5th European Conference on Social Networks
SNaples 2021

MANCHESTER 1824
The University of Manchester

# Multiplexity Analysis of Networks using Multigraph Representations

Termeh Shafie & David Schoch
The Mitchell Centre for Social Network Analysis
The University of Manchester

# multigraph representation of network data

MANCHESTER 1824 The University of Manchester

*graphs where multiple edges and self-edges are permitted*

- can appear directly in applications (although scarce)
- can be constructed by different kinds of aggregations in graphs


**multiplexity analysis with respect to both vertex and edge attributes**

- **exploratory**

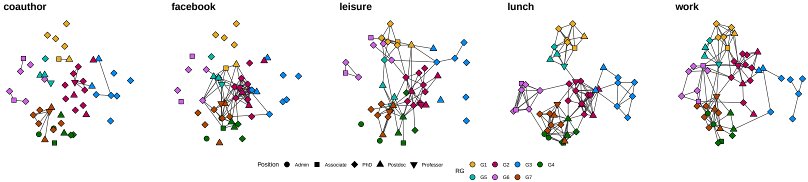
by using visual tools on the joint and marginal distribution of edge types

- **confirmatory**

by using multiplexity statistics under probability models for multigraphs

## the AUCS dataset

a multivariate network with multiple types of ties and vertex attributes

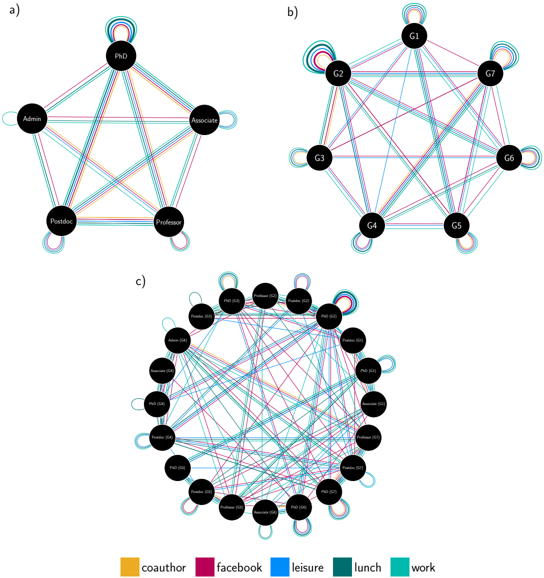



- five types of relations of the considered network dataset
- vertex attributes are research group (RG) and academic position

