

Enhanced extraction of weighted networks' backbones

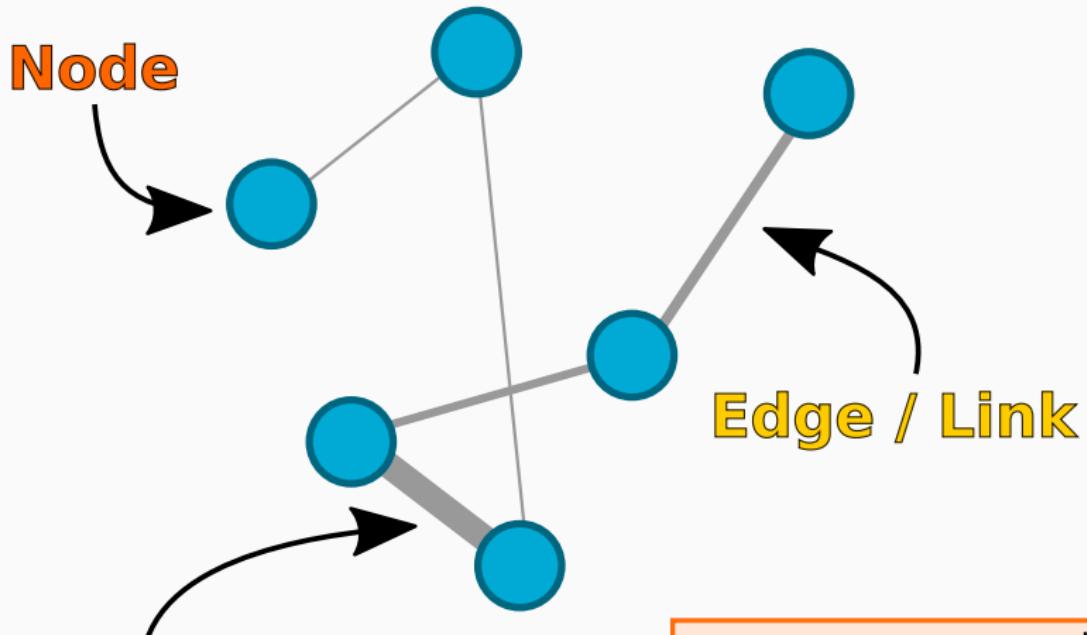
Alessio Cardillo (@a_cardillo)

Dept. Computer Science & Mathematics – Universitat Rovira i Virgili
Tarragona (Spain)

MPI-PKS – Dresden (Germany)
Wednesday, 11th March 2020



UNIVERSITAT ROVIRA i VIRGILI



degree = number of interactions
strength = sum of weights



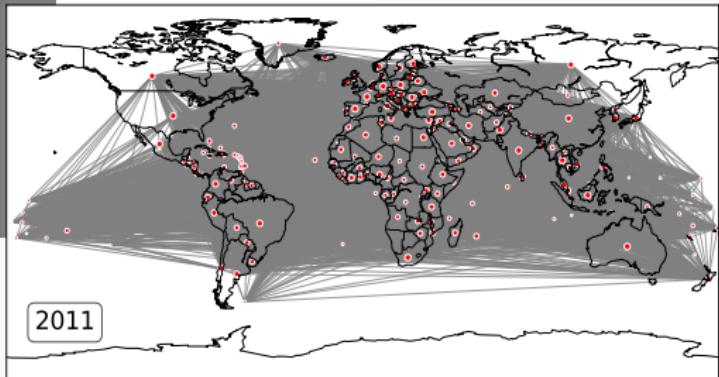
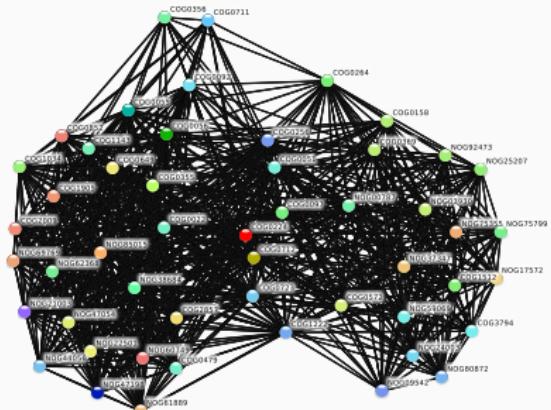
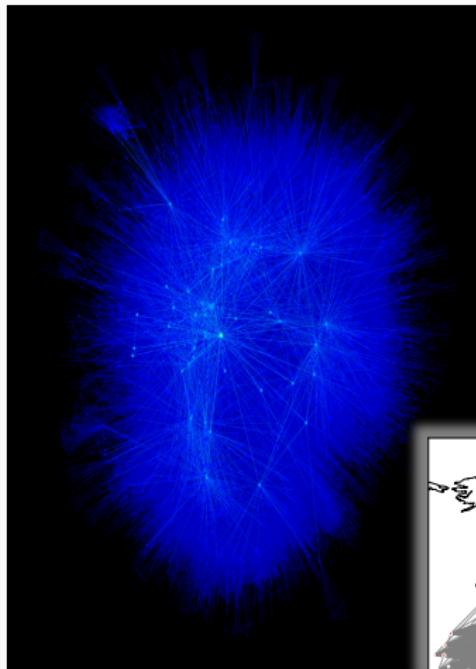
4cm

R

Motivation



Motivation



Motivation

Question:

What can we **learn** from a complex system
whose **network representation** is **noisy** and/or
extremely **dense**?

BRACE YOURSELVES



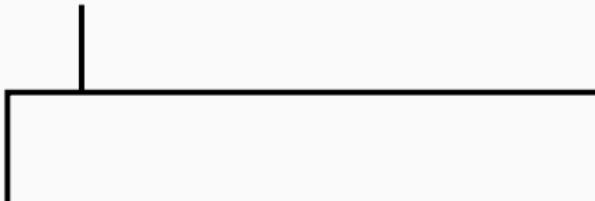
FILTERING IS COMING

- Motivation.
- ★ Taxonomy of filtering
- ★ The ECM-filter
- ★ Results
- Take home messages
- Questions

Taxonomy of filtering

Network Filtering

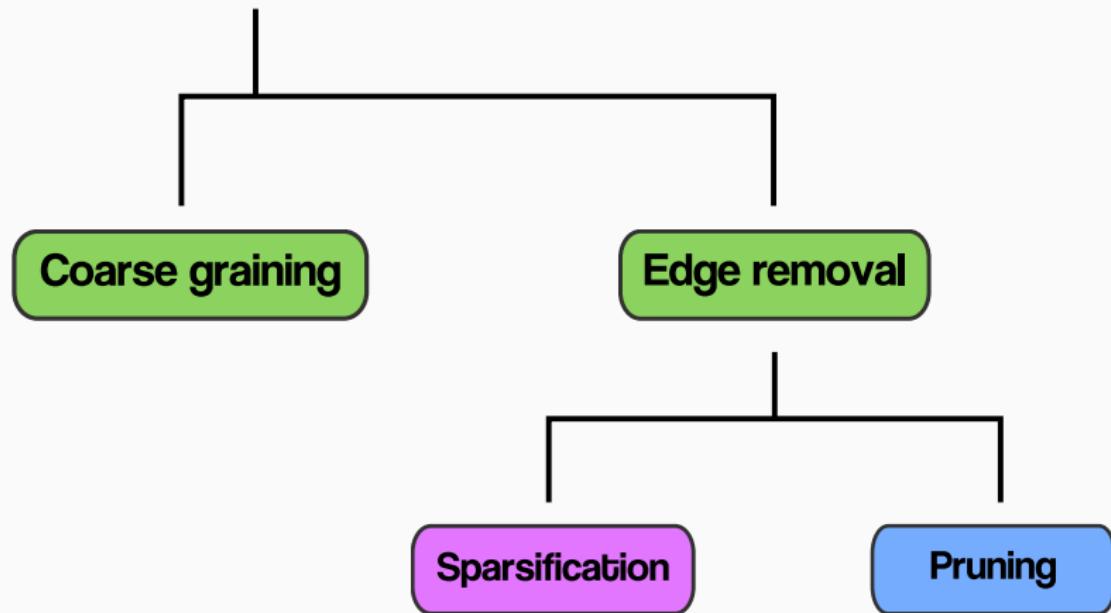
Network Filtering



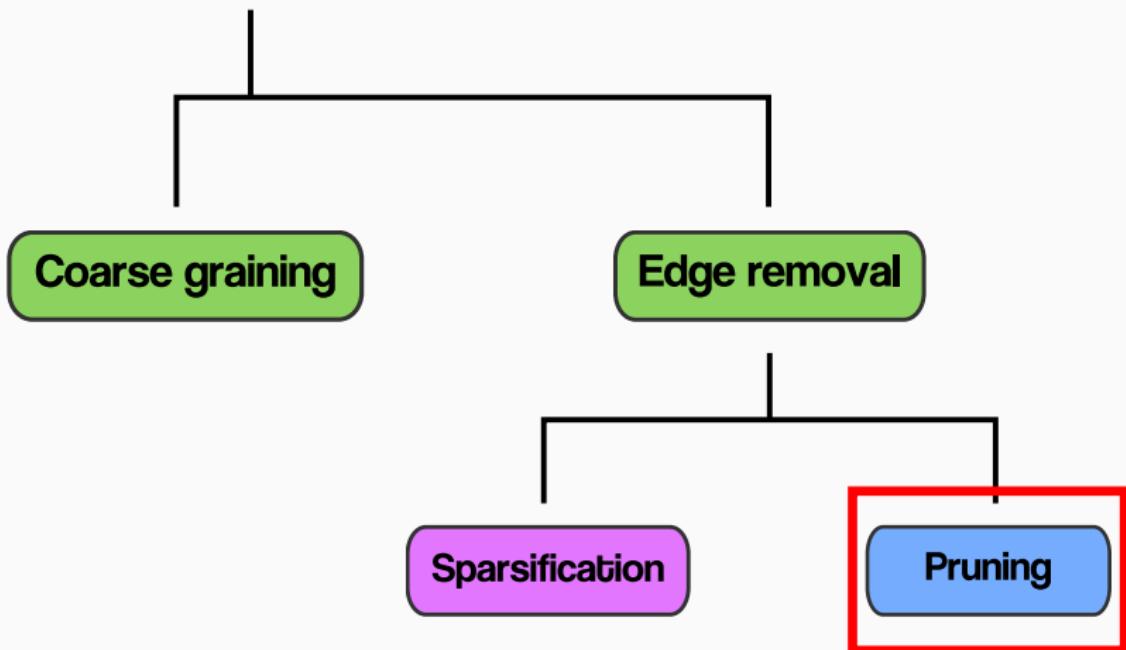
Coarse graining

Edge removal

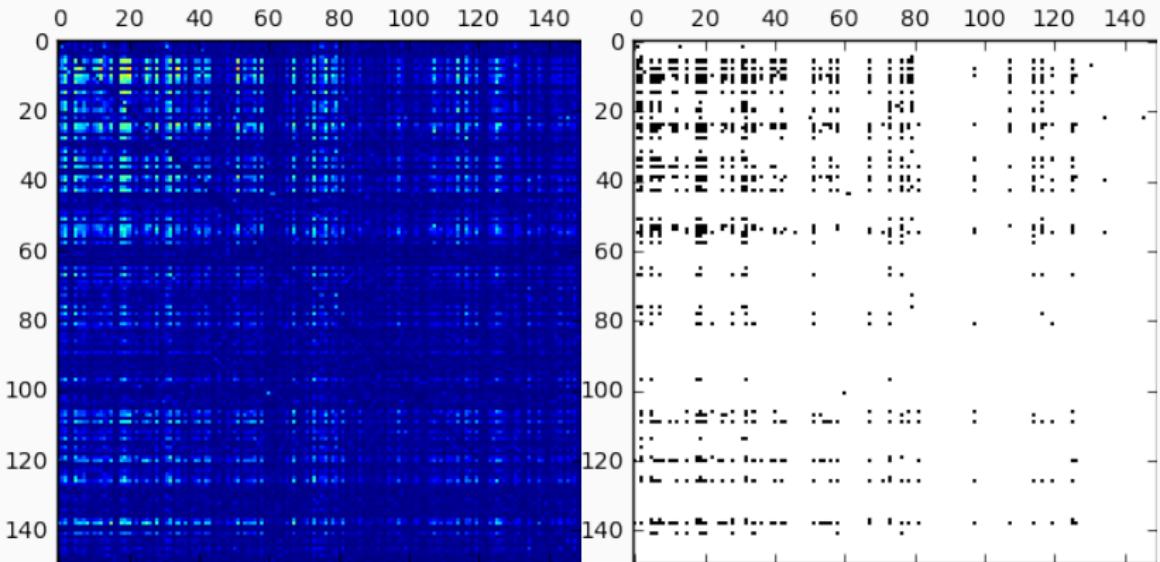
Network Filtering



Network Filtering



Thresholding



Thresholding

The screenshot shows the Proceedings of the National Academy of Sciences of the United States of America (PNAS) website. The main navigation bar includes Home, Articles (which is the active tab), Front Matter, News, Podcasts, and Authors. A search bar at the top right is labeled "Keyword, Author". Below the navigation, a section titled "NEW RESEARCH IN" lists "Physical Sciences" and "Social Sciences". The main content features a large title: "A tool for filtering information in complex systems". Below the title, the authors are listed as M. Tumminello, T. Aste, T. Di Matteo, and R. N. Mantegna. The publication details are PNAS July 26, 2005 102 (30) 10421-10426; <https://doi.org/10.1073/pnas.0500298102>. The article was edited by H. Eugene Stanley, Boston University, Boston, MA (received for review January 13, 2005). A small "Check for updates" icon is visible next to the title.

