Designing Experiments on Networks

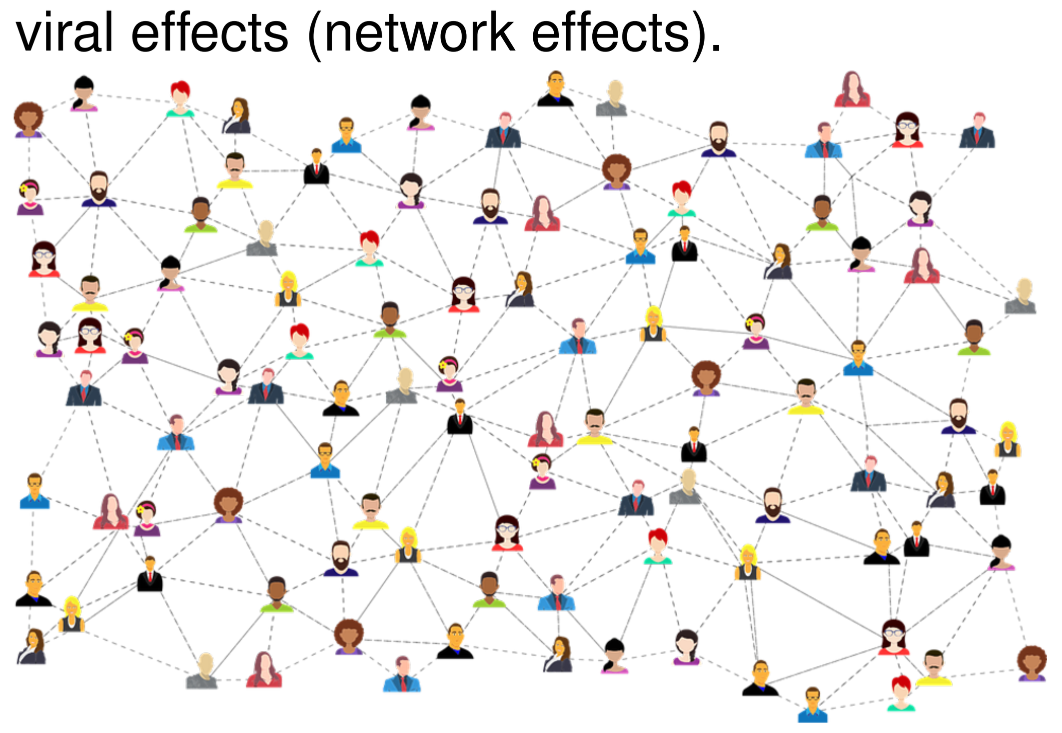
Vasiliki Koutra

Department of Mathematics, King's College London, UK

Motivation

Experiments on networks are everywhere in our everyday life

- A hospital campaign wishes to compare different patient information leaflets to choose which one will minimise infection rates.
- An experiment is run on a contact network of patients, with a survey used to identify interpersonal contacts.
- Patients might share the leaflets and therefore messages could have viral effects (network effects).



Credit: *pixabay*

1. How will the leaflets affect patients' responses? Are their responses affected by leaflets sent to their contacts?
2. If the network is large and complex, can we simplify it and then find an efficient way to arrange the leaflets to the patients?

Background

Fundamental challenge in science

- Good decisions depend on data.
- Gold standard for data collection is experimentation.
- Design of experiments maximises information content in data.
- Classical experimental designs assume "non-interference".
- Existing design approaches for large networks are too computationally expensive to be employed in practicable time.

Methodological issues and Objectives

Principled mathematical framework

Network: subjects available for experimentation form a network.

- Network (known) is represented by a graph.
- Adjacency matrix indicates the connections between vertices.

Symmetry in networks: a graph automorphism is a form of symmetry.

- Networks have nontrivial automorphism groups, within which permutation of vertices does not alter the network structure.
- Vertex orbits: groups of vertices play the same structural role.

Design approach: model treatment interference and optimise the design for that model.

