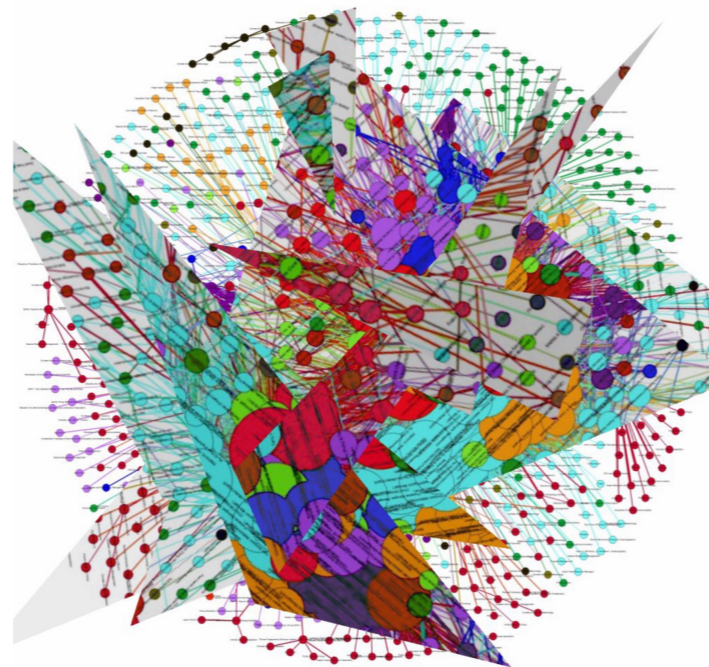


Daniele Frongia

SOCIAL NETWORK ANALYSIS

Una introduzione ai metodi
e agli strumenti



Data Science

Daniele Frongia

SOCIAL NETWORK ANALYSIS

Una introduzione ai metodi e agli strumenti

Data Science

Daniele Frongia

Social Network Analysis. Una introduzione ai metodi e agli strumenti

Roma : Dipartimento di Scienze statistiche [2015] 43 p.

ISBN 978-88-908757-3-1

© 2015, Daniele Frongia

Questo libro è stato realizzato con iBooks Author per la visualizzazione in Apple Mac con OS X Mavericks e iPad con iOS 5 o versioni successive. Per la segnalazione di errori e imprecisioni scrivere all'autore: Daniele Frongia, danielefrongia@gmail.com o al curatore: Luca Giuliano, luca.giuliano@uniroma1.it. Una copia statica in PDF è disponibile sul sito del dipartimento di Scienze statistiche - Sapienza Università di Roma: <http://www.dss.uniroma1.it/ricerca/pubblicazioni>

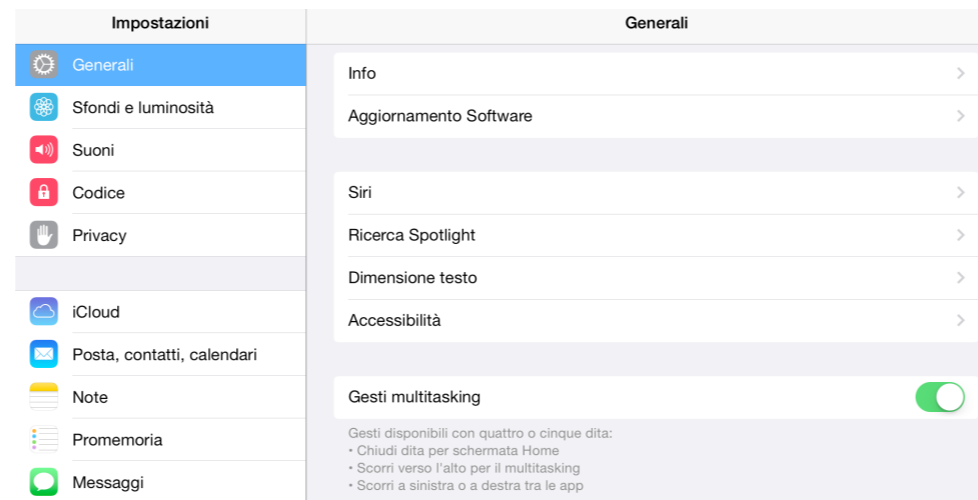


This book is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. This book may be reproduced, copied and distributed for non commercial purpose, provided the book remains in its complete original form.

Istruzioni sull'attivazione dell'opzione "Gesti Multitasking" su iPad

Per una lettura ottimale di questo libro interattivo e per una piena fruizione dei link esterni ci dobbiamo prima di tutto assicurare che l'opzione Gesti multitasking per iPad sia abilitata.

Aprirete **Impostazioni** e toccherete **Generali**. Scorrete verso il basso per trovare "Gesti multitasking" e posizionate lo su ON.



Utilizzando quattro o cinque dita per strisciare in orizzontale (come nella figura al centro, qui sotto) si vedranno scorrere le applicazioni aperte. Provate scorrendo da destra a sinistra, dal momento che molto probabilmente in questo momento siete nell'ultima applicazione utilizzata.



Questo gesto vi permetterà di ritornare facilmente alla lettura del libro nel punto in cui vi siete interrotti dopo aver seguito un link esterno nel browser Safari.

Le altre funzioni di Gestii multitasking, spesso poco note agli utenti di iPad, sono:

- Chiudere le applicazioni e tornare alla schermata principale con un gesto di “pizzico” effettuato con quattro o cinque dita (figura a sinistra).
- Rivelare la barra delle applicazioni aperte con un colpo in su effettuato con quattro o cinque dita (figura a destra). Ripetendo il colpo verso il basso la barra delle applicazioni si chiude.

Introduzione

Qual è il fattore più rilevante del successo dei social media?



1

Perché la Social Network Analysis?

Qual è il valore aggiunto di Facebook, Twitter, LinkedIn, Google+? La risposta è fin troppo ovvia: la risposta... siamo noi. Siamo noi che, con la nostra rete, il nostro vissuto e la nostra quotidianità, rendiamo questi social media luoghi vivi, ricchi di informazioni e di funzionalità.



Fig. 1.1 Un network di social media

Fenomeni come Facebook hanno di gran lunga superato il dominio dell'online e sono prepotentemente entrati nella vita (reale, tangibile, 3D) di tutti i giorni. Oltre venti milioni di italiani frequentano abitualmente i social network e la

penetrazione tra gli adolescenti (e non solo) è impressionante. Termini derivanti dai social network (mi piace, +1, taggare e, ahinoi, “laikare”) sono ormai entrati di fatto nel nostro dizionario.

Per comprendere pienamente il successo e le potenzialità di questi servizi partiamo dalle origini, e in particolare dalle origini delle reti sociali e dei suoi studi, la social network analysis (o SNA), dagli albori di Eulero fino alle reti di siti web.

Con una premessa molto importante: a che serve la SNA, quali sono, oggi, i campi di applicazione? I campi sono numerosi: ad esempio nella sociologia, nell'antropologia e la psicologia, nell'economia, nella comunicazione, nell'organizzazione aziendale, nella fisica, nella biochimica, nella genetica, nella medicina, nell'informatica, nelle investigazioni criminali e nel giornalismo investigativo.

L'esplosione del fenomeno dei social media ha dunque spinto aziende, istituzioni e ricercatori a un uso sempre più consistente di strumenti e metodi di analisi delle reti; in questa sezione ne vedremo alcuni.

Breve storia della Social Network Analysis

Partiamo dalla storia. E prima ancora
da una definizione.



2

Una definizione di rete sociale

Una rete sociale è una struttura formata da nodi (attori) e da legami che intercorrono tra di essi, come ad esempio l'amicizia, la collaborazione professionale o la parentela ecc. Le cinque persone in fig. 2.1 costituiscono un gruppo in cui Alice conosce Mario e Ludovico (a loro volta amici) e anche David e Leo (a loro volta amici).

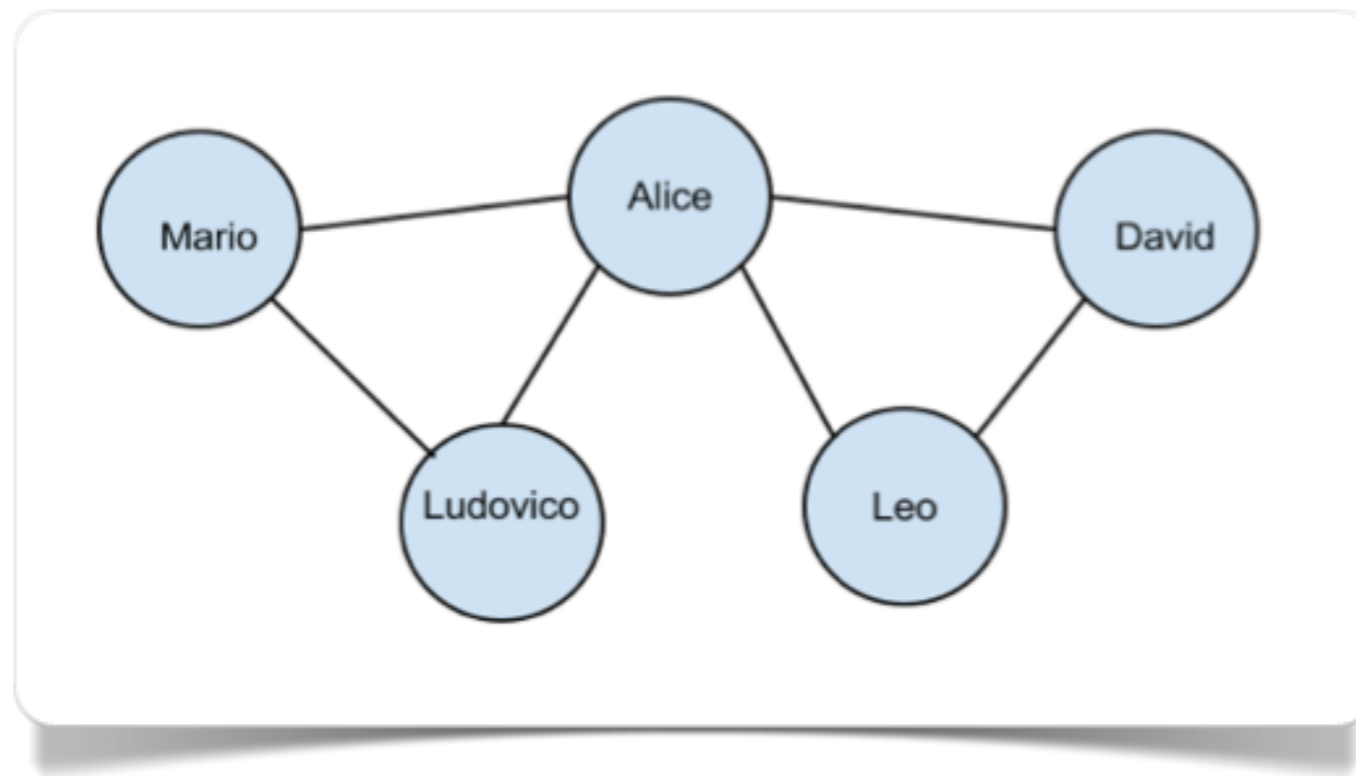


Fig. 2.1 Un esempio di rete sociale (grafo) con cinque nodi (persone) e sei legami

Dunque, nella SNA, per dato si intende l'elemento relazionale, e cioè le informazioni relative a una rete sociale composta da attori con i loro attributi, e i legami (link) tra i nodi.

In altre parole, l'analisi delle reti sociali considera le relazioni sociali in un'ottica di **network theory**, utilizzando quindi algoritmi e strumenti di analisi della teoria dei grafi. Proprio per questo motivo le radici metodologiche della SNA sono da attribuire a Eulero e ai suoi celebri studi sui **ponti di Königsberg**.

