Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Network Theory
Lecture 4.10

EEU45C09 / EEP55C09
Self Organising Technological Networks

Nicola Marchetti
nicola.marchetti@tcd.ie

Empirical Study of Social Networks

Social networks are networks in which the vertices are people or groups of people and the edges represent some sort of social interaction between them eg. friendship.

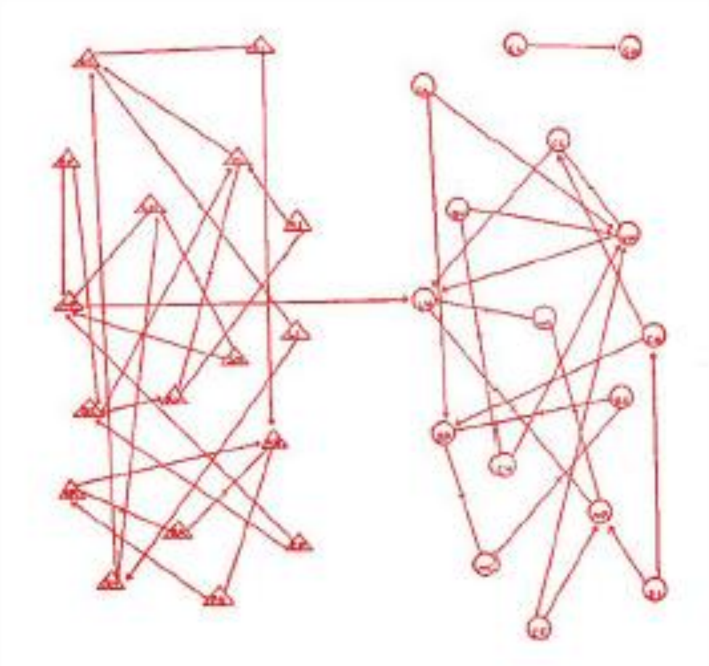


Figure 2: A hand-drawn image depicting friendships between girls (circles) and boys (triangles) in a class in the 1930s by Joseph Moreno.

Empirical Study of Social Networks

There are many different possible definitions of an edges in a social network. The definition used will depend on what questions one is trying to answer.

Edges may represent:

- friendships,
- professional relationships,
- exchange of goods or money,
- communication patterns,
- or romantic relationships.

Since these types of networks are quite different, the techniques used to form them must be quite different.

Social Networks: Surveys

The most common method for accumulating data on social networks is simply asking people questions. However, due to the large number of participants in many of these surveys, some care must be taken to guarantee consistent order, language, interpretation of question etc.

For this reason, often social network data is collected by a professional interviewer reading from a survey script. This significantly improves data collection:

- Survey takers approach questions more seriously.
- Misunderstandings and misinterpretations are reduced.

Social Networks: Surveys

To probe social networks, surveys typically employ a *name generator*:

```
My best friend at ______ Secondary School is:
My second-best friend at ______ Secondary School is:
My third-best friend at ______ Secondary School is:
```

In this case, the method may be flawed. No effort is made to deal with nodes outside the school which may affect any statistical analysis.

The main disadvantages of network studies based on direct questioning are that they are laborious and inaccurate.

Social Networks: Ego-Centred Networks

Sociometric networks in which all members of a community are surveyed are sometimes infeasible.

An alternative is ego-centric network where an individual and his/her immediate contacts are surveyed. No attempt is made to survey friends of friends etc. Gives reasonable measure of degree distribution.

Social Networks: Direct-Observation

By observing interactions between individuals over time, one can form a picture of the network of ties between those individuals.

A common example is being aware of friendships or enmities between friends and coworkers. Though again, direct observation networks tend to be very laborious to construct.

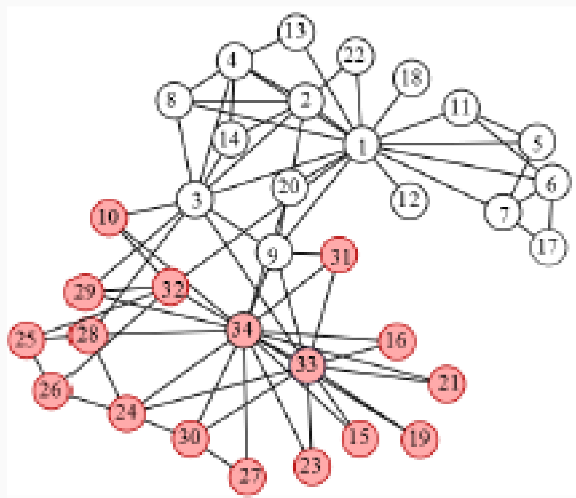


Figure 3: Network of direct observations of friendships within a karate club from 70's.

Social Networks: Archival or Third-Party Data

Archival records are often very large, free from human-memory error and in general much cheaper than direct observation or surveys.

