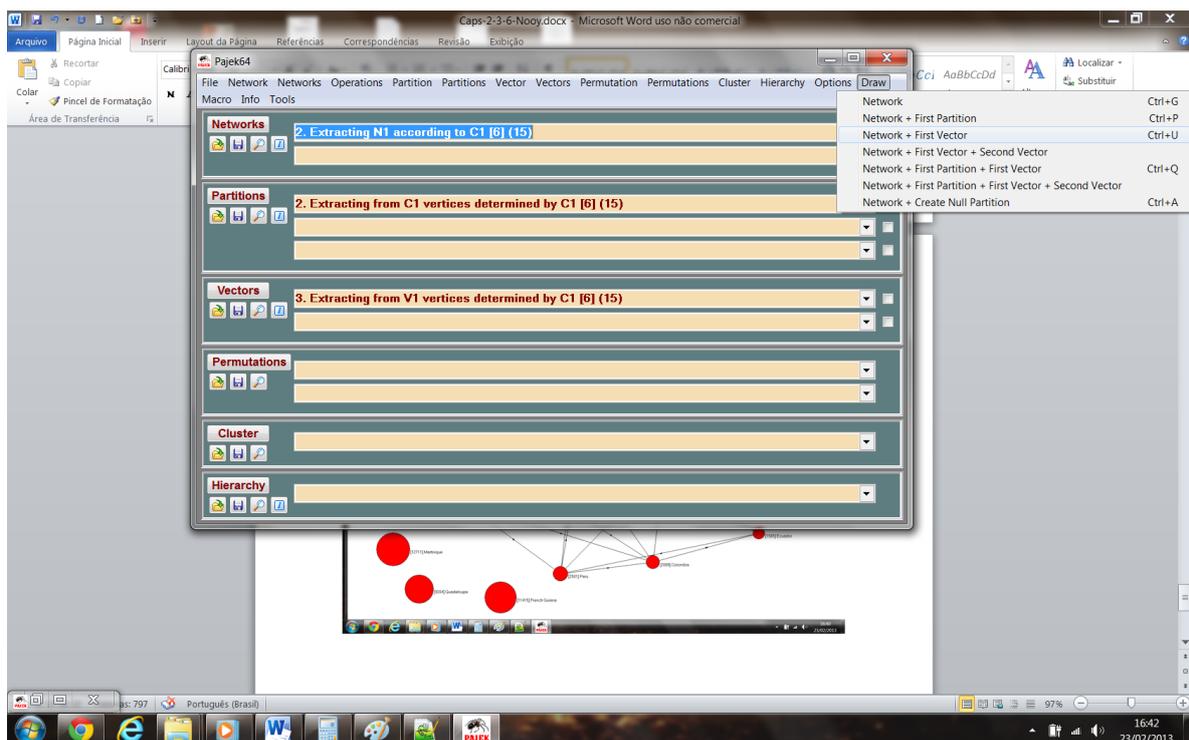


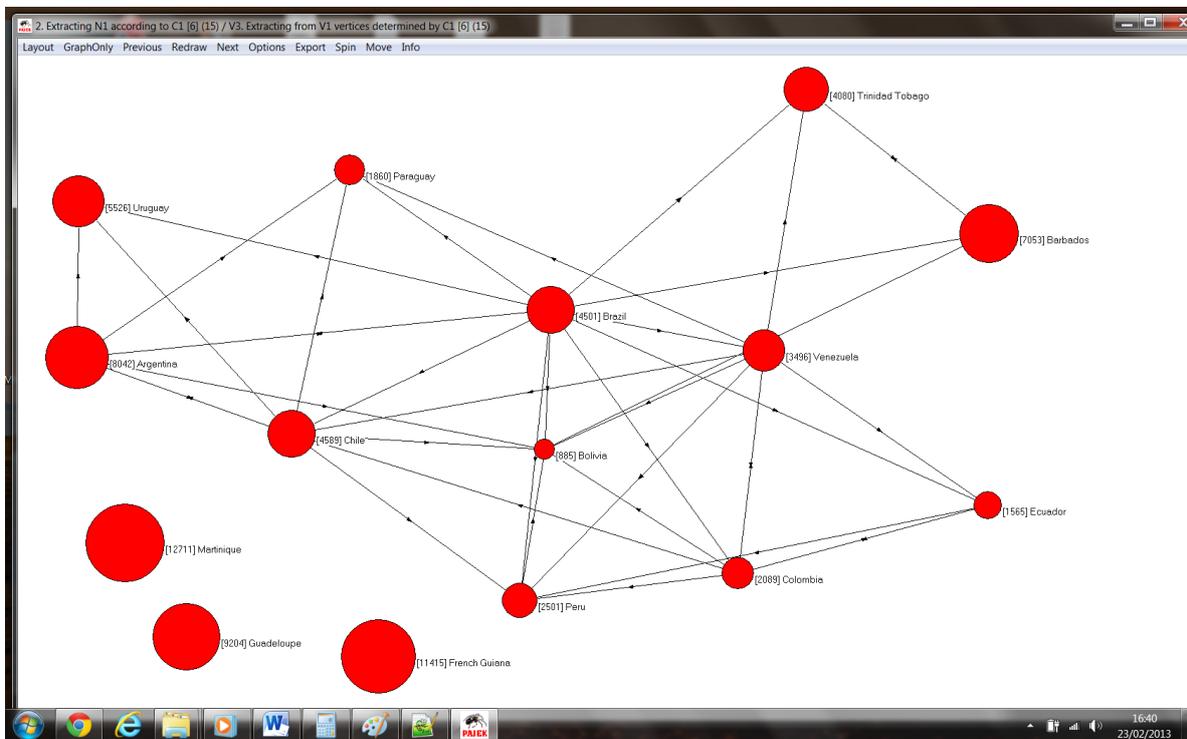
Para extrair um sub-vetor a partir de uma partição:

Operations/Vector+Partition/Extract Subvector



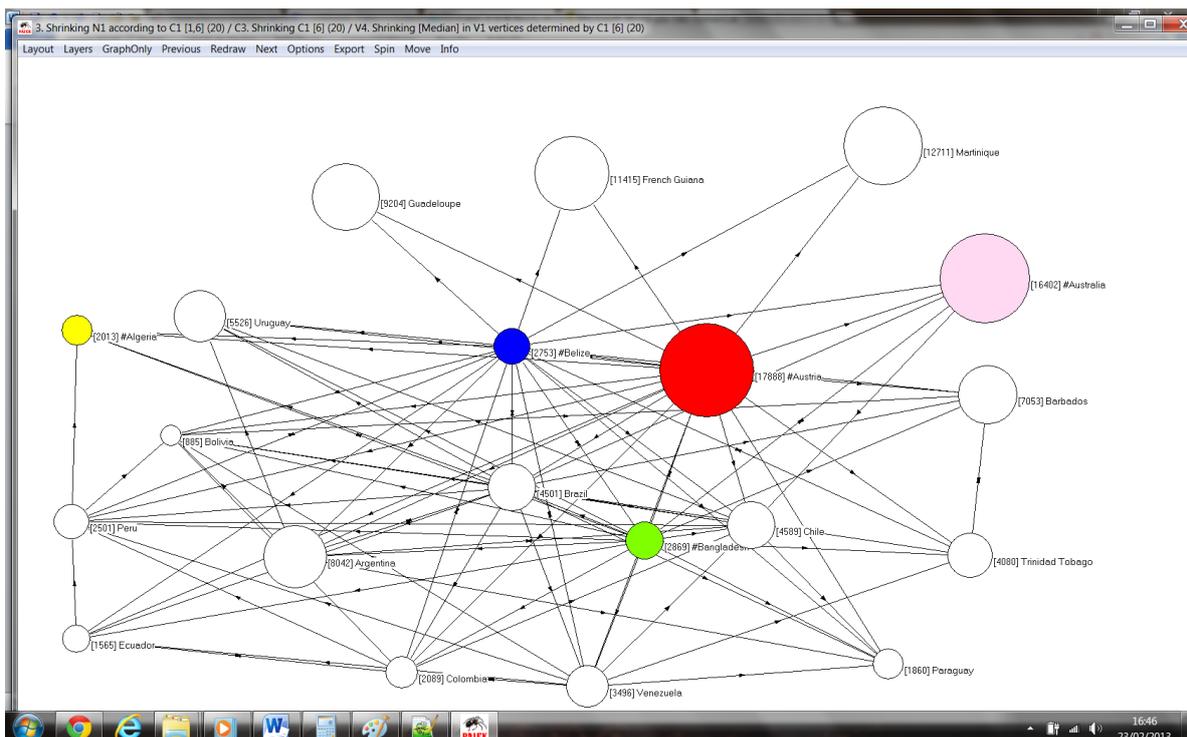
Usando como rede a partição existente associada à classe 6, pode-se obter o gráfico a seguir, onde o tamanho do vértice é associado ao sub-vetor GDP via Draw/Network+First Vector





O mesmo procedimento pode ser aplicado ao comando “Shrink” via

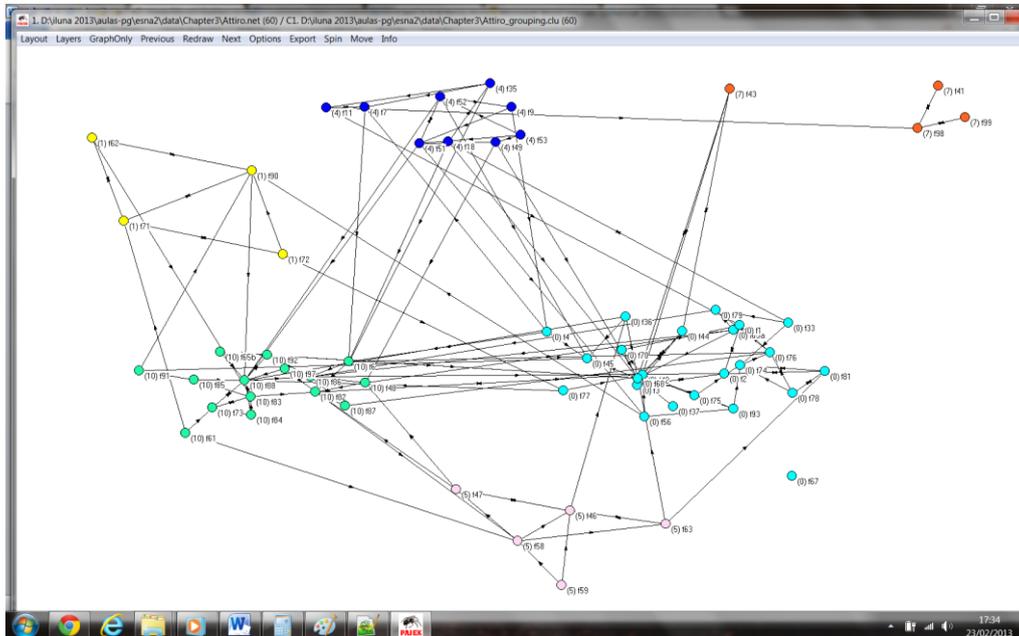
- Operations/Network+Partition/Shrink Network** para obter a rede na visão contextual (escolheu-se a classe 6 para análise)
- Operations/Vector+Partition/Shrink Vector/Median** para extrair o vetor referente à rede gerada e
- Draw/Network+First Partition + First Vector** para obter o desenho da rede considerando a visão contextual sobre a classe 6 e os PIBs de cada vértice (valores medianos para os continentes, existem outras opções de medidas centrais):



Medidas de Coesão

Objetivo: identificar grupos, diferenças e afinidades (homofilia).