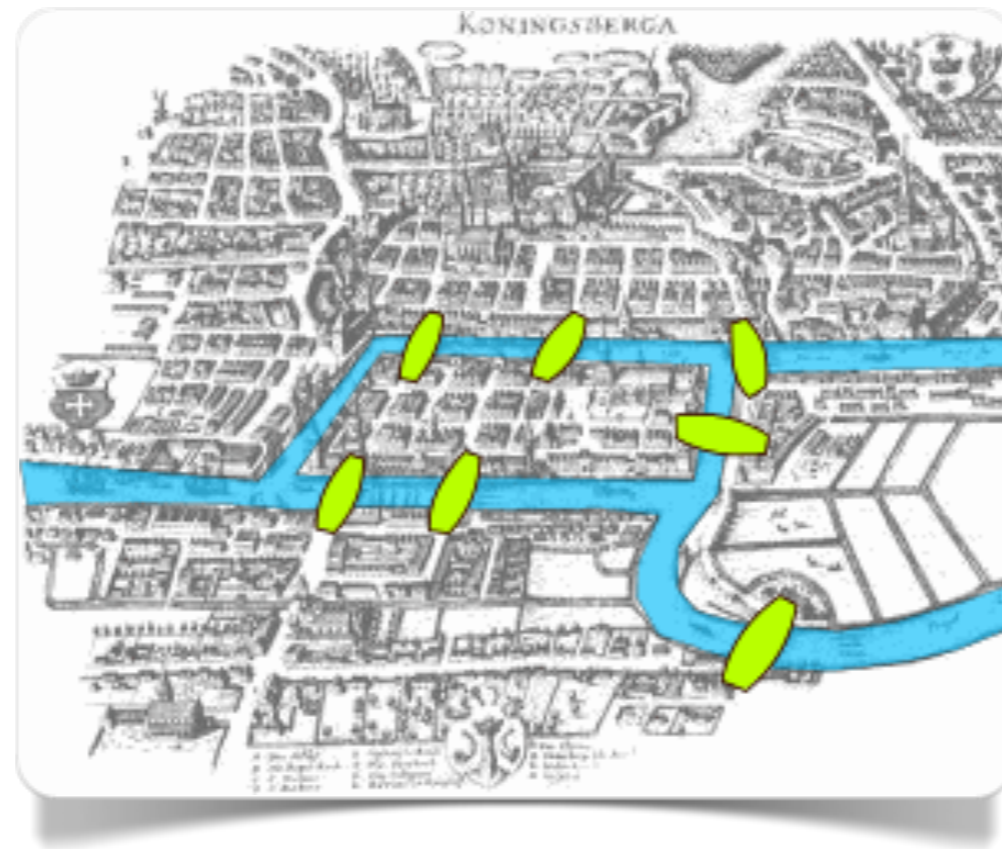


Fig. 2.2 Gli studi sui ponti di Königsberg

Successivamente, le prime analisi “sociali” furono condotte nell'Ottocento per gli scopi più vari: lo studio dell'interazione tra api, delle tribù nord americane, della scomparsa dei cognomi e delle dinamiche della folla.

Le dinamiche dell'interazione

Nei primi decenni del Novecento l'analisi delle reti sociali è stata poi applicata alle dinamiche di interazione tra bambini (inviti dei bambini alla festa, chi-ha-giocato-con-chi) e negli anni '30 **Jacob Levi Moreno** ha fondato la **sociometria**.

Tra gli anni '30 e '40 Harvard diede poi un significativo impulso e iniziarono a essere prodotti i primi studi a livello aziendale: furono eseguite, ad esempio, ricerche di antropologia sociale applicate allo studio della produttività degli operai, passando per le reti informali formate dalle amicizie e dai gruppi nelle catene di montaggio.

Fino agli anni '70 i contributi furono prevalentemente di natura matematica e gli attori principali furono studiosi di alcune università tra cui le università di Iowa, Michigan State University, MIT, Sorbona, Lund, Columbia, Manchester e la London School of Economics. Da segnalare, in questo lasso di tempo, lo studio di **Torsten Hagerstrand** sulla diffusione delle innovazioni nello spazio e nel tempo. **Claude Flament**, invece, fu il primo ad applicare formalmente la teoria dei grafi all'analisi delle reti sociali.

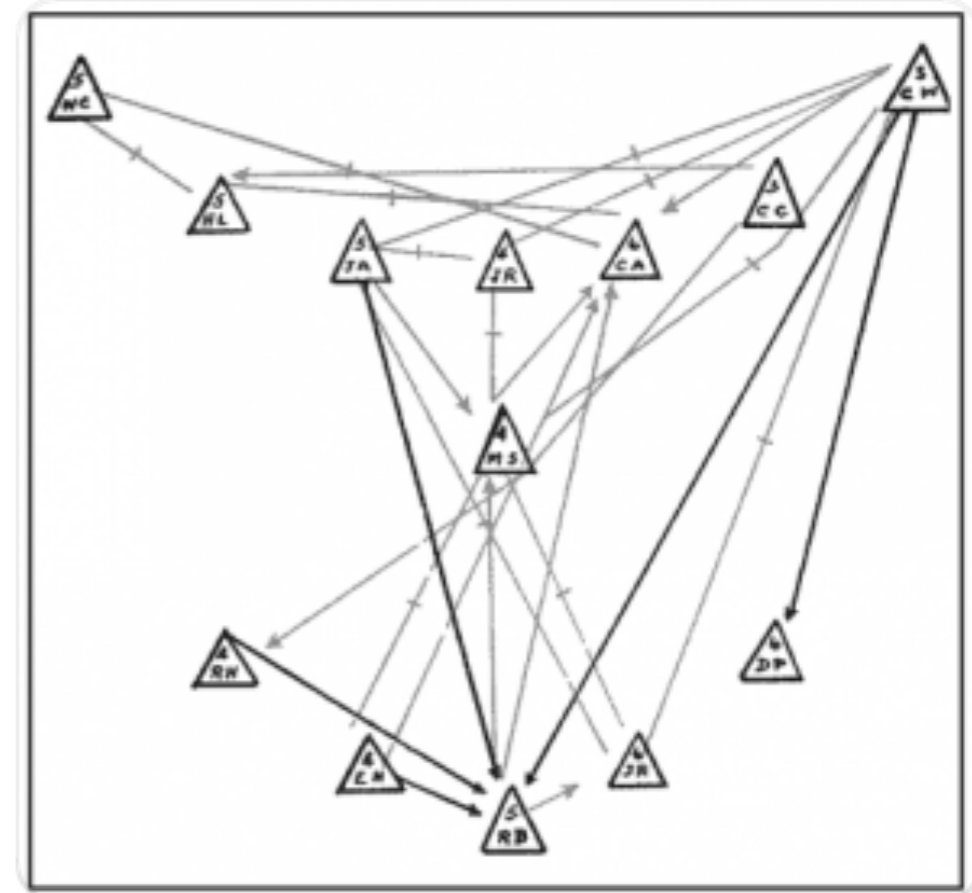


Fig. 2.3 Un grafo sociometrico



Fig. 2.4 Linton C. Freeman

Un altro significativo contributo alla SNA è stato dato dall'opera pluridecennale di **Linton C. Freeman** che, oltre ad essere il padre del software Ucinet, ha analizzato i processi decisionali all'interno di una azienda.

A partire dalla fine degli anni '70 la SNA viene riconosciuta dal mondo della sociologia e anche dalle altre scienze sociali e matematiche, essendo per sua natura multi-disciplinare.

La moderna SNA è basata sul seguente approccio strutturale:

- studio dei legami tra gli attori sociali,
- dati empirici sistematici,
- rappresentazione grafiche (ne vedremo diversi proprio in questa app),
- uso di modelli matematici.

Se vuoi saperne di più sulla storia della SNA, **qui trovi una sintesi** dalla Sociometria di Jacob Levi Moreno in poi.

Prima di andare avanti e vedere gli aspetti statistici di base della SNA, proviamo a ricapitolare con un **documentario sulle reti sociali**:



Fig. 2.5 A documentary on networks, social and otherwise

Statistiche di base della SNA

Vediamo alcune definizioni e misure
per analizzare una rete:

Densità

Centralità

A large, dark red number '3' is centered within a white circle. The circle is set against a background of concentric white circles and a grid of white lines, all on a dark red gradient background with small white specks.

Densità

La densità rappresenta una delle più importanti statistiche descrittive, spesso utilizzata come indicatore del livello generale di coesione della rete.

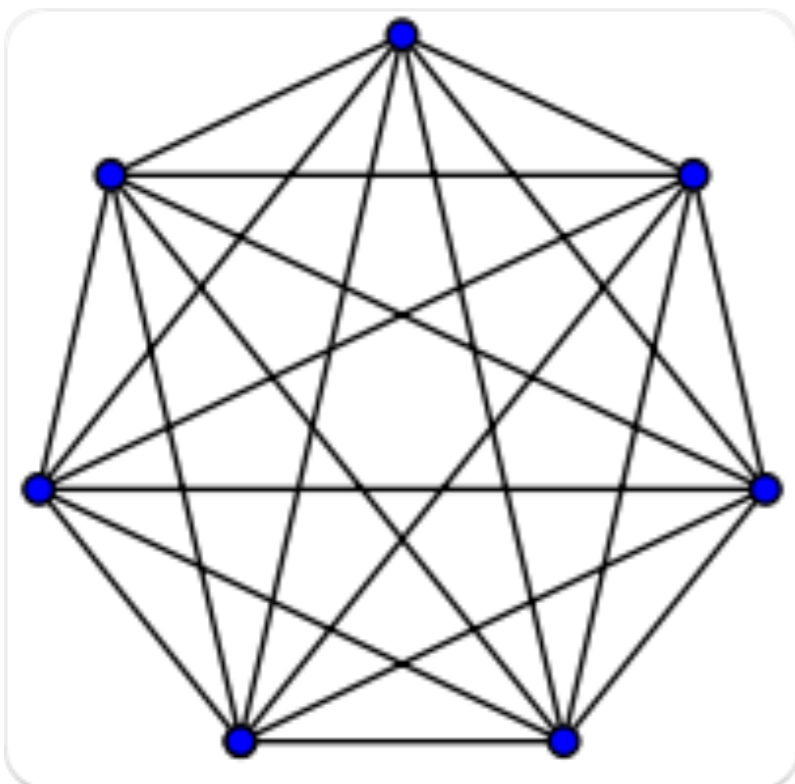


Fig. 3.1 Esempio di grafo completo: tutti i nodi sono connessi tra di loro

Può dunque essere utile quando un gruppo sia coeso, unito, collaborativo.

Da un punto di vista matematico, la densità di una rete è la proporzione dei legami presenti su tutti i legami possibili.

La densità di una rete assume dunque valori compresi tra 0 e 1: valori prossimi allo 0 indicano bassi livelli di coesione, mentre valori vicino ad 1 indicano alti livelli di coesione della rete.

La densità è pari ad uno in caso di **reti (grafi) completi**, quelli cioè dove tutti i nodi sono collegati agli altri, come in figura 3.1.

Centralità

Le **misure di centralità** permettono di definire il posizionamento di un attore nella propria rete in termini relazionali.

Questa misura è particolarmente importante ad esempio nel marketing sul Web, dove la ricerca di nodi centrali equivale, in alcuni casi, all'identificazione di **opinion leaders**.

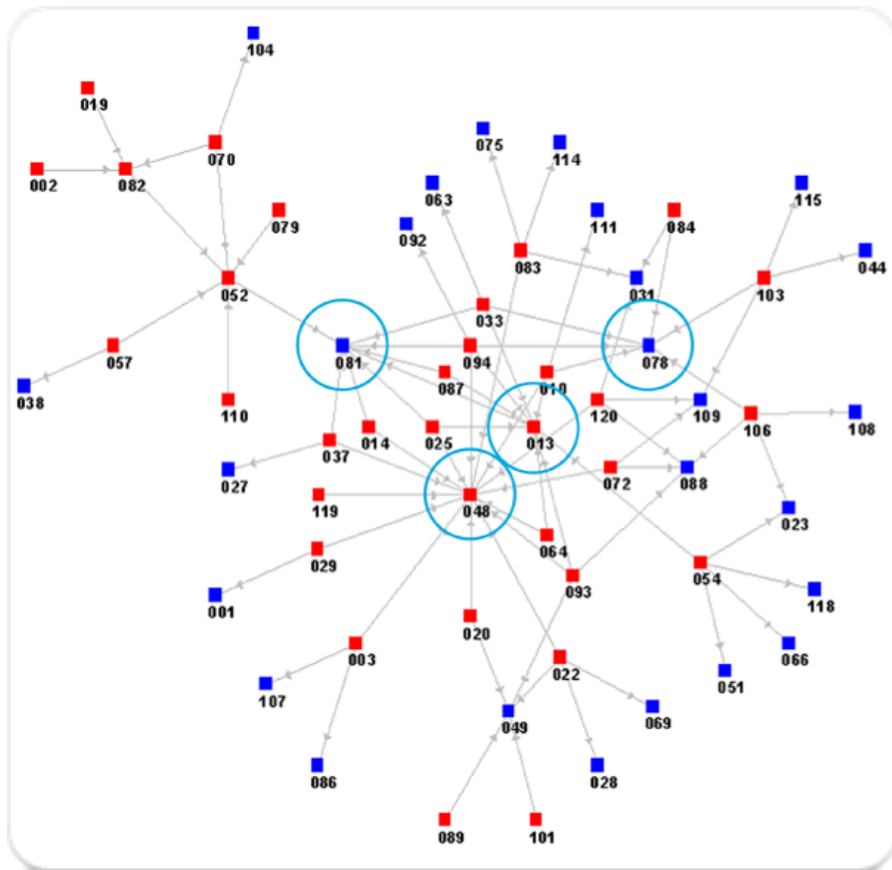


Fig. 3.2 Esempio di analisi degli *opinion leaders* (cerchiati in blu). Clicca sull'immagine per saperne di più

La centralità di un nodo può avere tre accezioni:

- centralità basata sul grado (quanti collegamenti in uscita/entrata);
- centralità basata sul suo essere intermedio tra i nodi;
- centralità basata sulla sua vicinanza agli altri nodi.

Due indici molto utilizzati si basano sui primi due significati di centralità e sono:

- *indegree*, che rappresenta la centralità basata sul grado di popolarità (numero di collegamenti ricevuti);
- *betweenness*, che misura la centralità basata sulla interposizione del nodo.

Vediamoli un po' meglio da vicino.

L'indice *indegree* è pari al valore assoluto delle “scelte” ricevute dal nodo. Pertanto, l'*indegree* di un nodo corrisponde al numero dei suoi legami in entrata.

