How the *k*-core decomposition helps in understanding the Internet Topology

J. Ignacio Alvarez-Hamelin[*,‡]
Luca Dall'Asta[*]
Alain Barrat[*]
Alessandro Vespignani[†,*]

[*] Laboratoire de Physique Théorique, Université Paris-Sud XI

[†] School of Informatics and Department of Physics, Indiana University

[‡] CONICET and Departamento de Electrónica, Universidad de Buenos Aires

Worshop on Internet Topology :: San Diego :: May 10-12

Outline

1. *k*-core decomposition
   - Definition
   - Examples
   - Applicability of *k*-cores
2. Applications
   - Properties analysis
   - Temporal analysis
   - Visualization

Outline

1. *k*-core decomposition
   - Definition
   - Examples
   - Applicability of *k*-cores
2. Applications
   - Properties analysis
   - Temporal analysis
   - Visualization

Outline

1. *k*-core decomposition
   - Definition
   - Examples
   - Applicability of *k*-cores
2. Applications
   - Properties analysis
   - Temporal analysis
   - Visualization

k-core decomposition
Definition

Given $G = \{V, E\}$ a **undirected** graph, where V is the vertices set and E is the edges set.

Definition (Seidman, 1983 [4]) :

A subgraph $H = (C, E|C)$ induced by the set $C \subseteq V$ is a k-core or a core of order k iff $\forall v \in C : \mathtt{degree}_H(v) \geq k$, and H is the maximum subgraph with this property.

Then, a **minimal degree** k is imposed to the core of order k.

k-core decomposition
Definition

Given $G = \{V, E\}$ a **undirected** graph, where V is the vertices set and E is the edges set.

Definition (Seidman, 1983 [4]) :

A subgraph $H = (C, E|C)$ induced by the set $C \subseteq V$ is a k-core or a core of order k iff $\forall v \in C : \mathtt{degree}_H(v) \geq k$, and H is the maximum subgraph with this property.

Then, a **minimal degree** k is imposed to the core of order k.

k-core decomposition
Definition

Given $G = \{V, E\}$ a **undirected** graph, where V is the vertices set and E is the edges set.

Definition (Seidman, 1983 [4]) :

A subgraph $H = (C, E|C)$ induced by the set $C \subseteq V$ is a k-core or a core of order k iff $\forall v \in C : \mathtt{degree}_H(v) \geq k$, and H is the maximum subgraph with this property.

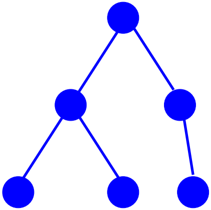
Then, a **minimal degree** k is imposed to the core of order k.

Outline

1. *k*-core decomposition
   - Definition
   - Examples
   - Applicability of *k*-cores
2. Applications
   - Properties analysis
   - Temporal analysis
   - Visualization

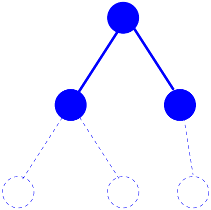
k-core decomposition
Examples

tree : 1-core



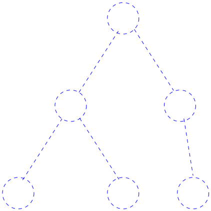
k-core decomposition
Examples

tree : 2-core ?

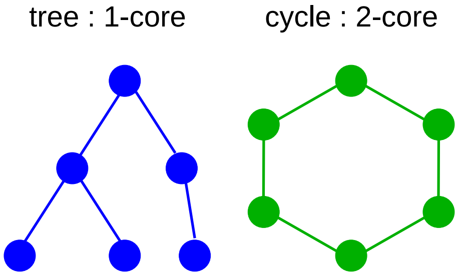


k-core decomposition
Examples

tree : ~~2-core~~ ?