PNAS

Proceedings of the
National Academy of Sciences
of the United States of America

Keyword, Author

Home Articles Front Matter News Podcasts Authors

NEW RESEARCH IN Physical Sciences Social Sciences

A tool for filtering information in complex systems

M. Tumminello, T. Aste, T. Di Matteo, and R. N. Mantegna

PNAS July 26, 2005 102 (30) 10421-10426; <https://doi.org/10.1073/pnas.0500298102>

Edited by H. Eugene Stanley, Boston University, Boston, MA (received for review January 13, 2005)

- Tumminello, M., et al. Proc. Nat. Acad. Sci., **102**, 10421–10426 (2005).

Thresholding

The screenshot shows a web page for a research article. At the top left is the PLOS logo with 'COMPUTATIONAL BIOLOGY' next to it. To the right are links for 'BROWSE', 'PUBLISH', and 'ABC'. Below the header, there are badges for 'OPEN ACCESS' and 'PEER-REVIEWED'. The article title is 'A Topological Criterion for Filtering Information in Complex Brain Networks'. Below the title, the authors listed are Fabrizio De Vico Fallani, Vito Latora, and Mario Chavez. A dropdown menu indicates 'Version 2'. The publication date is January 11, 2017, and the DOI is <https://doi.org/10.1371/journal.pcbi.1005305>.

PLOS COMPUTATIONAL BIOLOGY

BROWSE PUBLISH ABC

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

A Topological Criterion for Filtering Information in Complex Brain Networks

Fabrizio De Vico Fallani, Vito Latora, Mario Chavez

Version 2

Published: January 11, 2017 • <https://doi.org/10.1371/journal.pcbi.1005305>

- De Vico Fallani F., Latora V., & Chavez M. PLoS Comp. Bio. **13** e1005305 (2017).

PHYSICAL REVIEW E

covering statistical, nonlinear, biological, and soft matter physics

[Highlights](#)[Recent](#)[Accepted](#)[Collections](#)[Authors](#)[Referees](#)[Search](#)[Press](#)

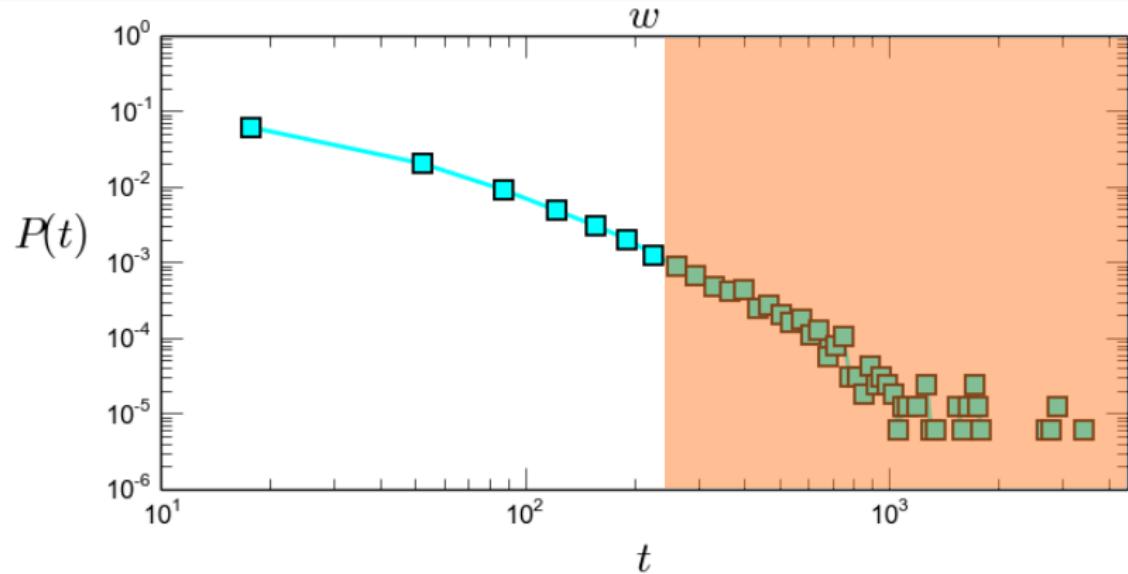
Weight thresholding on complex networks

Xiaoran Yan, Lucas G. S. Jeub, Alessandro Flammini, Filippo Radicchi, and Santo Fortunato
Phys. Rev. E **98**, 042304 – Published 8 October 2018

- Yan, X., et al. *Weight thresholding on complex networks* Phys. Rev. E **98**, 042304 (2018).



Thresholding



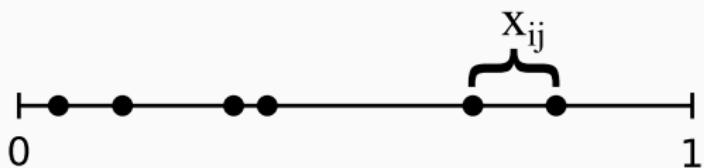
- Granovetter, M. S. *The Strength of Weak Ties*. Am. Jour. Soc., **78**, 1360 (1973).

Disparity/Pólya Filter

The screenshot shows the Proceedings of the National Academy of Sciences of the United States of America (PNAS) website. The main navigation bar includes links for Home, Articles (which is the active tab), Front Matter, News, Podcasts, and Authors. A search bar at the top right contains the placeholder "Keyword, Author, or phrase". Below the navigation, a "NEW RESEARCH IN" section is displayed, with "Physical Sciences" selected from a dropdown menu. Another dropdown menu next to it shows "Social Sciences". The main content area features a large, bold title: "Extracting the multiscale backbone of complex weighted networks". To the right of the title is a small "Check for updates" button with a circular arrow icon. Below the title, the authors' names are listed: M. Ángeles Serrano, Marián Boguñá, and Alessandro Vespignani. The publication details indicate the article was published in PNAS on April 21, 2009, volume 106, issue 16, pages 6483-6488, with a DOI of <https://doi.org/10.1073/pnas.0808904106>.

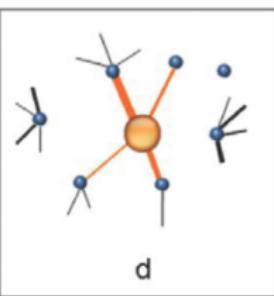
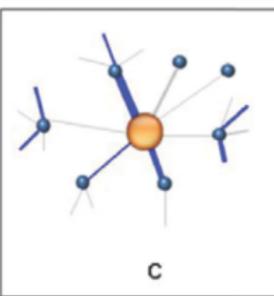
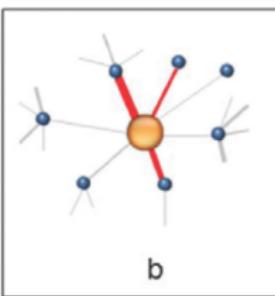
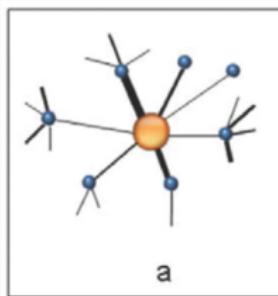
- Serrano M.A., Boguña M., & Vespignani A. Proc. Natl. Acad. Sci. (USA) **106** 6483 (2009).

Disparity/Pólya Filter



disparity

$$x_{ij} = \frac{w_{ij}}{\sum_j w_{ij}} \text{ or } \frac{w_{ij}}{\sum_i w_{ij}}$$



- Serrano M.A., Boguña M., & Vespignani A. Proc. Natl. Acad. Sci. (USA) **106** 6483 (2009).

Disparity/Pólya Filter

Pros:

- Comparison with null-hypothesis
- Easy to implement
- Computationally fast

Cons:

- Not Maxent
- Local hypothesis
- No unique p -value
- Bias towards heavier connections

- Serrano M.A., Boguña M., & Vespignani A. Proc. Natl. Acad. Sci. (USA) **106** 6483 (2009).



Disparity/Pólya Filter

The screenshot shows a digital journal cover for 'nature COMMUNICATIONS'. At the top left is a dropdown menu icon. The title 'nature COMMUNICATIONS' is displayed with a stylized orange and red wavy graphic next to it. Below the title, the word 'Article' is followed by 'OPEN' in red and 'Published: 14 February 2019' in blue. The main title of the article is 'A Pólya urn approach to information filtering in complex networks', written in large, dark font. Below the title, the authors' names are listed: 'Riccardo Marcaccioli & Giacomo Livan' with an envelope icon. At the bottom, the journal name 'Nature Communications', volume '10', article number '745', and year '(2019)' are shown, along with 'Download Citation' and a downward arrow icon.

- Marcaccioli, R., & Livan, G. *Nature Communications*, **10**, 745 (2019).

PHYSICAL REVIEW E

covering statistical, nonlinear, biological, and soft matter physics

[Highlights](#)[Recent](#)[Accepted](#)[Collections](#)[Authors](#)[Referees](#)[Search](#)[Press](#)[About](#)

Information filtering in complex weighted networks

Filippo Radicchi, José J. Ramasco, and Santo Fortunato

Phys. Rev. E **83**, 046101 – Published 1 April 2011

- Radicchi F., Ramasco J.J., & Fortunato S. Physical Review E, **83** 046101 (2011).



Pros:

- Comparison between real network and a null-model
- Unique p -value assigned to edges

Cons:

- Not Maxent
- Hard/strict constraints (topology + weight distro.)
- Very aggressive (limitations on p -value)

- Radicchi F., Ramasco J.J., & Fortunato S. Physical Review E, **83** 046101 (2011).

PHYSICAL REVIEW E

covering statistical, nonlinear, biological, and soft matter physics

[Highlights](#)[Recent](#)[Accepted](#)[Collections](#)[Authors](#)[Referees](#)[Search](#)[Press](#)[About](#)

Unwinding the hairball graph: Pruning algorithms for weighted complex networks

Navid Dianati

Phys. Rev. E **93**, 012304 – Published 11 January 2016

- Dianati, N. Phys. Rev. E, **93**, 012304 (2016).



Pros:

- Comparison between real network and a null-model based on Maxent
- Two versions: global and local

Cons:

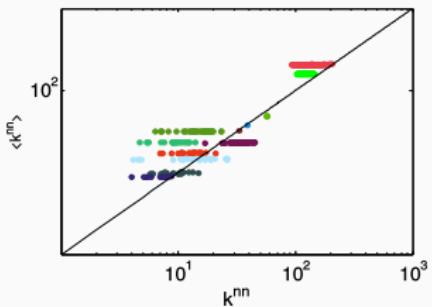
- Maxent based on conservation of $\{s_i\}$

- Dianati, N. Phys. Rev. E, **93**, 012304 (2016).

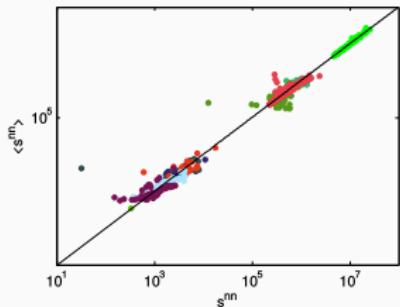
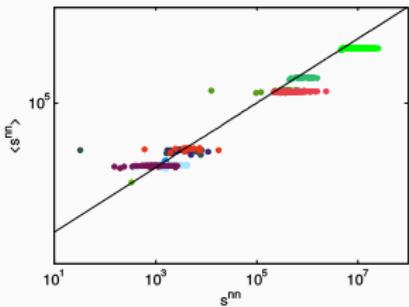
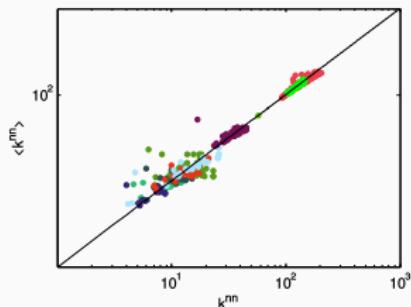
The ECM-filter

The Enhanced Configuration Model (ECM)

degree OR strength



degree AND strength



- Mastrandrea, R., et al. *Enhanced reconstruction of . . .* New Jour. Phys., **16**, 043022. (2014).

Main Features

1. Based on the comparison between the observed network and a null model one.

- Mastrandrea, R., et al. *Enhanced reconstruction of . . .* New Jour. Phys., **16**, 043022. (2014).

Main Features

1. Based on the comparison between the observed network and a null model one.
2. Null model: maximum-entropy canonical ensemble of networks satisfying given constraints.

Main Features

1. Based on the comparison between the observed network and a null model one.
2. Null model: maximum-entropy canonical ensemble of networks satisfying given constraints.
3. Constraints: $\{k_i\}$ and $\{s_i\}$ preserved (on average).

Main Features

1. Based on the comparison between the observed network and a null model one.
2. Null model: maximum-entropy canonical ensemble of networks satisfying given constraints.
3. Constraints: $\{k_i\}$ and $\{s_i\}$ preserved (on average).
4. Two versions: **local** (focus on links) and **global** (focus on entire network).

