Università
Ca'Foscari
Venezia

Corso di Laurea Magistrale in
Amministrazione, Finanza e Controllo

Tesi di Laurea

Ca' Foscari
Dorsoduro 3246
30123 Venezia

Studio formale e qualitativo delle reti aziendali attraverso strumenti della Social Network Analysis

Relatore
Ch. Prof. Andrea Borghesan

Correlatrice
Ch. Prof. ssa Daniela Favaretto

Laureando
Davide Tessari
Matricola 806501

Anno Accademico
2012 / 2013

Alla mia famiglia,

che in questi anni di studi

mi ha sempre sostenuto e supportato

Davide

INDICE

INTRODUZIONE p. 1

CAPITOLO PRIMO

LA SOCIAL NETWORK ANALYSIS – SNA
1.1 Introduzione p. 4
1.2 Approfondimento sulla Social Network Analysis p. 6
1.3 Social Network Analysis a livello italiano p. 7
1.4 Principi fondamentali della Social Network Analysis p. 8

CAPITOLO SECONDO

ANALISI E CENTRALITA' DI UNA RETE SOCIALE
2.1 Identificazione dei confini di una rete sociale p. 11
2.2 Definizione dell'unità di osservazione e di analisi p. 12
2.2.1 Approccio relazionale p. 12
2.2.2 Approccio posizionale p. 14
2.2.3 Approccio statistico p. 14
2.3 Le componenti subgraphs p. 15
2.4 Le triadi e le diadi p. 15
2.4.1 Triadi e le diadi con collegamenti diretti p. 16
2.5 I cicli p. 17
2.6 Gerarchia nel clustering p. 18
2.6.1 Suggerisci un'amico p. 19
2.7 Introduzione alla centralità p. 20
2.8 Il Bottleneck p. 21
2.9 Il Klout Score p. 22
2.9.1 Strumenti di Klout p. 23
2.9.2 Klout style p. 24
2.10 Il Page Rank di Google p. 26
2.11 Le Island Network p. 27

CAPITOLO TERZO

IL LINGUAGGIO DI PROGRAMMAZIONE R ALL'INTERNO DELLA SOCIAL NETWORK ANALYSIS

3.1 Introduzione al linguaggio R p. 29
3.2 Le matrici p. 33
3.3 Le Sociomatrici p. 33
3.4 La Connettività p. 36
3.5 I vertici isolati p. 37
3.6 La Raggiungibilità p. 39
3.7 La distanza Geodedica p. 39
3.8 La Gerarchia p. 42
3.9 La Densità p. 43
3.10 Il Grado p. 45
3.11 La Centralità p. 47
3.12 La Betweenness p. 49
3.13 La visualizzazione grafica in linguaggio R p. 49

CAPITOLO QUARTO

NODEXL E GEPHI

4.1 Introduzione a NodeXL p. 53
4.2 Principali funzionalità di NodeXL p. 54
4.3 L'applicazione NameGenWeb p. 62
4.4 Introduzione a Gephi p. 66
4.5 Principali funzionalità di Gephi p. 68
4.6 Gephi e gli elementi della Social Network Analysis p. 71
4.7 Aspetto grafico in Gephi p. 77

CAPITOLO QUINTO

LA POSIZIONE DEL LEADER ED IL LINGUAGGIO PYTHON

5.1 Il Leader p. 80
5.2 Il numero di Dunbar p. 82
5.3 Noah Friedkin ed Eugene Johnsen p. 83
    5.3.1 Il modello Noah-Friedkin p. 83
    5.3.2 Le condizioni del modello p. 85
5.4 Il modello del Leader in Python p. 86
5.5 Python oltre al Leader p. 90

CAPITOLO SESTO

CASO PRATICO: IL LEOPARDO NERO

6.1 Leopardo Nero: presentazione p. 92
6.2 Leopardo Nero: l'analisi p. 93
6.2.1 Gli influenzatori ed i leader p. 93
6.3.2 La Degree p. 95
6.3.3 Valori e grandezze a confronto p. 97
6.3.4 Average Degree e Graph Density p. 98
6.3.5 Modularity p. 100
6.4 Conclusioni caso pratico p. 102

CONCLUSIONI p. 103

BIBLIOGRAFIA p. 107

SITOGRAFIA p. 111

INTRODUZIONE

L’indole per eccellenza dell’uomo è quella di poter avere il comando e la possibilità di governare la totalità delle variabili possibili in natura. Questo aspetto è ben visibile nel campo gestionale all’interno delle aziende.

Alla luce di ciò, è palesemente visibile che non tutto può essere controllato; una fondamentale condizione per cui un aspetto o una variabile possa essere governata e tenuta sotto controllo ricade sul fatto che essa possa essere misurata.

La scelta della misura è un passo iniziale e fondamentale per lo studio di una variabile, al quale segue in modo quasi automatico la rappresentazione numerica in merito alla qualità e al giudizio sulla variabile in esame. Il concetto di rappresentazione numerica negli ambienti di studio economico e matematici viene rimarcato in moltissime opere, dando così poco valore al giudizio meramente qualitativo. Il giudizio qualitativo può avere un sostanziale significato se viene affiancato ad un giudizio numerico quantitativo.

Soffermandosi poi sull’ambito valutativo, l’attenzione si pone sulla possibilità di riuscire ad identificare e rendere misurabili gli aspetti sia materiali sia immateriali; laddove il processo di misurazione per gli aspetti materiali risulta nella quasi totalità dei casi elementare e semplice, lo stesso non si può affermare sulla valutazione degli aspetti immateriali, che spesso fuggono all’inquadramento numerico.

In ambito contabile, l’avviamento risulta essere l’esempio di come si possa quantificabile, in termini monetari gli aspetti immateriali di un’attività commerciale. Nell’avviamento identifichiamo aspetti quali marchio, prestigio, rete aziendale e commerciale ecc.

Proprio quest’ultimo elemento citato, la rete, sarà il tema veicolante e perno di questo elaborato.

La rete aziendale e commerciale di un utente – sia esso una persona fisica o una azienda – viene rappresentata in un modo diverso dalla solita ottica performativa,

incentrata sulla valutazione ex post degli elementi attraverso un sistema di indici prettamente di natura commerciale.

Gli strumenti designati in questi elaborati saranno diversi e di natura prettamente informatica e che poggiano su concetti di matematica, statistica e sociologia. Proprio quest'ultima è la macroarea in cui sono scaturiti i primi passi della Social Network Analysis.

La Social Network Analysis sarà il tema portante di questa tesi; dopo una introduzione storica degli sviluppi internazionali e italiani accorsi nel secolo precedente, verranno presentanti in prima battuta aspetti matematici e statistici.

La parte numerica che risulterà da una serie di formule, permetterà di individuare aspetti importanti per l'analisi formale e sostanziale di una rete sociale; i coefficienti fino a qui trovati, permetteranno di valutare l'ampiezza di una rete, l'importanza di un nodo in base al numero di collegamenti ecc.

Dalla parte numerica sarà quindi possibile porre le basi per una valutazione qualitativa degli elementi esaminati – nodi, collegamenti e sottoreti – che ci permetterà di identificare come la struttura della rete sociale possa avere pregi e difetti.

Dalla parte qualitativa si potrà individuare se i vari nodi hanno espresso tutte le proprie potenzialità, verificando se ci sia la possibilità di costruire ed instaurare nuovi collegamenti con nodi adiacenti; individuare la natura dei nodi, qualora essi possano essere dei "Colli di bottiglia" per una rete, un ponte di collegamento, o essere definiti Leader per la loro posizione di influenzatori e divulgatori.

La valutazione risulterà essere affiancata dall'utilizzo di software informatici, dove i linguaggi di programmazione risulteranno utili per visualizzare l'importanza dei nodi e collegamenti presenti all'interno della rete sociale.

Sempre attraverso stringhe di programmazione sarà possibile vedere come i principali elementi qualitativi di centralità, densità della rete, ecc., variano al variare di elementi quali il numero di nodi presenti e coefficiente di probabilità. Il tutto per cercare di individuare opportunità di crescita ed espansione.

Con altrettanti strumenti di programmazione informatica, sarà poi testata e visualizzata la posizione ed il comportamento che un nodo potrebbe trovarsi nella

posizione influente di Leader; si vedrà come una nota del Leader può condizionare o meno il pensiero dei propri seguaci.

L'aspetto informatico, oltre alla parte di test e simulazione, sarà presentato anche come mezzo di rappresentazione di rete sociale esistente. Reperendo dati forniti da social network come Facebook®, saranno testati alcuni software applicativi che permetteranno di vedere i risultati delle principali grandezze contenute all'interno della Social Network Analysis, e delle relative rappresentazioni grafiche; rappresentazioni che evidenzieranno elementi di raggruppamento tra i nodi presenti, e di clustering.

L'elaborato, si concluderà con una breve analisi di una rete sociale esistente, cui seguiranno delle brevi conclusioni basandosi sugli elementi della Social Network Analysis portati a conoscenza.

Lo scopo della seguente tesi, riprendendo alcuni concetti espressi precedentemente, vuol essere un modo di poter valutare un aspetto immateriale quale la rete sociale o aziendale, visto da un lato ex-ante e che si distingue dalla valutazione ex post contraddistinta rappresentazione dei Key Performance Indicators o KPI

CAPITOLO PRIMO

LA SOCIAL NETWORK ANALYSIS – SNA

1.1 INTRODUZIONE

L'acronimo SNA deriva dall'abbreviazione dei termini anglosassoni Social Network Analysis, sistema di analisi conosciuto anche come Analisi delle reti sociali o Teoria della rete sociale.

Nata e sviluppatosi dai contributi e dagli studi di Jacob Levi Moreno, la Social Network Analysis viene ritenuta una moderna metodologia di analisi delle relazioni sociali; trova applicazione in diverse scienze sociali quali la sociologia, l'economia e l'antropologia. Di recente utilizzo della Social Network Analysis in campo del management, dove trova applicazione nello studio dei fenomeni del commercio internazionale, nel funzionamento delle istituzioni e nella divulgazione delle informazioni.

I primi ricercatori che si cimentarono in questo campo furono alcuni studiosi di antropologia della scuola di Manchester, che rispondevano ai nomi di J. Clyde Mitchell, John Barnes, Elizabet Bott, Victor Turner, i quali già alla fine degli anni '40 si avvicinarono a questa metodologia di analisi sociale. Questo gruppo lavorava in collaborazione con il ricercatore Max Cluckman, il quale si appoggiava alle strutture offerte dall'allora istituto di ricerca Rhodes-Livingstone con sede a Lusaka, nell'Africa Centrale.

Con a capo Harrison White, ricercatori dell'università di Harward del calibro di Scott, M. Granovetter, R. Burt, B. Wellman[1] e S.D. Berkowitz, posero le basi della Social Network Analysis conosciuta fino ad ora; i loro sforzi portarono alla creazione

[1] http://homes.chass.utoronto.ca/~wellman/

strumenti che potessero analizzare e porre attenzione al comportamento degli attori all'interno della rete

Si riconosce a questo gruppo di ricerca l'importanza del lavoro svolto, grazie al quale si è cominciato a dare una interpretazione matematica ai fenomeni sociologici e relazionali tra le persone: lo testimoniano i concetti scaturiti in quegli anni di studi da cui si evince una sorta di fusione tra concetti prettamente matematici (formule, matrici, grafici, ecc) e concetti sociologici (posizione di una persona all'interno di un gruppo, tipologia di rapporto con i membri di un gruppo, ecc)

Rimane di fatto che negli ultimi decenni la Social Network Analysis ha trovato in diversi campi fonti importanti per contributi teorici, empirici e metodologici. Tra questi non possiamo non citare la rete web e internet, ma anche altre tipologie di ambienti come quello strategico, aziendale, economico e finanziario.

A questa enorme quantità di elementi e concetti portati alla luce dopo decenni di studi e ricerche, nei primi periodi di divulgazione e utilizzo dei personal computer, vi è stata l'introduzione anche dei primi software che riuscivano a dare un ausilio alla ricerca nel campo della Social Network Analysis. L'affiancamento informatico a questa scienza non è stato di fatto automatico e veloce; infatti fra gli esperti del settore si è dovuto trovare una sorta di accordo che permettesse la condivisione delle teorie e dei vari concetti fin lì creati per poter sviluppare ulteriormente gli studi anche grazie allo sfruttamento delle nuove ITC.

Oltre a questo aspetto di tipo informatico, lo sviluppo della Social Network Analysis è stato accompagnato dalla nascita di diverse riviste specializzate nel settore, che hanno rivestito il ruolo importante di strumento di divulgazione e scambio di idee e opinioni. Tra le testate più note nel settore possiamo citare Social Networks, Connections, Journal of Social Strucuture[2] e REDES.

Un fattore che porta la Social Network Analysis a non avere uno sorta di statuto epistemologico potrebbe ricadere proprio su una delle caratteristiche fondamentali di una rete sociale: l'aspetto casuale; attraverso la casualità – tipico concetto di estrazione statistico e matematica – un attore è libero di procedere in modo autonomo ed indipendente alla scelta con cui creare dei rapporti di comunicazione e di creare un rete sociale propria e personalizzata.

[2] http://www.cmu.edu/joss/

Da questa breve analisi introduttiva è possibile comprendere che la Social Network Analysis possa essere interpretata come una scienza che vuole essere non solo un mezzo descrittivo di una rete sociale, ma anche uno strumento che può essere allo stesso tempo predittivo ed esplicativo, dando all'utente che utilizza gli strumenti messi a disposizione, una risposta esauriente, completa di una serie di informazioni dettagliate e globali che tendono ad evitare il più possibile lacune.
Gli studi fin qui compiuti hanno sviluppato diversi concetti e aspetti, che si possono raggruppare in quattro categorie principali:

- identificazione dei confini di una rete sociale
- definizione dell'unità di osservazione e dell'unità di analisi
- identificazione del contenuto relazionale dei vari network
- identificazione della forma assunta dai legami reticolari.

1.2 APPROFONDIMENTO SULLA SOCIAL NETWORK ANALYSIS

La Social Network Analysis può essere vista come uno strumento atto ad analizzare modelli persistenti tra posizioni sociali (come studi effettuati da Edward O. Laumann [3] e *Franz U. Pappi* [4]). Gli studi effettuati da B. Wellman[5] e S.D. Berkowitz[6] definiscono però lo stesso contesto di riferimento, seppur analizzandolo da un diverso punto di vista, come un insieme di legami e interconnessione tra soggetti.
Con la dichiarazione di Georg Simmel secondo cui "la società esiste dove un certo numero di individui interagiscono tra di loro"[7], la Social Network Analysis pose le prima basi per nascita e la costruzione dei propri strumenti per l'analisi dei network sociali.
Con una sorta di primo "abbecedario" per questa materia inedita, Stanley Wasserman[8] e Katherine Faust[9] nel 1994 pubblicarono il libro "Social Network Analysis Methods

[3] http://sociology.uchicago.edu/people/faculty/laumann.shtml
[4] Piselli F., *Reti*, Donzelli, Milano, 1995
[5] http://homes.chass.utoronto.ca/~wellman/
[6] Piselli F., *Reti*, Donzelli, Milano, 1995
[7] Gianfranco Poggi, *Denaro e modernità. La "Filosofia del denaro" di Georg Simmel*, Bologna, 1998
[8] http://www.stat.indiana.edu/people/profile.phtml?id=stanwass

and Applications";ancora oggi questo libro definisce i quattro punti fondamenti su cui l'analisi dei network si focalizzano:

- gli attori e le loro azioni non sono visti come unità indipendenti ed autonome, ma tra loro interdipendenti
- gli attori comunicano tra loro attraverso specifici canali per il trasferimento di risorse materiali e non; queste comunicazioni sono frutti di legami relazionali definiti in termine tecnico linkages
- il network che si viene a creare tra i vari attori deve essere considerato come uno strumento di opportunità che sovrasti i limiti dati dall'azione individuale
- la struttura del network creato si prefigura che possa durare nel tempo.

1.3 SOCIAL NETWORK ANALYSIS A LIVELLO ITALIANO

A livello italiano, i primi passi che la Social Network Analysis è riuscita a compiere sono da ritenersi frutto degli studi compiuti dai sociologi Antonio Maria Chiesi[10] e Fortuna Piselli; la prima pubblicazione che si è espressa in merito a questa scienza è stata scritta nel 1996 da Fortuna Piselli nella rivista "Rassegna Italiana di Sociologia"[11], con la quale porta a termine uno studio fatto sulla Social Network Analysis iniziato con la pubblicazione della sua opera "Reti"[12] dell'anno precedente.
Da questo punto di partenza, all'interno dell'ambito sociologico italiano, le ricerche sulla Social Network Analysis presero piede e si svilupparono sensibilizzando il mondo accademico, che presto riconobbe il potenziale racchiuso da questo neonato settore di studi: la SNA poteva dare supporto a studi in campo imprenditoriale, nel mercato del lavoro e nell'ambito dello sviluppo economico; essa poteva inoltre produrre ottimi risultati anche in campo politico e di élite istituzionale, nell'ambito famigliare, nella gestione dell'immigrazione e della povertà, e non per ultimo, nel campo delle neonate

[9] http://www.faculty.uci.edu/profile.cfm?faculty_id=4660
[10] http://www.unimi.it/chiedove/schedaPersonaXML.jsp?matricola=16692
[11] Fortunata Piselli, "Esercizi di network analysis a Napoli", Rassegna italiana di sociologia, 37(1), 1996, pp. 83-106
[12] Piselli F., *Reti*, Donzelli, Milano, 1995

comunità virtuali (che all'epoca stavano muovendo i primi passi nel loro sviluppo globale).

Nell'ultimo decennio si sono visti diversi studiosi nel campo sociologico che hanno pubblicato numerose opere nel campo dell'analisi delle reti sociali. Il proliferarsi della carta stampata è stato il frutto di un'opera di convincimento sulla validità di questa scienza, sotto gli aspetti dell'efficacia metodologica ed euristica.

Dopo una prima fase esplorativa per decifrare e comprendere il campo e gli spazi di ricerca, con la Social Network Analysis si è cercato di percorre una mission più elevata, con l'obiettivo finale di individuare dei meccanismi utili alla valutazione qualitativo e quantitativo del capitale sociale.

Tra le opere e le pubblicazioni che si devono citare, frutto del'opera di Fortunata Piselli, ci sono "L'analisi delle reti sociali" del 1997 scritto da Enrica Amaturo[13], e "L'analisi dei reticoli" redatta da Antonio Maria Chiesi[14] nel 1999. Un altro importante lavoro, sintesi sull'uso della statistica e dell'algebra nelle reti vaste e complesse che sono nate con l'evoluzione tecnologica, è "Social Network Analysis", scritto dagli autori Alberto Trobia[15] e Veronica Milia[16] nel 2011.

1.4 PRINCIPI FONDAMENTALI DELLA SOCIAL NETWORK ANALYSIS

La Social Network Analysis, SNA, posa le sue attenzioni sulla base di studi e ricerche teoriche e metodologiche nell'ambito dello studio delle reti sociali. L'analisi strutturale di una rete sociale all'interno di questa scienza viene effettuata attraverso una serie di elaborazioni e calcoli algebrici e matematici.

L'analisi di una rete sociale può considerarsi una scienza che poggia su due temi fondamentali:

[13] http://www.dipsociologia.unina.it/docenti/amturo.htm
[14] http://www.unimi.it/chiedove/schedaPersonaXML.jsp?matricola=16692
[15] http://unipa.academia.edu/AlbertoTrobia
[16] http://unipa.academia.edu/VeronicaMilia

- la presa in considerazione che la società può essere vista come un intreccio di collegamenti tra soggetti che istaurano delle relazioni e si scambiano informazioni; questo rappresenta il cuore degli studi effettuati dai ricercatori all'interno della Social Network Analysis
- la possibilità di riuscire a interpretare e governare certe variabili relazionali attraverso l'applicazione di formule e rappresentazioni grafiche

L'utilizzo degli strumenti messi a disposizione della Social Network Analysis ha portato negli ultimi anni ad un incremento esponenziale della letteratura in merito a questo argomento; il propagarsi e il successo di questa scienza può essere attribuito al fatto di poter rappresentare in modo semplice e comprensibile anche a non addetti al settore dei fenomeni sociali spesso complessi e interdipendenti tra loro.

All'interno della Social Network Analysis, la società viene vista sottoforma di un concetto fondamentale rappresentato da una rete di relazioni (più o meno estese e strutturate) che intercorrono tra soggetti e istituzioni di diversa natura. Il presupposto fondamentale prevede che il comportamento dei soggetti, venga plasmato dalle relazioni poste in essere.

Scopo principale della Social Network Analysis è lo studio dei legami (ties) tra i vari soggetti o attori (nodes). Dall'analisi, derivano poi una serie di misurazioni e indici quali la coesione e la centralità – solo per citarne alcune – che risaltano le proprietà e le caratteristiche di un determinato network. L'analisi di una rete sociale non si esaurisce però con queste misurazioni, che sono infatti a loro volta approfondite da ricerche di sottoreti, di gruppi, somiglianze tra reti o altri studi specifici, fra i quali si cita in particolare l'analisi dell'equivalenza strutturale.

La Social Network Analysis, viene utilizzata in diversi campi scientifici come citato precedentemente: in ambito matematico ad esempio trova spazio all'interno della cosi detta "Teoria dei Grafi", anche se gli sviluppi di questa microarea ha generato lo sviluppo e l'utilizzo di una terminologie tecnica e specifica e indipendente da quella della SNA. A volte capita infatti che alcuni termini usati nella Social Network Analysis per esprimere concetti relativi alla teoria dei grafi siano diversi dal gergo tecnico di quest'ultima; per questo motivo si sta cercando di uniformare il linguaggio attraverso

la commensurabilità di termini, affinché i due ambiti, seppur autonomi, possano interagire con un linguaggio comune.

Il compito della Social Network Analysis diventa quindi quello di identificare le origini dei rapporti tra soggetti nel tentativo di rilevare le conseguenze a cui questi possono portare; fine ultimo sarà di creare degli schemi di rappresentazione che permettano di comprendere le reti sociali e il loro "comportamento".

In conclusione, soffermandosi sull'aspetto più prettamente informatico e della programmazione, tra i software più utilizzati nelle Social Network Analysis, si possono citare UCINET[17], NETMINER[18] e PAJEK[19] (quest'ultimo scaricabile gratuitamente dal web). Trovano inoltre spazio ulteriori software scaricabili gratuitamente quali AGNA[20] e NETDRAW[21], quest'ultimo sviluppato per la rappresentazione puramente grafica. A livello di software con natura open source (cari al mondo degli sviluppatori software) degno di nota è GEPHI[22], che mette a disposizione un supporto nella visualizzazione oltre che nell'analisi delle reti sociali.

[17] https://sites.google.com/site/ucinetsoftware/home
[18] http://www.netminer.com/index.php
[19] http://vlado.fmf.uni-lj.si/pub/networks/pajek/
[20] http://www.freewebz.com/benta/agna/download.htm
[21] https://sites.google.com/site/netdrawsoftware/
[22] https://gephi.org/

CAPITOLO SECONDO

ANALISI E CENTRALITA'DI UNA RETE SOCIALE

2.1 IDENTIFICAZIONE DEI CONFINI DI UNA RETE SOCIALE

In merito ai confini di una rete sociale è possibile notare come studi intrapresi da diversi ricercatori portino ad una sorta di brainstorming di idee e teoremi; E. O. Lauman[1], P Mardsen e D. Prensky furono fra quelli che posero le proprie attenzioni proprio sulla struttura di una rete sociale.

In poche righe, gli studi effettuati portarono alla definizione di due concetti diversi del confine e della struttura di una rete, sotto due diversi punti di vista.

Dal primo si può dedurre come la rete sociale possa essere vista mediante un approccio realista, secondo il quale i partecipanti riconoscono pienamente la rete di relazioni nella sua struttura, assecondandone pregi e difetti, potenzialità e debolezze.

Dal secondo invece ne deriva un approccio nominale, attraverso il quale i membri di una rete sociale ne approvano l'esistenza, ma non riconoscono, di fatto, i confini e i limiti da essa creati.

Questo è solo un esempio della molteplicità di studi che hanno portato allo sviluppo di approcci alternativi, che analizzano il lato strutturale e dei confini all'interno di una rete sociale; oltre a quelli già indicati in precedenza, si possono citare ad esempio quelli relativi all'aspetto posizionale e reputazionale.

L'approccio posizionale raggruppa tutte quelle analisi che permettono al ricercatore di ottenere come risultato finale la possibilità di definire la rete sociale attraverso l'individuazione di gruppi, zone o aree di interesse. In ambito reputazionale invece il

[1] http://sociology.uchicago.edu/people/faculty/laumann.shtml

ricercatore, grazie agli strumenti in possesso, riesce a dedurre l'importanza di un uno specifico nodo o di una sezione della rete relativamente alla sua globalità, questo attraverso all'indicazione delle relazioni più strette che si instaurano fra i diversi soggetti.

2.2 DEFINIZIONE DELL'UNITA' DI OSSERVAZIONE E DI ANALISI

In questo spezzone di studi, la ricerca e l'analisi si sposta dalla struttura globale della rete a quella relativa ai singoli elementi che la compongono e alle relazioni che li collegano. L'analisi analitica permetterà, invece, di parcellizzare la rete sociale, soffermandosi sullo studio di sottoreti locali.
All'interno di esse, potranno essere individuate delle reti egocentriche o ego-network, in base al comportamento di alcuni nodi, che hanno la particolare capacità di attrarre altri nodi ed istaurare nuovi collegamenti.
Per quanto riguarda lo studio e l'analisi delle reti in modo globale invece, è utile sottolineare la presenza di tre diverse modalità di analisi, di seguito elencate ed approfondite singolarmente:

- Approccio relazionale
- Approccio posizionale
- Approccio statistico

2.2.1 Approccio relazionale

L'approccio relazionale in una rete sociale permette di studiare come gli attori siano connessi l'uno all'altro, ponendo attenzione anche al flusso informativo che veicola tra i nodi, verificandone così anche l'intensità comunicativa.
Attraverso questa visione si individua come il punto focale di studio sia rappresentato dalla modalità con cui gli attori entrano in contatto fra loro. Questa analisi porta alla luce diverse possibilità di connessione e soggetti di analisi:

- legame singolo che lega un nodo ad un altro nodo, dal quale è possibile ulteriormente decifrare l'intensità o la debolezza della relazione;
- molteplicità dei collegamenti che un nodo intrattiene con altri (multiplexity);
- reciprocità dei collegamenti tra due nodi,

Rimanendo sempre nell'ambito dell'osservazione, è possibile affermare che il fulcro dello studio della Social Network Analysis non è il semplice soggetto o nodo; si può invece affermare con certezza che questa scienza moderna si discosti dalla sola analisi del singolo elemento sulla base di sondaggi e questionari (surveys) per porre l'attenzione, come già più volte ripetuto, sull'intera struttura, ponendo l'accento sulle relazioni di collegamento e sulla loro creazione.
Le relazioni sono un elemento dinamico che lasciano spazio a molteplici modalità di ricerca; questo è il risultato dovuto al fatto che esse rappresentano il frutto di una sorta di gioco strategico effettuato dai vari attori all'interno della rete sociale. Possiamo definire le relazioni la parte dinamica di una rete sociale e hanno talvolta il compito di condizionare i comportamenti tra attori e decifrarne l'importanza e il prestigio.
Dagli studi effettuati da R. Burt, rimanendo sempre nel campo dell'osservazione, si deduce infine come alcune variabili fisiche dei soggetti all'interno di una rete sociale (come età, sesso e caratteristiche individuali) siano delle rappresentazioni imperfette e poco rappresentative sotto l'ottica della Social Network Analysis. Sempre secondo R. Burt l'importanza deve essere posta sull'atteggiamento che i soggetti hanno all'interno di un ambiente e su come essi instaurino le proprie relazioni
Caratteristici di questo approccio sono, come in seguito vedremo per gli altri approcci, una serie di concetti e grandezze scaturite dagli studi e dalle ricerche effettuati nei vari decenni all'interno della Social Network Analysis; per questo specifico approccio si citano in particolar modo le principali grandezze di seguito riportate:
:

- Centralità
- Densità
- Connettività

- Incidenza
- Cicli

2.2.2 Approccio posizionale

In base a questa tipologia di analisi l'osservazione sarà concentrata su tutti quegli elementi che compongono la struttura della rete sociale, in particolar modo la disposizione spaziale dei nodi.
Dagli studi intrapresi attraverso questa metodologia di ricerca sono state introdotte misure quali:

- equivalenza strutturale
- equivalenza regolare
- Blockmodeling[2]

2.2.3 Approccio statistico

In quest'ultimo approccio, ma non per importanza, risulta fondamentale l'importanza di poter governare e calcolare le variabili delle reti sociali. L'aspetto statistico può rappresentare le classiche due facce di una medesima moneta; da un lato si può dire che esso abbia una natura di mero calcolo statistico e descrittivo, ma dall'altra rivela la capacità di poter governare le variabili di una rete sociale per poter costruire piccole reti elementari. Quest'analisi analitica può essere interpretata come la possibilità da parte del ricercatore di poter analizzare situazioni a livello micro, per poi definire una soluzione a livello macro.

