



Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

Introduction to Network Analysis using **Pajek**

5. Structure of networks: cohesion

Vladimir Batagelj

IMFM Ljubljana and IAM UP Koper

PhD and MS program in Statistics
University of Ljubljana, 2022



Outline

Cohesion

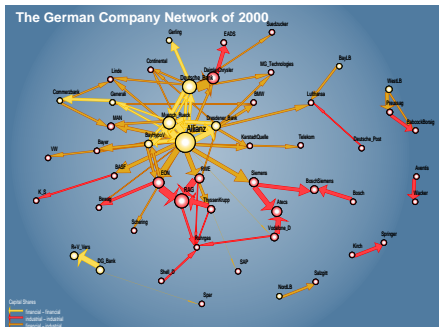
V. Batagelj

Islands

Cores

Generalized
cores

- 1 Islands
- 2 Cores
- 3 Generalized cores



L. Krempf, MPI.

Vladimir Batagelj: vladimir.batagelj@fmf.uni-lj.si

Current version of slides (March 23, 2022 at 00 : 19): [slides PDF](#)



Islands

Cohesion

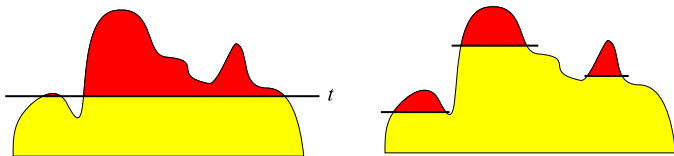
V. Batagelj

Islands

Cores

Generalized
cores

If we represent a given or computed value of nodes / links as a height of nodes / links and we immerse the network into a water up to selected level we get *islands*. Varying the level we get different islands.



We developed very efficient algorithms to determine the islands hierarchy and to list all the islands of selected sizes.
See [details](#).



... Islands

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

Islands are a very general and efficient approach to determine the 'important' subnetworks in a given network.

We have to express the goals of our analysis with a related property of the nodes or weight of the links. Using this property we determine the islands of an appropriate size (in the interval k to K).

In large networks we can get many islands which we have to inspect individually and interpret their content.

An important property of the islands is that they identify locally important subnetworks on different levels. Therefore they detect also emerging groups.

The set of nodes $\mathcal{C} \subseteq \mathcal{V}$ is a **local node peak**, if it is a regular node island and all of its nodes have the same value. Node island with a single local node peak is called a **simple node island**. In similar way we define simple link island.



... Islands

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

A set of nodes $C \subseteq \mathcal{V}$ is a *regular node island* in network $\mathcal{N} = (\mathcal{V}, \mathcal{L}, \rho)$, $\rho : \mathcal{V} \rightarrow \mathbb{R}$ iff it induces a connected subgraph and the nodes from the island are 'higher' than the neighboring nodes

$$\max_{u \in N(C)} \rho(u) < \min_{v \in C} \rho(v)$$

A set of nodes $C \subseteq \mathcal{V}$ is a *regular link island* in network $\mathcal{N} = (\mathcal{V}, \mathcal{L}, w)$, $w : \mathcal{L} \rightarrow \mathbb{R}$ iff it induces a connected subgraph and the links inside the island are 'stronger related' among them than with the neighboring nodes – in \mathcal{N} there exists a spanning tree \mathcal{T} over C such that

$$\max_{(u,v) \in \mathcal{L}, u \notin C, v \in C} w(u, v) < \min_{(u,v) \in \mathcal{T}} w(u, v)$$

Network/Create Partition/Islands/Line Weights
Operations/Network+Vector/Islands/Vertex
Property





Some properties of node islands

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

- The sets of nodes of connected components of node-cut at selected level t are regular node islands.
- The set $\mathcal{H}_p(\mathcal{N})$ of all regular node islands of network \mathcal{N} is a complete hierarchy:
 - two islands are disjoint or one of them is a subset of the other
 - each node belongs to at least one island
- Node islands are invariant for the strictly increasing transformations of the property p .
- Two linked nodes cannot belong to two disjoint regular node islands.



