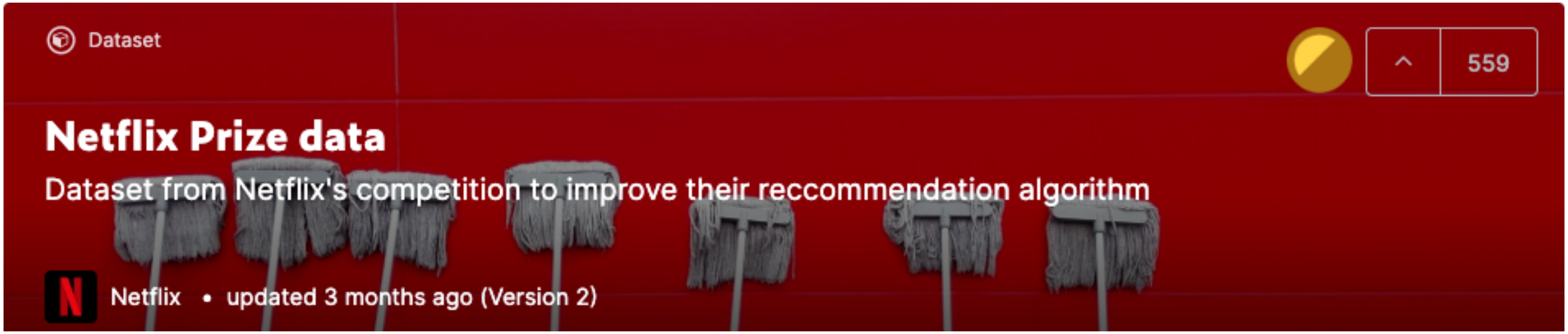
CISPA
HELMHOLTZ CENTER FOR INFORMATION SECURITY

Towards Plausible Graph Anonymization

Yang Zhang, Mathias Humbert, Bartlomiej Surma,
Praveen Manoharan, Jilles Vreeken, Michael Backes

Graph sharing

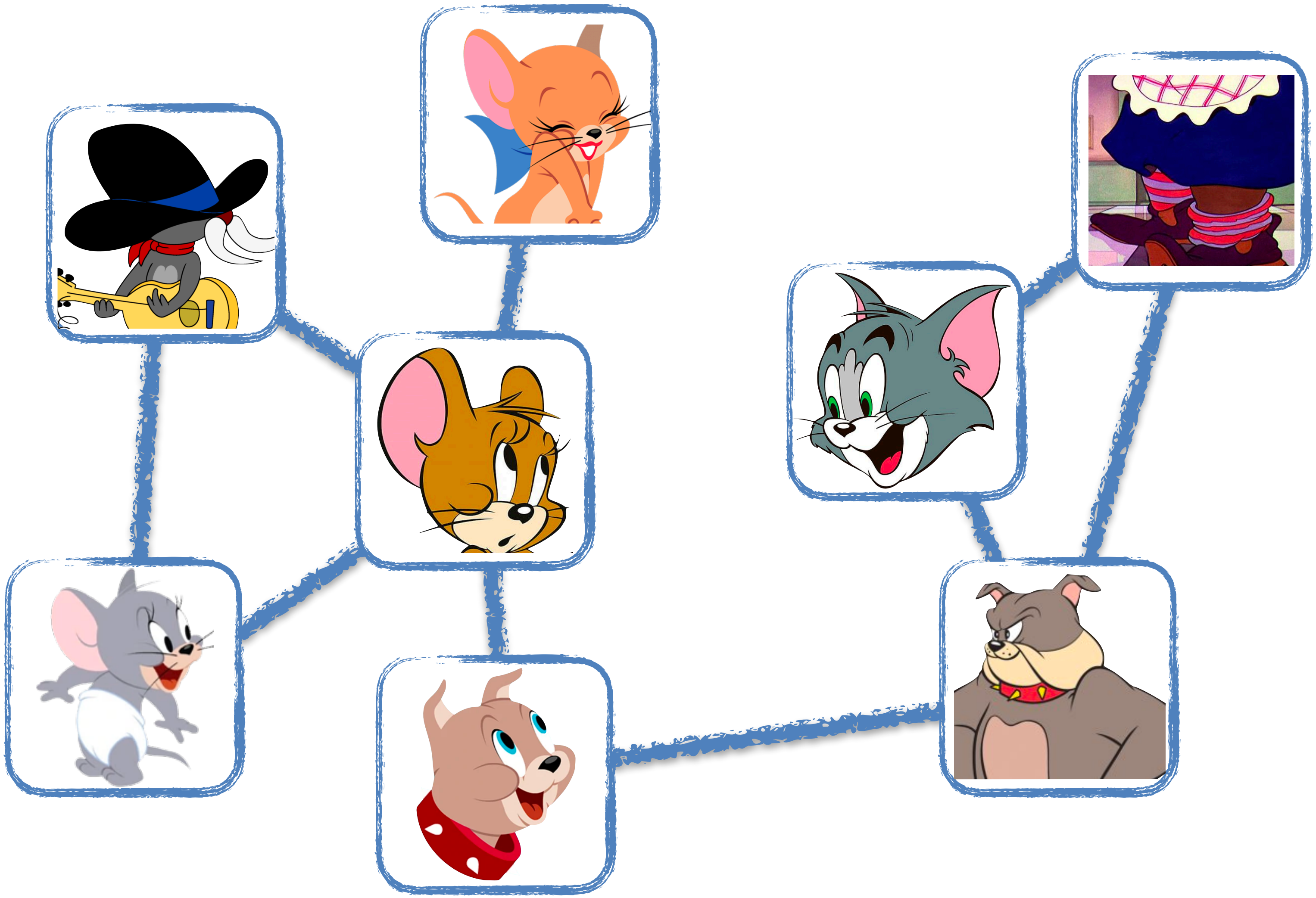


IJCNN Social Network Challenge

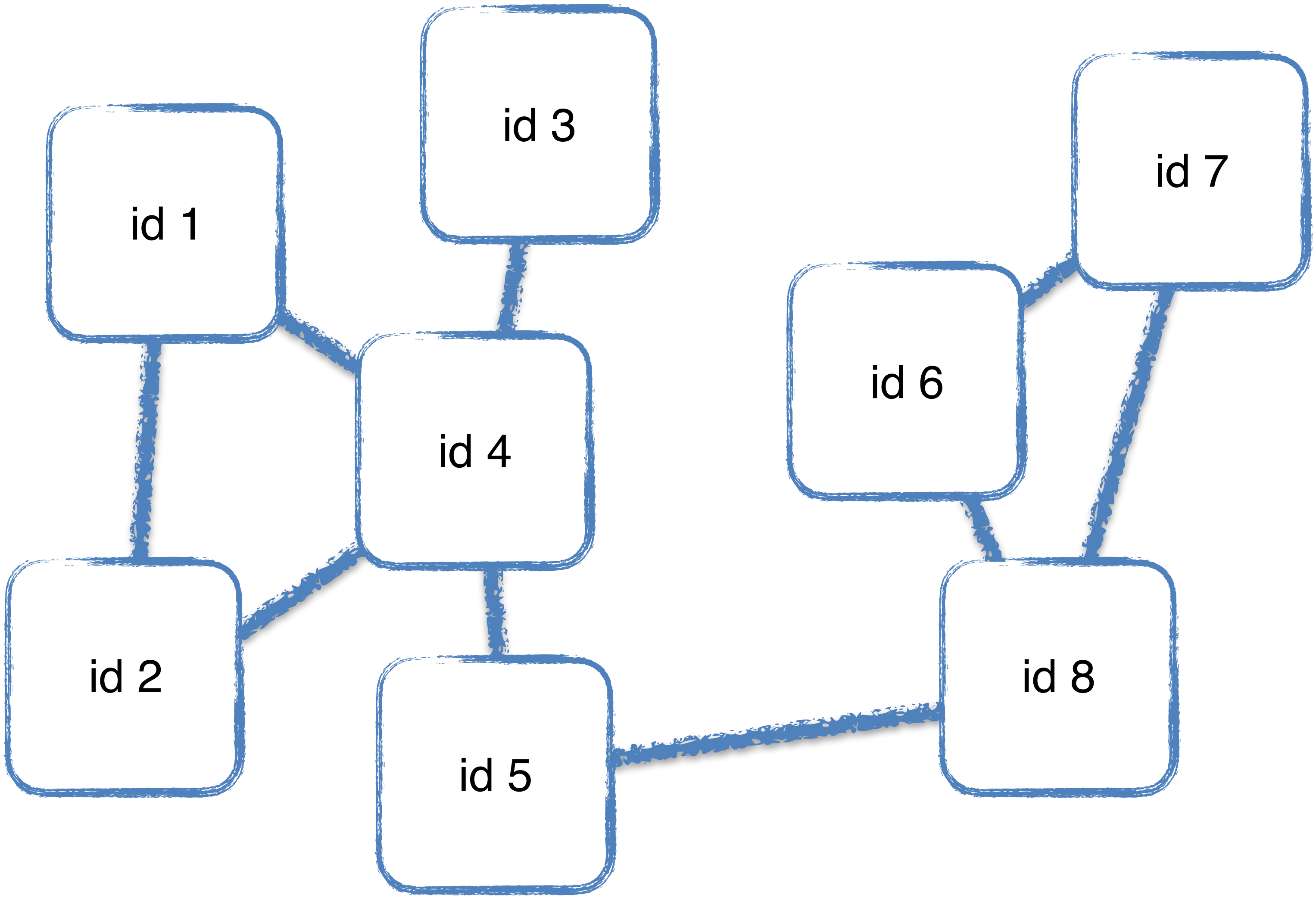
This competition requires participants to predict edges in an online social network. The winner will receive free registration and the opportunity to present their solution at IJCNN 2011.