[2] Strumento di comparazione tra reti sociali che si basa sull'utilizzo di matrici

2.3 LE COMPONENTI SUBGRAPHS

Soffermandosi sull'aspetto analitico della Sociale Network Analysis bisogna sottolineare e definire due concetti fondamentali: i Componenti e i cosiddetti Subgraphs. I Componenti vengono definiti come delle porzioni di network che sono disconnesse da il resto del network, invece i Subgraphs sono sottosistemi o "*subset*" di nodi all'interno di una rete; questi sottosistemi andranno a formare un gruppo di nodi ed edges che comunica tra loro e formeranno una parte fondamentale all'interno di qualsiasi network preso in esame.
All'interno dei Subgraphs, sono inseriti gli Ego Networks; sottoreti (o all'inglese subnetworks) che incentrano la loro attività di comunicazione attorno ad un particolare nodo di riferimento. Esempi di queste sottoreti sono rappresentati dalle rete di contatti che un soggetto ha nei principali social network come Facebook o Twitter

2.4 LE TRIADI E LE DIADI

Nello studio delle reti e della Social Network Analysis, trova spazio un approfondimento particolare sulle Triadi. Esse non sono altro l'estrema esemplificazione di una rete vista al microscopio. Le Triadi non solo altro che tre nodi connessi l'uno all'altro in diverse forme e modi:

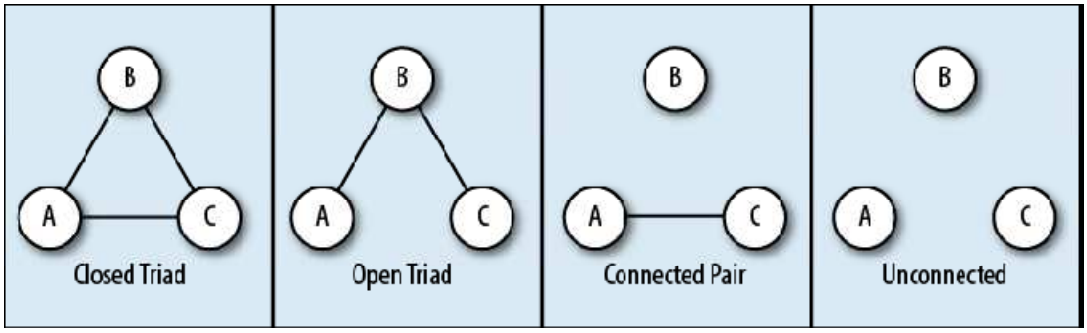


Figura 2.1: Tipologie di triadi

Il classico esempio di una Triade è rappresentato dalla comunicazione che intercorre tra il nodo A che trasmette al nodo B, e a sua volta a C; quest'ultimo nodo poi comunicherà poi nuovamente al nodo A formando una sorta di circolo dell'informazione.
Le Triadi si distinguono dalle Diadi o Dyads, reti nelle quali sono collegati fra loro solamente due nodi che comunicano l'un l'altro.

2.4.1 Triadi e diadi con collegamenti diretti

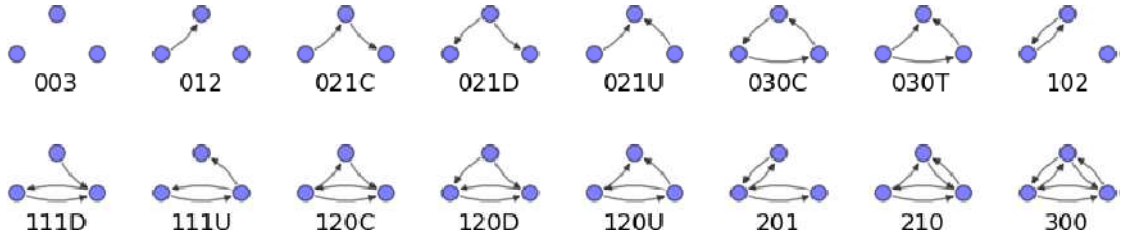


Figura 2.2: Tipologie di collegamenti all'interno dell'ambito delle triadi e diadi

Nella figura soprastante sono visualizzate le diverse combinazioni di direzioni tra nodi; con collegamenti a livello unidirezionali e bidirezionali; si possono formare complessivamente 16 diverse combinazioni.
Ad ognuna delle 16 tipologie scaturite dai diversi collegamenti, viene poi associato un codice di identificazione che permette di capire numero e natura dei collegamenti all'interno di questa semplice rappresentazione grafica di un network.
Procediamo per step alla comprensione e l'interpretazione del codice sottostante all'identificazione delle triadi. Il codice è composto in ordine da:

- un primo numero che identifica il numero dei collegamenti bidirezionali tra i vari nodi
- un secondo numero che si riferisce al numero dei collegamenti diretti ai singoli nodi
- un terzo numero che identifica il numero dei nodi non collegati fra loro

- un’eventuale lettera, che integra il codice e specifica la particolarità del collegamento tra i vari nodi. Le lettere utilizzate sono:
  - U – up-, se il collegamento è sviluppato su nodi superiori
  - D – down-, se il collegamento tra nodi è situato a livello inferiore
  - C – circle-, laddove i nodi sono interconnessi tra loro a forma di circuito
  - T – transitve -, qualora un unico nodo riesca a comunicare in modo diretto con due nodi diversi

2.5 I CICLI

All’interno della Social Network Analysis trova spazio un capitolo fondamentale che riguarda i Cicli.

I cicli sono rappresentati come un gruppo molto coeso di nodi o persone che appartengono ad uno stesso gruppo o rete; una caratteristica particolare dei cicli è che gli elementi che lo compongono non si interfacciano con elementi di altri gruppi esterni. Questo è in sostanza la descrizione e la definizione che si può porre a capo della materia della Social Network Analysis; in campo matematico ed algebrico, però esiste una definizione più dettaglia e rigorosa di ciclo.

Un’ulteriore aiuto per la comprensione del concetto di ciclo, deriva dalla terminologia tecnica della Social Network Analysis. Un ciclo viene definito come “maximal complete subgraph”: esso sottolinea , oltre al fatto che tutte le persone sono collegate fra loro a formale un gruppo, attraverso il termine maximal il fatto che un nodo o una persona non possa essere aggiunta al ciclo se essa ne preclude il minimo funzionamento e collegamento fra soggetti, compromettendone la comunicazione di informazioni.

Le caratteristiche fin qui descritte dei cicli, possono proprio per la natura di queste componenti nella Social Network Analys, comportare delle problematiche; non sono poco frequenti occasioni in cui cicli tra loro vicini in un network possano scaturire in una serie di conflitti, compromettendo la funzionalità di quest’ultimi.

2.6 GERARCHIA NEL CLUSTERING

Un altro elemento importante da valutare nella Social Network Analysis, è dato dalla definizione del ruolo del clustering. Rappresentato dal Coefficiente di Clustering, analizza il numero di contatti o amici all'interno della rete che sono a loro volta in contatto tra loro.

Questa analisi può essere applicata a tutte le tipologie di network, ma i risultati migliori vengono prodotti nelle reti a larga scala e di grandi volumi; in genere i valori medi di un coefficiente di clustering è di difficile interpretazione nelle reti più piccole.

La successione di collegamenti fra nodi all'interno di una rete sociale conduce la citazione del concetto di distanza. Nella Social Network Analysis, abbiamo due tipi di distanze; la prima definisce in modo elementare la distanza grafica che intercorre tra un nodo e un altro all'interno di uno stesso network. La seconda analizza la distanza prendendo i nodi a coppie; queste coppie verranno poi confrontate con altre per determinare se sussistono fra coppie di nodi distanze simili.

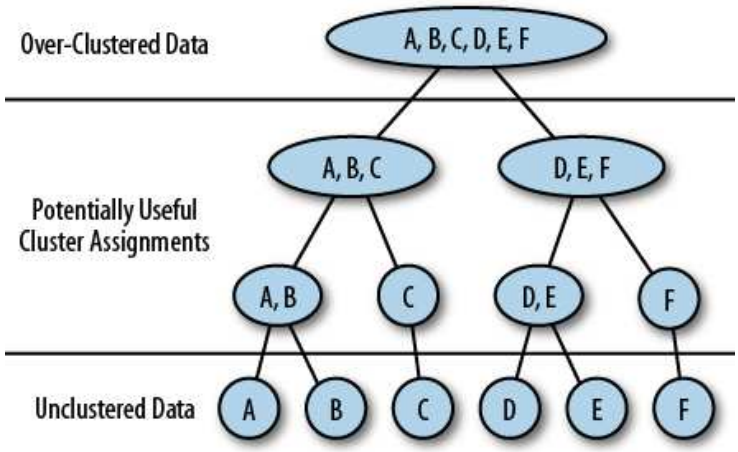


Figura 2.3: Gerarchia del clustering

2.6.1 Suggerisci un amico

Nel principali social network presenti nella rete troviamo con frequenza la funzione che permette di poter istaurare collegamenti con nuovi contatti; questa capacità, è il frutto di un calcolo a base di matrici.

Questa applicazione si attiva in diverse modalità nei principali social network, a seconda degli algoritmi che li sostengo; basti pensare che un “like” in Facebook o un hastag in Twitter – solo per citarne alcuni – dia già inizio all’inter algebrico di questa funzione.

Ora di seguito vediamo un caso pratico. Prendiamo per esempio una matrice 5X6 formata nel seguente modo:

	1	2	3	4	5
A	0	0	0	0	1
B	1	0	0	0	0
C	1	1	0	0	0
D	0	1	1	1	1
E	0	0	1	0	0
F	0	0	1	1	0

Procediamo ora nella comparazione dei nodi D ed F; essa avviene attraverso la mera moltiplicazione verticale e successiva somma in senso orizzontale dei valori presenti. Il risultato sarà il seguente:

D		1	2	3	4	5	Sum
	D	0	1	1	1	1	4
	F	0	0	1	1	0	2
	$F * G$	0	0	1	1	0	2

I seguenti calcoli permettono di individuare, tra il nodi D ed F, la presenza di 2 nodi in comune; procedendo alla stessa maniera per gli altri valori contenuti nella matrice precedente, avremo questo risultato:

$$\begin{array}{c|cccccc} & A & B & C & D & E & F \\ A & 0 & 0 & 0 & 1 & 0 & 0 \\ B & 0 & 0 & 1 & 0 & 0 & 0 \\ C & 0 & 1 & 0 & 1 & 0 & 0 \\ D & 1 & 0 & 1 & 0 & 1 & 2 \\ E & 0 & 0 & 0 & 1 & 0 & 1 \\ F & 0 & 0 & 0 & 2 & 1 & 0 \end{array}$$

In questa rappresentazione risulta infine che essa non è altro che la moltiplicazione della matrice originale per la sua trasposta

$$(A^t): AA = A * A^T$$

2.7 INTRODUZIONE ALLA CENTRALITA'

Attraverso la Social Network Analysis si riesce a fornire una serie di risposte grazie alla misurazione del potere e dell'influenza che le persone riescono ad avere nei confronti degli altri soggetti all'interno di una rete di contatti.

La Centralità può rispondere i quesiti quali: "Quanto è importante il mio network?", oppure "Qual è il punto di forza del mio network?". Questo strumento infatti cerca attraverso diverse metodologie di produrre un risultato finale con cui si assegna un'importanza finale globale alla rete analizzata.

L'importanza e la forza di un network potrebbe dipendere dalla mole di dati e dalla qualità di quest'ultimi ed altresì, dipendere dal numero di soggetti che sono all'interno della rete stessa.

All'interno del capitolo dedicato alla Centralità, non si può non citare il ruolo della cosiddetta Degree Centrality; con questo termine anglosassone si identifica l'analisi che permette di individuare il nodo all'interno di una rete che ha un significativo numero di collegamenti. Riportando a livello pratico questo concetto, si riscontra che nei più popolari social network, la Degree Centrality viene determinata dal numero

dei contatti in Facebook[3], followers in Twitter[4] o in Reddit[5], dal numero dei cosiddetti upvotes (concetto simile a quello dei karma nei più comuni blog).
Un successivo step nella conoscenza del concetto della Centralità, viene dato da quella che, con il termine anglosassone, viene definita come Closeness Centrality; quest'ultima viene calcolata attraverso una serie di step che prevedono:

- l'uso dell'algoritmo di Dijkstra, il quale analizza le distanze tra i vari nodi costruendo una sorta di tabella che riassumi i valori
- per ogni nodo poi vengono calcolati la distanza media e la distanza media che intercorre tra il nodo analizzato e il resto dei nodi

Il risultato finale sarà un numero compreso tra 0 e 1; più il risultato si avvicina all'unità più i nodi della rete sono in prossimità l'uno all'altro, evidenziando una distanza media tra i nodi bassa.

2.8 IL BOTTLENECK

Continuando nella comprensione del ruolo della Centralità all'interno della Social Network Anlysis, vi è un ulteriore aspetto su cui vale la pena concentrare le proprie attenzioni:il ruolo del Collo di Bottiglia, conosciuto anche con il termine inglese di Bottleneck.
La figura sotto riportata spiega graficamente come il nodo D sia il collo di bottiglia per tutto il network; si nota facilmente infatti come le trasmissioni dei nodi A, B e C con i nodi F, G e H (e viceversa) trovino nel punto D il solo ed unico snodo, punto di transito per la trasmissione delle informazioni.

[3] https://www.facebook.com/
[4] https://twitter.com/
[5] http://it.reddit.com/

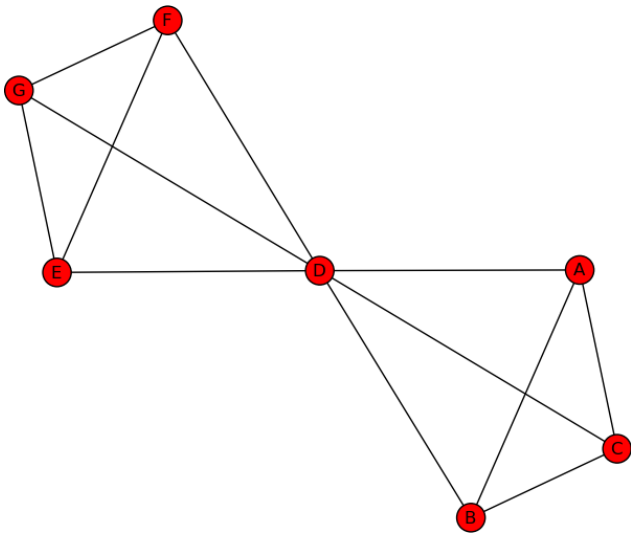


Figura 2.4: Esempio grafico di rete con la presenza di un Bottleneck

2.9 IL KLOUT SCORE

Un ulteriore indicatore che può essere di aiuto alla Social Network Analysis è quello ideato da Joe Fernandez e Binh Tran , fondatori del sito klout.com[6], attraverso il quale è possibile calcolare il Klout Score di un network. Attraverso questo indicatore ,Joe Fernandez e Binh Tran hanno voluto portare alla ribalta l'importanza e l'influenza degli utenti dei social network con i propri contatti. Questo coefficiente che ha un range tra 0 e 100, dipende, dall'ampiezza del network dell'utente, dal contenuto delle comunicazioni e dal livello di feedback creato.

Attraverso il sito klout.com è possibile vedere in modo semplice e basilare la misurazione delle centralità di una rete personale. Rimanendo in tema di social network, il sito supporta i blasonati Facebook®, Twitter®, Linkedin®[7], Google+®[8] e Foursquare®[9].

L'analisi di Klout si base sull'interrogazione periodica dei vari social network che un utente – sia privato che istituzionale – può avere; i social network supportati sono

[6] http://klout.com/home
[7] http://it.linkedin.com/
[8] http://www.google.com/+/learnmore/?hl=it
[9] https://it.foursquare.com/

quasi la totalità dei presenti nella rete web, Twitter®, Facebook®, Google+®, LinkedIn® e Youtube® solo per citarne alcuni dei più famosi.

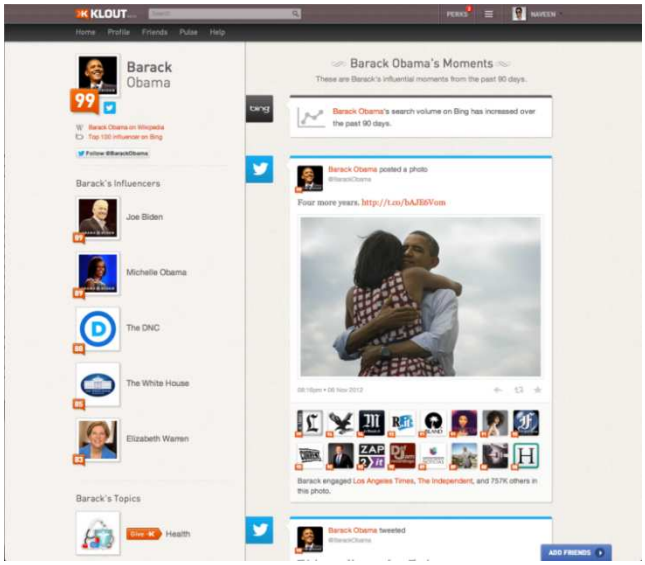


Figura 2.5: Dashbord di Klout

2.9.1 Gli strumenti di Klout

Le metriche che Klout utilizza per valutare l'influenza di un soggetto o di una azienda all'interno di una rete sono riassunte in tre tipologie:

- *True reach*: identifica il numero di follower che condividono i contenuti pubblicizzati da un soggetto;
- *Amplification*: rappresenta il numero di persone che rispondono ai contenuti pubblicati o che ri-condividono i contenuti pubblicati da altri follower
- *Network*: analizza in modo più approfondito il tema della True reach; sotto questo aspetto viene presa in considerazione la posizione dei top influencer che aumentano il punteggio complessivo di Klout

2.9.2 Klout Style

Dalle metriche fin qui esposte, e dagli anni di servizio offerto, la start-up americana Klout, è riuscita a determinare una matrice complessa per catalogare le diverse tipologie di utenti. Nelle quattro assi che compongo la matrice, troviamo:

- Dall'ascolto (*Listening*) alla partecipazione (*Partecipating*)
- Dalla condivisione (*Sharing*) alla creazione di contenuti particolare e non comuni (*Creativity*)
- Dalla frequenza sporadica delle pubblicazioni (*Casual*) a quella più continua di contenuti (*Consistent*)
- Dall'interesse verso temi specifici (*Focused*) a quelli di più ampio interesse (*Broad*)

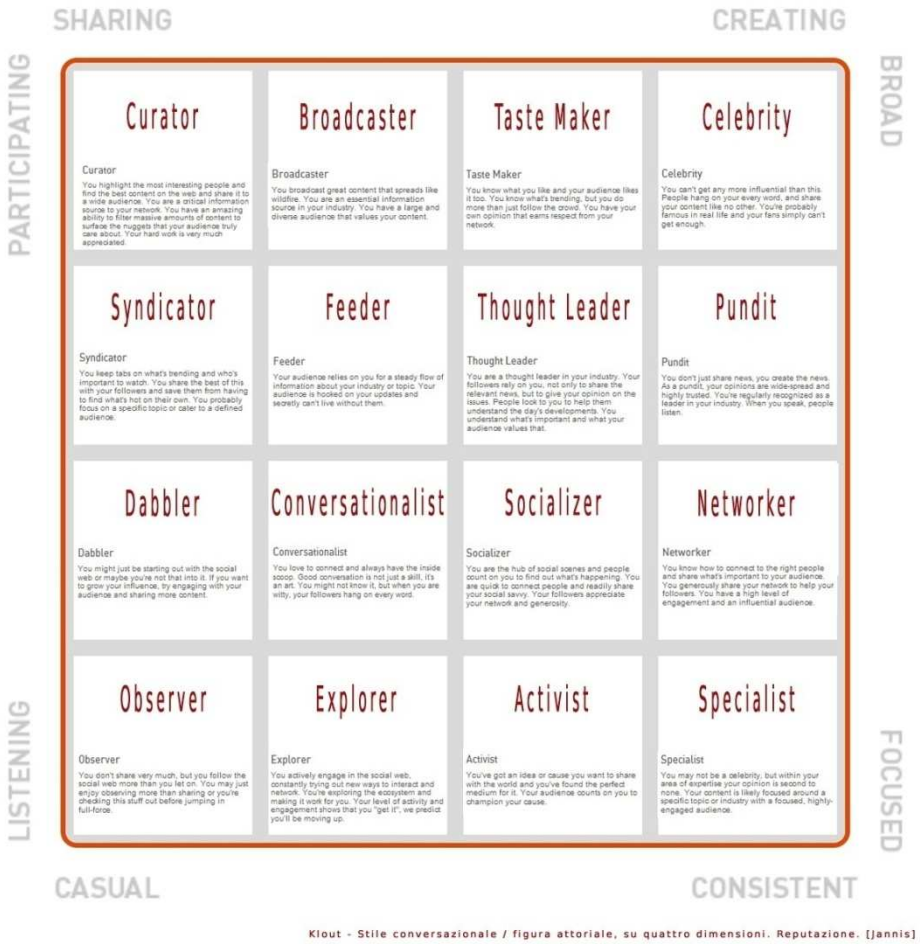


Figura 1.6: La matrice degli stili Klout

Dall'intersezione di questi assi, Klout riesce a scaturire ben 16 diverse posizioni:

- *Curator*: identifica chi riesce a selezionare le informazioni più rilevanti per poi metterle a disposizione dei rimanenti membri del gruppo
- *Broadcaster*: chi diffonde i contenuti ad una platea ampia di follower
- *Taste Maker*: coloro che danno un tocco di originalità nella diffusione dei contenuti che interessano al gruppo.
- *Celebrity*: coloro che hanno un numero molto ampio di follower, che a loro volta condividono i contenuti pubblicati.
- *Syndacator*: chi pubblicizza contenuti relativi a temi specifici e di nicchia.
- *Feeder*: identifica i soggetti che pubblicizzano un continuo flusso di informazioni; quest'ultime riceveranno una sorta di attenzione da parte dei follower attraverso un semplice click o un "mi piace"
- *Thought Leader*: in questa categoria sono identificate le persone che oltre a condividere contenuti fondamentali e con un audience elevata, rispondono in modo positivo ai vari stimoli dati dai propri follower
- *Pundit*: identifica tutte le persone che si esprimono come opinionisti su un determinato tema
- *Dabbler*:chi si è appena inserito all'interno di una rete sociale e trova difficoltà nell'affermarsi nella pubblicazione e condivisione dei contenuti
- *Conversationalist*: l'attività principale di questi soggetti è di conversare e avere frequenti interazioni attraverso i social media.
- *Socializer*: viene riscontrata a questi soggetti la capacità di essere una sorta di hub per la socializzazione tra utenti.
- *Networker*: chi ha costruito una rete di contatti di prestigio e di valore, tanto da poter essere data ad essa una valutazione economica
- *Observer*: chi non condivide molto, ma preferisce osservare le pubblicazioni degli altri utenti
- *Explorer*: i soggetti che rientrano in questa categoria sono quegli utenti che con la mera attività di condivisione di contenuti, iniziano a comprendere il funzionamento ed il meccanismo delle reti sociali

- *Activist*: chi si focalizza su un determinato tema, e pubblica solamente contenuti relativi allo stesso, riuscendo a coprire il ruolo di promotore dello stesso tema.
- *Specialist*: dalla parola stessa si identifica chi è tecnico del settore e pubblica nuove informazioni in merito a temi particolari

2.10 IL PAGE RANK DI GOOGLE

Un tipo particolare di analisi sulla centralità dei nodi all'interno di un proprio network, viene offerto dal servizio PageRank® di Google®. Questo servizio permette di soffermarsi sullo studio dei collegamenti in entrata e in uscita nei vari nodi; lo scopo è di ricavare dettagli relativi al flusso di informazioni che attraversano i nodi esaminati. Originariamente questo strumento fornito dal colosso Google era stato sviluppato per la mera analisi dei siti e portali web; successivamente è stato adattato allo studio delle reti, implementandolo anche nei siti a cui fanno capo i principali social network.
I valori che scaturiscono dall'analisi di PageRank® presentano un valore che varia da 0 a 1; in particolare se l'indice assume un valore è pari a 0,5 Significa che quel link ha probabilità pari al 50% di essere visitato da un utente del network.
Basato sul Processo di Markov ha come formula base e semplificata quella sotto stante riportata:

$$PR[A] = \frac{1-d}{N} + d\left(\sum_{k=1}^{n} \frac{PR[P_k]}{C[P_k]}\right)$$

In dettaglio le variabili identificate nella formula soprastante sono:

- *PR[A]*, identifica il livello di ranking che una pagine web (*A*) riesce a raggiungere
- *N*, rappresenta il numero totale di pagine presenti nel sito web

- *n*, raggruppa il numero dei link delle pagine web esterni verso la pagina web *A*
- P_k, il numero di pagine web esterne che contengono collegamenti verso la pagina web *A*
- $PR[P_k]$, rappresenta il valore di PageRank® delle pagine P_k
- $C[P_k]$, identifica il volume complessivo di link contenuti nella pagina web P_k
- *d*, il quale identifica il valore di *damping factor*; coefficiente prefissato da Google a 0,85, il quale rappresenta la parte di valore aggiunto che i PageRank® delle pagine web esterne apportano alla pagine web A esaminata

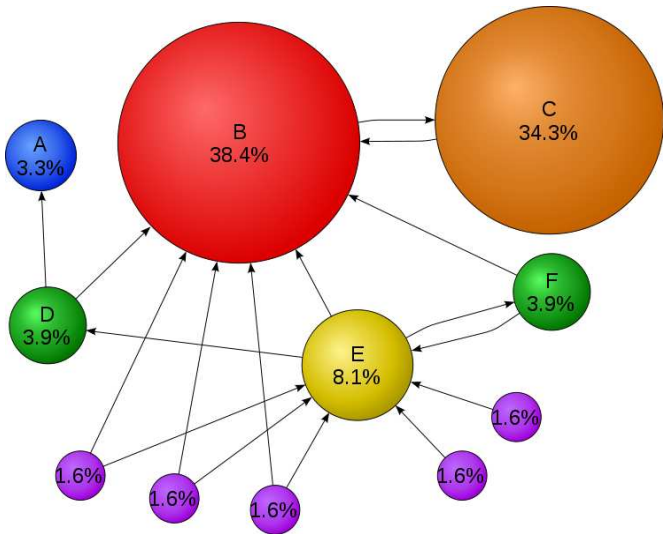


Figura 2.7 : Esempio di un calcolo del range a livello grafico

2.11 LE ISLAND NET

Una tecnica per analizzare i network è quella che viene chiamata Metodo delle Isole o Island Net.

Questa tecnica prevede che ad ogni nodo o punto della rete che si vuole analizzare, venga attribuito un valore; a livello grafico, l'insieme dei valori formerà una sorta di immagine stilizzata, a cui è possibile ricondurre la forma di un'isola. L'immaginario

livello del mare è dato dal valore “floor”, che permette di individuare i nodi con i valori attribuiti più alti. Con l’innalzamento del valore “floor” individueremo i nodi principali e rilevanti per la rete sociale presa in esame.

Questa tipologia di analisi, servirà poi a decifrare quali potranno essere i punti di forza e debolezza della rete stessa evidenziandone i nodi o gli edges – evidenziati dallo studio della centralità – che sono più importanti o fondamentali per la rete stessa.

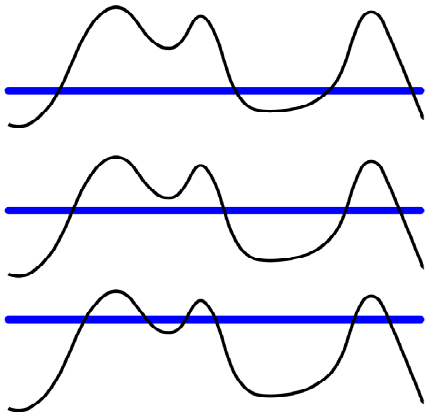


Figura 2.8: Il metodo delle Isole

CAPITOLO TERZO

IL LINGUAGGIO DI PROGRAMMAZIONE R ALL'INTERNO DELLA SOCIAL NETWORK ANALYSIS

3.1 INTRODUZIONE AL LINGUAGGIO R

Un ulteriore software che ci permette di esaminare studiare una rete sociale basandosi sui concetti forniti dalla Social Network Analysis, è dato dall'utilizzo di R e del suo relativo linguaggio di programmazione. Il linguaggio R[1], talvolta conosciuto anche con il nome di GNU S, è stato sviluppato sulle basi degli studi effettuati dai laboratori della AT&T Bell da Rick Becker, John Chambers e Allan Wilks.

GNU S è la versione "free" del linguaggio R, che permette di effettuare degli studi di analisi matematico-statistici e rappresentazione grafica; distribuito con licenza General Public License (GPL)[2], incarna in modo pieno lo spirito del movimento e della licenze Open Source[3].

Il fondatore del movimento Open Source, Richard Stallman[4], voleva evitare che "*finisse per essere utilizzato in progetti commerciali e proprietari, o in progetti eticamente discutibili*"[5].

Ciò va a dimostrare come il linguaggio R voglia discostarsi dall'utilizzo inopportuno da parte di persone come hacker e personaggi simili.

[1] http://cran.r-project.org/
[2] http://www.gnu.org/copyleft/gpl.html
[3] http://opensource.org/
[4] http://stallman.org/
[5] http://www.gnu.org/copyleft/gpl.html

In R sono possibili diverse operazioni di manipolazione dati, calcolo matematico algebrici e rappresentazioni grafiche, che si basano su:

- Possibilità di archiviare dati attraverso il salvataggio diretto nel database interno o esterno, utilizzando software quali MySQL[6], PostgreSQL[7], ecc.
- Operare con array e matrici attraverso suite di programmi
- Sistemi di strumenti atti all'analisi dei dati
- Produrre grafici sia a livello schermo e a livello cartaceo
- Capacità di produrre nuovi semplici software o personalizzare quelli già esistenti, attraverso un linguaggio di programmazione semplice nel suo utilizzo, ma completo nello stesso tempo

La dinamicità e la malleabilità del linguaggio R attraverso il suo frame work, permette di sviluppare una serie di pacchetti e set di funzioni per rispondere alla moltitudine di esigente che un utente può avere. Si vedrà in seguito, un esempio di installazione di un pacchetto di funzioni all'interno del software di programmazione,che permette la scrittura in linguaggio R. Il pacchetto denominato SNA, designato per lo studio delle rete sociali attraverso gli elementi della Social Network Analysis, è possibile trovarlo assieme in una serie di *repository*[8], assieme ad altri pacchetti tra i quali citiamo CRAN[9] e Bio-conductor[10].