Simple node islands

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

- The set of nodes $\mathcal{C} \subseteq \mathcal{V}$ is a **local node peak**, if it is a regular node island and all of its nodes have the same value.
- Node island with a single local node peak is called a **simple node island**.
- The types of node islands:
 - FLAT – all nodes have the same value
 - SINGLE – island has a single local node peak
 - MULTI – island has more than one local node peaks
- Only the islands of type FLAT or SINGLE are simple islands.



Some properties of link islands

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

- The sets of nodes of connected components of link-cut at selected level t are regular link islands.
- The set $\mathcal{H}_w(\mathcal{N})$ of all nondegenerated regular link islands of network \mathcal{N} is hierarchy (not necessarily complete):
 - two islands are disjoint or one of them is a subset of the other
- Link islands are invariant for the strictly increasing transformations of the weight w .
- Two linked nodes may belong to two disjoint regular link islands.



Simple link islands

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

- The set of nodes $\mathcal{C} \subseteq \mathcal{V}$ is a **local link peak**, if it is a regular link island and there exists a spanning tree of the corresponding induced network, in which all links have the same value as the link with the largest value.
- Link island with a single local link peak is called a **simple link island**.
- The types of link islands:
 - FLAT – there exists a spanning tree, in which all links have the same value as the link with the largest value.
 - SINGLE – island has a single local link peak.
 - MULTI – island has more than one local link peaks.
- Only the islands of type FLAT or SINGLE are simple islands.



US patents

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

US patents network (Nber, US Patents) has 3774768 nodes and 16522438 arcs (1 loop). Without the loop it is acyclic. The weight of an arc is the proportion of paths through the arc from some initial node to some terminal node. We determined 1 (2,90)-islands. The corresponding subnetwork has 470137 nodes, 307472 arcs and for different k : $C_2 = 187610$, $C_5 = 8859$, $C_{30} = 101$, $C_{50} = 30$, ... islands.

Rolex

[1]	0	139793	29670	9288	3966	1827	997	578	362	250
[11]	190	125	104	71	47	37	36	33	21	23
[21]	17	16	8	7	13	10	10	5	5	5
[31]	12	3	7	3	3	3	2	6	6	2
[41]	1	3	4	1	5	2	1	1	1	1
[51]	2	3	3	2	0	0	0	0	0	1
[61]	0	0	0	0	1	0	0	2	0	0
[71]	0	0	1	1	0	0	0	1	0	0
[81]	2	0	0	0	0	1	2	0	0	7