$950 · 117 teams · 9 years ago

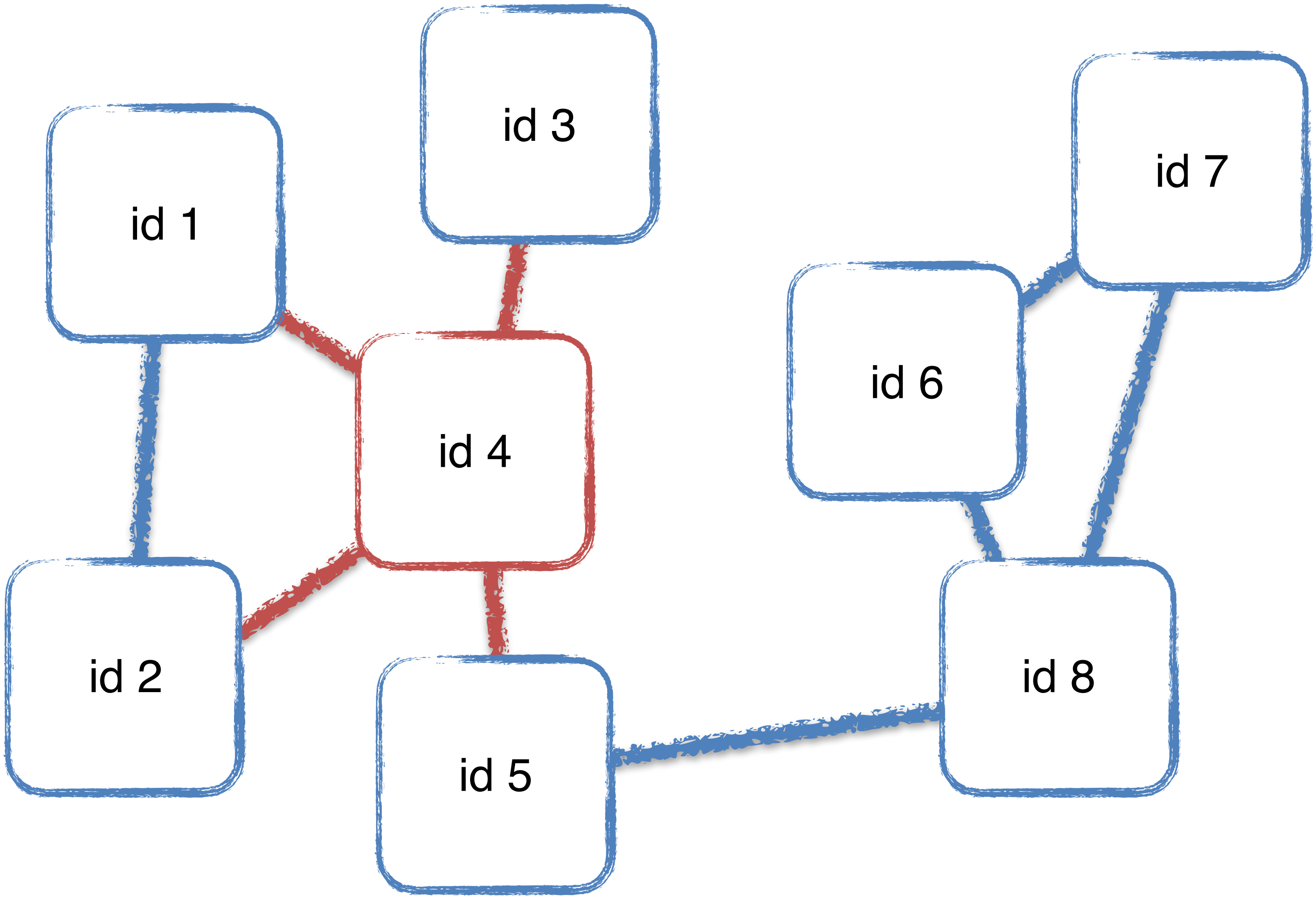
Graph anonymization



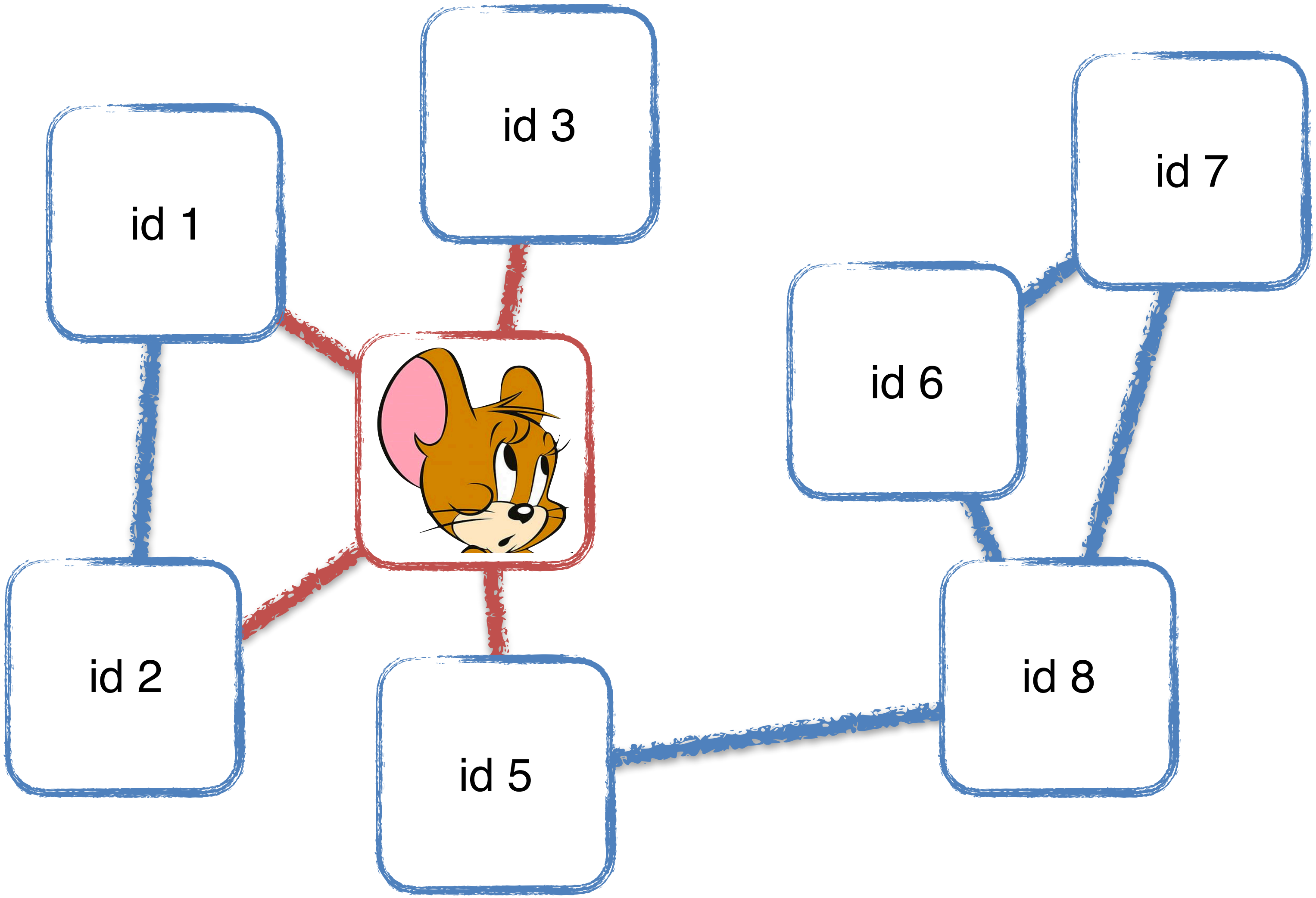
Graph anonymization



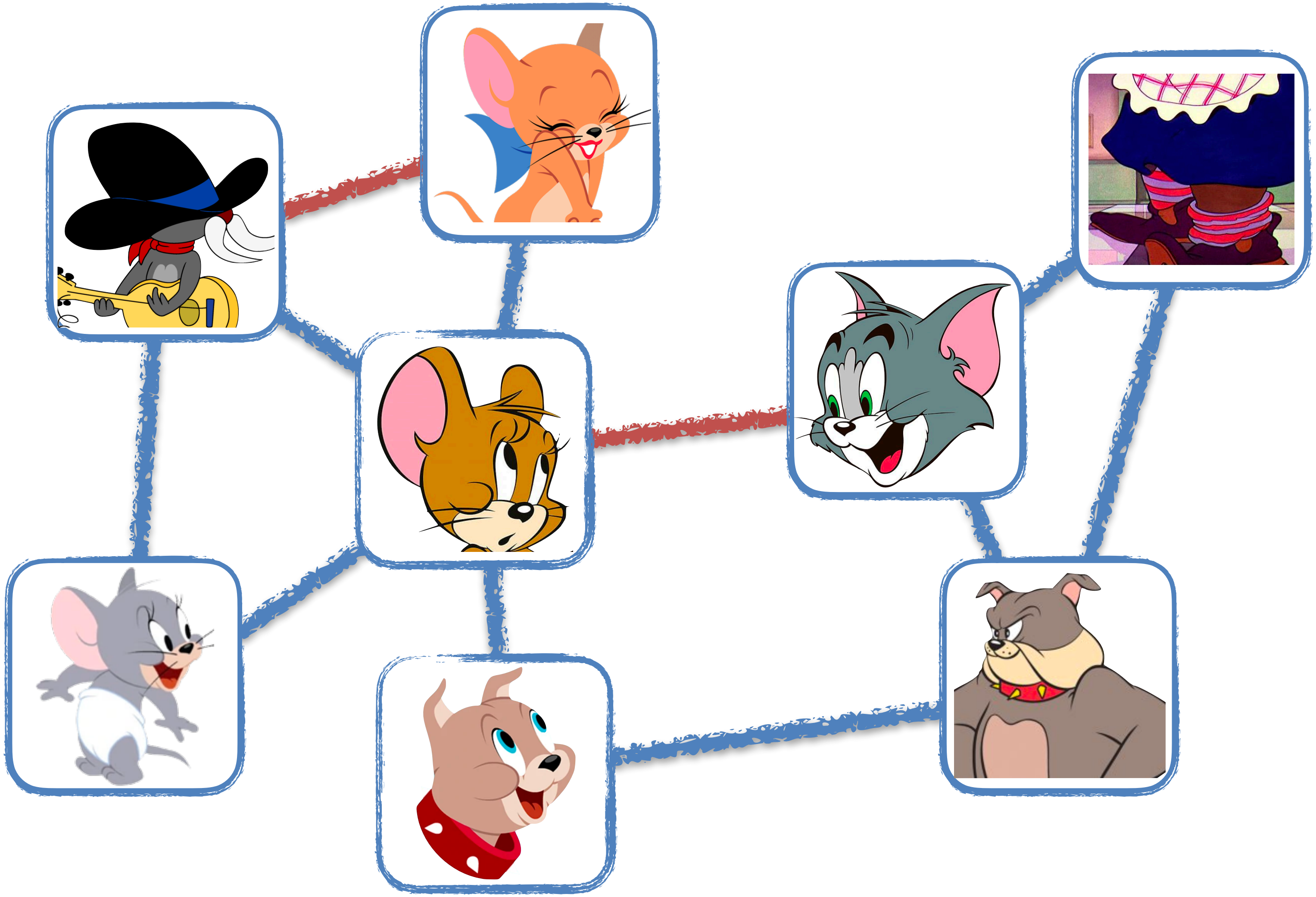
Graph anonymization



Graph anonymization



Graph anonymization



Our work

- Find a fundamental flaw in graph anonymization designs

Our work

- Find a fundamental flaw in graph anonymization designs
- Exploit it to recover original graph

Our work

- Find a fundamental flaw in graph anonymization designs
- Exploit it to recover original graph
- Use our findings to enhance anonymization designs

