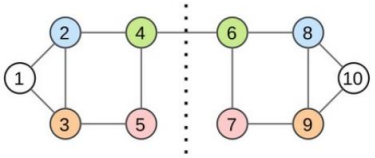
Universiteit Leiden
The Netherlands

Towards Statistical Disclosure Control for Complex Networks

R.G de Jong, M.P.J. van der Loo (mpj.vanderloo@cbs.nl), and F.W. Takes

New Techniques and Technologies for Statistics, Brussels, March 12 2025

Anonymity in Complex Networks



ANO-NET

Universiteit Leiden POPNET Centraal Bureau voor de Statistiek

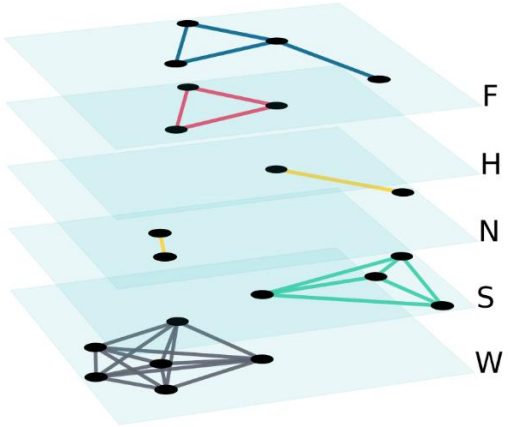
Rachel de Jong (1,2)

Mark van der Loo(2,1)

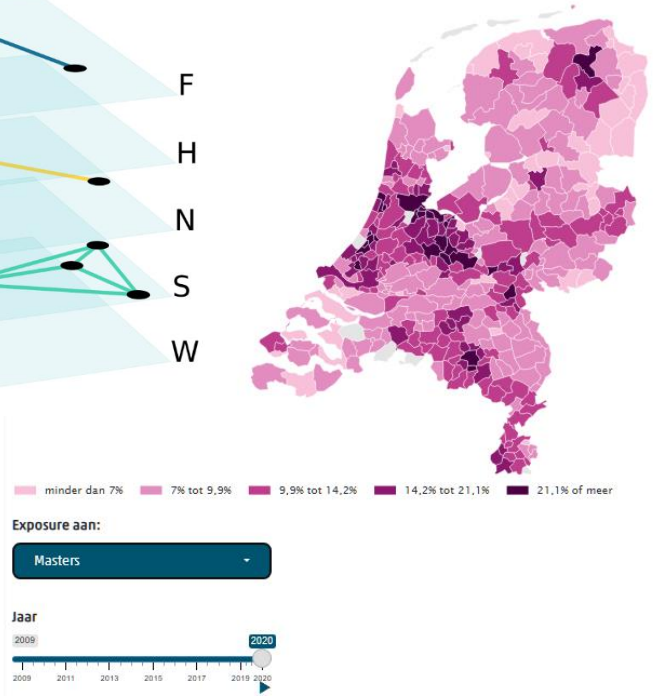
Frank Takes (1)

(1) Computational Network Science Group at the Leiden Institute of Advanced Computer Science
(2) Statistics Netherlands

Motivation



Family
Household
Neighbours
School
Work



https://dashboards.cbs.nl/v4/opl_segregatie/

European Sociological Review, 2023, **39**, 145–160
https://doi.org/10.1093/esr/jcac026
Advance access publication 11 June 2022
Original article

OXFORD

A Whole Population Network and Its Application for the Social Sciences

Jan van der Laan[1], Edwin de Jonge[1], Marjolijn Das[1,2], Saskia Te Riele[1] and Tom Emery[2,*]

scientific reports

Check for updates

OPEN

The anatomy of a population-scale social network

Eszter Bokányi[1,2✉], Eelke M. Heemskerk[1] & Frank W. Takes[2,3]

Menyhért et al. *EPJ Data Science* (2025) 14:8
https://doi.org/10.1140/epjds/s13688-025-00522-4

EPJ Data Science
a SpringerOpen Journal

EPJ.org

RESEARCH Open Access

Connectivity and community structure of online and register-based social networks

Márton Menyhért[1†], Eszter Bokányi[2*†], Rense Corten[3], Eelke M. Heemskerk[2], Yuliia Kazmina[2] and Frank W. Takes[1]

Contents lists available at ScienceDirect
Social Networks
journal homepage: www.elsevier.com/locate/socnet

ELSEVIER

Socio-economic segregation in a population-scale social network

Yuliia Kazmina [a,*], Eelke M. Heemskerk [a], Eszter Bokányi [a], Frank W. Takes [b]
[a] University of Amsterdam, The Netherlands
[b] Leiden University, The Netherlands

Research questions

(1) How does **network structure** affect risk of disclosure?