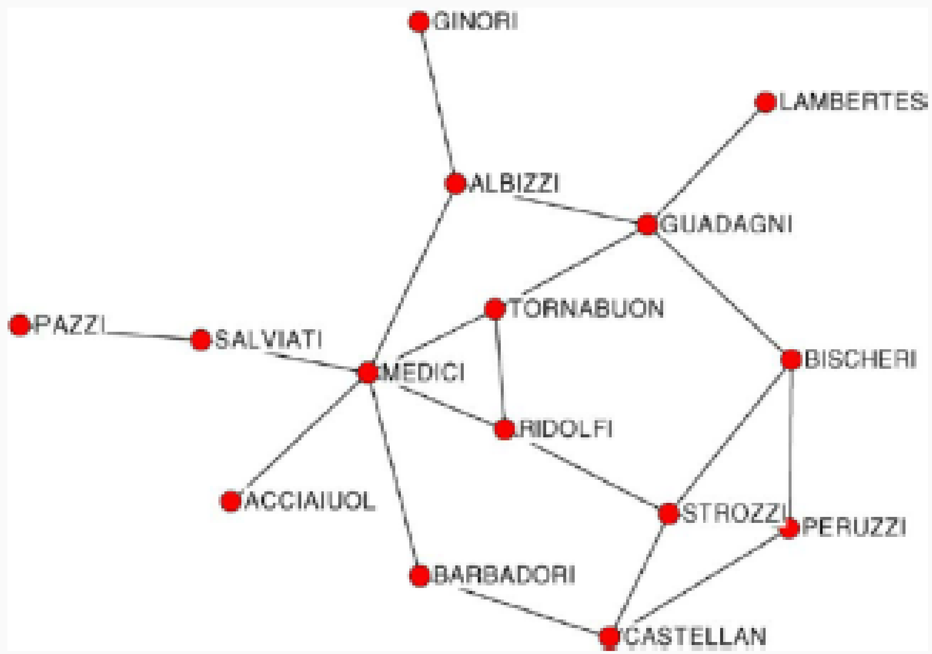


Figure 4: Marriages between Florentine families during the Renaissance.

Social Networks: Archival or Third-Party Data

Computers and online databases have greatly expanded the available archival datasets for social network analysis.

- Email logs
- Email address books
- Telephone Call graphs
- LinkedIn
- Forum threads and comments

Social Networks: Archival or Third-Party Data

Data from longitudinal studies eg. TILDA (The Irish Longitudinal Study on Ageing).

"Overarching aim of which is to make Ireland the best place in the world to grow old"

TILDA collects information on health, economic and social circumstances from people aged 50 and over in a series of data collection waves once every two years.

TILDA is unique amongst longitudinal studies in the breadth of physical, mental health and cognitive measures collected. Most comprehensive research studies of its kind both in Europe and internationally.

https://tilda.tcd.ie/

Social Networks: Affiliation Networks

An affiliation network is a network in which nodes are connected via comembership of some kind of group.

- People attending common social events
- CEOs of companies interacting via clubs
- Board members of companies
- Film actors: "Six degrees of Kevin Bacon"
- Co-authorship networks

https://en.wikipedia.org/wiki/Six_Degrees_of_Kevin_Bacon

Social Networks: Snowball sampling

Studies of some populations such as drug users or illegal immigrants present special problems to investigators. Members do not want to be identified or studied.

Snowball sampling: Get an identified member of the community to provide contact details for other after gaining their trust. Eventually, a network is built by being introduced to friends of friends. (Somewhat like a generalised ego-centric network). The process then "snowballs".

However this produces highly biased network samples with few analytical tools to estimate or correct for the bias.

Social Networks: Contact Tracing

Contact tracing is a form of snowball sampling applied to disease incidence. Some diseases, such as tuberculosis are sufficiently serious that an effort must be made to track down individuals having had significant contact with an infected person.

Contact tracing networks are very significantly biased networks since they contain only those who show symptoms/seek treatment/have been identified by others.

To correct for these sampling biases, we can perform random walks on the data since we know that sampling probabilities depend on degree only. In this way, we can scale the sample statistics to a population.

Interference graph

$$E(\sigma) = -J \sum_{\langle i,j \rangle} \delta_{\sigma_i \sigma_j}$$

$$J = \begin{cases} +1 & \text{Ferromagnetic,} \\ \boxed{-1 \quad \text{Antiferromagnetic.}} \end{cases}$$