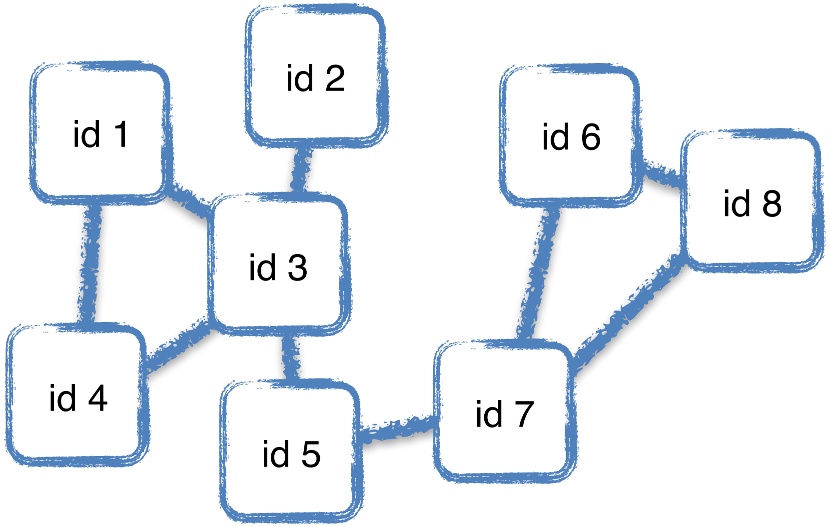
Our work

- Find a fundamental flaw in graph anonymization designs
- Exploit it to recover original graph
- Use our findings to enhance anonymization designs
- Evaluate privacy and usability of enhanced techniques on 3 real life datasets:
  - Enron, NO, Snap

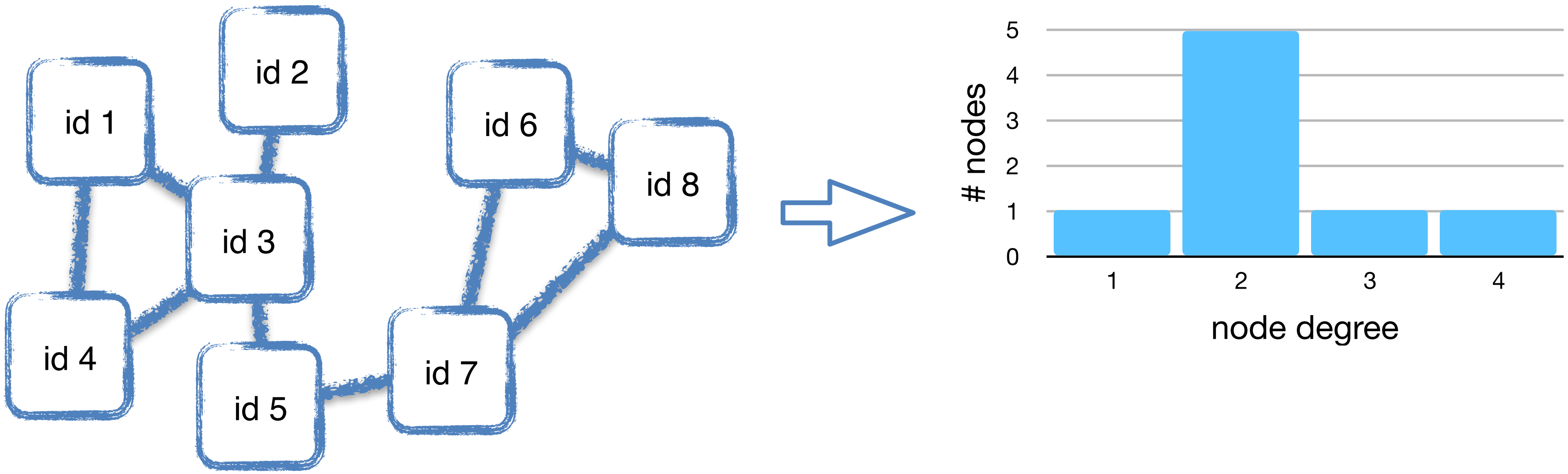
Graph anonymization methods

- **’08 Liu et al. - k-anonymity (k-DA)**
- ’08 Zhou et al. - k-anonymity (k-NA)
- ’10 Cheng et al. - k-anonymity (k-iso)
- **’11 Sala et al. - differential privacy**
- ’12 Mittal et al. - random walk privacy
- ’14 Xiao et al. - differential privacy