Il pacchetto appena citato, SNA, è frutto degli studi e del lavoro portato a termine da Carter Tribley Butts[11]; nella documentazione allegata al pacchetto di funzioni, si evince che il professore di Irvine voleva creare "*..un insieme di strumenti per la social network analysis, che include indici a livello sia di nodo sia a livello grafico, metodi di covarianza e di distanza strutturale, di equivalenza strutturale, di modulazione p*, di generazioni di reti causali e visualizzazione 2D e 3D*"[12].

[6] http://www.mysql.it/
[7] http://www.postgresql.org/
[8] Il repository è un ambiente informatico del tipo ERP, in cui avviene la gestione dei metadati mediante tabelle di tipo relazionale; il sistema di motori di calcolo, regole e di tabelle andrà a gestire i metadati.
[9] http://cran.r-project.org/
[10] http://www.bioconductor.org/
[11] http://www.faculty.uci.edu/profile.cfm?faculty_id=5057
[12] http://cran.r-project.org/web/packages/sna/sna.pdf

All'interno di SNA troviamo una serie di funzioni che permettono diversi calcoli a livello matematico, algebrico e statistico; in esso possiamo avere strumenti che coprono un range dalla modellazione gerarchica bayesiana, la regressione logistica di network (QAP[13] e CUG tests[14]), la rappresentazione 2D e 3D delle matrici, e molte altre funzioni che permettono di definire la centralità e l'analisi della struttura della rete.

Questo pacchetto, ci permette di poter avere la possibilità di utilizzo di uno strumento matematico utile alla costruzione di modelli idealtipo, attraverso i quali uno studioso nel campo della Social Network Analysis cerca di rappresentare la realtà.

Il pacchetto SNA, rispecchia al suo interno alcune funzioni particolari, che per completezza illustrativa, sorge l'obbligo di citare.

Le funzioni principali sono:

- Rgraph(): attraverso la quale possiamo ricavare delle matrici di adiacenza casuali che rispondono ad una distribuzione Bernoulina della rete sociale, condizionata dal valore *tprob* [15] . Condizione fondamentale di questa distribuzione Bernoulina è che ogni casella della matrice sia indipendente l'una dall'altra.
- Rguman(): funzione che permette la riproduzione di reti partendo da una distribuzione casuale uniforme, limitata dal dyad census prescelto dal ricercatore

[13] Acronimo di Quadratic Assigment Procedure, attraverso il quale vengono analizzare in modo continuativo *("scramble")*le varie variabili dipendenti all'interno di un dataset. I risultai che scaturiscono da questo processo vengono successivamente immagazzinati in un ulteriore dataset riassuntivo

[14] Acronimo di Conditional Uniform Graph, funzione in linguaggio R attivabile attraverso il comando `gug.test`, permette di poter uniformare e poter gestire a livello grafico i risultati delle elaborazioni eseguite.

[15] Il valore tprob può essere interpretato nel linguaggio R in diverse versioni:

1. Se è uguale ad un numero, esso rappresenta la probabilità di poter avere un collegamento tra due diversi nodi. Tale probabilità si espande a tutte le coppie di nodi all'interno della stessa rete
2. Se contiene un numero di valori tali da costituire un vettore, ogni valore in entrata dal vettore viene considerato singolarmente e processato come al punto precedente
3. Se contiene una matrice n*n, ogni valore di n rappresenta la probabilità relativa ad uno specifico arco di collegamento tra due nodi
4. Se il valore di riferisce ad un array tridimensionale, i valori contenuti in essi saranno riferiti alla probabilità esistente di un specifico arco di collegamento tra nodi in un'altrettanto specifico rete selezionato

- Rgnm(): simile alla precedente, dalla quale discosta dal fatto che la scelta del ricercatore ricade sulla distribuzione casuale uniforme condizionata sulla densità.

Di queste funzioni, quella che andremo a utilizzare sarà solamente la prima delle tre.
Per richiamare il pacchetto nella piattaforma R, dovremmo inserire una di una serie di comandi e brevissime stringhe di programmazione.
Per caricare SNA, dovremmo inserire in seguente comando:
>library(sna)
Una volta effettuato il caricamento della libreria, procediamo a effettuare un test sulla base del libro "*Analisi delle reti sociali – Teorie, metodi, applicazioni*" a cura di Andrea Salvini, che prevede una serie di inserimenti di comandi atti a produrre una sorta di breve tour e tutorial in questo linguaggio di programmazione.

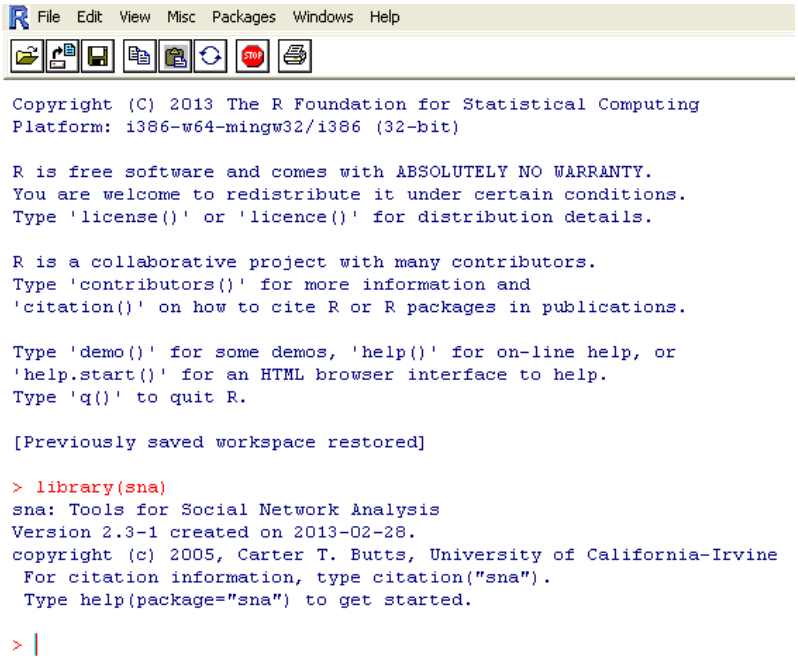


Figura 3.1 : R - caricamento libreria SNA

3.2 LE MATRICI

Anteposto il caricamento della libreria SNA, si andrà a creare una matrice fissando dei parametri preliminari, quali numero di colonne e righe ed il valore *tprob*. Salveremo a livello di programma questa matrice con un nome di fantasia, *g* in questo caso. Il comando in R che permette ciò, è riassunto in questa breve stringa.

>g<-rgraph(4,tprob=0.5)

Per visualizzare la matrice che si viene a creare, dobbiamo inserire il comando che permette di stampare a video i valori che si vengono a creare da essa. Di grandezza pari a n x n, all'interno di essa individueremo un collegamento tra nodi segnalato dal valore pari a 1, mentre nel caso opposto vedremo apparire un valore pari a 0.
Per questo utilizzeremo il seguente comando:

>print(g)

o semplicemente
>g

3.3 LE SOCIO MATRICI

>plot.sociomatrix(g)

Le sociomatrici sono elementi fondamentali della Sociometria; in particolar modo esse permettono di visualizzare a livello grafico il concetto di adiacenza tra nodi. Ad ogni riga corrispondono gli attori che creano ed inviano collegamenti ad altri attori, mentre le colonne rappresentano gli attori che ricevono i collegamenti in entrata.
Il risultato grafico del comando sarà il seguente:

```
> g<-rgraph(4,tprob=0.5)
> print(g)
     [,1] [,2] [,3] [,4]
[1,]    0    1    1    0
[2,]    1    0    0    1
[3,]    0    1    0    0
[4,]    0    0    0    0
> plot.sociomatrix(g)
```

Figura 3.2 : Stringhe in R per la stampa della socio matrice

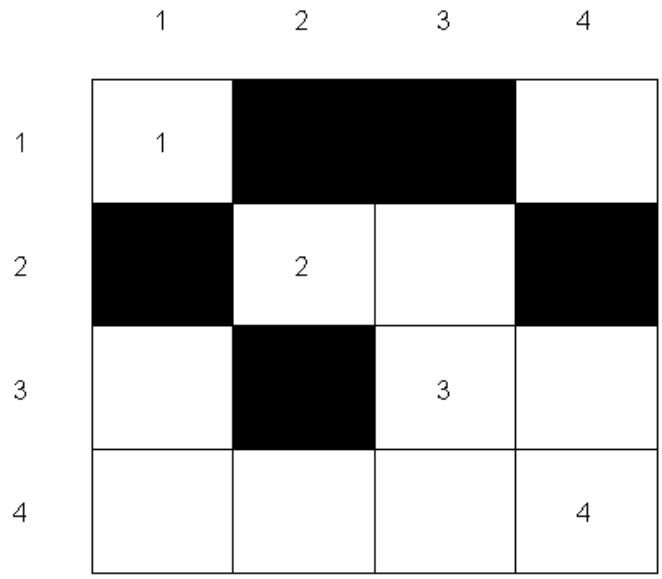


Figura 3.3 : Risultato matrice identità

Il risultato grafico della matrice mette in risalto come le celle colorate in nero evidenziano la possibile vicinanza di nodi, viceversa le celle in bianco mette in risalto i nodi lontani e che non potrebbero avere un collegamento.

Una precisazione ci sovviene fare in questo ambito nello studio delle matrici; se aumenteremo gradualmente il valore *tprob*, andremo a formare una sorta di matrice del tipo "*identità*".

Qui di seguito saranno illustrate una serie di immagini che evidenziano con valori intermedi questa affermazione.

```
> g<-rgraph(4,tprob=0.3)
> print(g)
     [,1] [,2] [,3] [,4]
[1,]    0    0    0    0
[2,]    0    0    1    0
[3,]    1    1    0    0
[4,]    1    0    1    0
```

Figura 3.3 : Stringhe in R con probabilità pari a 0.3

```
> g<-rgraph(4,tprob=0.6)
> print(g)
     [,1] [,2] [,3] [,4]
[1,]    0    0    0    1
[2,]    0    0    1    0
[3,]    0    1    0    1
[4,]    0    1    0    0
```

Figura 3.4 : Stringhe in R con probabilità pari a 0.6

```
> g<-rgraph(4,tprob=0.9)
> print(g)
     [,1] [,2] [,3] [,4]
[1,]    0    1    1    1
[2,]    1    0    1    1
[3,]    1    1    0    1
[4,]    1    1    1    0
```

Figura 3.5 : Stringhe in R con probabilità pari a 0.9

```
> g<-rgraph(15,tprob=0.6)
> print(g)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14] [,15]
 [1,]    0    1    1    0    1    1    0    0    0     0     0     0     1     1     1
 [2,]    1    0    0    0    1    0    1    1    0     0     0     1     1     1     1
 [3,]    1    0    0    1    1    0    0    1    1     0     1     1     1     1     1
 [4,]    1    0    0    0    0    1    0    0    0     0     1     1     0     1     0
 [5,]    1    0    1    1    0    0    0    1    0     1     0     0     1     1     0
 [6,]    0    0    0    1    1    0    0    0    0     1     1     1     0     1     1
 [7,]    1    1    1    1    1    0    0    1    1     0     1     1     1     0     0
 [8,]    0    0    1    1    1    1    1    0    1     0     0     0     1     1     0
 [9,]    1    0    1    0    0    1    1    0    0     1     0     1     1     1     0
[10,]    1    1    1    0    0    0    0    0    0     0     0     1     0     0     1
[11,]    1    0    0    1    0    0    0    1    1     1     0     1     1     1     0
[12,]    0    1    1    1    1    0    0    0    0     1     0     0     0     1     0
[13,]    0    0    0    0    0    1    1    1    1     1     1     1     0     1     1
[14,]    1    1    1    1    0    1    0    0    0     1     1     0     1     0     0
[15,]    1    0    1    1    0    1    1    0    1     1     0     0     1     0     0
```

Figura 3.6 : Esempio socio matrice con 15 nodi e probabilità di connessione pari a 0.6

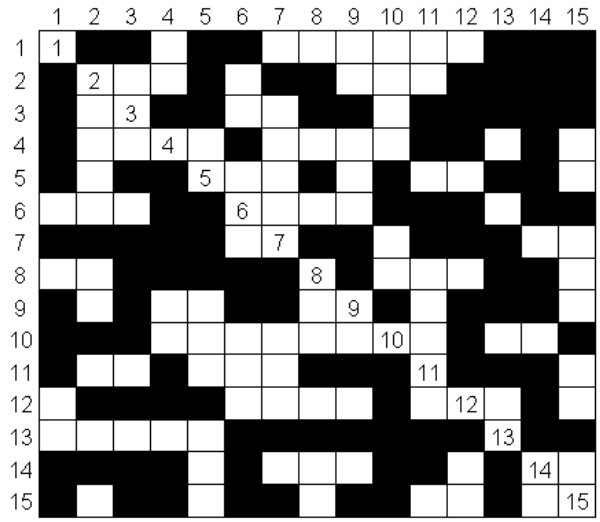


Figura 3.7 : Risultato grafico esempio precedente con il comando plot.sociomatrix

3.4 LA CONNETTIVITA'

Un'importante caratteristica che un nodo deve avere all'interno della rete sociale, che ne determina la qualità, è la connettività.

La funzione che permette di individuare questa caratteristica, è una vera e propria interrogazione da parte dell'utente all'interno di R, il quale darà una risposta di genere True e False.

Il comando dedicato a questo compito risulta essere,

>is.connected()

Di seguito le immagini del nostro esempio

```
> is.connected(g)
[1] FALSE
```

```
> g<-rgraph(4,tprob=0.5)
> is.connected(g)
[1] FALSE
>
```

Figura 3.8 : Esempio dell'interrogazione in R sulla connessione di una socio matrice

Se come nel caso precedente aumenteremo il valore *tprob* fino a portarlo ad un valore molto vicino all'unità, all'interrogazione con il comando *>is.connected()*, avremmo una risposta del tipo TRUE, con cui si afferma che tutti i nodi sono collegati fra loro

```
> g<-rgraph(4,tprob=0.9)
> print(g)
     [,1] [,2] [,3] [,4]
[1,]    0    1    1    1
[2,]    1    0    1    1
[3,]    1    1    0    0
[4,]    1    1    1    0
> is.connected(g)
[1] TRUE
```

Figura 3.9: Esempio completo sull'inserimento e interrogazione di una matrice

Da una semplice interrogazione del programma in merito alla socio matrice, possiamo avere ulteriori informazioni sugli elementi di una rete sociale, attraverso l'individuazione dei vertici isolati.

3.5 I VERTICI ISOLATI

Il comando dedicato, non esegue una mera somma dei nodi privi di collegamento, ma elenca in una semplice lista quest'ultimi. Il comando risulta essere il seguente,

>isolate(g)

Di seguito saranno presentate una serie di screen shot, con le quali si analizzerà come all'aumentare del valore *tprob*, il numero di nodi isolati si riduce.

```
> g<-rgraph(4,tprob=0.1)
> isolates(g)
[1] 1 2
> g<-rgraph(4,tprob=0.2)
> isolates(g)
[1] 1
> g<-rgraph(4,tprob=0.3)
> isolates(g)
integer(0)
```

Figura 3.10: Verifica in R dei vertici isolati con differenti probabilità

Rimanendo in tema di nodi isolati, R può essere interrogato attraverso il comando *is.isolate()*; a questa funzione sarà richiesto se il nodo x sia isolato. R risponderà in modo affermativo o meno attraverso la dicitura a video di TRUE o FALSE.
Ecco di seguito il nostro test.

```
> g<-rgraph(4,tprob=0.1)
> isolates(g)
[1] 1 2
> is.isolate(g,1)
[1] TRUE
> is.isolate(g,3)
[1] FALSE
```

Figura 3.11: Ulteriori test in R sui vertici isolati

Per effettuare la mera conta dei vertici isolati all'interno di una rete, il comando risulterà essere:

>lenght(isolate())

```
> g<-rgraph(4,tprob=0.1)
> g
     [,1] [,2] [,3] [,4]
[1,]    0    0    0    0
[2,]    0    0    0    0
[3,]    0    0    0    0
[4,]    0    0    0    0
> isolates(g)
[1] 1 2 3 4
> length(isolates(g))
[1] 4
```

Figura 3.12 : Conteggio in R dei vertici isolati

3.6 LA RAGGIUNGIBILITA'

Un concetto fondamentale per l'analisi dei una rete sociale, è la possibilità da parte di un vertice di essere raggiunto da altri nodi attraverso collegamenti diretti. Attraverso il comando *reachability()*, possiamo vedere l'assenza o la presenza di un collegamento tra nodi; nella fattispecie positiva del caso la matrice che risulterà, avere un valore pari ad 1, mentre nel caso opposto un valore pari a 0.

```
> g<-rgraph(10,tprob=0.1)
> g
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    0    0    0    0    0    0    0    0     0
 [2,]    0    0    0    0    0    0    0    0    0     0
 [3,]    0    0    0    1    0    0    0    0    0     0
 [4,]    0    0    0    0    0    1    0    0    1     0
 [5,]    0    0    0    0    0    0    0    0    0     1
 [6,]    0    0    0    0    0    0    0    0    0     0
 [7,]    1    0    0    0    0    0    0    0    0     0
 [8,]    0    0    0    0    0    1    0    0    0     1
 [9,]    0    0    0    0    0    0    0    0    0     1
[10,]    0    0    0    0    0    0    0    0    0     0
> reachability(g)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    1    0    0    0    0    0    0    0    0     0
 [2,]    0    1    0    0    0    0    0    0    0     0
 [3,]    0    0    1    1    0    1    0    0    1     1
 [4,]    0    0    0    1    0    1    0    0    1     1
 [5,]    0    0    0    0    1    0    0    0    0     1
 [6,]    0    0    0    0    0    1    0    0    0     0
 [7,]    1    0    0    0    0    0    1    0    0     0
 [8,]    0    0    0    0    0    1    0    1    0     1
 [9,]    0    0    0    0    0    0    0    0    1     1
[10,]    0    0    0    0    0    0    0    0    0     1
```

Figura 3.13 : Simulazione della raggiungibilità dei vari vertici

3.7 LA DISTANZA GEODEDICA

Per valutare in modo qualitativo una rete sociale, non dobbiamo sottovalutare un aspetto molto importante come la distanza tra un nodo ed un altro, andando ad analizzare i percorsi e la loro lunghezza. Molte volte capiterà all'interno di una rete sociale, di avere diversi percorsi che possono collegare due nodi. Quello che interessa in modo

particolare alla Social Network Analysis, è quello del percorso più breve percorribile, conosciuto anche con il termine di *shortest path* o *distanza /lunghezza geodedica*.
Strettamente collegati al concetto di *lunghezza geodedica* tra nodi, sono i concetti noti all'interno della Social Network Analysis come eccentricità e diametro.

- *Eccentricità*: caratteristica del nodo, che individua il percorso più lungo – geodedica – a cui viene legato agli altri nodi
- *Diametro*: caratterista del rete, attraverso la quale viene definita la lunghezza massima che può esistere tra due nodi.
-

Una precisazione bisogna effettuare in merito alla eccentricità e al diametro qualora una rete sociale sia del tipo orientato o non orientato.
Il comando in R che permette di rispondere a queste richieste, è dato la comando:

>geodist()

Questa funzione darà come risultato a video due matrici; la prima sarà simile a quella di output che si ricava dal comando *reachability()* – matrice *$counts* - , mentre la seconda – matrice *$gdist* – identifica la lunghezza geodedica per coppia di vertici.

```
> g<-rgraph(10,tprob=0.1)
> g
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    0    0    0    0    0    0    0    0     0
 [2,]    0    0    0    1    0    0    1    0    0     0
 [3,]    0    0    0    0    0    0    0    0    0     0
 [4,]    0    0    0    0    0    1    0    0    0     0
 [5,]    0    1    0    0    0    0    0    0    0     0
 [6,]    0    0    0    1    0    0    0    0    0     0
 [7,]    0    0    1    0    0    0    0    0    0     0
 [8,]    1    0    0    0    0    0    0    0    0     0
 [9,]    0    0    1    1    0    0    0    0    0     0
[10,]    0    0    1    0    0    1    0    0    0     0
> geodist(g)
$counts
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    1    0    0    0    0    0    0    0    0     0
 [2,]    0    1    1    1    0    1    1    0    0     0
 [3,]    0    0    1    0    0    0    0    0    0     0
 [4,]    0    0    0    1    0    1    0    0    0     0
 [5,]    0    1    1    1    1    1    1    0    0     0
 [6,]    0    0    0    1    0    1    0    0    0     0
 [7,]    0    0    1    0    0    0    1    0    0     0
 [8,]    1    0    0    0    0    0    0    1    0     0
 [9,]    0    0    1    1    0    1    0    0    1     0
[10,]    0    0    1    1    0    1    0    0    0     1

$gdist
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0  Inf  Inf  Inf  Inf  Inf  Inf  Inf  Inf   Inf
 [2,]  Inf    0    2    1  Inf    2    1  Inf  Inf   Inf
 [3,]  Inf  Inf    0  Inf  Inf  Inf  Inf  Inf  Inf   Inf
 [4,]  Inf  Inf  Inf    0  Inf    1  Inf  Inf  Inf   Inf
 [5,]  Inf    1    3    2    0    3    2  Inf  Inf   Inf
 [6,]  Inf  Inf  Inf    1  Inf    0  Inf  Inf  Inf   Inf
 [7,]  Inf  Inf    1  Inf  Inf  Inf    0  Inf  Inf   Inf
 [8,]    1  Inf  Inf  Inf  Inf  Inf  Inf    0  Inf   Inf
 [9,]  Inf  Inf    1    1  Inf    2  Inf  Inf    0   Inf
[10,]  Inf  Inf    1    2  Inf    1  Inf  Inf  Inf     0
```

Figura 3.14 : Simulazione della raggiungibilità dei vertici con uso delle funzioni $counts e $gdist

Se osserviamo la tabella $gdist, vediamo che tra i vertici [5] e [6], la distanza geodedica risulta essere pari a 3; qualora un nodo non avesse nessuna distanza geodedica, presenterebbe solamente valori nella matrice pari a `Inf`.

3.8 LA GERARCHIA

Rimanendo nel tema della connettività, oltre al fatto al concetto di efficienza, si deve introdurre un concetto inscindibile dal precedente che riguarda la gerarchia .

La gerarchia identifica il ruolo dell'attore; ruolo dell'attore visto in modo particolare rispetto agli altri attori. La gerarchia in R viene individuata attraverso il commando *hierarchy()*, dal quale ricaveremo un valore compreso tra i già noti 0 e 1.

Se la rete sociale presenta nodi con collegamenti asimmetrici, il valore che R riproduce sarà vicino a 1; nel caso opposto con valore risultato pari a 0, vedremo collegamenti simmetrici.

```
> g<-rgraph(10,tprob=0.1)
> g
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    0    0    0    0    0    0    0    0     0
 [2,]    0    0    1    0    0    0    0    0    0     0
 [3,]    1    0    0    0    1    0    0    0    0     0
 [4,]    0    0    0    0    0    0    0    0    0     0
 [5,]    0    0    0    0    0    1    0    0    0     0
 [6,]    0    0    0    0    1    0    0    0    0     1
 [7,]    0    0    0    0    1    0    0    0    0     0
 [8,]    0    0    0    0    1    0    0    0    0     0
 [9,]    0    0    0    0    0    0    1    0    0     0
[10,]    0    0    0    0    0    0    0    0    0     0
> hierarchy(g,measure="krackhardt")
[1] 0.9545455
```

Figura 3.15 : Esempio della funzione hierarchy di una socio matrice con probabilità pari ad 0.1

```
> g<-rgraph(10,tprob=0.8)
> g
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    1    1    1    1    1    1    1    0     1
 [2,]    1    0    1    0    1    1    1    1    1     1
 [3,]    1    1    0    0    1    1    0    1    1     0
 [4,]    1    1    1    0    1    1    1    0    0     0
 [5,]    1    1    1    1    0    1    0    1    1     1
 [6,]    1    1    0    0    1    0    1    1    1     1
 [7,]    1    1    1    1    1    1    0    1    1     1
 [8,]    1    1    0    1    0    1    1    0    1     1
 [9,]    1    0    1    1    1    1    1    1    0     1
[10,]    1    0    1    1    1    1    1    1    1     0
> hierarchy(g,measure="krackhardt")
[1] 0
```

Figura 3.16 : della funzione hierarchy di una socio matrice con probabilità pari ad 0.8

3.9 LA DENSITA'

Per determinare e studiare a livello qualitativo la struttura di una rete sociale, viene intrapreso andando a verificare la densità.
Una delle informazioni preliminare che dobbiamo avere per l'analisi di questo aspetto, è il computo dei possibili legami che un nodo può avere. Il conteggio di essi viene determinato da due semplici formule, che si differenziano dal fatto di essere applicate qualora i collegamenti tra nodi siano di tipo orientato o non orientato.
Per questo motivo la formula che determina questo valore, presenta al denominato un valore pari a 2, mentre nel caso di una rete orientata si vedrà l'assenza di un valore al denominatore.
Di seguito le formule:

- Rete orientata: $L = g(g-1)$
- Rete non orientata: $L = \frac{[g(g-1)]}{2}$

In R, in modo particolare, questo concetto viene riassunto da un semplice comando che ci conduce allo stesso risultato delle formule appena indicate; tutto ciò lo ritroviamo nel comando *nties()*

```
> g<-rgraph(10,tprob=0.2)
> g
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    0    0    1    1    0    1    0    0     1
 [2,]    0    0    0    0    0    0    1    0    0     0
 [3,]    0    0    0    0    0    0    0    0    0     0
 [4,]    0    0    0    0    0    0    0    0    0     0
 [5,]    0    1    1    1    0    0    0    0    0     0
 [6,]    0    0    0    0    0    0    1    0    0     0
 [7,]    0    0    0    1    1    0    0    0    0     0
 [8,]    1    0    0    0    1    0    0    0    1     0
 [9,]    0    0    0    0    0    0    0    0    0     0
[10,]    0    1    0    0    0    1    0    0    0     0
> nties(g,mode="digraph")
[1] 90
> nties(g,mode="graph")
[1] 45
```

Figura 3.17: Esempio del comando nties, sia nella modalità con rete orientata e non orientata

Come descritto in precedenza, il comando *nties(g,mode="digraph")* calcola i collegamenti orientati, dando come risultato un valore pari a 90; qualora volessimo identificare il valore dei collegamenti non orientati, dovremmo utilizzare il comando *nties(g,mode="graph")*; quest'ultimo commando riprodurrà il valore di 45, che non è altro che la metà del valore precedente.

La densità all'interno della Social Network Analysis, non è altro che il rapporto tra i collegamenti tra nodi esistenti all'interno di una rete sociale e quelli possibili.

In quanto alla valutazione della densità, il range di valori sarà compreso dagli ormai noti 0 e 1; qualora il valore fosse 0, tutti i vertici risulterebbero isolati, mentre nel caso opposto con valore pari a 1, risulterebbe che tutti i possibili collegamenti all'interno della rete sociale sono presenti.

In ambiente R, la funzione cui viene riposto questo calcolo, è definita da *gden()*.

Nell'applicazione pratica, avremo:

```
> g<-rgraph(10,tprob=0.7)
> g
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    1    1    1    1    1    1    1    1     1
 [2,]    1    0    1    1    0    1    1    1    0     0
 [3,]    1    0    0    1    1    1    0    0    1     0
 [4,]    1    1    1    0    1    0    1    1    1     0
 [5,]    1    1    0    1    0    1    1    0    1     1
 [6,]    0    0    0    0    1    0    1    1    1     0
 [7,]    1    0    0    0    0    1    0    1    0     1
 [8,]    1    1    0    1    1    1    1    0    1     1
 [9,]    1    1    1    1    0    1    0    1    0     1
[10,]    1    0    1    1    1    1    1    1    1     0
> gden(g,mode="digraph")
[1] 0.7222222
```

Figura 3.18 : Esempio funzione gden

3.10 IL GRADO

La centralità all'interno di una rete sociale trova la nella misura del grado, o *degree*, la sua massima e importante espressione.
In modo molto semplice, il valore di degree da parte dei vari nodi, viene computato come la semplice conta dei vari collegamenti che un nodo ha all'interno di una rete sociale con i restanti nodi e vertici affini.
A differenza dal solito copione nelle varie misure di valutazione delle reti sociali, in cui il valore varia tra 0 e 1, il degree ha un valore che intercorre tra lo 0 e g-1; nel caso di valore nullo, il nodo risulta essere isolato dal resto degli altri simili. Nel caso opposto, si avrà che il nodo preso in esame sarà interconnesso con tutti i nodi presenti.
All'interno del tema del degree, dobbiamo precisare che ci sono due tipologie di valutazione e conteggio dei collegamenti; in modo particolare avremo:

- *Outdegree*, attraverso la quale si avrà il conteggio dei collegamenti che scaturiscono dal nodo esaminato
- *Indegree*, analizzerà il numero dei collegamenti in entrata al nodo selezionato.