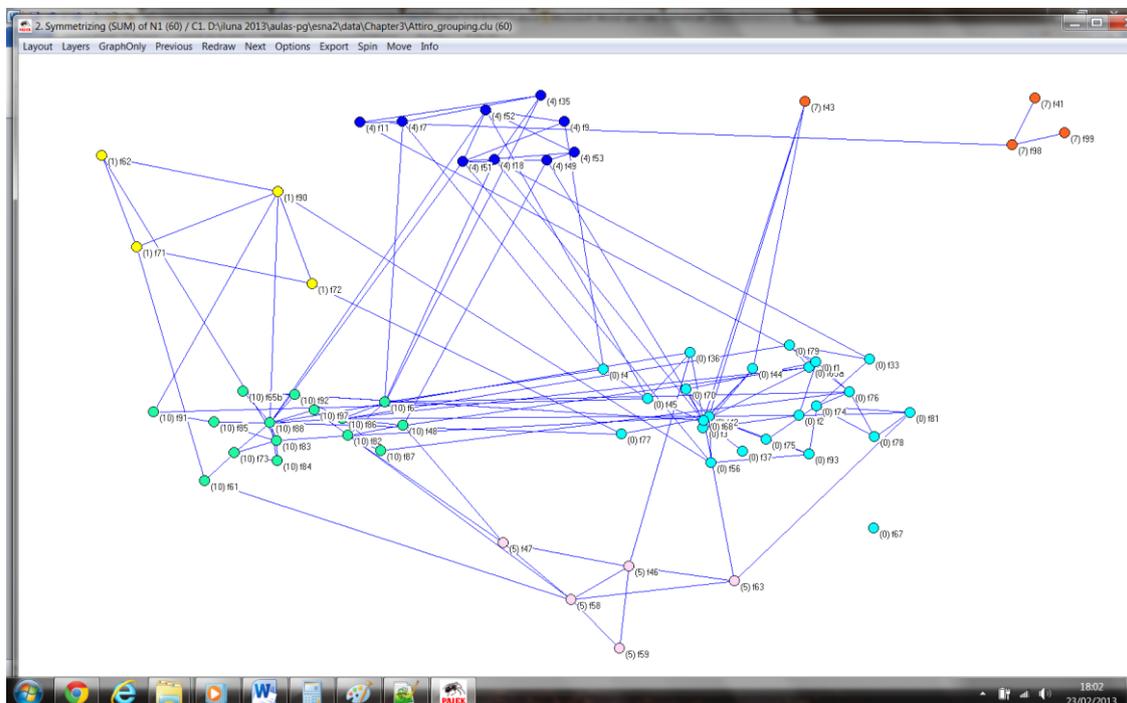
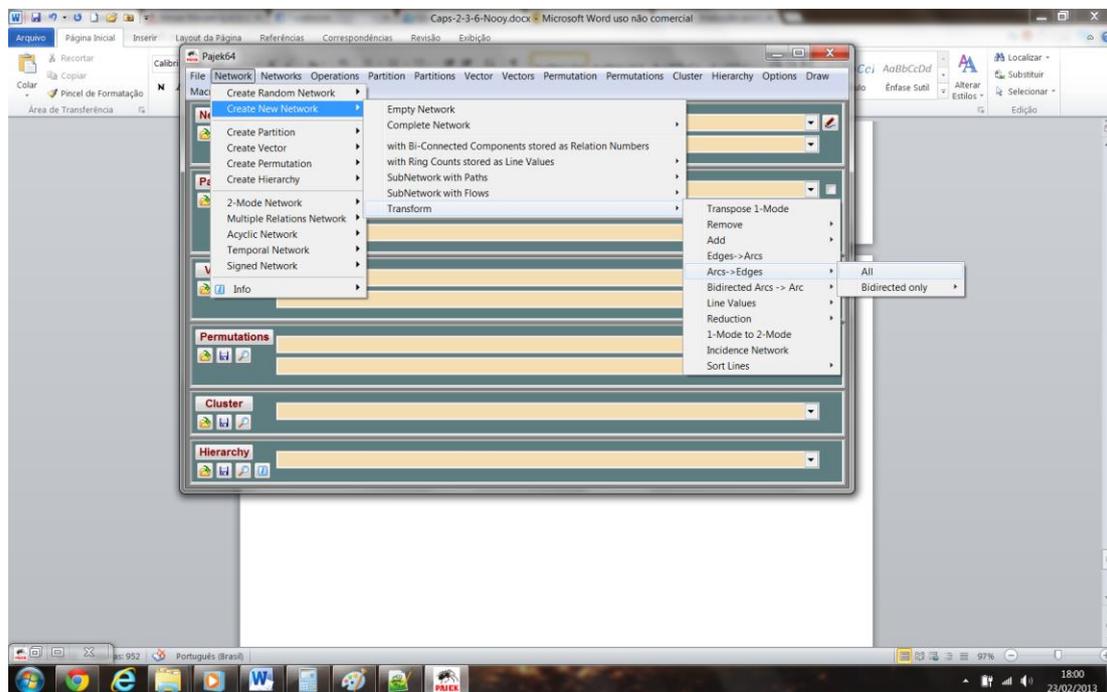
1. Carregar rede **Attiro.net**
2. Carregar Partição **Attiro_grouping.clu**
3. Plotar rede+partição e energizar por cluster.



Densidade

1. Para verificar a densidade da rede (considerada *undirected* e sem levar em conta os pesos dos arcos):
Network/Info/General
Obs.: a densidade de uma rede é afetada pelo seu tamanho $n \rightarrow$ não pode ser usada para fins de comparação entre redes de tamanhos diferentes. Alternativa: grau médio (*average degree*), pois elimina o efeito de n .
2. Transformar a rede em rede não direcionada e salvar a nova rede gerada:

Network/Create New Network/Transform/Arcs->Edges/All



3. Para armazenar o grau (entrada,saída ou all) de cada vértice:

Network/Create Partition/Degree/Input

4. Para ver a partição em uma tabela de frequência:

Partition/Info

| Cluster | Freq | Freq% | CumFreq | CumFreq% | Representative |
|---------|------|----------|---------|----------|----------------|
| 0 | 1 | 1.6667 | 1 | 1.6667 | f67 |
| 1 | 3 | 5.0000 | 4 | 6.6667 | f37 |
| 2 | 1 | 1.6667 | 5 | 8.3333 | f59 |
| 3 | 19 | 31.6667 | 24 | 40.0000 | f3 |
| 4 | 20 | 33.3333 | 44 | 73.3333 | f1 |
| 5 | 4 | 6.6667 | 48 | 80.0000 | f45 |
| 6 | 6 | 10.0000 | 54 | 90.0000 | f51 |
| 7 | 2 | 3.3333 | 56 | 93.3333 | f6 |
| 8 | 1 | 1.6667 | 57 | 95.0000 | f86 |
| 9 | 1 | 1.6667 | 58 | 96.6667 | f42 |
| 13 | 1 | 1.6667 | 59 | 98.3333 | f88 |
| 14 | 1 | 1.6667 | 60 | 100.0000 | f68 |
| Sum | 60 | 100.0000 | | | |