Local filter

1. Generate the null model networks' ensemble and compute:

$$q_{ij}(w) \equiv \frac{(x_i x_j)^{\Theta(w_{ij})} (y_i y_j)^{w_{ij}} (1 - y_i y_j)}{1 - y_i y_j + x_i x_j y_i y_j}.$$

Local filter

1. Generate the null model networks' ensemble and compute:

$$q_{ij}(w) \equiv \frac{(x_i x_j)^{\Theta(w_{ij})} (y_i y_j)^{w_{ij}} (1 - y_i y_j)}{1 - y_i y_j + x_i x_j y_i y_j}.$$

2. Compute the probability:

$$p_{ij}(w^*) = 1 - \sum_{w=0}^{w^*-1} q_{ij}(w).$$

Local filter

1. Generate the null model networks' ensemble and compute:

$$q_{ij}(w) \equiv \frac{(x_i x_j)^{\Theta(w_{ij})} (y_i y_j)^{w_{ij}} (1 - y_i y_j)}{1 - y_i y_j + x_i x_j y_i y_j}.$$

2. Compute the probability:

$$p_{ij}(w^*) = 1 - \sum_{w=0}^{w^*-1} q_{ij}(w).$$

3. Associate for each link a p -value, γ , such that $P(w_{ij} > w_{ij}^*)$.

Local filter

1. Generate the null model networks' ensemble and compute:

$$q_{ij}(w) \equiv \frac{(x_i x_j)^{\Theta(w_{ij})} (y_i y_j)^{w_{ij}} (1 - y_i y_j)}{1 - y_i y_j + x_i x_j y_i y_j}.$$

2. Compute the probability:

$$p_{ij}(w^*) = 1 - \sum_{w=0}^{w^*-1} q_{ij}(w).$$

3. Associate for each link a p -value, γ , such that $P(w_{ij} > w_{ij}^*)$.
4. Select a threshold $\tilde{\gamma}$ and remove all the links with $\gamma_{ij} > \tilde{\gamma}$.

Global filter

1. Find the subgraph Σ with L' links that minimizes the **likelihood** of being generated by chance.

Global filter

1. Find the subgraph Σ with L' links that minimizes the **likelihood** of being generated by chance.
2. Minimize the function:

$$P(\Sigma) = \prod_{i < j} [q_{ij}(w_{ij})]^{a_{ij}} .$$

Global filter

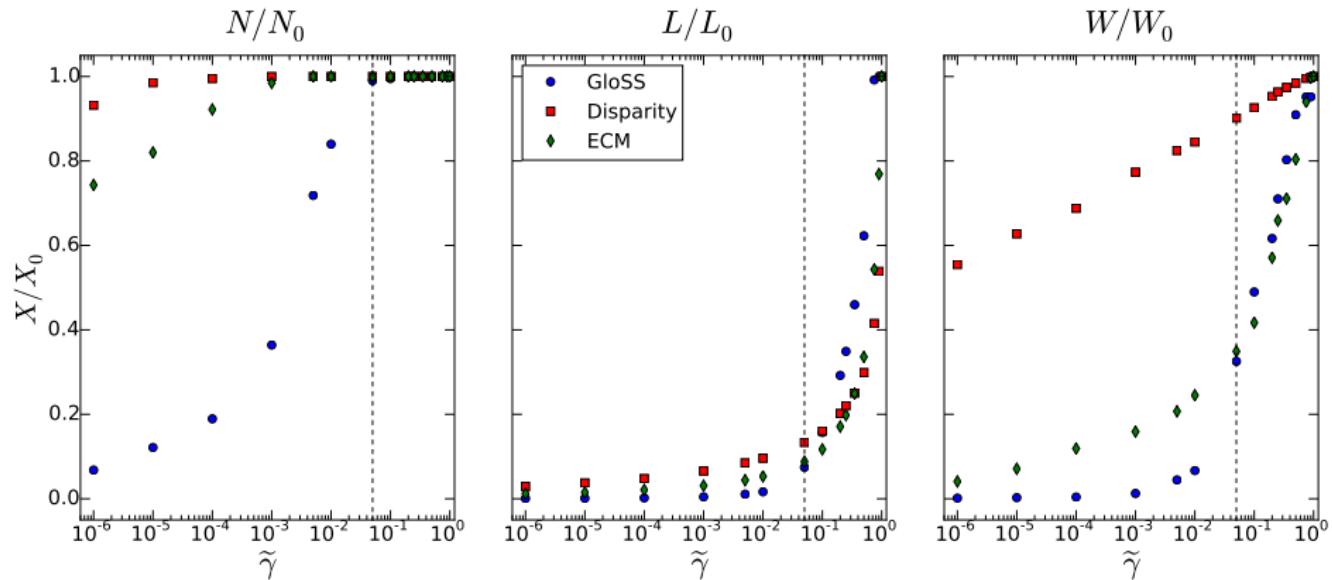
1. Find the subgraph Σ with L' links that minimizes the **likelihood** of being generated by chance.
2. Minimize the function:

$$P(\Sigma) = \prod_{i < j} [q_{ij}(w_{ij})]^{a_{ij}} .$$

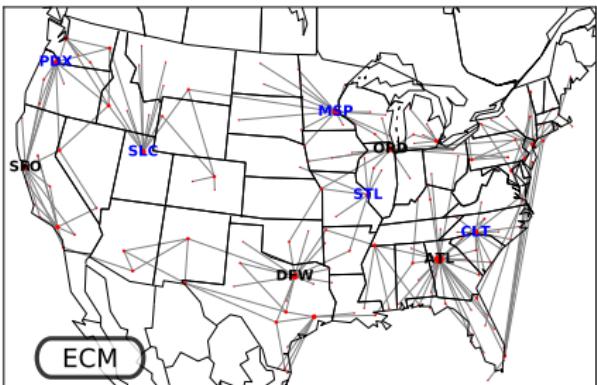
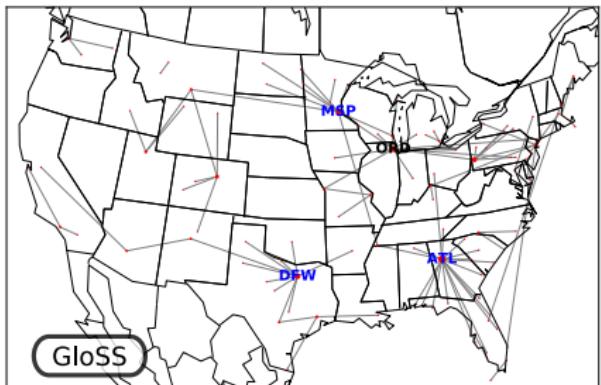
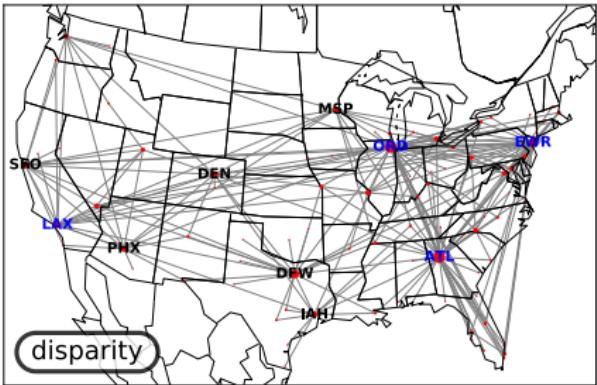
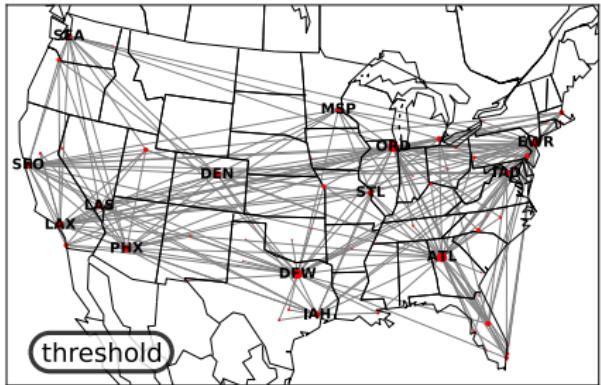
3. Rank edges upon their $q_{ij}(w)$ and add the L' smallest ones.