Da questa suddivisione binaria del degree, possiamo ricavare quattro tipologie di caratteristiche che identificano la natura di un nodo all'interno di una rete sociale; supportandoci con elementi di logica propri della matematica, possono definire che un nodo può essere:

- *Isolato*: se la Indegree e l'Outdegree è apri a zero
- *Trasmettitore*: in cui i nodi generano solo collegamenti e archi in uscita verso altri nodi
- *Ricevitore*: caso opposto al precedente, in cui un nodo vedrà ad essere protagonista di meri collegamenti in entrata
- *Portatore o Ordinario*: in cui un nodo si caratterizza dal fatto di essere sia trasmettitore e ricevitore all'interno di una rete sociale

Riprendendo un concetto preso in precedenza sul prestigio di un nodo, il valore di indegree rispecchia pienamente la qualità della popolarità di un nodo, visto da una diversa ottica.

All'interno del pacchetto SNA di R, la funzione dedicata a questo computo è svolta dal comando degree(), che in una applicazione pratica sarà simile a quanto segue:

```
> g<-rgraph(10,tprob=0.7)
> g
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    1    1    1    0    0    1    1    0     1
 [2,]    0    0    0    0    1    1    1    0    0     1
 [3,]    0    0    0    1    0    1    0    1    1     0
 [4,]    1    1    1    0    0    1    0    0    0     0
 [5,]    0    1    1    1    0    0    0    1    0     0
 [6,]    1    0    0    1    1    0    1    0    1     1
 [7,]    1    1    1    0    0    1    0    1    1     1
 [8,]    1    1    1    0    1    0    0    0    0     0
 [9,]    1    1    1    1    1    0    1    1    0     1
[10,]    0    1    1    0    0    0    1    0    1     0
> degree(g,gmode="digraph",cmode="freeman")
 [1] 11 11 11  9  8 10 12  9 12  9
```

Figura 3.19: Esempio calcolo degree

Da una prima analisi di un attento osservatore, la riga dei risultati che la funzione *degree()* riporta, sembra discostare da ciò che è stato sostenuto in precedenza; in precedenza si è sostenuto che il numero massimo che un vertice può avere in termine di degree è pari al numero dei nodi meno un unità – g-1.

In questo caso però, nell'applicazione di questo comando viene specificato una opzione particolare nel calcolo della *degree*; con la modalità *cmode="freeman"* viene indicato ad R di compiere la somma dei collegamenti sia di tipo indegree sia outdegree.

Per superare questo ostacolo, basta solamente sostituire al *"freeman"* nella stringa *cmode*, le parole *"indegree"* o *"outdegree"*, rispettivamente per verificare i possibili collegamenti e gli archi in entrata ed in uscita.

```
> g<-rgraph(10,tprob=0.7)
> g
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    1    1    0    0    0    1    0    0     1
 [2,]    1    0    1    1    1    1    1    1    0     0
 [3,]    1    0    0    1    1    1    0    0    0     1
 [4,]    1    0    1    0    1    0    1    1    1     1
 [5,]    1    0    1    1    0    1    0    1    1     1
 [6,]    0    1    1    1    1    0    0    0    1     1
 [7,]    1    1    1    1    1    1    0    1    1     1
 [8,]    1    0    1    1    1    1    1    0    1     1
 [9,]    0    0    1    1    1    0    0    1    0     1
[10,]    1    1    0    1    1    1    1    0    1     0
> degree(g,gmode="digraph",cmode="freeman")
 [1] 11 11 13 15 15 12 14 13 11 15
> degree(g,gmode="digraph",cmode="indegree")
 [1] 7 4 8 8 8 6 5 5 6 8
> degree(g,gmode="digraph",cmode="outdegree")
 [1] 4 7 5 7 7 6 9 8 5 7
```

Figura 3.20 : Esempio pratico del calcolo delle varie tipologie di degree

3.11 LA CENTRALITA'

Il concetto di centralità all'interno della Social Network Analysis, ricopre un altro grande tassello di questa scienza moderna.

All'interno di una rete sociale, la centralità di un nodo può essere vista sotto diverse ottiche:

- Quanto un attore è vicino agli altri attori
- La capacità di comunicare il più velocemente possibile tra attori, oltrepassando l'uso di nodi e collegamenti intermedi
- Avere delle distanze geodetiche molto brevi
-

Il range di valori che possiamo incontrare nel calcolo della centralità varia dallo 0, qualora un vertice evidenzia la impossibilità di essere raggiunto da un altro vertice; in caso contrario, dal valore che scaturisce dalla formula $(g-1)^{-1}$, dove il nodo preso in esame è collegato con il resto degli altri attori e in modo ravvicinato.

Un esempio grafico che esprime al meglio la centralità in modo semplice ed intuitivo, viene fornito dal rete a stella di una rete sociale.

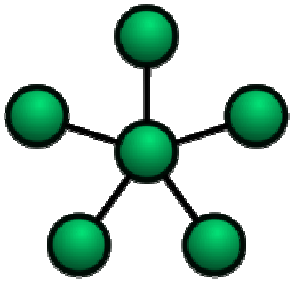


Figura 3.21 : Modello di centralità di un nodo

Alla funzione in R, *closeness()*, il dovuto compito di esaminare la centralità di un nodo all'interno di una rete sociale.

```
> g<-rgraph(10,tprob=0.7)
> g
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    1    1    1    0    1    0    0    0     0
 [2,]    0    0    1    1    0    1    1    0    1     1
 [3,]    1    1    0    1    0    1    1    1    1     0
 [4,]    0    1    1    0    1    1    0    1    1     1
 [5,]    1    1    1    1    0    1    1    1    1     1
 [6,]    1    0    1    0    1    0    1    1    0     1
 [7,]    1    1    0    1    1    0    0    1    1     1
 [8,]    0    1    1    0    1    1    1    0    1     0
 [9,]    1    1    1    0    1    1    0    1    0     1
[10,]    1    1    0    1    0    0    0    0    1     0
> closeness(g,gmode="digraph")
 [1] 0.6428571 0.7500000 0.8181818 0.8181818 1.0000000 0.7500000 0.8181818
 [8] 0.7500000 0.8181818 0.6428571
```

Figura 3.22 : Esempio del calcolo della centralità

3.12 LA BETWEENNEESS

Rimanendo nell'ambito della centralità, dobbiamo precisare come all'interno di una rete sociale, un attore può trovarsi al centro di numerosi collegamenti fra attori; attraverso il concetto della betweenness, viene analizzato come un nodo si in posizione di intermediario fra un nodo e un altro nodo.
Il concetto di betweenness, è stato fino dai primi momenti dello studio delle relazioni sociali un tema molto presente nelle ricerche; l'importanza di questi concetto non è mai stata messa in dubbi dai vari ricercatori anche quando questo concetto non si riusciva a determinare in modo quantitativo.
Questo ostacolo sulla determinazione della misura fu tuttavia superato con gli studi portati a termine da Robert Freeman[16] e Willie Eugene Pitts che si basava sulle distanze geodetiche all'interno di una rete sociale.

3.13 LA VISUALIZZAZIONE GRAFICA IN LINGUAGGIO R

Fin a questo punto dell'esposizione abbiamo elencato le caratteriste pratiche di R affiancate alla teoria, attraverso l'esposizione di formule e matrici.
L'ultimo step per concludere, vede l'analisi della visualizzazione grafica che è all'interno del pacchetto SNA; ne risulta di fondamentale importanza per le caratteristiche morfologiche che una rete sociale potrebbe avere.
L'ambito grafico all'interno della Social Network Analysis ha avuto una serie di sviluppi, che si sono susseguiti con le evoluzioni apportate dai ricercatori in questa scienza; di seguito ne vediamo in sintesi i principali.

- *Anni '30*: produzione manuale di grafi, dove l'aspetto che si risaltava era l'abilità grafica del disegnatore
- *Anni '50*: uso delle prime procedure computazionali, attraverso le quali il ricercatore veniva aiutato a decidere la miglior dislocazione possibile dei nodi all'interno di un rete

[16] http://it.wikipedia.org/wiki/Robert_Edward_Freeman

- *Anni ’70*: uso della grafica computerizzata, con l’avvenuta dei primi software che permettevano oltre ala dislocazione dei nodi, della creazione tra nodi e stampa del risultato finale
- *Anni ’80*: uso del personal computer, dove viene introdotta la visualizzazione a video delle reti
- *Anni ’90*: uso del World Wide Web, che viene visto oltre come oggetto di studio, anche come mezzo di condivisione tra gli utenti delle rappresentazioni grafiche.

Quello che andremo ora a citare sono le funzioni principali in cui R offre la possibilità di visualizzare a livello grafico le distribuzione dei nodi e dei collegamenti. La funzione più semplice ed elementare deputata a questo compito, viene attribuito al comando *gplot().*
All’interno di questa funzione possono essere specificati anche le coordinate dei vari vertici o applicare degli algoritmi che permettono di effettuare il posizionamento automatico degli elementi di una rete sociale.
Il comando che andremo ad applicare sarà il seguente,

>gplot(g,displaylabels=TRUE)

Nelle due immagini seguenti verranno evidenziati i calcoli a livello di matrice e la relativa immagine della probabile distribuzione dei nodi e dei collegamenti.

```
> g<-rgraph(10,tprob=0.7)
> g
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    1    1    1    0    1    0    0    0     0
 [2,]    0    0    1    1    0    1    1    0    1     1
 [3,]    1    1    0    1    0    1    1    1    1     0
 [4,]    0    1    1    0    1    1    0    1    1     1
 [5,]    1    1    1    1    0    1    1    1    1     1
 [6,]    1    0    1    0    1    0    1    1    0     1
 [7,]    1    1    0    1    1    0    0    1    1     1
 [8,]    0    1    1    0    1    1    1    0    1     0
 [9,]    1    1    1    0    1    1    0    1    0     1
[10,]    1    1    0    1    0    0    0    0    1     0
> gplot(g,displaylabels=TRUE)
Warning message:
In rep(label.lty, length = n) : 'x' is NULL so the result will be NULL
```

Figura 3.27 : Esempio calcolo matrice relativo al primo step a livello grafico

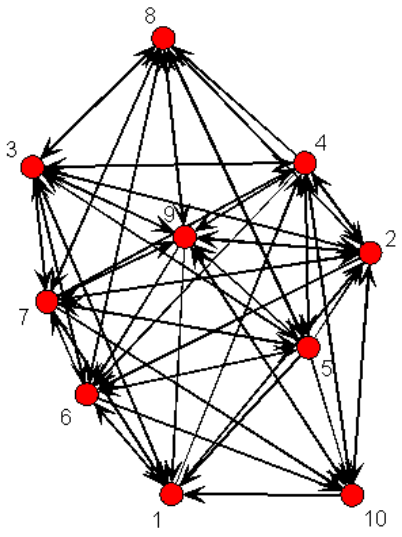


Figura 3.28 : Esempio primo risultato grafico

Oltre a questa tipologia di visualizzazione, R permette di avere una rappresentazione grafica in tre dimensioni; per ottenere a video questo risultato il pacchetto di funzioni contenute nella libreria SNA non è sufficiente. Dallo stessa pagine web che fornisce tutte le varie librerie per R, dovremmo scaricare la libreria denominata RGL.

Installata quest'ultima libreria non rimane che applicare il comando dedicato *,gplot3d(),* alla visualizzazione in tre dimensioni della rete sociale esaminata.

>gplot3d(g,displaylabels=TRUE)

```
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    0    1    1    1    0    1    0    0    0     0
 [2,]    0    0    1    1    0    1    1    0    1     1
 [3,]    1    1    0    1    0    1    1    1    1     0
 [4,]    0    1    1    0    1    1    0    1    1     1
 [5,]    1    1    1    1    0    1    1    1    1     1
 [6,]    1    0    1    0    1    0    1    1    0     1
 [7,]    1    1    0    1    1    0    0    1    1     1
 [8,]    0    1    1    0    1    1    1    0    1     0
 [9,]    1    1    1    0    1    1    0    1    0     1
[10,]    1    1    0    1    0    0    0    0    1     0
> gplot3d(g,displaylabels=TRUE)
Loading required package: rgl
```

Figura 3.29 : Esempio calcolo matrice relativo al step livello grafico in 3D

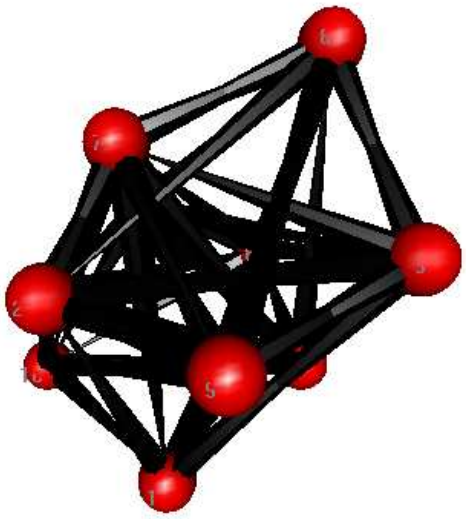


Figura 3.30 : Esempio visualizzazione grafica in 3D

CAPITOLO QUARTO

NODEXL E GEPHI

4.1 INTRODUZIONE NODE XL

All'interno di internet e della rete web, possiamo trovare diversi strumenti che possono essere d'aiuto all'analisi delle reti sociali.

Fra i vari strumenti, in questo capitolo vogliamo portare a conoscenza una plug-in di Microsoft Excel®, basata sul modello open source, conosciuta come NODE XL[1]

Attraverso NODE XL, possiamo analizzare le caratteristiche interne ed esterne della rete; con questa plug-in di Microsoft Excel®, possiamo studiare tutto quello che riguarda i flussi di mail e conversazioni attraverso gli Instant Messanger (IM)[2], cinguettii e followers in Twitter®[3], fan, amici e gli Like in Facebook®[4], followers e visite in Youtube®[5].

NODE XL, infine permette di analizzare in modo sintetico e in pochi step il traffico della rete nei siti web, soffermando le attenzioni sulle recensioni, valutazioni, tags, commenti e newsletter.

Questo flusso informativo si basa infine sulle definizione delle 3 R's, acronimo che definisce il flusso informativo sulla base alla:

- Rappresentanza
- Responsabilità
- Rispetto

[1] http://nodexl.codeplex.com/
[2] http://it.wikipedia.org/wiki/Messaggistica_istantanea
[3] https://twitter.com/
[4] https://www.facebook.com/
[5] http://www.youtube.com/?hl=it&gl=IT

Figura 4.1: Screen shot principali Social Network

4.2 PRINCIPALI FUNZIONALITA' DI NODEXL

In merito alle funzionalità di NODE XL (http://nodexl.codeplex.com/releases/view/83654), possiamo elencarle brevemente in pochi punti, evidenziando le principali potenzialità e analisi di questo strumento:

- Vertici: corrispondenti ai nodi di un grafo e rappresentai dalla persone della rete
- Edges: definiti come collegamento o relazioni tra i vari vertici
- Gruppi: rappresentati da un gruppo di vertici
- Tipologia di grafico: se di natura orientata o non orientata
- Graph Layouts e Graph Refresh: attraverso i quali possiamo visualizzare a livello grafico la struttura e I progressi effettuati dopo le serie di calcoli

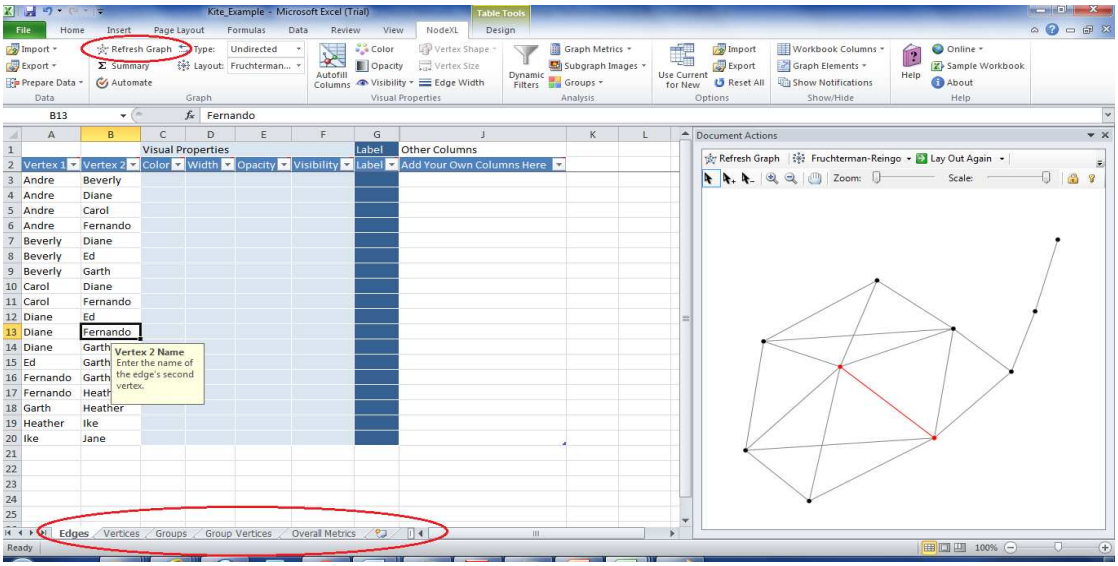


Figura 4.1: Esempio schermata iniziale NodeXL

NODE XL, permette inoltre tra le sue varie possibilità, di analizzare a livello grafico la tipologia di collegamento che si vuole studiare; le opzioni di questa plug-in prevedono l'analisi grafici sia a livello diretto sia indiretto.

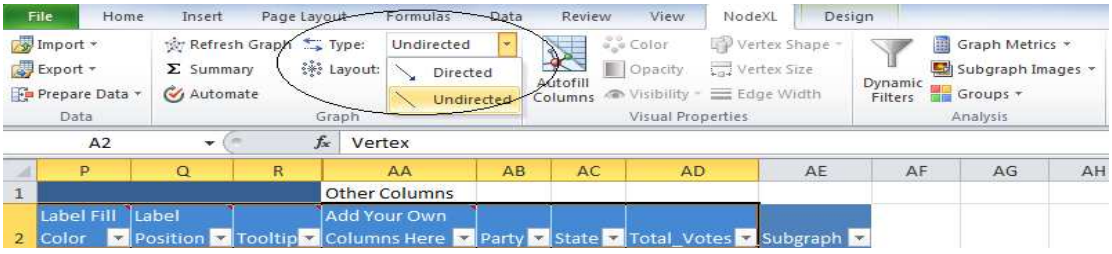


Figura 4.2: Selezione della tipologia dei collegamenti all'interno di una rete

La plug-in di Microsoft Excel®, permette inoltre una vasta gamma di calcoli e rappresentazioni a seconda delle esigenze dell'utente. Seguendo il tutorial messo a disposizione in rete, vedremo che sarà possibile analizzare la stessa rete, modificando alcuni parametri. Questi parametri permetteranno di verificare la forza dei legami e i collegamenti tra i vari nodi. Le screen shot qui sotto riportate saranno d'ausilio alla nostra spiegazione e al nostro test personale.

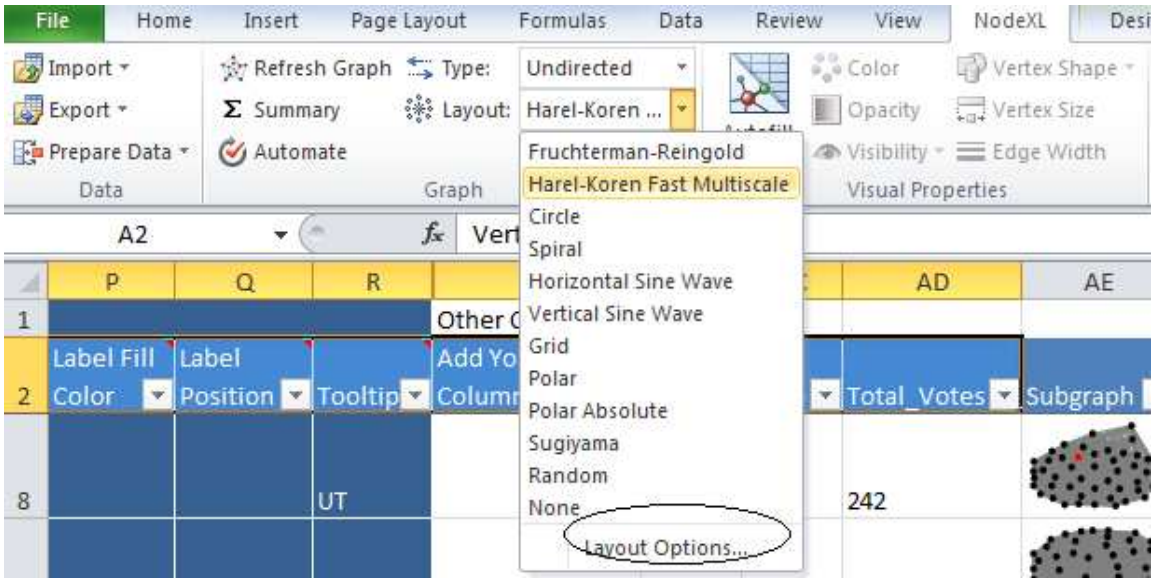


Figura 4.3: Setting grafico

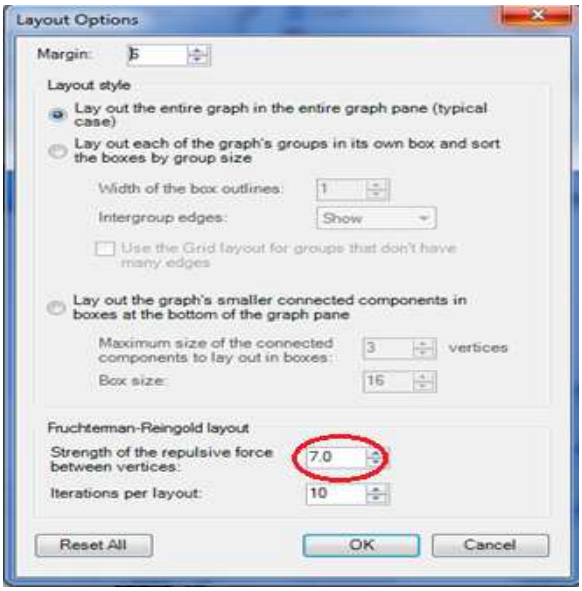


Figura 4.4: Setting a livello di layout

In NODE XL si evidenzia la possibilità di utilizzare un semplice processo di clustering dei dati contenuti nella rete sociale; essa permetterà un sommaria divisione dei contatti o edges in gruppi appartenenti al medesimo posto di lavoro, gruppo di amici o a compagni di squadra o di sport.

Nelle screen shot sotto stanti vedremo come attraverso la nostra plugin è possibile selezione e applicare questo processo di clustering e filtro.

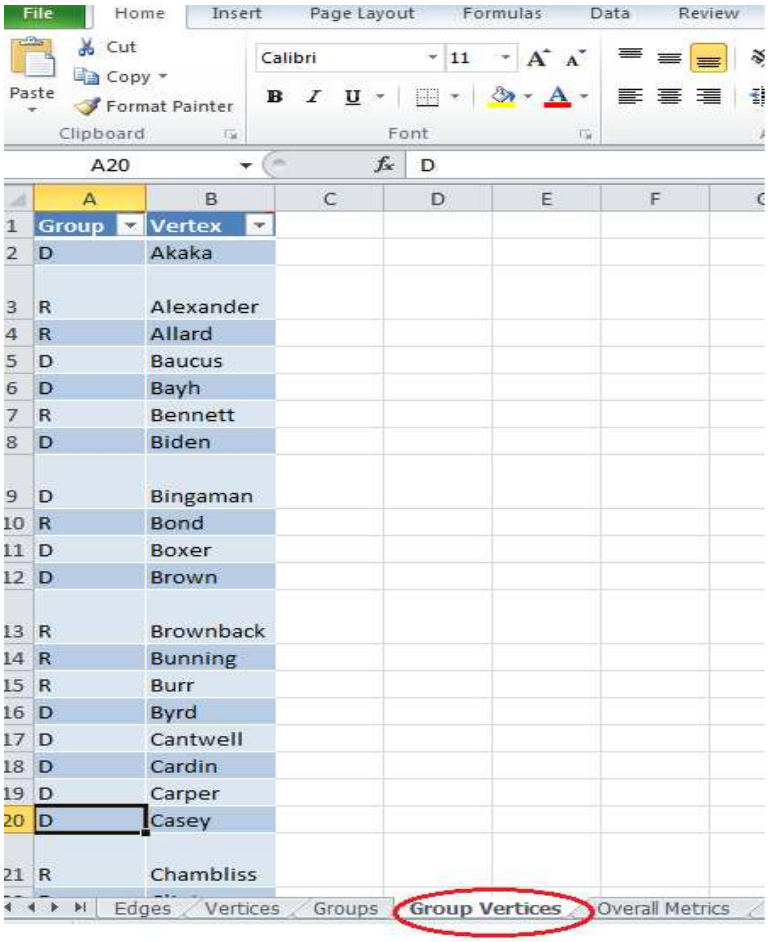


Figura 4.5: Raggruppamento per gruppi di vertici

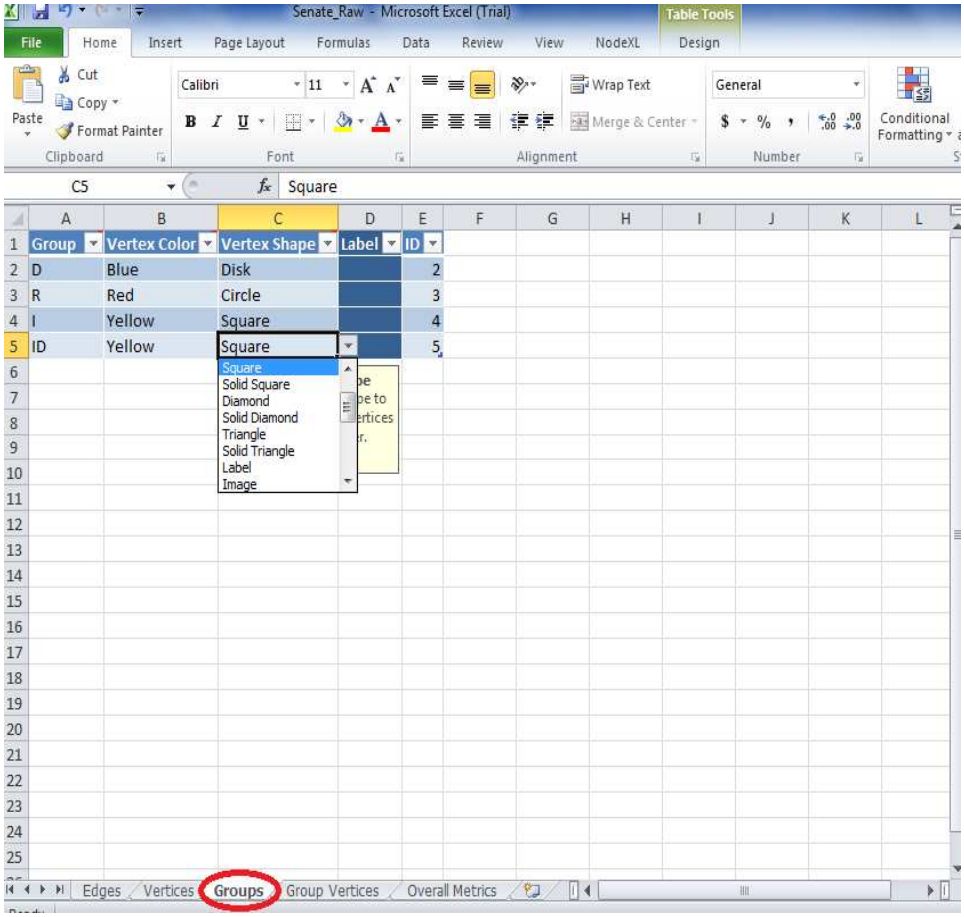


Figura 4.6: Raggruppamenti per gruppi

Le qualità di NODE XL non si fermano solo alla mera rappresentazione grafica; attraverso essa è possibile compiere una serie di calcoli che conducono alle misurazioni conosciute all'interno della Social Network Analysis; vedremo infatti che la plug-in Microsoft Excel® permette:

- Calcolo metrico e grafico: attraverso la Degree in entrata ed in uscita (eng.: in & out), la Betweenness, vicinanza ai nodi centrali (eng: Closeness Centrality) ecc.
- Generare dei sub-grafici per interpretare in modo migliore sessioni o sezioni della rete sociale.
- Possibilità di applicare diverse combinazioni di filtri che permettono di rispondere alle esigenze dell'utente.