Si può affermare che più è alto il numero di scelte ricevute da ciascun nodo, tanto maggiore è la sua popolarità (centralità) all'interno della rete in esame.

Questo indicatore viene ad esempio usato per misurare la popolarità di un blog, dove la rete è il Web e gli attori sono appunto i siti, mentre i link sono i collegamenti ipertestuali.



Fig. 3.3 Il Google PageRank, che misura la popolarità dei siti, si basa anche sull'indice *indegree*

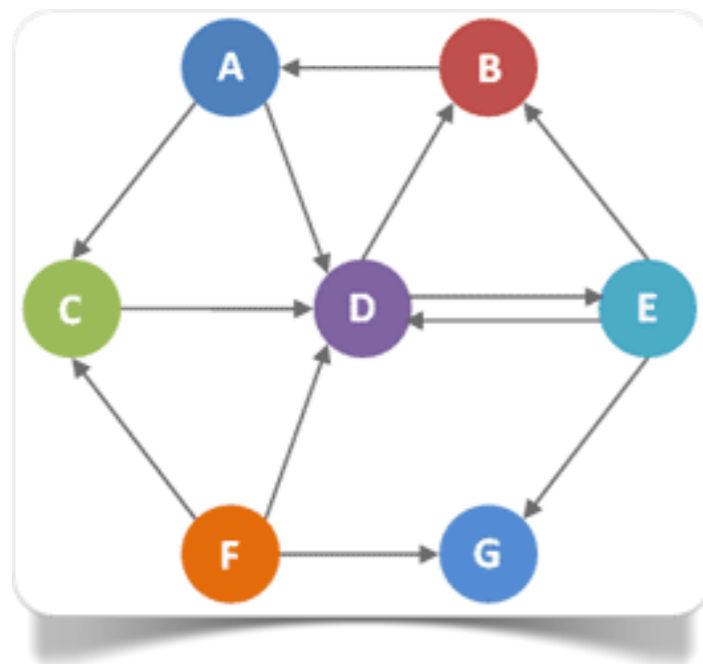


Fig. 3.4 Tra questi sette siti è evidente l'alta centralità del sito D

L'*indegree* varia tra 0 e $n-1$. Valori vicino allo 0 indicano una scarsa centralità del nodo all'interno della rete, ovvero il nodo è periferico; al contrario, valori prossimi a $n-1$ indicano invece una elevata centralità dell'individuo.

L'indice *betweenness* (centralità basata sulla interposizione del nodo) è basato invece sulla frequenza con cui ogni singolo nodo si trova nel percorso più breve che collega ogni altra coppia di nodi. Pertanto, la *betweenness* indica quanto una persona è “intermediaria” tra persone all'interno di un gruppo.

In alti termini, se il valore è alto probabilmente siamo in presenza di uno “snodo” importante di riferimento nelle comunicazioni, negli scambi, nel collegamento tra aree diverse della rete.

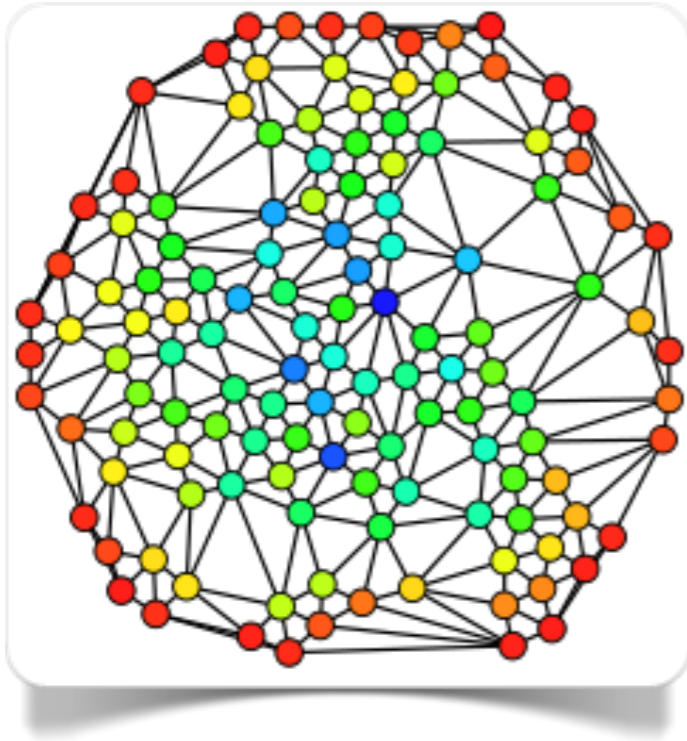


Fig. 3.5 I nodi rossi sono quelli con minore centralità (betweenness), quelli blu con la massima

Per ciascuna delle due misure di centralità può essere calcolata la corrispondente misura di centralizzazione della rete (*global centrality*).

La centralizzazione considera infatti la rete nel suo insieme e misura il grado in cui una rete si concentra intorno ad alcuni nodi centrali; in altre parole, descrive quanto la coesione della rete sia organizzata intorno a determinati punti focali. E' un altro modo per individuare gruppi di opinion leader.

Strumenti di analisi e visualizzazione delle reti

Ucinet

NetDraw

Pajek

SNA R

NodeXL

ORA

SocNetV

CFinder

Gephi



4

Strumenti di analisi e visualizzazione

Perché è importante visualizzare le reti? Torniamo alla domanda iniziale sul valore aggiunto dei social network: noi e le nostre reti. Ecco allora che creare, esplicitare, analizzare una rete assume un valore strategico per il marketing, la comunicazione, il personal branding e in molti altri campi.

E ovviamente non si può prescindere da una buona visualizzazione della rete.

Ma un **video** vale più di un Mb di parole:



Fig. 4.1. Social Networking in Plain English

In generale, la visualizzazione dati è lo studio della rappresentazione visuale di “dati”.

La quantità e la complessità delle informazioni prodotte nella scienza, nell'ingegneria, nel business tutti i giorni sta aumentando esponenzialmente: delle efficaci rappresentazioni visuali dei dati possono quindi aiutare nella comprensione e nella comunicazione.

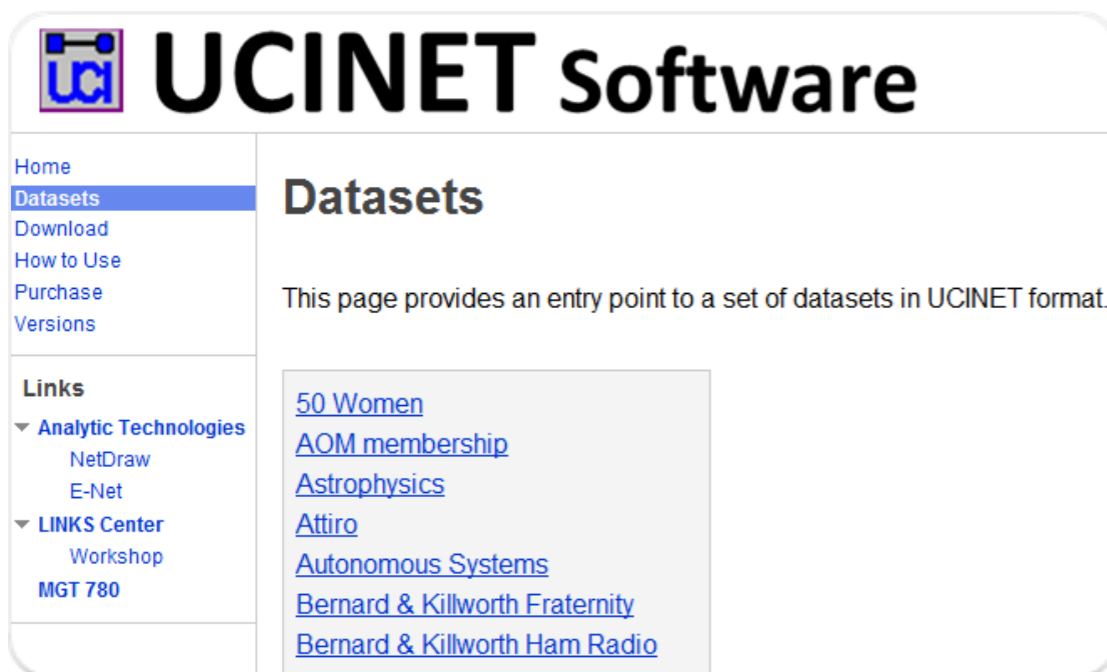
Questa considerazione vale a maggior ragione nel caso delle reti, laddove è difficile, se non impossibile, rappresentare informazioni con tradizionali grafici come istogrammi e torte.

Servono dunque soluzioni ad hoc per la SNA; vediamone alcune.

Ucinet

Il buon vecchio Ucinet è uno degli strumenti più utilizzati per la SNA. Offre un vasto insieme di funzioni e un'ottima documentazione.

Spesso è lo strumento scelto da chi muove i primi passi nella SNA: dopo aver scaricato la versione trial, caricate dei dati sotto forma di semplici **matrici di adiacenza**, iniziate a calcolare densità centralità e tirare fuori una rappresentazione grafica. Se volete dei dati di prova su cui esercitarvi, Ucinet ne mette a disposizione una **galleria** (fig. 4.2).



Sito: <https://sites.google.com/site/ucinetsoftware/home>

Caratteristiche: sistemi operativi MS Windows. Proprietario. Gratuito nella versione trial e circa 30 euro per l'acquisto.

Complessità: bassa.

Doc: <http://www.analytictech.com/networks.pdf>

Fig. 4.2 Ucinet Software / Data sets

NetDraw

Nel pacchetto Ucinet è incluso anche il freeware di visualizzazione NetDraw: un buon modo per iniziare a familiarizzare con la SNA.

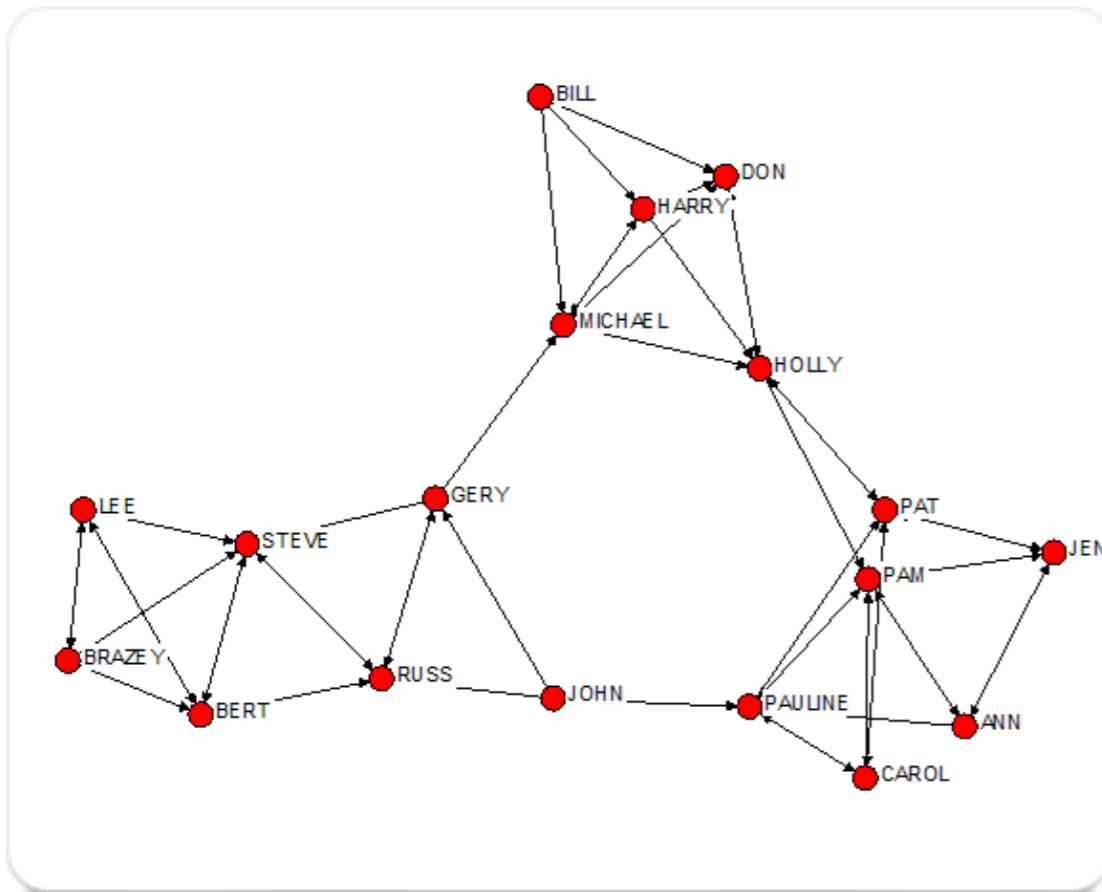


Fig. 4.3 Rete visualizzata con NetDraw

Sito: <https://sites.google.com/site/netdrawsoftware/home>

Caratteristiche: sistemi operativi MS Windows. Proprietario.
Gratuito.

Complessità: bassa

Doc: <http://www.analytictech.com/netdraw/netdrawguide.doc>

Proprio in quest'ultimo documento è possibile effettuare un'analisi (ottimo esercizio) che porta alla visualizzazione riportata in figura 4.3.