**aggregation based on single or combined vertex attributes**
$\Longrightarrow$ **three multigraphs**

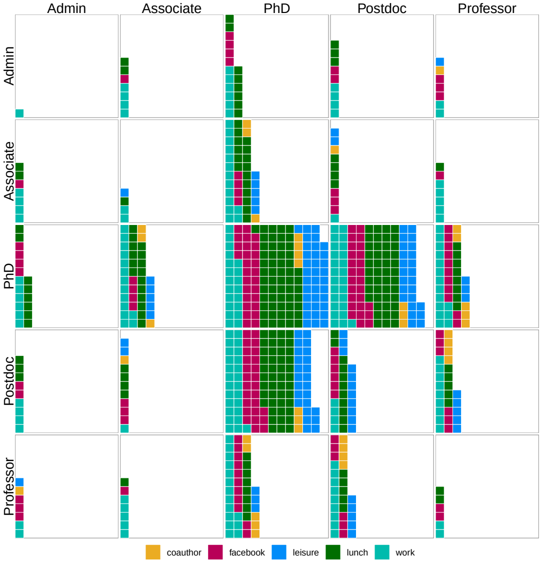
# example: aggregated multigraphs





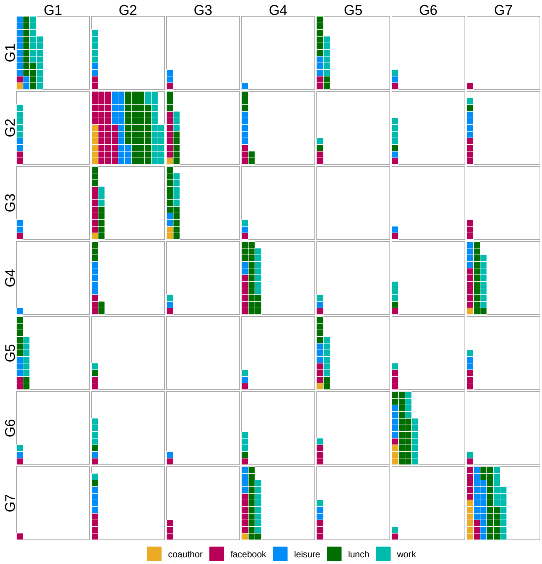
# example: waffle matrices

The University of Manchester



# example: waffle matrices

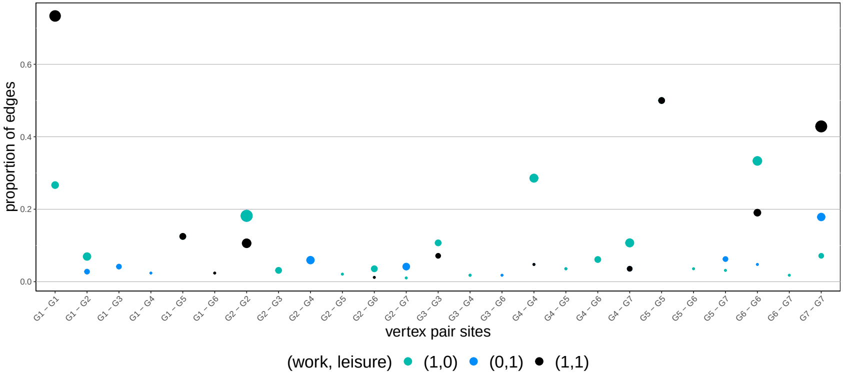
MANCHESTER 1824
The University of Manchester



# exploring distribution of edge types



systematically analyse the joint and marginal distribution of dyads within and between categories



# multiplexity statistics



- multigraphs represented by their edge multiplicity sequence

$$\mathbf{M} = (M_{ij} : (i,j) \in \mathcal{R})$$

where $\mathcal{R}$ is the canonical site space for undirected edges given by

$$\{(i,j) : 1 \leq i \leq j \leq n\}$$

- the number of vertex pair sites is given by $r = \binom{n+1}{2}$

- distribution of edge multiplicities as a sequence $\mathbf{R} = (R_1, R_2, \ldots, R_k)$ where

$$R_k = \sum_{i<j}\sum I(M_{ij} = k) \quad \text{for} \quad k = 0, 1, \ldots, m$$

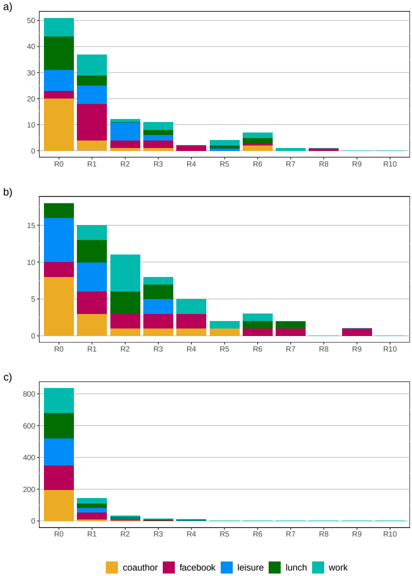
  - $R_0$ number of vertex pair sites with no edge occupancy
  - $R_1$ number of vertex pair sites with single edge occupancy
  - $R_2$ number of vertex pair sites with double edge occupancy

    ⋮

Shafie, T. (2015). A Multigraph Approach to Social Network Analysis. *Journal of Social Structure*, 16.