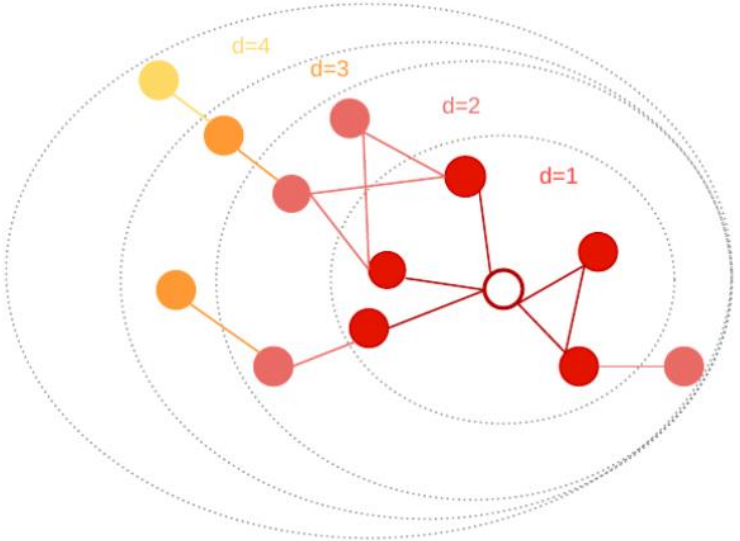
- **Identity**
- Attributes

(2) How can we mitigate disclosure risks?

- Utility

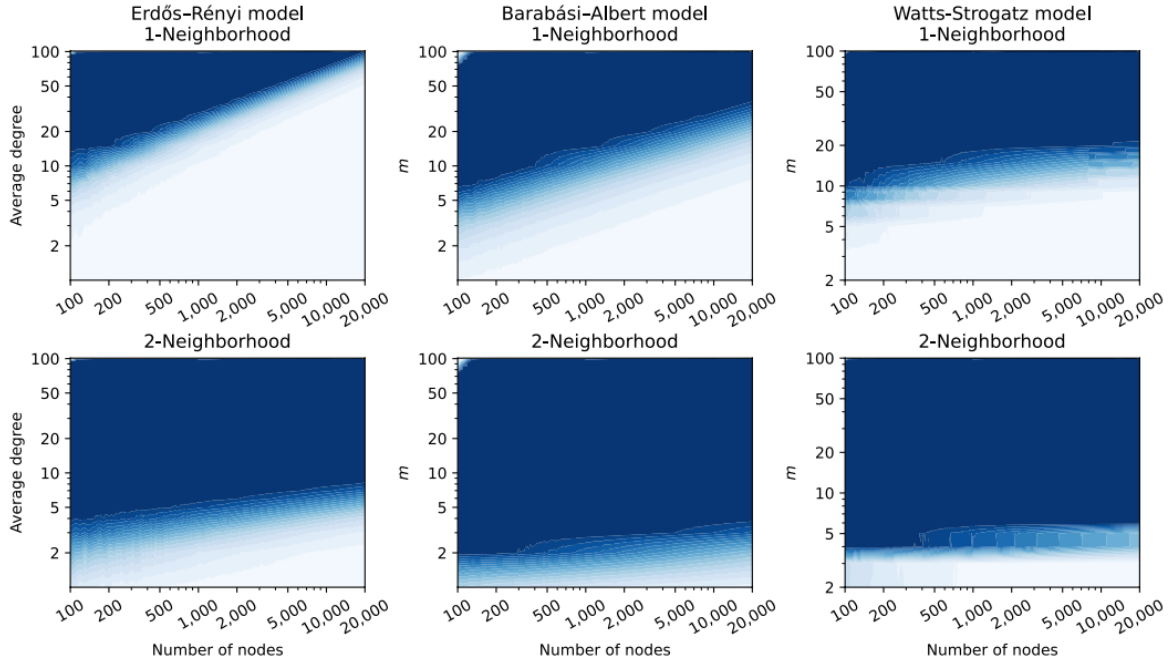
Graded *k*-anonymity: *d-k* anonymity

- Attacker knows *full* network structure surrounding a node, up to and including distance *d*
- *A node is k-anonymous if there are k-1 nodes with the same neighbourhood.*



d-k anonymity

If an attacker knows the neighbourhood up to and including *d*=2, most nodes are unique.