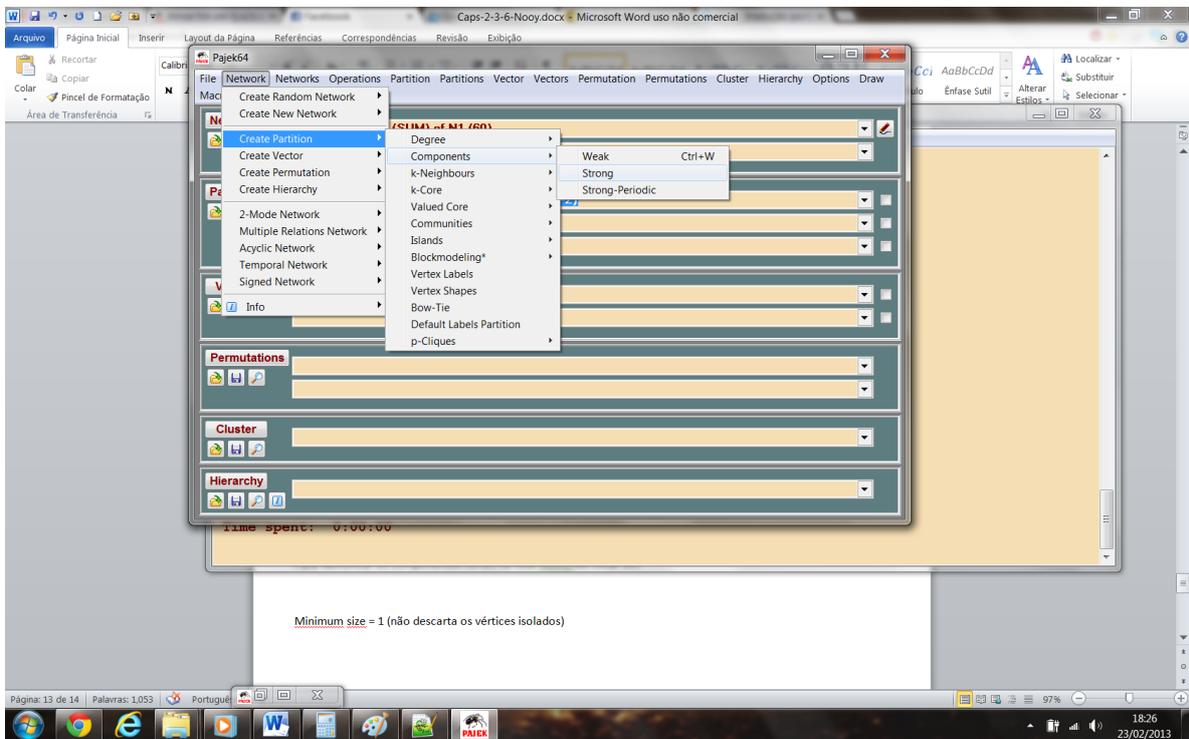
Componentes

- **Rede fracamente conectada:** todos os vértices estão conectados por um caminho (path), desconsiderando o sentido dos arcos.
- **Rede fortemente conectada:** todos os vértices estão conectados por um caminho, considerando ainda o sentido dos arcos (redes direcionadas).
- **Componente fraco:** subrede fracamente conectada
- **Componente forte:** subrede fortemente conectada

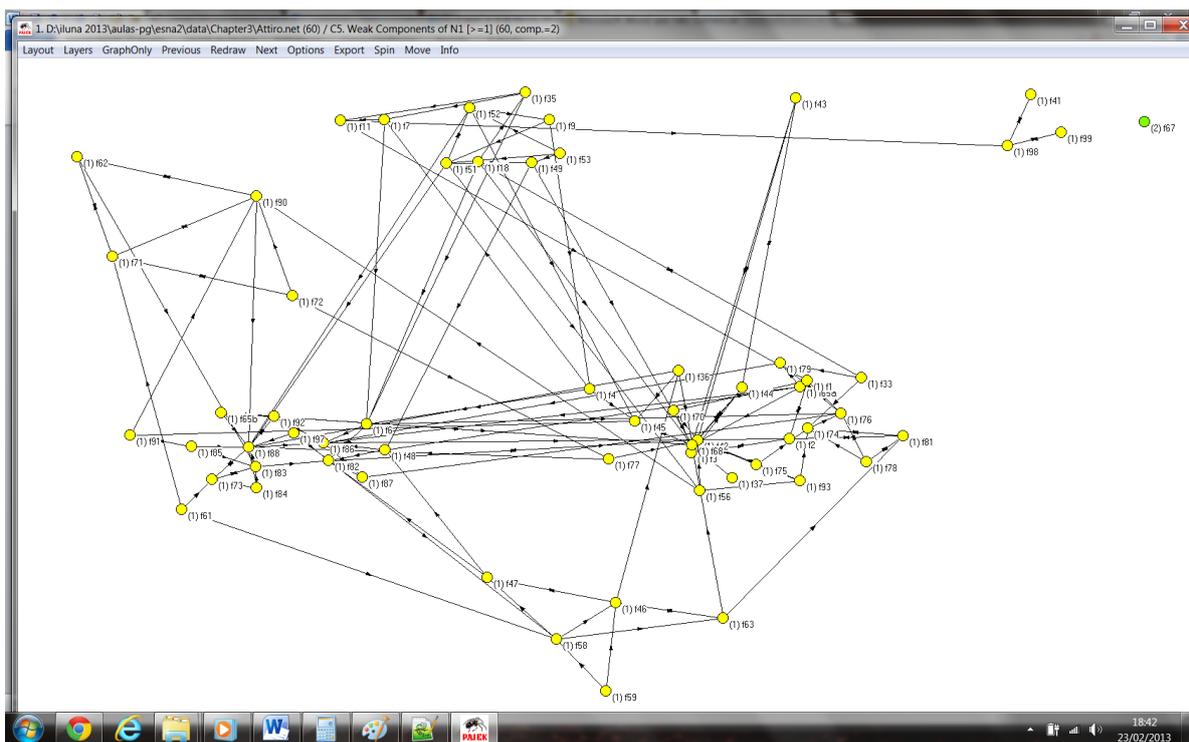
1. Para identificar os componentes fortes na rede Attiro.net clicar em