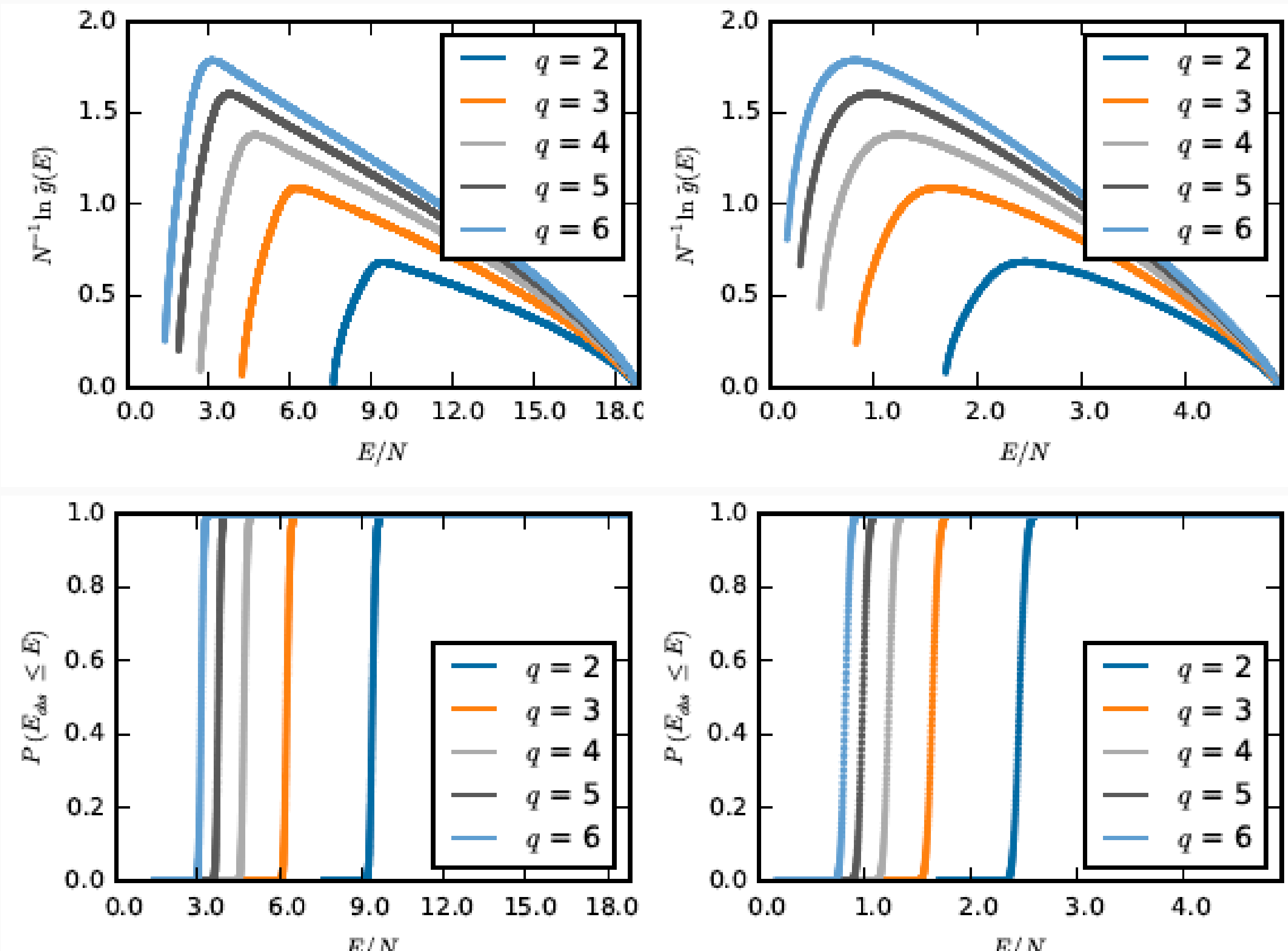
- Application of the Potts model to interference in wireless networks
- Spectrum allocation *microstates* contribute to interference level *macrostate*
- Multiple microstate configurations may result in the same macrostate (**degeneracy**)
- The degeneracy of the interference levels is numerically estimated
  - Results show that the majority of the configurations occur over a very short interference range
- Results are based on a real-world cellular network deployment

Degeneracy & interference

θ = 1 θ = 5

Degeneracy vs. Interference

Degeneracy CDF vs. Interference



N. McBride, J. Bulava, C. Galiotto, N. Marchetti, I. Macaluso, L. Doyle, "Degeneracy Estimation in Interference Models on Wireless Networks", *Elsevier Physica A: Statistical Mechanics and its Applications*, vol. 469, pp. 540–550, Mar 2017

Interference topology & design

- A Monte Carlo study of interference in real-world wireless networks using the Potts model, that maps the Potts energy to discrete interference levels

- The **CDF** of the resulting density of states **is steep**
  - Small changes in the acceptable level of interference will result in dramatically **larger number of acceptable spectrum configurations**

