

A Simulated Annealing Approach to Social Network Anonymization

E. Denisa Arsene¹
Rachel G. de Jong²
Frank W. Takes²
Anna L.D. Latour¹

¹Department of Software Technology,
Delft University of Technology

²Leiden Institute of Advanced Computer Science
(LIACS), Leiden University

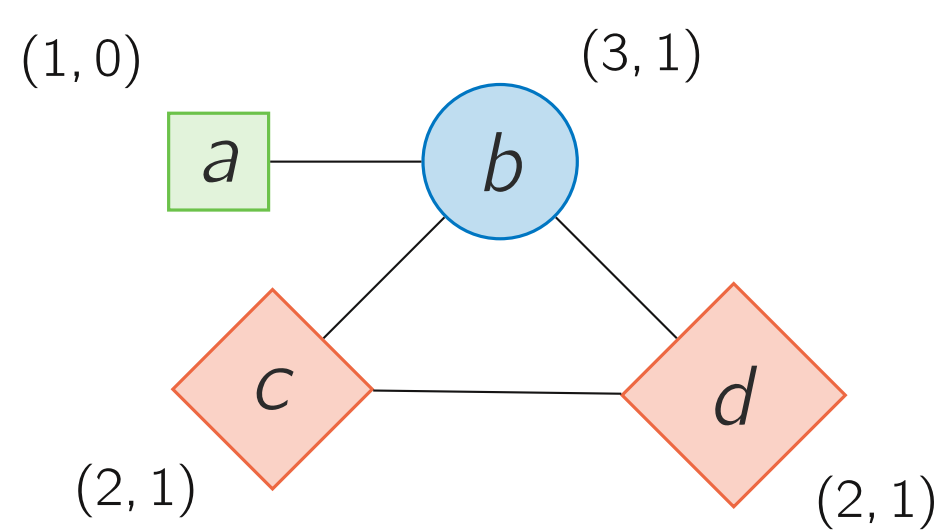
Motivation

- Social graphs risk leaking identity.
- Pseudonymization \neq privacy.
- Anonymize with minimal edge deletions.
- Budgeted anonymization.

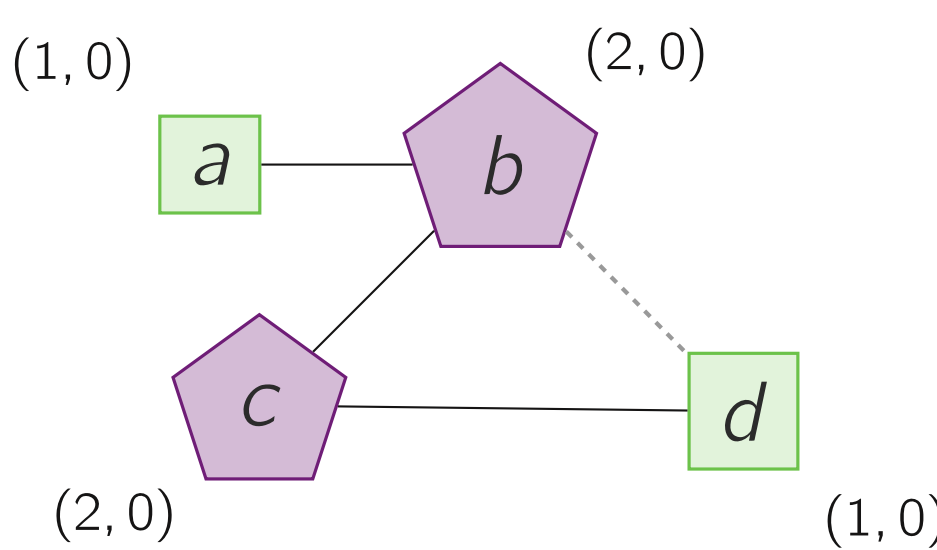
(n, m) -Anonymity

Nodes share degree n and m incident triangles.

Goal: each node in a group $k \geq 2$.