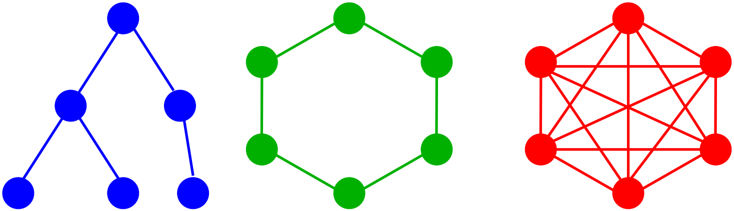


k-core decomposition
Examples

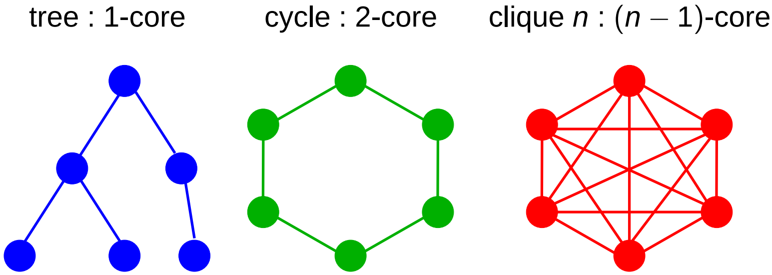


k-core decomposition
Examples

tree : 1-core — cycle : 2-core — clique n : $(n-1)$-core



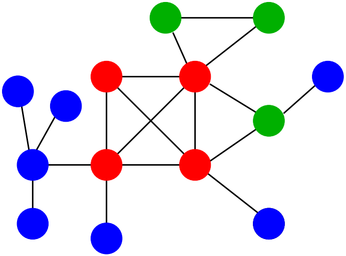
k-core decomposition
Examples



Remark : a *k*-connected graph **is** a *k*-core.

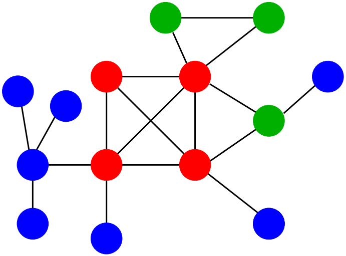
k-core decomposition
Examples

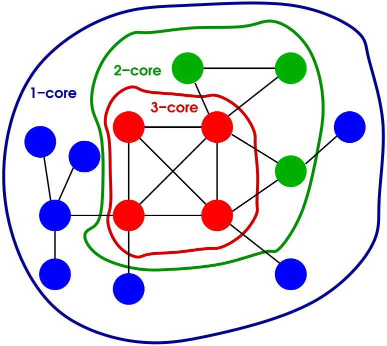
A graph :



k-core decomposition
Examples

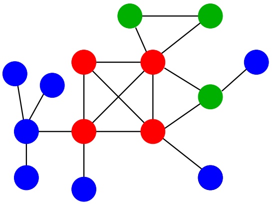
A graph :

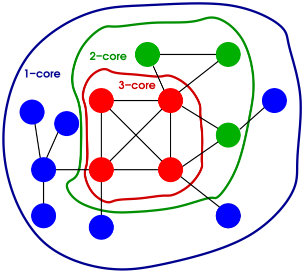




k-core decomposition
Examples

A graph :





Definition

A vertex i has a *shell index c*, if it belongs to the c-core but not to $(c+1)$-core.

Outline

1. *k*-core decomposition
   - Definition
   - Examples
   - Applicability of *k*-cores
2. Applications
   - Properties analysis
   - Temporal analysis
   - Visualization

Why the connectivity is interesting ?

k-core decomposition

The **connectivity** is mainly related to :

- robustness
  - faults,
  - attacks
- routing (to find a path between two vertices)
  - QoS,
  - efficiency (of several parameters).

Why the connectivity is interesting ?

k-core decomposition

The **connectivity** is mainly related to :

- robustness
  - faults,
  - attacks
- routing (to find a path between two vertices)
  - QoS,
  - efficiency (of several parameters).

Why the **connectivity** is interesting ?

k-core decomposition

The **connectivity** is mainly related to :

- robustness
  - faults,
  - attacks
- routing (to find a path between two vertices)
  - QoS,
  - efficiency (of several parameters).

relation between connectivity and *k*-cores ?

k-core decomposition

- *k*-connectivity ⇒ *k*-core

relation between connectivity and *k*-cores ?

k-core decomposition

- *k*-connectivity $\Rightarrow$ *k*-core

relation between connectivity and k-cores ?

k-core decomposition

- k-connectivity $\Rightarrow$ k-core
- k-core $\Rightarrow$ k-connectivity ?

relation between connectivity and *k*-cores ?

k-core decomposition

- *k*-connectivity ⇒ *k*-core
- ~~*k*-core ⇒ *k*-connectivity~~ ?

relation between connectivity and *k*-cores ?