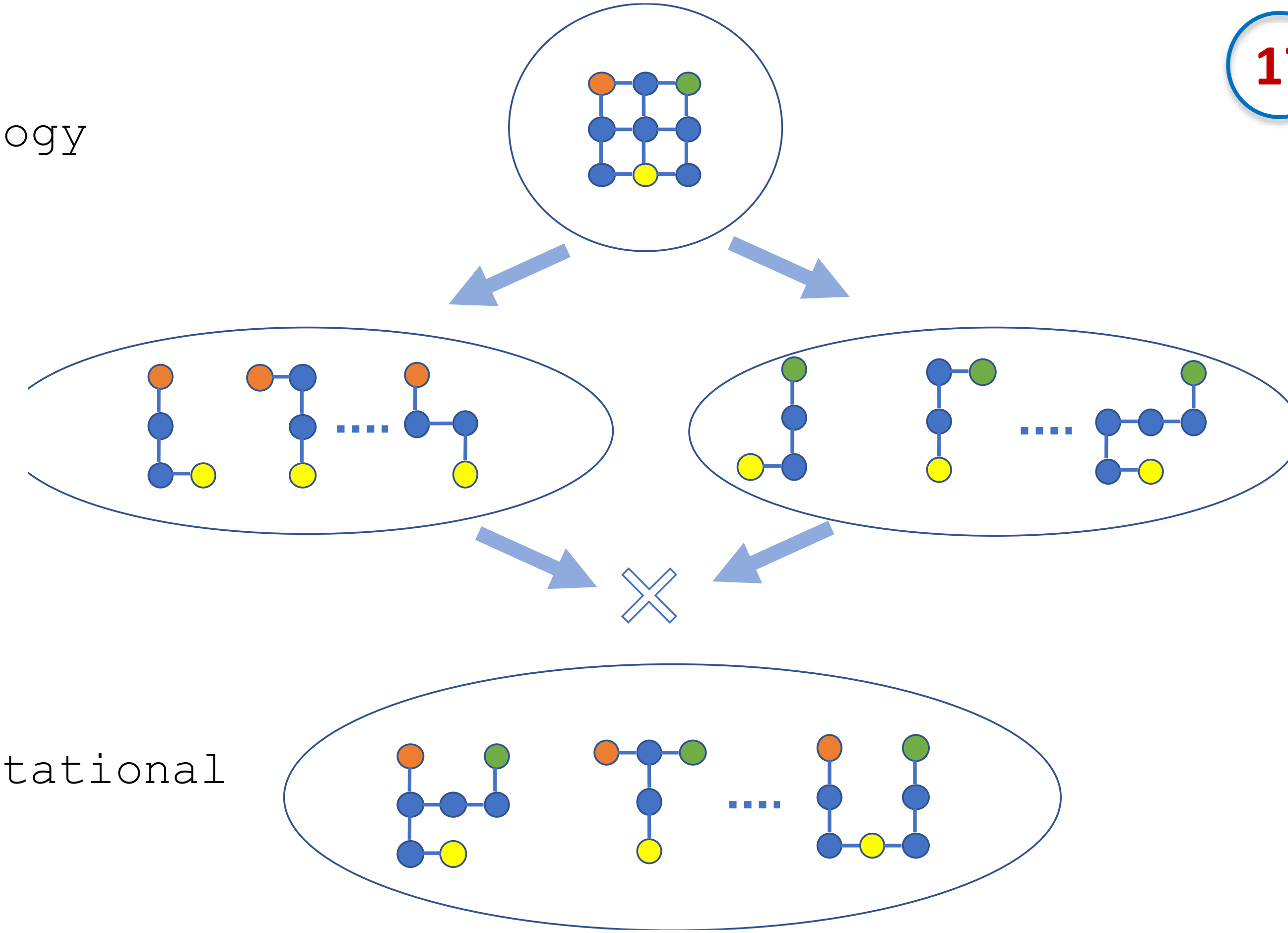
- Models with a greater number of available frequency channels and less dense interference networks result in:
  - The majority of configurations having lower interference levels
  - Their critical interference levels occurring at lower values

In-network computing in IoT

Physical topology

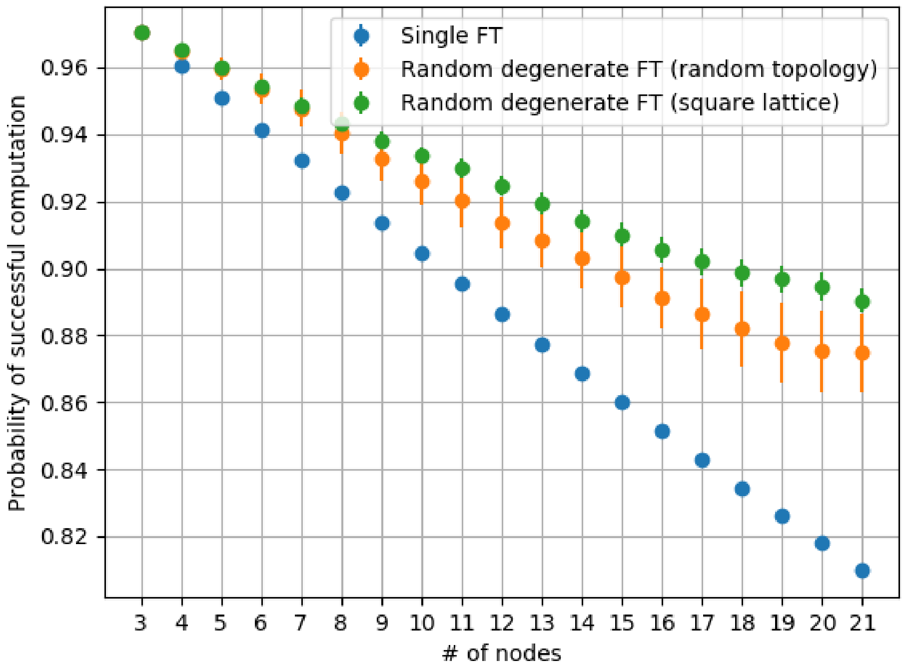
Set of paths between the sink and target (input) nodes