Pajek

Pajek è uno strumento di analisi e visualizzazione di (grandi) dataset, gratuito per un uso non commerciale. E' più complesso di Ucinet ma vale la pena considerarlo come valida alternativa open.

Pajek consente inoltre la visualizzazione tridimensionale dei dati.

Students networks

There were 32 students (22 female and 10 male) in the 4th class of Sociology, Social Science Informatics, Faculty of Social Sciences, Ljubljana. Students know each other very well.

Several networks were collected using the following questions:

1. **as1.net:** Suppose you have a birthday next week and you organize a party. Whom would you like to invite to the party? Choose as many students as you wish.
2. **as2.net:** Suppose there will be an organized trip by bus next week. With whom would you like to sit on the neighboring seats? Choose as many students as you wish.
3. **as3.net:** Whom would you select to inform you about important news in school, if you were out of school because of illness? Choose as many students as you wish.
4. **as4.net:** Sometimes you cannot be present at the lectures. From whom would you like to borrow the teaching material? Choose as many students as you wish.
5. **as5.net:** With whom do you discuss important personal issues? Choose as many students as you wish.
6. **as6.net:** From whom would you borrow small amount of money (e.g. 20 EUR)? Choose as many students as you wish.

Fig. 4.4 Alcuni dei dataset messi a disposizione sul sito di Pajek. Clicca l'immagine per navigare la galleria

Sito: <http://pajek.imfm.si/doku.php?id=pajek>

Caratteristiche: sistemi operativi MS Windows. Proprietario. Gratuito per uso non commerciale.

Complessità: media.

Doc: <http://vlado.fmf.uni-lj.si/pub/networks/pajek/doc/pajekman.pdf>

SNA R

R è un linguaggio di programmazione per l'analisi statistica open source e gratuito. È molto utilizzato in ambito accademico, ha una curva di apprendimento piuttosto ripida ma è estremamente potente. Dispone inoltre di pacchetti con cui effettuare analisi specifiche, come ad esempio il package **SNA**, che permette di effettuare analisi sia descrittive che **inferenziali**.

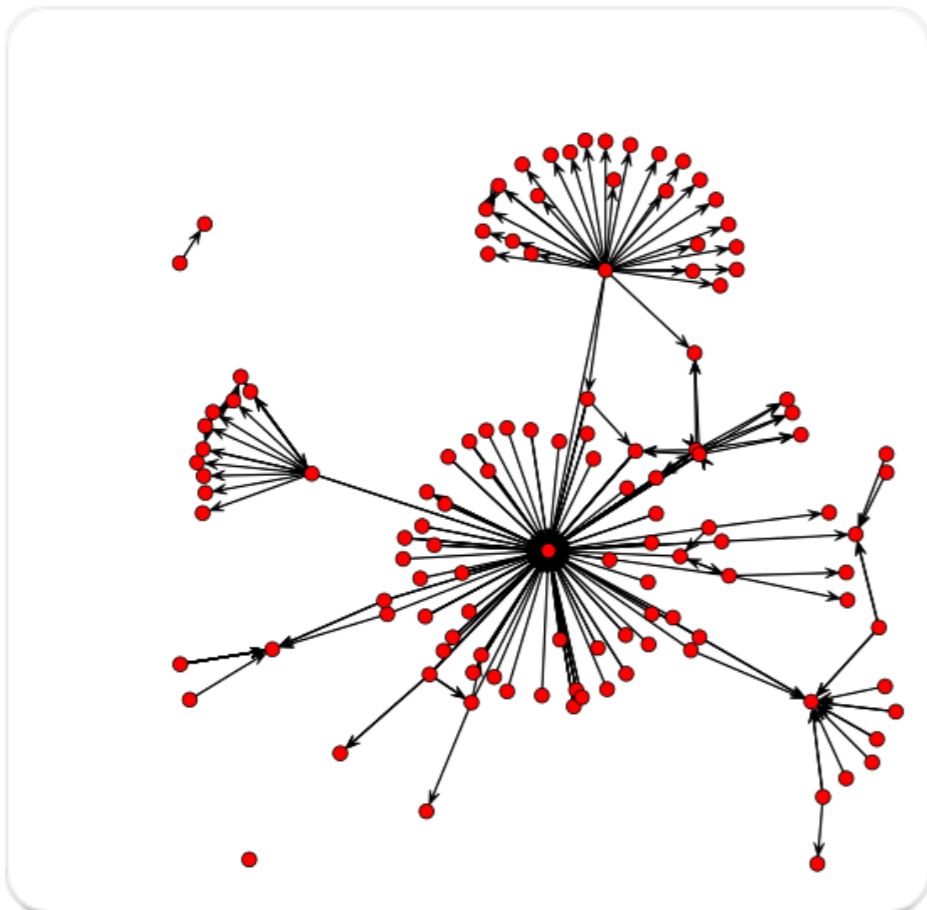


Fig. 4.5 Analisi del flusso di email (“da” e “per” di un server di posta IMAP); il nodo isolato rappresenta un’email inviata a sé stessi.

Sito: <http://cran.r-project.org/web/packages/sna/index.html>

Caratteristiche: multiplatforma. Open source. Gratuito.

Complessità: alta.

Doc: <http://cran.r-project.org/web/packages/sna/sna.pdf>

NodeXL

Il prodotto di Microsoft per la SNA si chiama NodeXL e consente analisi e visualizzazioni a partire da un semplice foglio Excel. L'ambiente di lavoro, riportato in fig. 13, è molto intuitivo e consente di importare dati dai formati classici della SNA (Ucinet, Pajek etc.) e anche direttamente da social media come Twitter, YouTube, Flickr, Facebook e porzioni del Web.

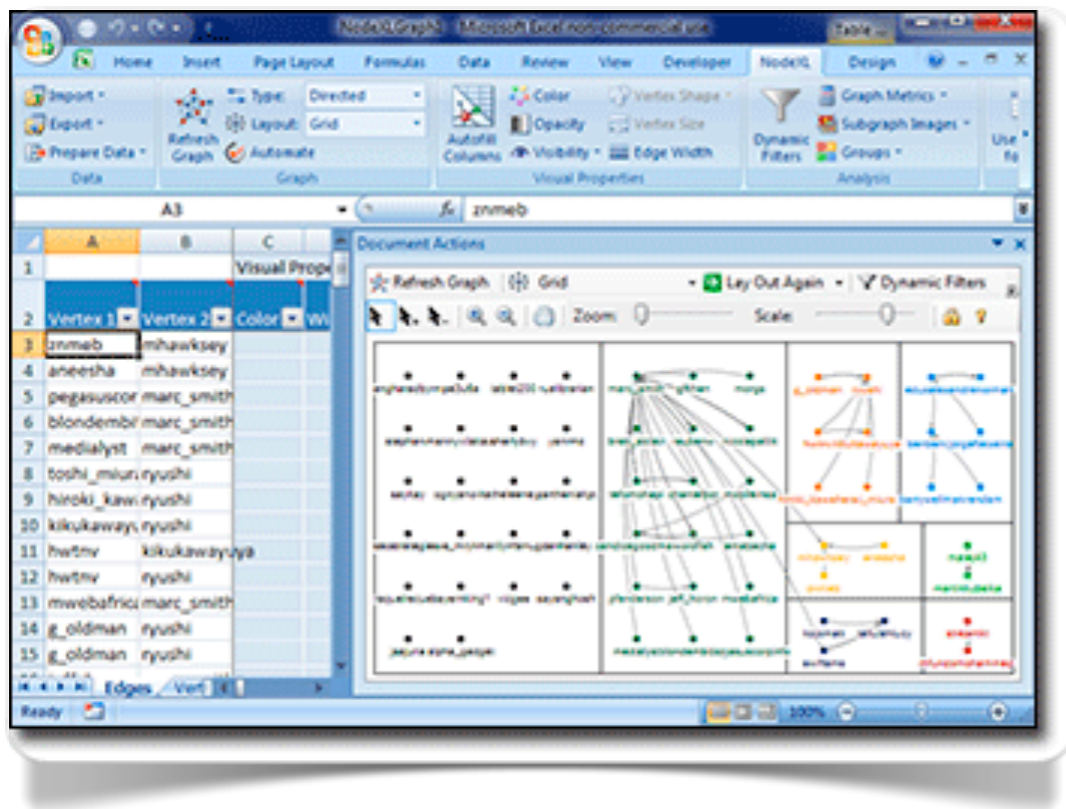


Fig. 4.6 L'ambiente di lavoro di NodeXL

Sito: <http://nodexl.codeplex.com/>

Caratteristiche: per sistema operativo Windows con Office 2007, 2010 o 2013. Open source. Gratuito per uso accademico e di ricerca.

Complessità: bassa.

Doc: presentazione del prodotto e dataset di prova; il **manuale**, invece, è a pagamento.

ORA

ORA nasce dall'approfondimento del carattere multidisciplinare della SNA, e in particolare dal terreno comune della teoria dei grafi, della psicologia sociale, della ricerca operativa e della teoria del management.

Oltre a tutte le funzionalità base che ci si aspetta da un tool di SNA, ci sono centinaia di funzioni avanzate, anche per il confronto tra reti, con un forte accento sulla dinamicità dei grafi (dunque, molto utile per le analisi di una rete nel tempo).

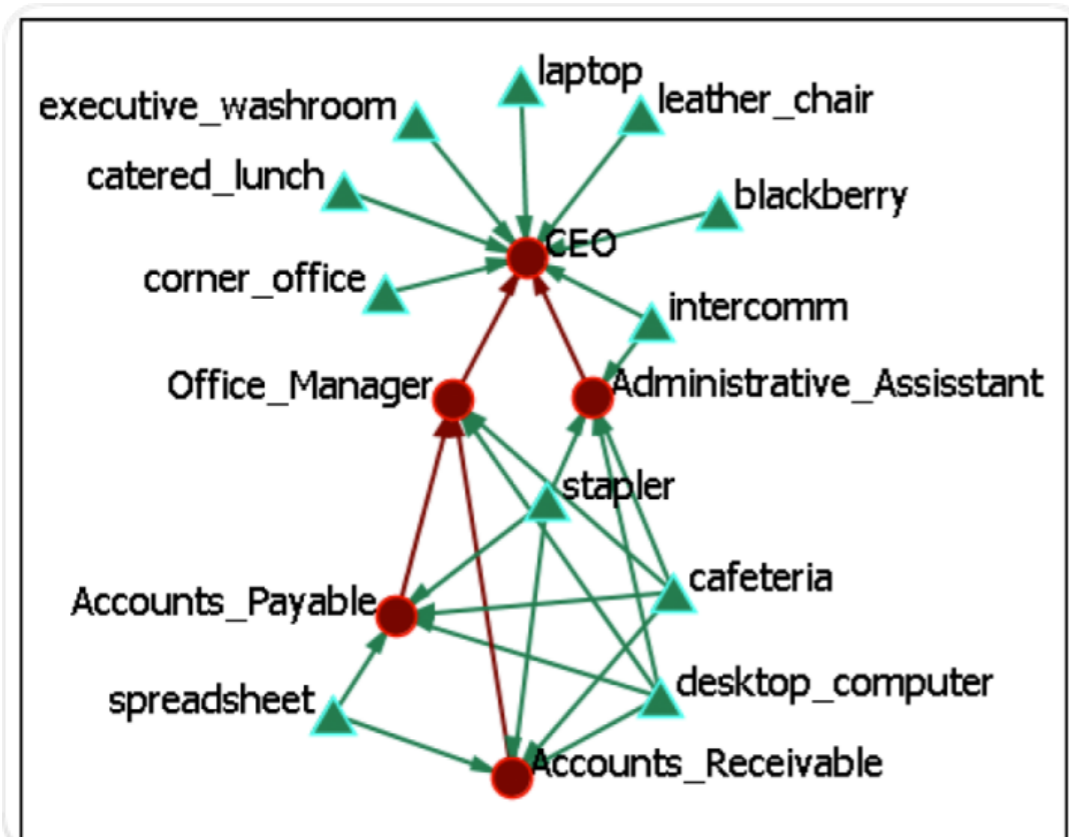


Fig. 4.7 Analisi di tipo *Organization Network Analysis* con ORA

I campi di applicazione di questo strumento sono numerosi, anche nell'analisi delle reti terroristiche e nel **metodo del percorso critico** del *project management*.

Sito: <http://www.casos.cs.cmu.edu/projects/ora/>

Caratteristiche: per sistema operativo Windows e Linux.
Proprietario. Gratuito per uso accademico e di ricerca.

Complessità: media.

Doc: <http://www.casos.cs.cmu.edu/publications/papers/ORA%20QuickStart-1.0.pdf>

SocNetV

Social Networks Visualizer (SocNetV) è un tool molto utile per chi deve “costruire” reti, ad esempio a partire da file di altri prodotti (GraphViz, GraphML, Adjacency, Pajek, UCINET etc). Una particolarità, molto gradita, di SocNetV, è quella di avere un suo *web crawler* che, con una URL di partenza, crea automaticamente reti sulla base dei collegamenti ipertertestuali presenti sulle pagine web.