k-core decomposition

- *k*-connectivity $\Rightarrow$ *k*-core
- Some results on *k*-core and *k*-connectivity :
  - Bollobas [1] : a **random graph** with large $< d >$ has a **minimum degree** *k* when $n \to \infty$, and it is *k*-connected.
  - Dimes project : study [2] the number of distinct paths as function of *k*-shell.
  - ***k*-edge-connectivity** : we show a theorem (k_{max}-core has diameter 2, and sets of vertices in *s*-shell are connected with at least *s* edges to higher cores).

relation between connectivity and k-cores ?

k-core decomposition

- k-connectivity $\Rightarrow$ k-core
- Some results on k-core and k-connectivity :
  - Bollobas [1] : a **random graph** with large $< d >$ has a **minimum degree** k when $n \to \infty$, and it is k-connected.
  - Dimes project : study [2] the number of distinct paths as function of k-shell.
  - k**-edge-connectivity** : we show a theorem (k_{max}-core has diameter 2, and sets of vertices in s-shell are connected with at least s edges to higher cores).

relation between connectivity and *k*-cores ?

k-core decomposition

- *k*-connectivity $\Rightarrow$ *k*-core
- Some results on *k*-core and *k*-connectivity :
  - Bollobas [1] : a **random graph** with large $< d >$ has a **minimum degree** *k* when $n \to \infty$, and it is *k*-connected.
  - Dimes project : study [2] the number of distinct paths as function of *k*-shell.
  - ***k*-edge-connectivity** : we show a theorem (k_{max}-core has diameter 2, and sets of vertices in *s*-shell are connected with at least *s* edges to higher cores).

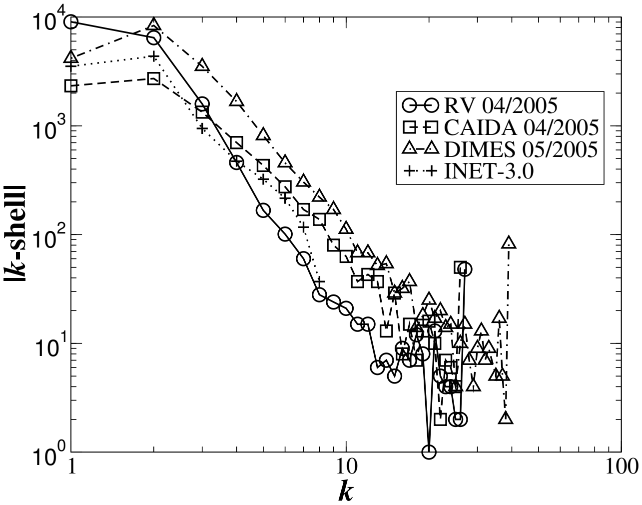
Outline

1. *k*-core decomposition
   - Definition
   - Examples
   - Applicability of *k*-cores
2. Applications
   - Properties analysis
   - Temporal analysis
   - Visualization

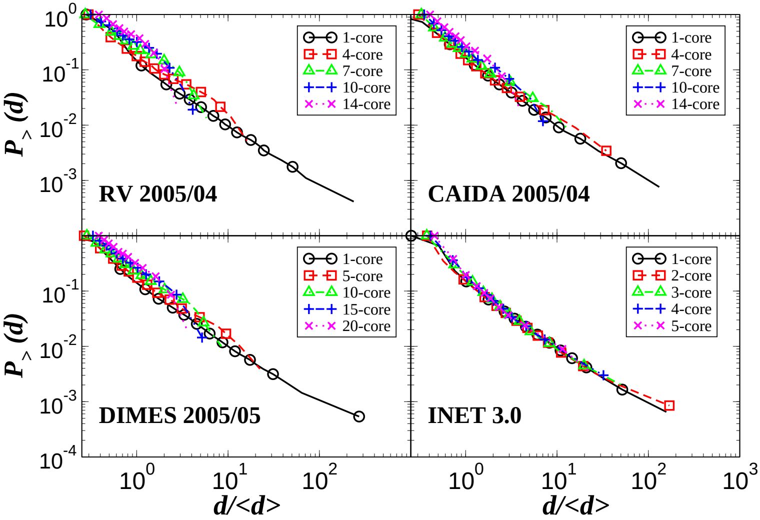
Outline

1. *k*-core decomposition
   - Definition
   - Examples
   - Applicability of *k*-cores
2. Applications
   - Properties analysis
   - Temporal analysis
   - Visualization

