Characterizing Weighted Complex Networks in Economics and Society

Jari Saramäki[1], Jukka-Pekka Onnela[1],
János Kertész[2], Kimmo Kaski[1]

[mobile communication network analysis: also
Jörkki Hyvönen[1], Gábor Szabó[3], Albert-László Barabási[3]]

[1]Laboratory of Computational Engineering,
Helsinki University of Technology, Finland

[2]Department of Theoretical Physics,
Budapest University of Technology and Economics, Hungary

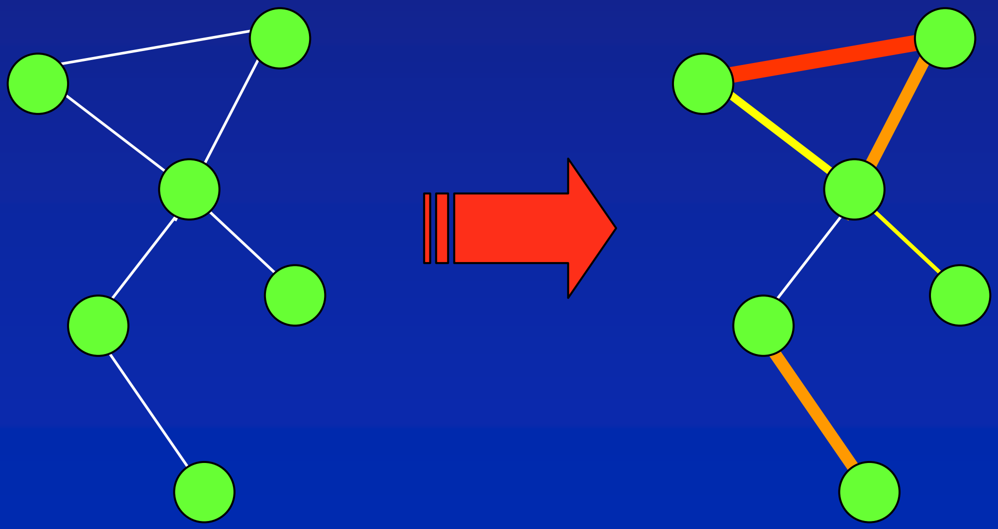
[3]Harvard University / University of Notre Dame, USA

Outline

1. **Beyond pure topology: weighted networks**
2. Clustering & subgraph intensity
3. Weighted motif statistics

From ”binary” networks to weighted networks

- Complex networks: a way of looking at complex systems
- Elements ⇒ nodes
- Interactions ⇒ links
- Interaction (or flow) strengths ⇒ link weights
- E.g. traffic networks, social networks, market interactions, metabolic networks, …



How to characterize weighted networks?

- Some usual characteristics are straightforward to generalize:

vertex degree $k_i \rightarrow$ vertex strength[1] $s_i = \sum_j w_{ij}$

- Others, like the clustering coefficient, less obvious

- In addition, novel ways of investigating weight-topology correlations are needed

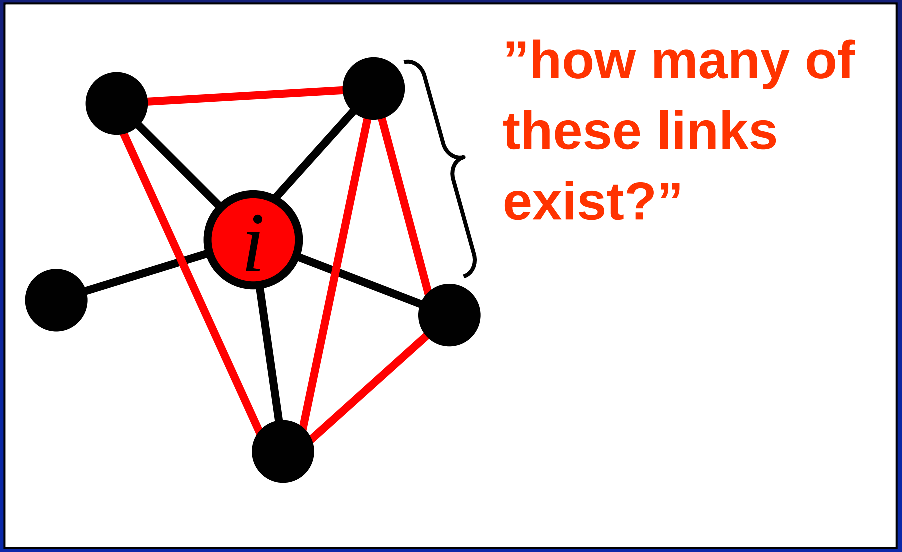
[1] A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, *Proc. Natl. Acad. Sci. USA* **101**, 3747 (2004)