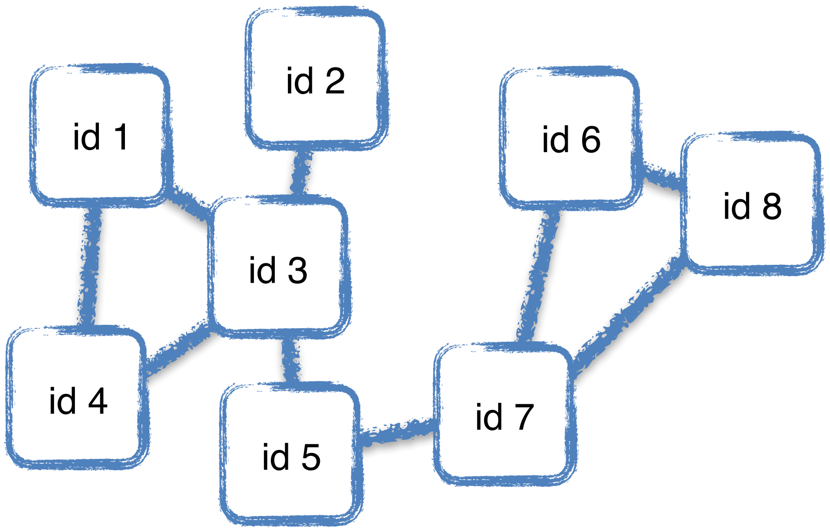
k-DA algorithm

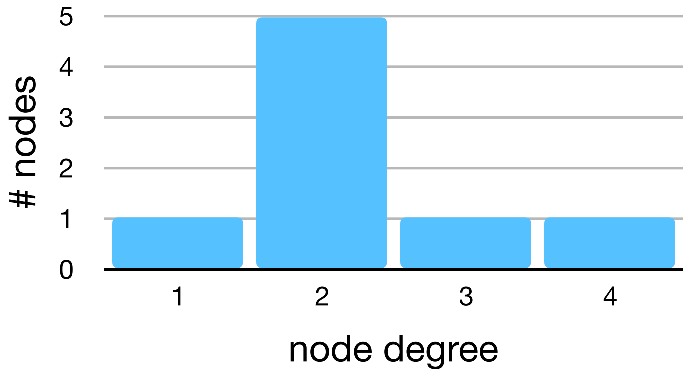


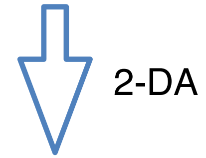
k-DA algorithm

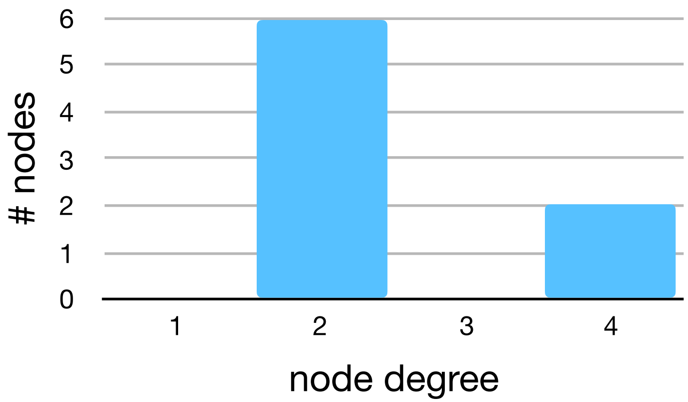


k-DA algorithm

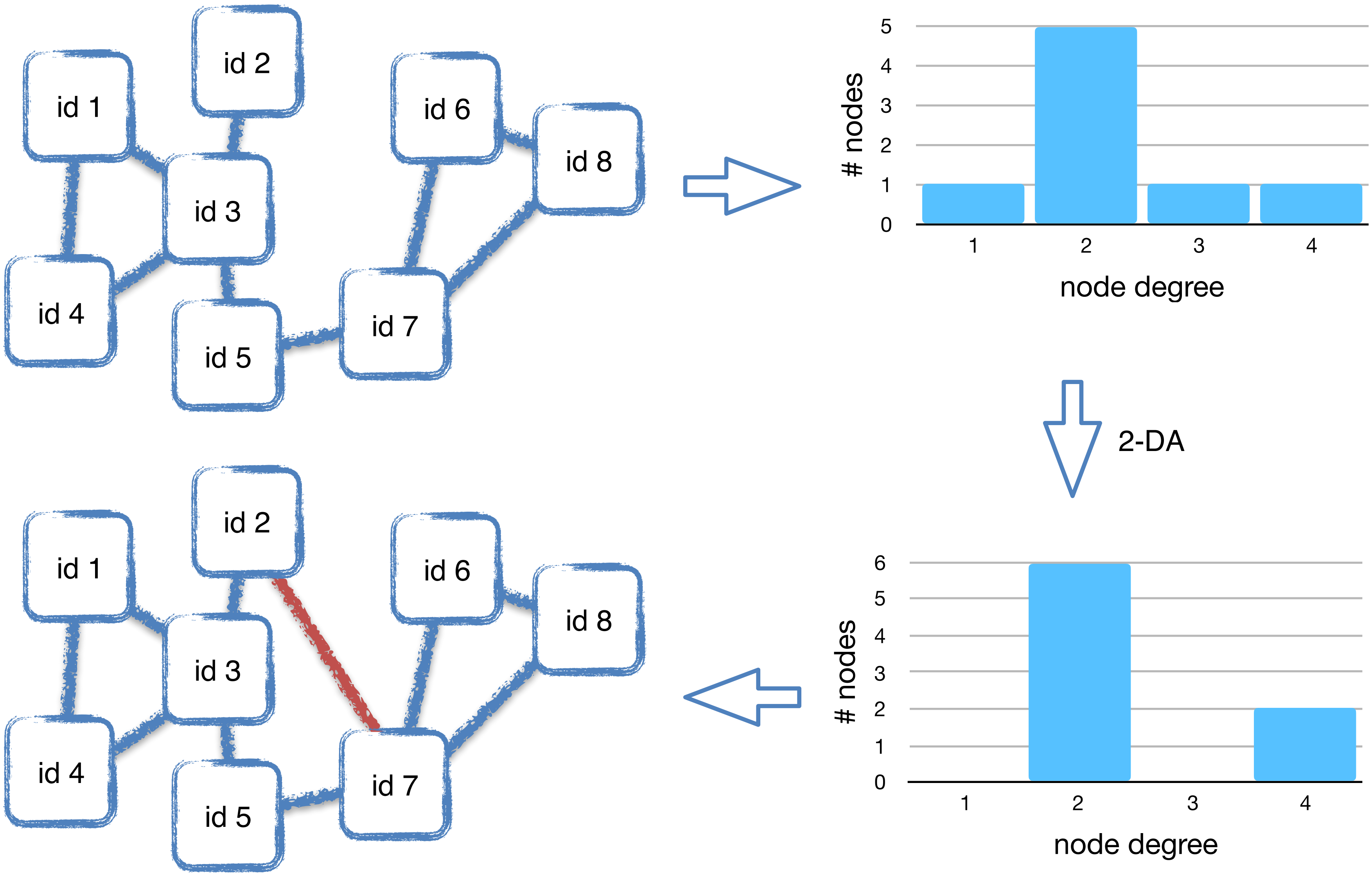




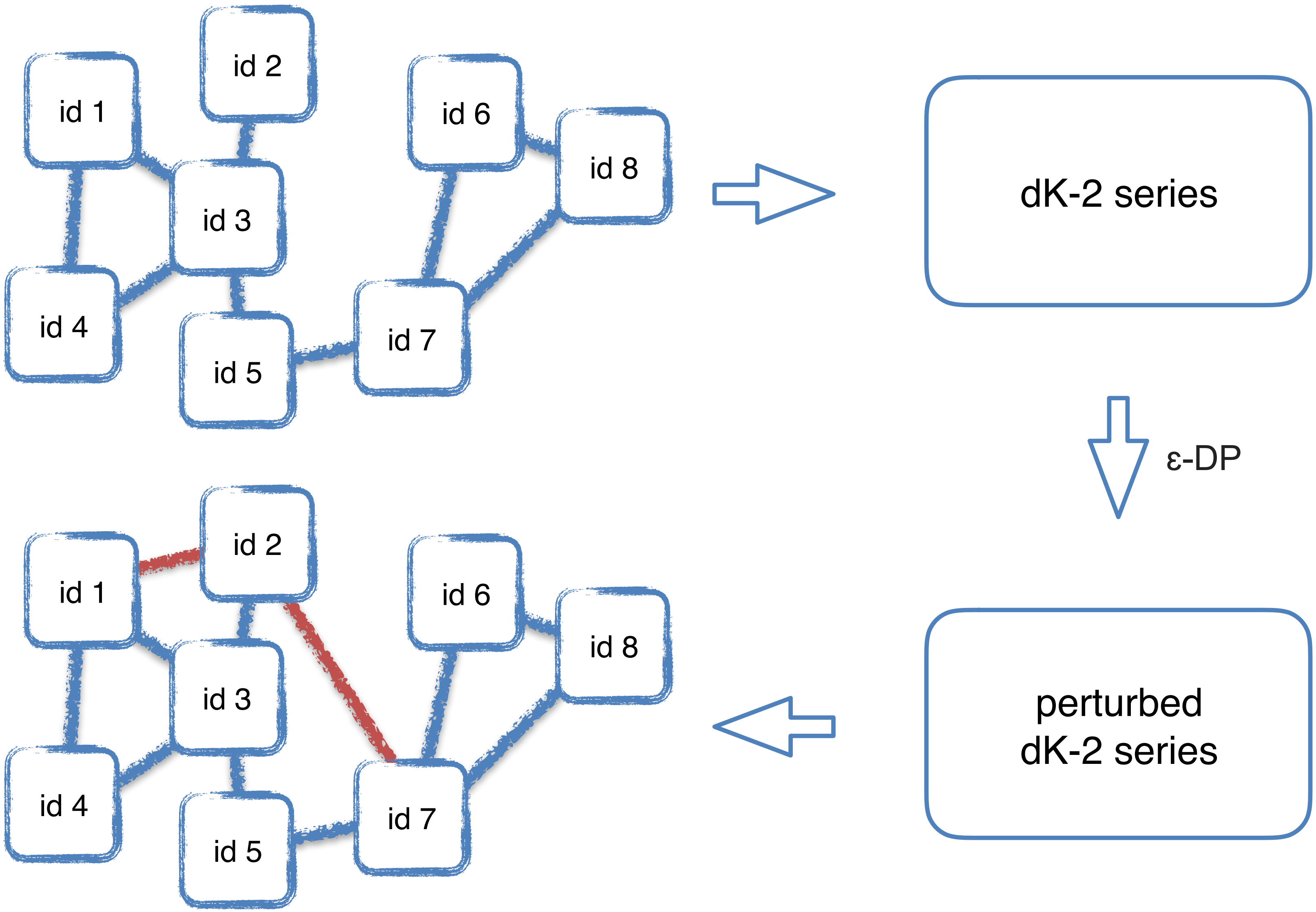




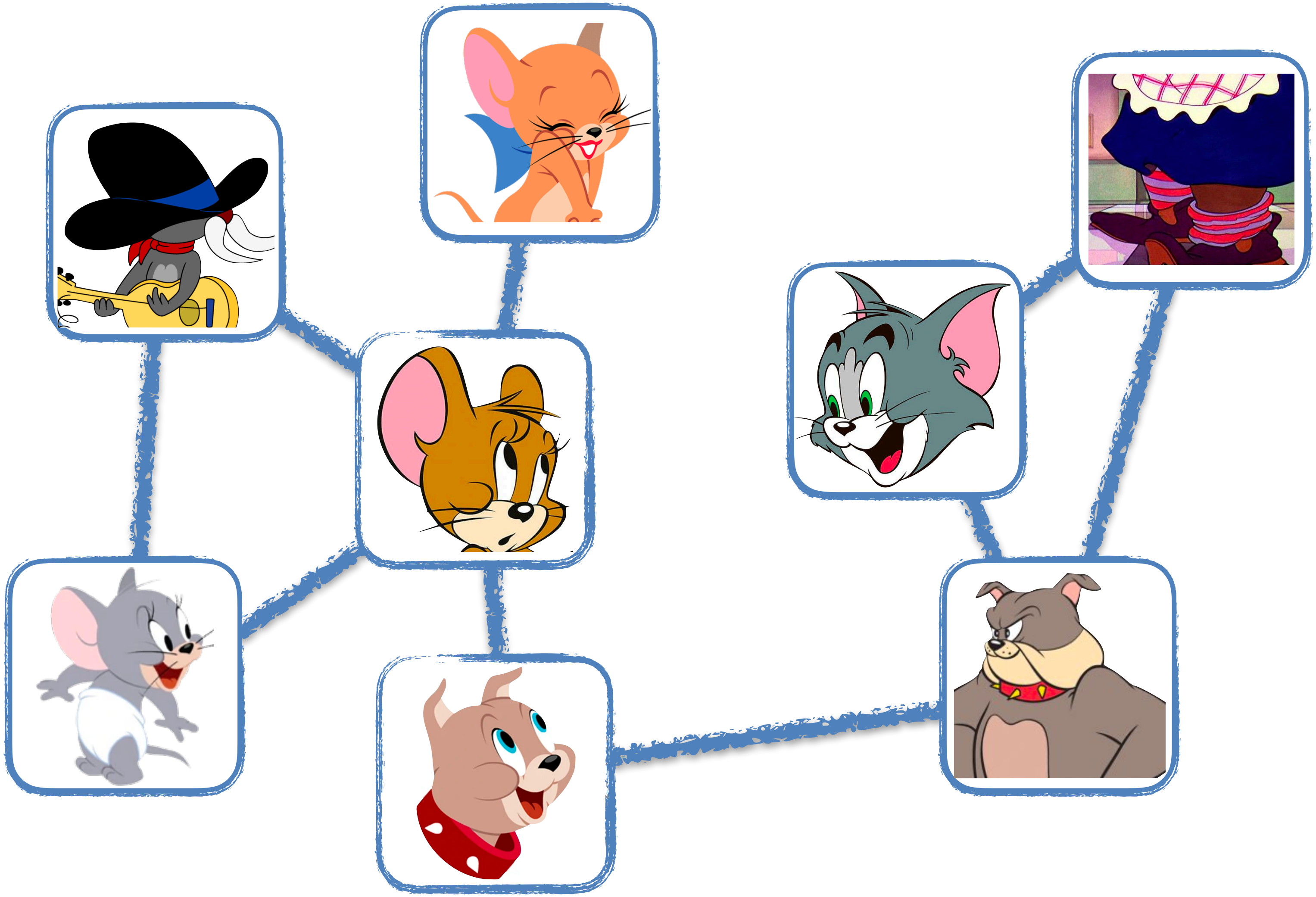
k-DA algorithm



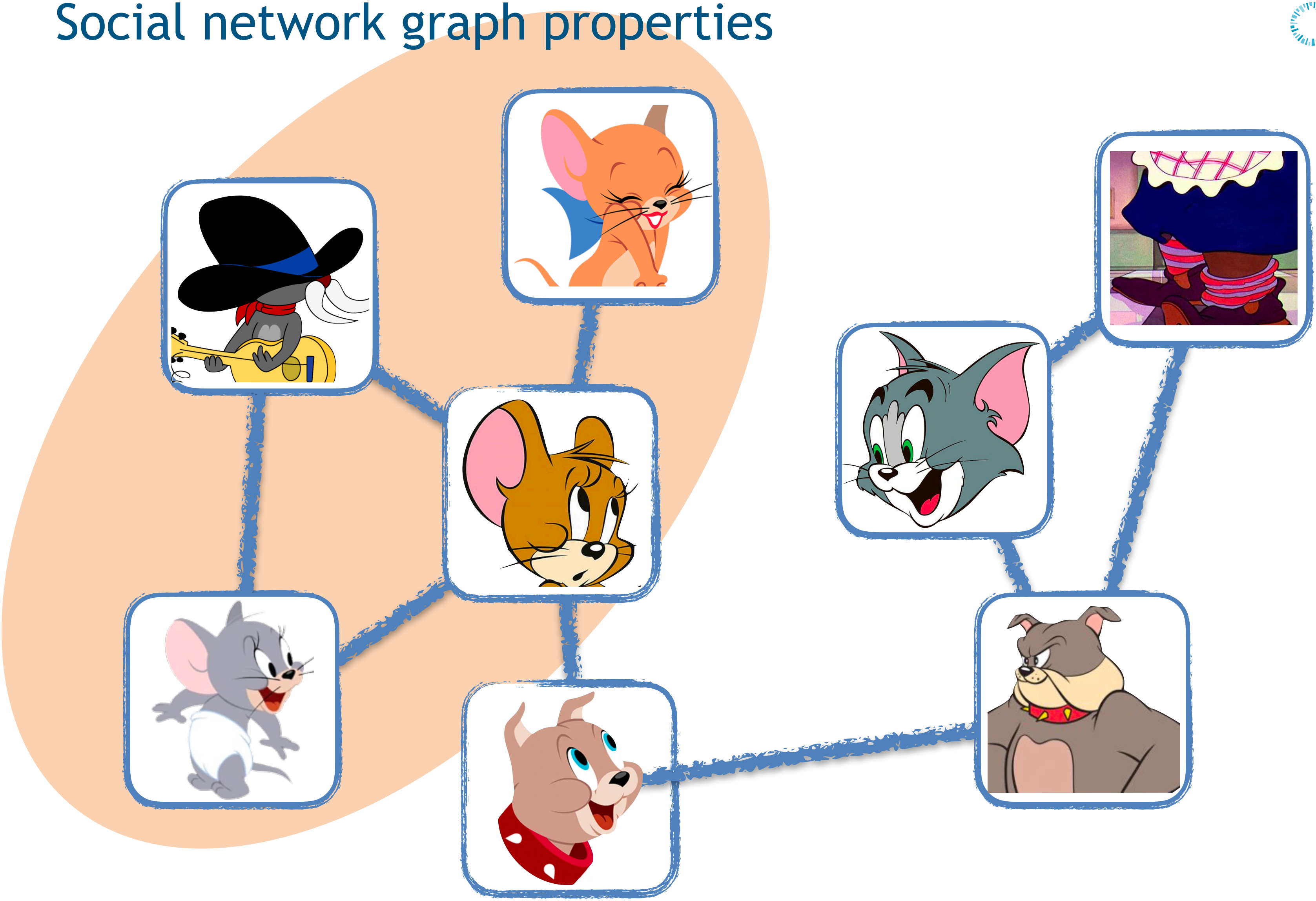