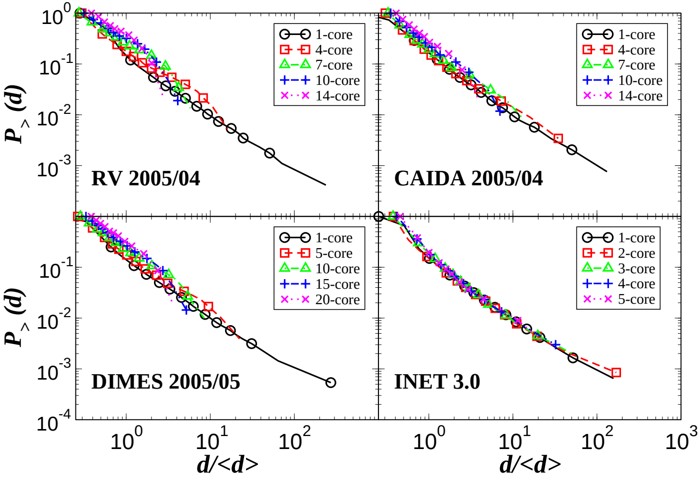
k-shell size vs *k*



cumulative degree distribution

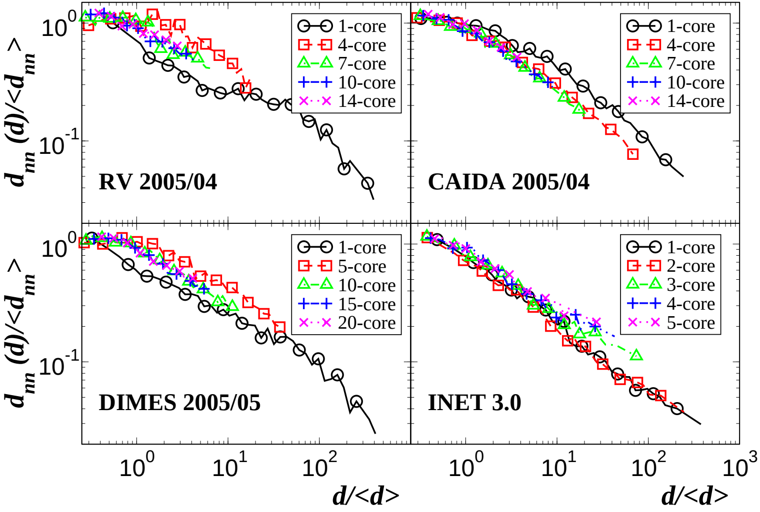


cumulative degree distribution

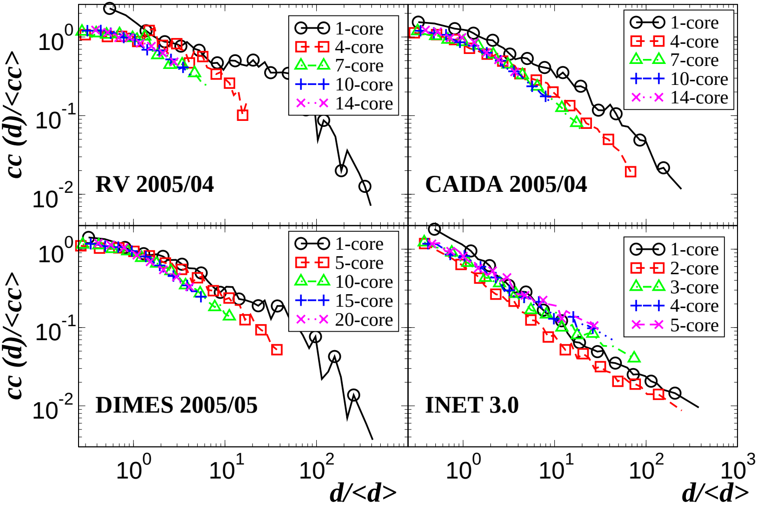


- **Self-similarity** of the *k*-**cores**

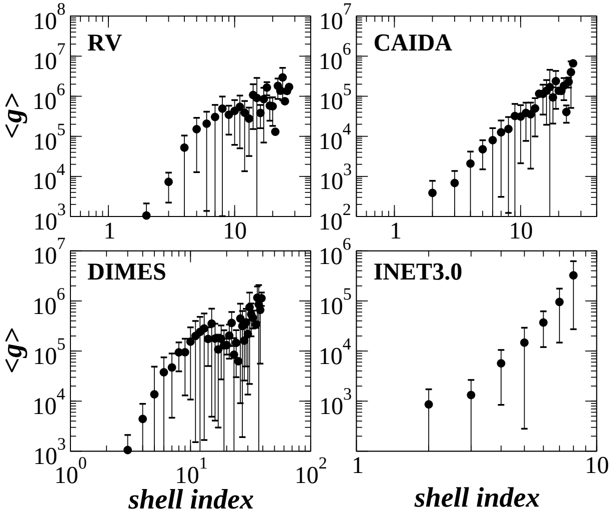
average neighbor degree vs degree



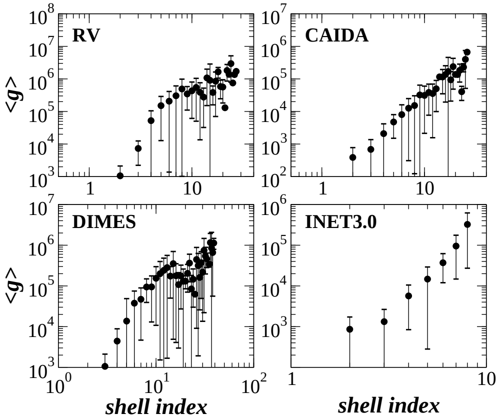
cluster coefficient vs degree



average betweenness vs shell index



average betweenness vs shell index