Functional/computational graph



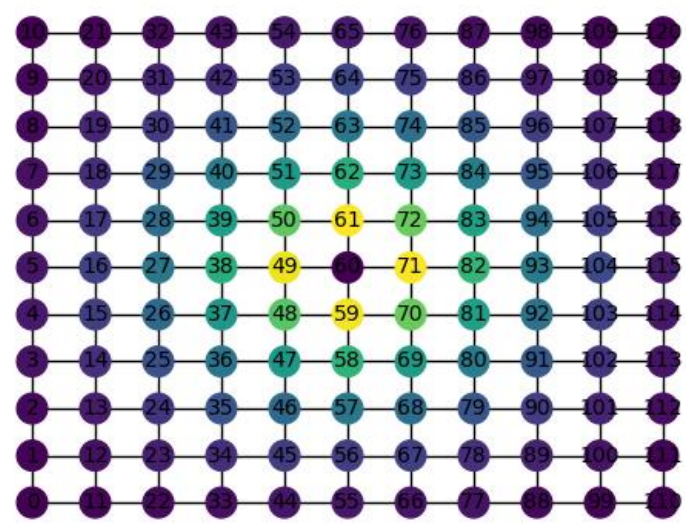
M. Dzaferagic, N. McBride, R. Thomas, I. Macaluso, N. Marchetti, “Improving In-Network Computing in IoT Through Degeneracy”, *IEEE Systems Journal*, vol. 15, no. 1, pp. 238-244, Mar 2021

In-network computing in IoT



$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_n$$

$$P_{S_c} = \prod_{i \in U} (1 - P_{f_i})$$



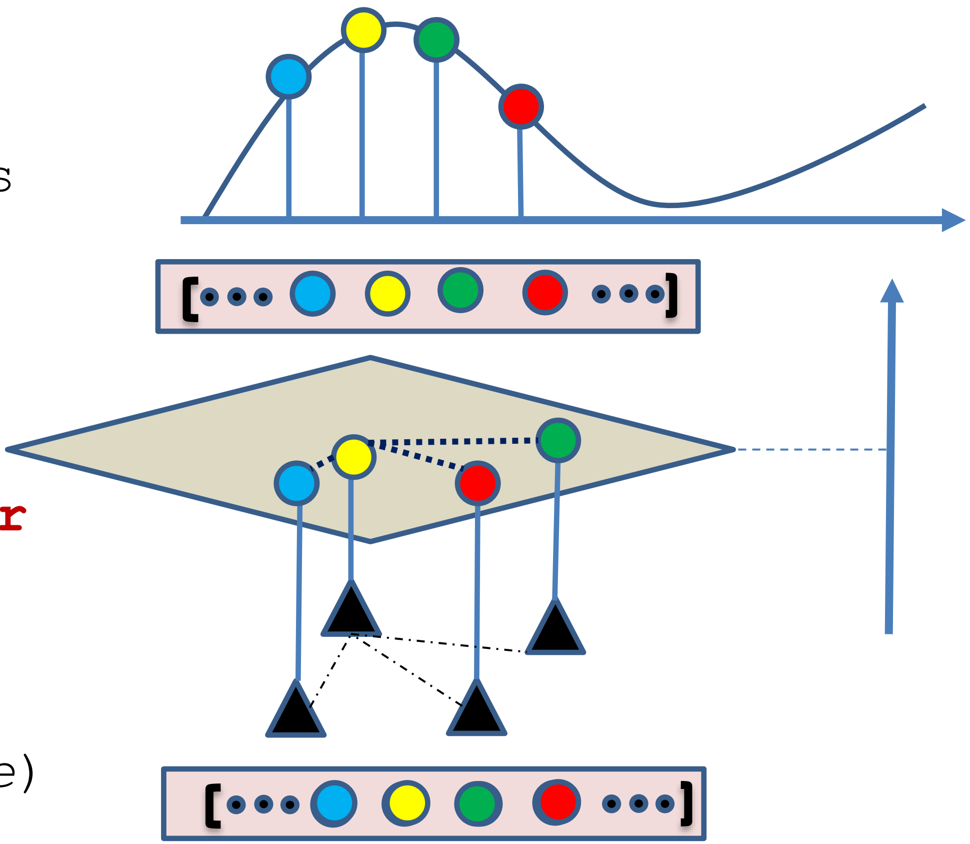
- Heat-map of the square lattice topology showing the probability of a node occurring in any feasible functional topology