Distribution of island size

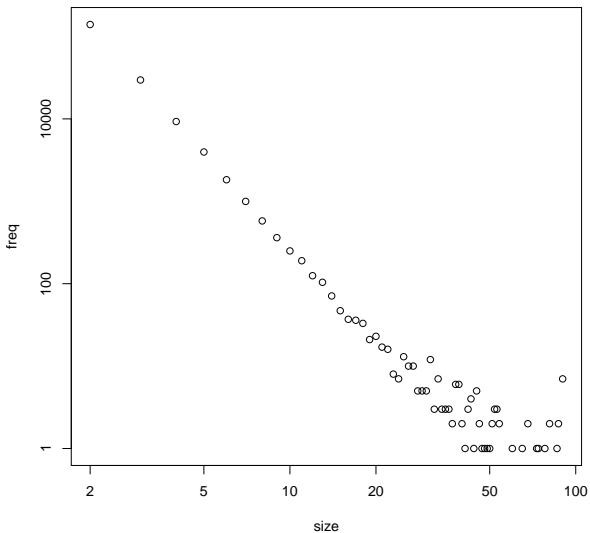
Cohesion

V. Batagelj

Islands

Cores

Generalized
cores



V. Batagelj

Cohesion



Main island Liquid Crystal display

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

Table 1: Patents on the liquid-crystal display

patent	date	inventor's name and title
2419489	Mar 14, 1976	Moyl: Electric light-polarizing sheet and the film and the formation and use thereof
2682542	Jun 29, 1954	Wooler, et al. Reflection of anisotropic carbide
3322485	May 20, 1975	Williams: Electro-optical elements utilizing an organic semiconducting compound
3626418	Jan 18, 1972	Josephson: Preparation of polycrystalline anisotropic compounds
3660484	May 30, 1972	Mockenhaus, et al. Liquid crystal vertical imaging system having an undistorted image on a distorted background
3675987	Jul 11, 1973	Hahn: Liquid crystal compositions and devices
3691715	Sep 10, 1972	Grazel: Clock with digital display
3691716	Sep 10, 1972	Wynick: Electro-optic system in which an electromechanical or dielectric material is dispersed throughout a liquid-crystal to reduce the transmittance
3737386	Oct 23, 1973	Avramis, et al. Phase trans-stilbene compounds, some displaying nematic mesophase at or near room temperature and others by a range up to 300°C
3773747	Nov 20, 1973	Solomonson: Substituted azoxy benzeno compounds
3795436	Mar 5, 1974	Huber, et al. Nematicotropic material which exhibits the Kerr effect at isotropic temperatures
3796479	Mar 12, 1974	Helfrich, et al. Electro-optical light modulation cell utilizing a nematicotropic material which exhibits the Kerr effect at isotropic temperatures
3872140	Mar 18, 1975	Klaernerman, et al. Liquid crystalline compositions and method
3872626	Apr 8, 1975	Deutscher, et al. Use of anisotropic liquid crystalline substances
3881866	May 6, 1975	Suzuki: Electro-optical display device
3891367	Jun 24, 1975	Tokumoto, et al. Phase control of the voltage applied to electro-optic display for a cholesteric to nematic phase transition display
3904725	Mar 30, 1976	Gray, et al. Liquid crystal materials and devices
3954653	Mar 4, 1976	Yoshino, et al. Liquid crystalline compositions having high dielectric anisotropy and display device incorporating same
3960732	Jun 1, 1976	Kilbinger, et al. Liquid crystal compositions
3973286	Aug 17, 1976	Oh: Low voltage actuated field effect liquid crystal compositions and method of synthesis
4000884	Dec 28, 1976	Block, et al. Liquid crystal mixtures for electro-optical display
4011713	Mar 9, 1977	Solomonson: Modified nematic mixtures with positive dielectric anisotropy
4013382	Mar 22, 1977	Barthelme: Liquid crystal compounds and electro-optic devices incorporating them
4017416	Apr 12, 1977	Block, et al. α -Cyano-2-(4-alkyl-4'-alkoxyphenyl)acetone for preparing same and liquid crystal compositions using same
4029595	Jun 14, 1977	Row, et al. Novel liquid crystal compounds and electro-optic devices incorporating them
4032749	Jun 28, 1977	Gray, et al. Diphenyl acetylene-biphenyl compounds and liquid crystal materials containing them
4072426	Mar 7, 1978	Gray, et al. Diphenyl acetylene-biphenyl compounds and liquid crystal materials containing them
4084226	Apr 4, 1978	Hsu: Liquid crystal composition and method