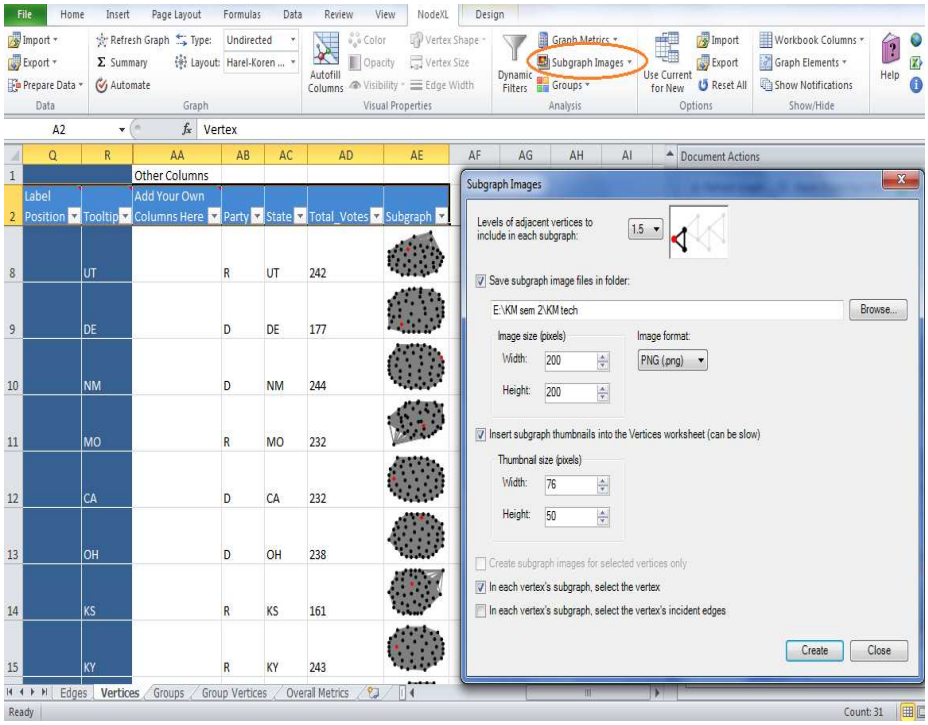


Figura 4.7: Opzioni di visualizzazione a livello di subgrafo

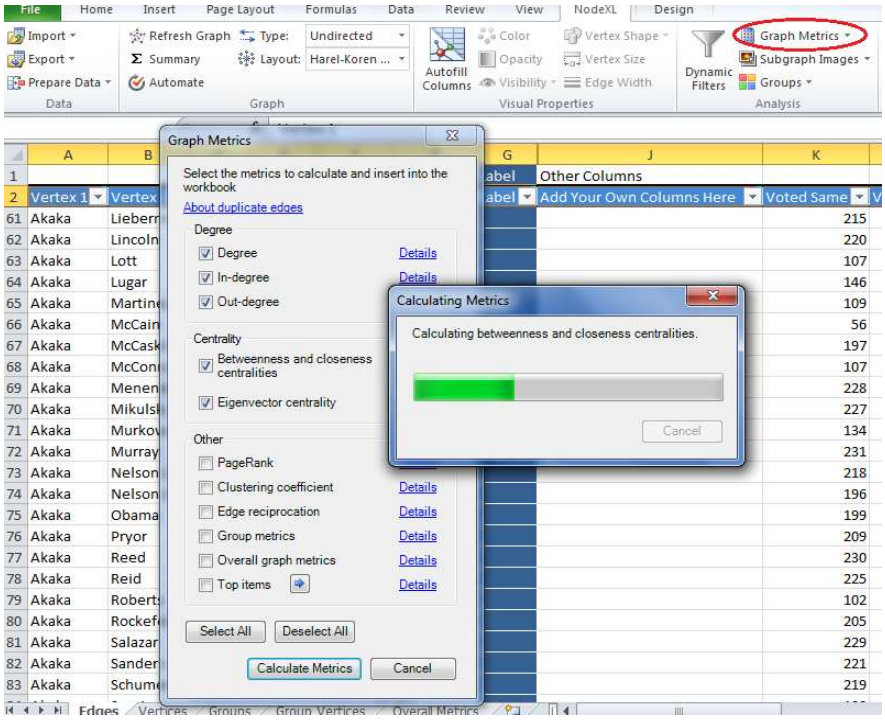


Figura 4.8: Calcolo principali misure della Social Network Analysis

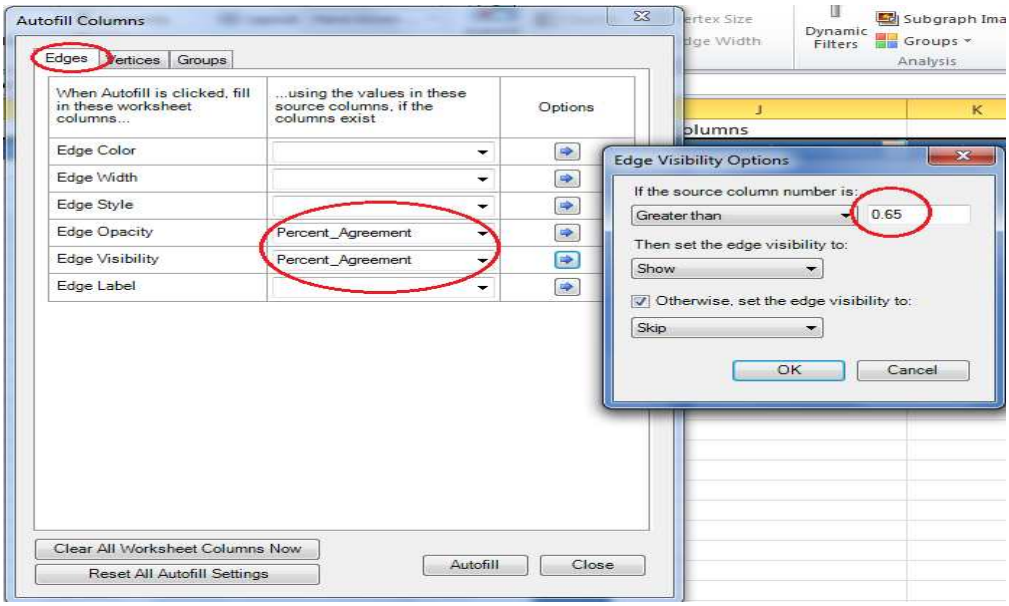


Figura 4.9 : Selezione opzioni visualizzazione dei nodi

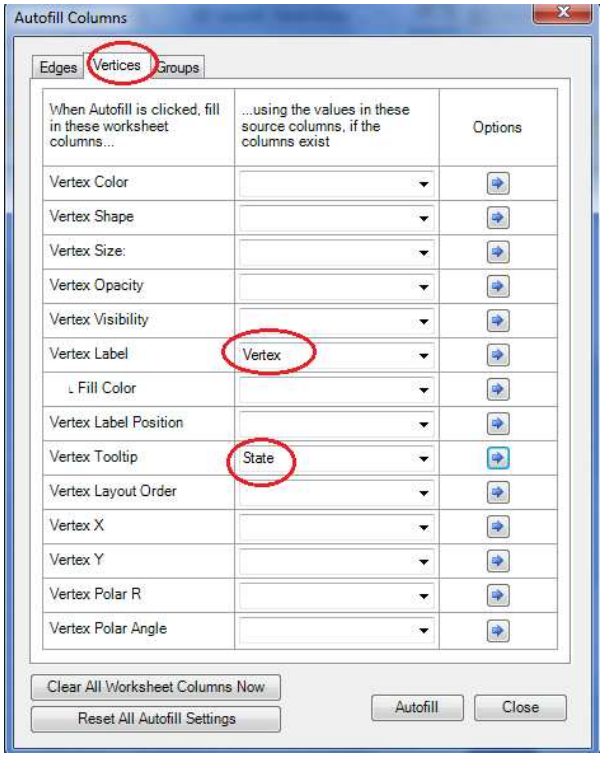


Figura 4.10: Selezione opzioni visualizzazione dei vertici

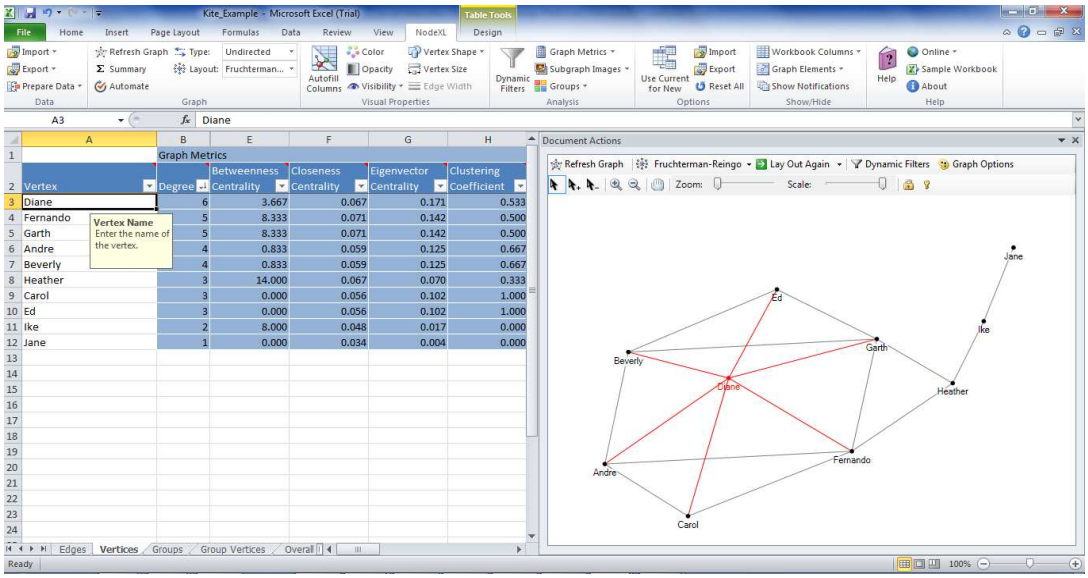


Figura 4.11: Esempio grafico - Kite net graph

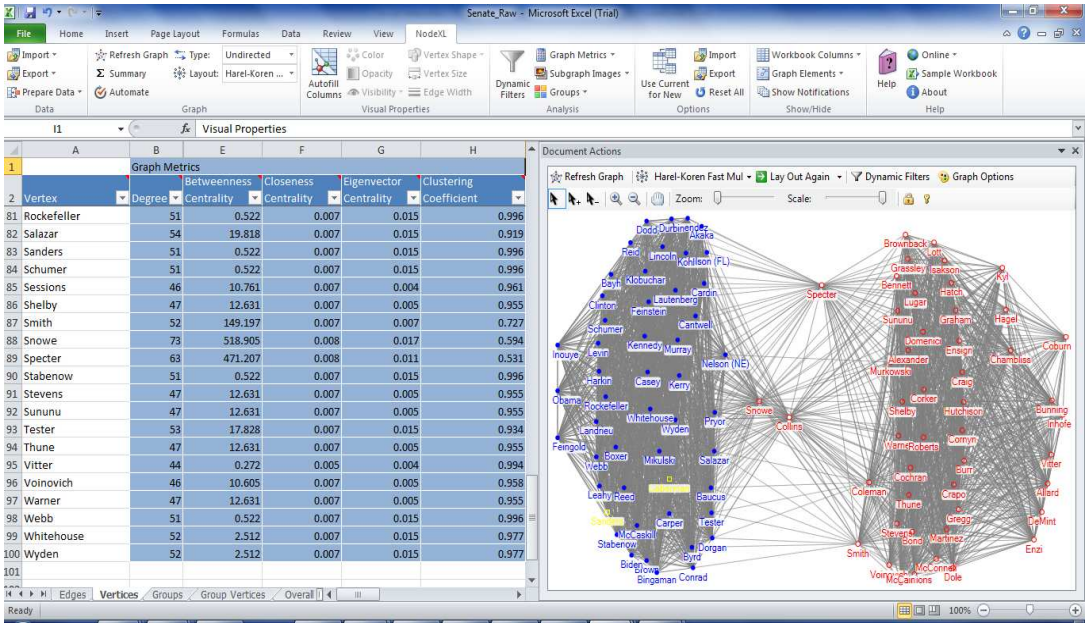


Figura 12 : Esempio grafico

4.3 L'APPLICAZIONE NAMEGENWEB

A supporto dell'analisi all'interno di NodeXL – e come si vedrà all'interno del software applicativo Gephi -, ricorreremo all'utilizzo di un'applicazione chiamata Namegenweb [6] .all'interno del social network Facebook® Questa applicazione, attraverso dei semplici step, riesce a creare in un file unico con tutte le informazione necessarie

Di seguito qui riportate, verranno illustrati i vari passi per la creazione del file unico attraverso una serie di screen-shot;

Figura 4.13 : NameGenWeb - Screen shot 1

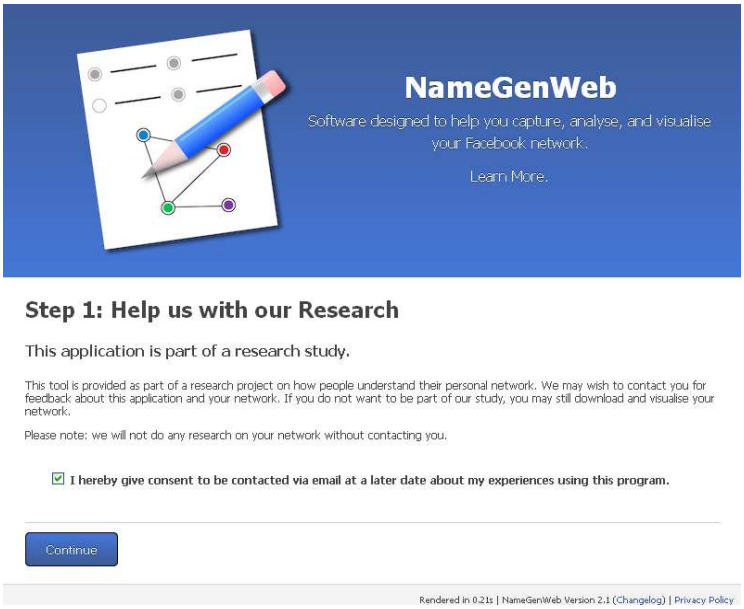


Figura 4.14 : NameGenWeb - Screen shot 2

[6] http://namegen.oii.ox.ac.uk/fb/

Figura 15 : NameGenWeb - Screen shot 3

Figura 16 : NameGenWeb - Screen shot 4

Figura 4.17 : NameGenWeb - Screen shot 5

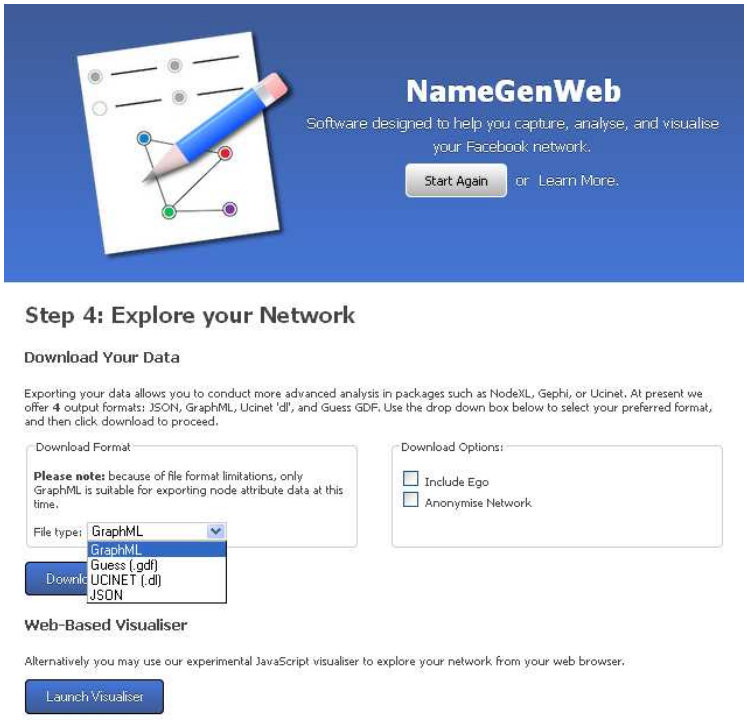


Figura 4.18 : NameGenWeb - Screen shot 6

Nella ultima slide è possibile, lanciare un'applicazione web per visualizzare la morfologia della rete che stiamo per analizzare; l'applicazione è stata creata e gestita dal team formato dall'Oxford Internet Institute[7] dell'omonima università britannica e dallo JISC – acronimo di Joint Information Systems Committee[8].

Di seguito qui riportate le screen-shot conclusive di questo capitolo, dove si vuole evidenziare le potenzialità di questa applicazione.

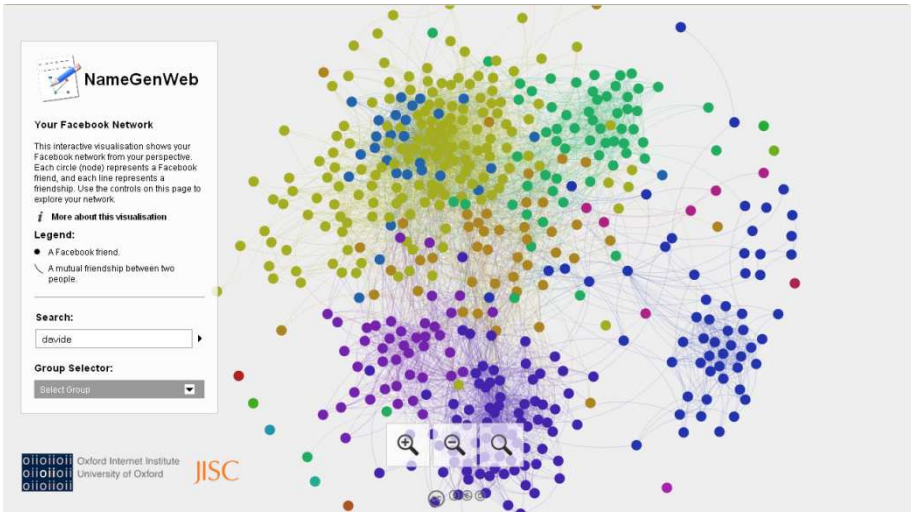


Figura 4.19 : NameGenWeb - Screen shot 7

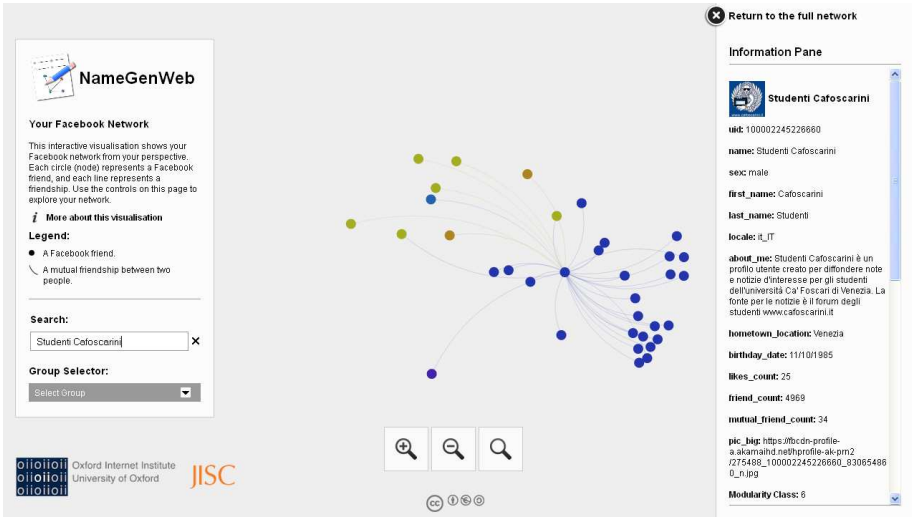


Figura 4.20 : NameGenWeb - Screen shot 8

[7] http://www.oii.ox.ac.uk/
[8] http://www.jisc.ac.uk/

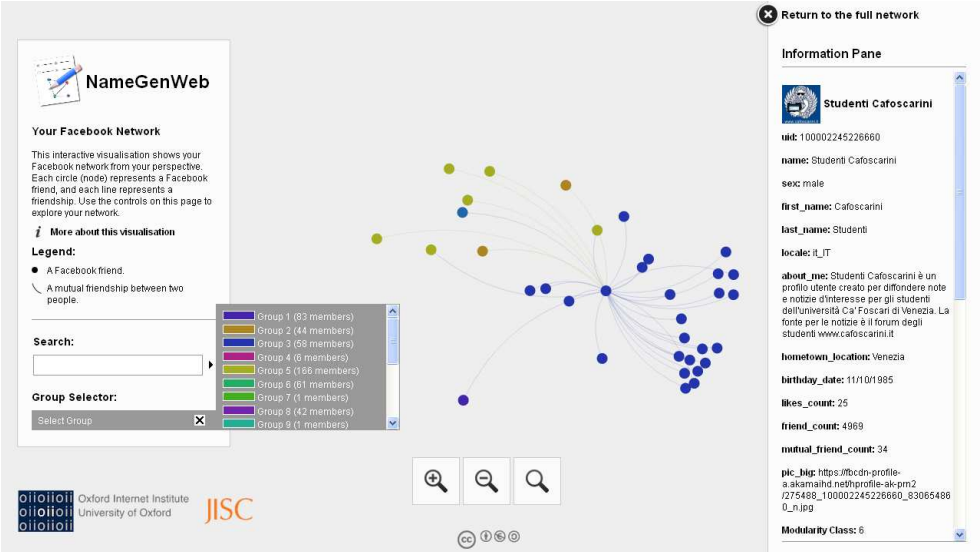


Figura4.21 : NameGenWeb - Screen shot 9

4.4 INTRODUZIONE A GEPHI

Un ulteriore software messo a disposizione dalle rete per valutare a livello grafico e a livello qualitativo una social network, è dato dal programma Gephi[9].

Gephi appartiene, come nel caso della plug-ing di Microsoft Execel® NodeXL, alla grande famiglia dei software open source; scaricabile dal sito permette un'analisi più complessa e completa rispetto alla plug-ing esaminata.

Come per NodeXL, Gephi necessita di un file unico tutte le informazioni necessarie alla elaborazione e valutazione della rete. Gephi permette di elaborare su una vasta gamma di estensioni di file; tra le principali vogliamo citare quelle relative a GEXF[10], GraphML[11], Pajek NET[12], GDF[13], GML[14], Tulip TLP[15], CSV e Compressed ZIP.

[9] https://gephi.org/
[10] http://gexf.net/format/
[11] http://graphml.graphdrawing.org/
[12] https://gephi.org/users/supported-graph-formats/pajek-net-format/
[13] https://gephi.org/users/supported-graph-formats/gdf-format/
[14] https://gephi.org/users/supported-graph-formats/gml-format/
[15] http://tulip.labri.fr/TulipDrupal/?q=tlp-file-format

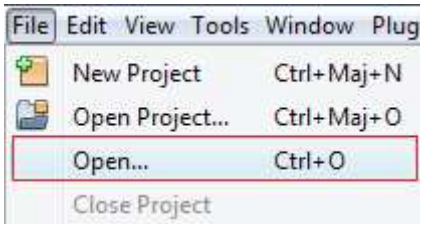


Figura 4.23: Importazione file dal menù a tendina

Una volta selezionato il file si aprirà una finestra che elencherà le caratteristiche principali della nostra rete, in particolar modo verranno resi visibili i numeri dei nodi e dei collegamenti; nella stessa schermata saranno poi possibili le selezioni riguardanti la tipologia dei tipi di collegamenti, se diretti o indiretti.

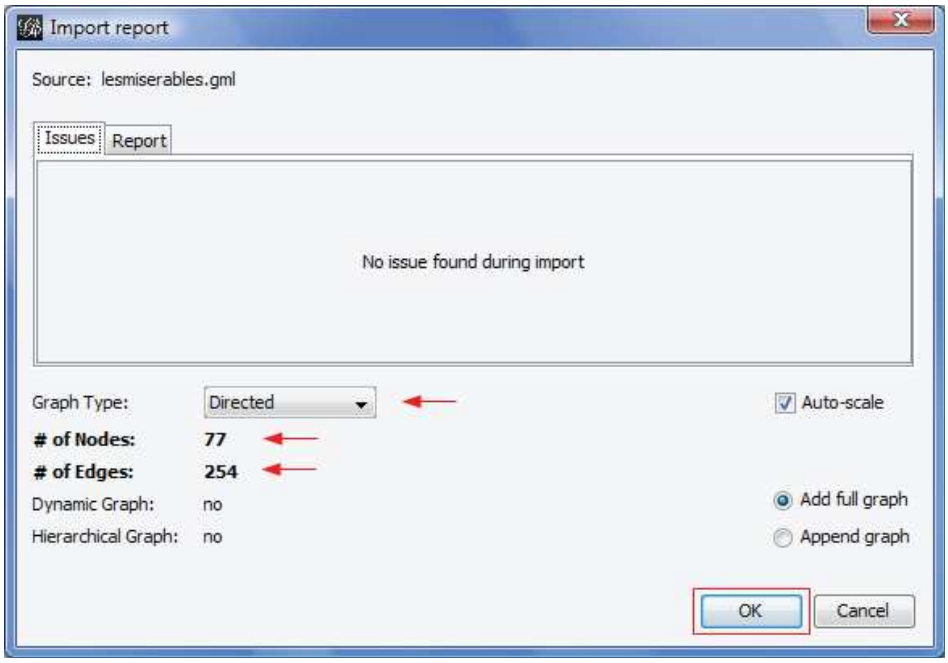


Figura 4.24: Specifiche nell'importazione dei dati

Una volta aperto il file, Gephi , provvederà ad elaborare una prima rappresentazione grafica.

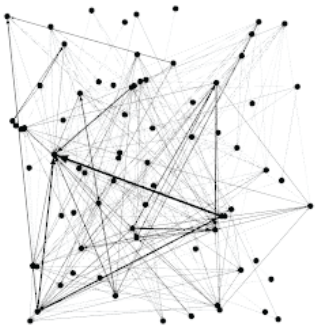


Figura 4.25: Prima visualizzazione in Gephi della rete

4.5 PRINCIPALI FUNZIONALITA' DI GEPHI

Qualora servisse, il programma permette di poter visionare in modo più dettagliato la struttura dei nodi e dei collegamenti a lui connessi, attraverso la funzione zoom.

Un primo passo in Gephi per visualizzare le sembianze di come una rete sociale realmente formata, viene fornita da una serie di funzioni a seconda delle esigenze di studio da parte dell'utente

Contenuta nel menù Layout, la funzione principale che risponde a questo scopo, sarà quella riconducibile all'applicazione della "*Force Atlas*[16]""

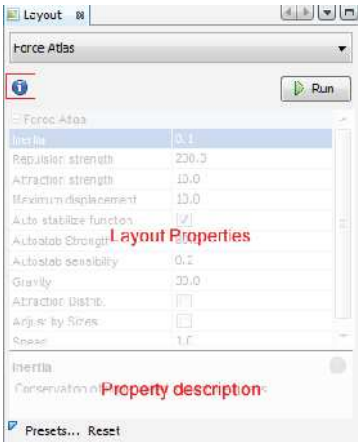


Figura 22.26: Applicazione della funzione Force Atlas

[16] Funzione di Gephi utilizzata per l'interpretazione dei dati che formano un network; alla base di questa funzione viene posto come ogni singolo nodo sia considerato come un magnete. Ogni singolo nodo sarà respinto e respingerà gli altri fino a formare una serie di agglomerati di nodi e gruppi di interesse comune. I collegamenti tra i vari nodi saranno così allungati al massimo in base anche alla forza di repulsione che l'utente imposta nelle funzione Force Atlas

Per rendere al meglio la rappresentazione grafica, viene consigliato di aumentare il coefficiente che determina la forza del collegamento; nel caso qui esaminato in Gephi , il coefficiente individuabile alla casella Repulsion Strenth, viene posto ad un valore consigliato pari a 1000.00.

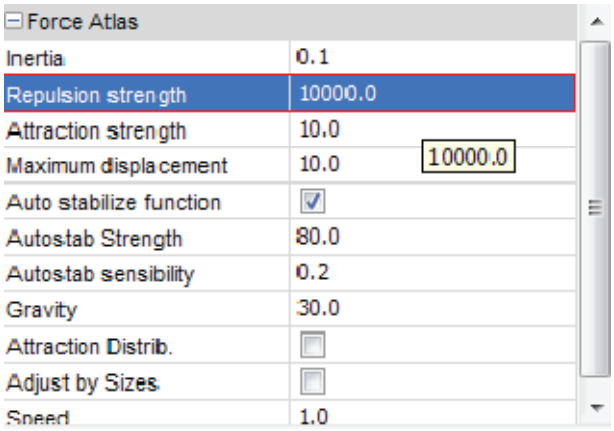


Figura 4.27: Impostazione del valore Repulsion Strenth

L'esempio qui sotto riportato, evidenzia il risultato del processo effettuato dalla funzione; si evincono fin da subito la natura di alcuni nodi, come la particolarità di essere membri di un sottografo o di essere da bridge fra una sessione e l'altra della rete sociale.

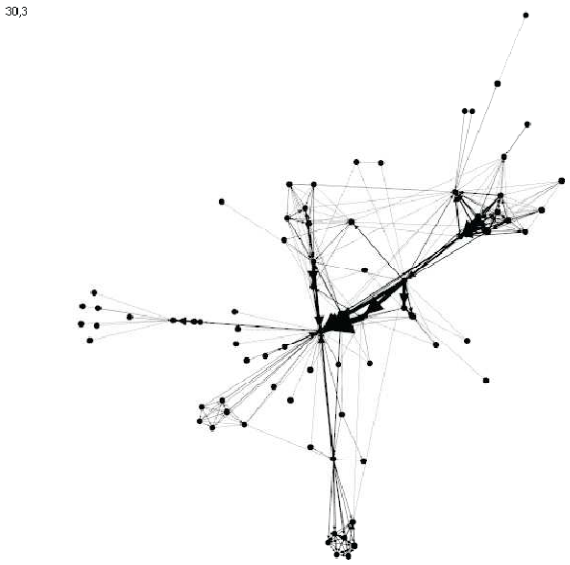


Figura 4.28: Rappresentazione grafica della rete dopo l'applicazione della Repulsion Strenght

Gephi permette di effettuare diverse modifiche a livello grafico per evidenziare l'importanza e la crucialità dei collegamenti e dei nodi. Dal menù Ranking, vedremo infatti che possiamo scegliere tra le varie opzioni, quella riferita alla Degree dei nodi; attraverso questa funzione possiamo scegliere la colorazione e la grandezza dei vari nodi.