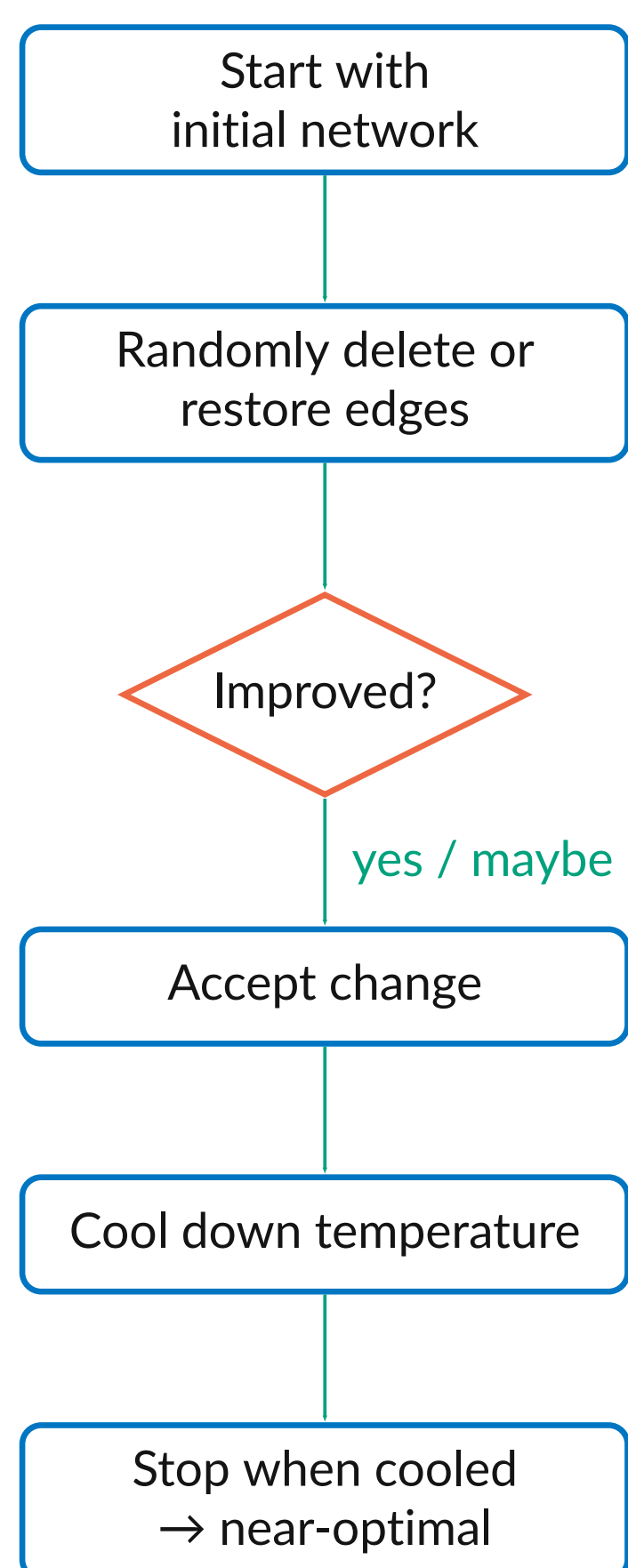


a and b are unique nodes — high re-identification risk.



Removing edge (b, d) makes the network 2-anonymous.

Approach



SANA method flow.

How to share
research-relevant
social network
data, while
preserving privacy?

SANA

Minimal changes.
Maximal privacy.

SANA

Simulated Annealing
for Network
Anonymization

Up to 18× more
anonymized nodes.
Same speed .
Same data utility.