Outline

1. Beyond pure topology: weighted networks
2. **Clustering & subgraph intensity**
3. Weighted motif statistics

The clustering coefficient

- The unweighted clustering coefficient measures to what extent the neighbours of a vertex are connected, i.e. how many triangles exist around the vertex



clustering coefficient at vertex *i*:

$$C_i = \frac{\#\ of\ triangles}{\#\ of\ possible\ triangles} = \frac{t_i}{\frac{1}{2}k(k-1)}$$

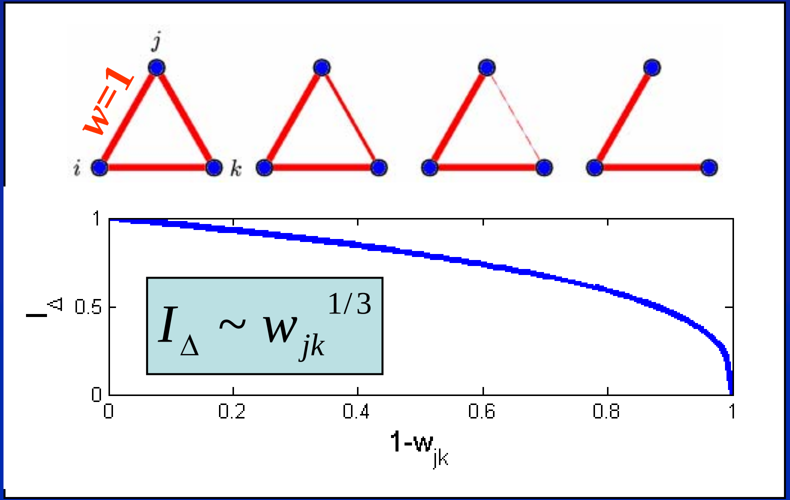
$$\langle C \rangle = \frac{1}{N}\sum_i C_i$$

Subgraph intensity

- Consider an undirected weighted network with $w_{ij}>0$
- Define <u>the intensity *I(g)*</u> of a particular subgraph g with vertices v_g and links l_g s.t. $|l_g|$ = the number of links in g as

$$I(g) = \left(\prod_{(ij)\in l_g} w_{ij} \right)^{1/|l_g|}$$

Schematic for a triangle: one low weight results in low intensity



The weighted clustering coefficient

- Normalize weights so that $I_{\Delta}=1$ if all weights equal to the maximum weight (of the network or local)

$$w_{ij} \leftarrow w_{ij} / \max_{ij}(w_{ij})$$

- Replace the number of triangles with the sum of their intensities:

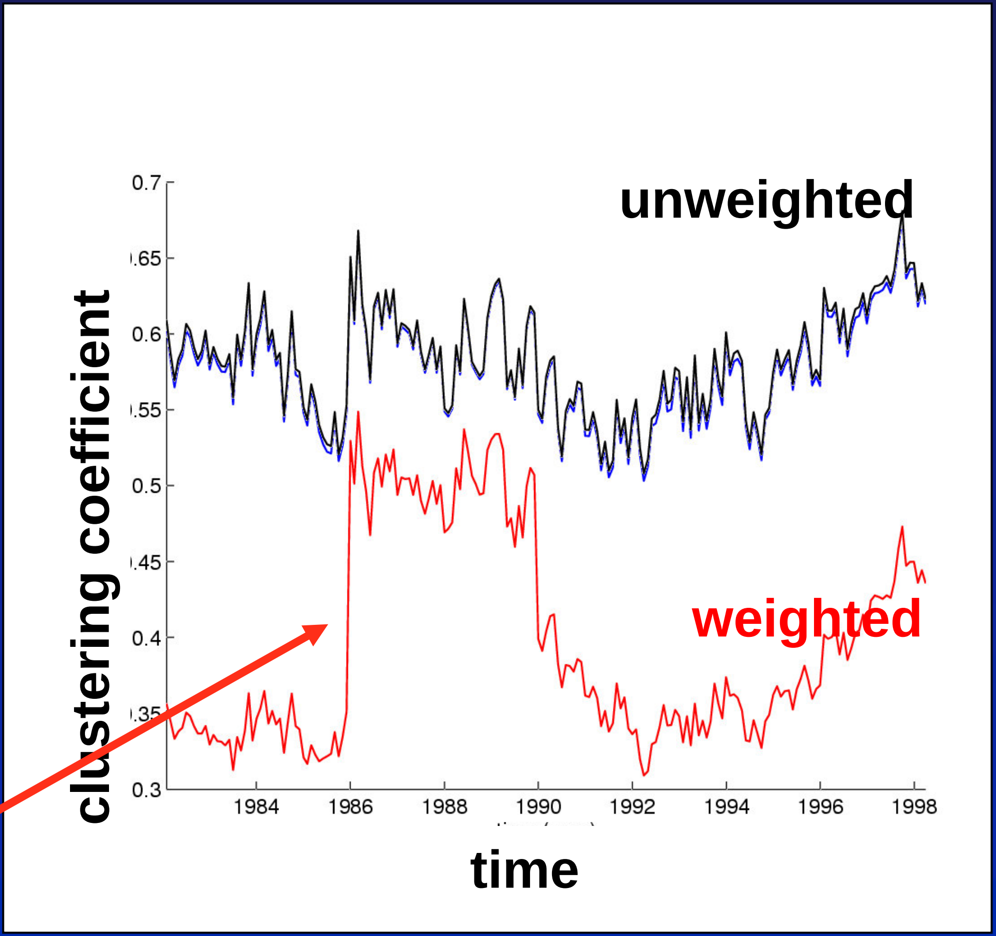
$$C_i = \frac{t_i}{\frac{1}{2}k(k-1)} \Rightarrow \tilde{C}_i = \frac{2I_{\Delta,i}}{k_i(k_i-1)} = \frac{2}{k_i(k_i-1)} \sum_{j,k} \left(w_{ij} w_{jk} w_{ki} \right)^{1/3}$$

- Range correct: $\tilde{C}_i \in [0,1]$
- Equals the unweighted C when all weights in the network are equal

Example: asset graph & Black Monday

- "Asset graph"[1] calculated from stock return time series: 477 NYSE stocks, daily returns within 4-year time windows → correlation matrix → weighted network (highest correlations only)
- Edge weights = correlations
- **Black Monday (10/19/1987) causes a temporary change in the network, affecting both topology and weights**

Average clustering coefficient



[1] J.P. Onnela et al, PRE **68**, 056110 (2003).

Outline

1. Beyond pure topology: weighted networks
2. Clustering & subgraph intensity
3. **Weighted motif statistics**

Characterizing Weighted Motifs

- Unweighted: ”Frequently occurring subgraphs”[1]
  - frequently occurring subgraphs are related to system functionality
- Weighted:

  Motif = a set of topologically equivalent subgraphs

  Motif intensity = sum of subgraph intensities

$$I_M = \sum_{g \in M} I(g)$$

  - Compare motif intensities (e.g. using Z-score), or intensity distributions of subgraphs, to randomized reference

[1] R. Milo et al, Science 298, 824 (2002)