SalaDP algorithm



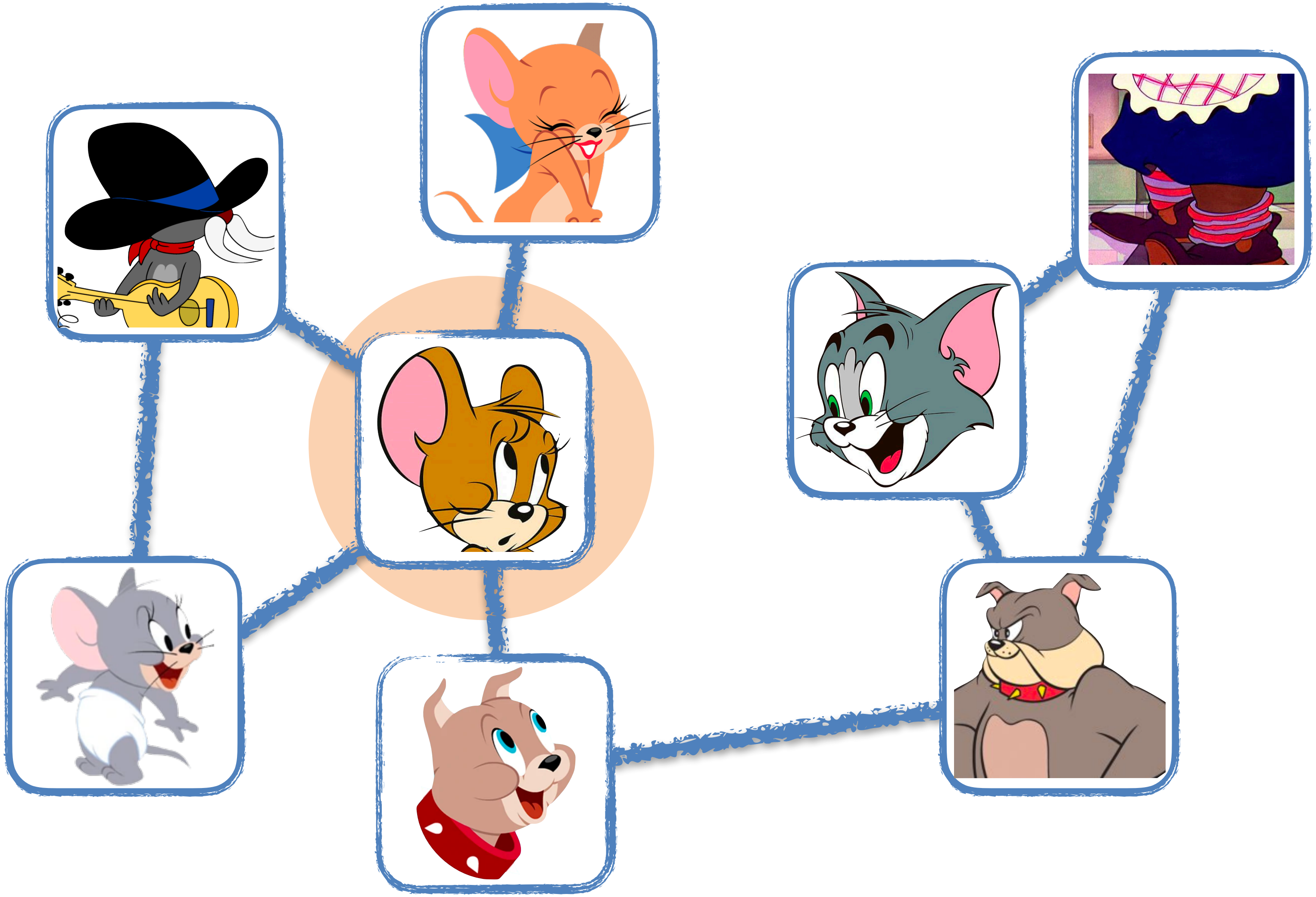
Social network graph properties



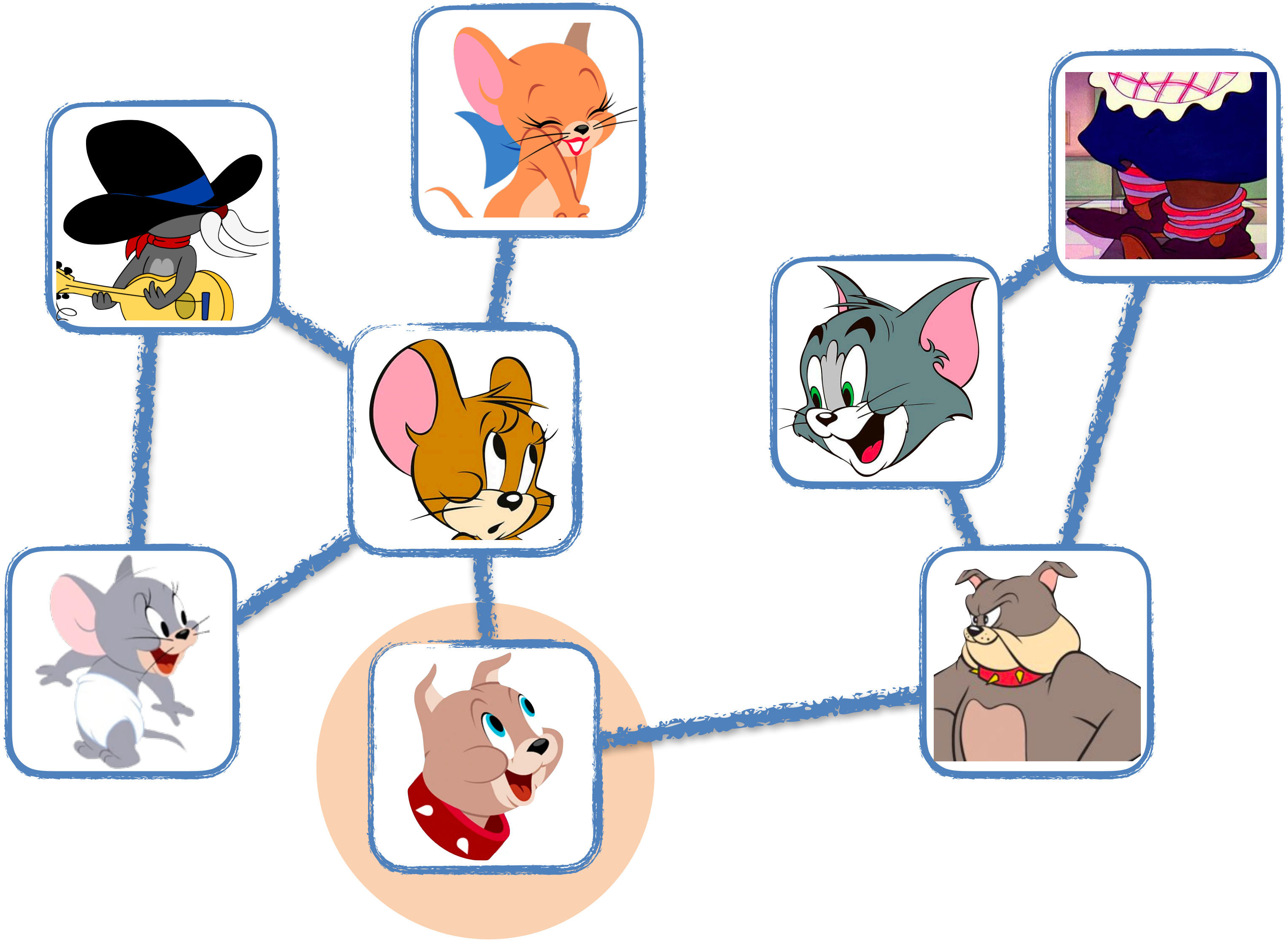
Social network graph properties



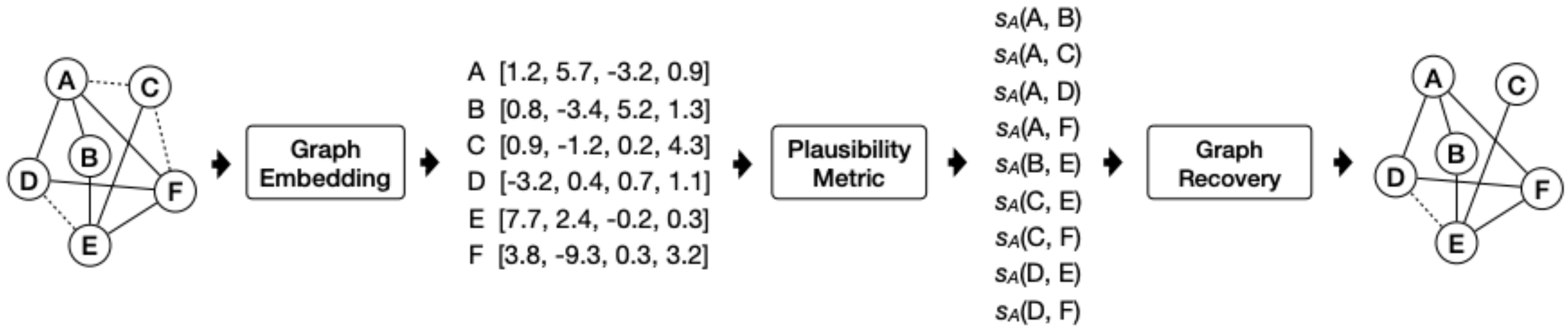
Social network graph properties



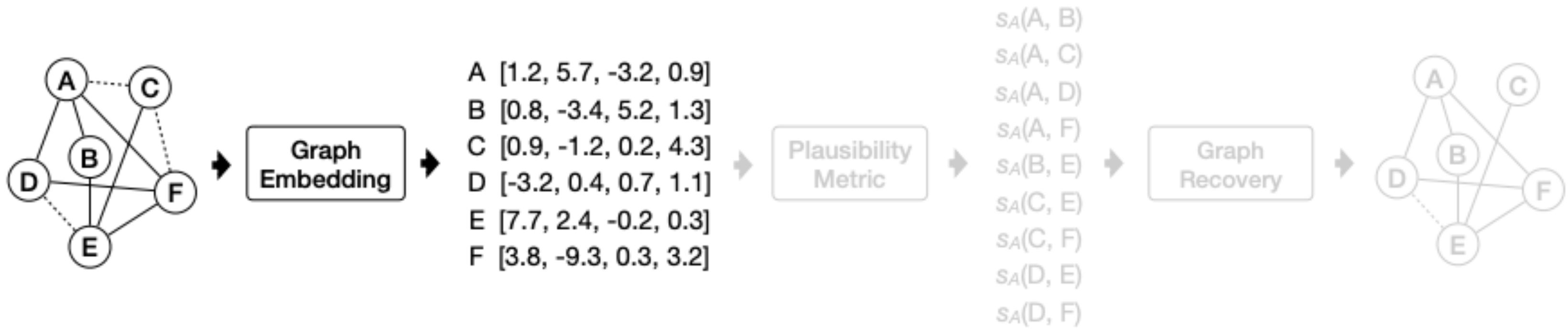
Social network graph properties



Graph recovery attack - overview

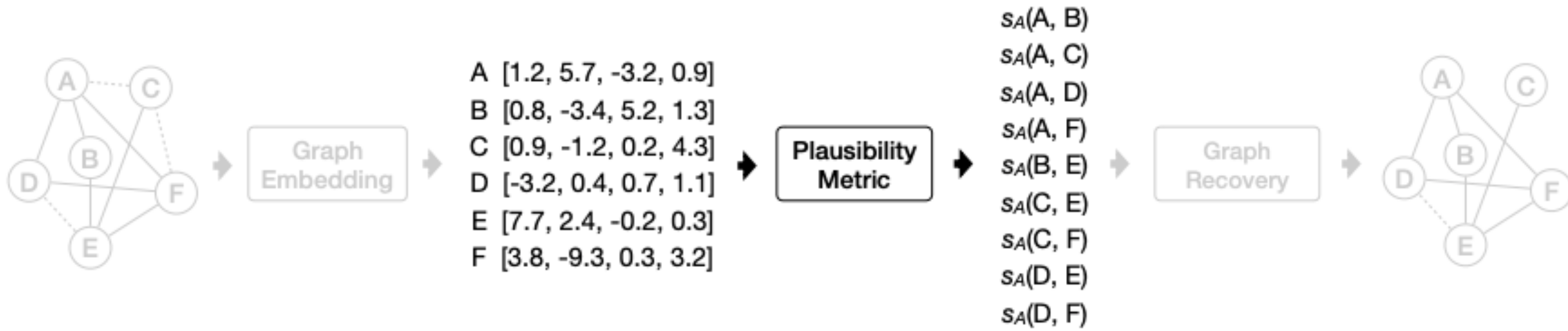