[Network/Create Partition/Components/Strong](#)

Minimum size = 1 (não descarta os vértices isolados)



2. Para visualizar os componentes: Draw/Network+Partition



Cores (núcleos)

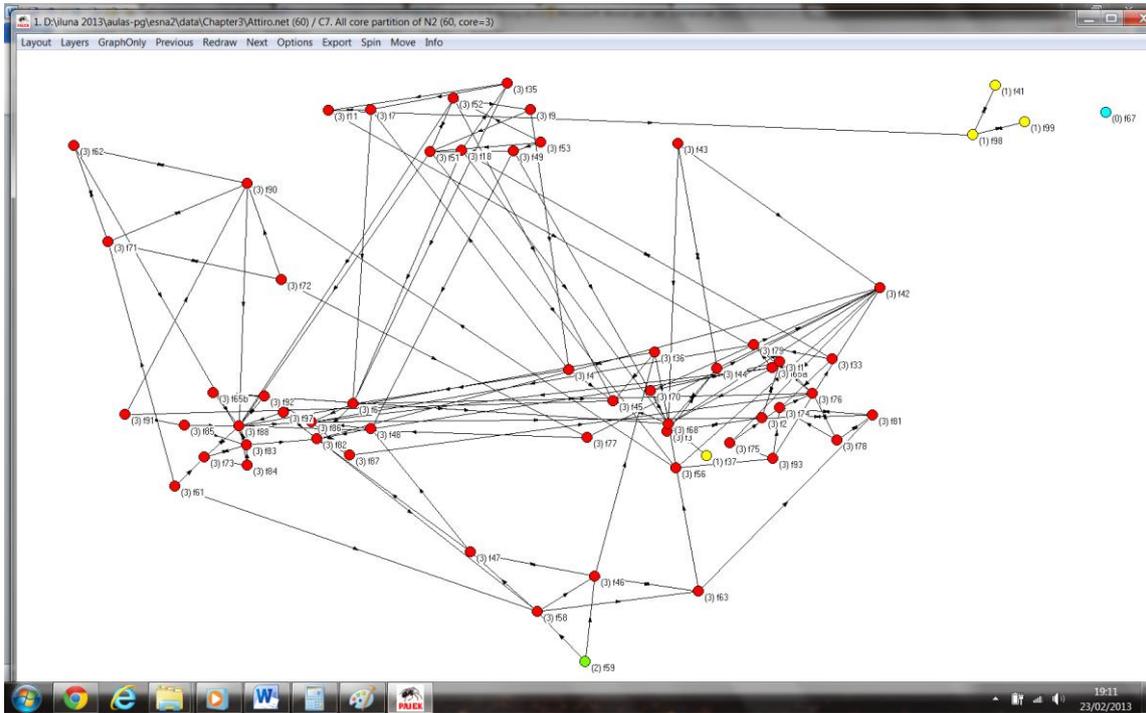
k -core: k indica o grau mínimo que todo elemento do “core” possui. Exemplo: 2-core implica que todos os elementos pertencendo a esse grupo tem pelo menos um grau igual a 2.

Rede k -core: todos os vértices têm grau mínimo igual a k . Vinculado principalmente a **redes não direcionadas** (alerta do próprio Nooy). No exemplo das famílias:

1. Para classificar os vértices por k-core clicar em

[Network/Create Partition/k-core](#)

2. Para visualizar, clicar em [Draw/Network+First Partition](#)

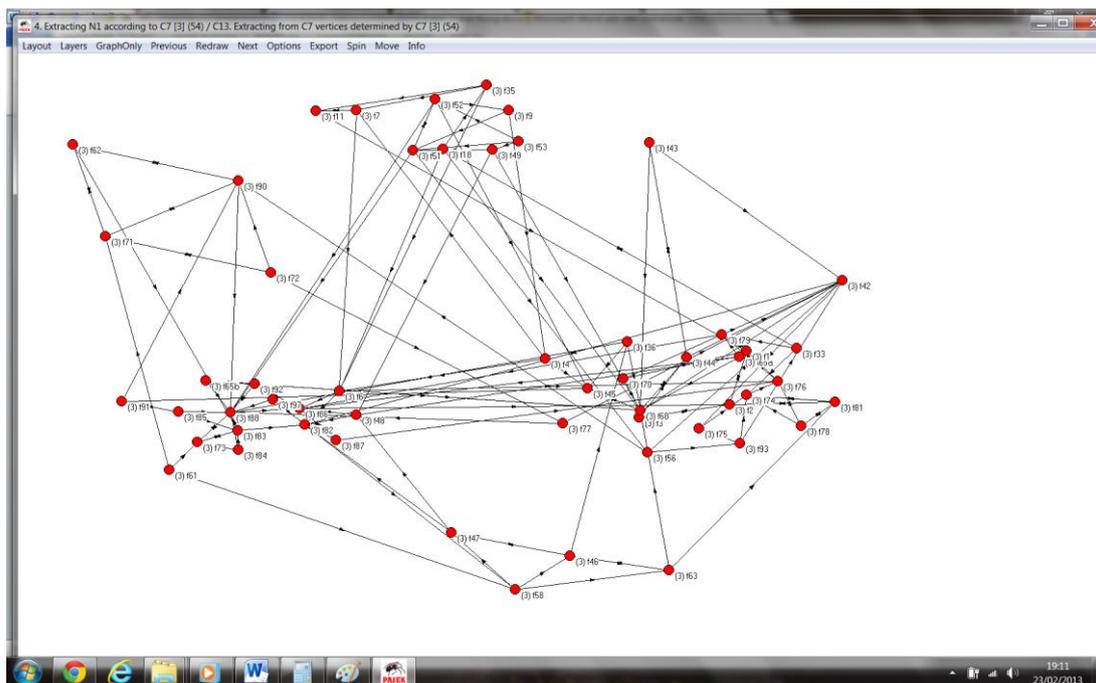


3. Para isolar o grupo com maior número de elementos ($k=3$)

[Operations/Network+Partition/Extract SubNetwork](#)