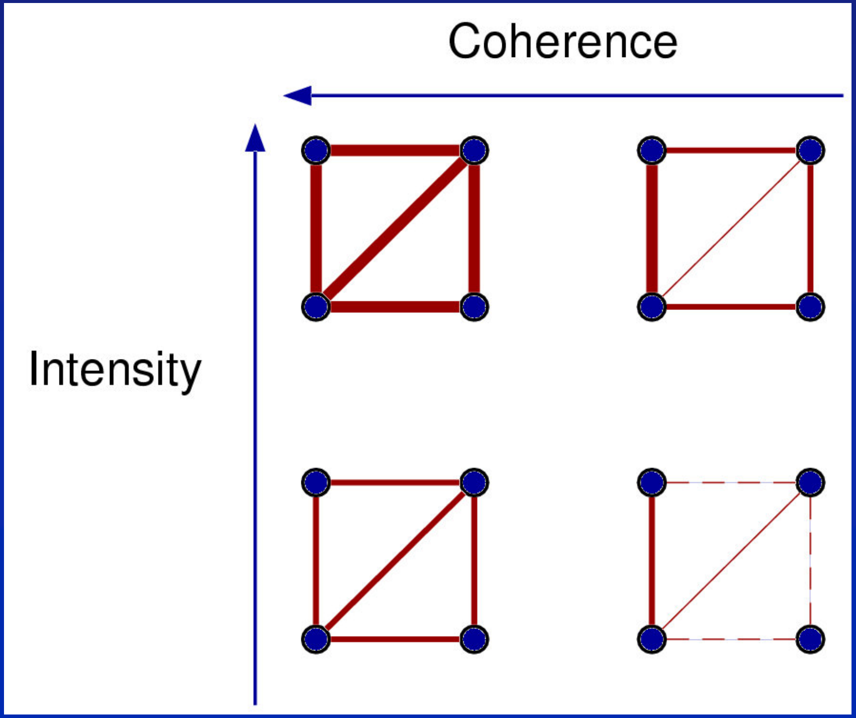
Intensity and Coherence of Subgraphs

INTENSITY

$$I(g) = \left(\prod_{(ij) \in l_g} w_{ij} \right)^{1/|l_g|}$$

COHERENCE

$$Q(g) = \frac{I(g)}{\frac{1}{|\ell_g|} \sum_{(ij) \in \ell_g} w_{ij}}$$



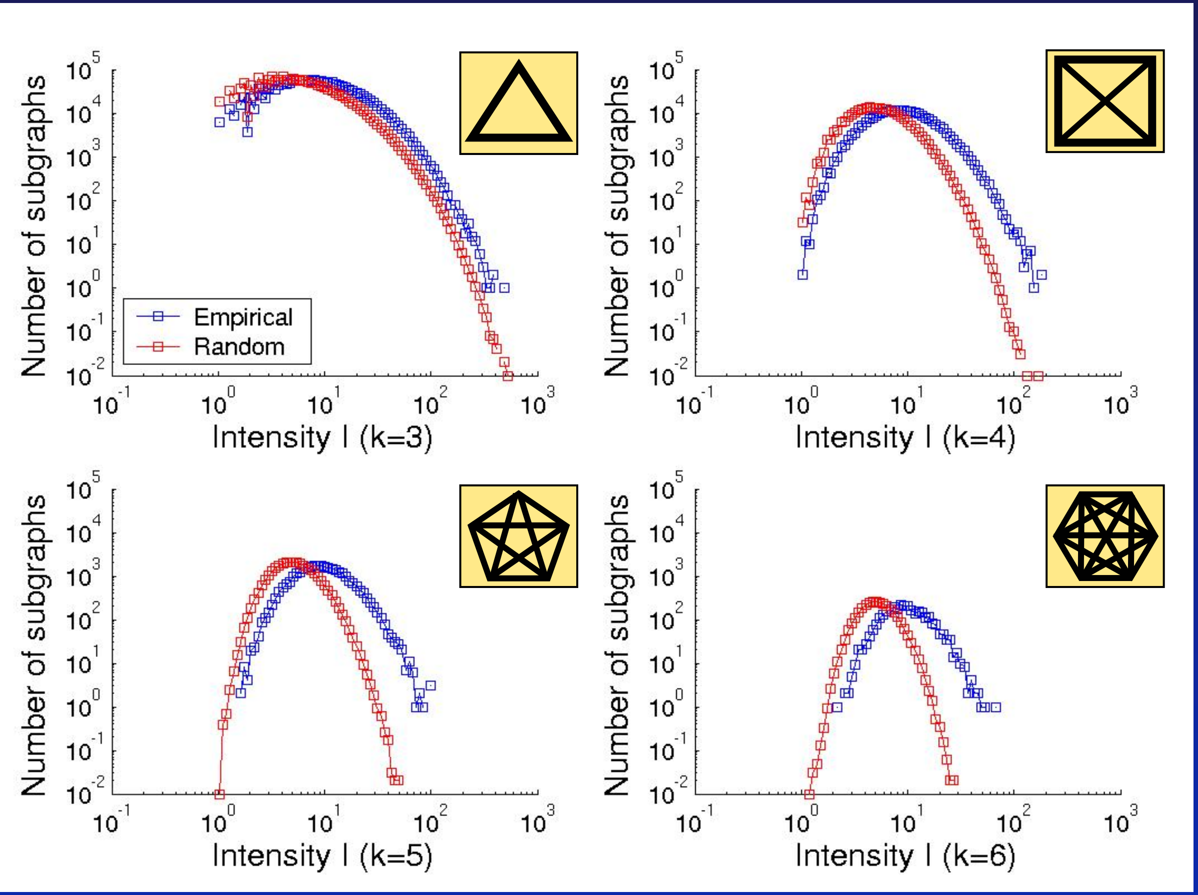
Intensity Distributions in the Mobile Communication Network

