

# Laboratorio di Social Network Analysis

Andrea Fronzetti Colladon – a@andreaafc.com

*Le presenti esercitazioni sono pensate per essere svolte con la presenza del docente e sono pertanto riservate a coloro che seguono le lezioni frontali in aula. I dati presenti in ogni esercizio sono di pura invenzione.*

## PREREQUISITI

Per poter svolgere gli esercizi proposti gli studenti dovranno scaricare ed installare i software gratuiti multiplatforma Pajek e Gephi, tramite i seguenti link:

- **Pajek** > <http://mrvar.fdv.uni-lj.si/pajek/>
- **Gephi** > <http://gephi.org/>  
Il software Gephi richiede che sia installata anche una versione aggiornata di **Java**  
<http://www.java.com/en/download/>

Sarà inoltre utile disporre del pacchetto **Microsoft Office o OpenOffice**, per l'elaborazione dei dati estratti dai software di analisi. Infine, per l'apertura dei file compressi, una possibilità è quella di scaricare il software gratuito **7zip** dal link <http://www.7-zip.org/download.html>

## ESERCIZIO 1 – PAJEK DATA INPUT

A partire dalla matrice di adiacenza riportata nel file Excel **Esercizio1.xlsx**, lo studente disegni la corrispondente **rete orientata** utilizzando il software Pajek.

Come parte dell'esercizio si richiede di:

1. elaborare un file di input per Pajek, formato .txt, rappresentativo della rete in oggetto;
2. ottenere un'immagine .jpg della rete, esportandola dal software Pajek (in cui siano visibili etichette dei nodi e valori sugli archi). Quale layout secondo voi è migliore?
3. raccogliere in un foglio di calcolo i valori, per ciascun nodo presente nella rete, degli indici di Indegree e Outdegree (non pesate), Alldegree (pesata), All Closeness Centrality e Betweenness Centrality. Che considerazioni possiamo dedurre dai valori ottenuti?
4. calcolare i valori di Aggregate Constraint. Chi tra Yattaman e Brontolo può beneficiare maggiormente della presenza di Structural Holes?

## ESERCIZIO 2 – GEPHI e PAJEK ANALISI

A partire dalla rete **orientata Esercizio2.net**:

1. fornire il grafico della All-Degree Distribution della rete. Indicare altresì se si tratta di una rete random o di una rete scale-free;
2. utilizzare il software Pajek per elaborare due immagini .jpg, rappresentative: la prima delle Strong Components (dimensione minima 3); la seconda della partizione di rete derivante dal algoritmo di VOS Clustering (Multi-Level Coarsening + Multi Level Refinement, parametri 1/100/200/200/500), rappresentata scegliendo il layout VOS Mapping 2D

(parametri 10/1000). Quali differenze è possibile notare tra i due metodi di partizionamento?

3. indicare per la rete in esame: il grado di separazione medio, considerando i cammini minimi che uniscono ciascuna coppia di nodi; il diametro; la densità; il grado medio dei nodi. Quali considerazioni possono essere fatte circa la rete in generale? Cosa si può osservare dal confronto di questa rete con quella dell'esercizio precedente?

### **ESERCIZIO 3 – GEPHI ANALISI E VISUALIZZAZIONE**

Elaborare in Gephi la rete **Esercizio3.net**, modificandola in modo da ottenere un layout che ne fornisca una più chiara rappresentazione.

Generare:

1. un'immagine esportata da Gephi, rappresentativa della rete, in cui i nodi abbiano colori diversi (e NON stesso colore più o meno intenso) a seconda della loro classe di modularità;
2. un'immagine esportata da Gephi, rappresentativa della rete, in cui i nodi abbiano dimensione più grande e colore più intenso (lo stesso, NON colori diversi), laddove cresca la loro indegree. Sull'immagine dovranno essere visibili le frecce che esprimono l'orientamento degli archi;
3. un file in cui, per ogni nodo, sia riportato il corrispondente valore di Eigenvector Centrality, calcolato attraverso il software Gephi.

### **ESERCIZIO 4 – ANALISI DI RETE**

*Lo scopo ultimo dell'analista è quello di dare un senso ai dati a sua disposizione, per metterne in evidenza il contenuto informativo.*

Si considerino i file contenuti in **Esercizio4.zip**:

- Foglio di calcolo con elenco dei nodi e dei loro attributi;
- Rete delle relazioni di amicizia (archi pesati da 1 a 5, dove 1 indica un conoscente, 5 uno dei migliori amici);
- Rete degli scambi di consigli in ambito lavorativo (archi pesati da 1 a 5 in base alla frequenza dell'interazione. 1 indica meno di una volta al mese, 5 diverse volte a settimana);
- Rete dei rapporti sessuali (archi pesati da 1 a 5 in base alla frequenza dell'interazione. 1 indica meno di una volta al mese, 5 diverse volte a settimana).

Si richiede allo studente di fornire un'analisi il più dettagliata possibile (visualizzazione e misure base per le reti offerte, legami tra nodi e attributi, legami tra reti differenti, etc..) a partire dai dati a disposizione. La soluzione per questo tipo di esercizi non è univoca e l'analisi può essere affrontata secondo prospettive differenti. Non si richiede quindi una trattazione esaustiva, anche perché il tempo a disposizione è limitato.

**Lo studente consegna i risultati delle analisi unitamente ad un riepilogo delle principali considerazioni che è riuscito ad estrarre dai dati stessi. Siate creativi!**

### **SOLUZIONI DEGLI ESERCIZI**

Per gli studenti che vogliano ripetere a casa i primi 3 esercizi, vengono messi a disposizione degli esempi di soluzione. Gli esempi stessi sono contenuti nel file **Soluzioni.zip** che è protetto da password. La password per aprire il file verrà comunicata dal docente a fine lezione.