Results

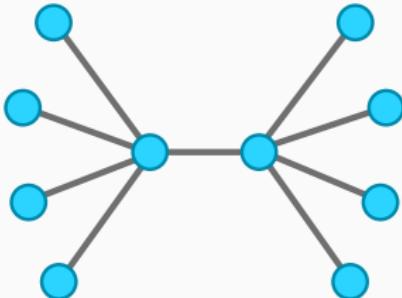
Comparison among methods



Comparison among methods

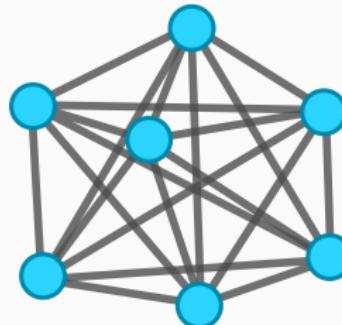


Comparison among methods



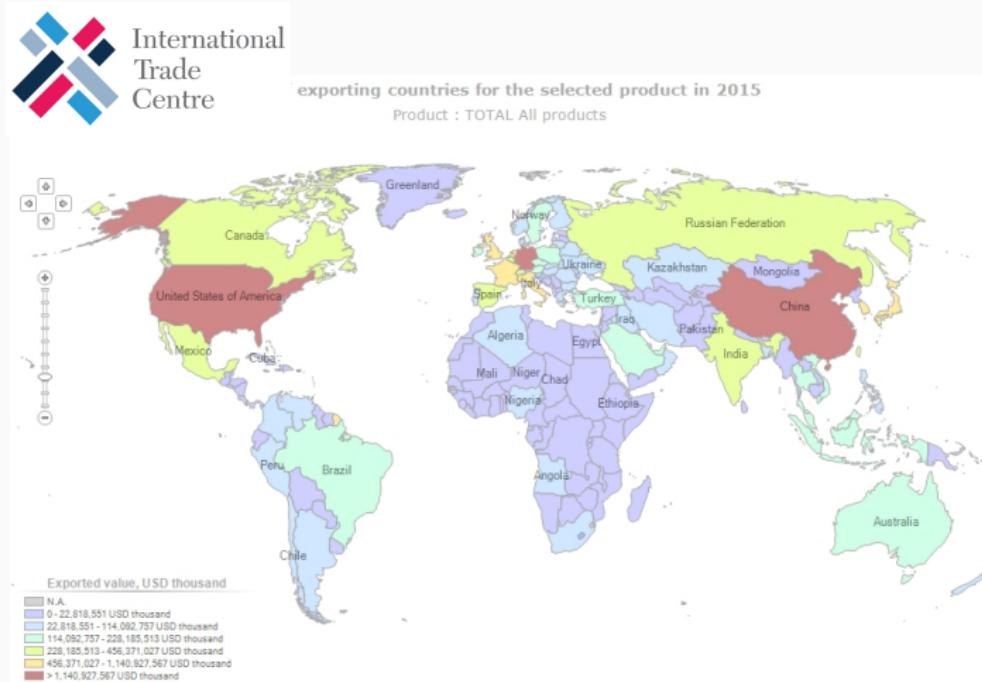
Hub and Spoke

Point to Point

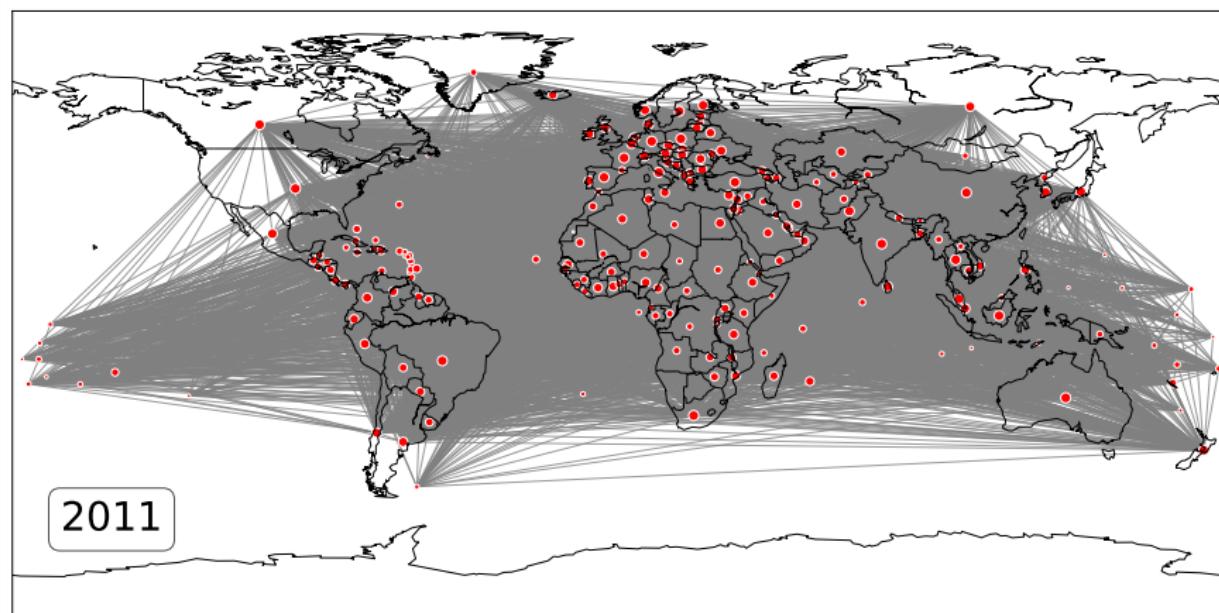


An example

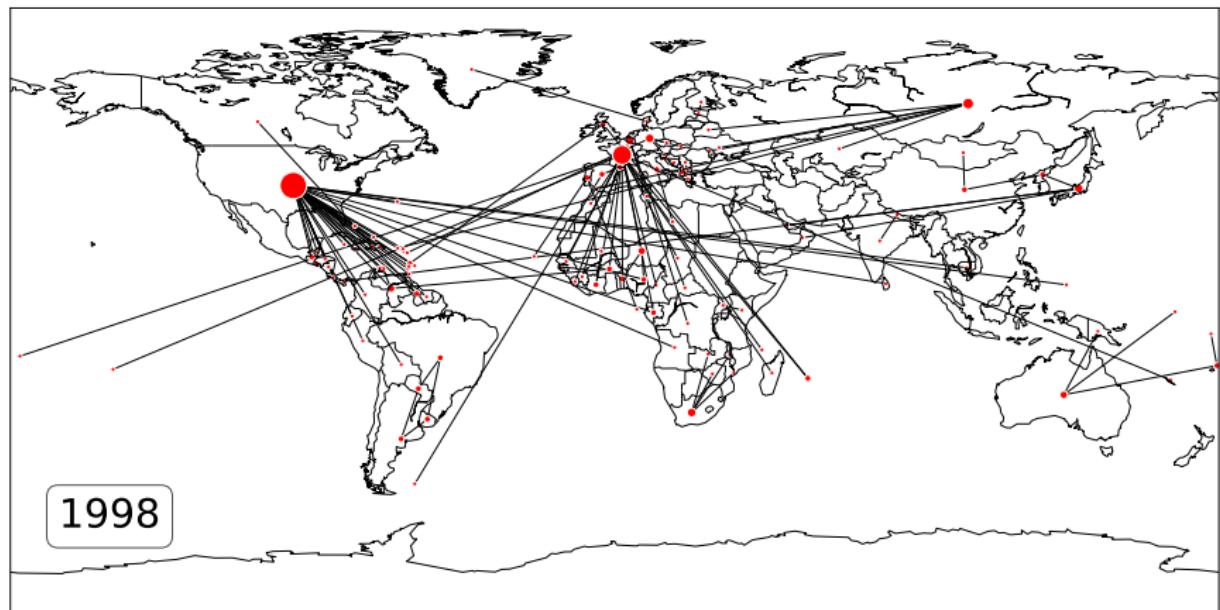
International Trade Network 1998 – 2011



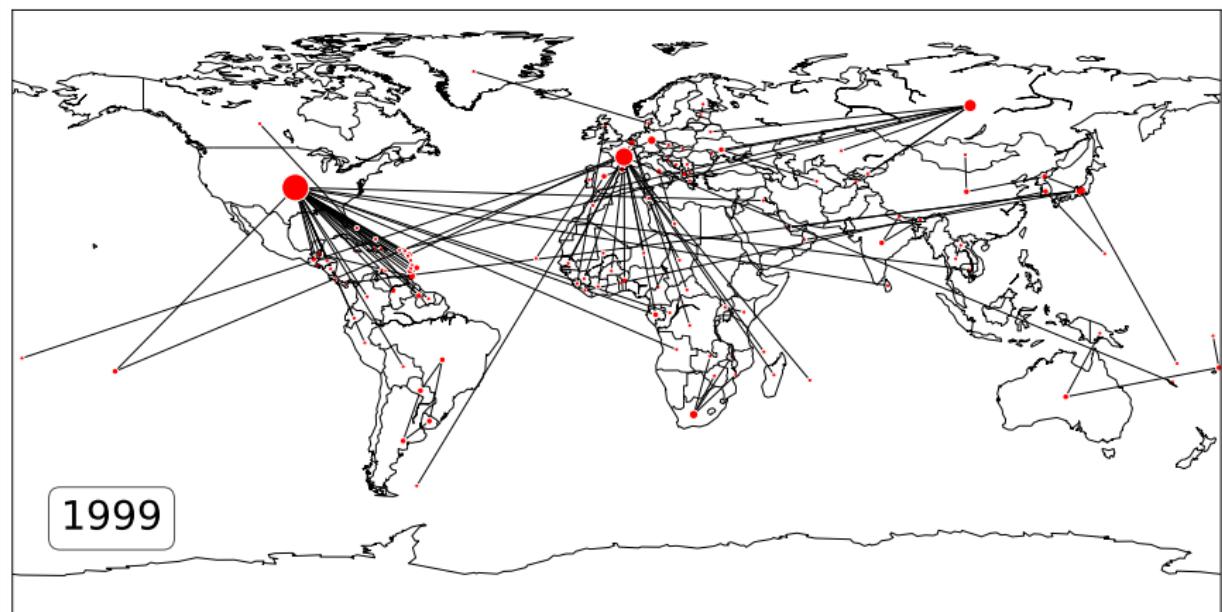
An example



