

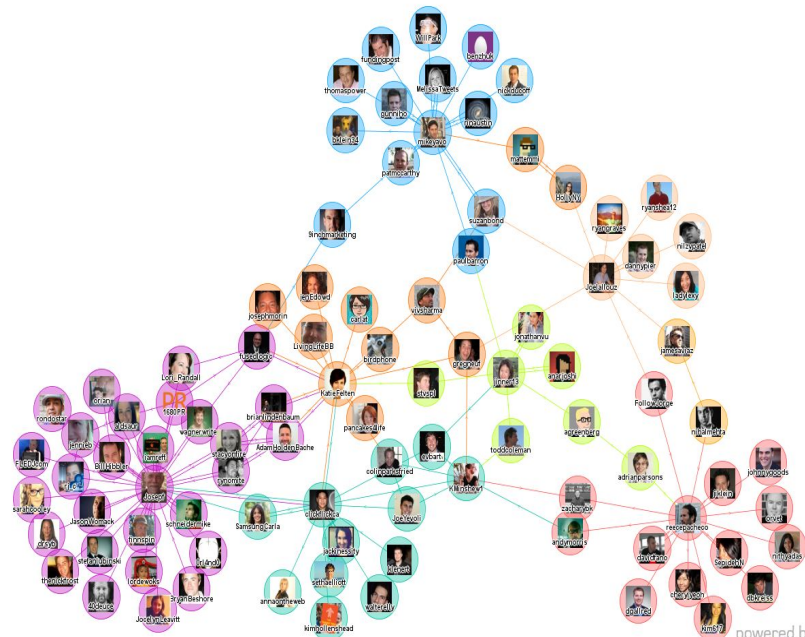
Data protection & privacy

ANONYMIZING SOCIAL NETWORKS

Authors:
Davide Senatori
Chengao Xia

Introduzione

Un social network rappresenta entità e connessioni fra di loro. Essi vengono rappresentati come grafi non diretti.



Introduzione



Per preservare l'identità del nodo nel grafo, identificatori sintetici sono introdotti per sostituire i nomi.
Questa procedura viene detta: Naive anonymization.

Naive anonymization

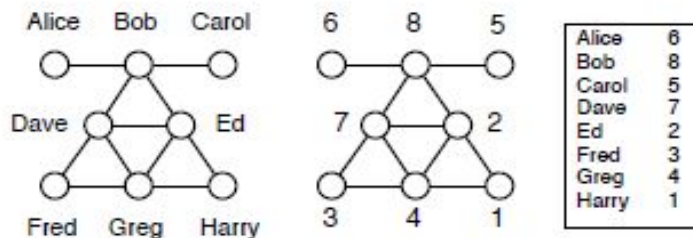


Figure 1: A social network, G ; the naive anonymization of G ; the anonymization mapping.

- Rende sicuro un grafo?



Un attaccante può scoprire molte informazioni attraverso domande VertexRefinement

Obiettivo del progetto

L'obiettivo del progetto è riuscire a rendere impossibile la re-identificazione.

Naive Anonymization



Random Perturbation

Mostreremo che tale tecnica può ridurre le possibilità di re-identificazione con accettabili distorsioni del grafo.

Random Perturbation

La procedura di Random perturbation viene effettuata in seguito ad una naive anonymization. Il nuovo grafo $G_p = (V_p, E_p)$ è costruito da G_{na} attraverso una sequenza di eliminazioni di m archi e un inserimento di m archi scelti in modo random (con $m \leq |E|$). I nodi invece non cambiano.

A seguito di ciò, l'avversario che tenterà di re-identificare i nodi dovrà considerare il set di possibili grafi (worlds) a cui è stata effettuata una perturbation attraverso m .