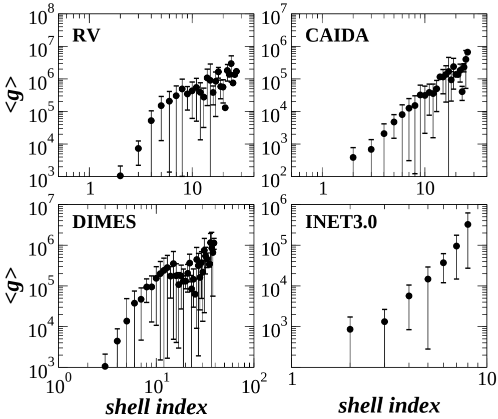


betweenness : $\mathcal{O}(n^2 \log n)$
k-core : $\mathcal{O}(e)$

k-**core** decomposition *establish* a **hierarchy** in the network :

- it can *identify* the vertex's **centrality**.

average betweenness vs shell index



betweenness : $\mathcal{O}(n^2 \log n)$
k-core : $\mathcal{O}(e)$

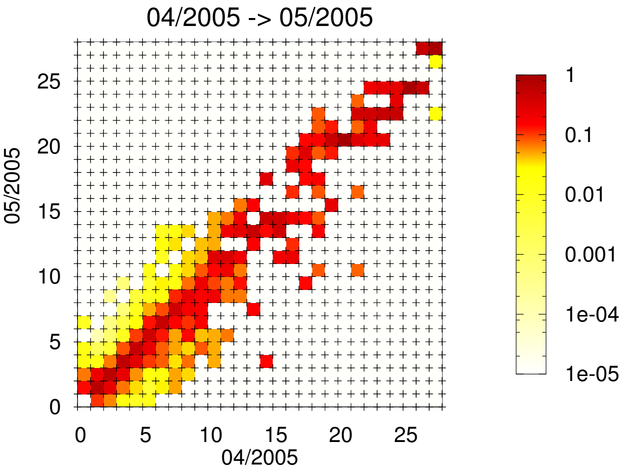
***k*-core** decomposition *establish* a **hierarchy** in the network :

- it can *identify* the vertex's **centrality**.