Diventa così particolarmente semplice analizzare, per esempio, un brand sul Web.

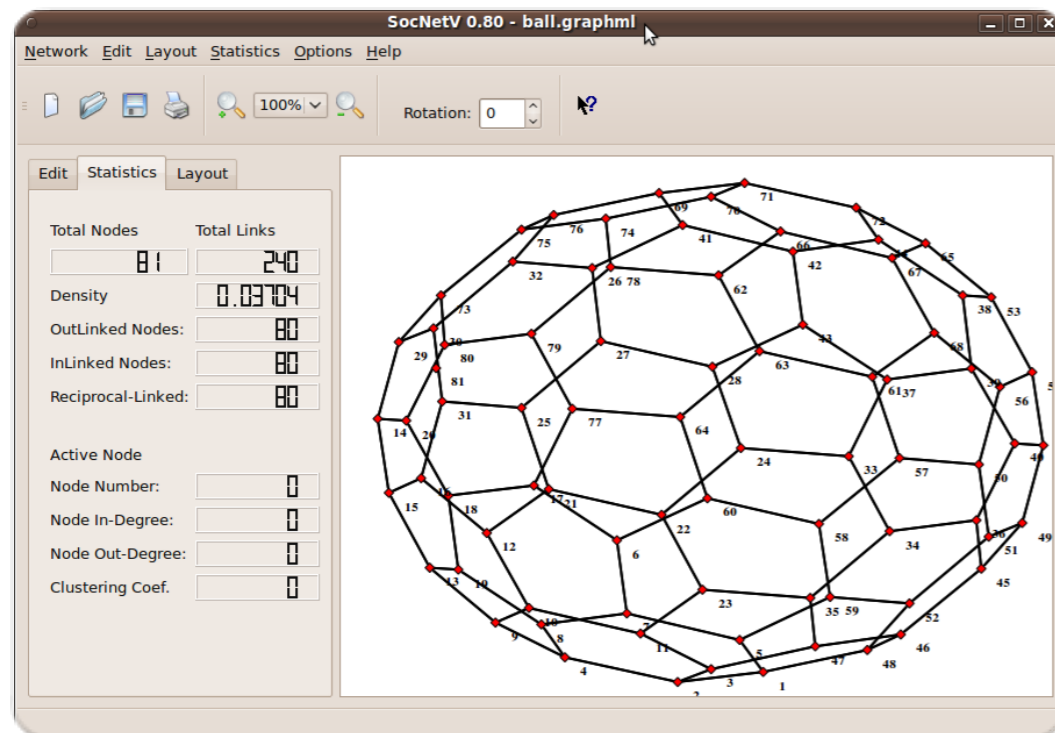


Fig. 4.8 Ambiente di lavoro di SocNetV

Il software è fermo dal 2010 ma conta ancora numerose migliaia di utenti attivi.

Sito: <http://socnetv.sourceforge.net/index.html>

Caratteristiche: Multiplatforma. Open source. Gratuito.

Complessità: media.

Doc: <http://socnetv.sourceforge.net/docs/manual.html>

CFinder

CFinder è un software utilizzato per identificare e visualizzare gruppi di nodi e clique in reti grandi e/o sparse. CFinder offre la possibilità di clusterizzare dati anche in reti molto estese, e quindi viene spesso utilizzato nell'ambito di studi genetici.

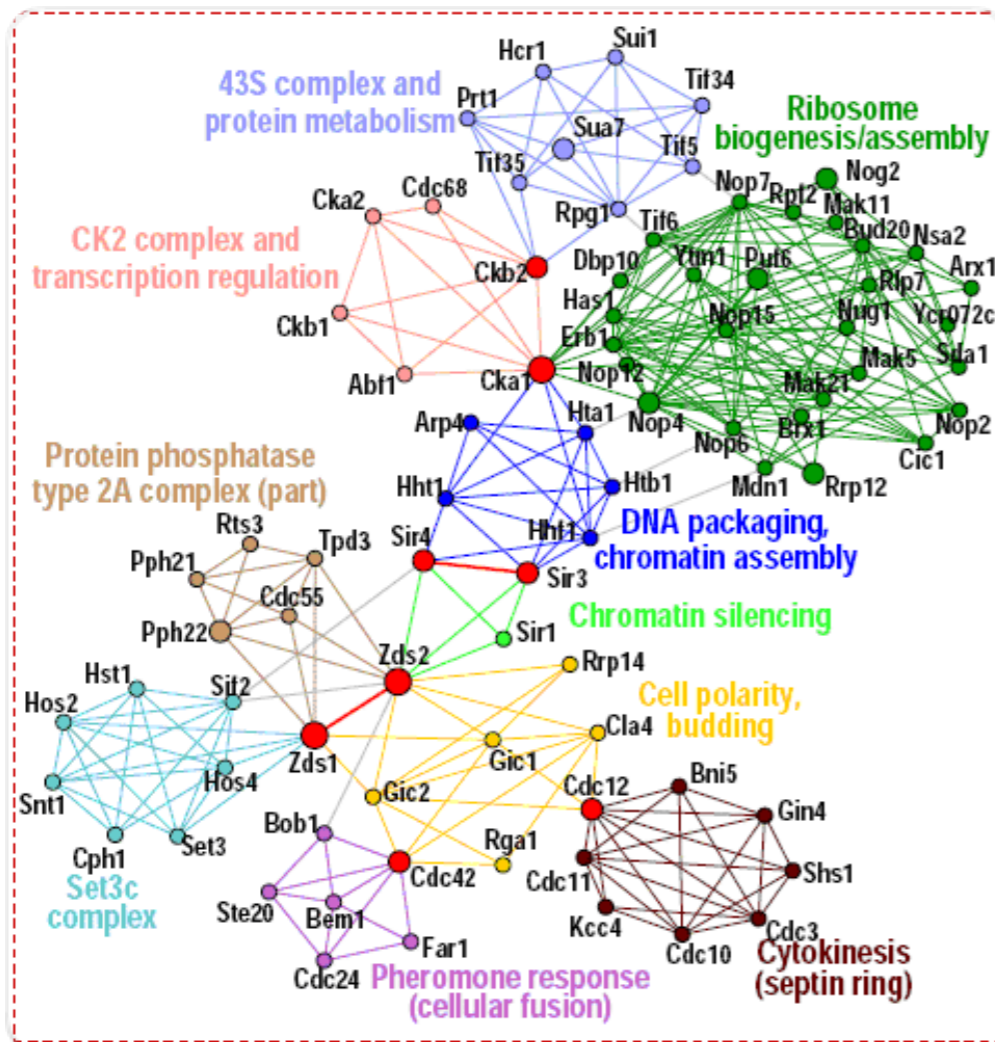


Fig. 4.9 Analisi della rete di interazione tra proteine condotta con CFinder

Sito: <http://www.cfinder.org/>

Caratteristiche: Multiplatforma. Proprietario. Uso gratuito per fini non commerciali.

Complessità: alta.

Doc: <http://hal.elte.hu/cfinder/wiki/?n=Main.Manual>

Gephi

Gephi è una soluzione che consente l'analisi e la visualizzazione di qualsiasi tipo di rete, di sistemi complessi e grafi dinamici.

E' una delle soluzioni più potenti sul mercato e non a caso Google la segue con molta attenzione.

Non offre tutte le statistiche dei tradizionali tool ma può essere un ottimo strumento aggiuntivo per l'esplorazione delle reti.



Fig. 4.10 Analisi di un insieme di siti web con Gephi

Supporta formati tipici della SNA come **GDF (GUESS)**, GraphML (NodeXL), **GML**, NET (Pajek), **GEXF** e altri ancora.

Sito: <http://gephi.github.io/>

Caratteristiche: Multiplatforma. Open source. Gratuito.

Complessità: media.

Come partire (con un esempio): <http://gephi.github.io/users/quick-start/>

Strumenti di analisi e visualizzazione dei social media

Touchgraph Navigator Facebook Browser

Mention Map

Circle Of Trust

Google+ Ripple

Pattern CLiPS

Strumenti ad hoc



5

Touchgraph Navigator Facebook Browser

Ci sono poi software specificamente pensati per eseguire analisi sui social media, eccone alcuni, tutti gratuiti:

Touchgraph Navigator Facebook Browser, strumento nato e pensato per analizzare e rappresentare le reti di Facebook in una modalità molto suggestiva, è in grado di caricare dati direttamente da un account oppure da un file o da un database.

Utili le funzioni di filtro e di esportazione delle immagini.

Da un punto di vista grafico è un software molto potente ma richiede l'installazione di una delle ultime versioni di Java e può appesantire la navigazione.

Questo strumento, che è gratuito, fa parte di una suite molto interessante di prodotti Touchgraph Navigator, che però sono a pagamento.

Sito: <http://www.touchgraph.com/facebook>

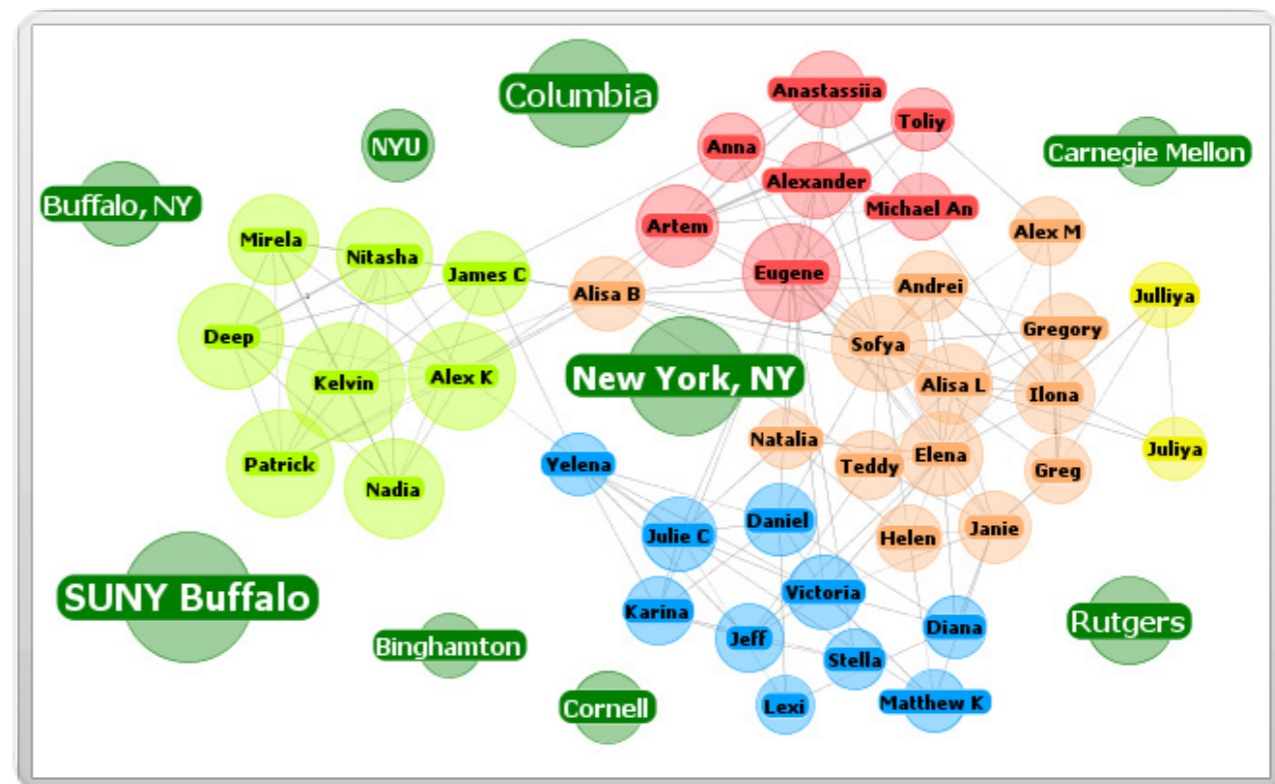


Fig. 5.1 Touchgraph Navigator Facebook Browser in azione per un utente Facebook. Sono presenti anche i riferimenti territoriali (nodi verdi).

Mention Map

Ecco come si presenta questo servizio: *See your Twitter community. We listen to the chatter and map the communities formed around your brand or topic.*

Una menzione, su Twitter, equivale ad una citazione: qualcuno parla di voi, riportando il vostro nome (il nome utente Twitter). Conoscere e analizzare la rete di diffusione delle menzioni è molto importante, ad esempio, per le aziende, per valutare l'efficacia di una campagna di comunicazione sui social media, per valutare lo stato di un brand e altro ancora.

Ecco perché Mentionmapp è uno degli strumenti che non può mancare nel toolbox di un **social media manager**.