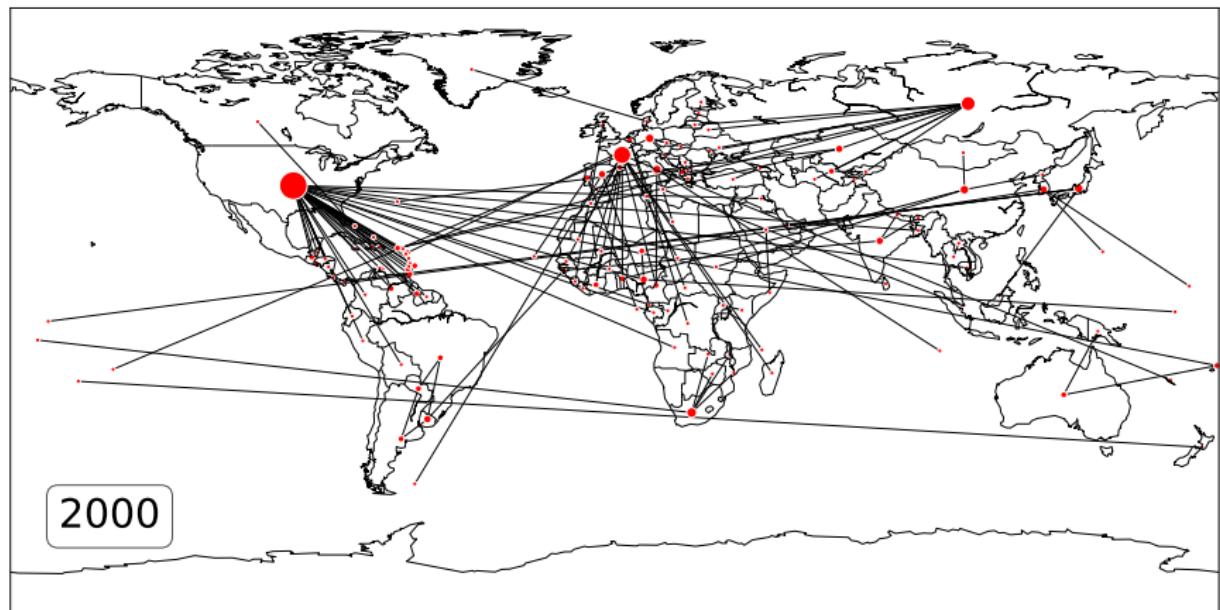
An example



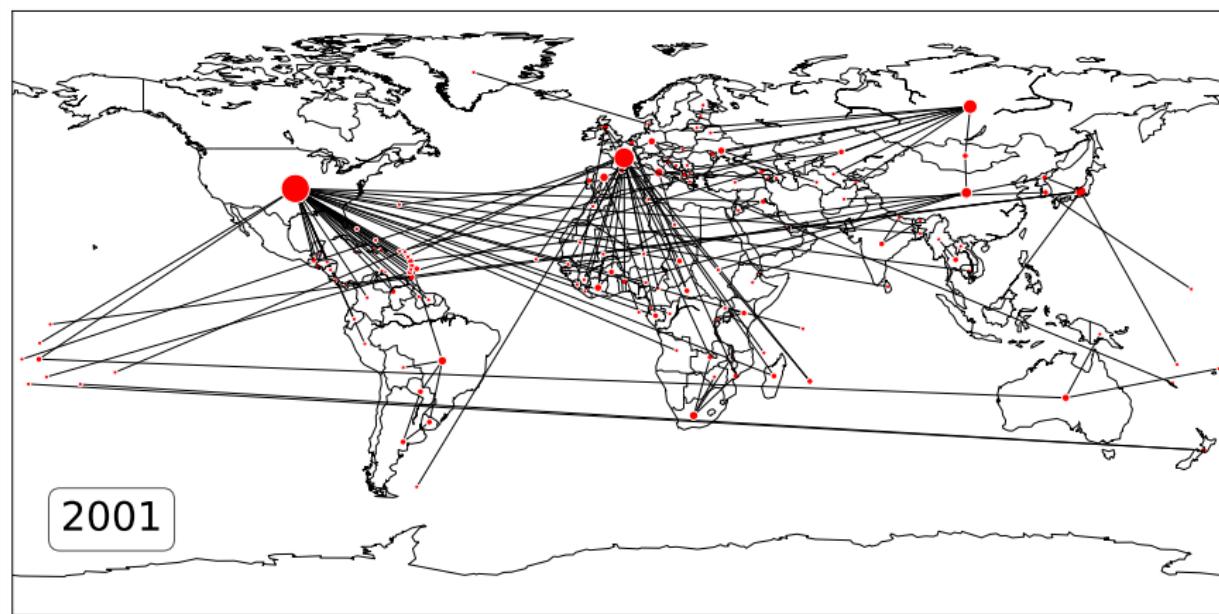
An example



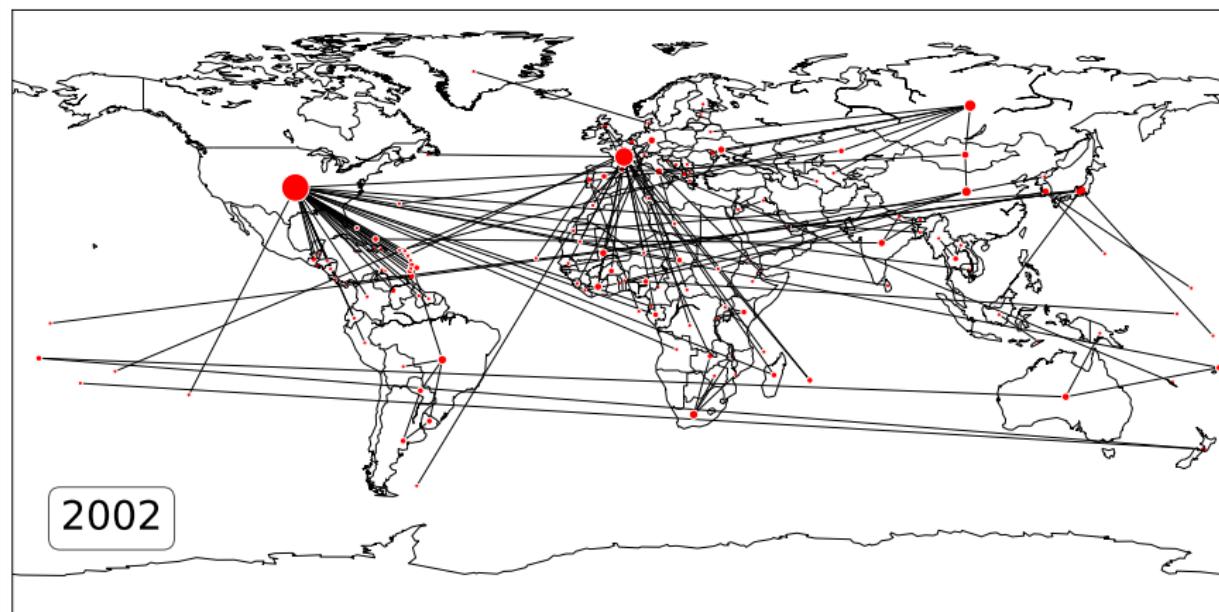
An example



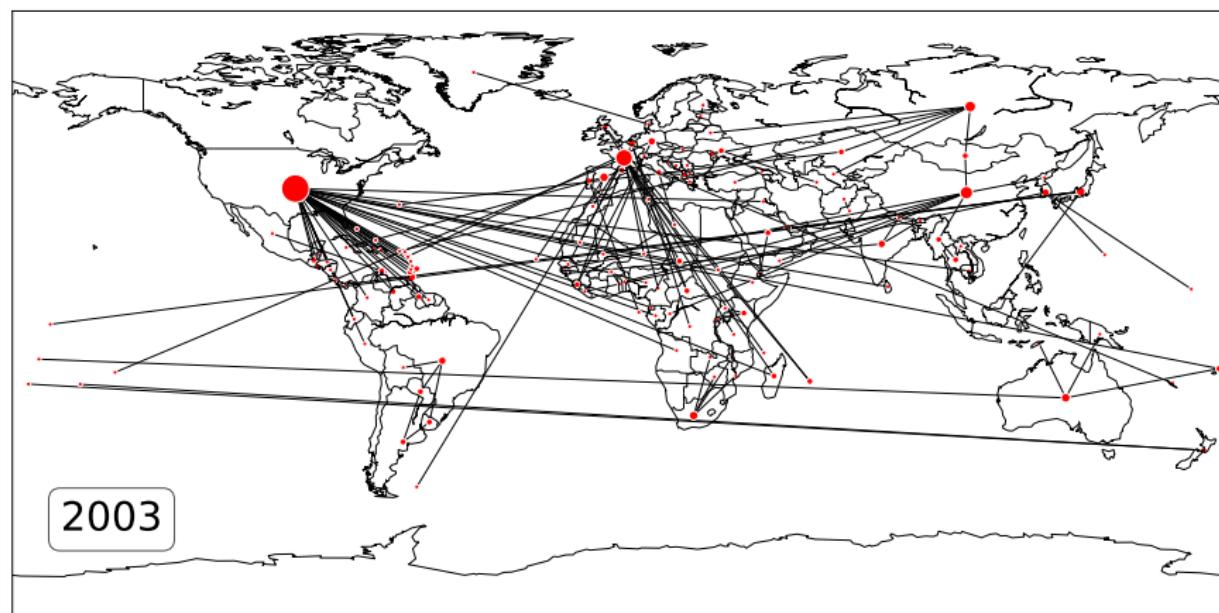
An example



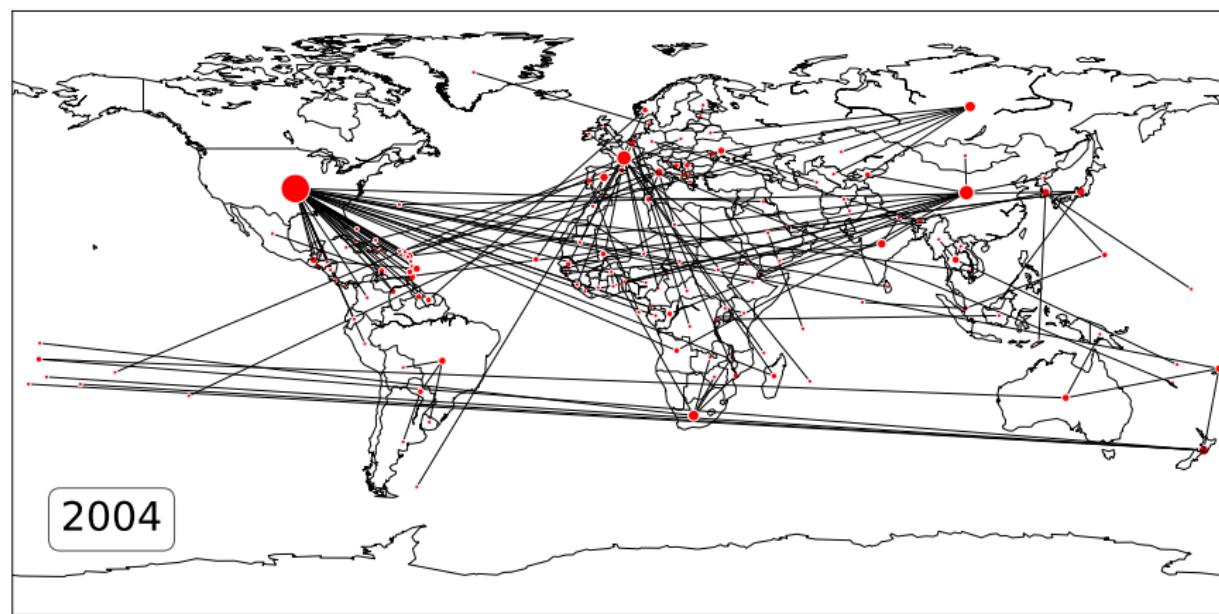
An example



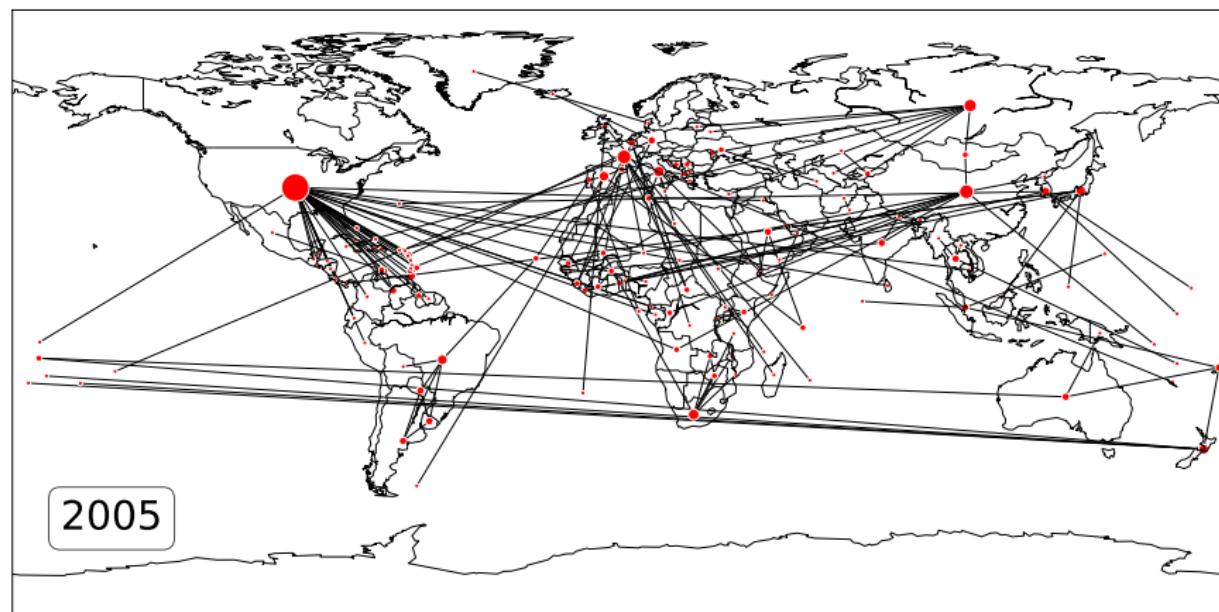
An example



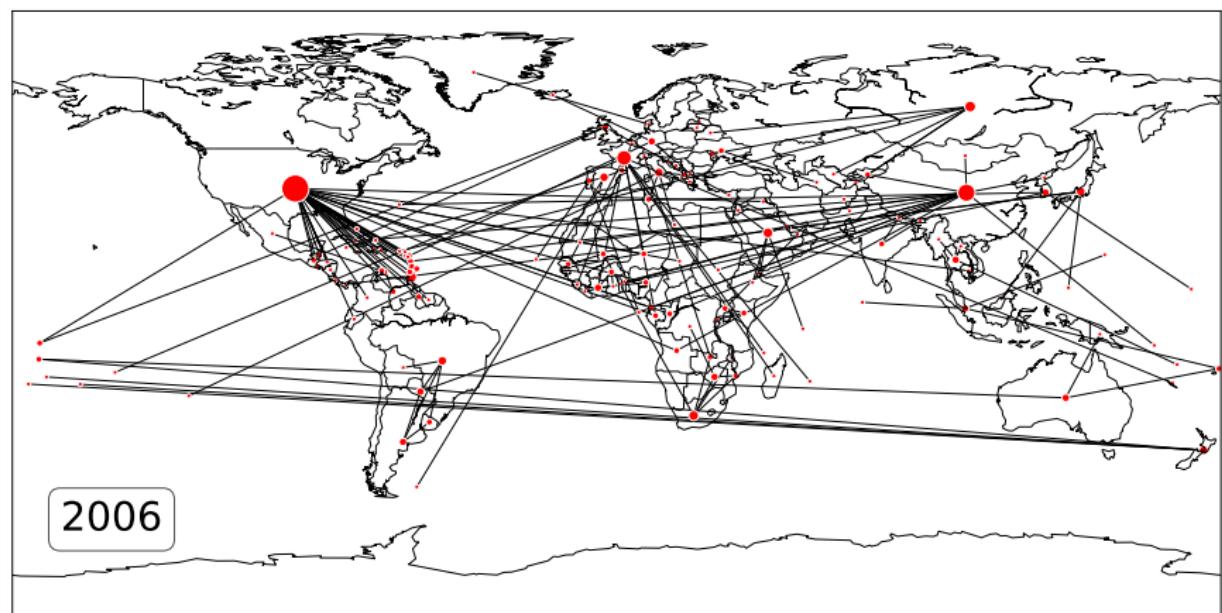
An example



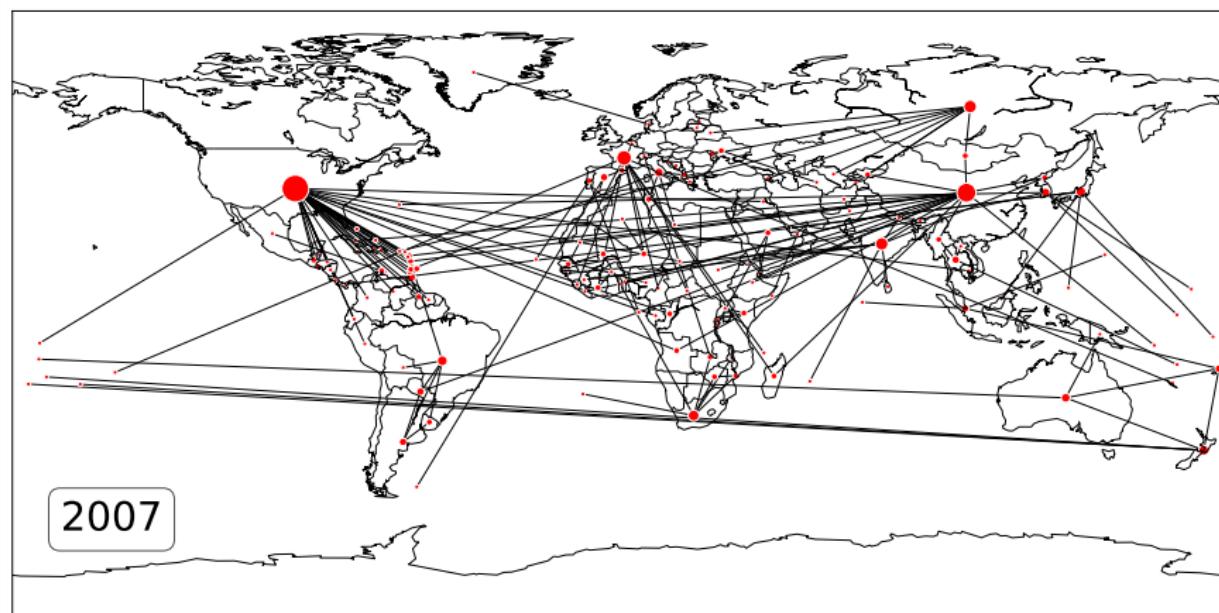
An example



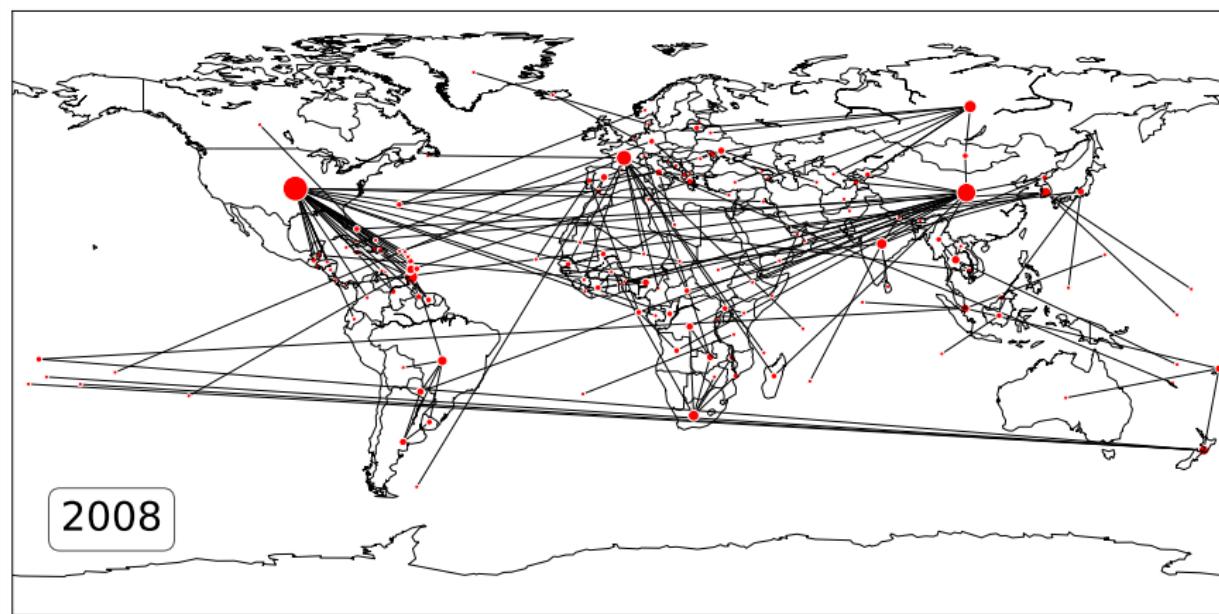