# example: observed edge multiplicities





# random multigraph models

MANCHESTER 1824 The University of Manchester

1. **random stub matching (RSM)**
   - edges are assigned to sites given fixed degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$
   - probability that an edge is assigned to site $(i,j) \in \mathcal{R}$

$$Q_{ij} = \begin{cases} \binom{d_i}{2} \Big/ \binom{2m}{2} & \text{for } i = j \\ d_i d_j \Big/ \binom{2m}{2} & \text{for } i < j \end{cases}$$

2. **independent edge assignment (IEA)**
   - edges are independently assigned to vertex pairs in site space $\mathcal{R}$
   - edge assignment probabilities $\mathbf{Q} = (Q_{ij} : (i,j) \in \mathcal{R})$
   - $\mathbf{M}$ is multinomial distributed with parameters $m$ and $\mathbf{Q}$
   - statistics for analysing local and global structure are easily derived
   - can be used as an RSM approximation when $\mathbf{Q} = \mathbf{Q}(\mathbf{d})$

---

Shafie, T. (2016). Analyzing local and global properties of multigraphs. *The Journal of Mathematical Sociology*, 40(4), 239-264.

# central moments of $R_k$ under IEA model



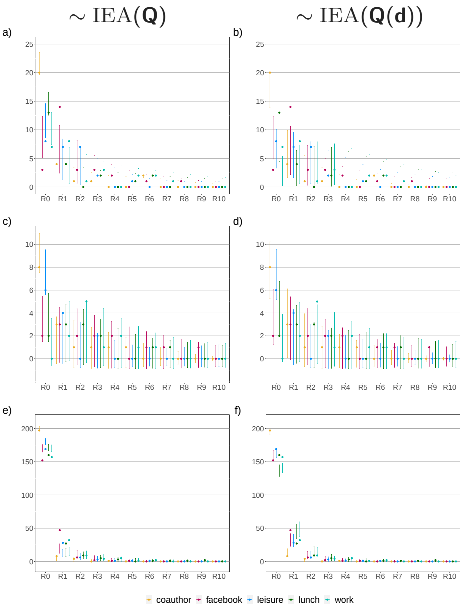
expected values and variance of $R_k$ are derived and estimated

- **the IEA model:** $\sim \mathrm{IEA}(\mathbf{Q})$
  MLE of the edge assignment probabilities given by the empirical fraction of each edge type
- **the IEA approximation of the RSM model:** $\sim \mathrm{IEA}(\mathbf{Q}(\mathbf{d}))$
  edge assignment probabilities given by the observed degree sequence of each edge type

intervals $\hat{E} \pm 2\sqrt{\hat{V}}$ illustrated

# example: interval estimates



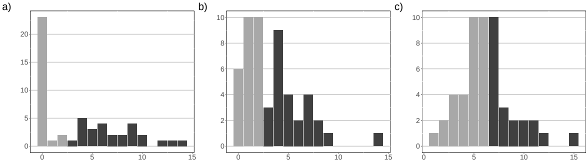


# extensions to presented framework

MANCHESTER 1824
The University of Manchester

**not contingent on the presence of vertex attributes**

observed edge variables can be transformed into vertex variables




define online and offline social influence by using degree distributions
a) facebook, b) leisure and c) lunch

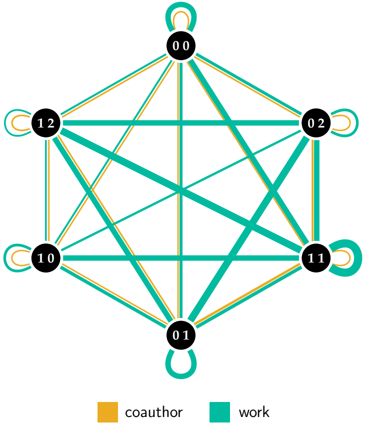
| | outcome space for social influence | |
|---|---|---|
| online = (facebook) | offline = (lunch, leisure) | (online, offline) |
| 0 | 0 = (0,0) | (0,0) |
| 1 | 1 = (0,1) or (1,0) | (0,1) |
| | 2 = (1,1) | (0,2) |
| | | (1,0) |
| | | (1,1) |
| | | (1,2) |

vertices in aggregated multigraph

# extensions to presented framework



multigraph with coauthor and work relations moving within and between categories based on online and offline social influence



# final remarks



**some cautionary words**

- let research question and social theories guide data transformations
- attention to density of various edges and vertex variable distributions

**limitations of presented framework**

- only applicable to undirected networks
- visual inspections of waffle matrices are only feasible for small networks
- direction of associations between different edge types not revealed

# implemented in R

https://github.com/termehs/multigraphr

```
# install.packages("devtools")
devtools::install_github("termehs/multigraphr")
```