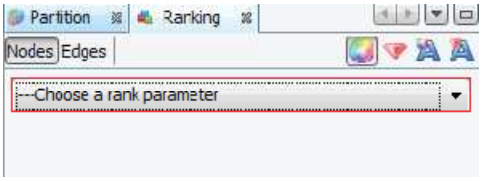


Figura 4.29: Scelta del parametri per il ranking

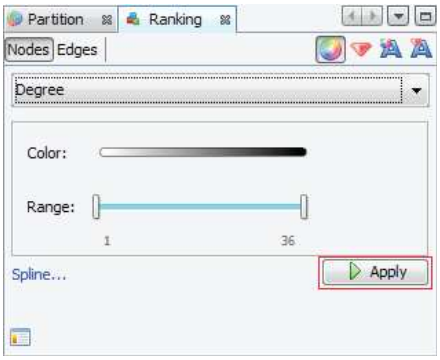


Figura 4.30: Scelta del parametro prescelto

Rimanendo nello stesso menù è possibile visualizzare un breve rapporto che indica il ranking dei vari nodi attributo dalla funzione Degree.

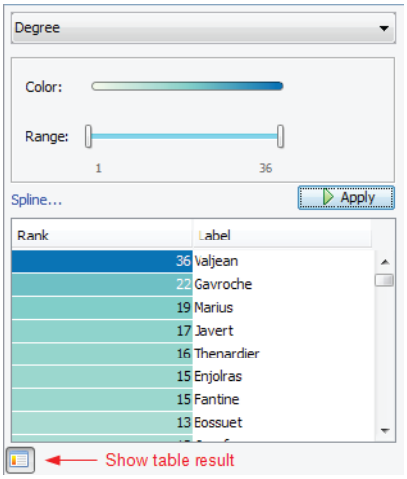


Figura 4.31: Prima rappresentazione del risultato del Ranking

4.6 GEPHIE GLI ELEMENTI DELLA SOCIAL NETWORK ANALYSIS

Le potenzialità di questo software, possono essere viste analizzando le funzioni disponibili dal menù Statistics; Gephi, provvederà automaticamente al calcolo di tutte le metriche secondo le esigenze informative che il cliente vuole estrarre dalla rete sociale che si presta ad analizzare.
In particolar modo, le funzioni che possiamo rintracciare in Gephi sono:

- Diametro
- Lunghezza media del percorso
- Coefficiente di clustering
- Page Rank
- HITS
- Centralità betweeness
- Centralità closeness
- Eccentricità
- Rilevazione di una comunità (attraverso il concetto e funzione Modularità)

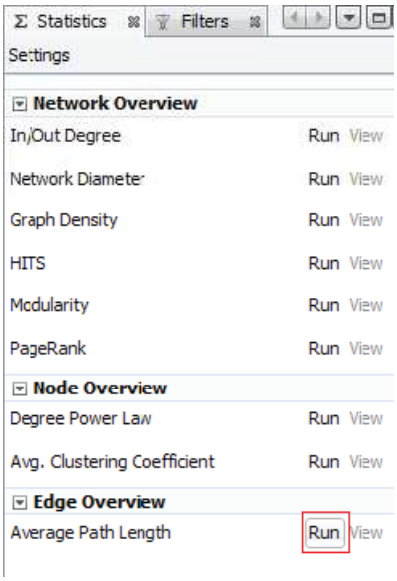


Figura 4.32: Tutte le varie opzioni di studi in Gephi attivabili attraverso il pulsante Run

Nello screen-shot qui di seguito viene rappresentato l'interfaccia che l'utente si troverà al momento del calcolo della densità.

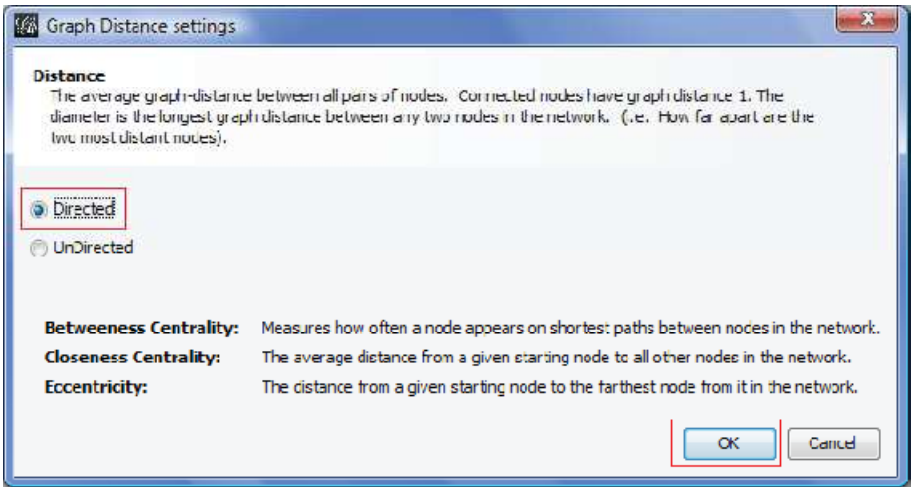


Figura 4.33: Selezione dell'opzione dei collegamenti diretti o indiretti tra nodi

Alla fine dell'algoritmo, l'utente verrà informato attraverso una serie di grafici dei risultati finali dell'analisi e della costruzione della densità all'interno di una rete sociale.

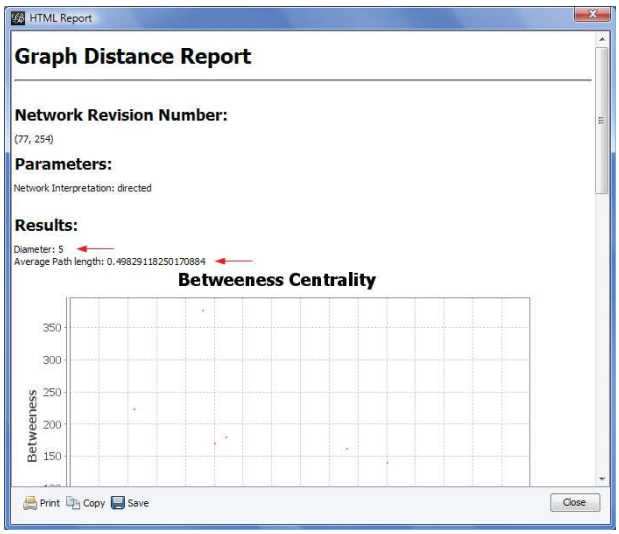


Figura 4.34: Report sulla distanza grafica

Incentrandosi sulle possibilità di personalizzazioni grafiche, dal menù Ranking possiamo cliccare sull'icona a forma di diamante. Da essa sarà possibile da parte dell'utente di rendere evidenti e visibili, i nodi più d'importanza strategica.

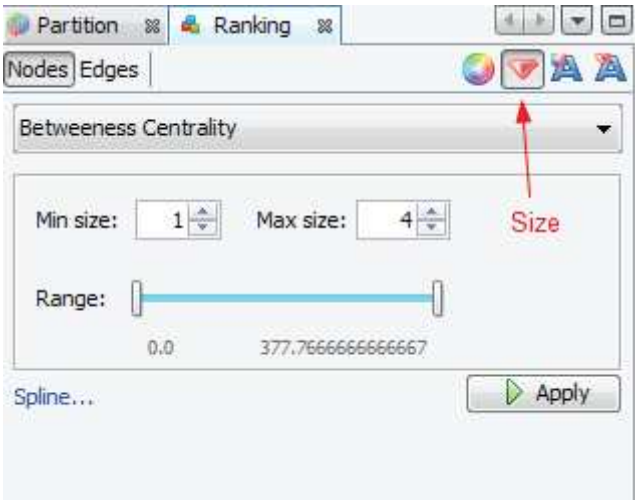


Figura 4.35: Selezione e applicazione del Ranking a livello grafico

Cliccando su Apple, una volta finite le modiche grafiche, il risultato sarà simile a quello di seguito riportato.

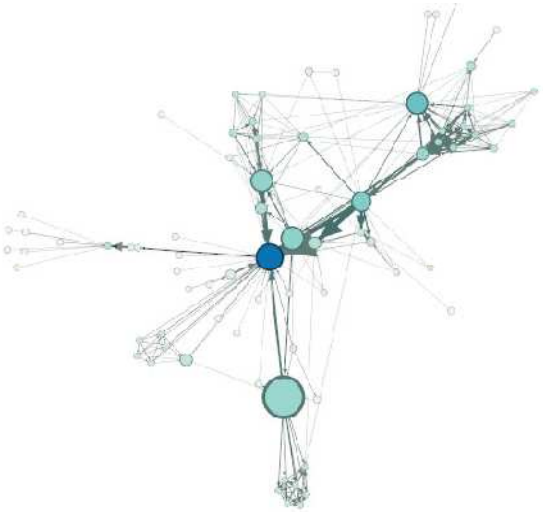


Figura 4.36: Primo risultato grafico dopo l'applicazione del applicazione grafica

Le funzioni a livello grafiche insite all'interno del software Gephi; possono premette di aggiungere alcuni dettaglia ai nodi come il nome, il carattere e le dimensioni di quest'ultimi.

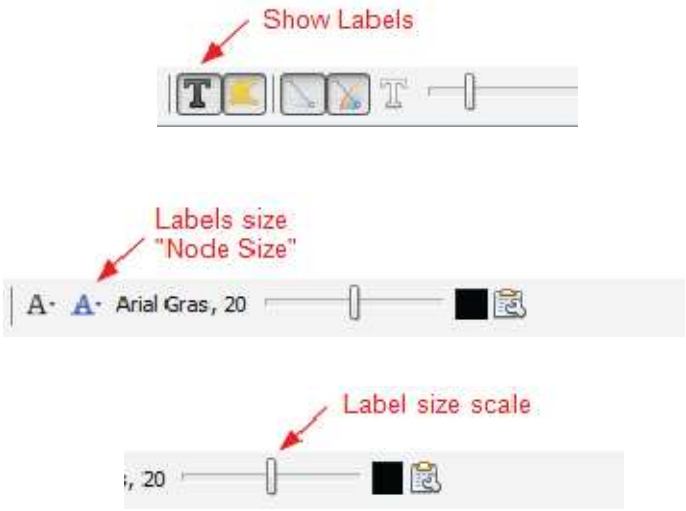


Figura 4.37: Possibilità di formattazione grafica in Gephi

Dalla precedente rappresentazione grafica, possiamo vedere di seguito i risultati finali dopo le nostre formattazioni appena descritte.

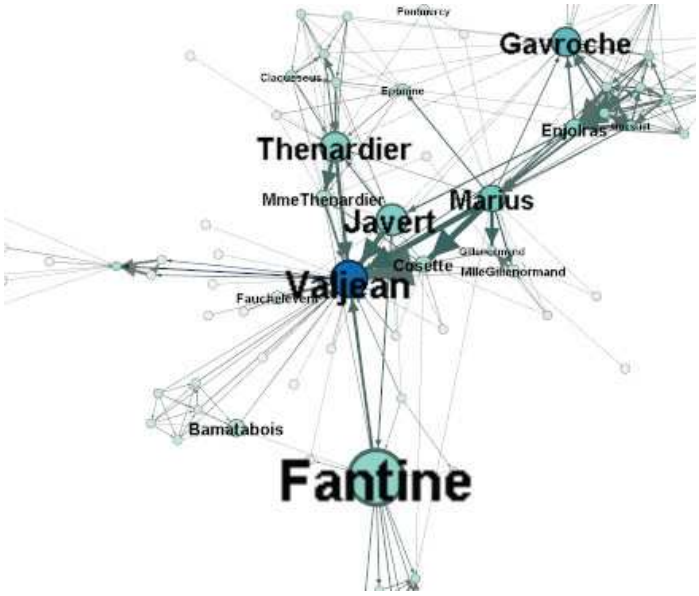


Figura 4.38: Rappresentazione grafica completa della struttura di una rete

Rimanendo in tema grafico, Gephi tra le sue funzioni permette la possibilità di clusterizzare i dati contenuti all'interno della nostra rete sociale. In particolare, questo software open source, procede alla clusterizzazione attraverso una modalità di calcolo conosciuta come Metodo di Louvain[17].

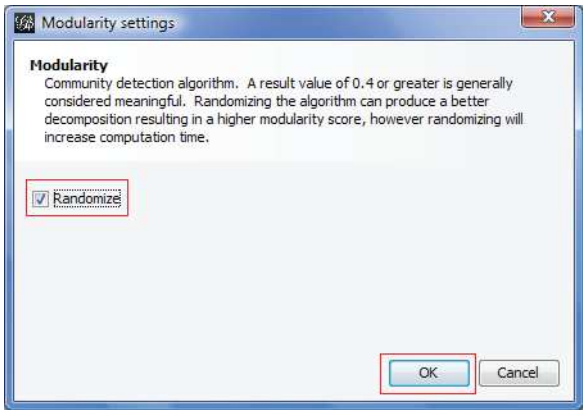


Figura 4.39: Selezione della funzione Modularity

Un volta concluso il processo matematico ed algebrico, è possibile avere una rappresentazione grafica ottimale per la comprensione della struttura dopo il processo di clusterizzazione. Dal menù Partition, si potrà evidenziare il peso specifico di ogni gruppo all'interno della rete sociale.

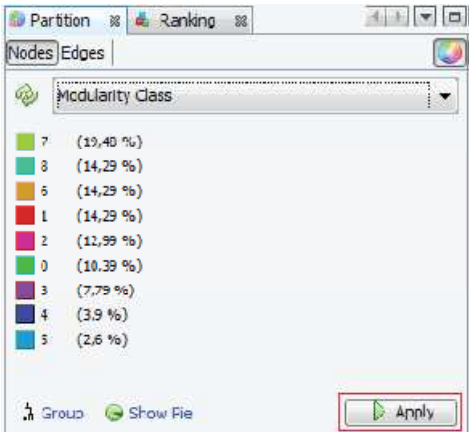


Figura 4.40: Risultati dell'applicazione della funzione Modularity

[17] Il metodo di Louvain è un metodo semplice ed efficiente per identificare i gruppi e le comunità all'interno di network molto estesi con milioni di nodi e miliardi di collegamenti. Questo metodo che si affianca al tema della Modularity, permette di aggregare di volta in volta i nodi che appartengono alle medesime comunità; successivamente, qualora la rete fosse di grandi proporzioni, i nodi aggregati sottoforma di comunità vengono considerati a loro volta nodi. Il metodo di Louvain continua in questo processo fintanto non viene raggiunto un rappresentazione comprensibile da parte dell'utente.

Dall'aspetto grafico, eseguendo un ulteriore step, andremo ad analizzare le funzioni filtro del programma Gephi. Dalla sezione menù a destra della schermata possiamo individuare la sezione dedicata ai vari filtri che un operatore può applicare a seconda delle sue esigenze di studio all'interno di una rete sociale.

L'applicazione dei filtri avviene attraverso il semplice drag del filtro nel menù sottostante, dedicato specificatamente all'applicazione di filtro.

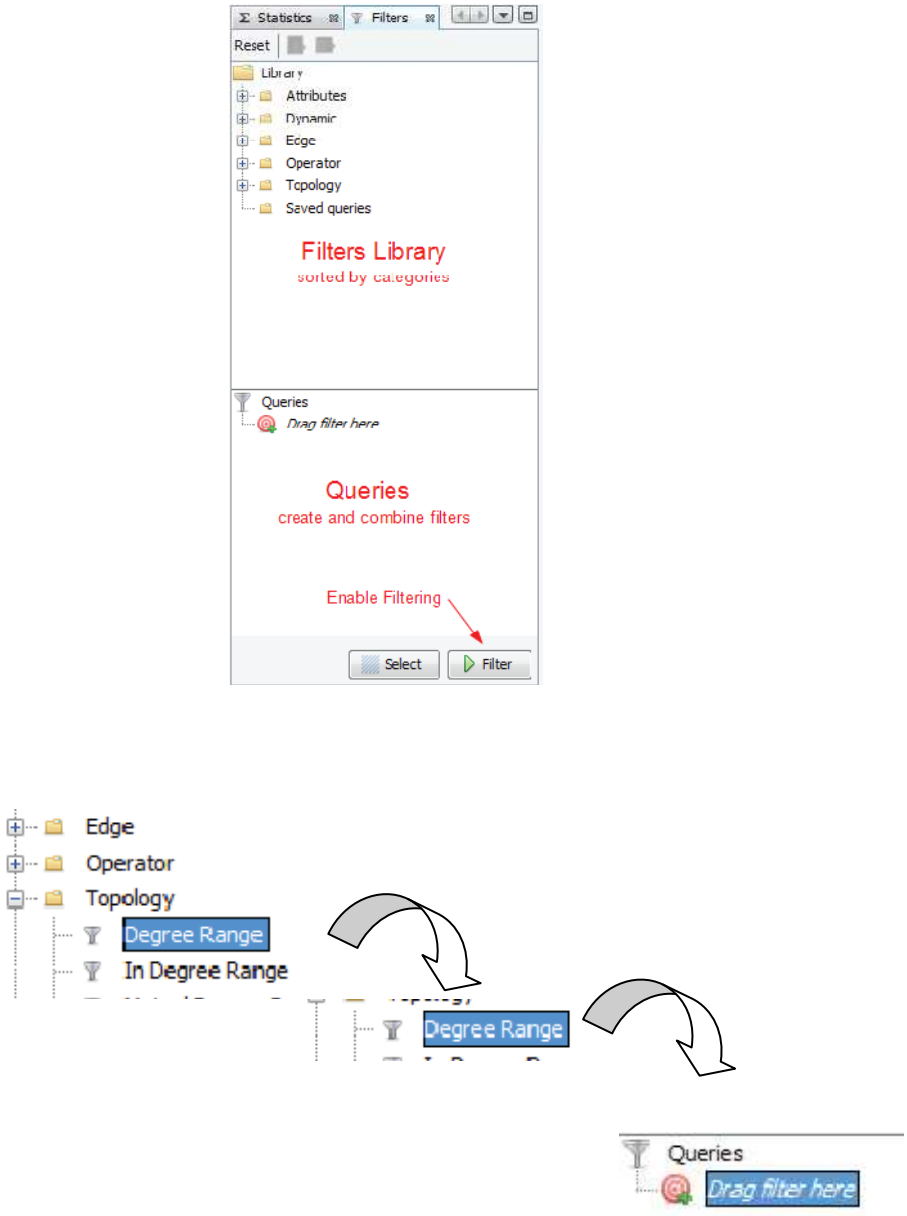


Figura 4.41: Esempio dell'applicazione delle diverse funzioni in Gephi

Un esempio di come svolgono la loro funzione di filtri, la possiamo vedere andando ad applicare uno dei tanti filtri presenti; tra quelli più semplici e intuitivi da comprendere è quello che troviamo con il nome di Degree Range all'interno della categoria Topology.

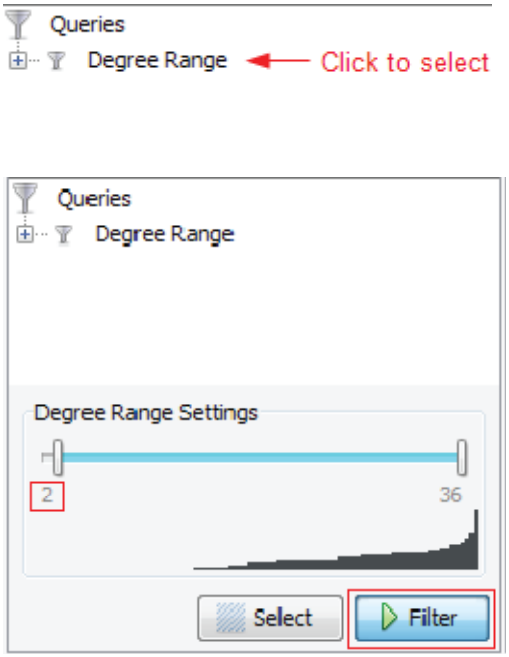


Figura 4.42: Applicazione della funzione Degree Range

Dopo questa serie step fondati sull'applicazione di formule matematiche, algebriche e visualizzazione grafica, procediamo alla conclusione del tour inerente alla conoscenza del software Gephi.
Come ultimo step andiamo a conoscere le funzioni che ci permetteranno di salvare i lavori svolti fin nella nostra rete sociale sotto forma di immagine.

4.7 L'ASPETTO GRAFICO IN GEPHI

Proprio nell'ultima sezione di Preview, poniamo le nostre ultime attenzioni nella spiegazione di questo software applicativo. Avremmo una serie di possibilità a disposizione dell'utente di poter modificare l'aspetto andando a toccare i principali elementi della rete sociale, come i già noti nodi, collegamenti, labels e il backgroud.

L’utente può successivamente salvare sottoforma si file immagine con estensione .PNG, .SVG o file .PDF.

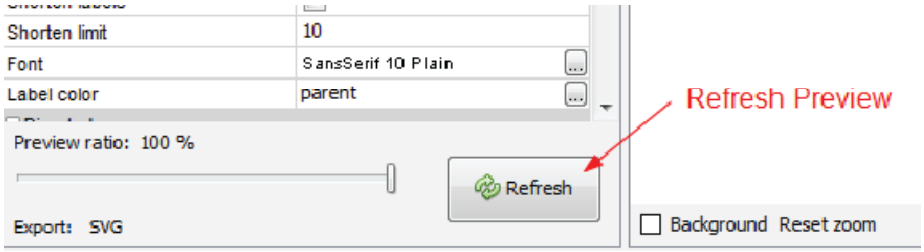


Figura 4.43: Ulteriore funzione per la visualizzazione grafica

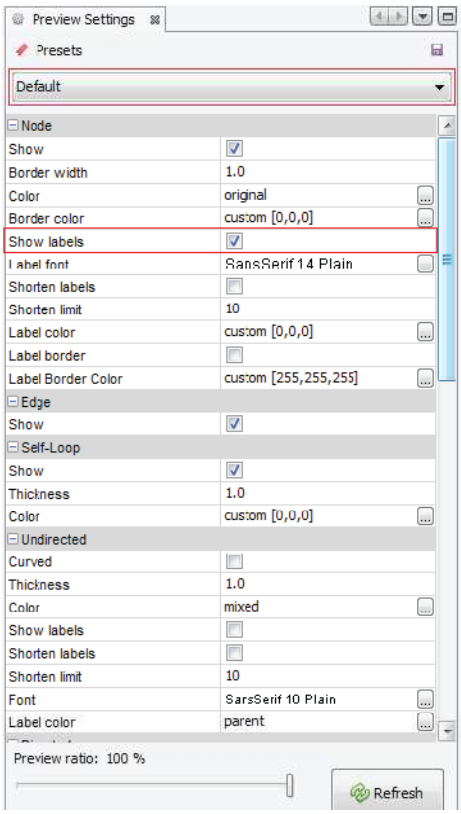


Figura 4.44: Ulteriori dettagli da apporre al risultato grafico finale

Una volta spuntato tutte le opzioni necessarie allo studio da parte dell’utente, si potrà cliccato il tasto Refresh; il risultato finale sarà simile all’immagine sottostante.

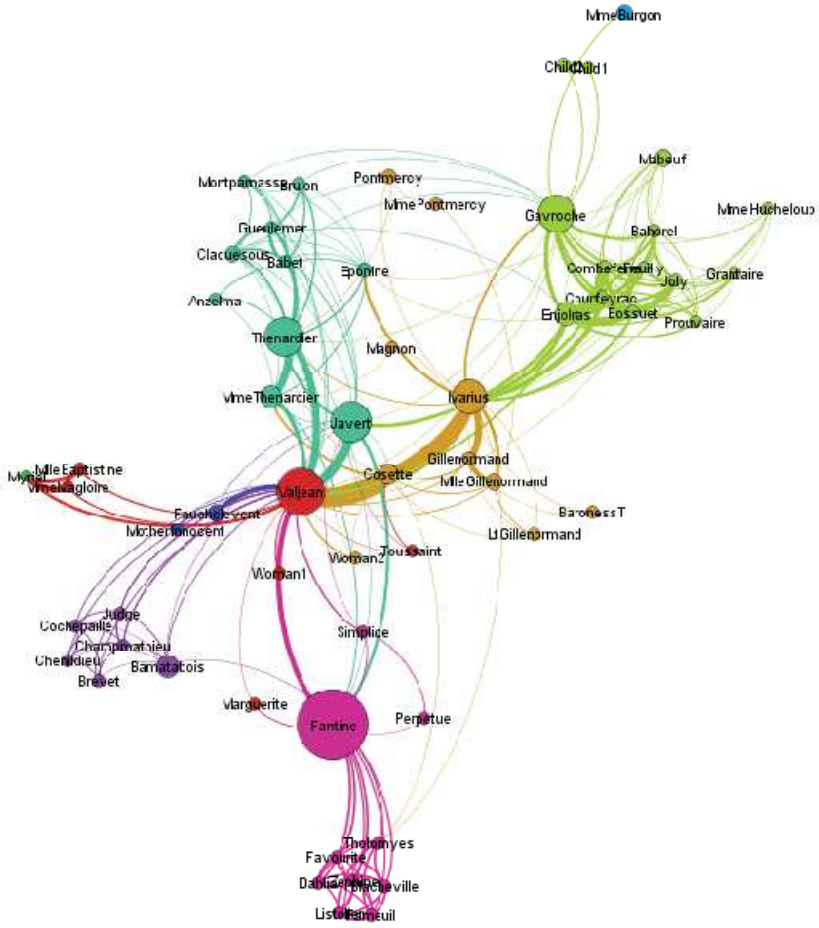


Figura 4.45: Risultato grafico ultimato

CAPITOLO QUINTO

LA POSIZIONE DEL LEADER ED IL LINGUAGGIO PYTHON

All'interno della Social Network Analysis trova spazio il concetto di leader, talvolta conosciuto anche con il termine di *opinion leader*. Tema su cui si incentra il lavoro di una leader, è finalizzato alla capacità di influenzare con le proprie azioni le idee dei soggetti che fanno riferimento a lui. Allo stesso tempo esso può dedicarsi alla comprensione dell'ambiente sociale, decifrandone le caratteristiche fondamentali che lo distinguono dal resto degli altri ambienti.

Figura 5.1: La mappatura degli obiettivi

5.1 IL LEADER

Il concetto di leader può essere visto sotto diversi punti di vista; un punto di partenza per l'analisi di questa posizione all'interno di un ambiente o di una rete sociale può essere data da una affermazione fatta da Gladwell nel suo libro" *Il punto critico: i grandi effetti dei piccoli cambiamenti* ". Gladwell infatti, ripropone sotto una diversa luce ed ottica il Principio di Pareto conosciuta anche come legge del 80/20. L'autore del

libro denota come l'80% di un lavoro o di una serie di flussi informativi siano compiuti dal solo 20% dei membri all'interno di una realtà sociale.
Gladwell collega in questo modo la posizione del leader all'interno del tema del marketing virale, e più precisamente all'interno della Legge dei Pochi.
Legge dei Pochi, la quale, oltre a definire come poche persone riescano a influenzare gruppi ben nutriti all'interno di una rete sociale, poggia su tre elementi fondamentali:

- Connettori
- Esperti di mercato
- Venditori

Attraverso questi tre elementi fondamentali, Gladwell individua tre tipologie fondamentali di persone; ai Connettori viene attribuito il ruolo di comunicatori con il resto della comunità grazie alla loro capacità di socializzazione e di intrattenere relazioni di tipo sociale.
Agli Esperti di Mercato si attribuisce la capacità di diffondere al gruppo di riferimento le nozioni tecniche e specifiche di un determinato settore; ad essi si possiamo dare il compito di comunicatori di conoscenza e sapere.
Infine, ma non ultimi per importanza, ai Venditori si attribuisce la skill di persuasione dei propri conoscenti, portandoli ad essere partecipi di un determinato progetto; il tutto riuscendo a dare ad essi una piena serenità e libertà al raggiungimento dell'obiettivo.
Un leader, oltre ad essere in possesso di questo mix di caratteristiche, deve avere un'ulteriore capacità, che si basa sull'ambito comunicativo. In Gadwell possiamo trovare tutto ciò all'interno del cosiddetto *Fattore Presa*, dal quale desumiamo come il messaggio che deve essere trasmesso dal leader ai propri followers deve essere chiaro e semplice. A coadiuvare il passaparola del messaggio devono intervenire alcune caratteristiche quali la capacità di essere ricordato e l'utilità da parte del ricevente.
All'interno di questo paragrafo possiamo citare come un leader deve avere anche capacità di *Seeding* e di *Influencer*; alla prima viene attribuita la capacità di un soggetto di poter infondere il messaggio in un ambiente "fertile". Alla *Influencer* invece viene affidata la capacità divulgativa e del passaparola all'interno di un determinato clima ed ambiente.

5.2 IL NUMERO DI DUMBAR

All'interno del tema del leader possiamo citare gli studi effettuati da Robin Dunbar, i cui frutti portarono alla ribalta il concetto conosciuto come Numero di Dunbar.
Ciò che permette di collegare la posizione di leader e il Numero di Dunbar, si rintraccia all'interno della definizione stessa; secondo Dunbar, un soggetto riesce ad avere all'interno di una comunità o di una rete sociale un numero massimo di 150 individui cui mantenere delle relazioni stabili e frequenti.
Oltre a questa definizione, Dunbar propone anche una rappresentazione grafica a livello piramidale, attraverso la quale suddivide in massimo 150 individui previsti, in quattro segmenti.
L'osservazione di questa piramide permette di notare come al vertice ci siano i collegamenti più stretti come quelli famigliari e gli amici più stretti, con un range teorico tra 3 e 7 persone. Allo step successivo vengono posti i famigliari quali zii, nonni, cugini, ecc. e tutte le persone che non hanno un contatto giornaliero con il soggetto principale. Con il terzo gradino troviamo quella serie di contatti cui il soggetto ha legami di lavoro e parentela lontana. Infine, alla base della piramide troviamo tutto il resto dei contatti cui il soggetto ha una relazione debole.