Graph recovery attack - graph embedding

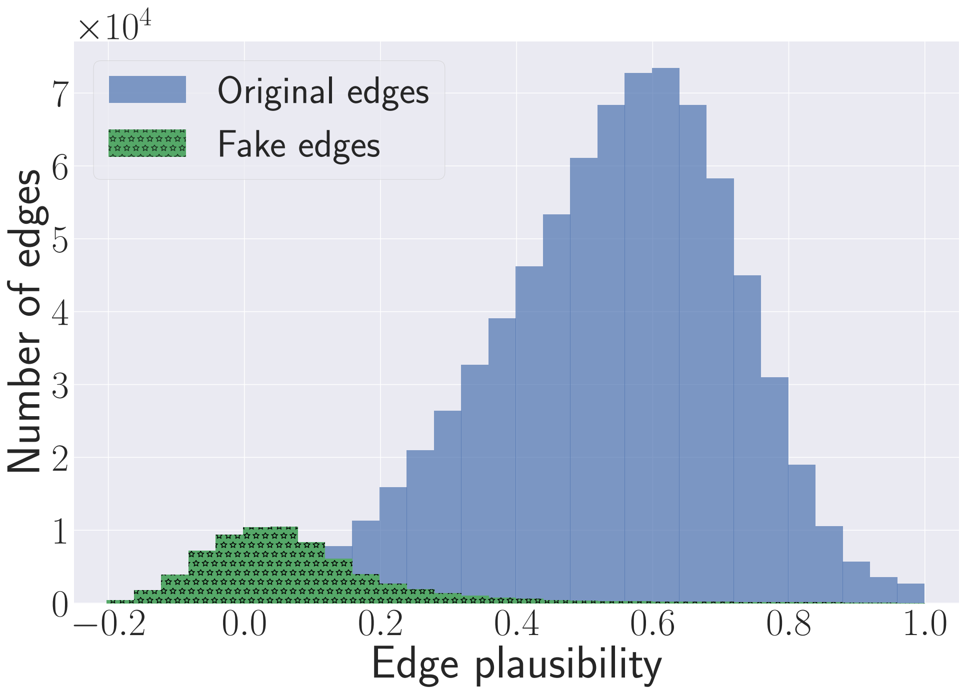


- Node embeddings with node2vec ’16 Grover and Leskovec
- Mapping users into continuous vector space
- User’s vector reflects structural properties

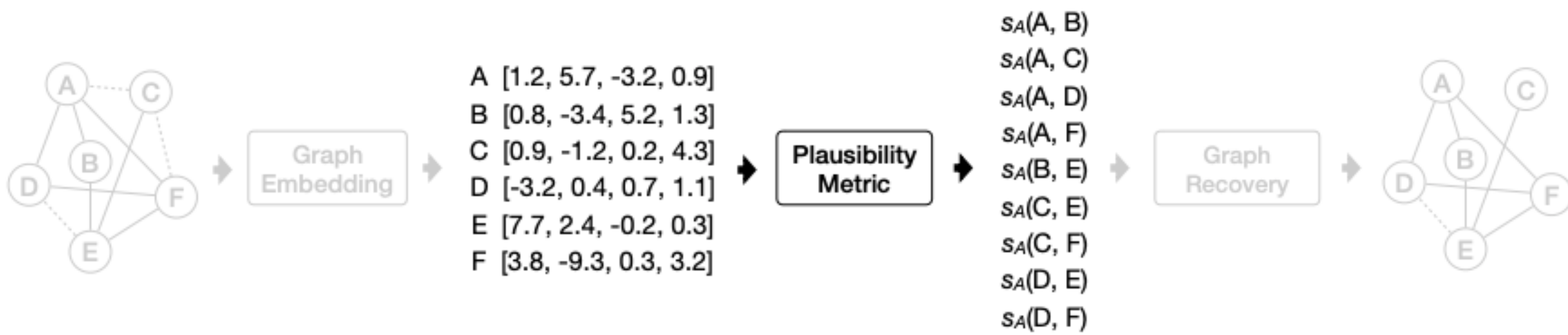
Graph recovery attack - graph embedding



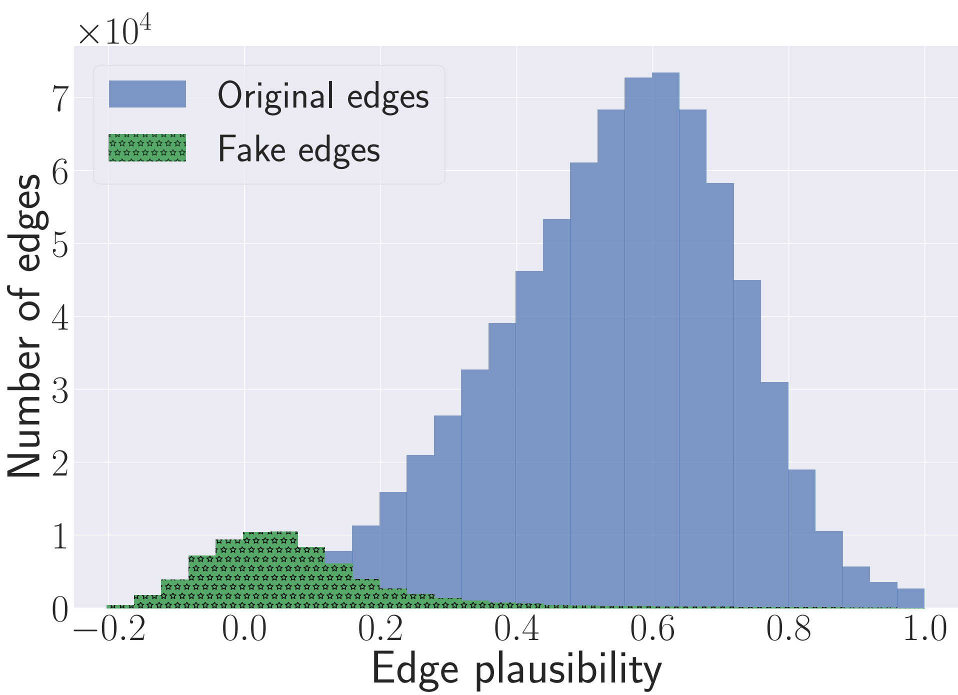
Plausibility is cosine similarity between embeddings