Sito: <http://mentionmapp.com/>

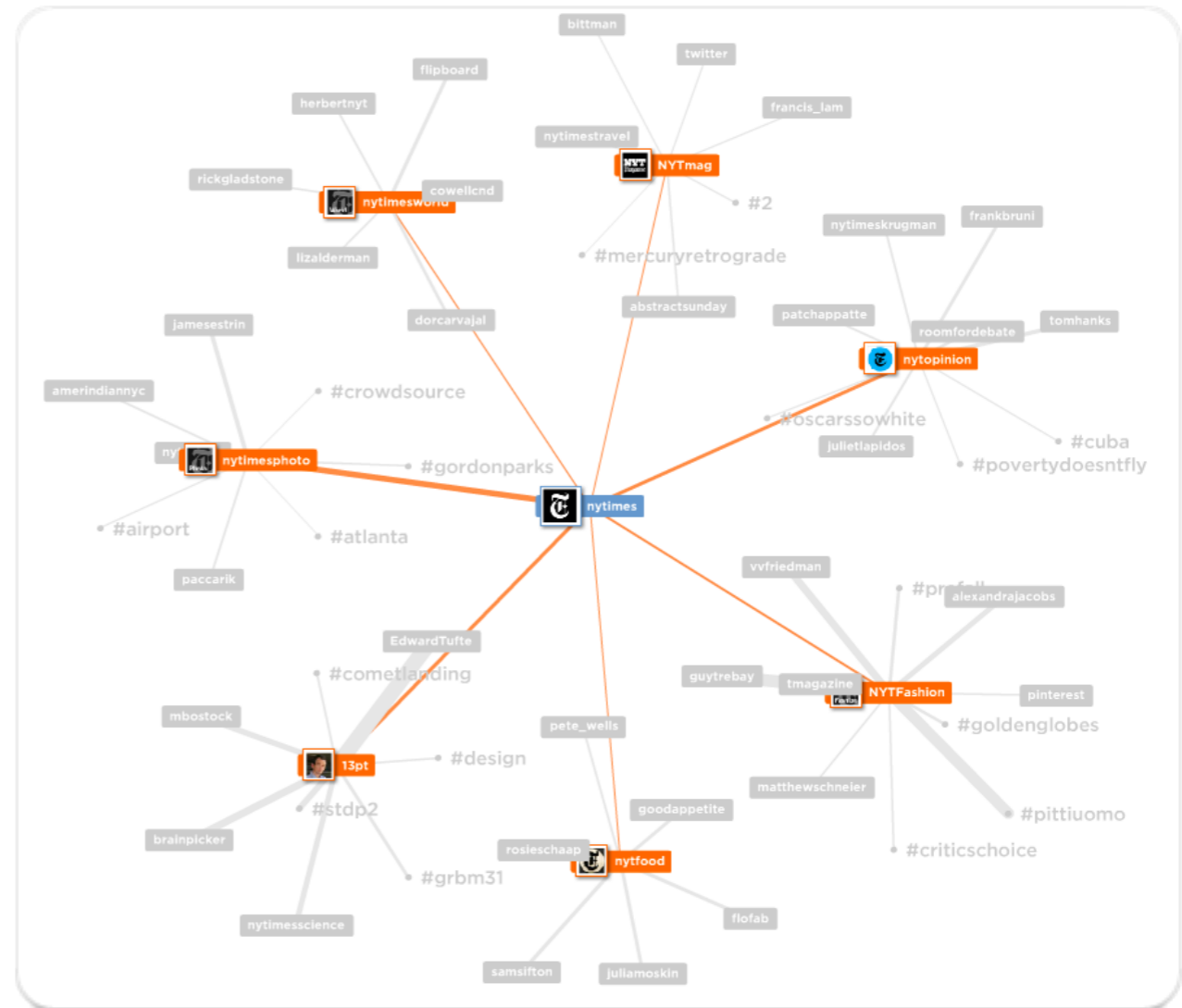


Fig. 5.2 Il network dei tweets del New York Times il 16 gennaio 2015 alle ore 17.00

Circle Of Trust

Circle Of Trust (per Google Plus) è uno strumento nato per rispondere a queste due domande:

- Quanto è asimmetrica la tua rete di relazioni su Google Plus?
- C'è un modo per visualizzare come gli utenti stanno usando e costruendo questo ambiente sociale?

Il sistema offre una suggestiva visualizzazione, basata sulle API di Google+ e sull'HTML5 (dunque, visibile anche sui prodotti Apple).

Per provare con la vostra utenza Google Plus, andate su <http://www.d3.do/labs/circleoftrust/>. Vedrete, distinte per colori, le persone che avete aggiunto voi, quelle che vi hanno aggiunto e i legami reciproci.

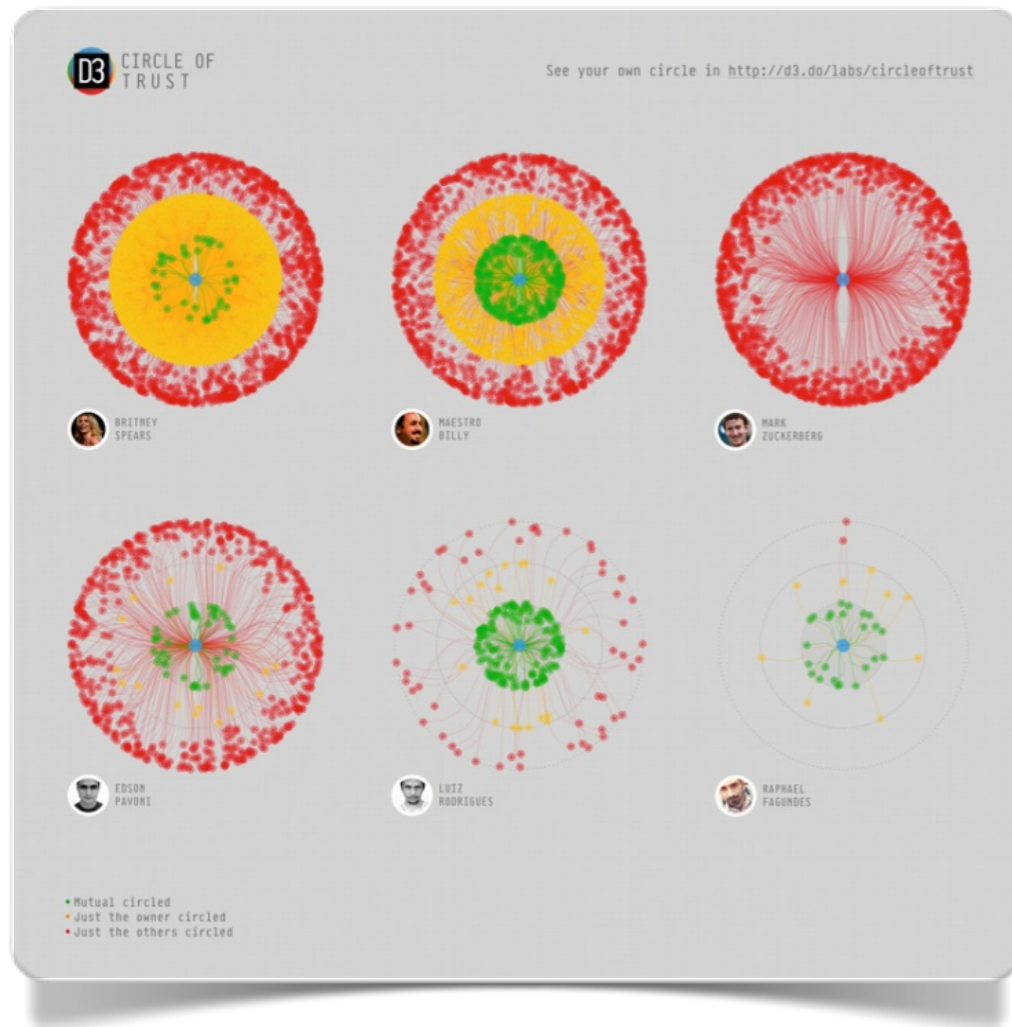


Fig. 5.3 Circle Of Trust in azione su alcuni personaggi noti

Google+ Ripple

Un altro servizio collegato a Google + è **Ripples**: è la funzione “eco” (diffusione della notizia) presente su ciascun post pubblico su Google +. Cliccando su eco si ottiene un grafico (un grafo) che mostra, dinamicamente, la diffusione di quella notizia. Oltre alla rappresentazione grafica, vengono mostrate alcune importanti statistiche come ad esempio gli opinion leader della rete di diffusione (quindi Ripples può essere utilizzato per trovare i nodi più importanti che parlando di un brand) e il numero di condivisioni medie.

Nell'**esempio di figura 5.4**, navigabile anche con iPad, viene mostrata la diffusione di un post del celebre blogger Pete Cashmore (**Mashable**).

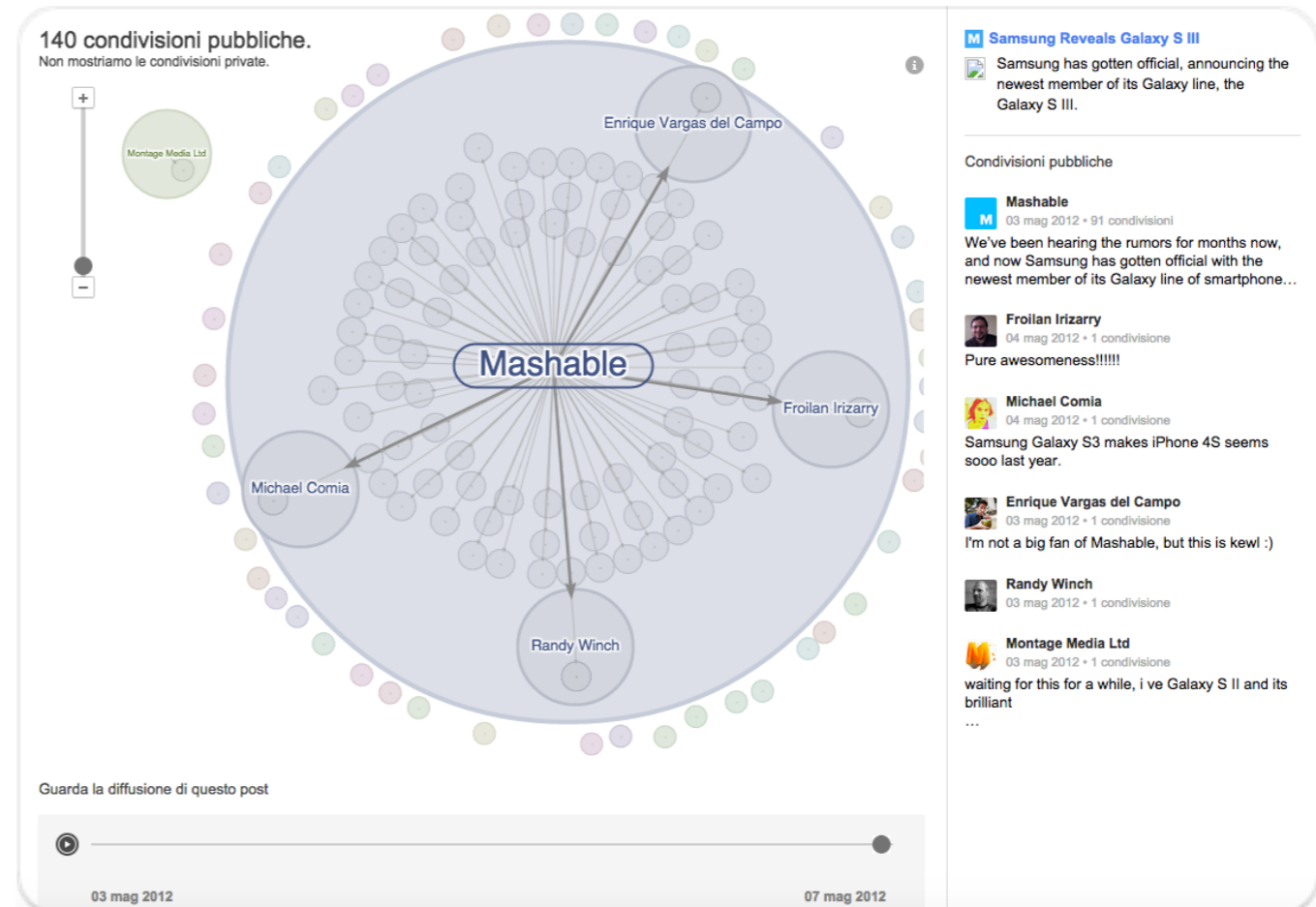


Fig. 5.4 Diffusione di un post su Google+ con Ripples

Pattern CLiPS



Fig. 5.5 Processo di analisi con Pattern clips

I punti di forza di questo prodotto sono i moduli di *web scraping*, *text mining* e *sentiment analysis* e la possibilità di un'analisi integrata di un brand sui social media (v. fig. 5.5).

Tra le numerose funzionalità, anche la rappresentazione grafica dei *network di parole* correlate semanticamente, come riportato in figura 5.6.