Table 2: Patents on the liquid-crystal display

patent	date	inventor's name and title
4085799	Apr 11, 1978	Oh: Nematic liquid crystal compositions
4112647	Sep 12, 1978	Coutts, et al. Liquid crystalline materials
4112652	Oct 3, 1978	Krause, et al. Liquid crystalline materials of reduced viscosity
4120592	Dec 10, 1978	Edelmann, et al. Liquid crystalline cyanoazobenzene derivatives
4149411	Apr 17, 1979	Gray, et al. Optically active liquid crystal mixtures and liquid crystal devices containing them
4154087	May 15, 1979	Edelmann, et al. Liquid crystalline benzodihydroxyphenyl derivatives
4195916	Apr 1, 1980	Coutts, et al. Liquid crystal compounds
4198130	Apr 15, 1980	Bolke, et al. Liquid crystal mixtures
4202791	May 13, 1980	Sato, et al. Nematic liquid crystalline materials
4229315	Oct 21, 1980	Krause, et al. Liquid crystalline cyanoazobenzene derivatives
4263052	Apr 14, 1981	Gray, et al. Liquid crystal compounds and materials and devices containing them
4269095	Sep 22, 1981	Deutscher, et al. Liquid crystal compounds
4292524	Oct 6, 1981	Kochle, Ester compound
4302252	Nov 24, 1981	Edelmann, et al. Fluorohydroxybenzenes, the preparation thereof and their use as components of liquid crystal dielectrics
4330426	Nov 18, 1982	Edelmann, et al. Cyanoalkylhydroxybenzenes, their preparation and use in dielectrics and electrooptical display elements
4340498	Jul 20, 1982	Sauer, Halogenated ester derivatives
4349452	Sep 14, 1982	Omura, et al. Cyanoalkylhydroxybenzenes
4357078	Nov 2, 1982	Gray, et al. Liquid crystal compounds containing an alkyloxy ring and exhibiting a low dielectric anisotropy and liquid crystal materials and devices incorporating such compounds
4361494	Nov 30, 1982	Omura, et al. Aminoxy cyanoalkyl cyanoalkylhydroxyphenyl ethers
4384135	Jun 11, 1983	Omura, et al. Aminoxy compounds with negative or positive DC-anisotropy and optically anisotropic
4386007	May 31, 1983	Krause, et al. Liquid crystalline azobenzene derivatives
4387638	Jun 7, 1983	Pikuli, et al. 4-(Transe-4'-alkoxyphenyl) benzoic acid α - α' -disubstituted esters
4387639	Jun 7, 1983	Solomonson, et al. Trans-4-(trans-4'-alkoxyphenyl)-cyanoazobenzene
4406293	Aug 23, 1983	Krause, et al. Liquid crystalline cyanoalkylhydroxyphenyl derivatives
4415470	Nov 12, 1983	Edelmann, et al. Liquid crystalline fluorine-containing hydroxyphenyls and derivatives and electro-optical display elements
4419263	Dec 6, 1983	Pradelle, et al. Liquid crystalline cyanoalkylcarboazobenzene derivatives
4422951	Dec 27, 1983	Solomonson, et al. Liquid crystal benzeno derivatives
4454545	Jan 10, 1984	Takato, et al. Nematic liquid compound
4456174	Jan 29, 1984	Cherian, et al. Benzimidazole benzeno compositions
4460770	Jul 17, 1984	Petrillia, et al. Liquid crystal mixtures
4472293	Sep 18, 1984	Solomonson: High temperature liquid crystal substance of low ring and liquid crystal compositions containing the same
4472590	Sep 18, 1984	Takato, et al. High temperature liquid-crystal mixtures
4486117	Oct 30, 1984	Takato, et al. Nematic liquid crystalline compounds
4502574	Mar 5, 1985	Takato, et al. High temperature liquid-crystal mixtures containing them
4510869	Apr 9, 1985	Edelmann, et al. Cyanoazobenzene derivatives