- Brighter colours represent higher probabilities

Graph Signal Processing (GSP) - Motivation

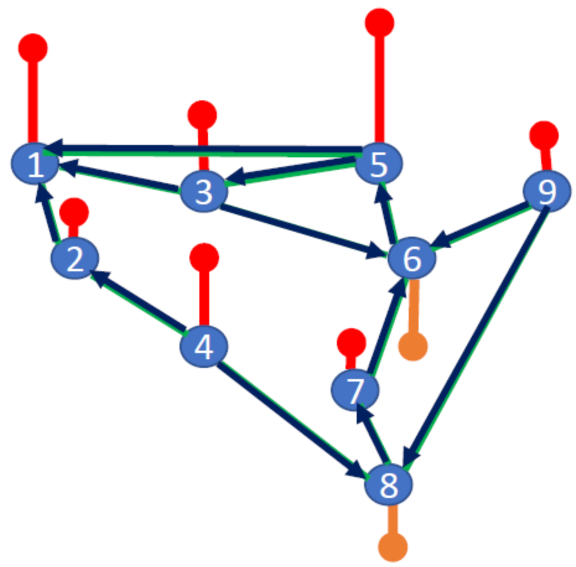
- Signal with a **regular** structure
  - Ordering (in *time*) is important

- Signal with an **irregular** structure
  - Ordering is not important
  - Correlation (in *space*) is important



GSP is a framework for network & signal modeling & processing

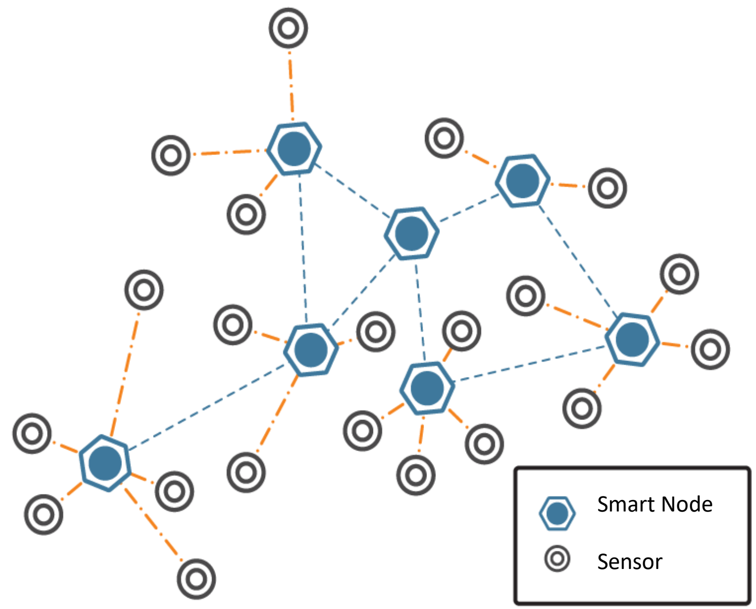
- **Graph**
  - A collection of edges and vertices
- **Signal on graph**
  - A function defined on the vertices of the graph
  - Irregular signal on graph
- **Processing tool: GSP**
  - Spectral interpretation



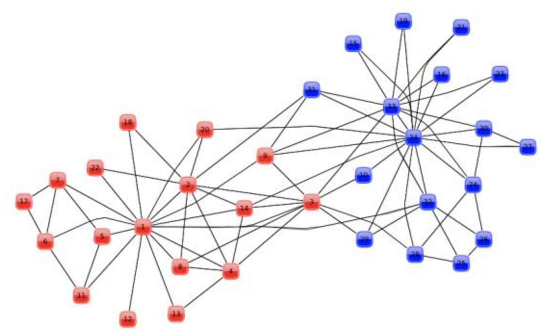
A. Miraki, H. Saeedi-Sourck, N. Marchetti, A. Farhang, "Spectral Domain Spline Graph Filter Bank", *IEEE Signal Processing Letters*, vol. 28, pp. 469-473, Feb 2021

Applications

- **Sensor networks**
  - ➢ Wireless sensor networks, Energy grid networks
  - ➢ Relative positions of sensors