- Search across many possible designs (exchange algorithm).
- Assess and select design based on minimising the variance of treatment comparisons (objective function).

Objectives:

- Find optimal designs for estimating, either separately or jointly, the direct and indirect treatment effects.
- Speed up the search for finding efficient designs in large and complex networks by utilising network symmetry.

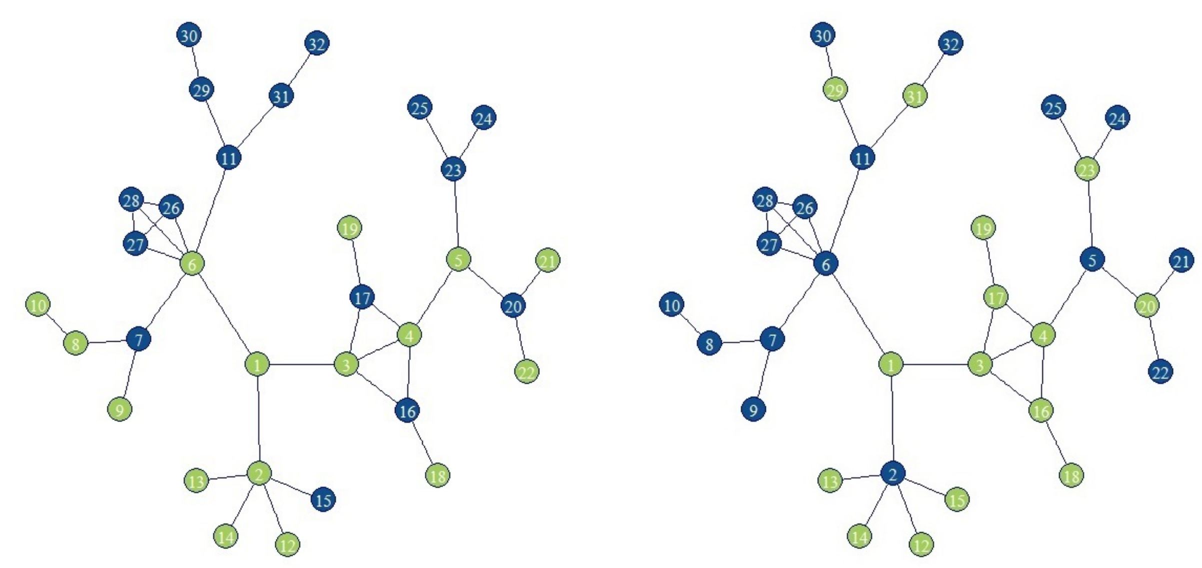
Example

Critical difference from previous methods

Optimal designs with network effects

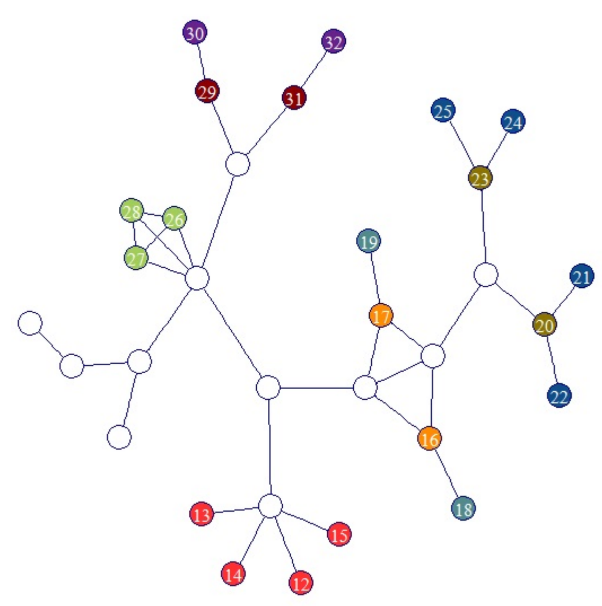
Efficient designs for estimating with minimum variance:

1. the difference in treatment effects;
2. the difference in network effects.



