

Draft di presentazione progetto del corso di:

“Analisi delle reti sociali applicata ad Internet”

Davide Nunzio Maccarrone

Matricola: 843825

Sommario

| | |
|-----------------------------|----------|
| Abstract | 3 |
| Network datasets | 3 |
| Facebook | 3 |
| Youtube | 4 |
| Academia | 4 |
| Strumenti di analisi | 5 |

Abstract

La proposta di progetto consiste nell'analisi di differenti reti sociali reali scelte dal network repository networkrepository.com.

Il progetto sarà volto all'analisi di tre diverse reti sociali le quali sono state scelte, appositamente, in base al loro diverso contesto di utilizzo:

- **Facebook** è un social network basato sulla condivisione di stati e/o contenuti multimediali tra utenti che hanno una connessione tra loro ("*friendship*"). Gli utenti creano profili che solitamente contengono fotografie e liste di interessi personali, scambiano messaggi privati o pubblici. Possono creare e prendere parte a "Gruppi", in cui si condividono interessi, e a "Pagine" legate a cose o persone di loro gradimento. Gli iscritti a Facebook possono scegliere di aggregarsi a una o più reti, organizzate per città, posto di lavoro, scuola e religione.
- **Youtube** è un social network in cui ogni utente puoi condividere video, su un proprio *canale*, consultabile in ogni momento dagli utenti. Ogni utente può decidere di iscriversi ad un *canale* per ricevere aggiornamenti sugli eventuali nuovi video pubblicati.
- **Academia** è un sito web per ricercatori dedicato alla condivisione delle pubblicazioni scientifiche. La piattaforma può essere utilizzata per condividere articoli, monitorare il proprio *impact factor* e seguire studiosi di tematiche specifiche.

Lo scopo del progetto è quello di analizzare le peculiarità che contraddistinguono le diverse reti per determinarne la struttura ed eventuali comportamenti specifici degli utenti.

Network datasets

Facebook

Nel tipo di dataset preso in considerazione per questa rete sociale i *nodi* rappresentano i singoli utenti, che sono stati resi anonimi e quindi identificabili soltanto mediante in ID. I collegamenti (*edges*) tra loro rappresentano le relazioni sociali tra gli utenti, ovvero la stretta di amicizia tra due persone. La rete analizzata è di tipo *diretta* e *non pesata*.

Esempio network dataset (Facebook)

```
2 1
603 2
624 2
637 2
654 2
665 2
697 2
764 2
791 2
830 2
889 2
```

Youtube

Nel tipo di dataset preso in considerazione per questa rete sociale i *nodi* rappresentano i singoli utenti, che sono stati resi anonimi e quindi identificabili soltanto mediante in ID. I collegamenti (*edges*) tra loro rappresentano le relazioni sociali tra gli utenti. Una relazione permette ad un *follower* di ricevere aggiornamenti non appena l'utente scelto pubblica un nuovo video sul *canale Youtube*. La rete analizzata è di tipo *diretto* e *non pesata*. Il dataset, inoltre, presenta anche un *timestamp* che permette di stabilire quando un determinato *edge* si è stabilito tra due utenti.

Esempio network dataset (Youtube)

```
1,3,1,1165705200
1,4,1,1165705200
1,5,1,1165705200
1,6,1,1165705200
1,7,1,1165705200
1,8,1,1165705200
1,9,1,1165705200
1,10,1,1165705200
1,11,1,1165705200
```

Academia

Nel tipo di dataset preso in considerazione per questa rete sociale i *nodi* rappresentano i singoli utenti, che sono stati resi anonimi e quindi identificabili soltanto mediante in ID. I collegamenti (*edges*) tra loro rappresentano le relazioni tra gli utenti. Stabilendo una relazione con un altro utente, il *follower* riceverà aggiornamenti ad ogni nuova pubblicazione dell'utente scelto. La rete analizzata è di tipo *diretto* e *non pesata*.

Esempio network dataset (Academia)

1 39132
2 174662
2 16721
2 178567
2 125884

Strumenti di analisi

Per l'analisi delle reti verranno calcolati e presi in considerazione i seguenti parametri:

- Number of nodes
- Number of edges
- Maximum degree
- Average degree
- Assortativity Coefficient
- Number of triangles
- Average triangles formed by a edge
- Maximum number of triangles formed by a edge
- Average clustering coefficient
- Clustering coefficient
- Maximum k-core number
- Lower bound on the size of the maximum clique

Per la determinazione di tali parametri verrà realizzato uno script in linguaggio *Python* che utilizzerà la libreria *Network X*. Successivamente i risultati ottenuti verranno analizzati e confrontati tra loro al fine di evidenziare eventuali diversità tra le reti che possano dare risalto alle modalità di creazione di reti degli utenti in uno specifico *social network*.

Tali dati potranno anche essere confrontati con i risultati forniti dal network repository: *networkrepository.com*.