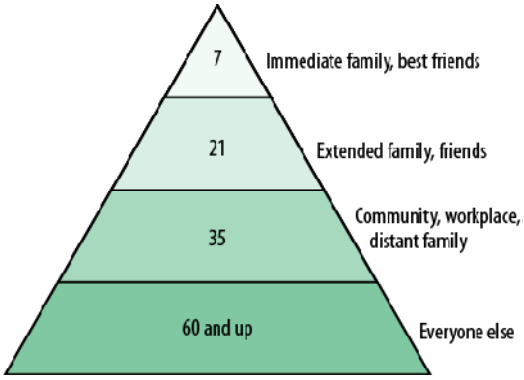


Figura 5.2: Piramide di Dunbar

5.3 NOAH FRIEDKIN ED EUGENE JOHNSEN

All'interno del tema della posizione del leader all'interno di una rete sociale, soggiunge necessario eseguire un passo successivo alla mera spiegazione della posizione centrale che una persona può assumere. Ciò avviene affiancando alla parte espositiva a quella di calcolo matematico e algebrico.

Il sunto di questi due aspetti sono alla base dell'opera "*Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*" di Noah Friedkin, che assieme a Eugene Johnsen, riuscirono a dare una risoluzione matematica all'influenza che un leader potrebbe avere. Gli studi apportati all'interno della Social Network Analysis da questi ricercatori, hanno avuto come campo di studi le forme di reti elementari quali le diadi e le triadi.

5.3.1 Il modello Noah Friedkin

Il modello matematico che scaturisce dagli studi di Noah Friedkin e Eugene Johnsen poggia sulle basi di un processo diviso in due fasi; nella prima, dove si esaminano le opinioni più influenti all'interno di una rete, alla quale sussegue un calcolo ponderato.
Quindi, l'influenza sociale che può avere un individuo all'interno di un gruppo viene misurata in base ai risultati che essa apporta negli individui; la formula pesa infatti l'opinione dei vari individui precedentemente e successivamente le dichiarazioni di un leader.
A livello di espressione matematica-algebrica, possiamo riassumere il lavoro effettuato da Noah Friedkin e Eugene Johnsen nel modo seguente:

$$y_i^{(t+1)} = a_{ii} \sum_{j=1}^{n} w_{ij} y_j^{(t)} + (1 - a_{ii}) y_1^{(1)}$$

Analizzando la seguente formula matematica possiamo individuare:

- ✓ Con t, il numero di iterazioni tra soggetto e leader per poter raggiungere il consenso
- ✓ a_{ii}, rappresenta suscettibilità di un soggetto (i); a_{ii} può essere intesa come il reciproco della influenza di un leader w_{ij}, descritta come $a_{ii} = (1 - w_{ij})$. La suscettibilità ha un valore compreso tra 0 e 1 ($0 \leq a_{ii} \leq 1$). All'interno del tema suscettibilità bisogna ricorrere ad una precisazione:
  - o Con valori di $a_{ii} = 0$, il soggetto (i) dimostra di essere irremovibile nella sua idea iniziale
  - o Mentre con valori di $a_{ii} = 1$, il soggetto (i) non attribuisce nessun peso alla propria idea iniziale
- ✓ w_{ij}, identifica la l'influenza del leader (j) nei confronti di un' altro soggetto presente nella rete sociale (i). Il valore di w_{ij} deve avere un range di valori compreso tra 0 e 1. Le seguenti condizioni possono essere riassunte nel seguente modo.

$$0 \leq w_{ij} \leq 1 \text{ e } \sum_{j=1}^{n} w_{ij} = 1$$

- ✓ $y_i^{(1)}$, rappresenta l'opinione iniziale del soggetto (i).

La formula fin qui vista analizza la mera posizione di un individuo soggetto all'influenza di un leader. Qualora l'analisi venga rapportata in un ottica di gruppo e di rete sociale, dobbiamo integrare la formula di Noah Friedkin e Eugene Johnsen con elementi matematici più complessi come le matrici.
Alla luce di ciò, gli studi portati a termine da questi due ricercatori, riuscirono a determinare la seguente formula

$$y^{(t+1)} = A\,W\,y^{(t)} + (I - A)y^{(1)}$$

Analizzando la seguente formula matematica possiamo individuare:

- ✓ A, la matrice diagonale n x n$(a_{11}, a_{22}, \dots, a_{nn})$, che raccoglie le suscettibilità dei soggetti alle comunicazioni e indicazioni degli altri soggetti.
- ✓ W, la matrice $[w_{ij}]$ delle influenze che intercorrono di un gruppo o di una rete sociale.
- ✓ $y^{(t)}$, rappresenta il vettore n x 1 relativo all'atteggiamento dei soggetti membri di una rete sociale nei confronti di un determinato argomento al tempo t
- ✓ I, rappresenta la matrice identità

5.3.2 Le Condizioni del modello

All'interno di una rappresentazione quantitativa di alcuni fenomeni - come quello sociale - utilizzando modelli matematici, dobbiamo prendere in considerazione alcune condizioni per rendere valido e credibile i modelli.
Basandosi su questi presupposti, gli studi effettuati da Noah Friedkin e Eugene Johnsen individuano una serie di condizioni basilari, che possono essere riassunti in sei punti.

- ✓ Un soggetto esprime una propria opinione all'interno di una rete sociale, nasce dalla capacità di ponderare le varie informazioni che provengono dal resto dei componenti della rete. In questo contesto vengono presi in considerazione sia aspetti di tipo endogeni ed esogeni del contesto sociale esaminato *(Wighted Averaging)*
- ✓ Il contesto della struttura sociale considerata deve mantenere le medesime caratteristiche in termini di soggetti, suscettibilità, influenza, ecc. *(Fixed social structure)*
- ✓ Il cambio di opinione di un soggetto all'interno di una rete sociale è determinata sulla base di un modello predefinito *(Determinism).*
- ✓ Il processo per la formazione di una opinione può essere continuo fintanto tutti i soggetti avranno preso una posizione in merito alle notizie immesse dai soggetti più influenti. *(Continuance)*

- ✓ La tempistica per il processo della formazione di una opinione può essere determinata in varie fasi; fasi che possono avere durate diverse da soggetto a soggetto *(Decomposability)*
- ✓ La possibilità di avere una sorta di report di tutte le influenze in atto in un determinato istante, utile a determinare una futura previsione delle opinioni da parte dei soggetti all'interno di una rete *(Simultaneity)*

5.4 IL MODELLO DEL LEADER IN PYTHON

Le teorie fin qui esposte sulla posizione del leader all'interno di una rete sociale o di un gruppo, può essere simulata in modo circoscritto a dei numero ridotto di elementi e di collegamenti. Utile a ciò è l'ausilio di software di programmazione come Python[1], attraverso il suo linguaggio di programmazione.

Oltre a Python ci si dovrà affidare ad alcune librerie di comandi che affiancano il già ricco linguaggio di programmazione. Le librerie in questione sono quelle di NetworkX[2] e Matplotlib[3].

Nelle immagine seguenti saranno evidenziati tutta una serie di stringhe di programmazione che diverranno d'aiuto allo studi e alla a simulazione della posizione del leader. Il sistema sviluppato su Python è stato ideato da Maksim Tsvetovat[4], alla quale possono essere apportate alcune variazioni in termini di valori, a discrezione dell'utente e del ricercatore.

[1] http://www.python.org/
[2] http://networkx.github.io/
[3] http://matplotlib.org/
[4] http://www.css.gmu.edu/~maksim/

```
#!/usr/bin/env python
# encoding: utf-8
"""
friedkin.py

Created by Maksim Tsvetovat on 2011-08-08.
Copyright (c) 2011 __MyCompanyName__. All rights reserved.
"""

import sys
import os
import networkx as net
import matplotlib.pyplot as plot
import matplotlib.colors as colors
import random as r
```

Figura 5.3 : Stringhe iniziali in Python per lo studio della posizione del Leader

La prima parte del complesso algoritmo creato da Maksim Tsvetovat, oltre riconoscere il diritto di copyright all'autore, indica tutti gli strumenti e le librerie che saranno utilizzati; oltre alle canoniche importazioni di Python con import *sys* e import *os*, noteremo che sono state importate, le librerie di Network X e Matplotlib.

```
class Person(object):

    def __init__(self, id):
        #Start with a single initial preference
        self.id=id
        self.i = r.random()
        self.a = self.i
        #we value initial opinion and subsequent information equally
        self.alpha=0.85

    def __str__(self):
        return(str(self.id))
```

Figura 5.4 : Stringhe in Python contenenti gli studi apportati da Friedkin e Johnsen

In questa immagine possiamo notare il momento il cui vengono inseriti i comportamenti che un soggetto assume al momento della comunicazione di un'informazione da parte del leader. Rispecchiando la formula di Noah Friedkin[5] e Eugene Johnsen[6] possiamo vedere come a livello di programmazione, le posizioni iniziali e finali dei soggetti sono rappresentate. Infatti:

[5] http://www.soc.ucsb.edu/faculty/friedkin/
[6] http://math.ucsb.edu/department/emeriti.php

- *Self.i*, rispecchia la posizione iniziale che un soggetto ha prima della ricezione delle informazioni da parte del leader.
- *Self.a*, stringa che permette di poter riassumere gli input che derivano dagli altri attori con cui il soggetto entra in relazione all'interno della rete sociale.
- *Self.alpha*, che rispecchia il valore di credulità che un leader ha all'interno di una rete sociale (*gullibility factor*). Il valore in oggetto spazia in un range tra 0 ed 1; nel nostro caso abbiamo impostato per praticità 0,85

```
def step(self):
    #loop through the neighbors and aggregate their preferences
    neighbors=g[self]
    #all nodes in the list of neighbors are equally weighted, including self
    w=1/float((len(neighbors)+1))
    s=w*self.a
    for node in neighbors:
        s+=w*node.a

    # update my beliefs = initial belief plus sum of all influences
    self.a=(1-self.alpha)*self.i + self.alpha*s
```

Figura 5.5 : Stringhe in Pyhton che analizzano le diverse opinioni dei membri all'interno di una rete

In questa immagine vediamo la successiva parte delle stringhe di programmazione; all'interna di essa possiamo individuare come le opinioni dei vari membri siano pesate allo stesso modo.

Nella stringhe successive viene introdotta la formula di Noah Friedkin e Eugene Johnsen, già introdotta nel capitolo precedente e fondamentale per studiare la posizione del leader all'interno di una rete sociale.

$$Self.a = self.alpha * s + (1 - self.alpha) * self.i$$

$$y_i^{(t+1)} = a_{ii} \sum_{j=1}^{n} w_{ij} y_j^{(t)} + (1 - a_{ii}) y_1^{(1)}$$

Figura 5.6: Formule scaturite dagli studi effettuati da Friedkin e Johnsen

Sulla base di queste informazione inseriamo una serie di stringhe che permetto di simulare a livello grafico il consenso di volta unanime tra i vari soggetti verso una determinata informazione derivante dal leader. Di seguito saranno riportate le stringhe in riferimento alla simulazione di una numero pari a 15 attori, ed il relativo grafico. Grafico che viene reso disponibile alla funzione *plot.show()* della libreria NetworkX di Python.

```
col=[n.a for n in g.nodes()]
pos=net.spring_layout(g)
net.draw_networkx(g,pos=pos, node_color=col,cmap=plot.cm.Reds)
plot.figure(2)

for i in range(15):
    for node in g.nodes():
        node.step()

    col=[n.a for n in g.nodes()]
    print col
    plot.plot(col)
plot.figure(i)
net.draw_networkx(g, pos=pos ,node_color=col, cmap=plot.cm.Reds)
plot.show()
```

Figura 5.7 : Stringhe in Python orientate all'aspetto grafico

La rappresentazione grafica permette di evidenziare come i vari soggetti tendono a convergere verso una medesima posizione in merito ad una determinata informazione, ma non raggiungendo mai e in qualunque caso il consenso unanime

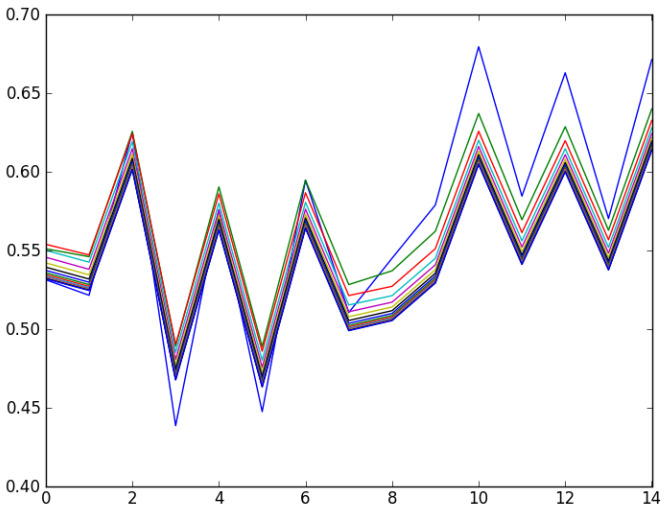


Figura 5.8: Convergenza delle idee attraverso i parametri impostati sulla posizione del leader

5.5 PYTHON OLTRE AL LEADER

Il linguaggio di programmazione Python riesce ad analizzare temi della Social Network Analysis in modo analitico; oltre all'importante apporto che fornisce allo studio della posizione del Leader, esso risulta essere uno strumento utile per decifrare in modo morfologico le reti sociali di piccole dimensioni. Python da questo lato può essere ritenuto uno strumento simile ad R con il suo linguaggio di programmazione.
Avendo come base le stesse stringhe dello studio della posizione del Leader, aggiungeremo ulteriori righe di programmazione utili al calcolo della rete sociale.

```
density=0.90
g=net.Graph()

## create a network of Person objects
for i in range(15):
    p=Person(i)
    g.add_node(p)

##this will be a simple random graph
for x in g.nodes():
    for y in g.nodes():
        if r.random()<=density: g.add_edge(x,y)
```

Figura 5.2 : Stringhe in Python per delineare la visualizzazione grafica

In dettaglio, nel nostro esempio sono stati impostati due elementi fondamentali di una rete sociale:

- *densità*,
- *numero di attori*,

Impostati questi due elementi, rimane come step successivo quello della rappresentazione grafica,affidandoci ad ulteriori stringhe qui esposte.

```
col=[n.a for n in g.nodes()]
pos=net.spring_layout(g)
net.draw_networkx(g,pos=pos, node_color=col,cmap=plot.cm.Reds)
plot.figure(2)
```

Figura 5. 3: Successive stringhe in Python per delineare la visualizzazione grafica

Le stringhe sopra esposte hanno il mero compito di imporre a Python di preparare tutti i calcoli per la rappresentazione grafica della rete sociale. Per avere a video dell'utente la simulazione della rete sociale con i parametri inseriti, dovremmo inserire un comando fondamentale contenuto nella libreria di NetworkX. Questo comando è *plot.show()*, ed il risultato grafico è il seguente:

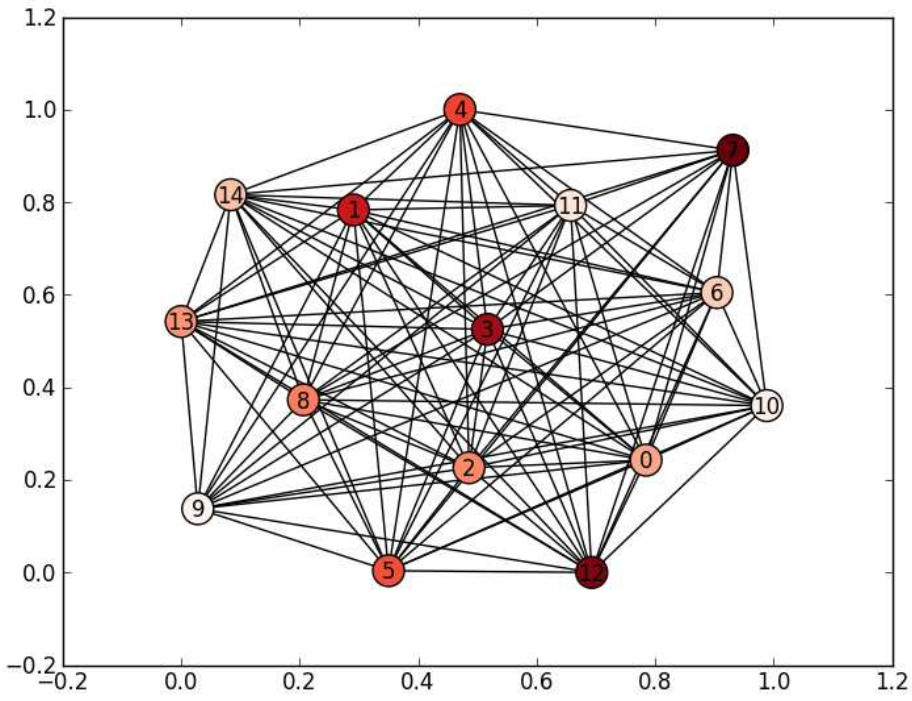


Figura 5.4: Risultato grafico dalla serie di stringhe in Python

CAPITOLO SESTO

CASO PRATICO: IL LEOPARDO NERO

Alla luce dei diversi concetti che sono stati esposti fino ad ora nel campo della Social Network Analysis, viene necessario e di obbligo dimostrare essi all'interno di un caso pratico. La scelta del caso pratico ricade nella mera presa in considerazione di un'attività commerciale situata nel territorio limitrofo, e che possa avere una mole di contatti considerevole per mettere alla prova i software d'analisi presentati.

6.1 LEOPARDO NERO: PRESENTAZIONE

La scelta dell'attività commerciale è ricaduta sul locale Leopardo Nero; il locale in questione si trova al centro delle più rinomate zone della Riviera del Brenta nei pressi di Mira.

Sin dalla sua nascita, i proprietari hanno condiviso la stessa mission comune nel rendere il Leopardo Nero, un locale giovane e che potesse offrire alla clientela una atmosfera diversa, distinguendosi dal resto dei locali presenti nella medesima zona rivierasca e limitrofa.

Tutt'ora, dopo una diecina di anni dalla sua apertura, mantiene inalterata la sua offerta di divertimento e luogo di aggregazione, restando per la zona, un locale dove poter consumare aperitivi, vini di diversa gamma, feste e cene.

6.2 LEOPARDO NERO: L'ANALISI

La base dati che siamo riusciti a ricavare attraverso l'applicazione web NameGen Web all'interno di Facebook®, è stata analizzata attraverso due soli programmi fin qui esposti in questo elaborato. I suddetti software sono Gephi e NodeXL; quest'ultimi, testati con una grossa mole di dati – 2826 nodi e 99652 collegamenti nel nostro caso – non hanno segnalato segni di instabilità. Ciò non può essere detto con software applicativi quali i già citati StOCNet e Ucinet.

6.2.1 Gli influenzatori ed i leader

Dalla prima analisi della morfologia reticolare dei contatti del Leopardo Nero, notiamo una caratteristica molto particolare. A livello grafico, applicando la funzione Force Atlas di Gephi, ribalza al primo impatto come la quasi totalità dei contatti si accentrata formando un unico blocco.

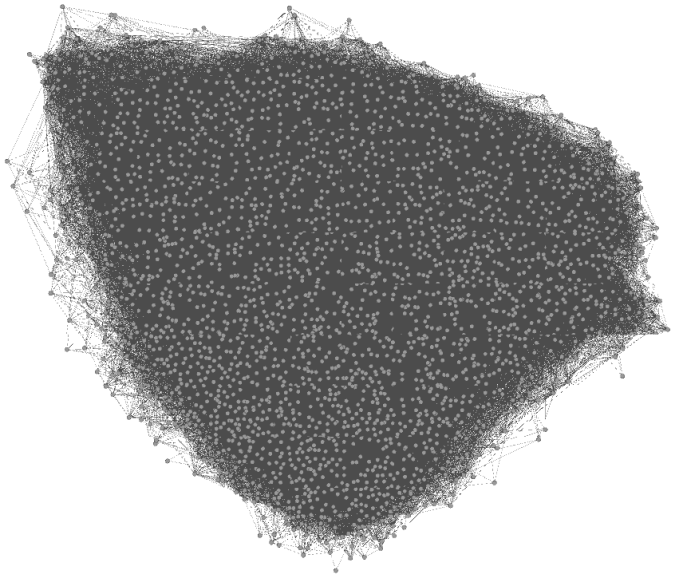


Figura 6.1: La rete del Leopardo Nero dopo l'applicazione della funzione Force Atlas

A questo punto scatta d'obbligo individuare per una corretta analisi, il ranking delle persone più influenti all'interno della rete, calcolato attraverso l'uso della funzione Degree di Gephy.

Rank	Label
742	Belafonte-Mira Valentina ...
736	AgroDolce Mira
576	Manuele Beda
570	Gran VarietÃ
546	Davide Dax Niero
537	Molocinque VE
497	O Bianco O Rosso
461	Marco Sergio Frezza
461	Francesco Ranieri
451	Moda Bella
448	Paolo Besenzon
441	Alessia Zampieri
432	Alessandro Zabeo
428	AL Numero Quattro Mirano
420	Giuseppe Leoncin
416	Mattia Effe Steve Murphy
406	Nottimagiche Pubbliche R...
406	Gos SipMira
399	Creazione Di Michele Parr...
399	Marco Canton

Figura 6.2: Ranking dei nodi influenti calcolato attraverso la funzione Degree

Conseguentemente a questo calcolo, la prima rappresentazione grafica della rete, denota fin da subito due importanti "leader" – Belafonte e AgroDolce - che possono influenzare le opinioni e i pensieri all'interno del network "Leopardo Nero".

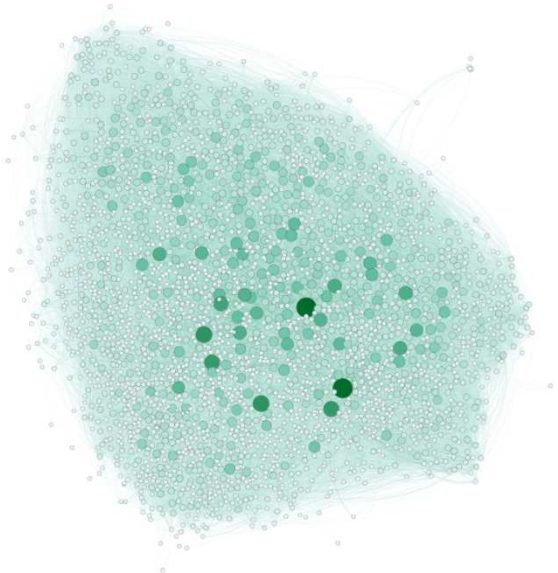


Figura 6.3: Il network "Leopardo Nero" visto dall'ottica dei leader

In modo più analitico possiamo vedere come i principali influenzatori e leader del network “Leopardo Nero” siano dislocati a poca distanza; ci affideremo alle potenzialità di NodeXL il livello morfologico delle sottoreti e dei sub-grafi presi in esame.

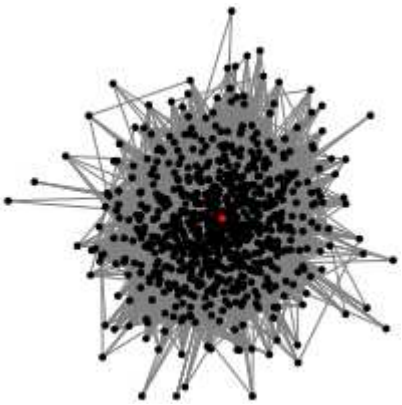


Figura 6.4: Sub-grafo del nodo Bela Fonte

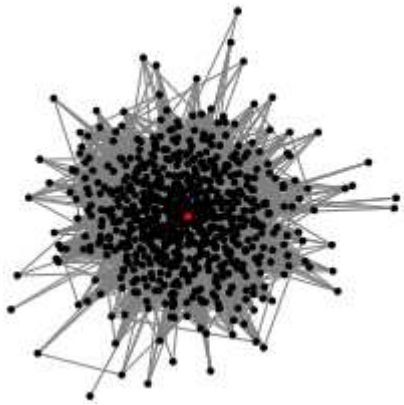


Figura 6.5: Sub-grafo Agro Dolce

6.2.2 La Degree

L’importanza di pochi e ristretti elementi influenzatori all’interno del Network “Leopardo Nero” ci viene ulteriormente confermata dal calcolo della Degree.

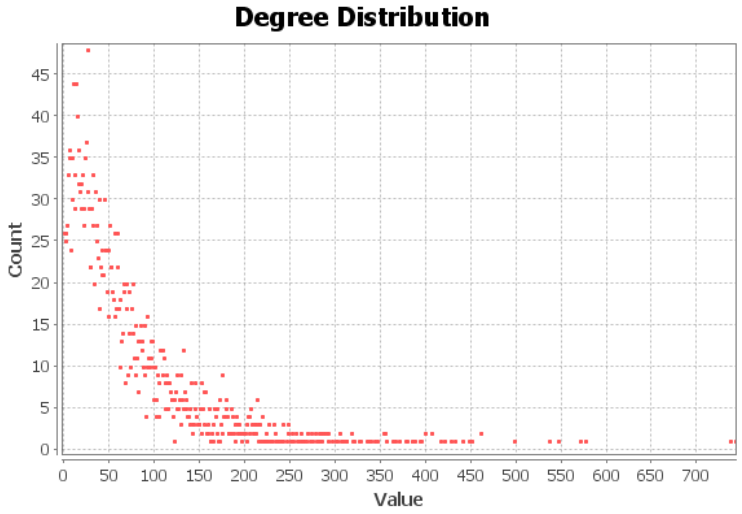


Figura 6.6: Distribuzione su grafico cartesiano della Degree

Nell’immagine soprastante si evidenzia, la marcata presenza di circoscritti nodi influenzatori all’interno del network esaminato; sull’asse orizzontale delle ascisse troviamo i valori che il software Gephi ha attribuito ai vari nodi - in termini di Degree – mentre in quello delle ordinate, troviamo la mera lista degli elementi.

La ristretta cerchia di influenzatori, può essere visibile in un ulteriore modo, andando ad analizzare tre step grafici, modificando di volta in volta il valore minimo di range del Degree all’interno del software Gephi.

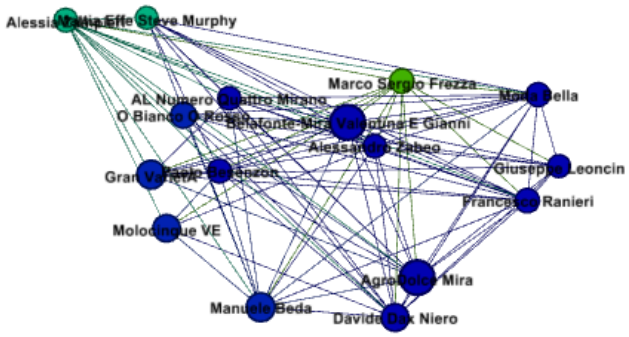


Figura 6.7: Risultato grafico con range Degree minimo pari a 411

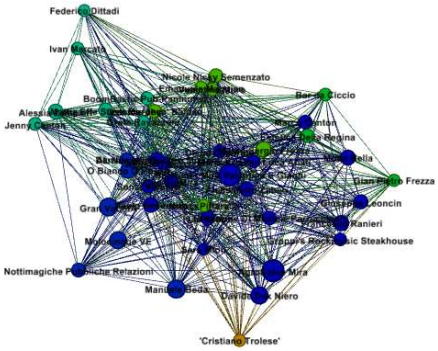


Figura 6.8: Risultato grafico con range Degree minimo pari a 321

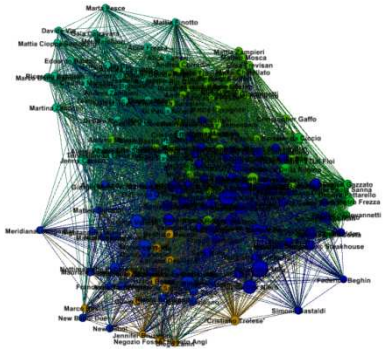


Figura 6.9: Risultato grafico con range Degree minimo pari a 203

6.2.3 Valori e grandezze a confronto

Come riferito in precedenza, Gephi e NodeXL hanno dato grossomodo gli stessi risultati in termini numerici per il calcolo delle principali grandezze della Social Network Analysis. Vedremo nelle immagini seguenti come le grandezze *Connected Components*[1], *Network Diameter*[2], *Graph Density*,[3] ecc. daranno i medesimi risultati.