Esempio Random Perturbation

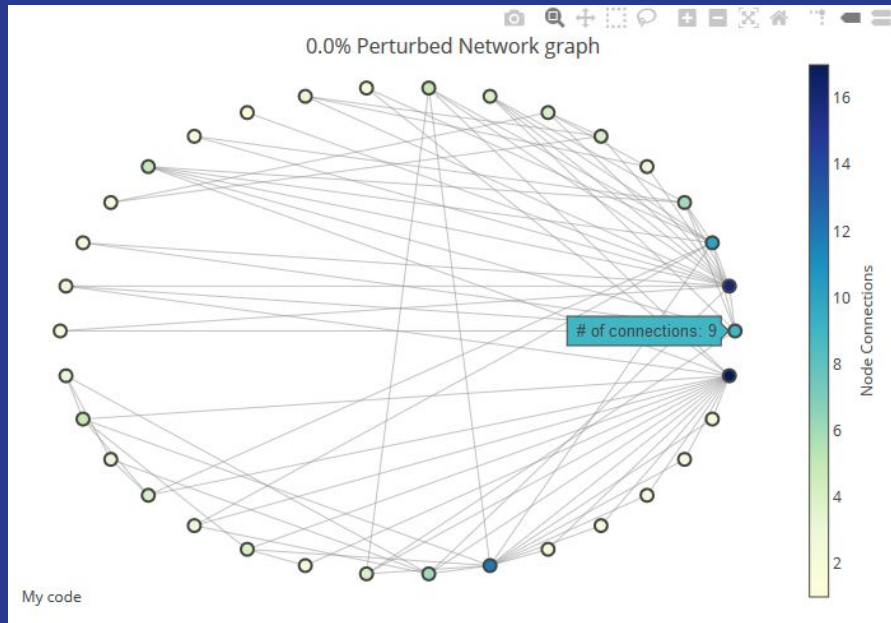
Come dimostrazione di questo studio abbiamo analizzato un dataset rappresentato come un grafo. I dati vengono presi da un file di testo contenente tutti i nodi connessi fra di loro.

karate.txt - Blocco note

File Modifica Formato Visualizza ?

```
2 1
3 1
3 2
4 1
4 2
4 3
5 1
6 1
7 1
7 5
7 6
8 1
8 2
8 3
8 4
9 1
9 3
10 3
11 1
11 5
11 6
12 1
13 1
<
```

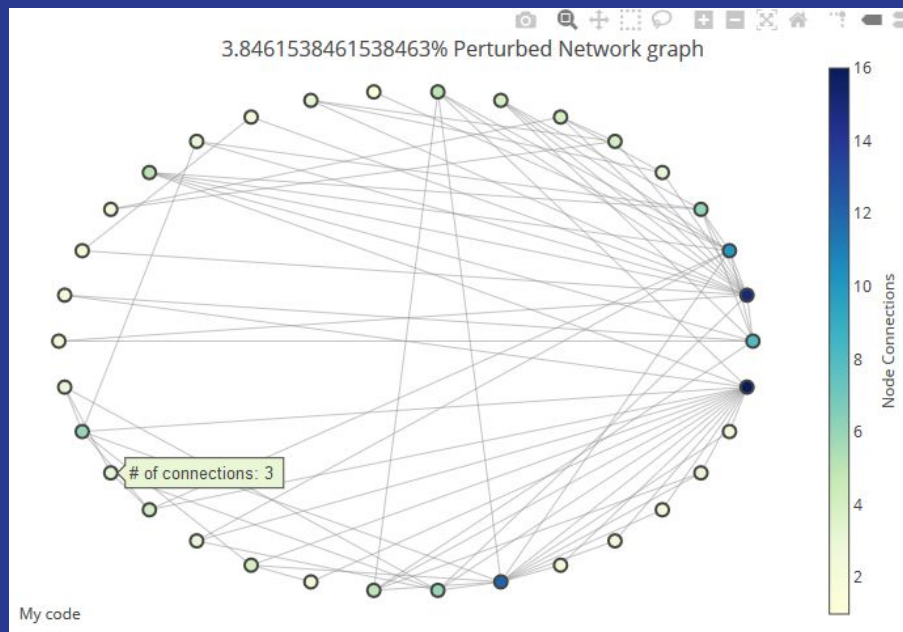
Esempio Random Perturbation



Tools:

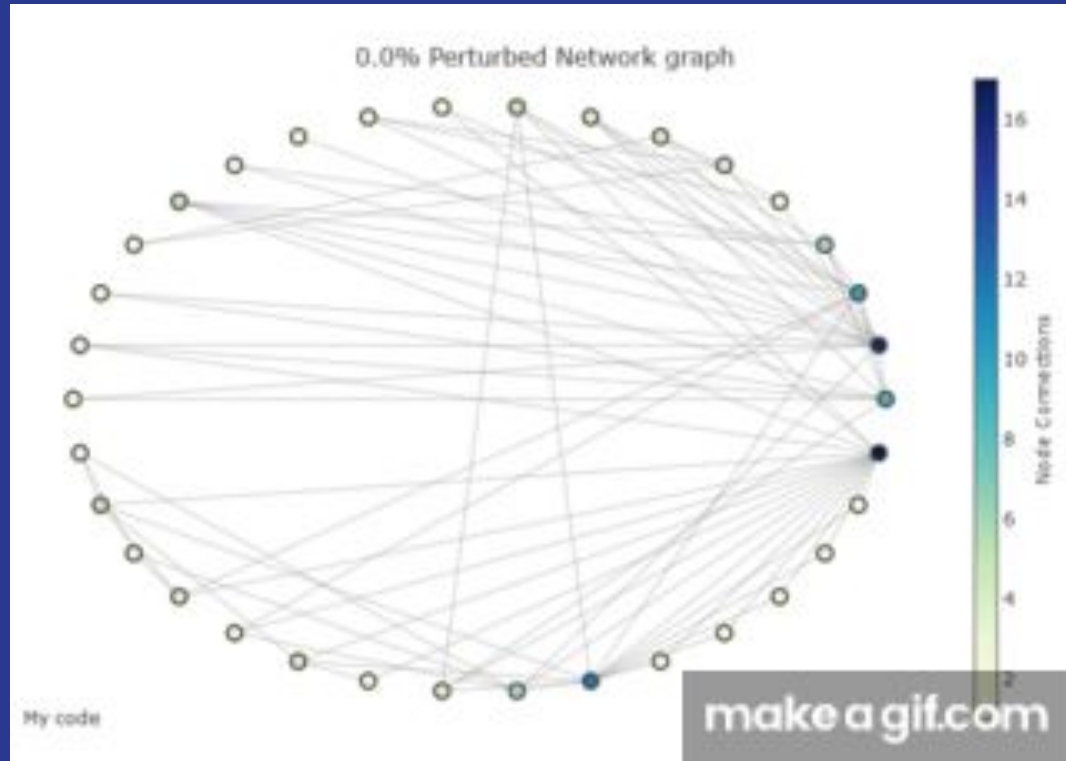
- Python
- NetworkX
- Plot.ly

Esempio Random Perturbation



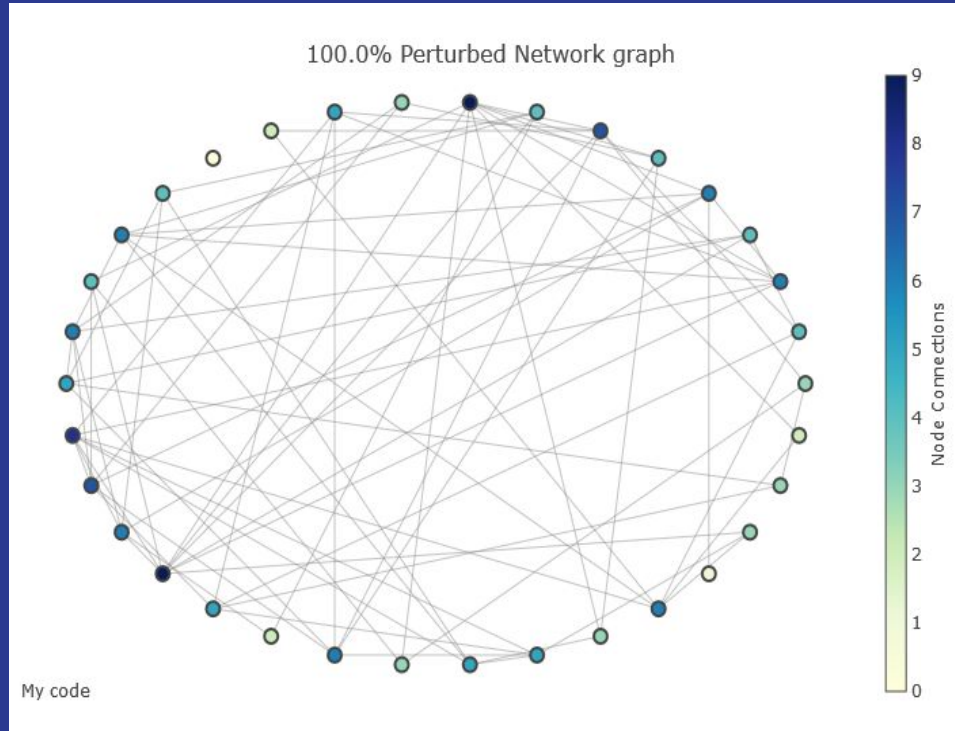
In seguito è stata effettuata una perturbazione del 5% al nostro grafo. Come possiamo notare dall'immagine sono stati rimossi ed inseriti circa il 5% degli archi.