Table 3: Patents on the liquid-crystal display

patent	date	inventor's name and title
4510441	Apr 30, 1985	Grigoras, et al. 1-(Transe-4-alkoxyphenyl)-2-(trans-4'-[p-alkoxy-ortho]phenyl)cyanoalkylbenzene and liquid crystal mixtures
4526704	Jul 2, 1985	Petrillia, et al. Multilayer liquid crystal elements
4530981	Nov 5, 1985	Petrillia, et al. Liquid crystalline esters and mixtures
4534515	Dec 10, 1985	Takato, et al. Nematic liquid crystalline compounds
4538286	Apr 22, 1986	Petrillia, et al. Fluorethanes
4621361	Nov 11, 1986	Petrillia, et al. Novel liquid crystal mixtures
4630986	Dec 23, 1986	Petrillia, et al. Benzimidazoles
4631705	Apr 14, 1987	Sato, et al. Substituted propanolone
4650962	Apr 21, 1987	Watson, et al. Ethane derivatives
4659131	Sep 22, 1987	Bullbeck, et al. Dimethylated ethane and their use in liquid crystal materials and devices
4704227	Nov 2, 1987	Krause, et al. Liquid crystal compounds
4709020	Nov 28, 1987	Petrillia, et al. Novel liquid crystal mixtures
4711317	Dec 1, 1987	Schulz, et al. Aminoxy compounds and liquid crystal mixtures therein
4713197	Dec 12, 1987	Edelmann, et al. Nitrogen-containing heterocyclic compounds
4719026	Jan 12, 1988	Waldner, et al. Cyanoazobenzene derivatives
4721367	Jan 26, 1988	Yoshino, et al. Liquid crystal devices
4725414	Jan 27, 1988	Edelmann, et al. Nitrogen-containing heterocyclic compounds
4739063	Jan 13, 1988	Bachewer, et al. Liquid crystalline compounds
4757519	Sep 23, 1988	Vandekerckhove, et al. 2,2'-bis(4-alkoxy-4'-hydroxyphenyl) ether and their derivatives, their production process and their use in liquid crystal display devices
4797228	Jun 10, 1989	Goto, et al. Cyanoazobenzene derivative and liquid crystal compositions containing same
4826830	Apr 11, 1989	Krause, et al. Nitrogen-containing heterocyclic ethers
4832462	May 23, 1989	Clark, et al. Liquid crystal devices
4837147	Oct 31, 1989	Wolter, et al. Liquid crystal element
4957749	Sep 18, 1990	Chen, et al. Active matrix screen for the color display of television pictures, control system and process for producing said screen
5010086	Mar 21, 1991	Edelmann, et al. Liquid crystal display device with a birefringent compensator
5010180	Mar 21, 1991	Ohbayashi, et al. Liquid crystal element with improved contrast and lightness
5122295	Jun 16, 1992	Waldner, et al. Matrix liquid crystal display
5124824	Jun 23, 1992	Kaneko, et al. Liquid crystal display device comprising a retardation compensator layer having a maximum principal refractive index in the thickness direction
5137480	Dec 15, 1992	Herrlich, et al. Liquid-crystal matrix display
5280877	Feb 9, 1994	Edelmann, et al. Liquid crystal display with crossed rings between terminal groups
5306238	Mar 3, 1994	Solomonson
5327154	Dec 20, 1994	Wolter, et al. Superpixel liquid-crystal display
5545077	Aug 6, 1996	Flanagan, et al. Nematic liquid-crystal composition
5561316	Sep 10, 1996	Iskakov, et al. Liquid crystal display having adjacent electrode terminals of equal size length
5582626	Nov 4, 1997	Schjerve, et al. Liquid crystal compositions
5858314	Jun 8, 1999	Edelmann, et al. Liquid crystal compositions and liquid crystal display elements

V. Batagelj

Cohesion





The Edinburgh Associative Thesaurus

$n = 23219$, $m = 325624$, transitivity weight

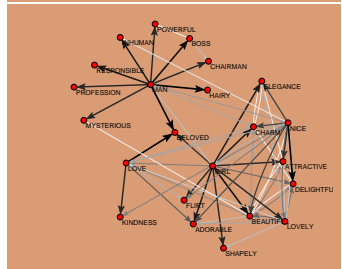
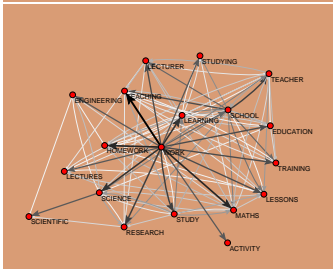
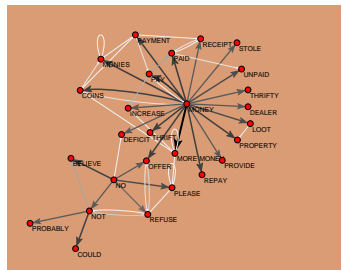
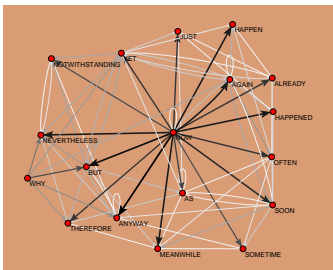
Cohesion

V. Batagelj

Islands

Cores

Generalized
cores





Dense groups

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

Several notions were proposed in attempts to formally describe dense groups in graphs.

Clique of order k is a maximal complete subgraph (isomorphic to K_k), $k \geq 3$.

s -plexes, LS sets, lambda sets, cores, . . .

For all of them, except for cores, it turned out that they are difficult to determine.