Preprint
paper



GitHub
code

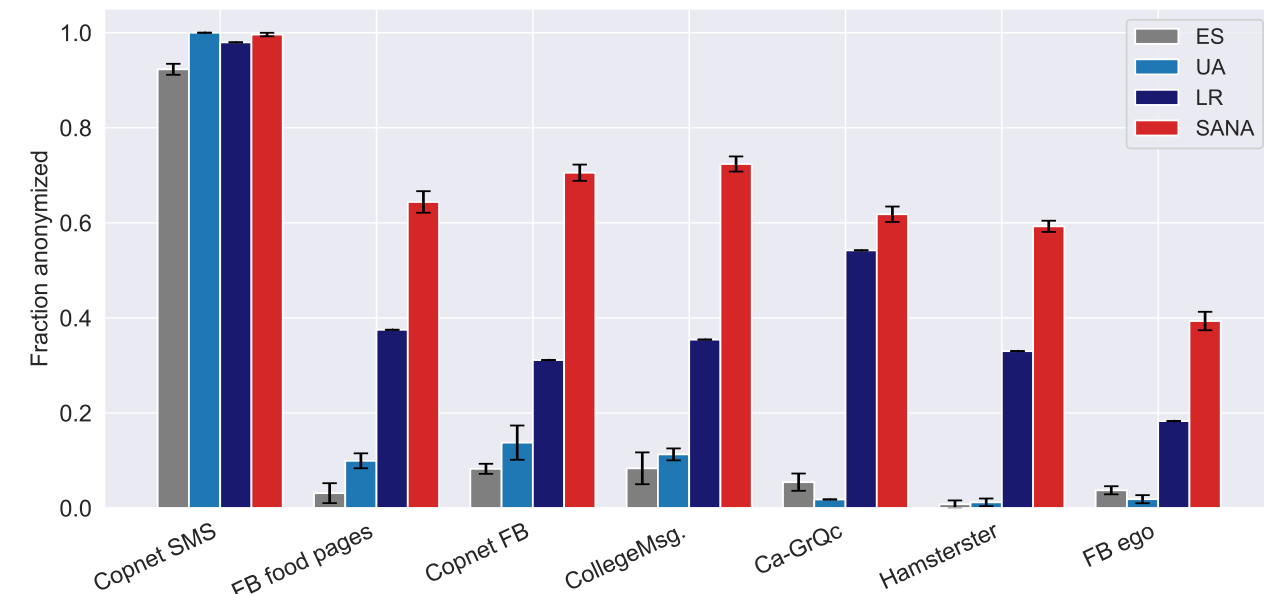
E.D.Arsene@student.tudelft.nl

<https://github.com/arsenedenisa/>

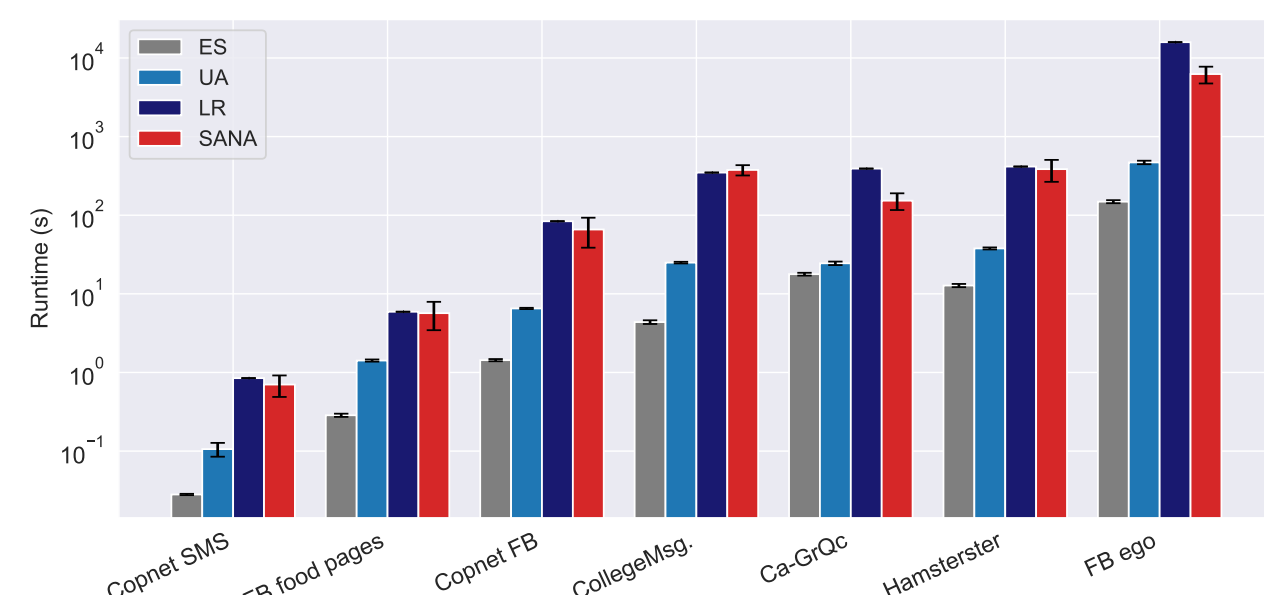
Simulated-Annealing-for-Network-Anonymization