[Select clusters = 3](#)

4. Visualizando:



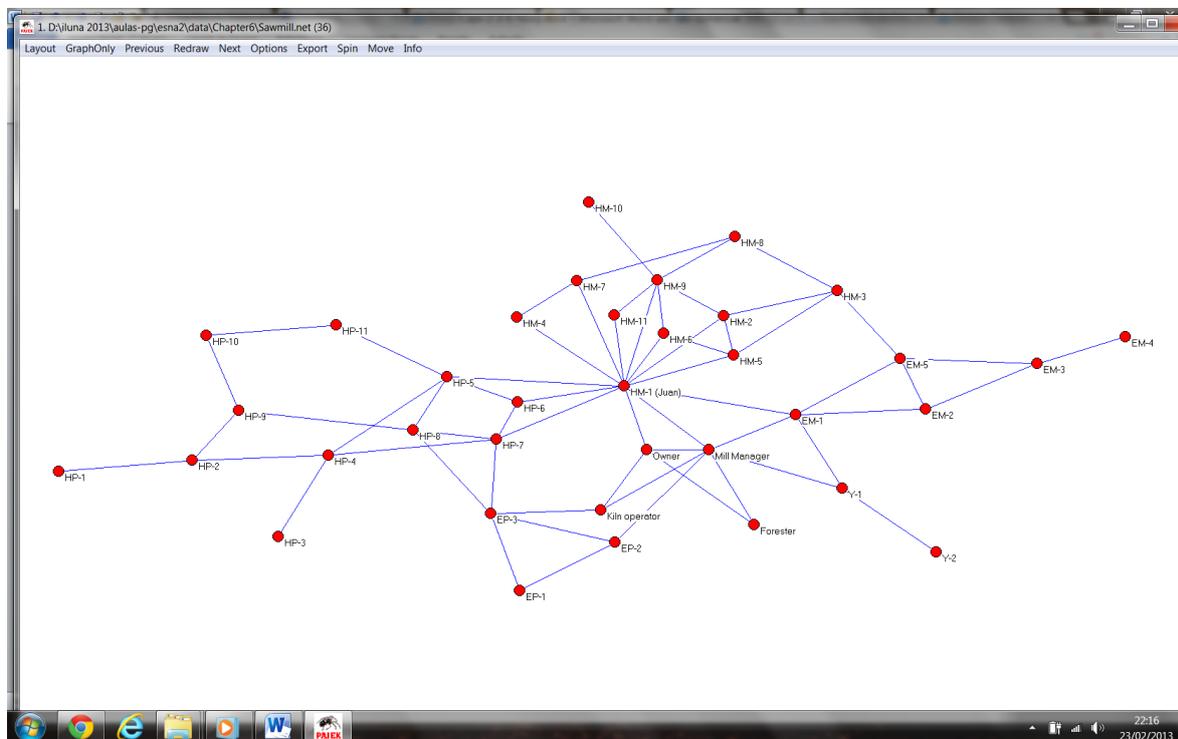
A partir dos comandos anteriores, é possível verificar se o k-core selecionado possui mais de um componente. Ainda, combinando um vetor de graus + uma partição de etnografia, podemos identificar os atores centrais de cada grupo tendo como critério o grau dos vértices.

Centro e periferia (Centralidade)

Voltado para redes não direcionadas. Para redes direcionadas ver parte IV do livro (Ranking, prestígio). Identifica atores centrais na rede.

A rede Sawmill.net é composta por 36 empregados classificados de acordo com a origem e o tipo de emprego

- H: Spanish speaking
- E: English speaking
- M: Mill section
- P: Planer section
- Y: Yard



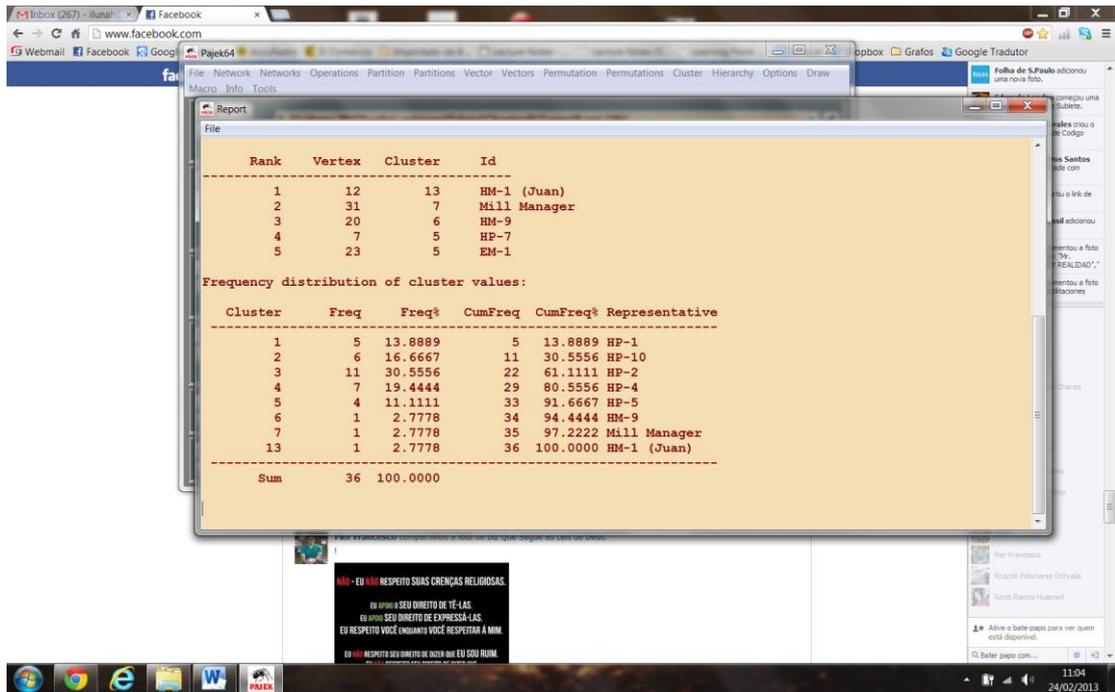
Intuitivamente, podemos afirmar que Juan encontra-se em uma posição central da rede, pois conta com um alto número de vizinhos diretos, e a partir dessas relações, consegue se comunicar com todos na rede. Ainda, entre os seus vizinhos diretos, encontram-se o dono e o gerente da serraria.