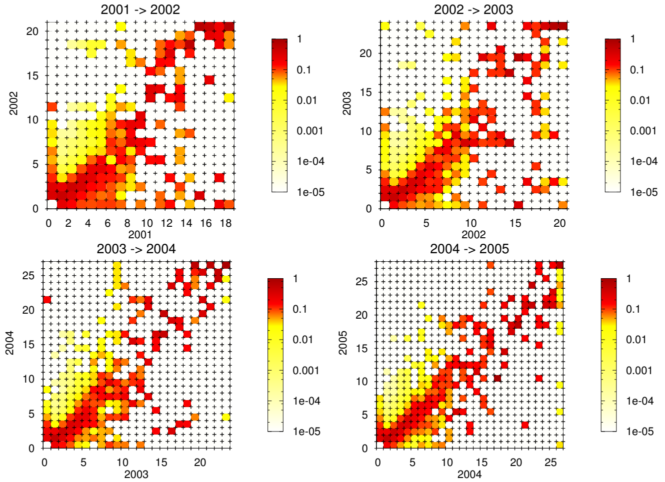
Outline

1. *k*-core decomposition
   - Definition
   - Examples
   - Applicability of *k*-cores
2. Applications
   - Properties analysis
   - Temporal analysis
   - Visualization

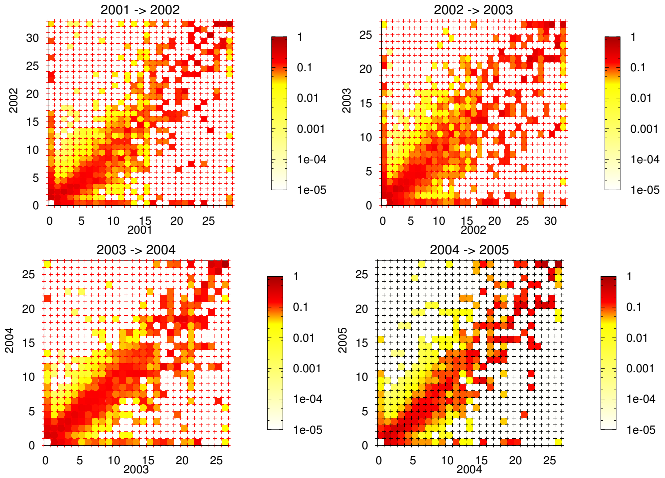
AS Oregon R. V. : april -> may 2005



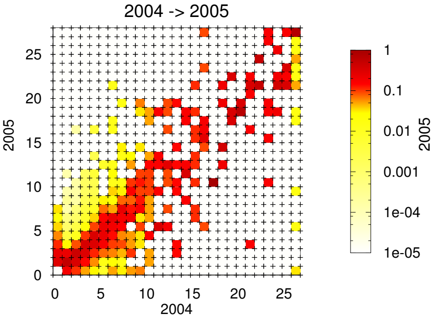
AS maps of Oregon R. V.



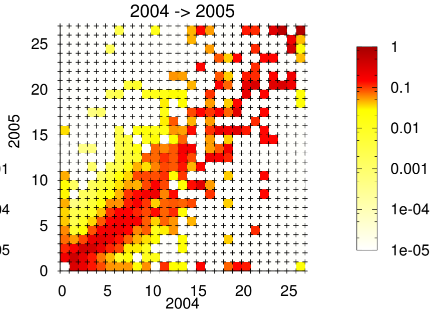
AS maps of CAIDA



AS maps : 2004->2005



Oregon_RV

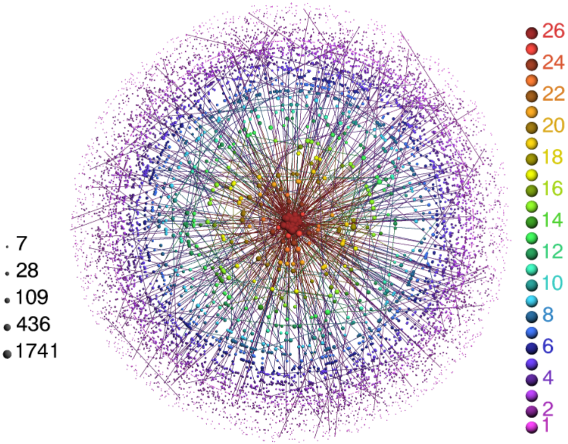


CAIDA

Outline

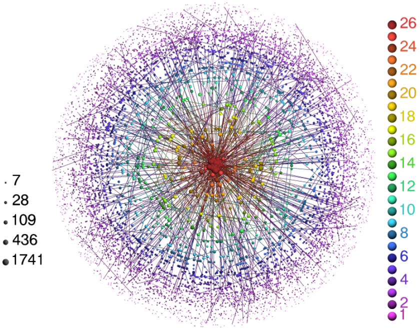
1. *k*-core decomposition
   - Definition
   - Examples
   - Applicability of *k*-cores
2. Applications
   - Properties analysis
   - Temporal analysis
   - Visualization