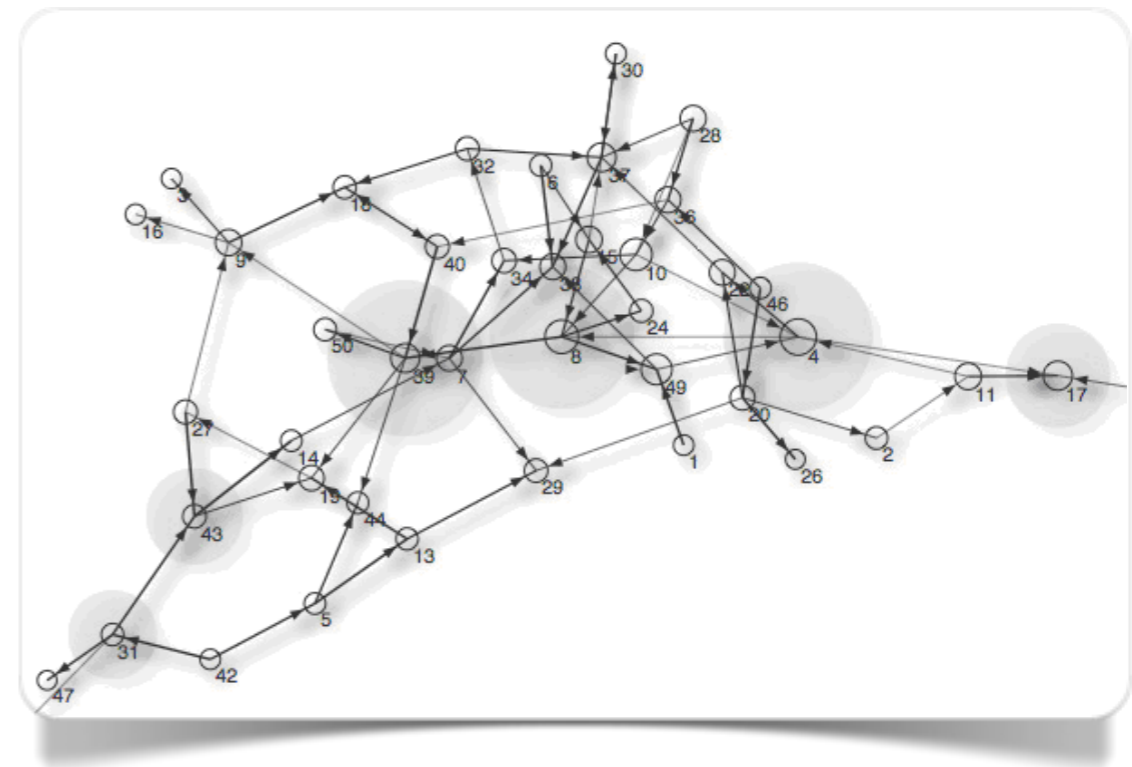


Fig. 5.6 Esempio di analisi di un testo dove i nodi indicano le parole e i collegamenti la correlazione semantica

Strumenti ad hoc



Fig. 5.7 Mappa del potere in Italia, strumento di SNA realizzato dallo studio Casaleggio Associati

È possibile poi prendere in considerazione l'ipotesi di sviluppare dei prodotti ad hoc. È questo il caso della **Mappa del potere** dello studio Casaleggio e Associati, a lungo pubblicizzata sul sito **Blog di Beppe Grillo**. Il successo di questo sito dipende da molti fattori tra i quali la consapevolezza dell'importanza delle reti e della loro cura: la scelta dei collegamenti presenti sul suo sito, la gestione delle "reti territoriali" (MoVimento 5 Stelle e MeetUp) e anche, appunto, il sistema di visualizzazione interattiva della mappa del potere, uno strumento molto apprezzato da cittadini e giornalisti ma anche dagli stessi manager nonché, a volte, dai pubblici ministeri.

È uno strumento ad hoc anche quello sviluppato da Gianluca Leboni e Lucia Passaro del Laboratorio di Linguistica Computazionale del Dipartimento di Filologia, Letteratura e Linguistica dell'Università di Pisa. **Guarda come Parlano** è un tool sperimentale con il quale si possono conoscere le relazioni tra le parole (e quindi anche il mondo lessicale dei parlanti) di alcune personalità del mondo politico e dello spettacolo. Nella versione attualmente online sono stati analizzati 30 personaggi sulla base di un insieme di testi tratti da interviste, dichiarazioni, discorsi ecc.

raccolti tra agosto e settembre 2013. Si tratta di 533 documenti per un totale di 466.755 parole. I raggruppamenti di parole spazialmente vicini individuano connessioni di ruoli e di temi. Per esempio, Grillo, Vendola e Berlusconi, distanti per moltissimi aspetti, hanno in comune la parola “cittadino”. Nello spazio semantico rappresentato dalle relazioni tra queste parole si individuano raggruppamenti di personaggi che condividono ruoli sociali, come i politici, i calciatori, le persone di spettacolo. Beppe Grillo si posiziona, non a caso, al confine tra diversi “mondi”.

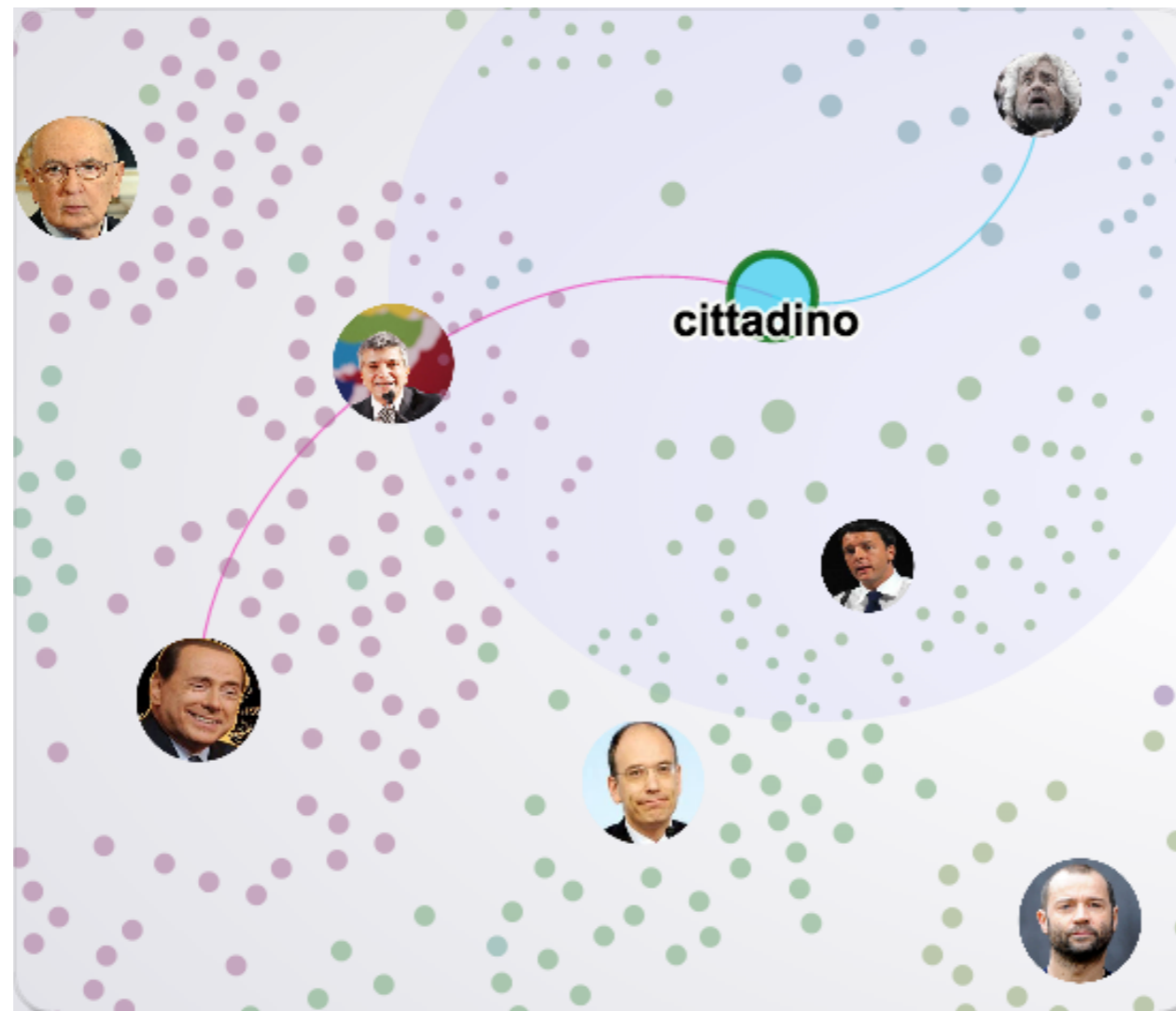


Fig. 5.8 Guarda come Parlano (Laboratorio di Linguistica computazione dell'Università di Pisa) - Clicca sull'immagine per accedere alla demo

Riepilogo

6

Una lista ripilogativa

Una lista (quasi) completa di tutti gli strumenti di visualizzazione di reti è presente su [Wikipedia](#) e su [Software for social network analysis](#), ma attenzione: alcuni software sono fermi da molti anni e quindi non più supportati da aziende, team di sviluppo e community.

Sette cose che non possono mancare nella cassetta degli attrezzi di chi inizia ad occuparsi di SNA:

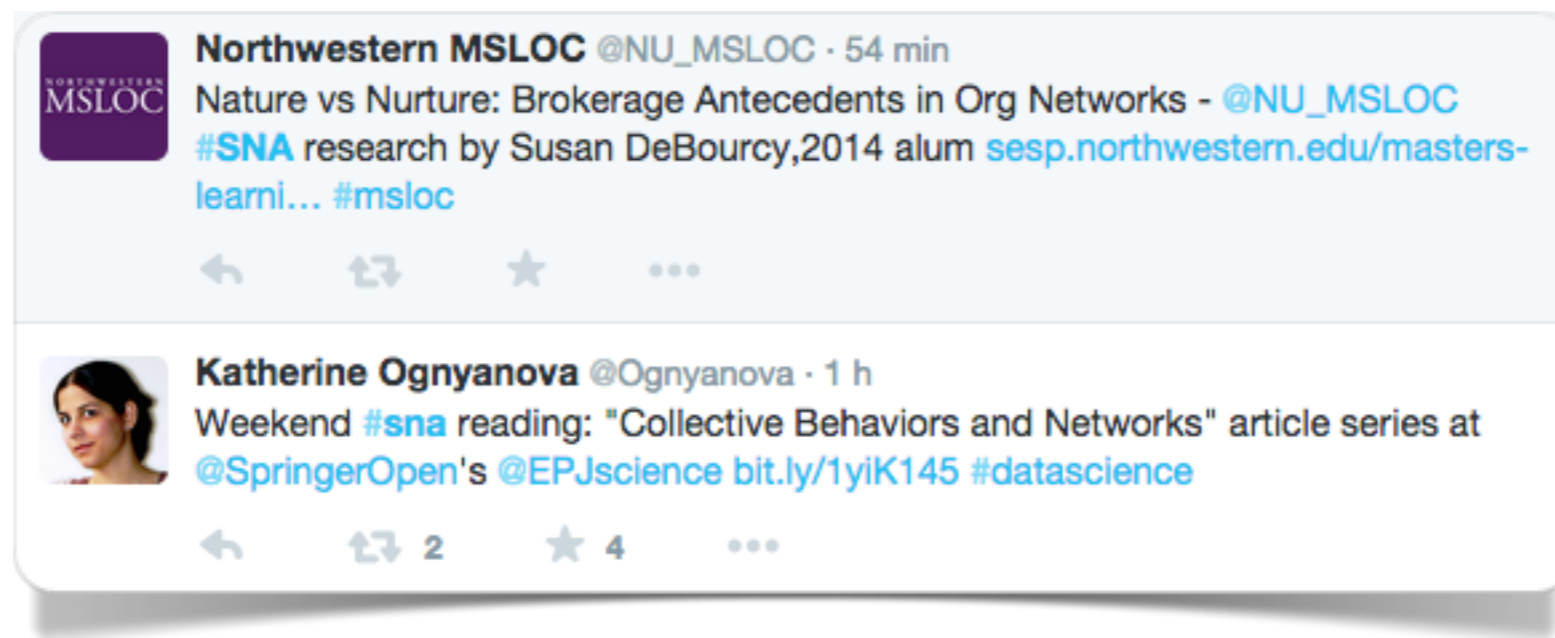


Fig. 6.1 Tweets con l'hashtag #sna

- 1) L'[articolo sulla SNA](#) presente su Wikipedia, al quale hanno contribuito numerosi studiosi del settore.
- 2) Un buon testo sulla storia dell'analisi delle reti sociali : Linton C. Freeman, *The Development of Social Network Analysis. A Study in the Sociology of Science* ([recensione](#) di Charles Kadushin), disponibile anche in lingua italiana: [Lo sviluppo delle reti sociali. Uno studio di sociologia della Scienza](#) (FrancoAngeli).

- 3) Un testo teorico e pratico, con numerosi esempi, per mettersi subito alla prova: [Social Network Analysis. Theory and Applications](#).
- 4) Qualche software gratuito e semplice per iniziare: [Ucinet](#), [NetDraw](#), [NodeXL](#).
- 5) Essere presenti sui vari gruppi professionali su LinkedIn (gruppo [Social Network Analysis in practise](#)), Facebook (pagina [Social Network Analysis](#)), Twitter (seguendo l'hashtag #sna) e non perdersi le news sul sito dell'[International Network of Social Network Analysis](#).