Algorithms for Efficiently Computing Structural Anonymity in Complex Networks

RACHEL G. DE JONG and MARK P. J. VAN DER LOO, Leiden University, Statistics Netherlands, The Netherlands
FRANK W. TAKES, Leiden University, The Netherlands

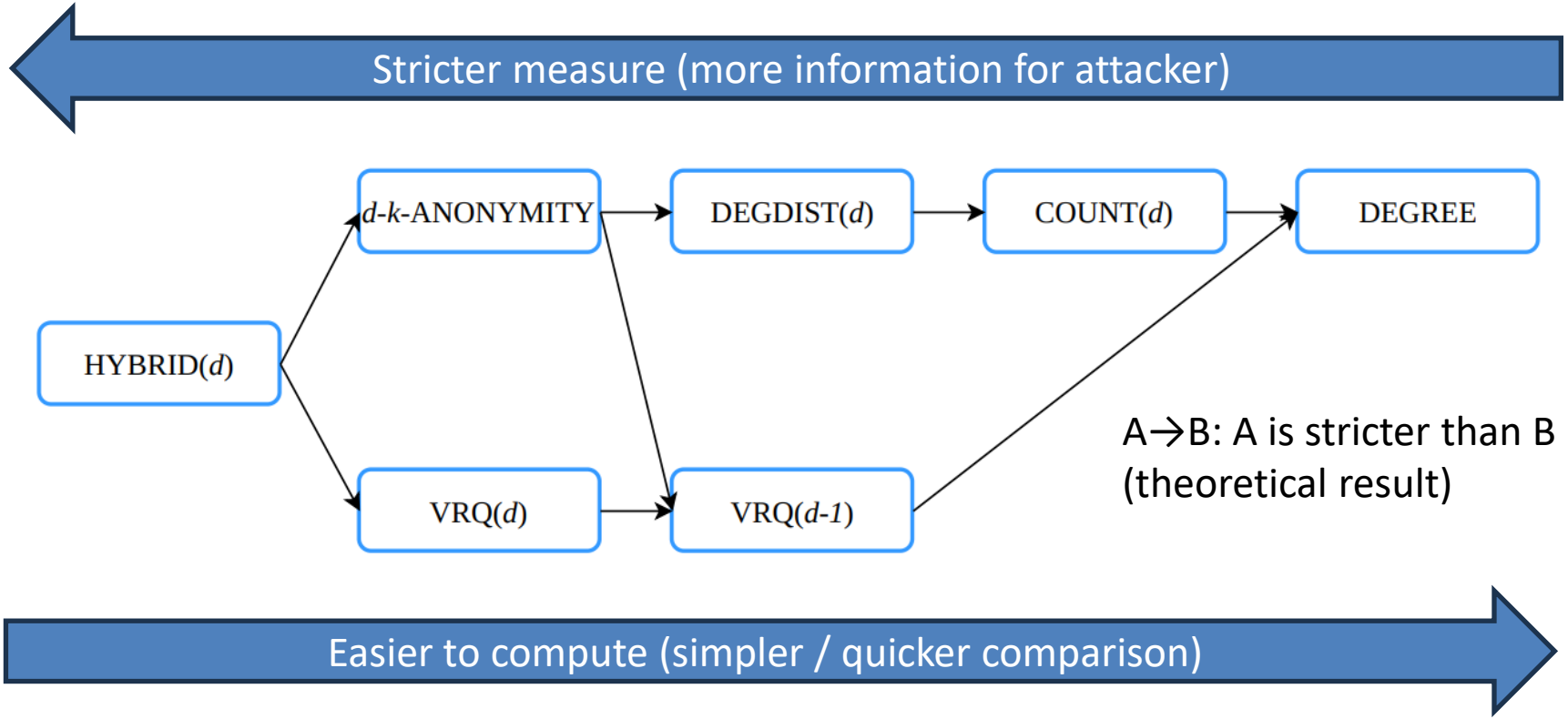
scientific reports

OPEN The effect of distant connections on node anonymity in complex networks

Rachel G. de Jong[1,2✉], Mark P. J. van der Loo[1,2] & Frank W. Takes[1]