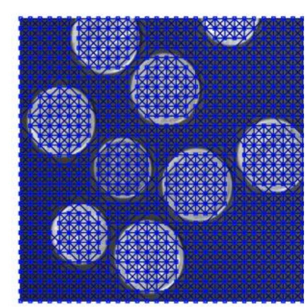


- **Social networks**
  - ➢ Friendship relationship



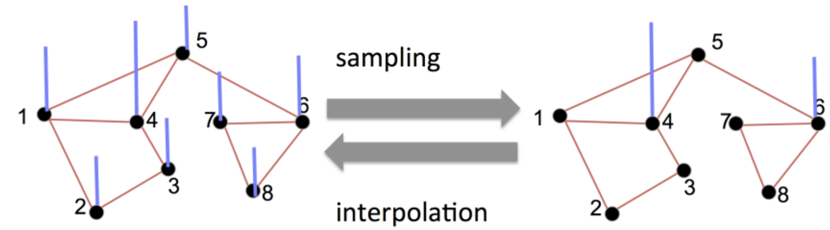
- **Biological networks**

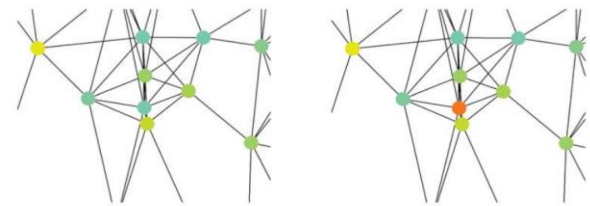
- **Images and computer graphics**
  - ➢ Pixels values and similarity



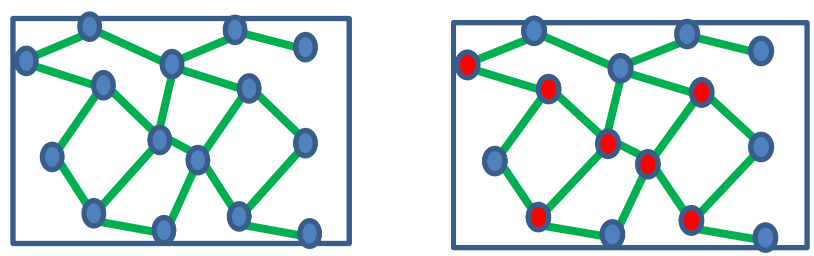
Applications of GSP to Wireless Sensor Networks

- Sampling theory
  - Turn on/off sensors
  - Sensor position selection
- Anomaly detection
- Topology inference
  - Brain signal processing
- *Compression*