- Network: ~4 mill. mobile telephone users, edge weights = communication frequencies

BLUE: EMPIRICAL

RED: REFERENCE ENSEMBLE = EMPIRICAL WITH PERMUTED WEIGHTS

- **Fully connected cliques have <u>larger intensities</u> in the empirical network**



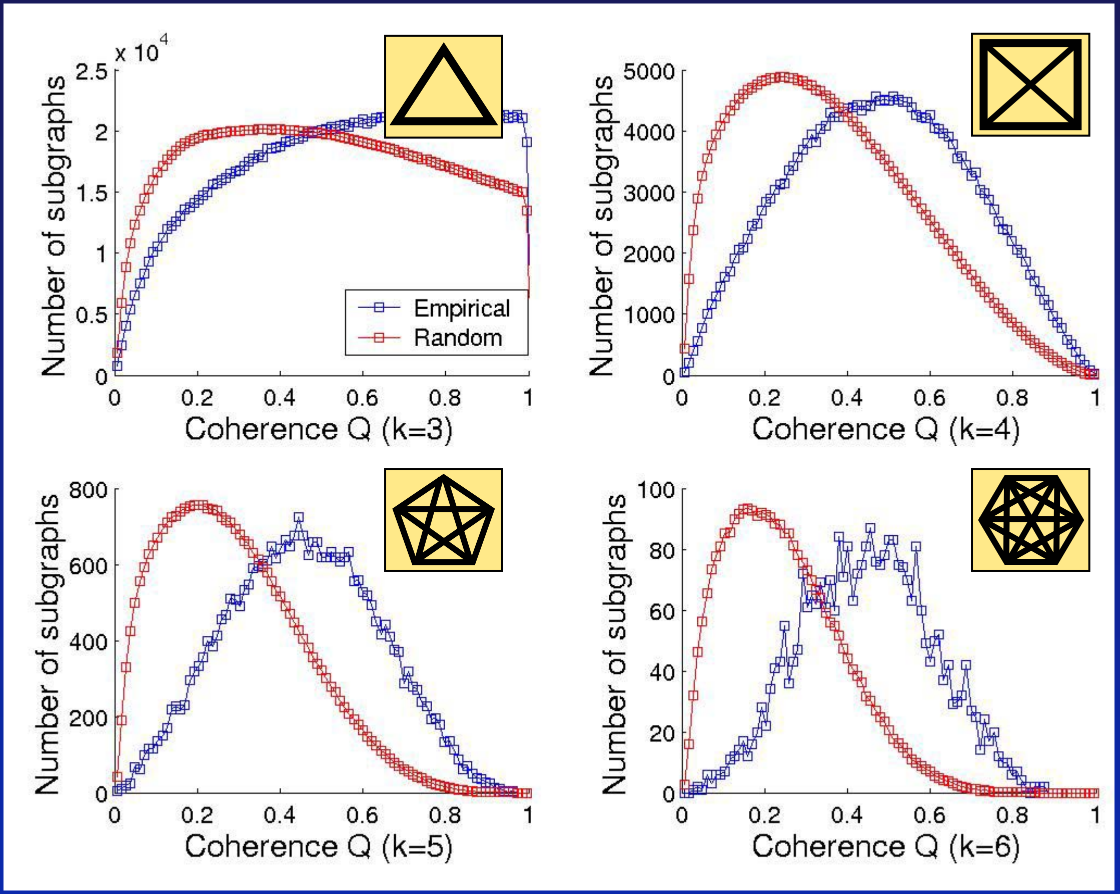
Coherence Distributions in the Mobile Communication Network

- Network: ~4 mill. mobile telephone users, edge weights = communication frequencies

BLUE: EMPIRICAL

RED: WEIGHTS PERMUTED

- **Fully connected cliques are <u>more coherent</u> in the empirical network**
- **→ high weight edges are concentrated in cliques**



THANK YOU!

J.-P. Onnela, J. Saramäki, J. Kertész, and K. Kaski:

Intensity and coherence of motifs in weighted complex networks, *Phys. Rev. E* 71, 065103 (2005)