Cores and generalized cores

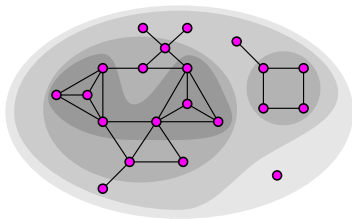
Cohesion

V. Batagelj

Islands

Cores

Generalized
cores



The notion of core was introduced by Seidman in 1983. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph. A subgraph $\mathcal{H} = (W, \mathcal{E}|_W)$ induced by the set W is a *k-core* or a *core of order k* iff $\forall v \in W$: $\deg_{\mathcal{H}}(v) \geq k$, and \mathcal{H} is a maximal subgraph with this property. The core of maximum order is also called the *main* core.

The *core number* of node v is the highest order of a core that contains this node. The degree $\deg(v)$ can be: in-degree, out-degree, in-degree + out-degree, etc., determining different types of cores.



Properties of cores

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

From the figure, representing 0, 1, 2 and 3 core, we can see the following properties of cores:

- The cores are nested: $i < j \implies \mathcal{H}_j \subseteq \mathcal{H}_i$
- Cores are not necessarily connected subgraphs.

An efficient algorithm for determining the cores hierarchy is based on the following property:

If from a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ we recursively delete all nodes, and edges incident with them, of degree less than k , the remaining graph is the k -core.



Generalized cores

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

The notion of core can be generalized to networks. Let $\mathcal{N} = (\mathcal{V}, \mathcal{E}, w)$ be a network, where $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a graph and $w : \mathcal{E} \rightarrow \mathbb{R}$ is a function assigning values to edges. A *node property function* on \mathbf{N} , or a p -function for short, is a function $p(v, U)$, $v \in \mathcal{V}$, $U \subseteq \mathcal{V}$ with real values. Let $N_U(v) = N(v) \cap U$. Besides degrees and (corrected) clustering coefficient, here are some examples of p -functions:

$$p_S(v, U) = \sum_{u \in N_U(v)} w(v, u), \text{ where } w : \mathcal{E} \rightarrow \mathbb{R}_0^+$$

$$p_M(v, U) = \max_{u \in N_U(v)} w(v, u), \text{ where } w : \mathcal{E} \rightarrow \mathbb{R}$$

$$p_t(v, U) = \frac{|\mathcal{L}(U) \cap \mathcal{L}(K(N^+(v)))|}{|\mathcal{L}(K(N^+(v)))|}$$

$$p_k(v, U) = \text{number of cycles of length } k \text{ through node } v \text{ in } (U, \mathcal{E}|U)$$

The subgraph $\mathcal{H} = (C, \mathcal{E}|C)$ induced by the set $C \subseteq \mathcal{V}$ is a *p -core at level* $t \in \mathbb{R}$ iff $\forall v \in C : t \leq p(v, C)$ and C is a maximal such set.



Additional p -functions

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

relative density

$$p_\gamma(v, \mathcal{C}) = \frac{\deg(v, \mathcal{C})}{\max_{u \in N(v)} \deg(u)}, \text{ if } \deg(v) > 0; 0, \text{ otherwise}$$

diversity

$$p_\delta(v, \mathcal{C}) = \max_{u \in N^+(v, \mathcal{C})} \deg(u) - \min_{u \in N^+(v, \mathcal{C})} \deg(u)$$

average weight

$$p_a(v, \mathcal{C}) = \frac{1}{|N(v, \mathcal{C})|} \sum_{u \in N(v, \mathcal{C})} w(v, u), \text{ if } N(v, \mathcal{C}) \neq \emptyset; 0, \text{ otherwise}$$



Generalized cores algorithms

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

The function p is *monotone* iff it has the property

$$C_1 \subset C_2 \Rightarrow \forall v \in \mathcal{V} : (p(v, C_1) \leq p(v, C_2))$$

The degrees and the functions p_S , p_M and p_k are monotone. For a monotone function the p -core at level t can be determined, as in the ordinary case, by successively deleting nodes with value of p lower than t ; and the cores on different levels are nested

$$t_1 < t_2 \Rightarrow \mathcal{H}_{t_2} \subseteq \mathcal{H}_{t_1}$$

The p -function is *local* iff $p(v, U) = p(v, N_U(v))$.