An example



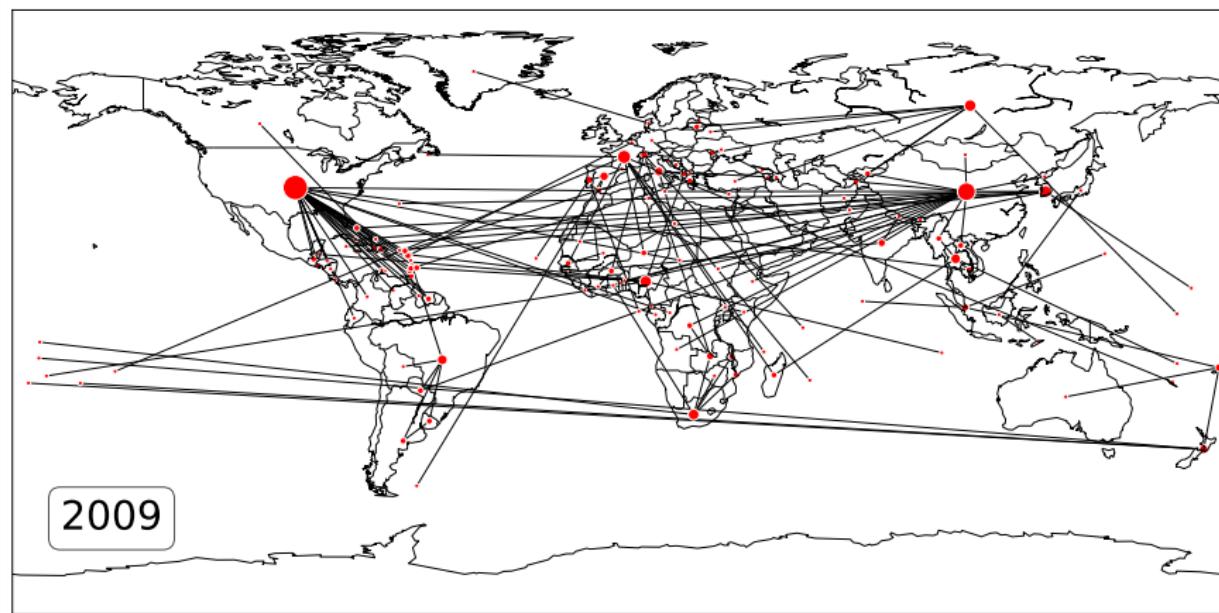
An example



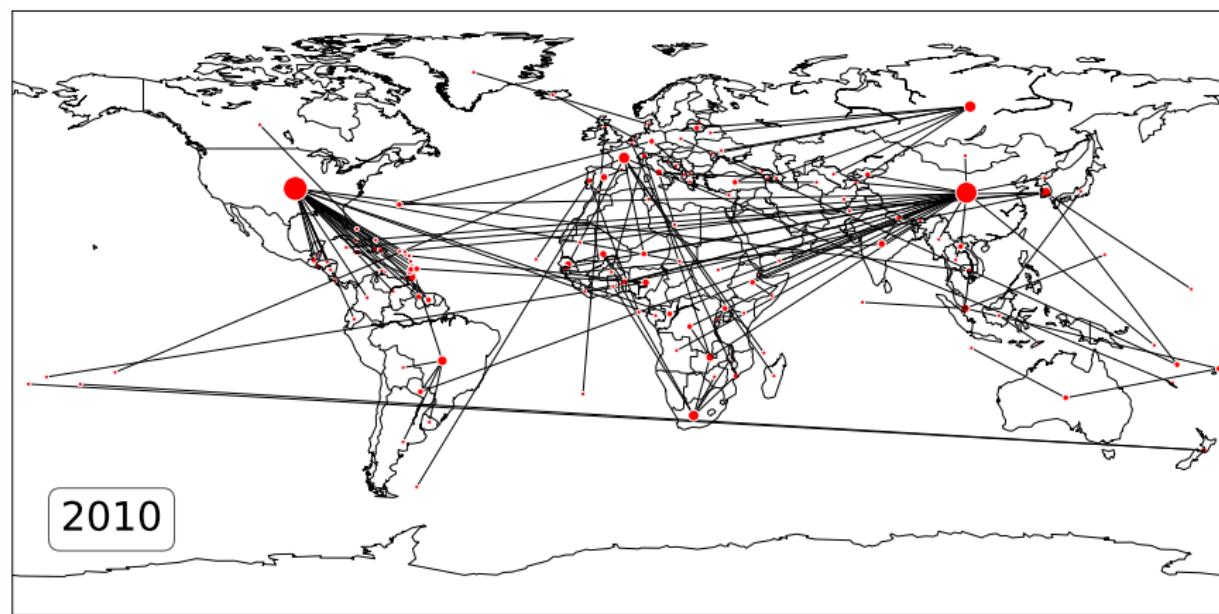
An example



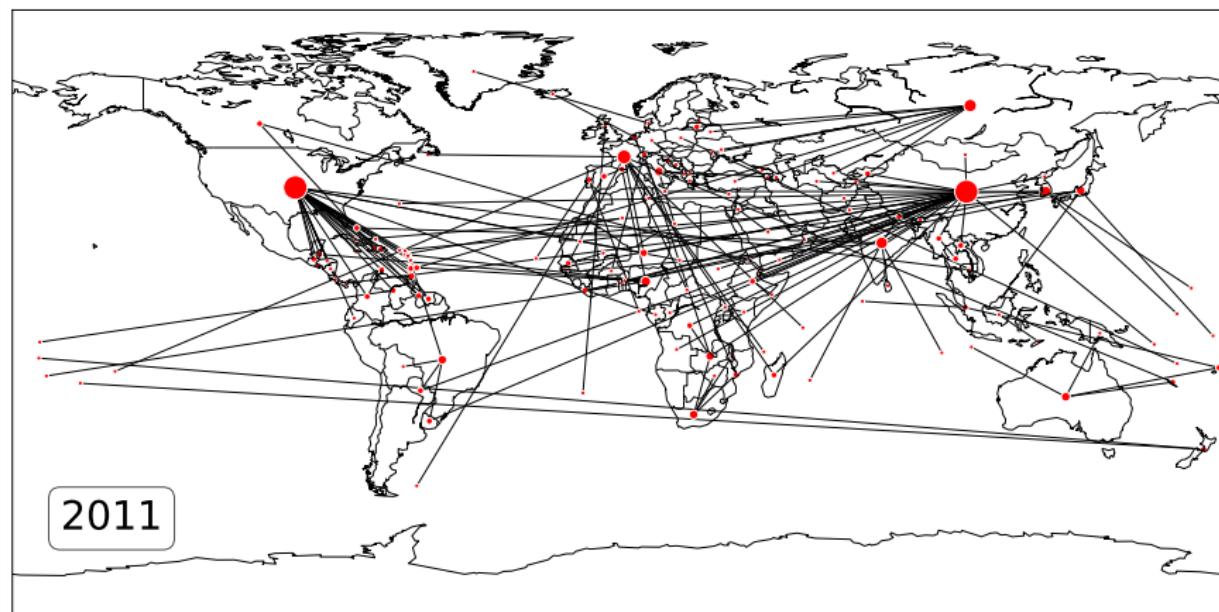
An example



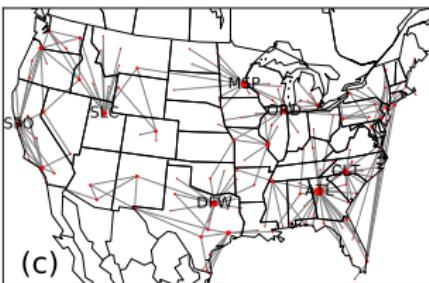
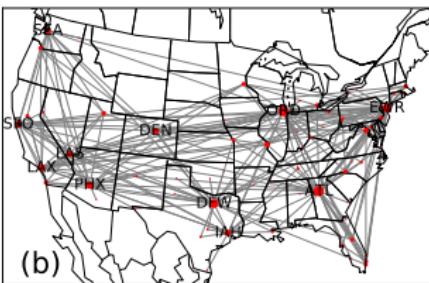
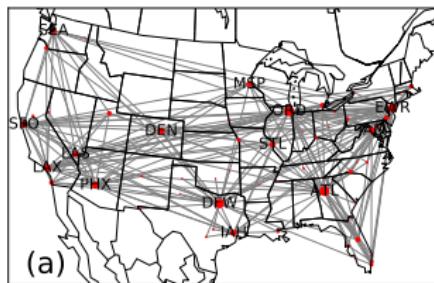
An example



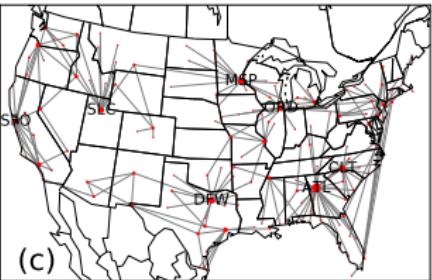
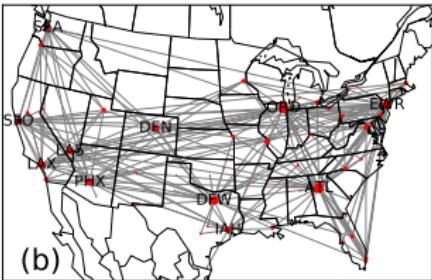
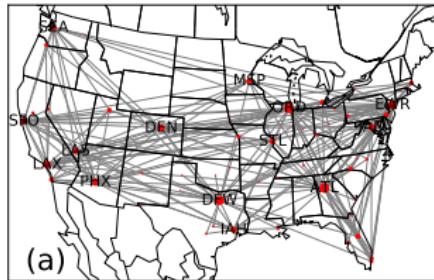
An example