Symmetry-breaking and design

- The skeleton, comprising the white vertices, is asymmetric.
- Vertices of same colour correspond to vertex orbits.
- The skeleton is significantly smaller than the original network.



Design efficiencies for estimating the treatment and network effects, are 100% and 91% compared to the global optimal designs, respectively.

Computational complexity is considerably decreased using symmetry-breaking, as the size of the design space is reduced.

Summary

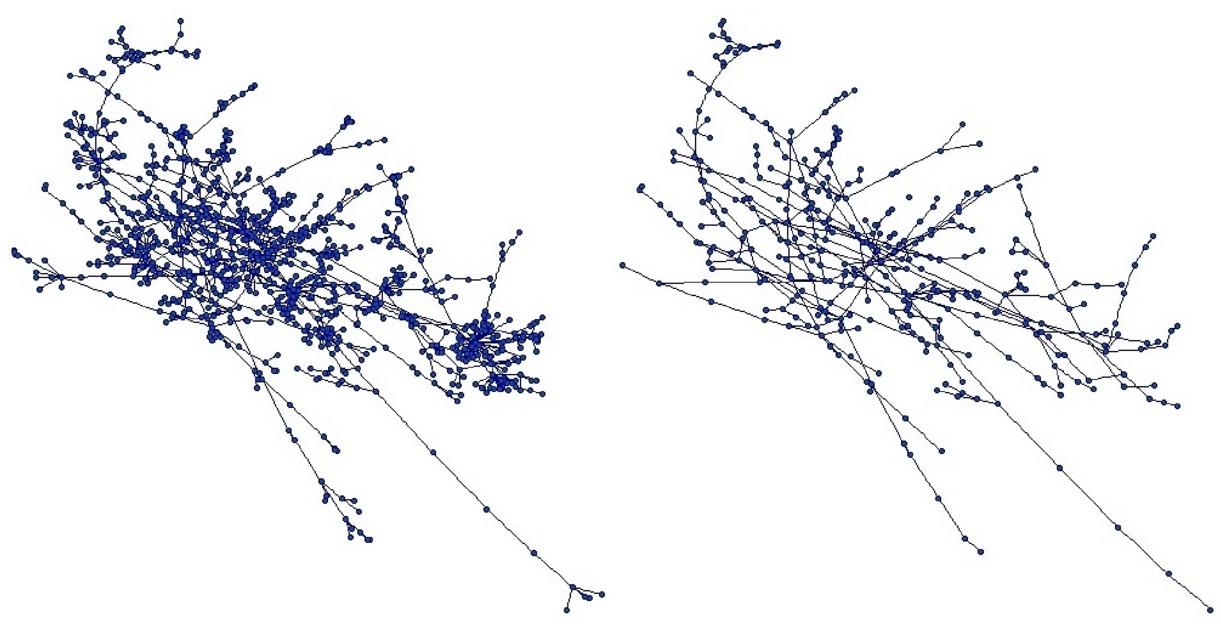
Benefits and impact

Benefits:

- Effective decision making on large networks in seconds.
- High quality data, low costs.
- Both direct and indirect comparisons between treatments.

Impact:

- Design methodology for a wide class of networked experiments.
- Wide academic, economic and societal impact.
- Potential to revolutionise the way experiments on networks are conducted. Real-world networks are richly symmetric.



Bibliography

- Koutra, V., Gilmour, S.G., and Parker, B.M. (2021). Optimal block designs for experiments on networks, Journal of the Royal Statistical Society, Series C, 70, 596–618.
- MacArthur, B., Sánchez-García, R., and Anderson, J. (2008). Symmetry in complex networks. Discrete Applied Mathematics, 156, 3525–3531.
- Mead, R., Gilmour, S.G. and Mead, A. (2012). Statistical Principles for the Design of Experiments. Applications to Real Experiments. Cambridge University Press.
- Parker, B.M., Gilmour, S.G., and Schormans, J. (2017). Optimal design of experiments on connected units with application to social networks, Journal of the Royal Statistical Society, Series C, 66, 455–480.

vasiliki.koutra@kcl.ac.uk