Medidas de centralidade

1. **Centralidade de grau:** Mede o máximo número de fontes de informação direta do vértice a partir da cardinalidade da vizinhança. Nesse sentido, a estrutura mais eficiente é a estrela. Assim, a centralidade de uma rede indica a eficiência da rede em termos de fluxo de informação entre todos os elementos que a compõem.

Obs.: a definição de centralidade da rede dada pelo livro difere do grau “médio” na rede. É altamente recomendado trabalhar com redes não direcionadas e eliminar os links múltiplos e loops.

- a. Para calcular o grau dos vértices ir para **Network/Create Partition/Degree/All**
- b. Para visualizar a distribuição de grau: **Partition/Info**



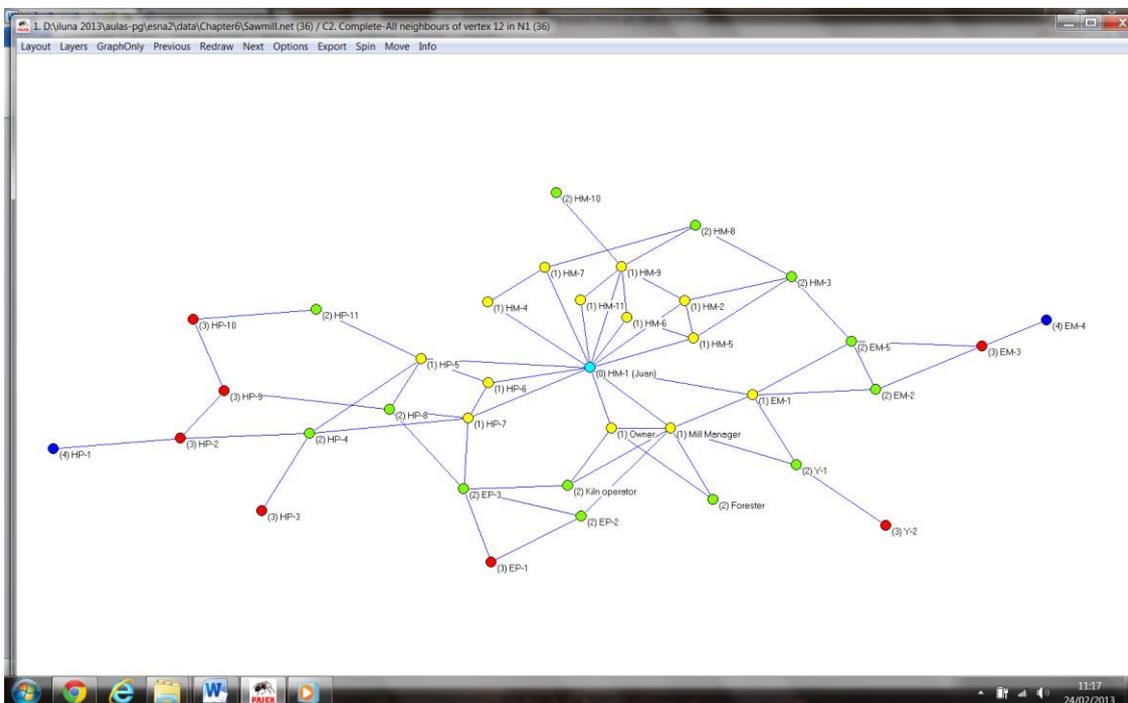
- c. Para identificar os vizinhos (k-neighbors) de um determinado vértice, clicar em

Network/Create Partition/k-neighbours/All

Selected Vertex = 12 (Juan; pode ser outro)

Maximum distance (0: no limit) = 0 (útil para redes muito grandes – limita a busca)

Draw/Network+First Partition



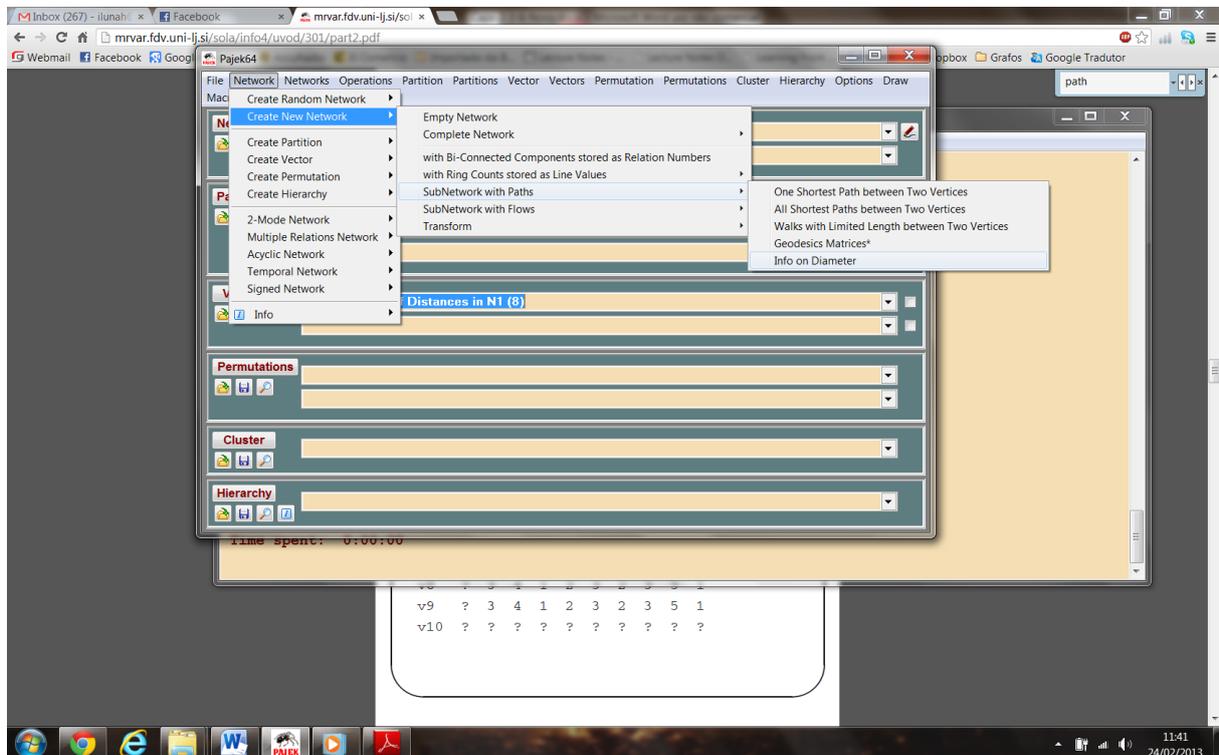
Obs.:

- Input: calcula a distância a partir do vértice selecionado;
- Output: calcula a distância para o vértice selecionado;

- All: desconsidera o sentido dos arcos.
 - 999999998: valor usado para indicar que dois nós não são alcançáveis entre si (geodésica infinita).
- d. Para calcular as geodésicas entre os vértices, escolher antes na tela de Draw, a opção de marcar vértices com "Labels" **Options/Mark Vertices Using/Labels**

Entrar no menu **Network/Create New Network/Subnetwork with Paths**

E escolher a opção mais adequada



Escolhendo por exemplo **All Shortest Paths between Two Vertices**,

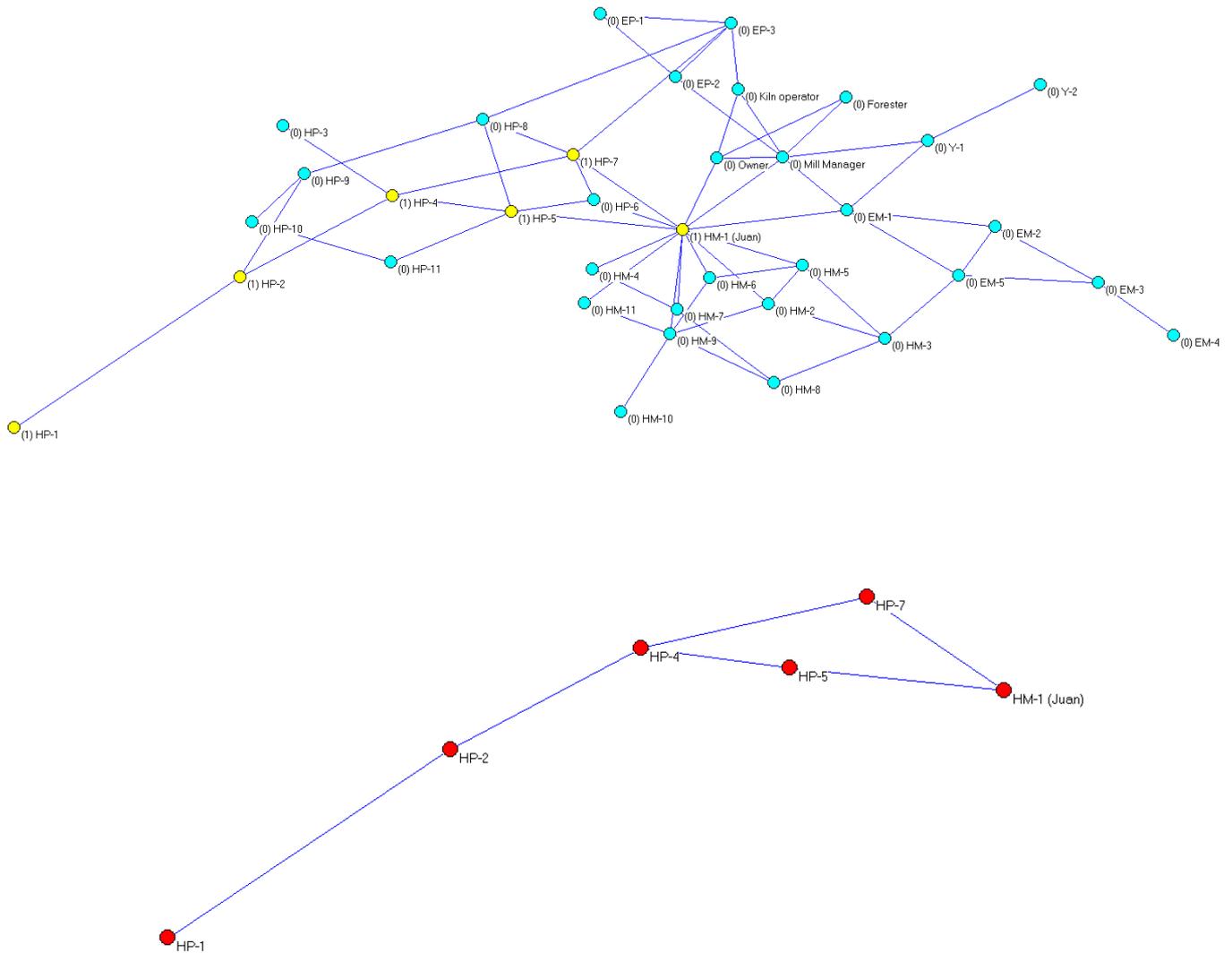
Initial Vertex = 1

Terminal vertex= 12

Forget values of lines? = Yes (caso os pesos dos vértices não sejam de interesse)

Identify vertices in the source network? = Yes (cria uma partição de duas classes 0 e 1. Coloca na classe 1 a todos os vértices que compõem a geodésica), **No** (apresenta o resultado no relatório)

Draw/Network+First Partition



2. **Centralidade de proximidade:** Por adotar valores reais, este encontra-se na opção **Network/Create Vector/Degree/Closeness/**

O seu cálculo é computacionalmente intensivo. Após o cálculo, a distribuição do vetor pode ser visto via **Vector/Info**

3. **Centralidade de intermediação:** **Network/Create Vector/Degree/Betweenness**

Sites de apoio:

SNA: <http://mrvar.fdv.uni-lj.si/pajek/be2.htm>

MIT Course 14 (Economics) - Networks: <http://ocw.mit.edu/courses/economics/14-15j-networks-fall-2009/>

FAQ

Send Pajek mailing list submissions to
pajek@list.fmf.uni-lj.si

To subscribe or unsubscribe via the World Wide Web, visit
<http://list.fmf.uni-lj.si/cgi-bin/mailman/listinfo/pajek>