Local vs Global



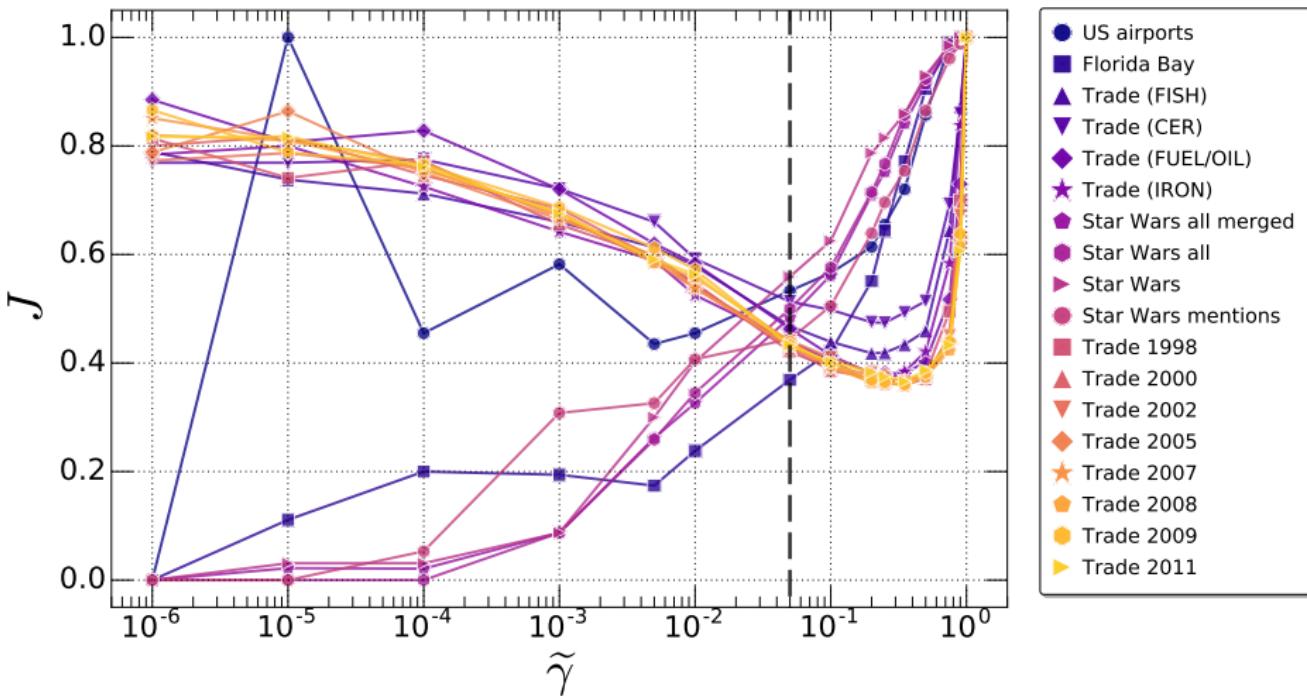
Local vs Global



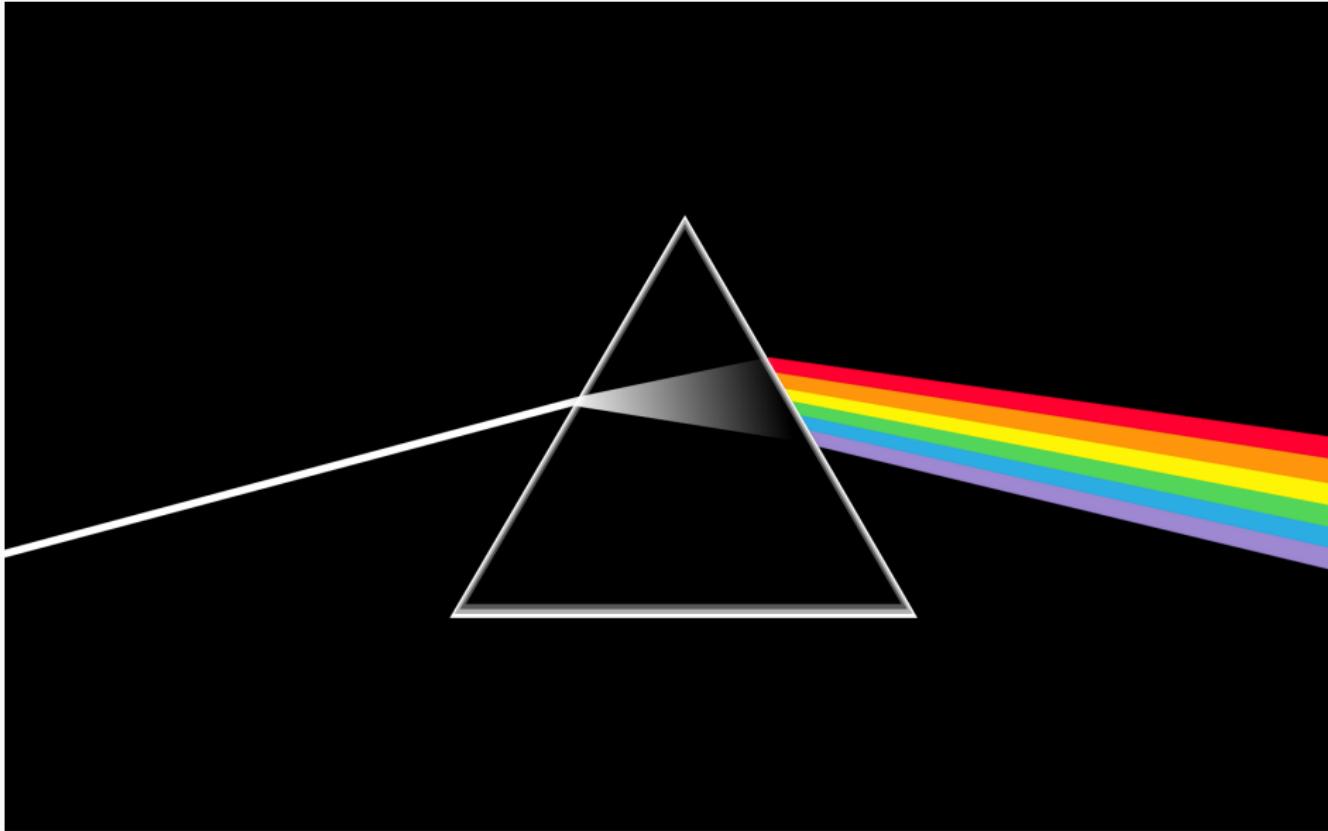
Jaccard Score

$$J = \frac{|A \cap B|}{|A \cup B|} \quad J \in [0, 1].$$

Local vs Global



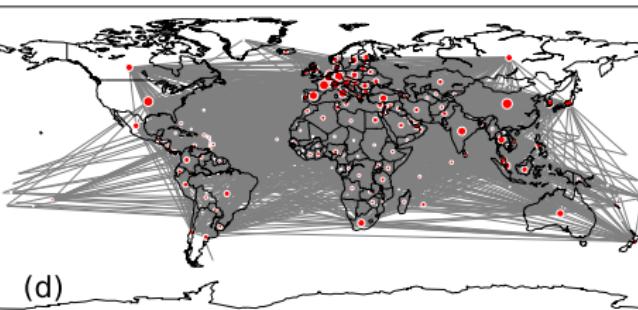
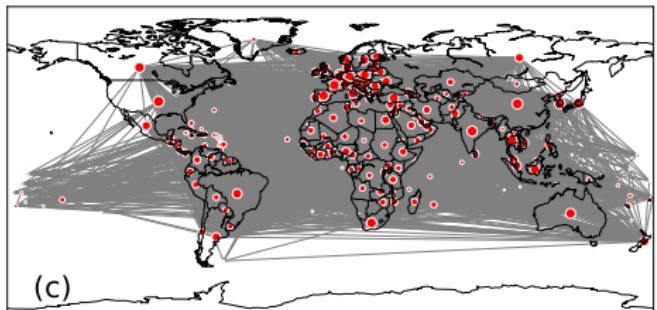
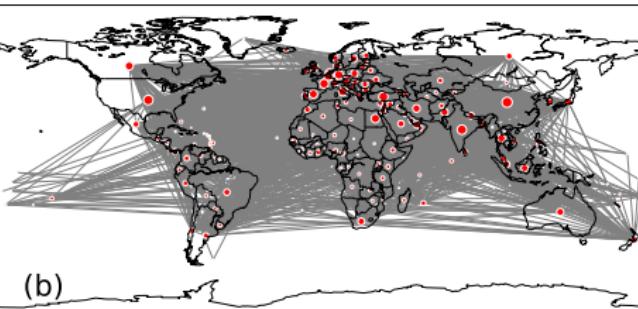
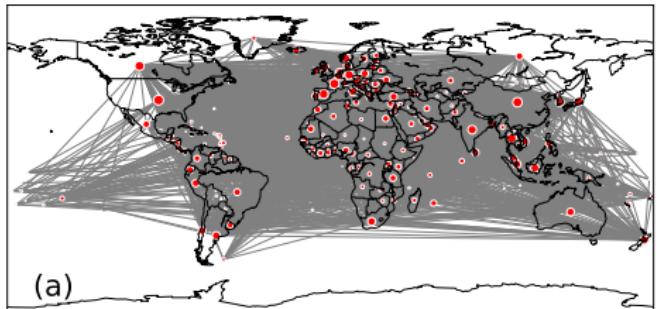
ECM and Multiplex



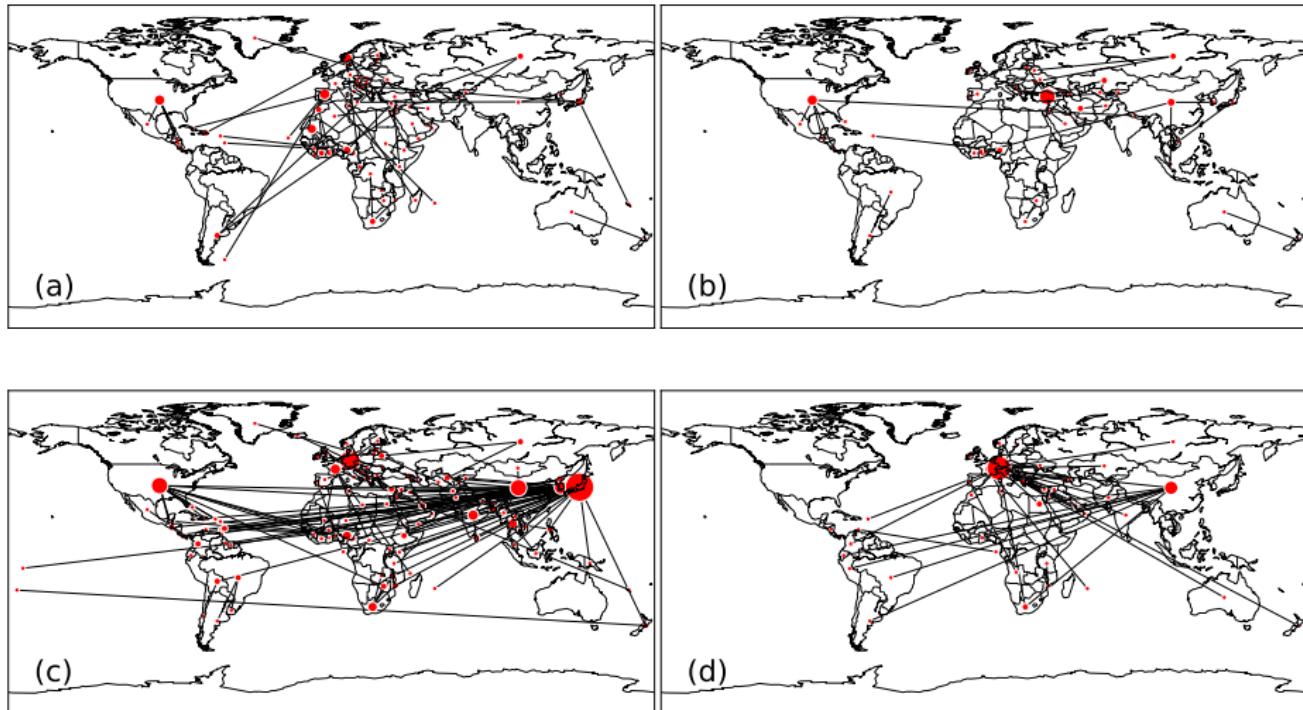
ECM and Multiplex



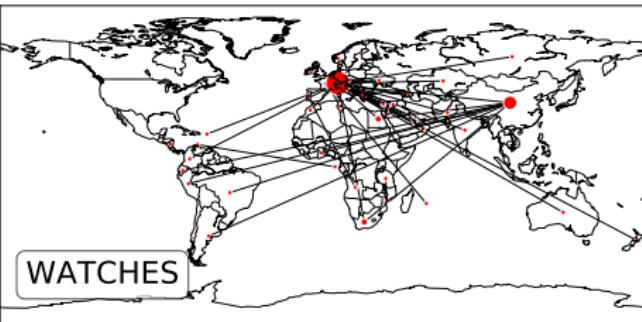
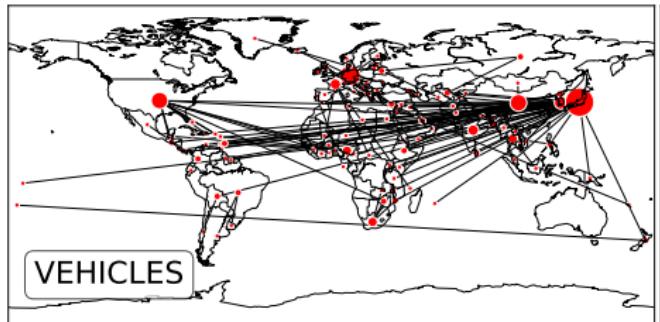
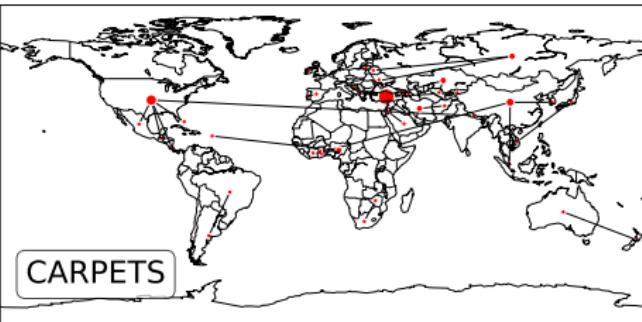
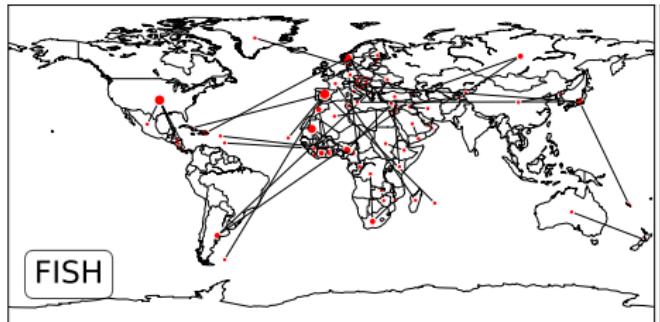
ECM and Multiplex



ECM and Multiplex



ECM and Multiplex



The screenshot shows a web page from **nature COMMUNICATIONS**. At the top, there is a dropdown menu icon and the journal logo. Below the header, the text "Article | OPEN | Published: 15 January 2019" is displayed. The main title of the article is "The structured backbone of temporal social ties". Below the title, the authors are listed as "Teruyoshi Kobayashi, Taro Takaguchi & Alain Barrat". A small "✉" icon is next to the name of the first author. At the bottom of the page, there is a link to "Nature Communications 10, Article number: 220 (2019)" and a "Download Citation" button.

- Kobayashi, T., Takaguchi, T., & Barrat, A. *Nature Communications*, **10**, 220 (2019).

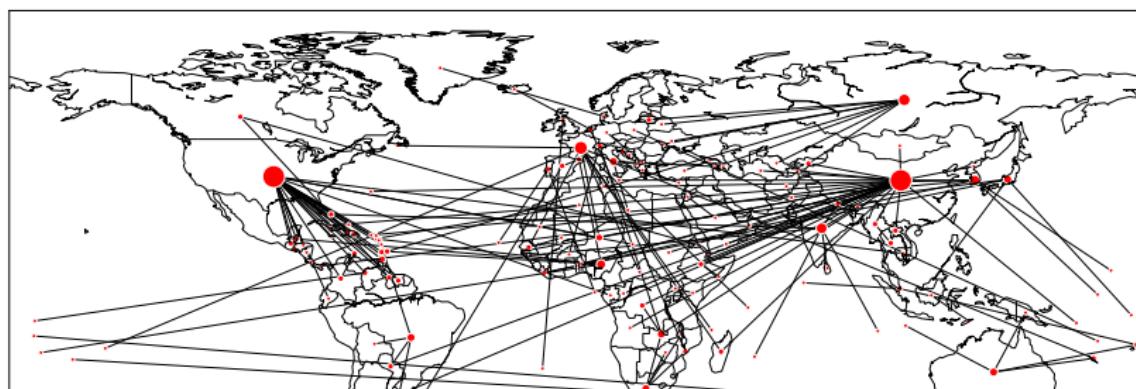
Conclusions

Take home messages

Filtering is becoming a more and more **required** step to continue using networks to study complex systems.



Take home messages



ECM-filter overcomes limitations of previous approaches, and retrieves non trivial features.

Acknowledgements



Diego Garlaschelli

Valerio Gemmetto



Irreducible network backbones: unbiased graph filtering via maximum entropy

Valerio Gemmetto, Alessio Cardillo, Diego Garlaschelli

(Submitted on 1 Jun 2017 (v1), last revised 9 Jun 2017 (this version, v2))

Networks provide an informative, yet non-redundant description of complex systems only if links represent truly dyadic relationships that cannot be properties such as size, importance, or coordinates in some embedding space. In any real-world network, some links may be reducible, and others

arXiv:1706.00230



alessio.cardillo@urv.cat



<http://www.bifi.es/~cardillo/>



@a_cardillo

Bibliography i

-  Kim, B. J. (2004). Geographical Coarse Graining of Complex Networks. *Phys. Rev. Lett.* **93**, 168701
-  Gfeller, D., & De Los Rios, P. (2007). Spectral Coarse Graining of Complex Networks. *Phys. Rev. Lett.*, **99**, 38701.
-  Gionis A., Rozenshtein P., Tatti N., & Terzi E. (2017) Community-aware network sparsification Proc. of the 2017 SIAM International Conference on Data Mining. 2017, 426–434
-  Tumminello, M., Aste, T., Di Matteo, T., & Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences*, **102**, 10421–10426.