Results

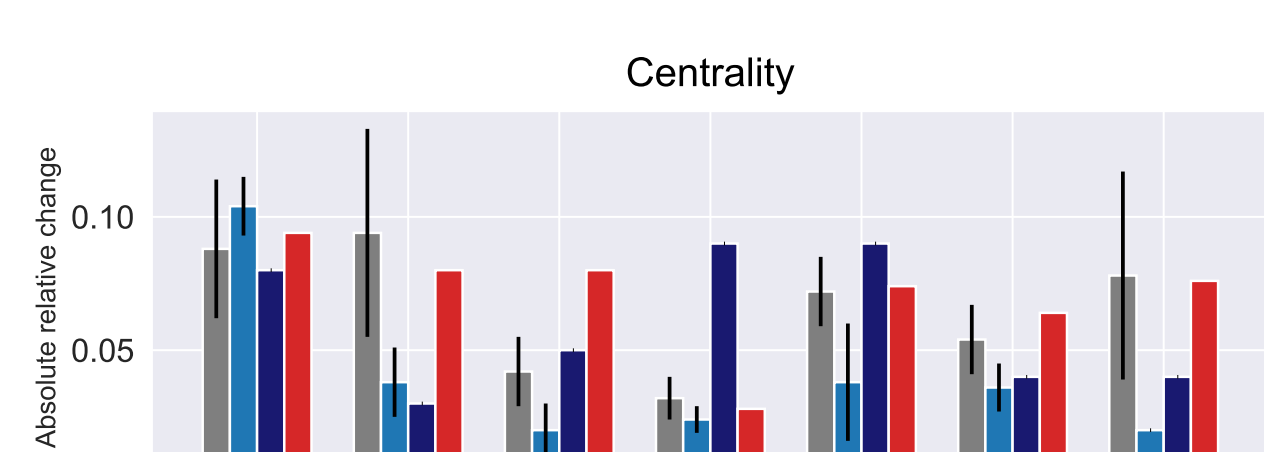
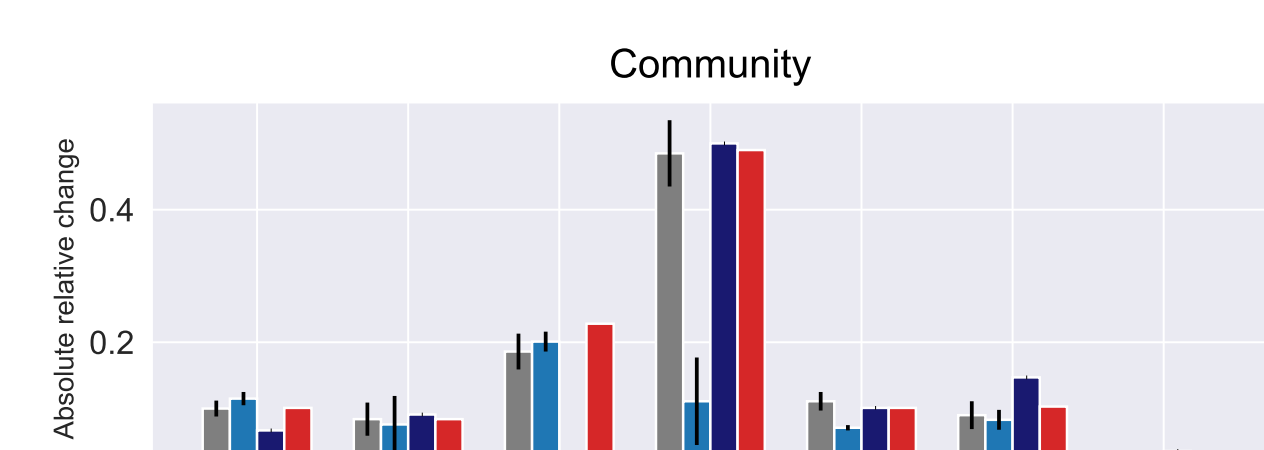
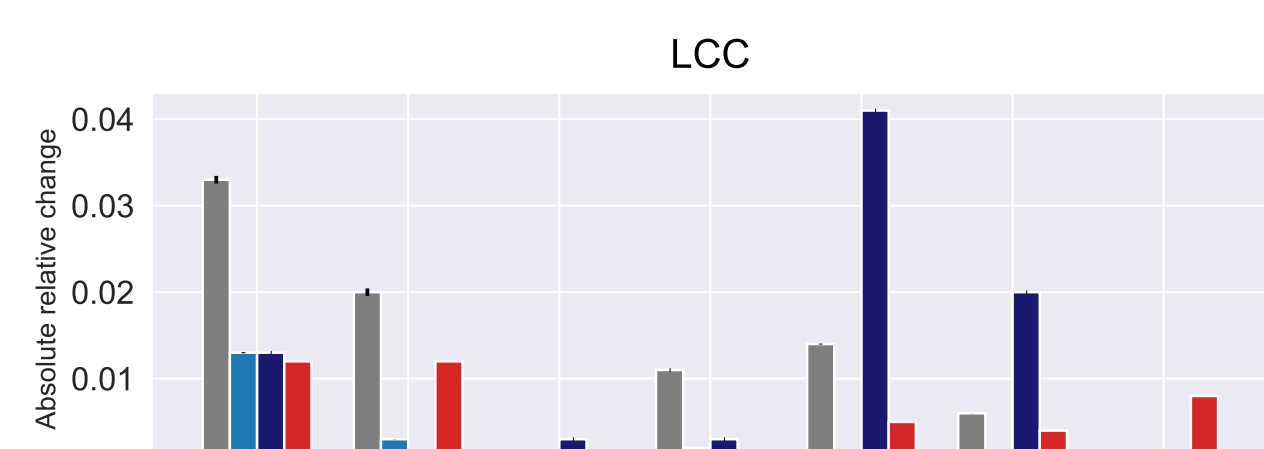
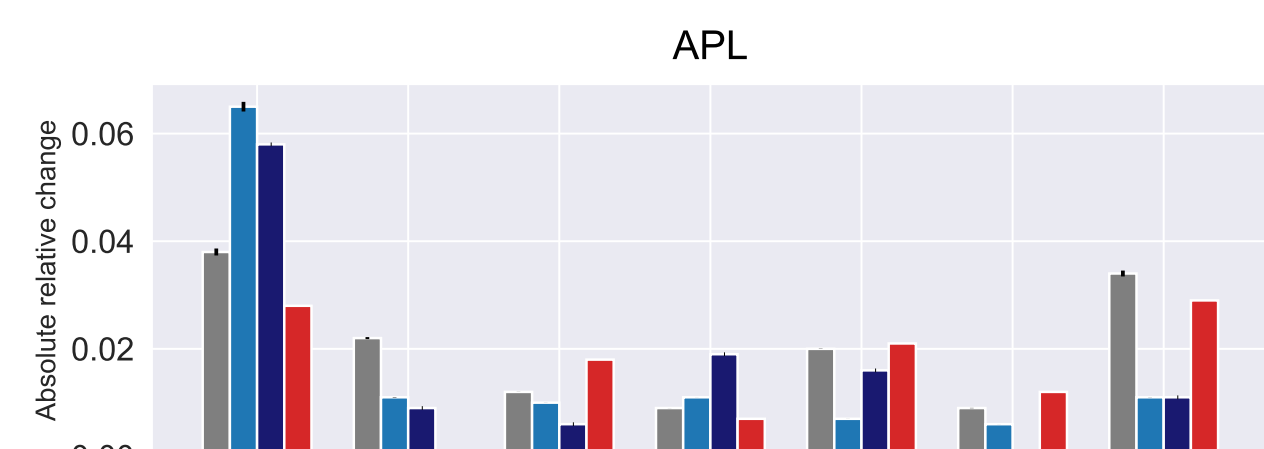
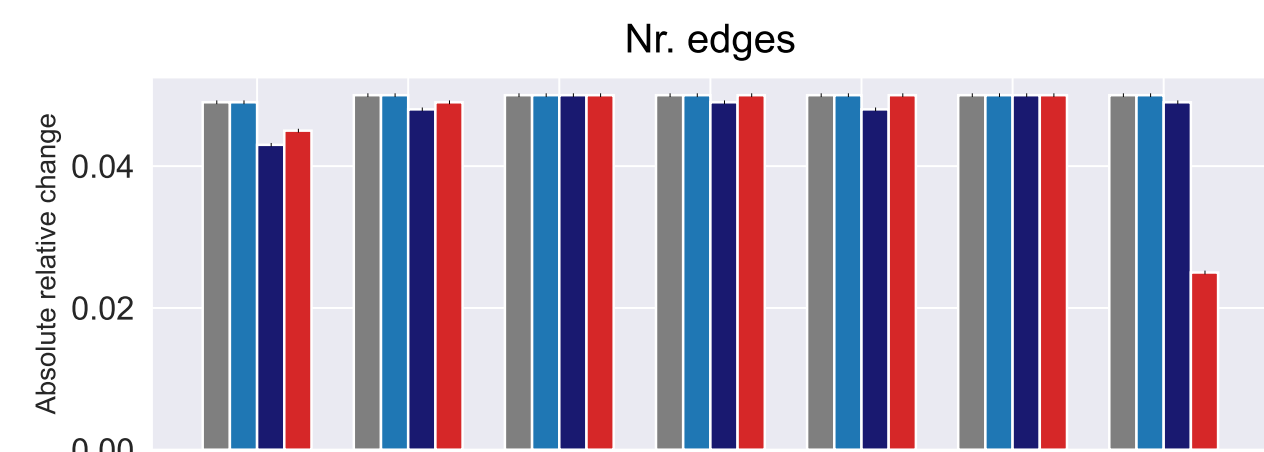
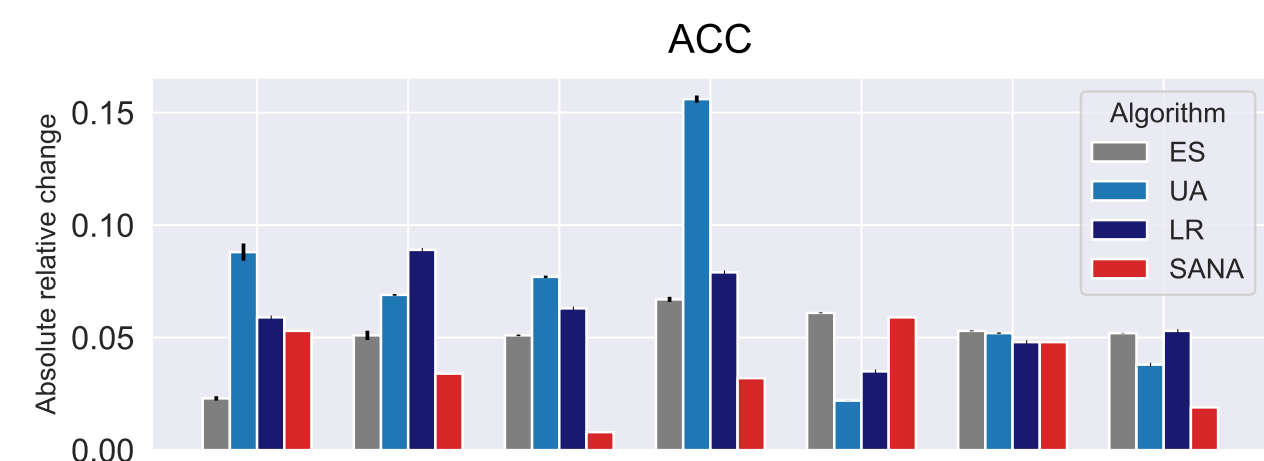
- **Datasets:** 7 social networks, 5% edge-deletion budget.
- **Anonymization:** up to 18× more nodes anonymized than heuristics.
- **Utility:** Metrics change by <5%.
- **Runtime:** Similar to Logistic-Regression baseline.



Anonymization comparison.



Runtime comparison.



Utility comparison.

Takeaway

- SANA = fast, high-quality anonymization.
- Best for dense & medium networks.

Next challenges

- Multi-objective anonymity + utility.
- Dynamic (real-time) anonymization.