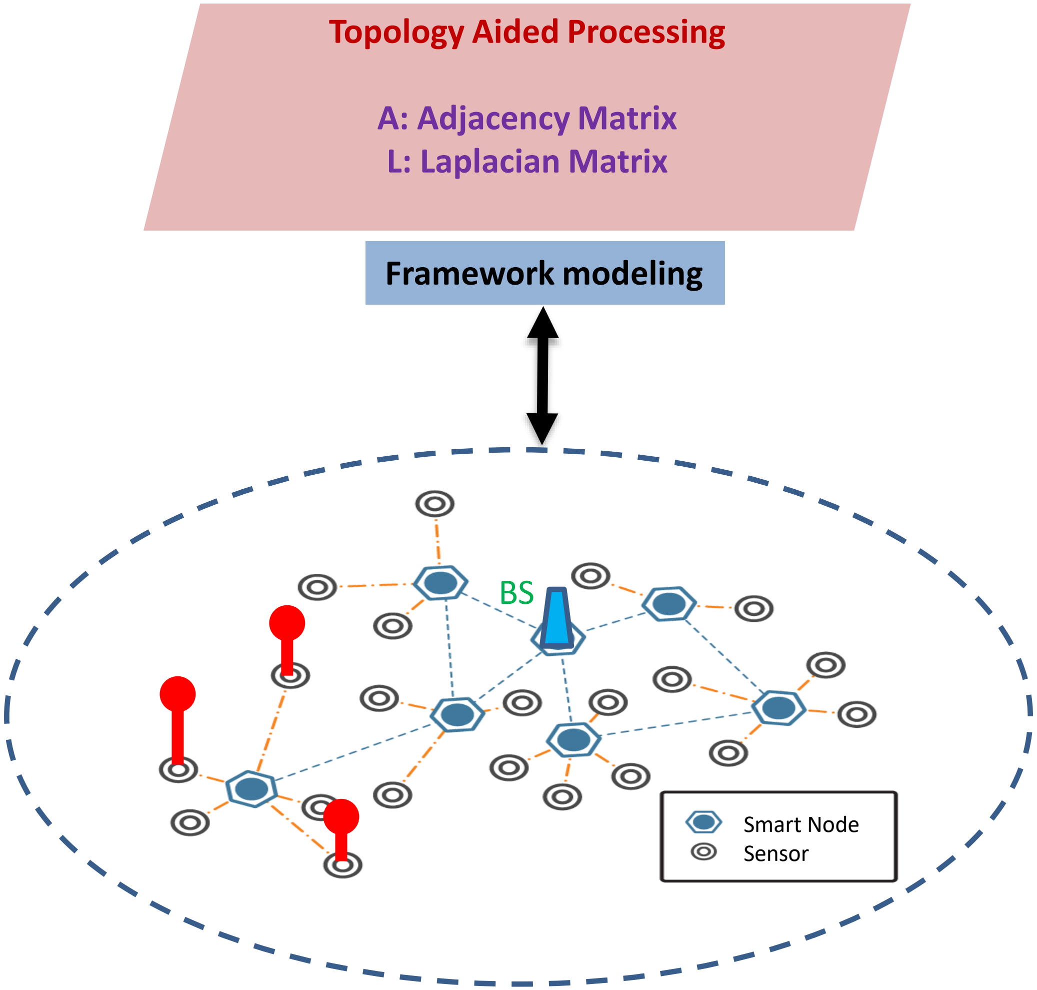


(a) True measurement (b) Corrupted measurement



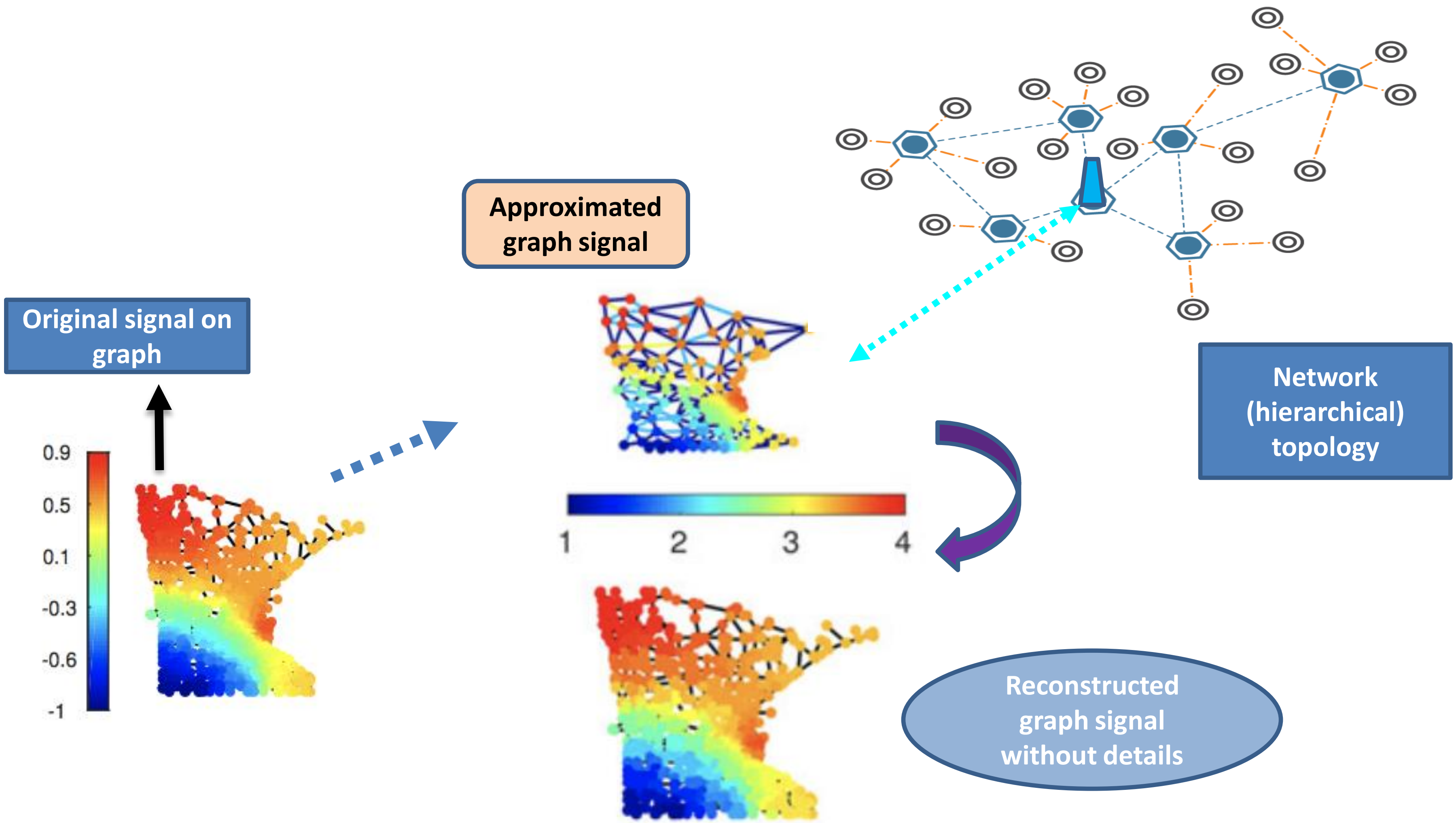
Topology-aided signal processing

- Assumption
  - *Smooth* signal on graph
- Goals
  - Distributed processing using smart nodes
  - Improving energy efficiency



A. Chiumento, N. Marchetti, I. Macaluso, “Energy Efficient WSN: a Cross-layer Graph Signal Processing Solution to Information Redundancy”, *International Symposium on Wireless Communication Systems (ISWCS)*, Aug 2019

Topology-aided signal processing



Ackowledgement

- Prof Hamid Saeedi-Sourck, Yazd University, Iran
- Dr Irene Macaluso, CableLabs, USA
- Dr Merim Dzaferagic, TCD
- Dr Neal McBride, Arista Networks