Esempio Random Perturbation

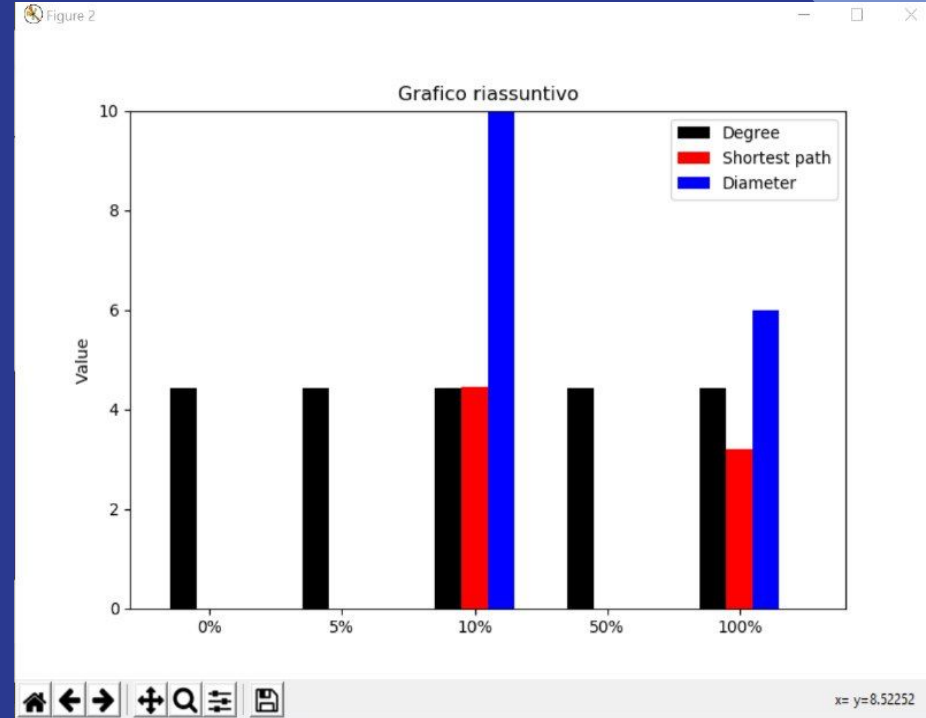
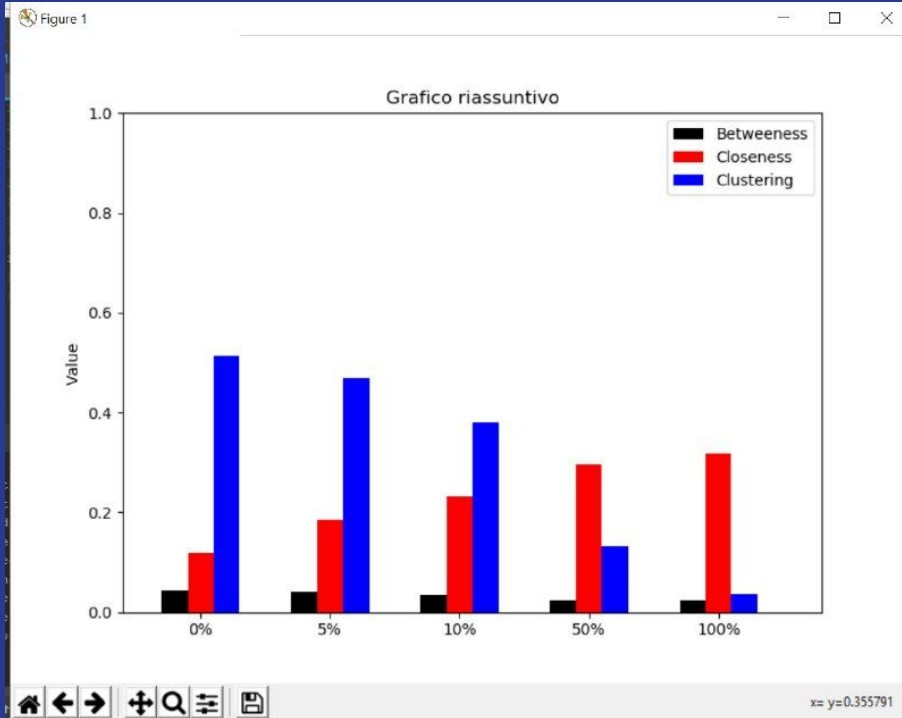


Esempio Random Perturbation

100% Perturbato:



Esempio Random Perturbation



Conclusione

	Enron			
Measure	Original	Perturbed 5%	Perturbed 10%	Random (100%)
Degree	5.0	4.5	4.6	5.0
Diameter	9.0	8.7	7.6	6.1
Path length	4.0	3.2	3.0	3.0
Closeness	0.276	0.293	0.304	0.337
Betweenness	0.005	0.009	0.010	0.014
Clust. Coeff.	0.286	0.242	0.191	0.000

	OUR DATASET			
Measure	Original	Perturbed 5%	Perturbed 10%	Random (100%)
Degree	4.5	4.5	4.6	4.5
Diameter	5	4	5	5
Path length	2.4	2.3	2.3	2.4
Closeness	0.426	0.437	0.429	0.414
Betweenness	0.044	0.041	0.043	0.045
Clust. Coeff.	0.570	0.467	0.427	0.136

Anonimizzazione:



GRAZIE PER L'ATTENZIONE