Bibliography ii

-  De Vico Fallani, F., Latora, V., & Chavez, M. (2017). A Topological Criterion for Filtering Information in Complex Brain Networks. *PLoS Comp. Bio.*, **13**, e1005305.
-  Yan, X., Jeub, L. G. S., Flammini, A., Radicchi, F., & Fortunato, S. (2018). Weight thresholding on complex networks. *Phys. Rev. E*, **98**, 042304.
-  Serrano, M. Á., Boguñá, M., & Vespignani, A. (2009). Extracting the multiscale backbone of complex weighted networks. *Proc. Natl. Acad. Sci. USA*, **106**, 6483–8.

Bibliography iii

-  Coscia M., & Neffke F. M. H. (2017) Network Backboning with Noisy Data, 2017 IEEE 33rd International Conference on Data Engineering (ICDE), San Diego, CA, 2017, 425–436.
-  Radicchi, F., Ramasco, J. J., & Fortunato, S. (2011). Information filtering in complex weighted networks. *Phys. Rev. E*, **83**, 046101.
-  Dianati, N. (2016). Unwinding the hairball graph: Pruning algorithms for weighted complex networks. *Phys. Rev. E*, **93**, 012304.

Bibliography iv

-  Neal, Z. (2014). The backbone of bipartite projections: Inferring relationships from co-authorship, co-sponsorship, co-attendance and other co-behaviors. *Social Networks*, **39**, 84–97.
-  Kobayashi, T., Takaguchi, T., & Barrat, A. (2019). The structured backbone of temporal social ties. *Nature Communications*, **10**, 220.
-  Marcaccioli, R., & Livan, G. (2019). A Pólya urn approach to information filtering in complex networks. *Nature Communications*, **10**, 745.

Bibliography v

-  Gemmetto, V., Cardillo, A., & Garlaschelli, D. (2017) Irreducible network backbones: unbiased graph filtering via maximum entropy. arXiv e-prints, arXiv:1706.00230.
-  Kojaku, S., & Masuda, N. (2019) Constructing networks by filtering correlation matrices: A null model approach. Proc. Roy. Soc. A, **475**, 20190578.

Bipartite Networks



ELSEVIER

Social Networks

Volume 39, October 2014, Pages 84-97



The backbone of bipartite projections:
Inferring relationships from co-authorship,
co-sponsorship, co-attendance and other co-
behaviors

Zachary Neal  

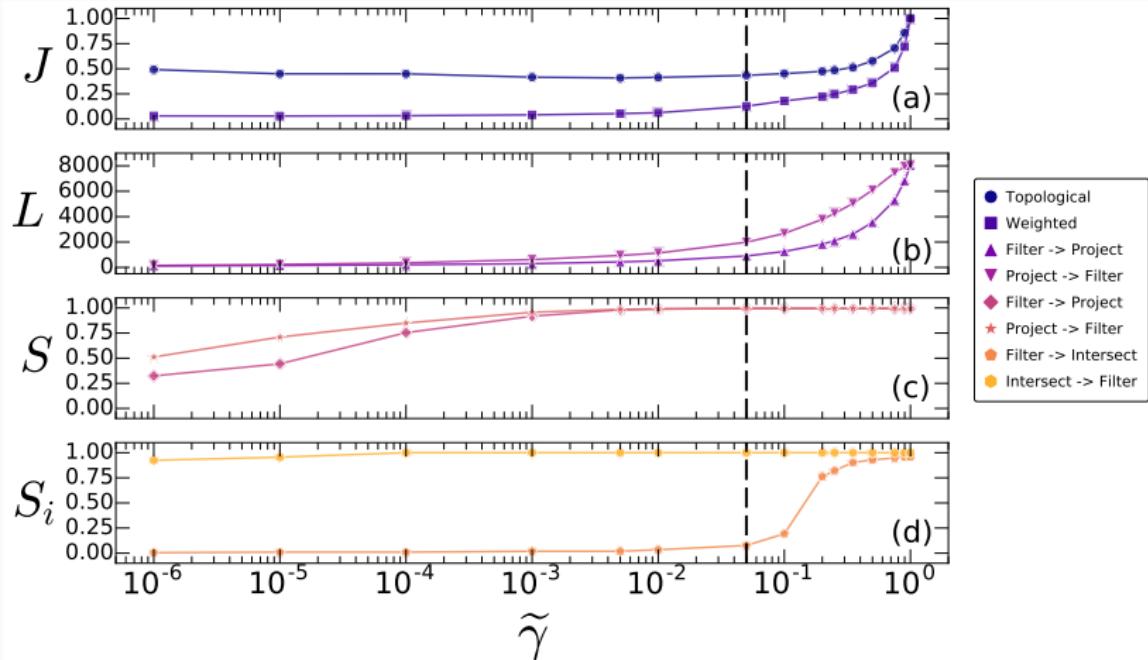
Show more

<https://doi.org/10.1016/j.socnet.2014.06.001>

[Get rights and content](#)

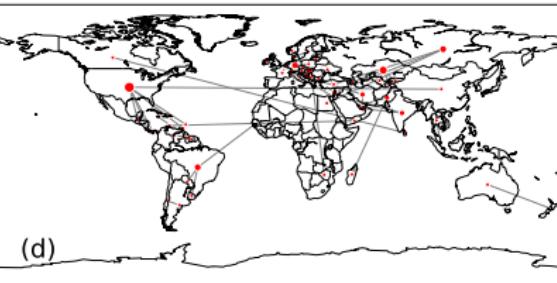
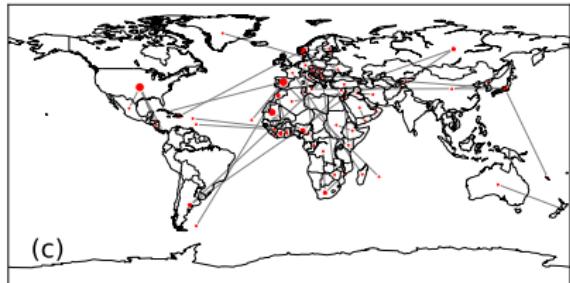
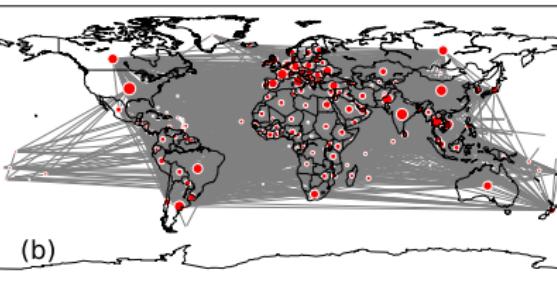
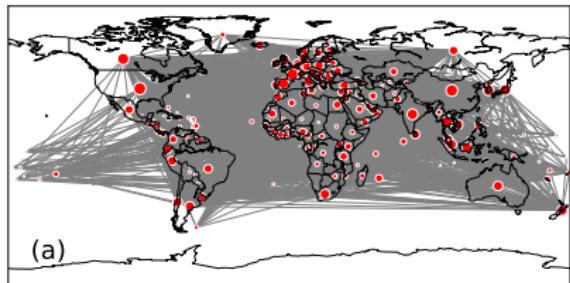
- Neal, Z. Social Networks, 39, 84–97 (2014).

Filtering & Multiplexity



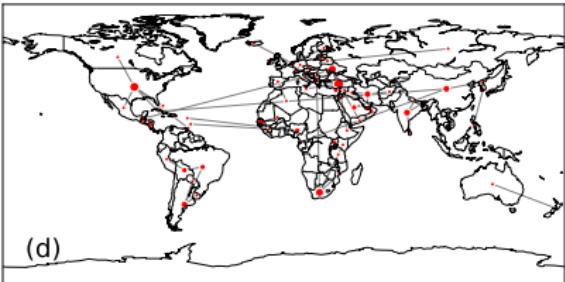
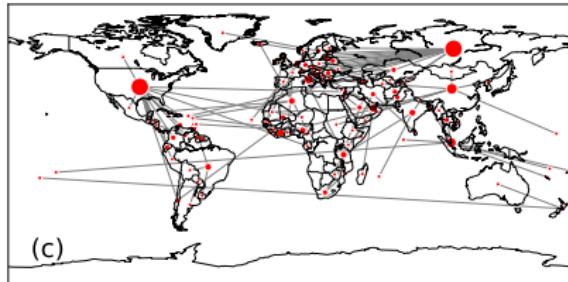
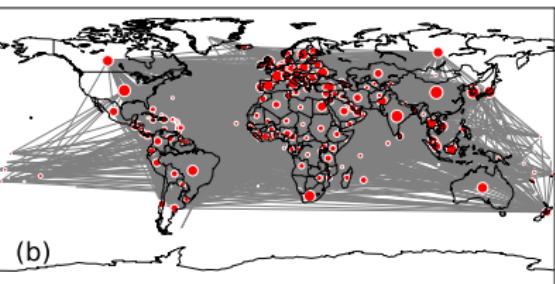
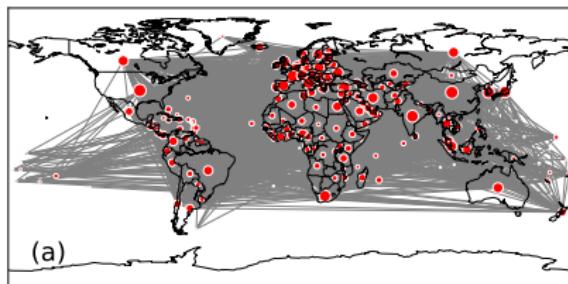
Trade: other commodities

FISH – CEREALS in 2011

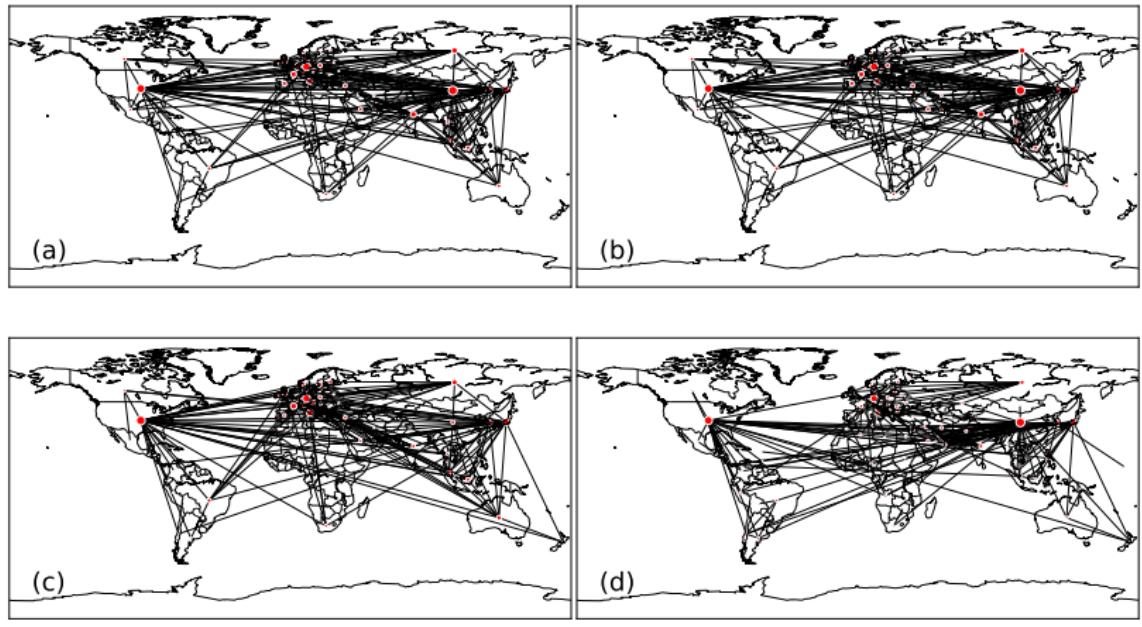


Trade: other commodities

FUEL – IRON in 2011



Trade: comparison of methods



PROCEEDINGS OF THE ROYAL SOCIETY A

MATHEMATICAL, PHYSICAL AND ENGINEERING SCIENCES

You have access

 Check for updates

 View PDF

Tools

Share

Cite this article ▾

Research articles

Constructing networks by filtering correlation matrices: a null model approach

Sadamori Kojaku and Naoki Masuda

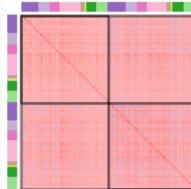
Published: 13 November 2019 | <https://doi.org/10.1098/rspa.2019.0578>

- Kojaku, S., & Masuda, N. Proc. Roy. Soc. A, **475**, 20190578 (2019)

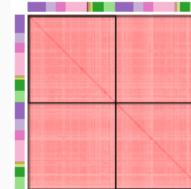


Sample correlation matrices

(a) '72-'91

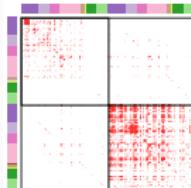


(b) '92-'11

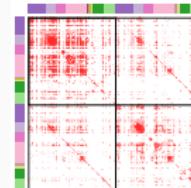


Thresholding

(c) '72-'91

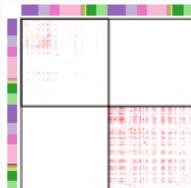


(d) '92-'11

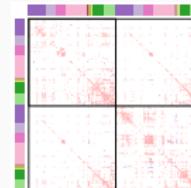


Scola

(e) '72-'91



(f) '92-'11



■ Machinery & transport equipment (\$327b)
■ Chemicals (\$327b)
■ Miscellaneous manufactured articles (\$241b)
■ Manufactured goods by material (\$227b)
■ Miscellaneous (\$161b)

■ Mineral fuels, lubricants & related materials (\$159b)
■ Beverages & tobacco (\$145b)
■ Animal & vegetable oils, fats & waxes (\$125b)
■ Food & live animals (\$122b)
■ Raw materials (\$102b)