Network **fingerprints**

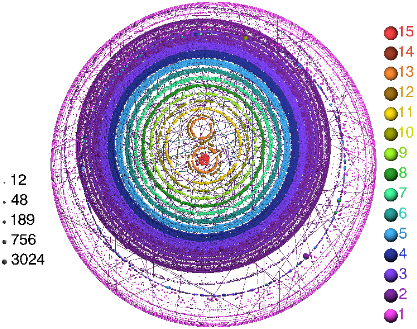


CAIDA AS map

Network **fingerprints**

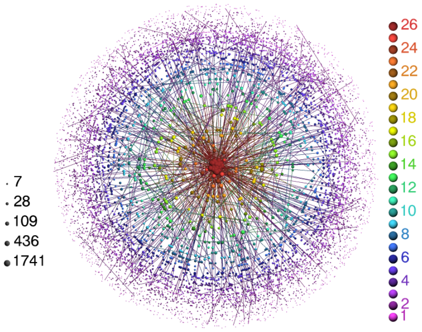


CAIDA AS map

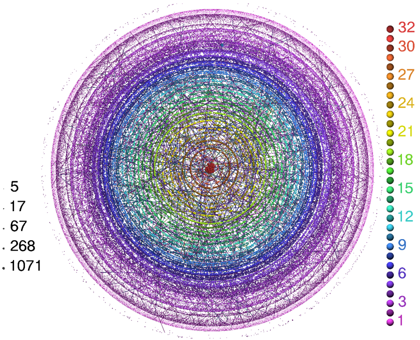


a portion of www.***.fr

Network **fingerprints**

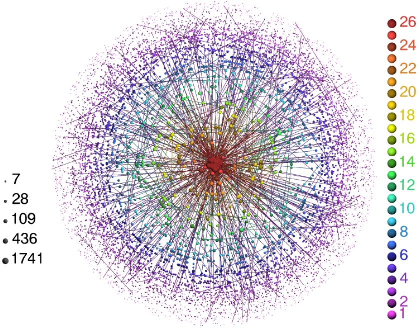


CAIDA AS map

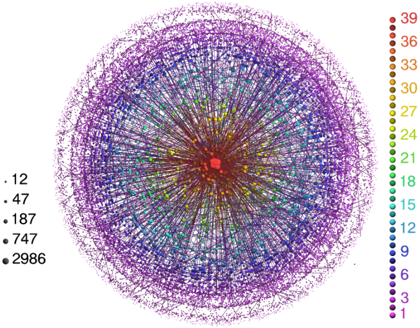


CAIDA Router level map

Network **fingerprints**



CAIDA AS map



Dimes AS map

Summary

- Using *k*-core decomposition, we can determine :
  - **a <u>hierarchical</u> and <u>self-similar</u> structure**
  - **centrality**
  - **connectivity**
- **Comparative** analysis of **different** maps
- **Fingerprints** (Visualization)
  http ://xavier.informatics.indiana.edu/lanet-vi/
- http ://arxiv.org/abs/cs.NI/0504107/
  (Large scale networks fingerprinting and visualization using the k-core decomposition, Advances in Neural Information Processing Systems 18, 2006),
  and http ://arxiv.org/abs/cs.NI/0511007/

Summary

- Using *k*-core decomposition, we can determine :
  - **a hierarchical and self-similar structure**
  - **centrality**
  - **connectivity**
- **Comparative** analysis of **different** maps
- **Fingerprints** (Visualization)
  http ://xavier.informatics.indiana.edu/lanet-vi/
- http ://arxiv.org/abs/cs.NI/0504107/
  (Large scale networks fingerprinting and visualization using the k-core decomposition, Advances in Neural Information Processing Systems 18, 2006),
  and http ://arxiv.org/abs/cs.NI/0511007/

Summary

- Using *k*-core decomposition, we can determine :
  - **a hierarchical and self-similar structure**
  - **centrality**
  - **connectivity**
- **Comparative** analysis of **different** maps
- **Fingerprints** (Visualization)
  http ://xavier.informatics.indiana.edu/lanet-vi/
- http ://arxiv.org/abs/cs.NI/0504107/
  (Large scale networks fingerprinting and visualization using the k-core decomposition, Advances in Neural Information Processing Systems 18, 2006),
  and http ://arxiv.org/abs/cs.NI/0511007/