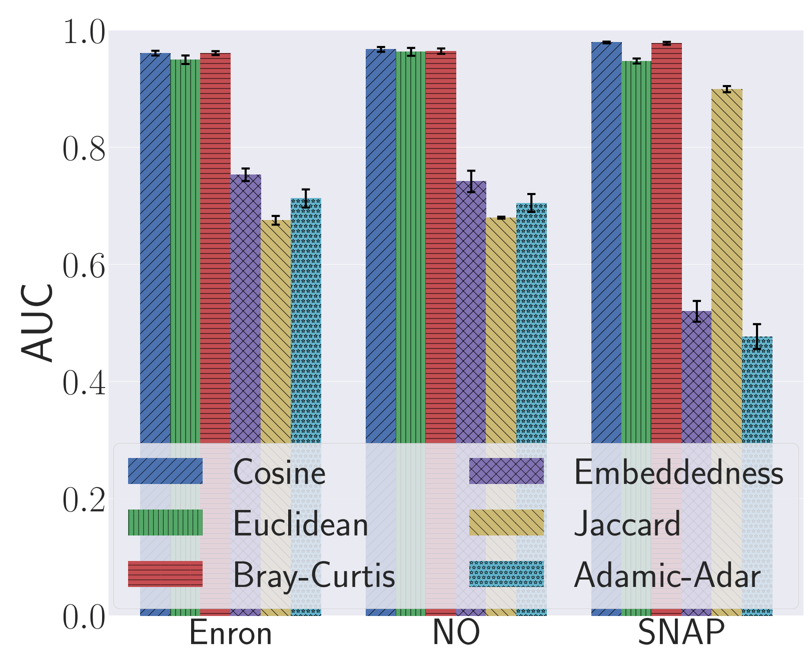


Graph recovery attack - graph embedding

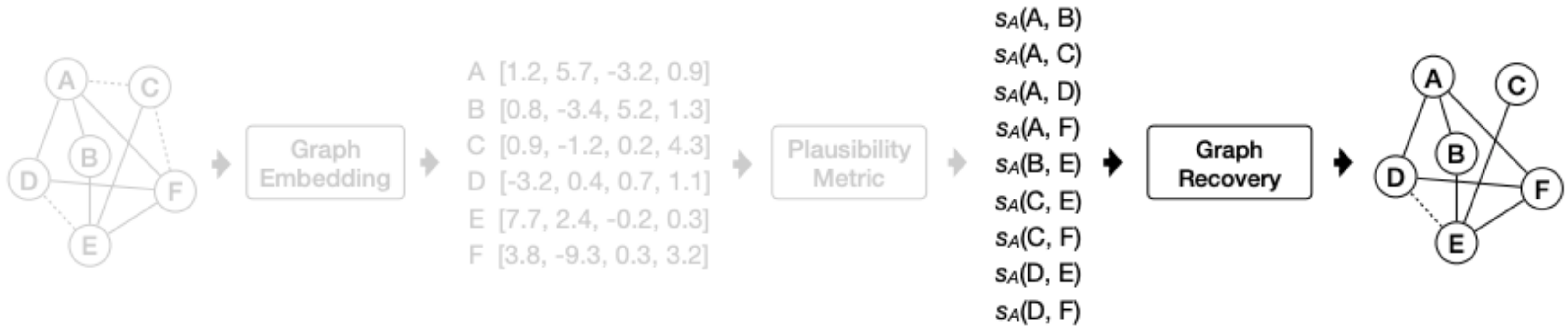


Plausibility is cosine similarity between embeddings

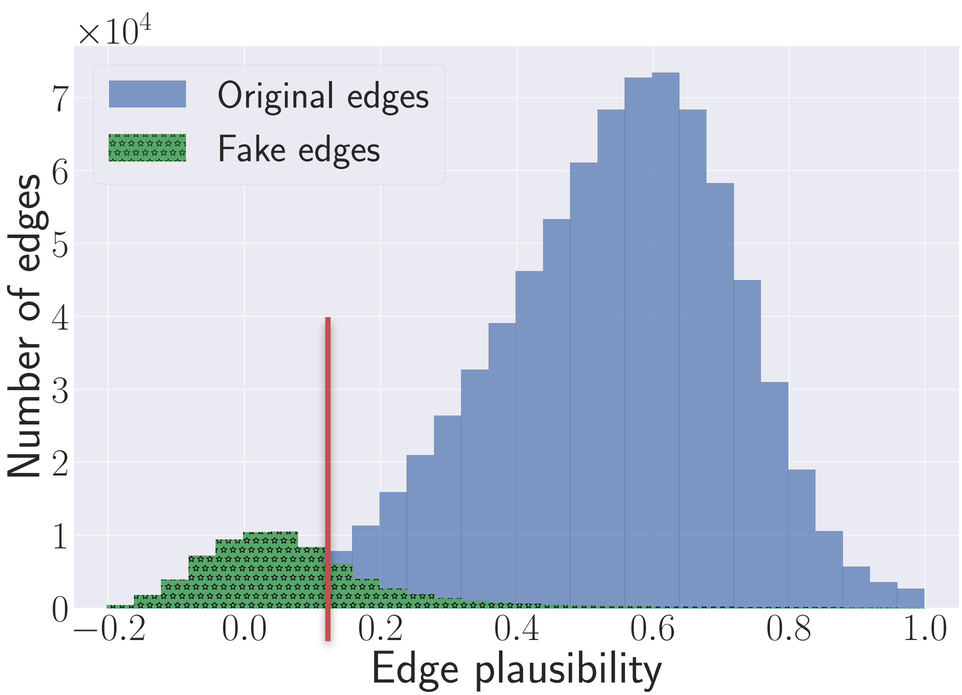




Graph recovery attack - graph embedding



Find a cutoff point and remove non-plausible edges



	Enron	NO	SNAP
k-DA ($k = 50$)	0.792	0.642	0.857
k-DA ($k = 75$)	0.796	0.710	0.869
k-DA ($k = 100$)	0.812	0.761	0.881
SalaDP ($\epsilon = 100$)	0.672	0.712	0.853
SalaDP ($\epsilon = 50$)	0.750	0.723	0.835
SalaDP ($\epsilon = 10$)	0.819	0.876	0.802

F1 score

Enhancing anonymization

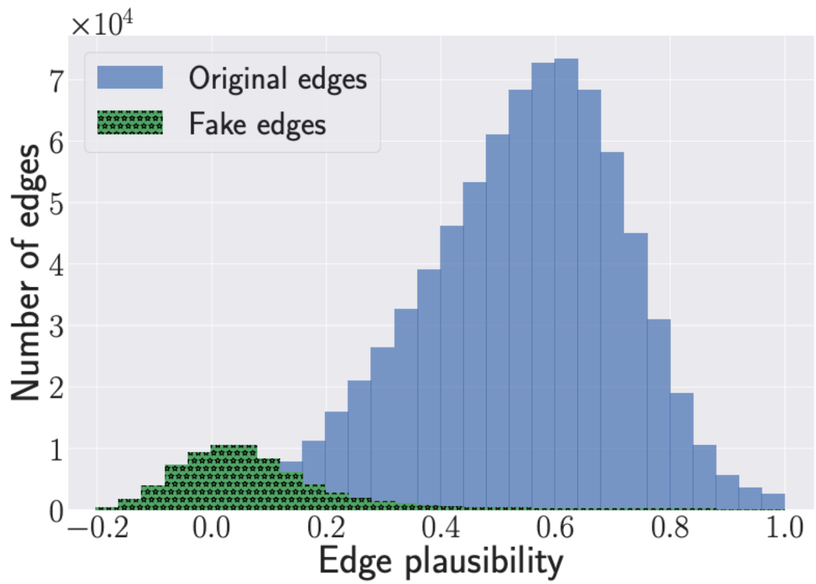
- get fake edges with highest plausibility?
  - the distribution will look unnatural

Enhancing anonymization

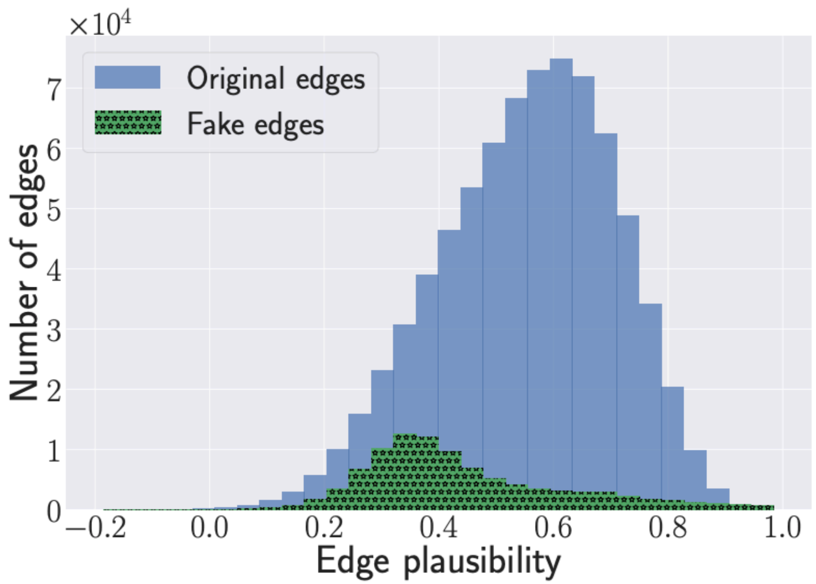
- get fake edges with highest plausibility?
  - the distribution will look unnatural
- draw fake edges from same plausibility distribution?

Enhancing anonymization

- get fake edges with highest plausibility?
  - the distribution will look unnatural
- draw fake edges from same plausibility distribution?



k-DA (k=100)



Enhanced k-DA (k=100)

Resilience to graph recovery attack

- F1 score for original anonymizations

	Enron	NO	SNAP
k-DA ($k=50$)	0.792	0.642	0.857
k-DA ($k=75$)	0.796	0.710	0.869
k-DA ($k=100$)	0.812	0.761	0.881
SalaDP ($\epsilon=100$)	0.672	0.712	0.853
SalaDP ($\epsilon=50$)	0.750	0.723	0.835
SalaDP ($\epsilon=10$)	0.819	0.876	0.802