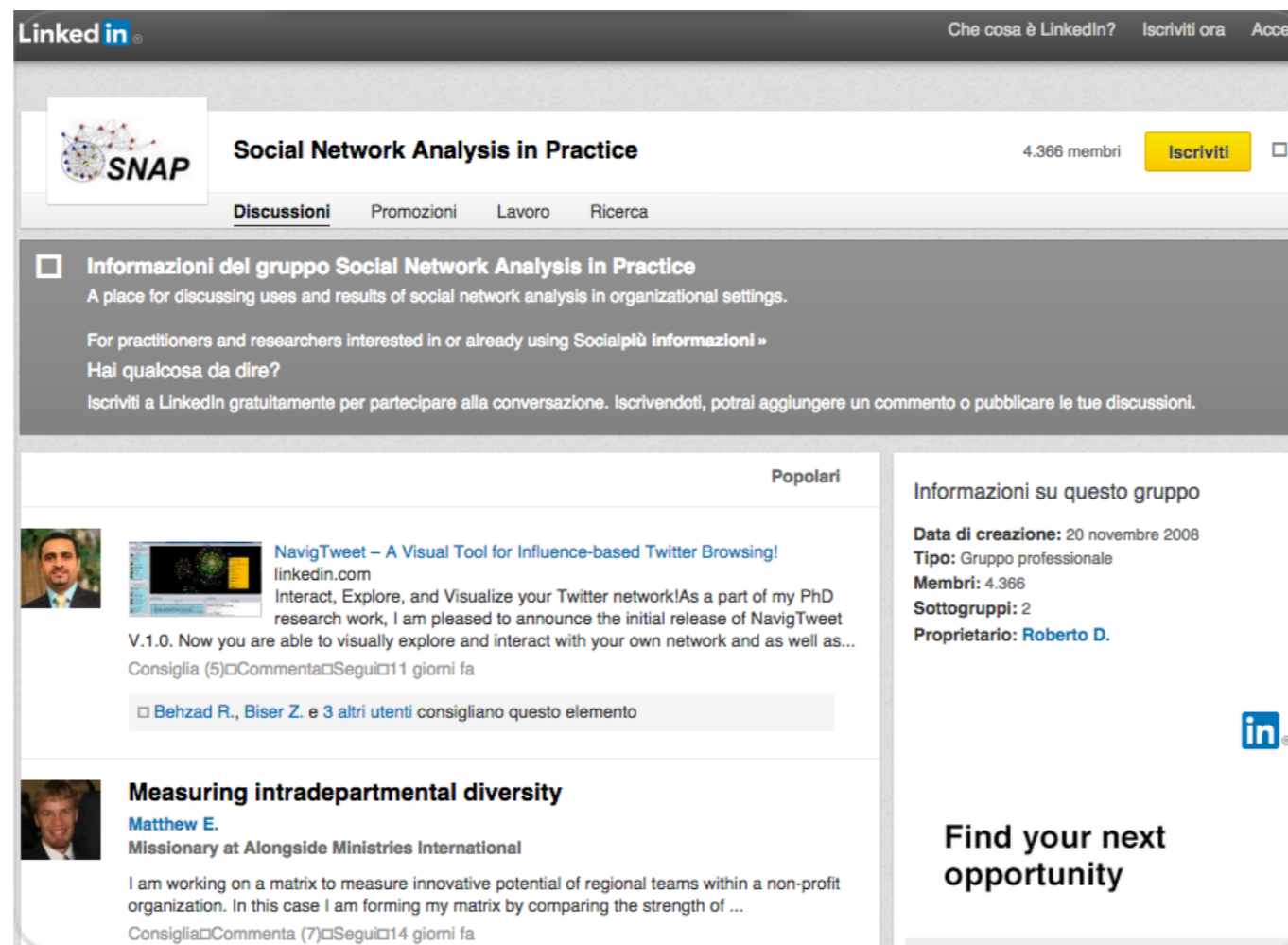


Fig. 6.2 Social Network Analysis in Practice

6) Una sito/bibliografia sempre a portata di mano. Un'ottima bibliografia aggiornata al 2004 sul tema della SNA è su <http://www.socialnetworks.org/>. Una selezione che comprende, tra gli altri, alcuni libri di SNA e ONA pubblicati negli ultimi anni è invece [qui](#).

7) Uno sguardo al futuro verso le soluzioni più innovative e orientate al web; due progetti di particolare interesse per rappresentare reti in modo dinamico e con una grafica di alto livello (compatibile anche con il mondo Apple) sono [Arbor.js](#) e [SigmaJs](#) (nella figura 5.3 una demo navigabile di [Atlas](#) realizzata con SigmaJs).

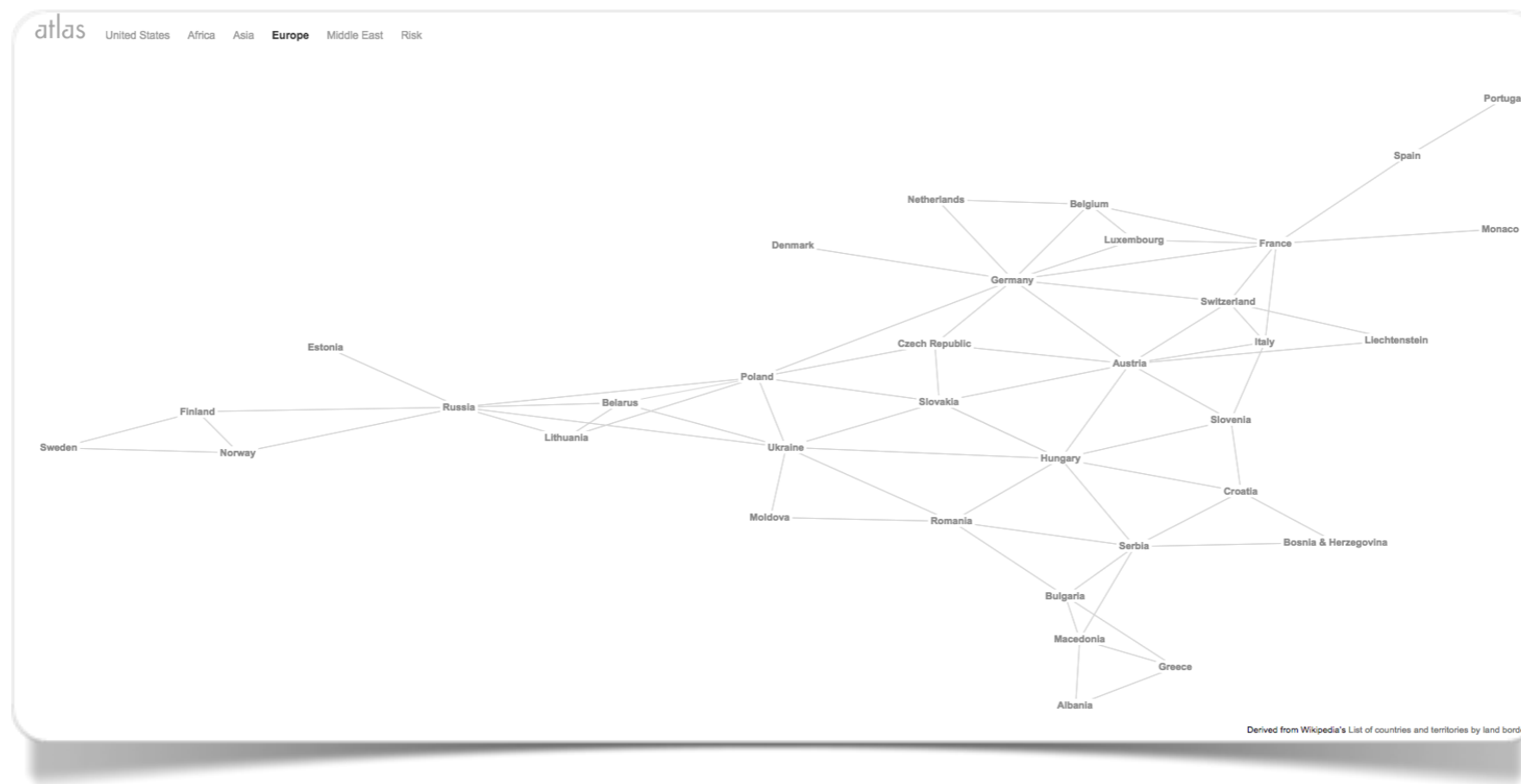


Fig. 6.3 Demo navigabile di Atlas (SigmaJs)

Bibliografia e risorse in web

7

Bibliografia

- Borgatti S. P., Everett M. G., Freeman L.C. (2002) *Ucinet for Windows: Software for social network analysis*. Harvard (MA): Analytic Technologies.
- Campbell M., Dagli C. K., Weinstein C. J. (2013) Social Network Analysis with Content and Graphs. *Lincoln Laboratory Journal*, 20, 1, [MIT Online](#) (cons. 28/1/2015).
- Can F., Özyer T., Polat F. (2014) *State of the Art Applications of Social Network Analysis*. London: Springer.
- Carrington P. J., Scott J., Wasserman S. eds. (2005) *Models and methods in social network analysis*. New York: Cambridge University Press.
- Cross R. L., Thomas R. J. (2009) *Driving Results Through Social Networks: How Top Organizations Leverage Networks for Performance and Growth*. Hoboken (NJ): Jossey-Bass (Wiley & Sons).
- Cross R. L., Thomas, R. J., Singer J., Colella S., Silverstone Y. (2010) *The Organizational Network Fieldbook. Best Practices, Techniques and Exercises to Drive Organizational Innovation and Performance*. Hoboken (NJ): Jossey-Bass (Wiley & Sons).
- Doreian P., Stokman F. eds. (1997) *Evolution of Social Networks*. Amsterdam: Routledge.
- Faust K., Wasserman S., Iacobucci D. (1994) *Social Network Analysis. Methods and Applications (Structural Analysis in the Social Sciences)*. Cambridge: Cambridge University Press.
- Freeman L. C. (2007) *Lo sviluppo dell'analisi delle reti sociali. Uno studio di sociologia della scienza*. Milano: Franco Angeli.
- Garson G. D. (2012) *Network Analysis*. North Carolina State University, Statistical Associate Publishers, Blue Book Series.
- Garton L., Haythornthwaite C., Wellman B. (1997) Studying Online Social Networks. *Journal of Computer-Mediated Communication*, 3, 1 (June), [Wiley Online Library](#) (cons. 28/1/2015).

Goldbeck J. (2013) *Analyzing the Social Web*. Barlington (MA): Morgan Kaufmann ([Table of Contents](#)).

Hanneman R.A., Riddle M. (2005) *Introduction to social network methods*. Riverside (CA): University of California ([disp. online 28/1/2015](#))

Newman M. E. J (2010) *Networks: An Introduction*. New York: Oxford University Press.

Oriani G. (2008) *La forza delle reti di relazioni informali nelle organizzazioni. L'organizational Network Analysis*. Milano: Franco Angeli.

Scott J. (2013) *Social Networks. Theory and Analysis*. London: Sage (Third Edition).

Tronca L. (2013) *Sociologia relazionale e social network analysis. Analisi delle strutture sociali*. Milano: Franco Angeli.

Risorse in web

Arbor.js (<http://arborjs.org/>)

CFinder (<http://www.cfinder.org/>)

Circle Of Trust (<http://www.d3.do/labs/circleoftrust/>)

Gephi (<http://gephi.github.io/>)

Google+ Ripple (Ripple (<https://support.google.com/plus/answer/1713320?hl=it>))

Guarda come parlano (<http://www.unipi.it/index.php/tutte-le-news/item/3364-scopri-come-parlano-i-politici-e-non-solo>)

Mappa del potere (<http://mappadelpotere.casaleggio.it/>)

Mention Map (<http://mentionmapp.com/>)

NetDraw (<https://sites.google.com/site/netdrawsoftware/home>)

NodeXL (<http://nodexl.codeplex.com/>)

ORA (<http://www.casos.cs.cmu.edu/projects/ora/>)

Pajek (<http://pajek.imfm.si/doku.php?id=pajek>)

Pattern CLiPS (<http://www.clips.ua.ac.be/pages/pattern>)

R - The R Project for Statistical Computing (<http://www.r-project.org/>)

SigmaJs (<http://sigmajs.org/>)

SNA R (<http://cran.r-project.org/web/packages/sna/index.html>)

SocNetV (<http://socnetv.sourceforge.net/index.html>)

Touchgraph Navigator Facebook Browser (<http://www.touchgraph.com/facebook>)

Ucinet (<https://sites.google.com/site/ucinetsoftware/home>)



SAPIENZA
UNIVERSITÀ DI ROMA

Dipartimento di Scienze statistiche

Data Science Series

1. L. Giuliano, *Il valore delle parole. L'analisi automatica dei testi in Web 2.0*
2. D. Schiavon, L. Giuliano, *Wizard grafico. Una guida alla visualizzazione dei dati numerici*
3. L. Giuliano, *The Value of Words. Automatic Text Analysis Tools in Web 2.0*
4. D. Frongia, *Social Network Analysis. Una introduzione ai metodi e agli strumenti*
5. D. Schiavon, L. Giuliano, *Chart Wizard. A guide to the graphical representation of numeric data* (in preparazione)

Daniele Frongia

Social Network Analysis. Una introduzione
ai metodi e agli strumenti

Roma : Dipartimento di Scienze statistiche,
[2015] 43 p.

ISBN 978-88-908757-3-1