https://arxiv.org/abs/2407.02290

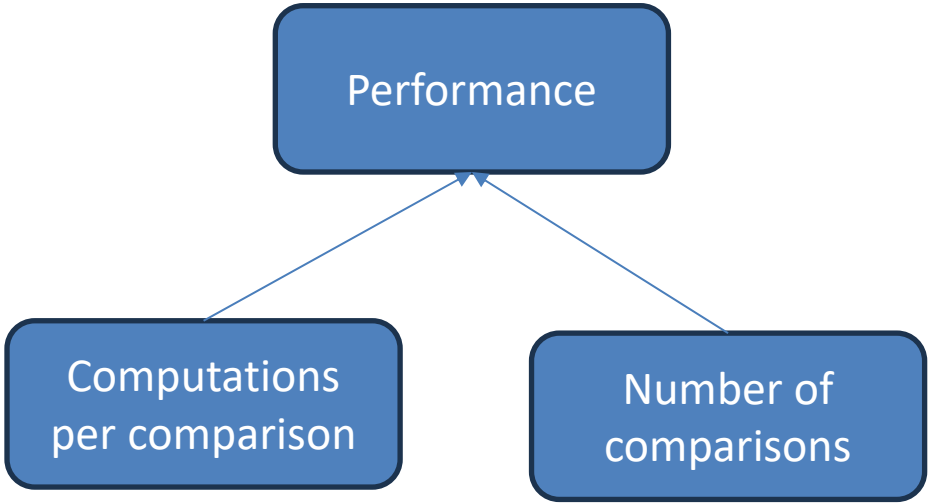
Approximating *d-k* anonymity



https://arxiv.org/abs/2407.02290

Performance matters
(at 18 mln nodes and 1.4 bln edges)

Depending on chosen measure and distance:



COUNT(*d*) is a reasonable approximation for

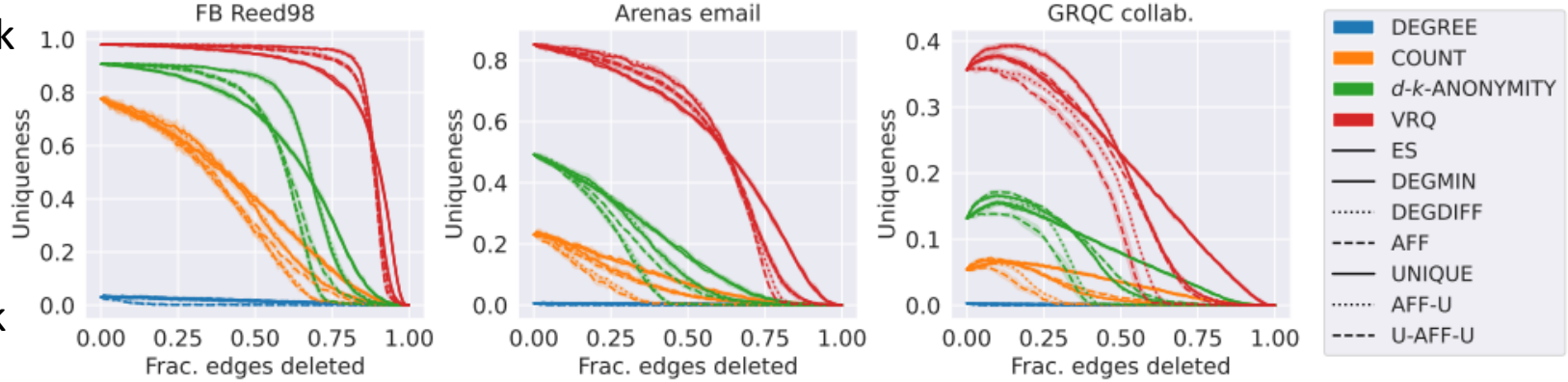
d-*k* anonymity and faster at d=1

https://arxiv.org/abs/2409.16163

Experimenting with anonymization

High risk

Low risk



Fraction of links deleted

- Link deletion is better than adding or swapping
- Targeted deletion works *much* better than random deletion

Take-home messages

Network data is highly interesting for official statistics, and of great interest to the scientific community

We are working on measuring and mitigating disclosure risk when sharing network data

Thank you!

computationalnetworkscience.org
Computational Network Science at Leiden University

markvanderloo.eu
mpj.vanderloo@cbs.nl