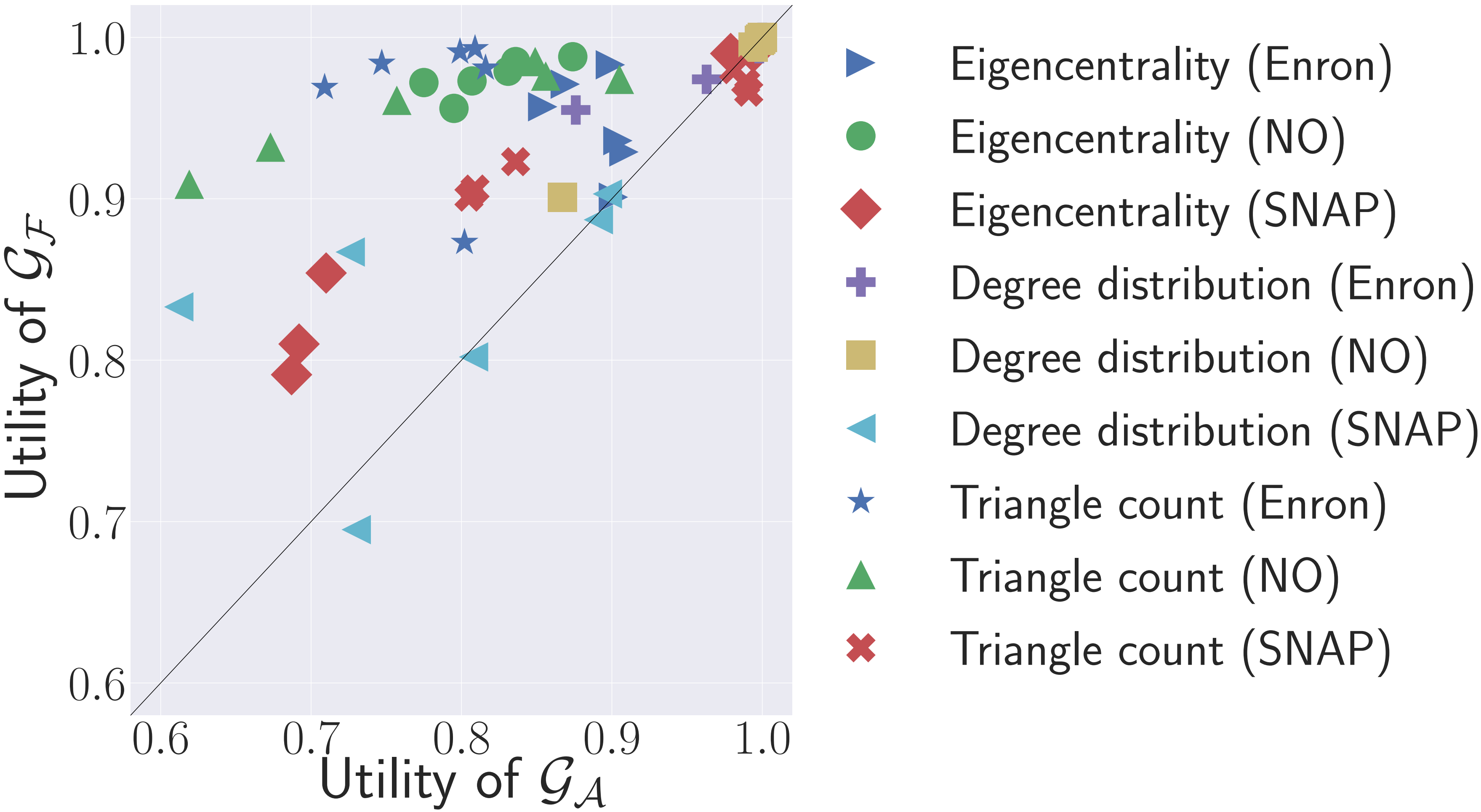
k-DA drops by:
26~51%

SalaDP drops by:
37~48%

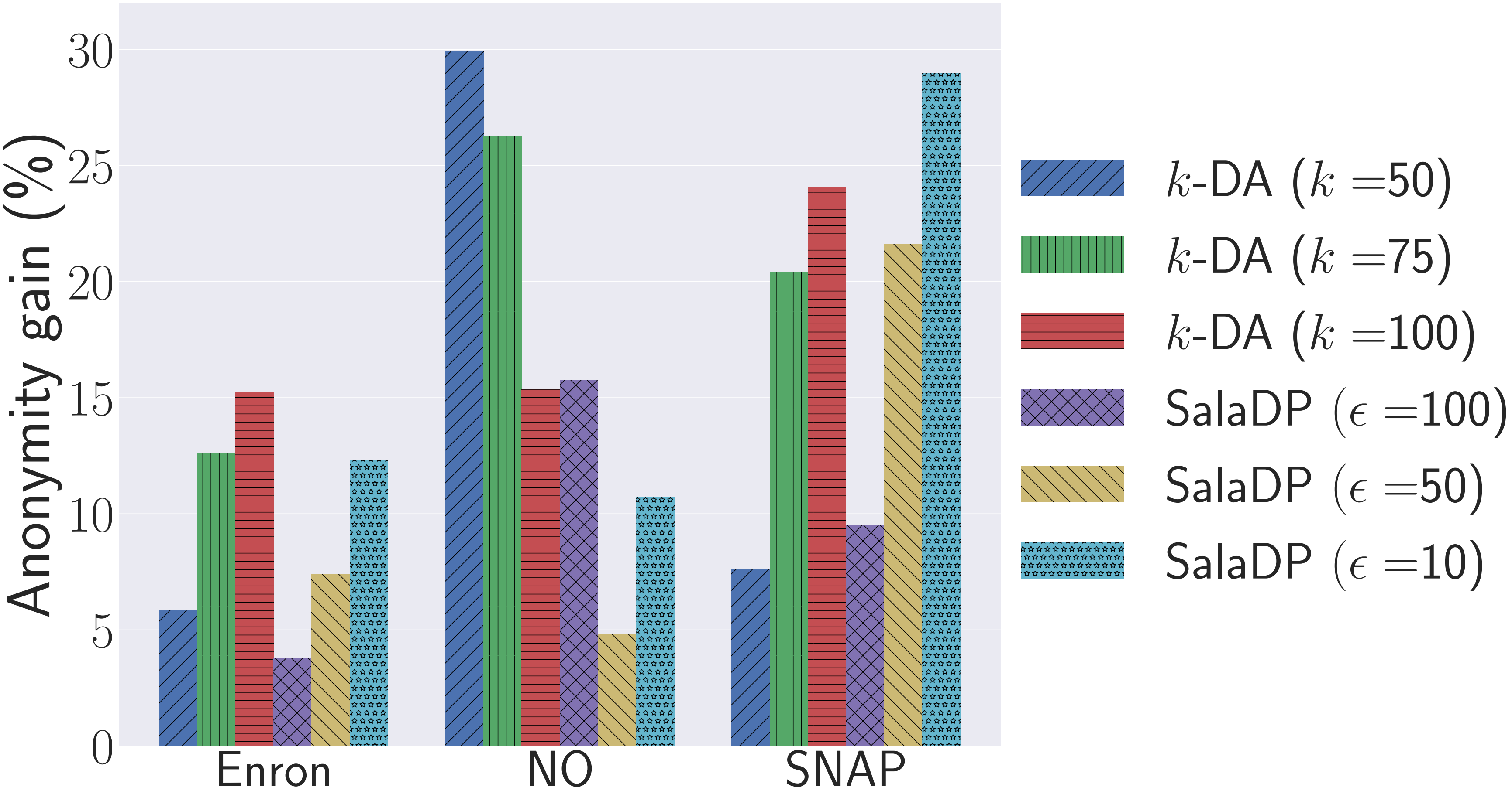
- F1 score for enhanced anonymizations

	Enron	NO	SNAP
k-DA ($k=50$)	0.531	0.391	0.632
k-DA ($k=75$)	0.428	0.433	0.609
k-DA ($k=100$)	0.510	0.501	0.597
SalaDP ($\epsilon=100$)	0.422	0.370	0.515
SalaDP ($\epsilon=50$)	0.390	0.411	0.522
SalaDP ($\epsilon=10$)	0.439	0.527	0.490

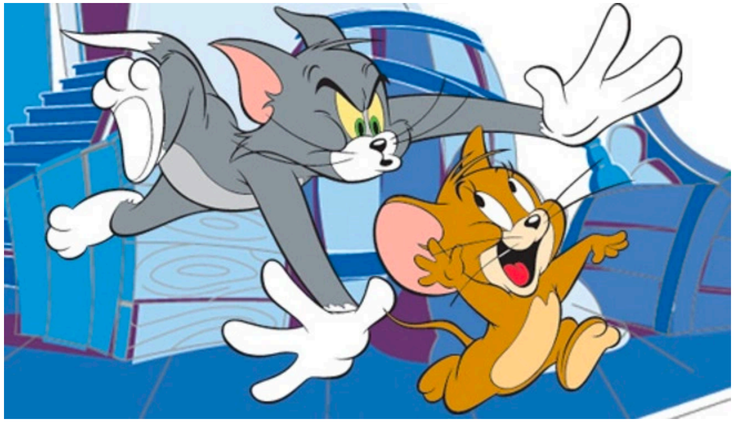
Utility of Enhanced anonymization



Resilience to deanonymization attack

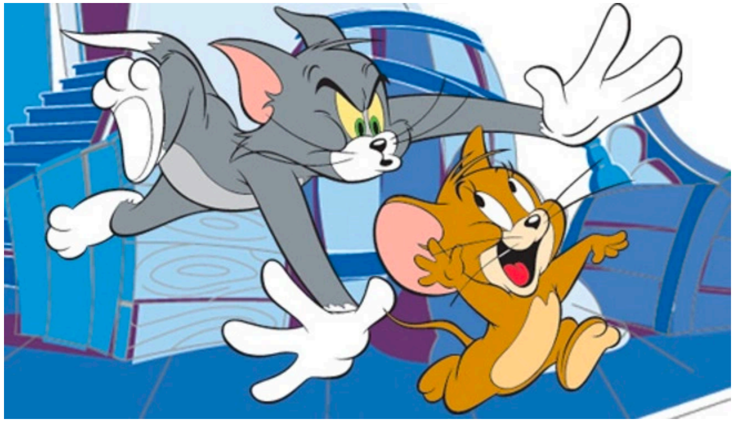


Conclusion



We find flaws in current graph anonymizations

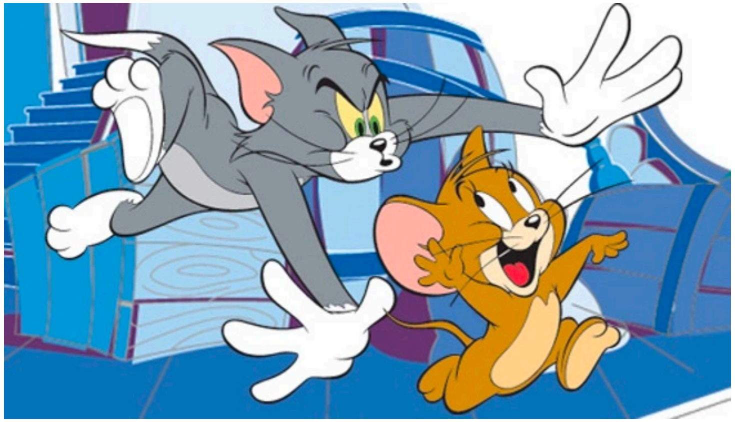
Conclusion



We find flaws in current graph anonymizations

We recover the original, pre-anonymized graph

Conclusion

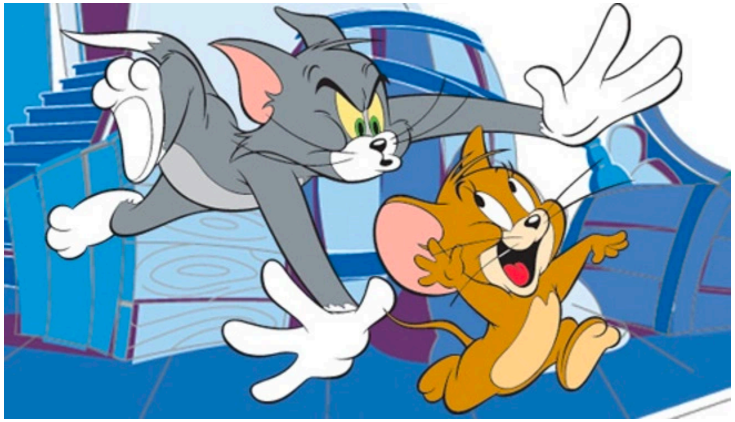


We find flaws in current graph anonymizations

We enhance the anonymization techniques

We recover the original, pre-anonymized graph

Conclusion



We find flaws in current graph anonymizations

We enhance the anonymization techniques

We recover the original, pre-anonymized graph

We evaluate privacy and utility of enhanced anonymization