The degrees, p_S and p_M are local; but p_k is **not** local for $k \geq 4$.

For a local p -function an $O(m \max(\Delta, \log n))$ algorithm for determining the p -core levels exists, assuming that $p(v, N_C(v))$ can be computed in $O(\deg_C(v))$.

For details see the [paper](#).



Cores and generalized cores / Pajek commands

Cohesion

V. Batagelj

Islands

Cores

Generalized
cores

```
File/Network/Read [Geom.net]
Network/Create Partition/k-Core/All
Info/Partition
Operations/Network+Partition/Extract Subnetwork [13-*]
Draw/Network+First Partition
Layout/Energy/Kamada-Kawai
Options/Values of lines/Similarities
Layout/Energy/Kamada-Kawai
Operations/Network+Partition/Extract Subnetwork [21]
Draw/Network
Layout/Energy/Kamada-Kawai
Options/Values of lines/Forget
Layout/Energy/Kamada-Kawai
[select Geom.net]
Network/Create Vector/Generalized Cores/Sum/All
Info/Vector
Vector/Make Partition/by Intervals/Selected Thresholds [4]
Info/Partition
Operations/Network+Partition/Extract Subnetwork [2]
Draw/Network
Options/Values of lines/Similarities
Layout/Energy/Fruchterman-Reingold
```



Cores of orders 10–21 in Computational Geometry

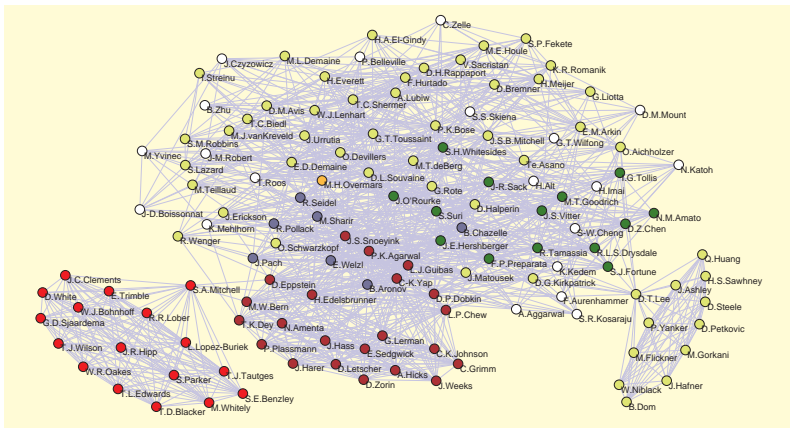
Cohesion

V. Batagelj

Islands

Cores

Generalized
cores



V. Batagelj

Cohesion



p_S -core at level 46 of Geombib network

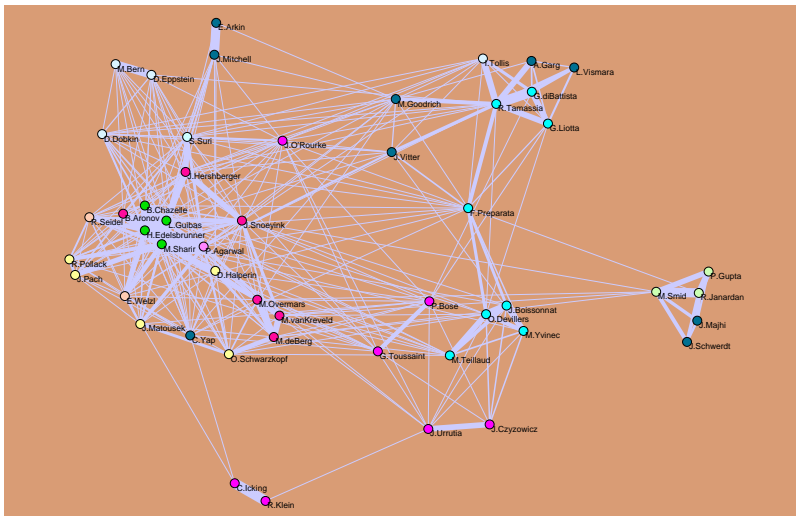
Cohesion

V. Batagelj

Islands

Cores

Generalized
cores



V. Batagelj

Cohesion