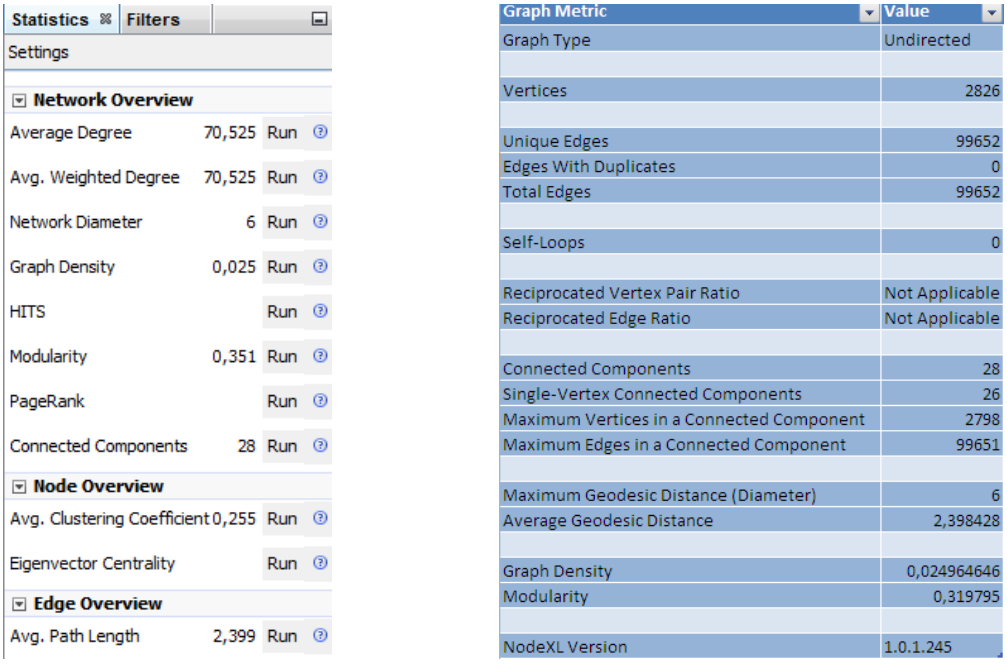
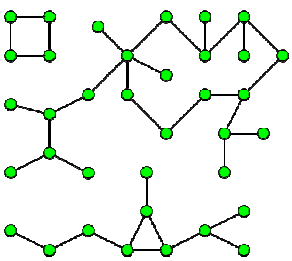


Figura 6.10: Gephi vs NodeXL - grandezze a confronto

[1] Grandezza all'interno della Social Network Analysis che individua una serie di sottografi (eng.: subgraph) in cui i vari nodi sono collegati in modo unitario e diretto con uno ed uno solo nodo; Il sistema di nodi che ne scaturisce, prevede che il nodo in esame può interagire con i restanti nodi solo attraverso il passaggio dell'informazione attraverso step di nodi



Nell'immagine qui sopra riportata, si possono evidenziare tre Connected Components

[2] Distanza geodedica massima raggiungibile che collega due nodi all'interno della stessa rete

[3] Esprime il rapporto tra il numero dei collegamenti presenti e realmente esistenti ed il numero di tutti i collegamenti possibili all'interno di una rete

6.2.1 Average Degree e Graph Density

Andando ad analizzare la *Average Degree*[4], notiamo come mediamente ogni nodo dovrebbe essere collegato ad altri nodi attraverso un numero di 70,525 collegamenti. Questo valore porta ad una prima affermazione in cui possiamo sostenere che, ogni nodo all'interno del network "Leopardo Nero", sia scarsamente legato con il resto degli altri nodi.

Questa nostra affermazione è sostenuta dal fatto che il valore della *Graph Density*[5] è relativamente basso, tanto da avvicinarsi allo 0. Il valore che ne ricaviamo è attorno al 0,025.

Il valore relativo *all'Average Weighted Degree*[6] risulta essere pari *all'Average Weighted*, poiché ad ogni nodo e è stato attribuito un medesimo peso pari all'unità.

I valori fin qui esaminati tenderebbero agli occhi dei meno esperti, affermare che il network "Leopardo Nero" sia poco sviluppato; ciò non può essere vero, vista l'impossibilità logica che, un numero elevato di nodi e collegamenti, sia collegato tra loro nelle medesime modalità.

Possiamo avere un'ulteriore conferma di questa nostra affermazione analizzando il ratio dell'*Average Clustering Coefficient*[7]; il valore che scaturisce dalla nostra analisi, - 0,255 - risulta essere basso, tanto da poter sostenere che i nodi non siano molto collegati tra loro. Questo dato purtroppo rispecchia i valori molto bassi della gran parte dei nodi che non riescono a istaurare delle reti elementari, all'interno della network "Leopardo Nero"

[4] Misura di centralità di un nodo dato dal semplice rapporto tra collegamenti e i nodi presenti all'interno di una rete;il valore può avere un valore minimo pari a 0 se il nodo è isolato, ed un massimo di (n-1) se il nodo è collegato con tutti i nodi restanti. In maniera molto pratica, nel nostro esempio, il valore 70,525 conta sia i collegamenti in entrata (in - degree) che in uscita (out – degree) visti in modo separato l'uno dall'altro.

[5] Misura che viene determinata dal rapporto tra il numero dei nodi e i collegamenti presenti all'interno di una rete sociale

[6] Misura di centralità che permette di accostare al valore della Average Degree il peso dei singoli nodi e collegamenti in termini di importanza e prestigio.

[7] Coefficiente di degree che misura la tendenza di un nodo di aggregarsi con altri nodi all'interno di una rete sociale permettendo il processo di clustering; alla base del calcolo di questo processo è l'analisi delle triadi nella loro composizione. Il coefficiente ha un range variabile tra i classici 0 e 1 ed è derivabili dalla seguente formuta

$$C = \frac{N° \; triadi \; chiuse}{N° \; triadi \; connesse \; attraverso \; vertici}$$

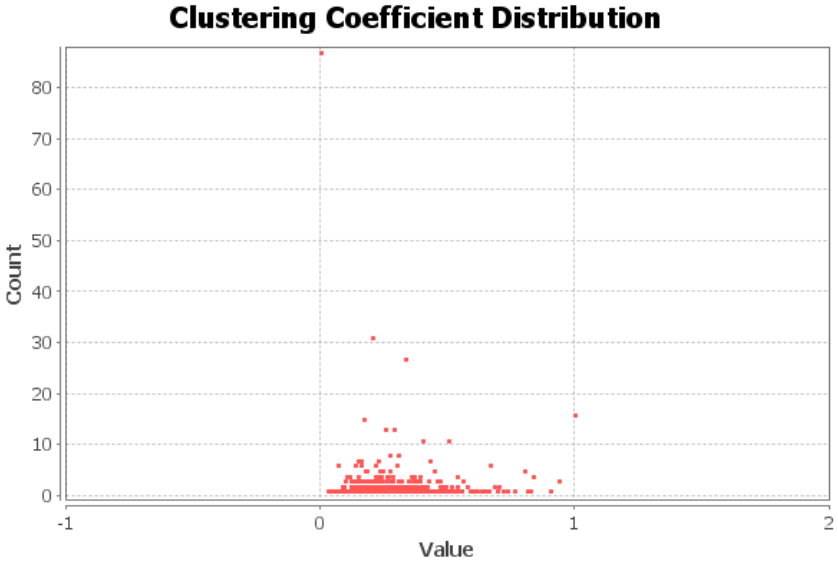


Figura 6.11: Distribuzione statistica della Average Clustering Coefficient

L'efficienza dei collegamenti fra i diversi nodi, ci viene dato *dall'Avarege Path Length*[8]; con il valore di 2,399 è nettamente inferiore al valore del *Network Diameter* – 6. Questo ci porta ad affermare che la comunicazione tra nodi viene trasmessa attraverso poco più di due collegamenti, sinonimo di una buona comunicazione e stretti legami tra i vari vertici all'interno dei vari gruppi.

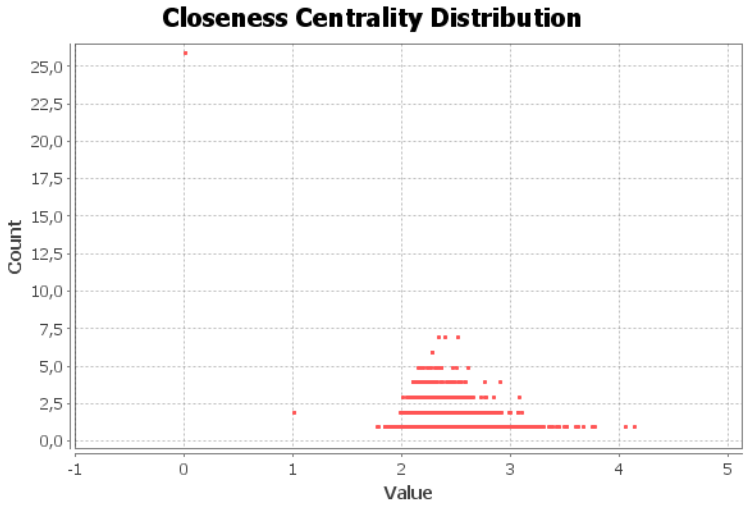


Figura 6.12: Grafico statistico prodotto dal Average Path length di Gephi

[8] Concetto di natura topologica a livello di rete sociale, attraverso la quale si analizzano in media, i numeri minimi di passaggi che intercorrono tra nodi di uno stesso network.

6.2.5 Modularity

Per individuare la dinamica della comunicazione tra i vari gruppi presenti all'interno del network "Leopardo Nero", utilizziamo la funzione *Modularity* [9] dalla quale oltre individuare i vari gruppi di interesse, verrà eseguito un semplice processo di clustering. Il valore che ne determiniamo è pari a 0,351, che individua una buona comunicazione e interscambio di informazioni tra gli appartenenti ai gruppi.

Di seguito vedremo come la rappresentazione prodotta dalla funzione *Modularity* del software Ghephy.

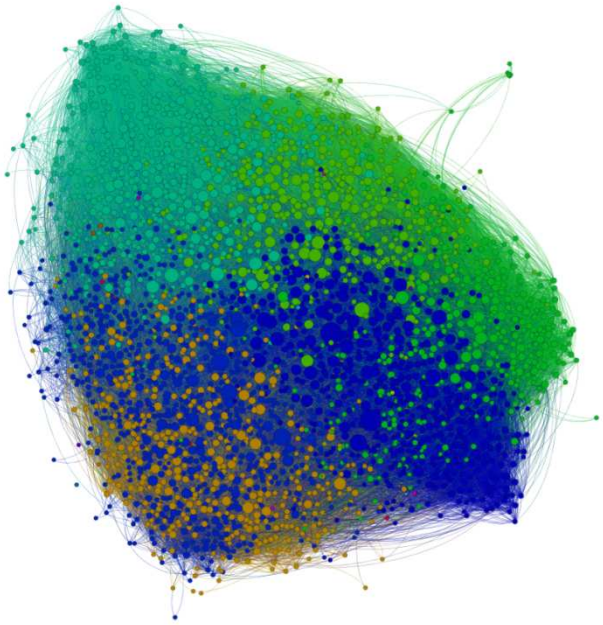


Figura 6.13: Rappresentazione grafica della modularity con Gephi

Dalla presente visualizzazione è possibile individuare la suddivisione dei principali gruppi all'interno del Network "Leopardo Nero"; notiamo a colpo d'occhio come riusciamo ad individuare principalmente quattro macro gruppi che si distinguono dagli altri. La prova numerica del numero effettivo dei gruppi viene dall'analisi svolta da

[9] Misura che determina la qualità della comunicazione e interazione all'interno dei gruppi all'interno di una rete sociale. Codesto valore può avere un range compreso tra -0,5 ed 1; se positivo, si determina che il numero dei collegamenti è superiore alla probabilità impostata ad inizio analisi. Nel nostro caso abbiamo lasciato il valore della probabilità inalterata e fissando il fattore pari ad 1.0 come da software applicativo

NodeXL; da essa, si evidenziano quattro gruppi principali che racchiudono la quasi totalità dei nodi.

	A	B	C
1		Visual Properties	
2	Group	Vertex Color	Vertex Shape
3	G1	0, 12, 96	Disk
4	G2	0, 136, 227	Disk
5	G3	0, 100, 50	Disk
6	G4	0, 176, 22	Disk
7	G5	191, 0, 0	Disk
8	G6	230, 120, 0	Disk
9	G7	255, 191, 0	Disk
10	G8	150, 200, 0	Disk
11	G9	200, 0, 120	Disk
12	G10	77, 0, 96	Disk
13	G11	91, 0, 191	Disk
14	G12	0, 98, 130	Disk
15	G13	0, 12, 96	Solid Square

Figura 6.14: Individuazione dei gruppi attraverso NodeXL

In un'altra forma di rappresentazione, viene fornita da Gephi. Nell'asse delle ordinate troviamo il numero dei nodi raggruppati dopo il processo di clusterizzazione, e nell'asse delle ascisse i vari gruppi che l'algoritmo ha derivato.

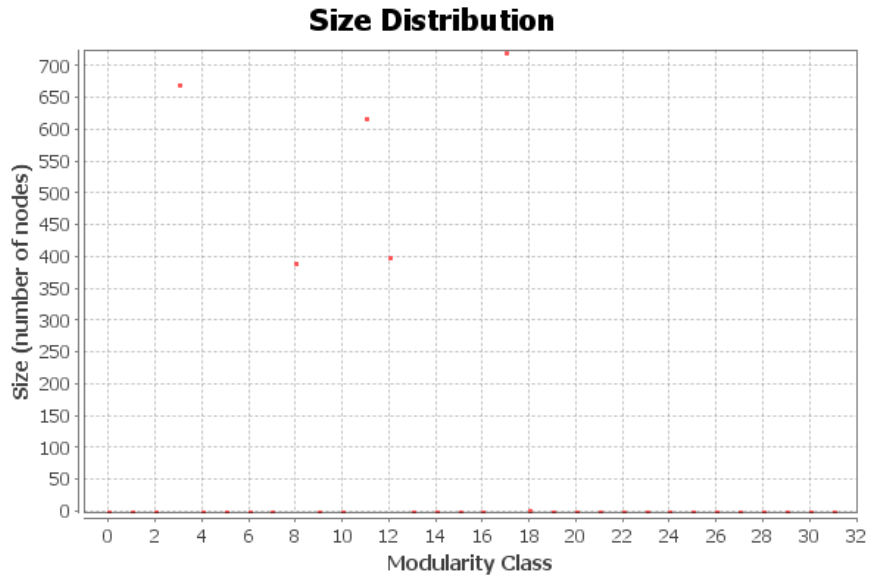


Figura 6.15: Risultato grafico della suddivisione dei gruppi attraverso la funzione Modularity Class

6.3 CONCLUSIONI CASO PRATICO

Dall'analisi fin qui effettuata sul network "Leopardo Nero" possiamo trarre delle brevi conclusioni per determinare una mera valutazione qualitativa.

Dai primi risultati grafici e numerici, possiamo notare come il network preso in esame sia ben denso di contatti legati da una fitta rete di collegamenti; l'articolata distribuzione dei all'interno della rete ha tuttavia riscontrato la presenza di alcuni leader.

La particolare vicinanza dei due leader può essere interpretata dai molti contatti in comune, cui offrire la propria proposta di divertimento; ciò è dovuto a fatto che la maggior parte dei nodi collegati, sono attività commerciali del settore bar, ristorazione e discoteche.

Le caratteristiche morfologiche della rete esposte pocanzi, inducono a proporre una iniziativa "pilota" in cui poter offrire ai clienti delle attività commerciali una serie di eventi o serate all'interno di un calendario. Ciò potrebbe portare ad un utilizzo ampio e intensificato della rete esistente; a tal proposito si potrebbe configurare la nascita di nuovi collegamenti tra nodi che andrebbero a implementare i canali comunicativi esistenti.

Questo calendario di eventi che offre una diversa gamma di possibilità di divertimento, può a sua volta implementare, ai già presenti nodi e contatti, il numero di nuovi contatti che si aggregheranno al network "Leopardo Nero".

L'intensificazione comunicativa e utilizzo della rete sociale fin qui esaminata porta a sostenere che essa può ricadere nell'ambito dei temi relativi al Marketing Virale, accennato in questo elaborato all'interno dell'analisi della posizione del leader nei network.

CONCLUSIONI

La stesura di questo elaborato ha portato a toccare diversi aspetti tecnici di una rete sociale, relativi ad attività commerciali o di un qualsiasi profilo social. Lo scopo principale è stato quello di poter introdurre gli elementi caratteristici della Social Network Analysis come la Degree o la Centralità – solo per citarne alcuni – in modo teorico e vederne in seguito una serie di applicazioni pratiche.

La Social Network Analysis, come si è visto all'interno dei vari capitoli, è una scienza nata dagli sforzi di alcuni studiosi provenienti dal campo sociale ed economico; i risultati di questi studi hanno portato ad avere una gamma ampia di strumenti che possono rendere visibile e governabili le relazioni e i collegamenti sociali.

Gli strumenti di analisi individuabili all'interno della Social Network Analisys, derivano da un excursus storico, che ha visto vari tecnici nel settore applicare diversi approcci di studio sulle relazioni e sui nodi.

Se da un lato l'approccio relazionale, posizionale e statistico, hanno apportato strumenti di analisi globali come il rapporto di densità o la centralità, da un altro lato gli approcci di analisi analitica hanno evidenziato l'importanza delle sub- networks e delle Triadi/Diadi

Le simulazioni che sono state compiute anche all'interno di questo testo, hanno reso evidente come l'immaterialità dei rapporti tra soggetti possa essere valutabile qualitativamente con strumenti numerici.

Senza l'apporto degli sviluppi informatici nel settore della Social Network Analysis, non si potrebbe aver avuto la possibilità di simulare contesti diversi di una rete sociale; si può ora ben vedere il comportamento dei vari nodi e dei collegamenti, al variare di condizioni come il numero dei nodi presenti e probabilità di connessione tra di essi.

Quest'ultimo fine è stato reso possibile grazie all'apporto di strumenti e linguaggi di programmazione in R, al cui interno possiamo trovare librerie e tools di strumenti per il nostro scopo.

I linguaggi di programmazione hanno fornito un'ulteriore possibilità di utilizzo all'interno di un elemento chiave della Social Network Analysis; qual è la posizione del Leader.

In questo caso Python ha fornito assieme ad alcune plug - in come NetworkX e Matplotlib, strumenti per testare la particolare posizione di trascinatore di un nodo all'interno di una rete sociale.

I linguaggi di programmazione appena citati hanno dato un'ulteriore apporto oltre a quello relativo ai calcoli e alle simulazioni. A tal proposito citiamo come questi programmi hanno la possibilità di rendere visibile graficamente le simulazioni delle rete sociali analizzate.

I pregi del comparto informatico si sono visti anche da un altro lato; se fino ad ora si sono specificati gli apporti dell'informatica solo a livello di linguaggio di programmazione, ci viene l'obbligo di citare i risultati dei software applicativi nel settore della Social Network Analysis

I software applicati presi in esame, NodeXL e Gephi in particolare, hanno reso possibile l'analisi puramente descrittiva di una rete sociale. Grazie alla loro natura di strumenti informatici open source, si sono dimostrati, durante le fasi di test ed elaborazione del caso pratico, programmi semplici da gestire anche dall'utente meno esperto.

Se da un lato abbiamo NodeXL che con la sua peculiarità di essere una plug-in di Microsoft Excel, si rende all'utente uno strumento basilare e maneggevole, dall'altro abbiamo l'apporto di Gephi che risulta essere un software per utenti più esperti.

Dalla simulazione del caso pratico del Leopardo Nero, si è riscontrato che, dove l'esigenza informativa e di analisi che non era fornita da un programma, poteva essere colmata dall'altro.

NodeXL e Gephi fornendo senza discrepanza i valori in uscita come Degree, Centralità, ecc – solo per citarne alcuni – si sono resi complementari nell'analisi grafica, relativamente alla rappresentazione di output.

Gephi si è reso uno strumento adatto a rispondere alle esigente grafiche qualora la rete sociale presa in esame sia di modeste e considerevoli dimensioni; in questo programma applicativo si ha possibilità di evidenziare i nodi cruciali e i collegamenti di relativa importanza.

NodeXL, altresì è uno strumento applicativo che permette di analizzare più in dettaglio le reti sociali, andando ad evidenziare quali sono le sottoreti e la posizione del nodo prescelto in sede di esame.

Proprio per questo motivo, possiamo ritenere che il programma applicativo Gephi risponde di più alle esigenze informatiche che scaturiscono da un approccio di analisi globale.

Come non ultimo per importanza nell'ambito informatico, la stessa rete web fornisce strumenti utili e semplici per l'analisi della propria rete sociale. Fra tutti gli strumenti, abbiamo messo in evidenza le funzionalità di Klout.

Klout, come è stato evidenziato, risulta uno strumento semplice ed intuitivo che permette di analizzare la propria presenza in diversi portali social. L'utente che avrà più registrazioni ai canali di informazioni e sociali – Facebook®, LinkedIn®, Youtube®, ecc – avrà più opportunità di visibilità e di avere seguaci nella rete.

Pertanto, anche se poco conosciuto, Klout risulta essere uno strumento su cui le aziende e i piccoli imprenditori dovrebbero farne uso, grazie al facile utilizzo e interpretazione del valore di rating che viene attribuito (score)

Dopo l'excursus storico, individuazione degli elementi portanti e dei software dedicati alla Social Network Analysis, si è cercato di applicare ad una realtà imprenditoriale i concetti espressi.

La scelta è ricaduta su una attività commerciale che potesse avere un numero considerevole di collegamenti e nodi. La rete sociale del profilo Facebook de "Il Leopardo Nero", rispondeva alla nostre esigenze.

Con il supporto di alcune applicazioni insite all'interno del social network fondato da Mark Zuckerbeg, è stato reso possibile estrarre i dati necessari per poter testare ed avere informazioni utili alla conformazione morfologica delle rete sociale de "Il Leopardo Nero".

Da dati grezzi e scarni di informazioni, abbiamo tratto le dovute conclusioni utilizzando gli software applicativi NodeXL e Gephi.

Con i dati alla mano e le rappresentazioni grafiche, siamo riusciti ad avere elementi per poter valutare una rete sociale, evidenziando in modo semplice quali potevano essere i pregi e i limiti.

Si è accertato quindi, come in un esempio pratico si è potuto passare dalla impossibilità di vedere la struttura di una rete sociale, alla valutazione meramente qualitativa, andando a evidenziare i punti di forza e di debolezza.
Punti di forza e debolezza che fanno parte della tabella S.W.O.T, uno strumento fondamentale per la programmazione in un'ottica strategica ed organizzativa.
Gli studi fino a qui riportati in questo elaborato possono essere la base per un piano di sviluppo di una realtà imprenditoriale; assistita da relativo budget, l'azienda potrebbe individuare nuove opportunità di crescita e spunti dalla propria struttura reticolare.

BIBLIOGRAFIA

Ajith Abraham, Aboul-Ella Hassanien, Vaclav Snásel, *Computational Social Network Analysis*, Springer Vienna, 2010

Berkowitz, Stephen, Wellman, Barry, *Social structure: a network approach*, Cambridge University press, 1988

Bonacich P., *"Factoring and weighting approaches to status scores and cluque identification"*, Journal of Mathematical Sociology, 1972

Bonacich P., *"Power and Centrality: A Family of Measures"*, American Journal of Sociology, 1987

Burts R.S., *"Toward a Structural Theory of Social Action"*, Academic Press, New York 1982

Butts C.T., *"The complexity of Social Networks: theoretical and empirical findings"*, Social Networks, 2001

Carrington P., Scott J., Wasserman S., *Models and Methods in Social Network Analysis*, Cambridge University Press, 2005.

Chiesi A., *"L'analis dei reticoli"*, Franco Angeli, 2009

Corbisiero F., *Social Network Analysis. Tendenze, metodi e tecniche dell'analisi relazionale*, Franco Angeli, 2007.

Dehmer, Matthias, Basak, Subhash C, *Statistical and machine learning approaches for network analysis*, Hoboken Wiley, 2012

Forsé M., Tronca L., *Capitale sociale e analisi dei reticoli*, Franco Angeli, 2005.

Fortunata;Barnes, John A., *Reti : l'analisi di network nelle scienze sociali - Piselli*, Donzelli - Roma, 2001

Freeman L., "*A set of measeures of centrality based on betweenness"*, Sociometry, 1977

Freeman L., "*Visualizing Social Networks"*, Journal of Social Structure, 2000

Friedkin Noah, "*Social Influence Network Theory*", Cambridge University Press, New York, 2011

Friedkin, Noah E.;Johnsen, Eugene C., *Social influence network theory: a sociological examination of small group dynamics*, Cambridge University press, 2011

Freeman, Linton C.;Memoli, Rosanna, *Lo sviluppo dell'analisi delle reti sociali: uno studio di sociologia della scienza*, Milano F. Angeli, 2007

Gladwell Malcom, "*Il punto critico: i grandi effetti dei piccoli cambiamenti*", Rizzoli, Milano, 2006

Katz L., "*A new status index derived from sociometric analysis"*, Psychometrika, 1953

Katz L., Powell J.H. "*Measurement of the tendency toward reciprocation of choice"*, Sociometry, 1955

Knoke D., Kulinski J.H. "*Network Analysis"*, Sage University Paper series on Quantitative Applications in Social Sciences – Sage Publication London, 1982

Kolaczyk, Eric D, *Statistical analysis of network data : methods and models*, Springer New York, 2009

Lee Giles, Marc Smith, John Yen, Haizheng Zhang, *Advances in Social Network Mining and Analysis*, Springer Vienna, 2010

Lin N, "*Foundations of Social Research"*, McGraw Hill – New York, 1976
Maksim Tsetovat & Alexander Kouznetsov, "*Social Network Analysis for Startups*", O'Reilly, 2011.

Piselli F., *Reti*, Donzelli, 1995.

Prell, Christina, *Social Network Analysis: history, theory and methodology*, ed. Los Angeles Sage, 2012

Pryke, Stephen, *Social Network Analysis in Construction*, Hoboken John Wiley & Sons., 2012

Rashotte Lisa, "Social Influence", *http://www.sociologyencyclopedia.com/fragr_image/media/*

Reis Pinheiro, Carlos Andre, *Social network analysis in telecommunications*, Hoboken John Wiley & Sons Inc., 2011

Salvini A. (a cura di); G. Conaldi, D. Cordaz, S. Gabbriellini, E. Giuliani, A. Givigliano, E. Mazzoni, S. Milella, G. Rivellini, A. Salvini, L. Savoia, G. Tomei, S.Zaccarin, "*Analisi delle reti sociali – Teorie, metodi, applicazioni*", Franco Angeli, 2007.

Salvini, A., *L'analisi delle reti sociali. Risorse e meccanismi*, PLUS, 2005.

Scott, John, *L'analisi delle reti sociali*, NIS, 1997

Scott J., *Social Network Analysis. A Handbook*, London Sage, 2000.

Scott, John, Wasserman, Stanley, Carrington, Peter J., *Models and methods in Social Network Analysis*, Cambridge University Press, 2005

Tansel Özyer, Jon Rokne, Gerhard Wagner, Arno H.P. Reuser, *The Influence of Technology on Social Network Analysis and Mining*, Springer Vienna, 2013

Vincenzo Cosenza, "*Social Media ROI*", Apogeo, 2012.

Trobia A., Milia V., *Social Network Analysis. Approcci, tecniche e nuove applicazioni*, Carocci, 2011.

Vargiu A., *Il nodo mancante*, FrancoAngeli, 2001.

Wasserman S., Faust K., "*Social Network Analysis*", Cambridge University Press, 1998.

Wasserman S., Berkovitz S., "*Social Structure. A network approach*", Cambridge University Press, 1988.

Wasserman, Stanley, Faust, Katherine, *Social network analysis : methods and applications*, Cambridge university, 1994

SITOGRAFIA

Butts C.T., *"sna: Tools for Social Networks Analysis. R package version 2.3-1"*, *http://cran.r-project.org/web/packages/sna/sna.pdf*, 2013

Cug.test {sna} - *http://www.inside-r.org/packages/cran/sna/docs/cug.test*

Exploratory social network analysis with pajek - *http://www.slideshare.net/plotti/exploratory-social-network-analysis-with-pajek*

Facebook Network Analysis using Gephi - *http://www.slideshare.net/persuasion/facebook-network-analysis-using-gephi-11390089?ref=http://persuasionradio.wordpress.com/2010/05/06/using-netvizz-gephi-to-analyze-a-facebook-network/*

Gephi Plugin Developer Workshop - *http://www.slideshare.net/gephi/gephi-plugin-developer-workshop?from_search=1*

IEEE SocialCom 2009 NetViz Nirvana (NodeXL Learnability) - *http://www.slideshare.net/ebonsign/ieee-socialcom-2009-netviz-nirvana-nodexl-learnability#btnNext*

Il Seeding tutto quello che volevate sapere Viralbeat – Social Media Company - *http://viralblog.viralbeat.com/2011/05/il-seeding-tutto-quello-che-volevate-sapere/*

Louvain method for community detection - *http://perso.uclovain.be/vicente.blondel/research/louvain.html*

M. Jacomy, S. Heymann, T. Venturini, M. Bastian, *"ForceAtlas2, A Graph Layout Algorithm for Handy Network Visualization"*, *http://webatlas.fr/tempshare/ForceAtlas2_Paper.pdf#page=3&zoom=auto,0,668*

Remix Your Data: Visualizing Library Instruction Statistics - *http://www.slideshare.net/notsosternlibrarian/visualizing-library-instruction-statistics*

Repository - *http://it.wikipedia.org/wiki/Repository*

Social network analysis intro part I - *http://www.slideshare.net/plotti/social-network-analysis-intro-part-i*

Social network Analysis - whys and hows (with NODEXL) - *http://www.slideshare.net/muthukum/social-networks-next-gen-knowledge-platforms*

Social network analysis course 2010/2011 - *http://www.slideshare.net/ereteog/social-network-analysis-5800120*

Social Network Analysis – QAP Procedure, *www.lsu.edu/faculty/bratton/networks/qap.ppt*

StOCNET - *http://www.gmw.rug.nl/~stocnet/StOCNET.htm*