Summary

- Using *k*-core decomposition, we can determine :
  - **a hierarchical and self-similar structure**
  - **centrality**
  - **connectivity**
- **Comparative** analysis of **different** maps
- **Fingerprints** (Visualization)
  http ://xavier.informatics.indiana.edu/lanet-vi/
- http ://arxiv.org/abs/cs.NI/0504107/
  (Large scale networks fingerprinting and visualization using the k-core decomposition, Advances in Neural Information Processing Systems 18, 2006),
  and http ://arxiv.org/abs/cs.NI/0511007/

Summary

- Using *k*-core decomposition, we can determine :
  - **a <u>hierarchical</u> and <u>self-similar</u> structure**
  - **centrality**
  - **connectivity**
- **Comparative** analysis of **different** maps
- **Fingerprints** (Visualization)
  http ://xavier.informatics.indiana.edu/lanet-vi/
- http ://arxiv.org/abs/cs.NI/0504107/
  (Large scale networks fingerprinting and visualization using the k-core decomposition, Advances in Neural Information Processing Systems 18, 2006),
  and http ://arxiv.org/abs/cs.NI/0511007/

Summary

- Using *k*-core decomposition, we can determine :
  - **a hierarchical and self-similar structure**
  - **centrality**
  - **connectivity**
- **Comparative** analysis of **different** maps
- **Fingerprints** (Visualization)
  http ://xavier.informatics.indiana.edu/lanet-vi/
- http ://arxiv.org/abs/cs.NI/0504107/
  (Large scale networks fingerprinting and visualization using the k-core decomposition, Advances in Neural Information Processing Systems 18, 2006),
  and http ://arxiv.org/abs/cs.NI/0511007/

B. Bollobasr.
Graph Theory and Combinatorics, conference in honor of Paul Erdos, Academic Press, 35–57, 1983.

Shai Carmi, Shlomo Havlin1, Scott Kirkpatrick, Yuval Shavitt and Eran Shir.
MEDUSA – New Model of Internet Topology Using k-shell Decomposition
arXiv, cond-mat/0601240, January 2006.

V. Batagelj and M. Zaversnik.
Generalized Cores.
CoRR, cs.DS/0202039, 2002.

S. B. Seidman.
Network structure and minimum degree.
Social Networks 5, 269–287, 1983.

LArge NETwork VIsualization tool.
`http ://xavier.informatics.indiana.edu/lanet-vi/`.

Router-Level Topology Measurements" "Cooperative Association for Internet Data Analysis.
`http ://www.caida.org/tools/measurement/skitter/ router_topology/`.

University of Oregon Route Views Project.
`http ://www.routeviews.org/`.

"Distributed Internet MEasurements and Simulations".
`http ://www.netdimes.org`.

Jared Winick and Sugih Jamin.
Inet-3.0 : Internet topology generator.
Technical Report UM-CSE-TR-456-02, Department of EECS, University of Michigan, 2002.

J. I. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani.
k-core decomposition : a tool for the visualization of large scale networks.
arxiv.org, cs.NI/0504107, 2005.

M. A. Serrano, M. Boguña and A. Díaz-Guilera.
Competition and Adaptation in an Internet Evolution Model.
Phys. Rev. Letters, 94 :038701, 2005.

J. I. Alvarez-Hamelin and N. Schabanel.
An Internet Graph Model Based on Trade-Off Optimization.
Eur. Phys. J. B, special issue on "Applications of networks", 38(2) :231–237, march II 2004.

P. Erdös and A. Rényi.
On random graphs I.
Publ. Math. (Debrecen), 6 :290–297, 1959.

A. Fabrikant, E. Koutsoupias, and C. H. Papadimitriou.
Heuristically Optimized Trade-Offs : A New Paradigm for Power Laws in the Internet.
LNCS, 2380 :110–, Jun 2002.

k-edge-connectivity

- k_{max}-core is *k*-edge-connected (or diameter 2)
- cluster in *s*-shell is *s*-edge-connected and it has at least *s* edges towards $(s+1)$-core.
- then, the graph is *k*-edge-connected